Analysis of Allostery in the Tryptophan Synthase Complex by Horizontal and Vertical Approaches

Dissertation zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
der Fakultät für Biologie und vorklinische Medizin
der Universität Regensburg

vorgelegt von

Michael Schupfner

aus

Regensburg

im Jahr 2019

Das Promotionsgesuch wurde eingereicht am: 12.7.2019

Die Arbeit wurde angeleitet von: Prof. Dr. Reinhard Sterner

Unterschrift: _____

Michael Schupfner

“Causarum enim cognitio cognitionem eventorum facit” – Cicero, *Topica 67*

“Thus the knowledge of causes produces a knowledge of effects”

Abstract

The tryptophan synthase (TS) is a linear αββα complex that catalyzes the last two steps of tryptophan biosynthesis in primary metabolism. It is a model system for subunit interactions in multi-enzyme complexes with a long-standing history of research. The α-subunit TrpA and the β-subunit TrpB display a rather low catalytic activity in isolation but strongly stimulate each other in the TS complex in an allosteric manner. In this thesis, allostery in the TS was studied by means of two different approaches.

In the first manuscript, the TrpA subunit from *Zea mays*, which is dependent on activation by the respective TrpB subunit, was compared to a homologue of TrpA, called BX1. BX1 catalyzes the same reaction in secondary metabolism as does TrpA in primary metabolism but is highly active in the absence of a TrpB-like partner protein. The TrpA - BX1 comparison identified two differing amino acids within a loop region known for its role in allosteric activation of TrpA. The transfer of the corresponding BX1 loop residues into TrpA yielded variants with turnover numbers that were increased by at least one order of magnitude and up to 520-fold. This corresponds to the activation of TrpA by TrpB in the wild-type TS complex. At the same time substrate affinity was reduced drastically, an effect that could be reversed by binding of wild-type TrpB to the TrpA variants. These findings contribute to the understanding of the allosteric activation of the α-subunit by the β-subunit of TS and suggest an evolutionary trajectory that describes the transition from a primary metabolic enzyme regulated by an interaction partner to a self-reliant stand-alone secondary metabolic enzyme.

In the second manuscript, the allosteric network of TS was analyzed by means of ancestral sequence reconstruction (ASR), which is an *in silico* method to resurrect extinct ancestors of modern proteins. In previous work, the sequences of TrpA and TrpB from the last bacterial common ancestor (LBCA) were computed by means of ASR. The corresponding primordial proteins were produced in *Escherichia coli*, purified, and characterized. The results showed that LBCA-TS is reminiscent of modern TS by forming a αββα complex with indole channeling taking place. However, LBCA-TrpA decreases the activity of LBCA-TrpB by a factor of 5 whereas, for example, the modern ncTrpA from *Neptuniibacter caesariensis* increases the activity of ncTrpB by a factor of 30. In order to identify those amino acid residues that are responsible for this large difference, all six evolutionary TrpA and TrpB intermediates that stepwise link LBCA TS with *N. caesariensis* TS were produced and characterized. Remarkably, the switching from TrpB-inhibition to TrpB-activation by TrpA occurred between two successive TS intermediates. The comparison of these intermediates and the mutual

exchange of residues by iterative rounds of site-directed mutagenesis allowed for the identification of four (out of 413) residues from TrpB that are necessary and sufficient for its allosteric activation by TrpA. These findings demonstrate that ancestral sequence reconstruction can efficiently identify residues essential for allosteric communication and contribute to our understanding of signal propagation in TS.

References of Published Manuscripts

This thesis contains a manuscript that is published and one chapter with unpublished data that is in preparation for submission.

Manuscript A Adapted with permission from "**Schupfner, M.**, Busch, F., Wysocki, V. and Sterner, R. (2019), Generation of a Stand-Alone Tryptophan Synthase α-Subunit by Mimicking an Evolutionary Blueprint. *ChemBioChem.* doi:10.1002/cbic.201900323"

Manuscript B **Schupfner, M.**, Straub, K., Merkl, R. and Sterner, R. (2019), Analysis of Allosteric Communication in a Multi-Enzyme Complex by Ancestral Sequence Reconstruction. *In preparation*.

In the course of this work, I contributed to another publication, which is not part of this thesis:

Manuscript C Fleming, J.R., **Schupfner, M.**, Busch, F., Baslé, A., Ehrmann, A., Sterner, R. and Mayans, O. (2018), Evolutionary Morphing of Tryptophan Synthase: Funcional Mechanisms for the Enzymatic Channeling of Indole. *Journal of Molecular Biology* 430, 5066-5079.

Personal Contributions

Manuscript A

The experiments were conducted by myself. All figures, except **Figure 11** and **Figure 12**, and all tables were generated by myself. Florian Busch performed native mass spectrometry experiments and generated **Figure 11** and **Figure 12**. Reinhard Sterner and Vicky H. Wysocki supervised research. The manuscript was written by myself and Reinhard Sterner with contributions from Florian Busch.

Manuscript B

The experiments were conducted by myself. Kristina Straub performed *in silico* studies. All figures, except **Figure 20**, and all tables were generated by myself. Kristina Straub contributed to **Figure 15A** and created **Figure 20**. Rainer Merkl and Reinhard Sterner supervised research. All authors contributed to the writing of the manuscript.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Enzymes

BX1	homologue of a-subunit of the tryptophan synthase, catalyzes same reaction, independent of allosteric activation by b-subunit of tryptophan synthase like activator
ImGPS	imidazole glycerol phosphate synthase
LOX	lysyl oxidase
PKA	protein kinase A
TrpA	α-subunit of the tryptophan synthase
TrpB(1)	β-subunit of the tryptophan synthase
TS	tryptophan synthase

Organisms

AncXnc	intermediary ancestor between LBCA and *Neptuniibacter caesariensis*, sequences are calculated by ASR
ec	*Escherichia coli*
LBCA	last bacterial common ancestor, sequences are calculated by ASR
nc	*Neptuniibacter caesariensis*
pf	*Pyrococcus furiosus*
sf	*Shewanella frigidimarina*
st	*Salmonella typhimurium*
tm	*Thermotoga maritima*
tt	*Thermus thermophilus*
zm	*Zea mays*

Chemicals

DIMBOA-Glc	2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one-glucoside
DTT	Dithiothreitol
EPPS	4-(2-Hydroxyethyl)-1-piperazinepropanesulfonic acid
GAP	glyceraldehyde 3-phosphate
IGP	indole-3-glycerol phosphate
IPTG	Isopropyl-β-D-thiogalactopyranosid
LB	lysogeny broth media
PLP	pyridoxal 5'-phosphate

Methods and Techniques

ASR	ancestral sequence reconstruction
HPLC	high performance liquid chromatography
IMAC	immobilized metal ion chromatography
MD	molecular dynamics
MS	mass spectrometry
MSA	multiple sequence alignment
NMR	nuclear magnetic resonance
PNK	site-directed mutagenesis method according to the NEB Q5® site-directed mutagenesis kit
QCM	QuikChange™ site-directed mutagenesis method
SEC	size-exclusion chromatography

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

Nucleotides are abbreviated by the 1-letter code.

1 General Introduction

1.1 Significance of Allostery

Cells across all superkingdoms require a vast number of mechanisms to sustain life. Cell cycle control, transcription and translation, metabolism and signaling pathways are just a few of them (Musnier et al. 2010, Lane and Fan 2015, Washburn and Gottesman 2015, Dever et al. 2016, Reyes-Lamothe and Sherratt 2019). In order to control the enormous complexity and to response reasonably and flexible to external factors like stress factors or stimuli, regulatory mechanisms are needed that act on different levels within a cell. Many regulatory mechanisms involve control on RNA level, like initiation and termination of transcription, mRNA processing or mRNA stability. This kind of cellular control is effective on a long timescale, since mRNA once synthesized can be the foundation for protein synthesis long after a gene is silenced (Desvergne et al. 2006). For control of cellular processes on a shorter timescale, regulatory mechanisms on the protein level take place. Here, mechanisms like post-translational modifications and protein degradation are known, e.g. the human LOX enzyme is translated as inactive preproprotein and needs to be cleaved in order to obtain its functionality by crosslinking two of its residues yielding an essential cofactor (Lucero and Kagan 2006). However, a mechanism called allostery is likely the most interesting one. Allostery describes the binding of a ligand to a specific site of a protein that results in functional, structural or dynamical alterations at a distant site (Monod et al. 1963). The specific mechanisms that lead to signal propagation between distant sites are uncharacterized for most biological systems. Yet allostery is an intense field of research, due to its role in disease and drug discovery (Nussinov and Tsai 2013, Lisi and Loria 2017). This makes the elucidation of molecular mechanisms, which facilitate allostery, a task of major relevance. Since allostery manifests in functional, structural or dynamical changes, a broad spectrum of methods is employed. For the identification of specific structural states that might occur during allosteric communication, high-resolution structures were solved by x-ray crystallography (Niks et al. 2013). Since allosteric communication cannot just be reduced on static structures, NMR studies are performed to reveal structural and dynamically changes in solution (Lisi and Loria 2016). Such experiments can be complemented with *in silico* methods, including molecular dynamics (MD) simulations (Liang et al. 2019). In addition, the identification of allosteric positions, which modulate allosteric

communication, has proven to facilitate the understanding of allosteric mechanisms (List et al. 2012, Lisi and Loria 2017).

A popular model system for allosteric communication is the tryptophan synthase complex (TS) (Miles 1991, Dunn et al. 2008, Raboni et al. 2009, Dunn 2012). Despite a long-standing history of extensive research, TS is still the focus of ongoing investigation. On the one hand, the TS could serve as potential drug target. In a recent study, a small-molecule allosteric inhibitor for the TS of *Mycobacterium tuberculosis* was presented as potential antibiotic against the pathogen that causes tuberculosis. This is of high interest, since *M. tuberculosis* displays multidrug resistance, while surpassing HIV as world's leading cause of death due to infectious disease (Wellington et al. 2017). On the other hand, a subunit of the TS complex proved to be a valuable catalyst for synthesizing precursors for potential agrochemical or pharmaceutical useful compounds (Herger et al. 2016, Romney et al. 2017, Blei et al. 2018, Boville et al. 2018). This research was accompanied by studies on allosteric activation of the catalyst (Buller et al. 2015, Murciano-Calles et al. 2016, Buller et al. 2018). These studies in addition with a recent study that illuminated the evolutionary history of TS (Busch et al. 2016) sparked the light for the following work on allosteric interactions within the TS.

1.2 The Tryptophan Synthase as Model System for Allosteric Communication

The TS is a hetero-tetrameric complex consisting of a central dimer of β-subunits (TrpB) and one α-subunit (TrpA) on each side (**Figure 1A**).

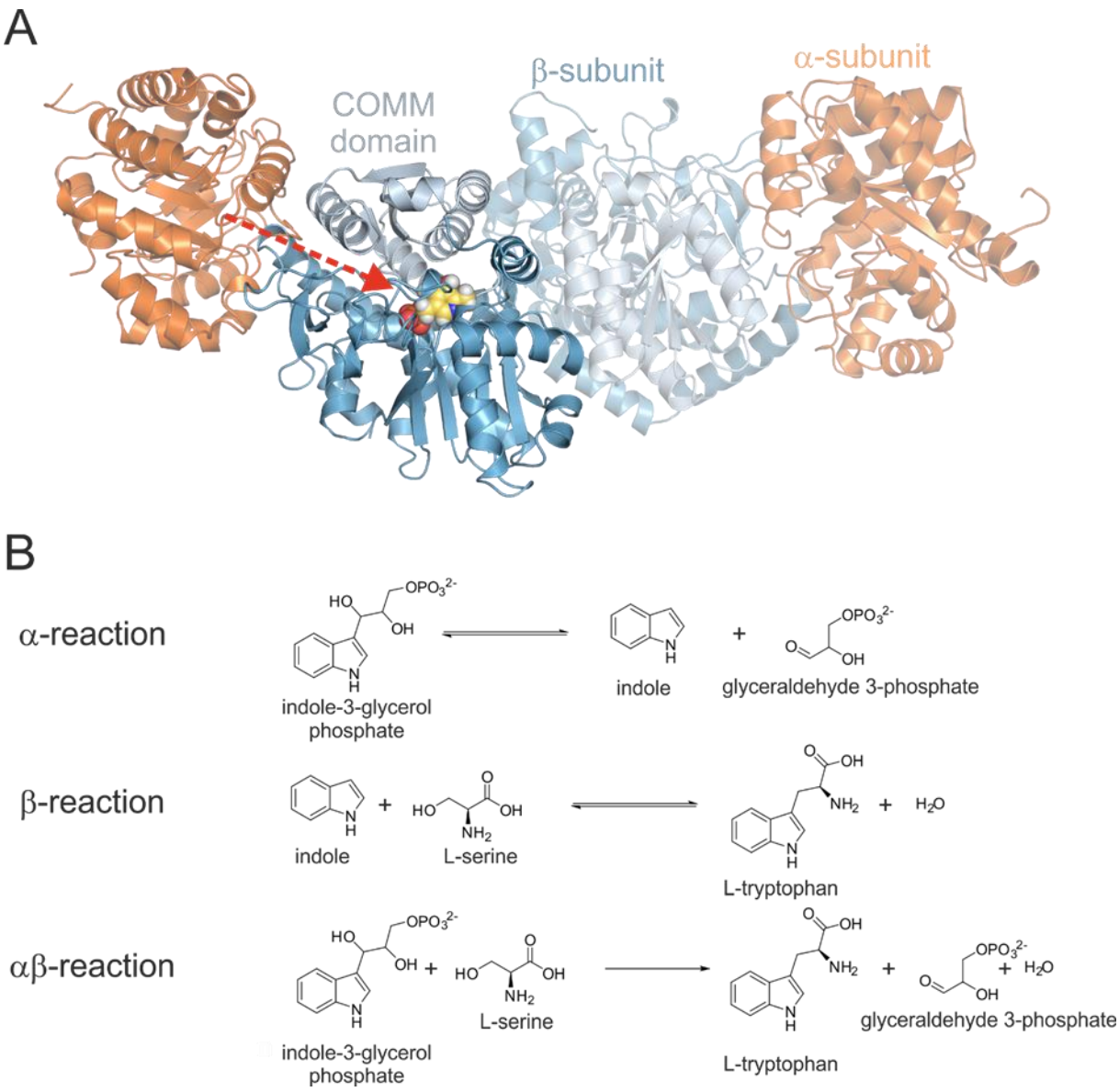


Figure 1: Structure of the tryptophan synthase and its catalyzed reactions.

(**A**) Structural model of the αββα-complex. The α-subunits are colored in orange, the β-subunits are colored in blue. The COMM domain within the β-subunits is colored in light blue. The active site of the β-subunit is indicated by its PLP cofactor depicted in yellow spheres. The red arrow indicates the interconnecting channel from the active site of the α-subunit to the active site of the β-subunit. (**B**) Reactions catalyzed by the α- and β-subunit, and combined αβ-reaction of both subunits.

Within this linear αββα complex, the last two steps of tryptophan biosynthesis are performed (Yanofsky and Rachmeler 1958). The α-subunit catalyzes the aldolytic cleavage of indole-3-glycerol phosphate (IGP) into indole and glyceraldehyde 3-phosphate (GAP) (**Figure 1B**). The intermediate indole is channeled from the α-subunit active site via 25 Å-long tunnel to the active site of the β-subunit (Hyde et al. 1988, Miles et al. 1999) where it reacts mediated by pyridoxal-5'-phophate (PLP) with L-serine yielding the final product L-tryptophan (**Figure** 1**B**) (Dunn 2012). In order to coordinate such a complex system of two reactions both subunits of the TS have to be regulated in an allosteric manner.

1.2.1 The α-Subunit

The active site of the α-subunit is constituted of the two catalytic residues E49 and D60 (numbering from *Salmonella typhimurium*) (Yutani et al. 1987, Nagata et al. 1989). Within the second shell, a residue of high relevance is T183 in Loop6. The side chain of this residue forms an H-bond that orients the catalytic residue D60 (Kulik et al. 2002, Axe et al. 2015) (**Figure 2A**). For the reaction of the α-subunit a step-wise and a concerted mechanism have been proposed (**Figure 2B**).

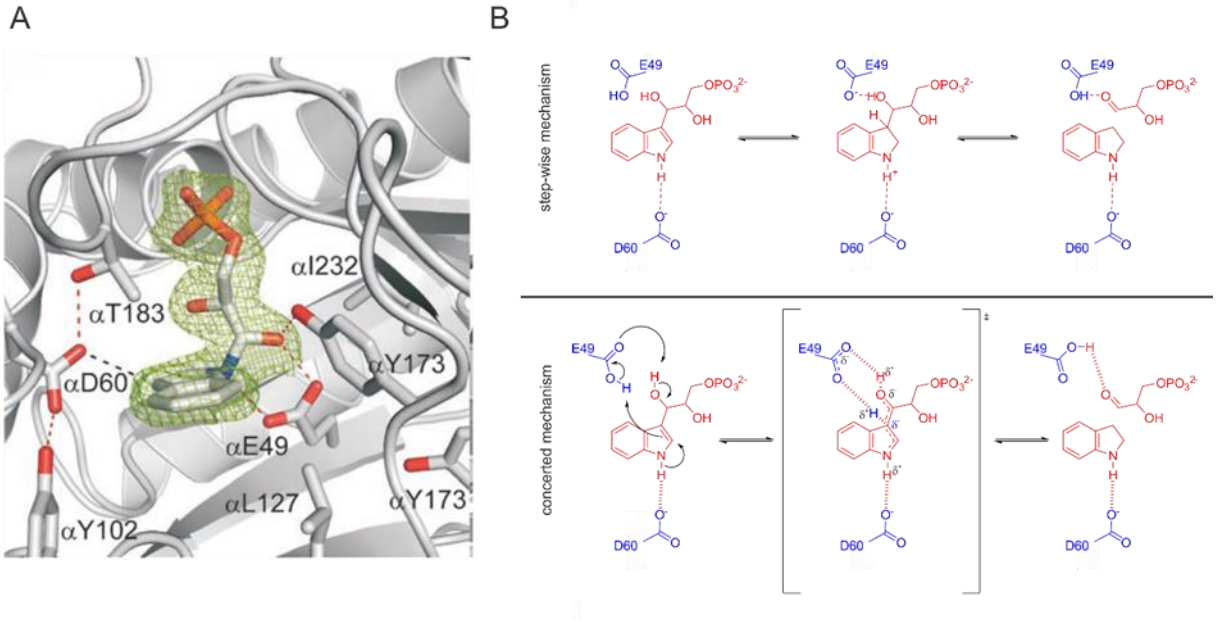


Figure 2: Active site structure and proposed reaction mechanisms of the α-reaction.

(**A**) Active site of the α-subunit from *S. typhimurium* with indoline-G3P adduct as IGP analogue. Figure modified from (Barends et al. 2008). (**B**) Postulated mechanisms of α-reaction. Reaction intermediates are colored in red, catalytic residues are colored in blue. Figure modified from (Dunn 2012). Residue numbering is according to the α-subunit of *S. typhimurium*.

Within both mechanisms, the strictly conserved residue E49 (numbering according to *S. typhimurium*) acts as acid-base catalyst for proton transfers that lead to the cleavage of the indole ring system from the glycerol-3-phosphate chain. The strictly conserved residue D60, which is part of Loop2, is responsible in both proposed mechanisms for charge stabilization at the N-atom of the indole ring system (Dunn 2012). If the product indole or GAP is not removed, the reaction equilibrium is on the side of IGP (Weischet and Kirschner 1976). The rate constants of the single steps were determined by employing quenched-flow and stopped-flow methods (Anderson et al. 1991). The reaction velocity is drastically improved by complex formation of the α-subunit with the β-subunit (Hettwer and Sterner 2002, Kriechbaumer et al. 2008). The interaction networks at the active site during catalysis are topic of ongoing research (O'Rourke et al. 2018, O'Rourke et al. 2019).

1.2.2 The β-Subunit

The reaction of the β-subunit contains multiple intermediates (**Figure 3**).

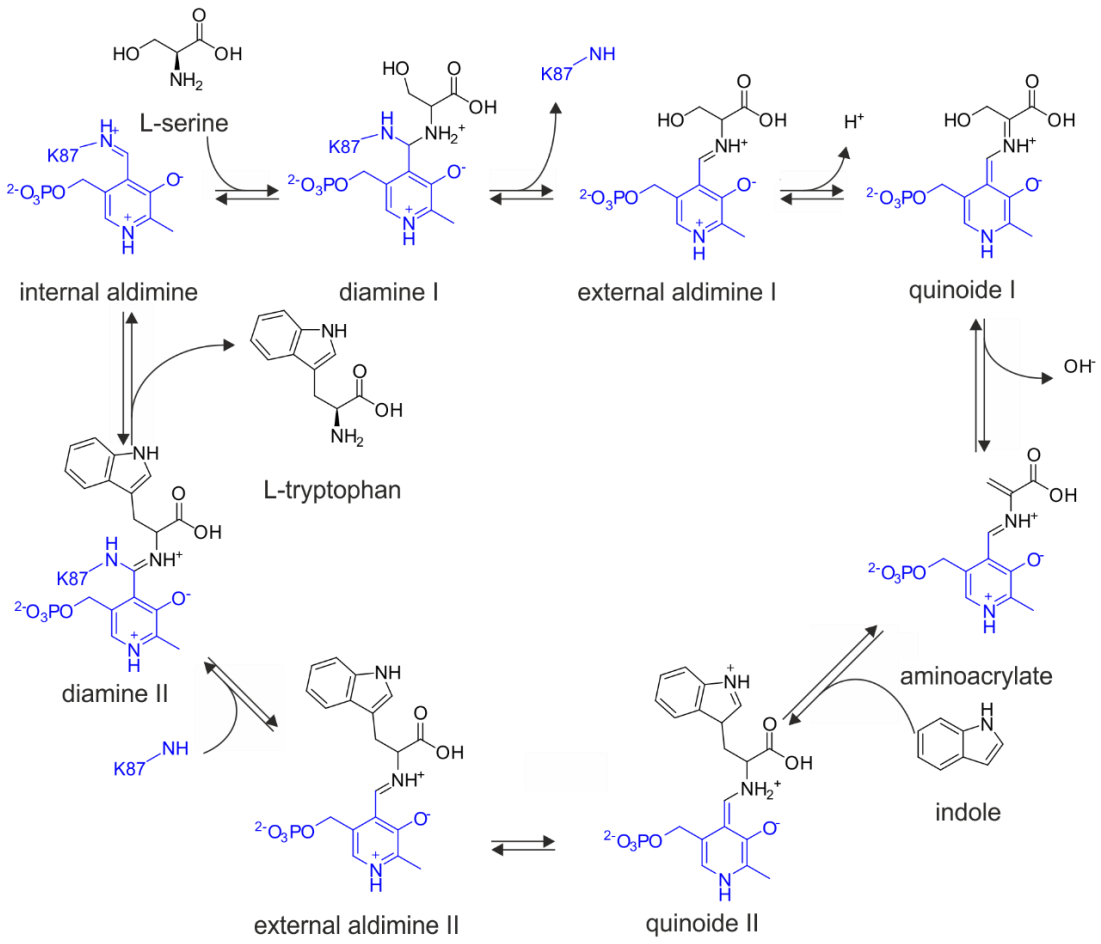


Figure 3: Reaction cycle of the β-reaction.

The PLP cofactor and the catalytic lysine of the β-subunit (K87) are colored in blue. Figure modified from (Schiaretti et al. 2004, Busch 2015).

The reaction cycle starts with an internal aldimine formed by the PLP cofactor with the side chain of the strictly conserved K87 (numbering according to *S. typhimurium*). The binding of the amino moiety of L-serine leads to the diamine I intermediate. With K87 leaving the intermediate, the external aldimine I with L-serine bound to PLP is formed. A dehydratization step leads via quinoide I to the reactive intermediate aminoacrylate. The nucleophilic attack of indole at the aminoacrylate leads to quinoide II that is further transformed into the external aldimine II where L-tryptophan is bound to PLP. In a next

step, K87 binds to PLP forming the diamine II intermediate, L-tryptophan is replaced by K87 and subsequently released from the active site.

Each of the shown intermediates can be observed spectroscopically at different wavelengths (Dunn 2012). This allowed for the determination of the rate constants for each reaction step (Anderson et al. 1991). The equilibrium between external aldimine I and the reactive aminoacrylate is regulated by various allosteric regulators, like complex formation with the α-subunit, ligands at the bound α-subunit and monovalent cations bound to the β-subunit (Ferrari et al. 2001, Hettwer and Sterner 2002, Ngo et al. 2007, Dierkers et al. 2009).

1.2.3 The Allosteric Communication Within Tryptophan Synthase Complex

The monomeric α-subunit and the homo-dimeric β-subunits are active enzymes, but display rather low catalytic activities without their respective interaction partner (Hettwer and Sterner 2002). In order to synchronize the coupled αβ-reaction (Figure 1B) mutual allosteric activation of the α- and β-subunits takes place (Brzovic et al. 1992). Within the formed complex, both the α- and β-subunit display increased catalytic activities (Hettwer and Sterner 2002) that are accompanied with conformational transitions in both subunits. The α-subunit as well as the β-subunit switch from an open state with low activity to a closed state with high activity (Ngo et al. 2007) (**Figure 4A**). While the closed state of the α-subunit is defined by closure of Loop2 and Loop6, the closed state of the β-subunit is characterized by closure of the COMM domain (Dunn et al. 2008) (**Figure 4B, C**)

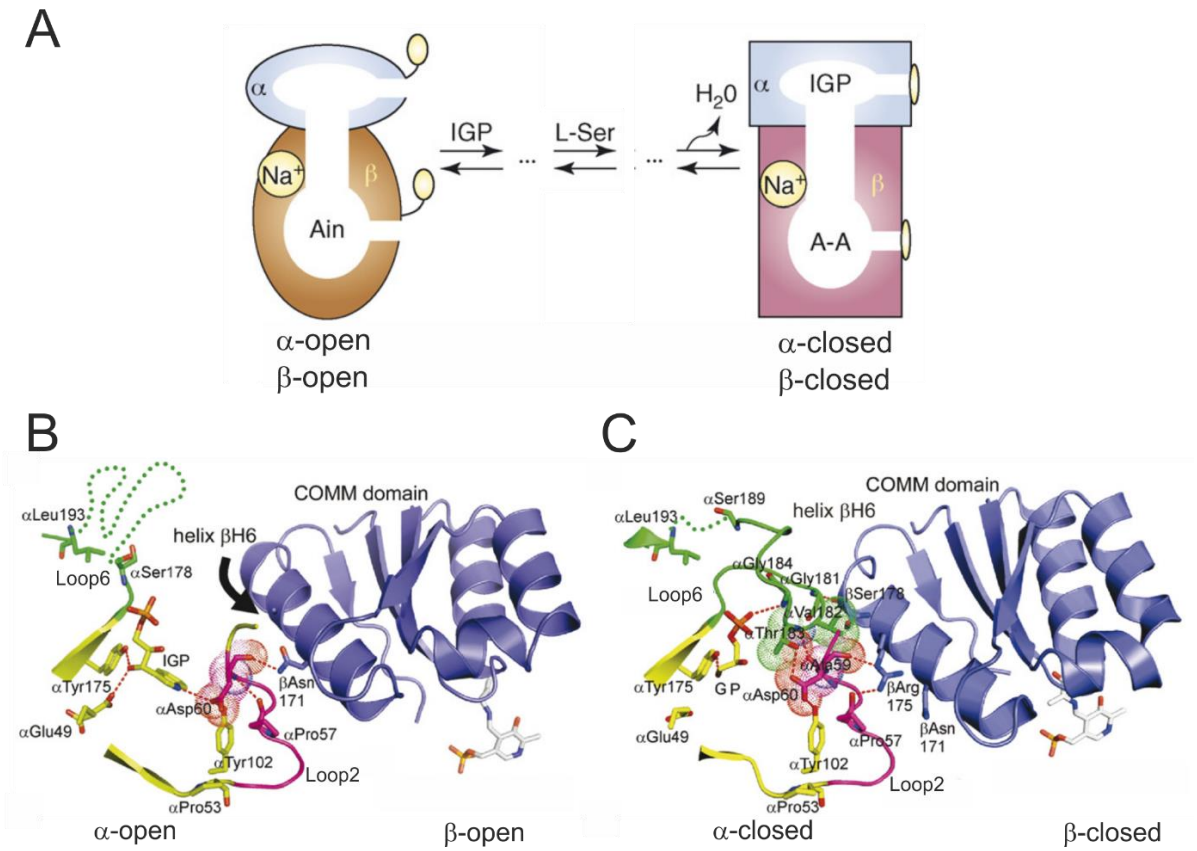


Figure 4: Open and closed state of α- and β-subunit.

(**A**) Scheme of a protomer of TS; binding of IGP to the α-subunit and L-serine to the β-subunit leads to subsequent formation of the aminoacrylate (A-A) inducing the closed state of both subunits. (**B**) Structural details of the open state of α- and β-subunits. The open state of the α-subunit is characterized by a disordered Loop6 that is colored in green and represented by the dotted line connecting Ser178 and Leu193. For the β-subunit, the black arrow indicates the movement of the COMM domain relative to the α-subunit during transition from open state to closed state. (**C**) Structural details of the closed state of α- and β-subunit. The closed state of the α-subunit is defined by an ordered and closed Loop6 with an H-bond formed between Thr183 (located in Loop6) and the catalytic residue Asp60 (located in Loop2 that is colored in pink). For the β-subunit the COMM domain moved closer to the active site indicated by the PLP cofactor (light grey). Figure modified from (Dunn et al. 2008).

1.3 Aims of This Work

In order to deepen our understanding of the mutual allosteric communication within TS, recently a “stand-alone” variant of the β-subunit was generated by directed laboratory evolution. This activated β-subunit in isolation showed catalytic properties comparable to a wild-type β-subunit in presence of a α-subunit. The activating mutations within the β-subunit were identified by random mutagenesis and subsequent screening (Buller et al. 2015). On the other hand, a “stand-alone” variant of the α-subunit, which does not require a β-subunit for allosteric activation, did not exist prior to this work. Since there is no established cell-based selection system for the α-reaction available and screening of variants for enhanced α-activity would be extremely time-consuming, a directed evolution approach like the one conducted for creating the stand-alone β-subunit (Buller et al. 2015) is not feasible.

Therefore, another approach had to be followed for the generation of a stand-alone α-subunit. This approach was based on the comparison of the α-subunit with the homologous BX1 protein, which catalyzes the same IGP cleavage reaction but is highly active in absence of a β-subunit-like partner protein (Frey et al. 2000). Hence, in the present work BX1 should be exploited as an evolutionary blueprint for the creation of a α-subunit that displays enhanced activity in absence of a β-subunit (**Chapter 2**) Such a procedure, which aims at the identification of crucial residue differences between extant homologous proteins, has been termed “horizontal” approach (**Figure 5A**).

Still another approach was followed to identify residues that are crucial for the allosteric activation of the β-subunit by the α-subunit. This approach is based on ancestral sequence reconstruction (ASR), which allows one to calculate the sequences of extinct proteins on the basis of a multiple sequence alignment (MSA) of extant homologues and a phylogenetic model (Merkl and Sterner 2016). Previously, the sequences of the α- and the β-subunits of the last bacterial common ancestor (LBCA) and of all intermediates connecting LBCA with the extant α- and β-subunits, were determined by ASR (Straub and Merkl 2019). The reconstructed LBCA α- and β-subunits were produced and experimentally characterized. LBCA-TS displayed similar properties as the extant TS from *Escherichia coli* or *S. typhimurium*, with one crucial exception. Unlike the above-mentioned mutual activation of the α- and β-subunit, only the activation of the LBCA α-subunit by the β-subunit was observed. In contrast, the LBCA β-subunit was deactivated by the α-subunit (Busch et al. 2016). The sequence basis for this functional difference

should be identified by the characterization of all evolutionary TS-intermediates linking LBCA-TS with an extant TS (**Chapter 3**). Such a procedure, which aims at the identification of crucial residue differences between extinct and extant homologous proteins, has been termed “vertical” approach (**Figure 5C**).

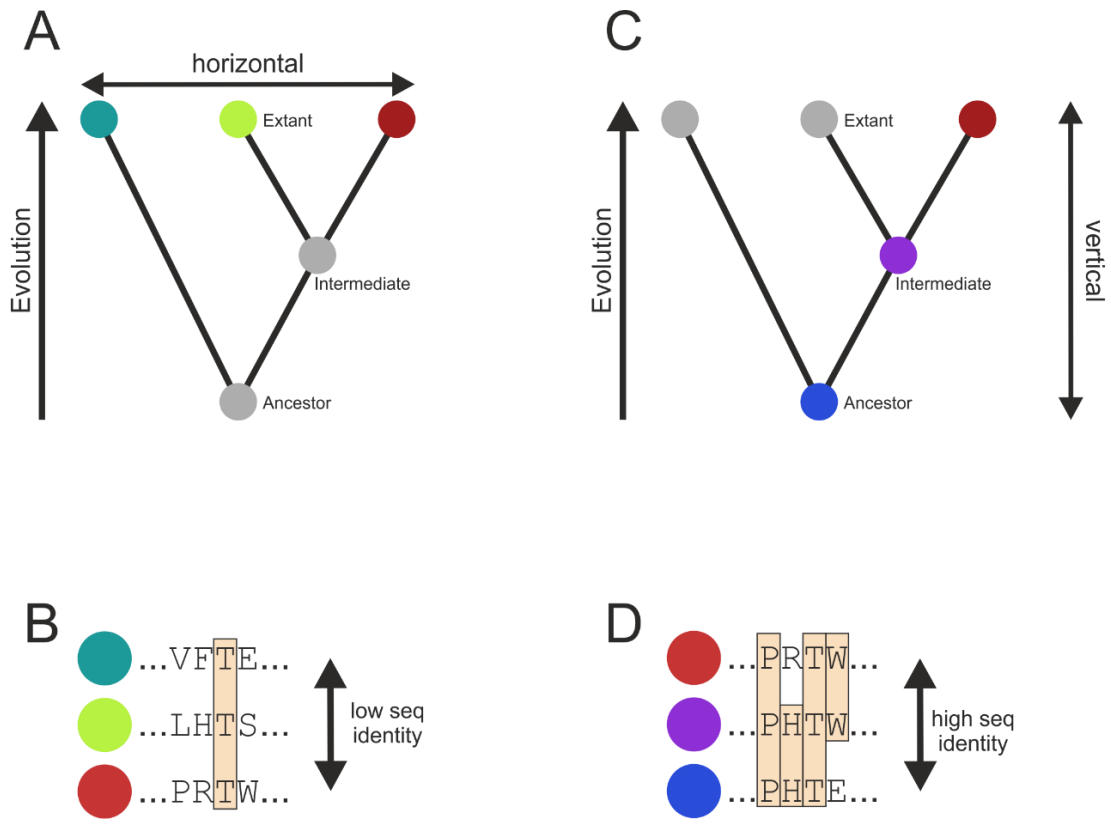


Figure 5: Schemes of horizontal and vertical approaches.

(**A**, **C**) Schematic phylogenetic trees containing homologous proteins with different sequences and functional characteristics (circles). (**B**, **D**) Exemplary multiple sequence alignments (MSA) with identical amino acids indicated with a beige colored background. (**A**, **B**) In a horizontal approach, extant homologous proteins are compared in order to identify residues that cause functional differences. (**C**, **D**) In a vertical approach, extant homologous proteins are compared with extinct ones in order to identify residues that cause functional differences. (**B**, **D**) Extant sequences are more diverse among each other than extant and extinct sequences. As a consequence, a horizontal approach is not as straightforward for sequence-function correlations as a vertical approach.

1.4 Guide to the Following Chapters

Each of the following two chapters corresponds to a manuscript.

The published manuscript ***Generation of a stand-alone tryptophan synthase a-subunit by mimicking an evolutionary blueprint*** reports on a protein engineering approach for the activation of the α-subunit of TS from *Zea mays*. The generated zmTrpA variants became partially independent of the presence of zmTrpB.

The manuscript ***Analysis of allosteric communication in a multi-enzyme complex by ancestral sequence reconstruction*** contains unpublished data that are in preparation for submission. Here an evolutionary approach was followed to identify residues in TrpB that are crucial for its activation by TrpA.

2 Generation of a Stand-Alone Tryptophan Synthase α-Subunit by Mimicking an Evolutionary Blueprint

Michael Schupfner, Florian Busch, Vicki H. Wysocki, and Reinhard Sterner

ChemBioChem 10.1002/cbic.201900323

Key words: enzyme, protein design, molecular evolution, primary metabolism, secondary metabolism

2.1 Abstract

The αββα tryptophan synthase (TS), which is part of primary metabolism, is a paradigm for allosteric communication in multi-enzyme complexes. In particular, the intrinsically low catalytic activity of the α-subunit TrpA is stimulated several hundred-fold by the interaction with the β-subunit TrpB1. The BX1 protein from *Zea mays* (zmBX1), which is part of secondary metabolism, catalyzes the same reaction as its homologue TrpA, but with high activity in the absence of an interaction partner. We found that the intrinsic activity of TrpA can be significantly increased by the exchange of several active site loop residues, mimicking the corresponding loop in zmBX1. The subsequent identification of activating amino acids in the generated “stand-alone” TrpA contributes to the understanding of allostery in TS. Moreover, our findings suggest an evolutionary trajectory that describes the transition from a primary metabolic enzyme regulated by an interaction partner to a self-reliant stand-alone secondary metabolic enzyme.

2.2 Introduction, Results and Discussion

Allosteric communication is a central mechanism for the regulation of protein-based biological systems. A well-characterized model system for studies on allosteric communication within enzymes is the tryptophan synthase (TS). The TS, an $\alpha\beta\beta\alpha$ hetero-tetrameric complex that is crucial for primary metabolism in archaea, bacteria, and plants (Crawford 1975, Merkl 2007), consists of a central dimer of β-subunits (TrpB1) and two peripheral α-subunits (TrpA). TrpA catalyzes the cleavage of indole-3-glycerol phosphate (IGP) into glyceraldehyde 3-phosphate (GAP) and indole. The latter is channeled to the active site of TrpB1 where the cofactor pyridoxal-5'-phosphate (PLP) facilitates its condensation with L-serine to the final product L-tryptophan (Dunn 2012). TrpA and TrpB1 mutually stimulate each other (**Figure 6a**) but the underlying allosteric pathways are not fully understood. Recently, a partial comprehension of the activation of TrpB1 by TrpA has been achieved by the directed evolution towards a "stand-alone" TrpB1 subunit, which contains amino acid exchanges that mimic its stimulation by TrpA in wild-type TS (Buller et al. 2015).

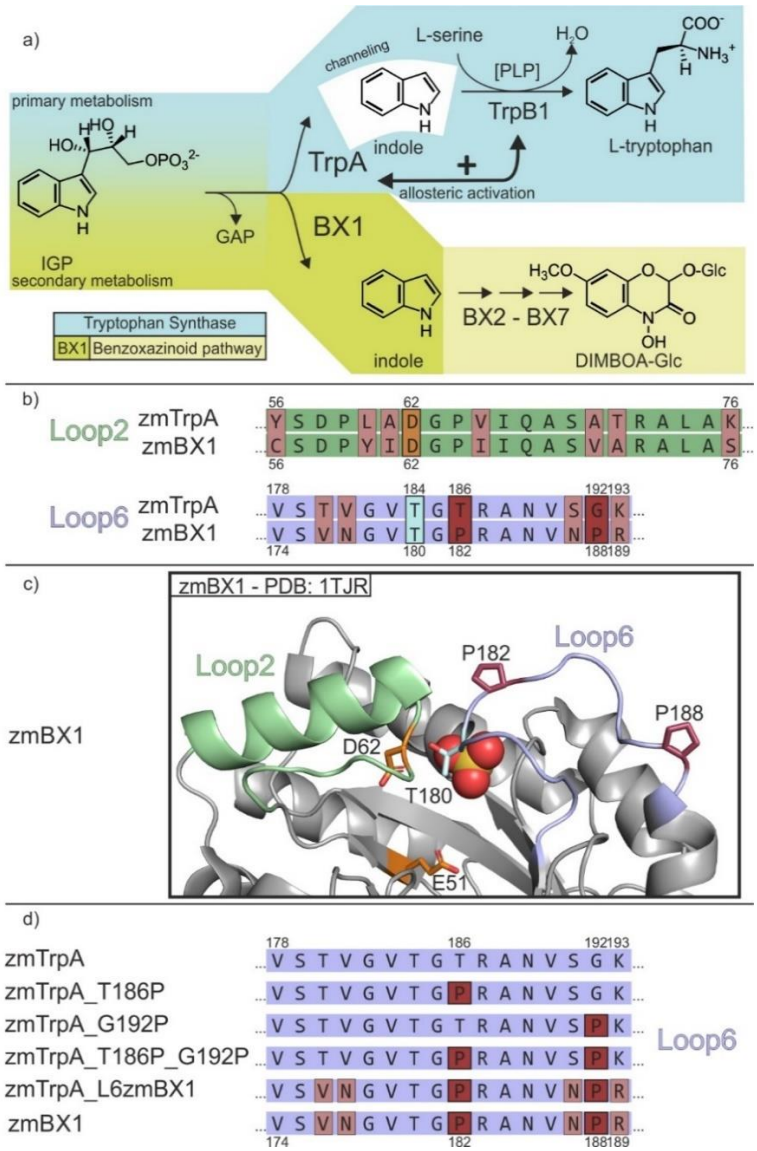


Figure 6: Reaction of TS and BX1, sequence comparison of zmTrpA and zmBX1, structural detail of BX1 active site and Loop6 sequence of studied variants

a) The cleavage of IGP by the paralogous enzymes TrpA and BX1 in *Zea mays* marks a branch point between primary (tryptophan biosynthesis) and secondary metabolism (benzoxazinoid biosynthesis). In primary metabolism, TrpA is part of the tryptophan synthase and is allosterically activated by the TrpB1 subunit. In secondary metabolism, BX1 is independent of an activating interaction partner (Frey et al. 2009). **b**) Sequence comparison of Loop2 and Loop6 from zmTrpA and zmBX1. In Loop2 the catalytic residue D62 is colored orange. In Loop6 the conserved T184 is colored cyan. The residues that differ between zmTrpA and zmBX1 are colored light red. The positions that carry a proline in zmBX1 are colored dark red. The entire sequences of zmTrpA and zmBX1 are shown in **Figure 9**. **c**) Crystal structure of zmBX1 (PDB ID 1TJR, chain A). Loop2, consisting of a loop segment and α helical segment, is colored green and Loop6 is colored light blue. The proline residues in Loop 6 at positions 182 and 188 are colored dark red. The catalytic residues E51 and D62 (Kulik et al. 2005) are colored orange. The conserved T180 is colored cyan. A sulfate ion, which is represented as sphere, marks the phosphate-binding pocket at the active site. **d**) Loop6 sequences of the experimentally characterized enzymes. For the zmTrpA variants, differences to wild-type zmTrpA are colored dark red (substitutions to proline) or light red (other substitutions).

Interestingly, a blueprint for a “stand-alone” TrpA protein already exists in nature. This protein is named BX1 and is a paralogue of TrpA that plays a role within the secondary metabolism of, for example, plants like *Zea mays*. BX1 shares with TrpA the ubiquitously encountered $(\beta\alpha)_8$-barrel fold (Sterner and Höcker 2005) and the enzyme from *Z. mays* (zmBX1) shares a sequence identity of 63.3% with zmTrpA (**Figure 9**). Remarkably, BX1 catalyzes the same IGP lyase reaction as TrpA (**Figure 6a**) using the identical active site residues (Kulik et al. 2005), but with high efficiency in the absence of an interaction partner (Frey et al. 2000, Kriechbaumer et al. 2008). Indole generated by BX1 is not used for tryptophan biosynthesis but is fed into the benzoxazinoid pathway where it is finally converted to the plant protective agent DIMBOA-Glc (Frey et al. 2009).

It has been postulated that enzymes of secondary metabolism have evolved by gene duplication from enzymes of primary metabolism (Vining 1992, Frey et al. 1997, Firn and Jones 2000). In line with this hypothesis, the high sequence identity as well as structural and functional similarities between the two IGP lyases suggest that zmTrpA is the progenitor of zmBX1. Based on this hypothesis, we reasoned that it should be possible to remove the zmTrpB1-dependence of zmTrpA by amino acid exchanges that mimic the situation in zmBX1. The resulting “stand-alone” zmTrpA variant should provide insights into the allosteric communication of the TS complex and into the evolutionary relationship between zmTrpA and zmBX1. In order to identify crucial differences between zmTrpA and zmBX1, we compared the amino acid sequences of the central catalytic $(\beta\alpha)_8$-barrel domains, without considering the N-terminal subcellular localization sequences. An initial analysis showed that the differences in amino acid sequence were not clustered but distributed over the whole proteins (**Figure 9**). We then focused on sequence stretches from Loop2 (residues 56-76) and Loop6 (residues 174-189 for zmBX1 and residues 178-193 for zmTrpA) (**Figure 6b**), because these two loop regions are known to be important for the catalytic activity and the allosteric activation of TrpA (Dunn 2012, Axe et al. 2015). Moreover, both loops together form a lid covering the active site (**Figure 6c**). Amino acids in Loop2, which carries one of the two conserved catalytic residues (D62), differ in several positions between zmTrpA und zmBX1. However, these differences seem to be rather insignificant with respect to the physico-chemical properties of the corresponding residues (i.e. A61 and V65 in zmTrpA vs. I61 and I65 in zmBX1) or are rather remote from the active site (i.e. Y56 and K76 in zmTrpA vs. C56 and S76 in zmBX1). Loop6 contains a conserved threonine residue (T184 in zmTrpA, T180 in zmBX1), which is known to form a crucial hydrogen bond with D62 (Yang and Miles 1992, Kulik et al. 2002, Axe et al. 2015). There are six positions in Loop6

that differ between zmTrpA and zmBX1. The two most prominent ones are T186 in zmTrpA vs. P182 in zmBX1 as well as G192 in zmTrpA vs. P188 in zmBX1. In addition, the crystal structure of zmBX1 (PDB ID 1TJR, chain A) shows a fully resolved Loop6. In contrast, crystal structures of TrpA enzymes display a poorly defined Loop6, both in the presence and in the absence of TrpB1. In light of these findings, it has been argued that the higher IGP lyase activity of zmBX1 compared to TrpA is based on the stabilization of a specific conformation of Loop6 (Kulik et al. 2005). We speculated that the two proline residues in zmBX1 are responsible for this stabilization and hence for the increased catalytic activity of zmBX1 compared to zmTrpA. In order to test this hypothesis, we generated in zmTrpA the T186P and G192P exchanges, separately and in combination. In addition, we replaced the entire Loop6 from zmTrpA with Loop6 from zmBX1 (**Figure 6d**).

Wild-type and mutant genes were cloned (see **Table 2** for mutagenesis primers) and expressed in *Escherichia coli*. The recombinant proteins (**Table 3**) were purified and analyzed by size-exclusion chromatography (SEC) and native mass spectrometry (MS). In SEC experiments, each of the proteins eluted as a single peak indicating a well-defined oligomeric state (**Figure 10**). MS showed that zmTrpA and its variants are monomers, and that zmBX1 and zmTrpB1 are dimers (**Figure 11, Figure 12**). Steady-state enzyme kinetics were recorded for the enzymes zmBX1, zmTrpA, zmTrpA_T186P, zmTrpA_G192P, zmTrpA_T186P_G192P, and zmTrpA_L6zmBX1 (**Table 1**). The determined turnover numbers (k_{cat}) and Michaelis constants for IGP (K_m^{IGP}) of zmBX1 and zmTrpA are comparable to reported values (Frey et al. 1997, Kriechbaumer et al. 2008) with zmBX1 exhibiting a k_{cat} (2.8 s^{-1}) that exceeds the k_{cat} of zmTrpA by three orders of magnitude. In comparison to zmTrpA, the k_{cat}-values of the single mutants zmTrpA_T186P and zmTrpA_G192P were enhanced 10-15 fold. The k_{cat}-value of the double mutant zmTrpA_T186P_G192P was enhanced 40-fold, while the k_{cat}-value of zmTrpA_L6zmBX1 was enhanced 520-fold. Thus, the k_{cat}-value of zmTrpA_L6zmBX1 is close to those of zmBX1 as well as zmTrpA in complex with its activating binding partner zmTrpB1. While all zmTrpA variants show a significantly enhanced k_{cat}-value, they also display an increased K_m^{IGP} value compared to zmBX1 and zmTrpA (**Table 1**). These results demonstrate that the exchanges T186P and G192P as well as the incorporation of the whole Loop6 of zmBX1 are sufficient to accelerate turnover rates for zmTrpA, albeit negatively affect the binding of the substrate IGP.

Table 1: Steady-state enzyme kinetic parameters for the TrpA reaction of zmBX1 and zmTrpA and its variants, in absence and presence of zmTrpB1

Protein(s)	k_{cat} / s^{-1}	K_m^{IGP} / mM	k_{cat}/K_m^{IGP} / $M^{-1}s^{-1}$
zmBX1	2.8 ± 0.1	0.005 ± 0.001	$5.4\times10^5 \pm 9.4\times10^4$
zmTrpA	0.0020 ± 0.0001	0.170 ± 0.025	$1.2\times10^1 \pm 2.2\times10^0$
zmTrpA + zmTrpB1	3.9 ± 0.1	0.175 ± 0.014	$2.2\times10^4 \pm 2.2\times10^3$
zmTrpA_T186P	0.021 ± 0.001	3.2 ± 0.3	$6.5\times10^0 \pm 8.0\times10^{-1}$
zmTrpA_T186P + zmTrpB1	2.3 ± 0.1	0.043 ± 0.005	$5.5\times10^4 \pm 7.6\times10^3$
zmTrpA_G192P	0.031 ± 0.004	4.0 ± 1.2	$7.7\times10^0 \pm 3.3\times10^0$
zmTrpA_G192P + zmTrpB1	1.34 ± 0.04	0.051 ± 0.006	$2.6\times10^4 \pm 4.0\times10^3$
zmTrpA_T186P_G192P	0.08 ± 0.01	4.01±0.75	$2.0\times10^1 \pm 5.2\times10^0$
zmTrpA_T186P_G192P + zmTrpB1	2.4 ± 0.1	0.036 ± 0.006	$6.8\times10^4 \pm 1.4\times10^4$
zmTrpA_L6zmBX1	1.04 ± 0.05	1.8 ± 0.3	$5.7\times10^2 \pm 1.2\times10^2$
zmTrpA_L6zmBX1 +zmTrpB1	0.46 ± 0.01	0.080 ± 0.006	$5.8\times10^3 \pm 5.8\times10^2$
zmTrpA_T186M	0.057 ± 0.001	3.9 ± 0.2	$1.5\times10^1 \pm 1.1\times10^0$
zmTrpA_T186I	0.024 ± 0.001	3.0 ± 0.3	$8.1\times10^0 \pm 1.1\times10^0$

Experimental conditions: 100 mM EPPS/KOH (pH 7.5), 180 mM KCl, 40 mM PLP, 6 mM NAD^+, 20 mM $NaAsO_4$, varying concentrations of IGP, 100 mM L-serine (in presence of zmTrpB1), and 5.5 µM GAP dehydrogenase from *Thermotoga maritima*. The reaction was followed at 30 °C by monitoring the cleavage of IGP to GAP and indole with a coupled enzymatic assay (Creighton 1970). The mean and the standard error were calculated from at least three independent measurements.

We next used SEC and activity titrations to test whether the activated zmTrpA variants are still able to bind zmTrpB1. The results showed that all variants form a complex with zmTrpB1 (**Figure 13**). The apparent dissociation constants (K_D^{app}) as determined by measuring TrpA activity as a function of zmTrpB1 concentration are between 0.3 µM and 0.7 µM, comparable to the K_D^{app} value of 0.5 µM for the zmTrpA:zmTrpB1 complex (**Table 4**). The effect of zmTrpB1 binding on the steady-state kinetic parameters of zmTrpA and its activated variants are summarized in **Table 1**. Wild-type zmTrpA shows a 2000-fold increased k_{cat}-value in the presence of zmTrpB1 while the K_m^{IGP}-value is unaffected. The variants zmTrpA_T186P, zmTrpA_G192P, and zmTrpA_T186P_G192P display 30-100-fold improved k_{cat}-values in the presence of zmTrpB1 while the K_m^{IGP}-values are decreased 70-110-fold. Interestingly, the variant zmTrpA_L6zmBX1 displays a two-fold reduced k_{cat}-value while the K_m^{IGP}-value is lowered 20-fold. Apparently, an optimal catalytic activity for zmTrpA is achieved by the exchange of its Loop6 with that of zmBX1 and cannot be further increased by the presence of zmTrpB1. It has been shown for TrpA from *Salmonella typhimurium* that the transition from a rather inactive to a highly active conformation is the rate-limiting step for the IGP lyase reaction (Anderson et al. 1991). This conformational change, which is triggered by the formation of the aminoacrylate intermediate in TrpB1, includes Loop6. It is therefore plausible to assume

that zmTrpA_L6zmBX1 has become independent of zmTrpB1, since Loop6 stemming from zmBX1 constitutively adopts the active conformation. In contrast, substrate binding as reflected by the K_m^{IGP}-value is impaired in the zmTrpA variants. It can, however, be restored to wildtype level by the interaction with zmTrpB1, for reasons that are unclear at this point.

The improved catalytic activities of the variants zmTrpA_T186P and zmTrpA_G192P raised the question whether residues other than proline have a similar activating effect at positions 186 and 192. We therefore tested, at both positions, a set of amino acids that sample a range of diverse physicochemical properties (Pommie et al. 2004). Mutagenesis at position 186 led to two variants, zmTrpA_T186M and zmTrpA_T186I, that display increased IGP lyase activity (**Figure 7**). Comparable to zmTrpA_T186P, zmTrpA_T186M and zmTrpA_T186I show a 10-30 fold enhanced k_{cat}-value, but also a 20-fold increased K_m^{IGP}-value compared to zmTrpA (**Table 1**). Analytical SEC and MS demonstrated that both variants are homogeneous monomers (**Figure 10, Figure 11**) that form a complex with zmTrpB1 (**Figure 13**). Screening of position 192 identified no activating amino acid exchanges other than G192P.

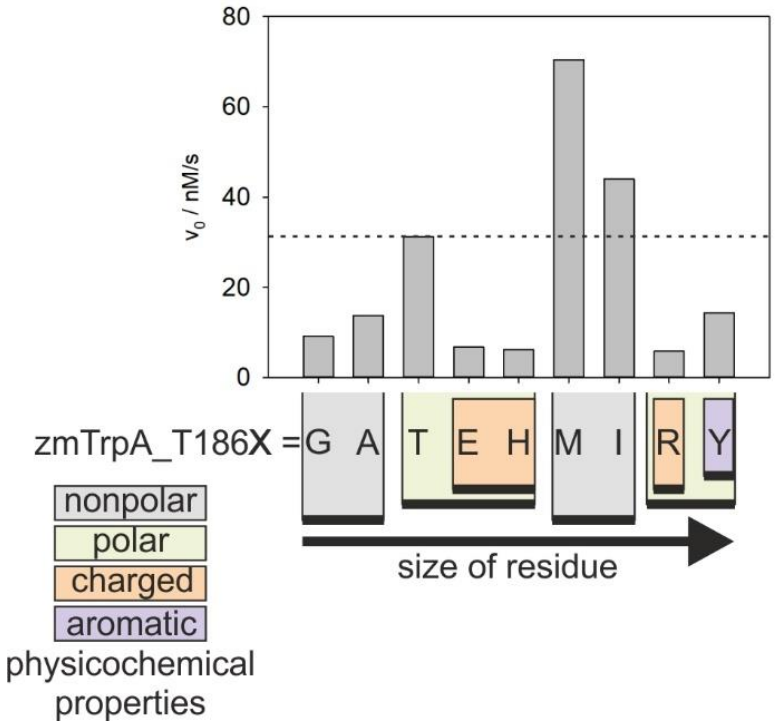


Figure 7: Reaction rates of zmTrpA variants with different amino acids at position 186.

The dashed line indicates wild-type activity. Experimental conditions: 100 mM EPPS/KOH (pH 7.5), 180 mM KCl, 40 mM PLP, 6 mM NAD^+, 20 mM NaAsO4, 5.5 µM GAP dehydrogenase from *T. maritima*, 5 µM zmTrpA variant. The reaction was followed at 30 °C by monitoring the cleavage of 1 mM IGP to GAP and indole with a coupled enzymatic assay (Creighton 1970).

Our work establishes that zmTrpA can be significantly activated by the exchange of two residues in Loop6 with the corresponding residues of zmBX1. This activating effect is further enhanced when the entire Loop6 of zmTrpA is replaced with Loop6 from zmBX1. Thus, our findings reinforce the importance of Loop6 for the allosteric activation of TrpA by TrpB1 (Dunn 2012). However, although the turnover number of zmTrpA_L6zmBX1 is close to that of zmBX1, the Michaelis constant for IGP is increased by several orders of magnitude. Currently, the structural basis for a weaker substrate affinity of the activated variants is unknown and might result from other sequence differences between zmTrpA and zmBX1 (Frey et al. 2000). Remarkably, high substrate affinity is restored by the interaction of zmTrpA variants with zmTrpB1, illustrating the decoupling of stand-alone activity and allosteric activation.

Based on our experimental findings, we propose a plausible model for the evolution of zmBX1 from a progenitor zmTrpA (**Figure 8**). In this model, gene duplication allows for the evolution of one copy towards zmBX1 while the other copy of zmTrpA maintains its role in the canonical TS. On the route towards zmBX1, initial single exchanges at key positions 186 and 192 in Loop6 would have increased the turnover number. The combination of such activating exchanges plus the accumulation of further substitutions would have led to a complete rearrangement of Loop6 with a further drastic increase of activity. For the final transition from zmTrpA to zmBX1, two major changes in properties were necessary: the improvement of affinity for IGP and the formation of the dimer. It is unclear whether these changes in properties occurred in a stepwise manner or along with the improvement of the k_{cat}-value. In any case, our model is in accordance with the situation found in extant *Zea mays*. Here, a zmTrpA enzyme is activated by zmTrpB1 in the canonical TS whereas a zmBX1 enzyme displays a high IGP lyase activity without an interaction partner. Taken together, our findings demonstrate that a small number of amino acid exchanges can be sufficient to drive the evolution of an enzyme from primary metabolism into one of secondary metabolism with altered catalytic and regulatory properties (Plach et al. 2015).

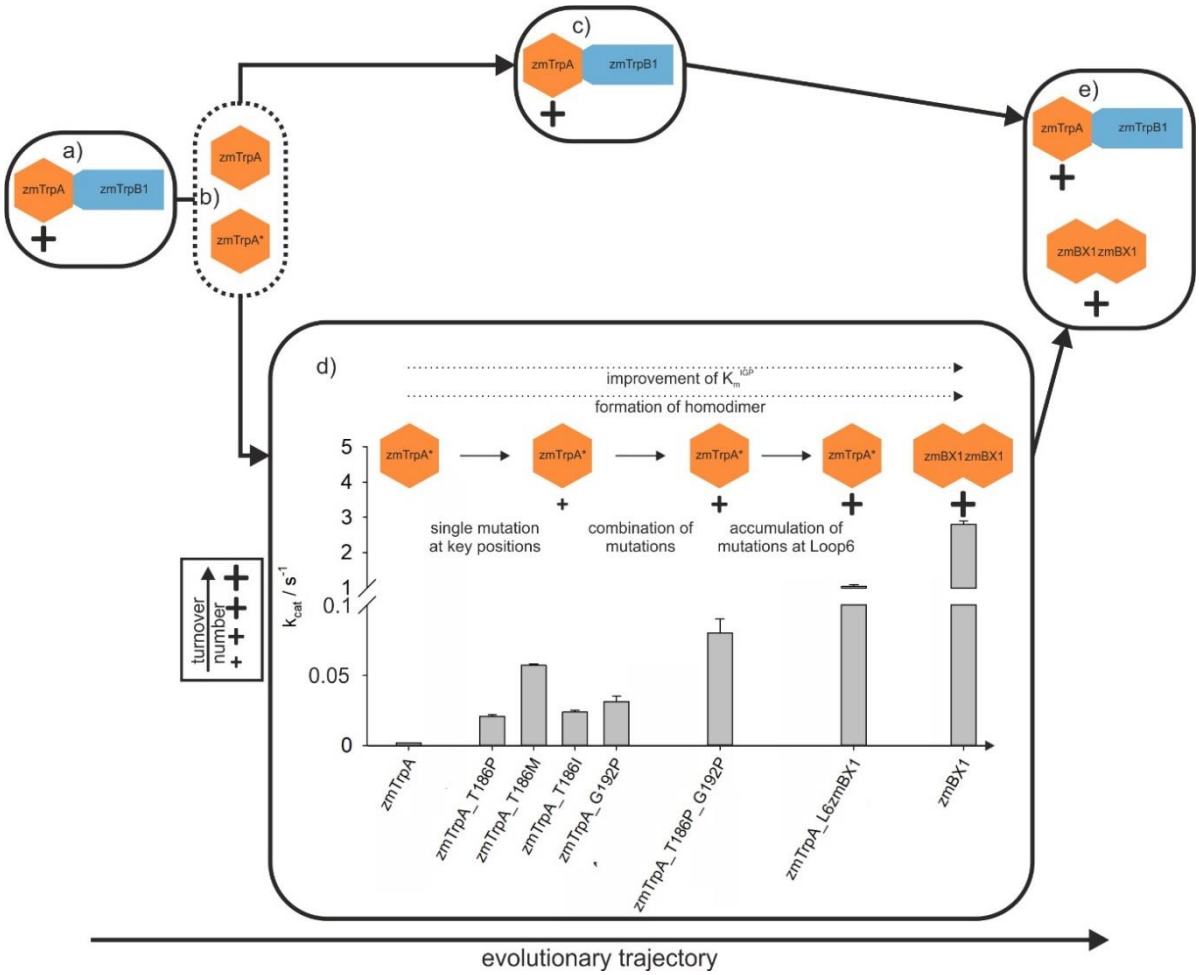


Figure 8: Model for the evolution of zmTrpA to zmBX1.

Model for the evolution of zmTrpA to zmBX1. The '+' sign size indicates the relative IGP lyase activity of the depicted variants. **a**) Initial situation: TS complex with the IGP lyase activity of zmTrpA being stimulated by zmTrpB1. **b**) The gene for zmTrpA is duplicated. **c**) The TS complex remains unaltered whereas **d**) zmTrpA* sequentially accumulates beneficial mutations that enhance its IGP lyase activity towards that of zmBX1. The bars indicate the experimental determined k_{cat}-values. Solid arrows describe experimentally confirmed steps, while dashed arrows indicate the steps necessary for the final evolution of extant zmBX1. **e**) Situation in extant *Z. mays*: The TS complex exists in parallel with zmBX1. For the TS complex, only zmTrpA:zmTrpB1 is shown instead of the entire zmTrpA:zmTrpB1:zmTrpB1:zmTrpA complex.

Acknowledgments

This work was supported by a grant of the Deutsche Forschungsgemeinschaft to R.S. (STE 891/ 11-1) and a grant by the National Institutes for Health to V.H.W. (NIH P41GM128577). The cDNA of *Zea mays* W22 was a generous donation of Dr. Kevin Begcy (University of Regensburg). We thank Sonja Fuchs, Christiane Endres, and Jeannette Ueckert for expert technical assistance, and Sandra Schlee, Patrick Babinger, and Rainer Merkl for critical reading of the manuscript.

2.3 Experimental Materials and Methods

2.3.1 Cloning of zmTrpA, zmBX1 and zmTrpB1

The genes coding for zmBX1 and zmTrpA were amplified from cDNA of *Zea mays* W22 leaves, using the primers 5'_zmBX1_*Bsa*I:
5'-AAAAAA**GGTCTC**A|GATGAGGAGCCGTCCGGTGTCGGACACCATGGCG-3'
and 3'_zmBX1_*Bsa*I:
5'-AAAAT**GGTCTC**T|GTCATGGCAGCGCGTTCTTCATGCCCCTGGC-3' for zmBX1, and the primers 5'_zmTrpA_*Nde*I:
5'-AAAAA**CATATG**GACAAGCGCAGCATCTCCGGCACCTTCGCC-3'
and 3'_zmTrpA_*Hind*III:
5'-AAAAA**AAGCTT**TCATGGCAATGCGGCCTTCAGGTTCTTGGC-3' for zmTrpA (bases in bold indicate the endonuclease recognition site). The "|" symbol indicates the cleavage site of the *Bsa*I endonuclease. The amplicon of zmBX1 was cloned into a pUR28a vector using the *Bsa*I cloning technique (Rohweder et al. 2018) for expression with an N-terminal His_6-tag. The amplicon of zmTrpA was cloned into pET28a via *Nde*I and *Hind*III cleavage sites for expression with an N-terminal His_6-tag. The gene coding for zmTrpB1 was synthesized by Thermo Fisher Scientific (GeneArt Strings DNA Fragments) and cloned into a pUR28a vector applying the *Bsa*I cloning technique for expression with an N-terminal His_6-tag. For all constructs Sanger sequencing was applied to confirm the desired sequence.

2.3.2 Mutagenesis of zmTrpA and zmBX1

The zmTrpA variants studied in this work were generated using QuikChange™ site-directed mutagenesis (QCM) (Wang and Malcolm 1999) or the method according to the NEB Q5® site-directed mutagenesis kit (PNK). The templates, primer pairs, and the methods for the generation of all variants are shown in **Table 2**. For all constructs Sanger sequencing was applied to confirm the desired sequence.

2.3.3 Expression and Purification of Recombinant Proteins

For expression of zmBX1, zmTrpA and its variants, and zmTrpB1 *E.coli* BL21 (DE3) Gold cells were transformed with the respective plasmids. The cells were grown at 37 °C in LB medium supplemented with 75 mg/ml kanamycin and 20 mM potassium phosphate (pH 7.5) for zmBX1, zmTrpA and its variants or 20 µM PLP for zmTrpB1. At a cell density

of OD_{600} = 0.6, protein expression was induced by the addition of 0.5 mM IPTG. The cells were grown overnight at 20 °C and then harvested by centrifugation (Beckman Coulter, Avanti J-26S XP, JLA-8.1, 20 min, 4000 rpm, 4 °C). Cell pellets were suspended in 100 mM potassium phosphate (pH 7.5), 300 mM KCl and 10 mM imidazole for zmBX1, zmTrpA and its variants or 50 mM potassium phosphate (pH 7.5), 300 mM KCl, 10 mM imidazole and 40 µM PLP for zmTrpB1. The cells were disrupted by sonication (Branson Sonifier W-250D; amplitude 50 %; 2 min, 30 s pulse/30 s pause). The insoluble fraction was removed by centrifugation (Beckman Coulter, Avanti J-26S XP, JA-25.50, 30 min, 14000 rpm, 4 °C) and soluble target proteins were purified by IMAC (immobilized metal ion chromatography) (GE Healthcare, HisTrap FF Crude). The proteins were eluted in 100 mM potassium phosphate (pH 7.5), 300 mM KCl using a linear gradient of 10 to 500 mM imidazole for zmBX1, zmTrpA and its variants or in 50 mM potassium phosphate (pH 7.5), 300 mM KCl using a linear gradient of 10 to 500 mM imidazole for zmTrpB1. Fractions with sufficiently pure protein were pooled and subjected to preparative size-exclusion chromatography (GE Healthcare, HiLoad 26/600 Superdex 75 PG for zmBX1, zmTrpA and its variants or GE Healthcare, HiLoad 26/600 Superdex 200 PG for zmTrpB1). The proteins were eluted in 100 mM potassium phosphate (pH 7.5), 300 mM KCl, 1 mM DTT for zmBX1, in 100 mM EPPS/KOH (pH 7.5), 1 mM DTT for zmTrpA and its variants, or in 100 mM EPPS/KOH (pH 7.5) for zmTrpB1. The main peak fractions were pooled, flash frozen in liquid nitrogen and stored at -80 °C. Protein concentrations were determined by absorbance spectroscopy at 280 nm (Thermo Scientific, NanoDrop One) for zmBX1, zmTrpA and its variants using molar extinction coefficients that were calculated from the amino acid sequence (Gasteiger E. 2005), and by the Bradford assay (Bio-Rad, Bradford protein assay) for zmTrpB1.

2.3.4 Analytical Size-Exclusion Chromatography

Analytical size-exclusion chromatography was performed at room temperature with a chromatographic device (GE Healthcare, ÄKTAmicro) coupled with an autosampler (Spark Holland, Alias GE Bio Cool) providing a cooled sample compartment at 8 °C. Proteins were eluted from an analytical size exclusion column (GE Healthcare, Superdex S75, 10/300 GL or GE Healthcare, Superdex S200, 10/300 GL) in 50 mM KP (pH 7.5), 300 mM KCl at a flow rate of 0.5 ml/min for the S75 column and 0.3 ml/min for the S200 column.

2.3.5 Analytical Steady-State Enzyme Kinetics, Activity Titrations, and Activity Screening of Indole-3-Glycerol-Phophate Lyases

The activity of zmBX1 and zmTrpA and its variants in absence and presence of zmTrpB1 was measured at 30 °C by monitoring the cleavage of IGP to GAP and indole (TrpA-reaction) with a coupled enzymatic assay in a V-650 spectrophotometer (JASCO) and analyzed using $\Delta\varepsilon_{340[NADH-NAD^+]}$= 6.22 $mM^{-1}cm^{-1}$ (Creighton 1970). The reactions were performed with varying concentrations of IGP for steady-state kinetics or a fixed concentration of IGP for activity titrations (0.5 mM) and activity screenings (1 mM). For all reactions 100 mM EPPS/KOH (pH 7.5), 180 mM KCl, 40 µM PLP, 6 mM NAD^+, 20 mM $NaAsO_4$, 100 mM L-serine (when zmTrpB1 was present), and 5.5 µM GAP dehydrogenase from *Thermotoga maritima* was added to the reaction mixture. For steady-state enzyme kinetics, zmBX1 was assayed at a final monomer concentration of 0.1 µM. zmTrpA was assayed at a final monomer concentration of 15 µM. zmTrpA_T186P was assayed at a final monomer concentration of 3 µM. zmTrpA_G192P was assayed at a final monomer concentration of 5 µM. zmTrpA_T186P_G192P and zmTrpA_L6zmBX1 were assayed at a final monomer concentration of 0.5 µM. zmTrpA and its variants in presence of 15 µM zmTrpB1 were assayed at a final monomer concentration of 0.5 µM. zmTrpA_T186M and zmTrpA_T186I were assayed at a final monomer concentration of 5 µM. For the activity screening each variant of zmTrpA was assayed at a final monomer concentration of 5 µM. For activity titrations the final monomer concentration for zmTrpA and its variants was 0.5 µM. The monomer concentration of zmTrpB1 was varied.

For steady-state kinetics, the initial velocities were plotted as a function of the IGP concentration. The resulting data points were fitted with a hyperbolic function for the determination of v_{max} and K_m^{IGP}. The k_{cat} value was determined by dividing v_{max} by the total enzyme concentration E_0. For activity titrations, the initial velocities were plotted as a function of the added zmTrpB1 concentration. The resulting data points were fitted with a hyperbolic function with an offset for the TrpA activity in absence of zmTrpB1. This allowed for the determination of K_D^{app} values, which correspond to the zmTrpB1 concentrations at which zmTrpA or its variants reach their half-maximum activity values. For activity screening, the initial velocities were taken.

2.3.6 Native Mass Spectrometry (MS)

Sample purity and integrity was analyzed by on-line buffer exchange MS using an UltiMate™ 3000 RSLC (Thermo Fisher Scientific) coupled to an Exactive Plus EMR Orbitrap instrument (Thermo Fisher Scientific) modified to incorporate a quadrupole mass filter and allow for surface-induced dissociation (VanAernum et al. 2019). 50 pmol protein were injected and on-line buffer exchanged to 200 mM ammonium acetate, pH 6.8 (AmAc) by a self-packed buffer exchange column (Waitt et al. 2008) (P6 polyacrylamide gel, BioRad) at a flow-rate of 0.1 ml/min. Mass spectra were recorded for 1000 – 8000 m/z at 17500 resolution as defined at 200 m/z. The injection time was set to 100 ms. Voltages applied to the transfer optics were optimized to allow ion transmission while minimizing unintentional ion activation. Mass spectra were deconvoluted with UniDec version 2.6.5 (Marty et al. 2015) using the following processing parameters: sample mass every 1 Da; peak FWHM 1 Thompson, Gaussian peak shape function.

For size exclusion MS experiments, proteins were first manually buffer-exchanged to 200 mM AmAc, pH 6.8. 150 pmole were injected onto a bioZen 1.8 µm SEC-2 column (150 x 4.6 mm) equipped with a security guard column, and eluted with 200 mM AmAc at a flow-rate of 0.2 ml/min. Total ion chromatograms were smoothed (Boxcar; 7 points) and overlaid.

2.4 Supplementary Figures

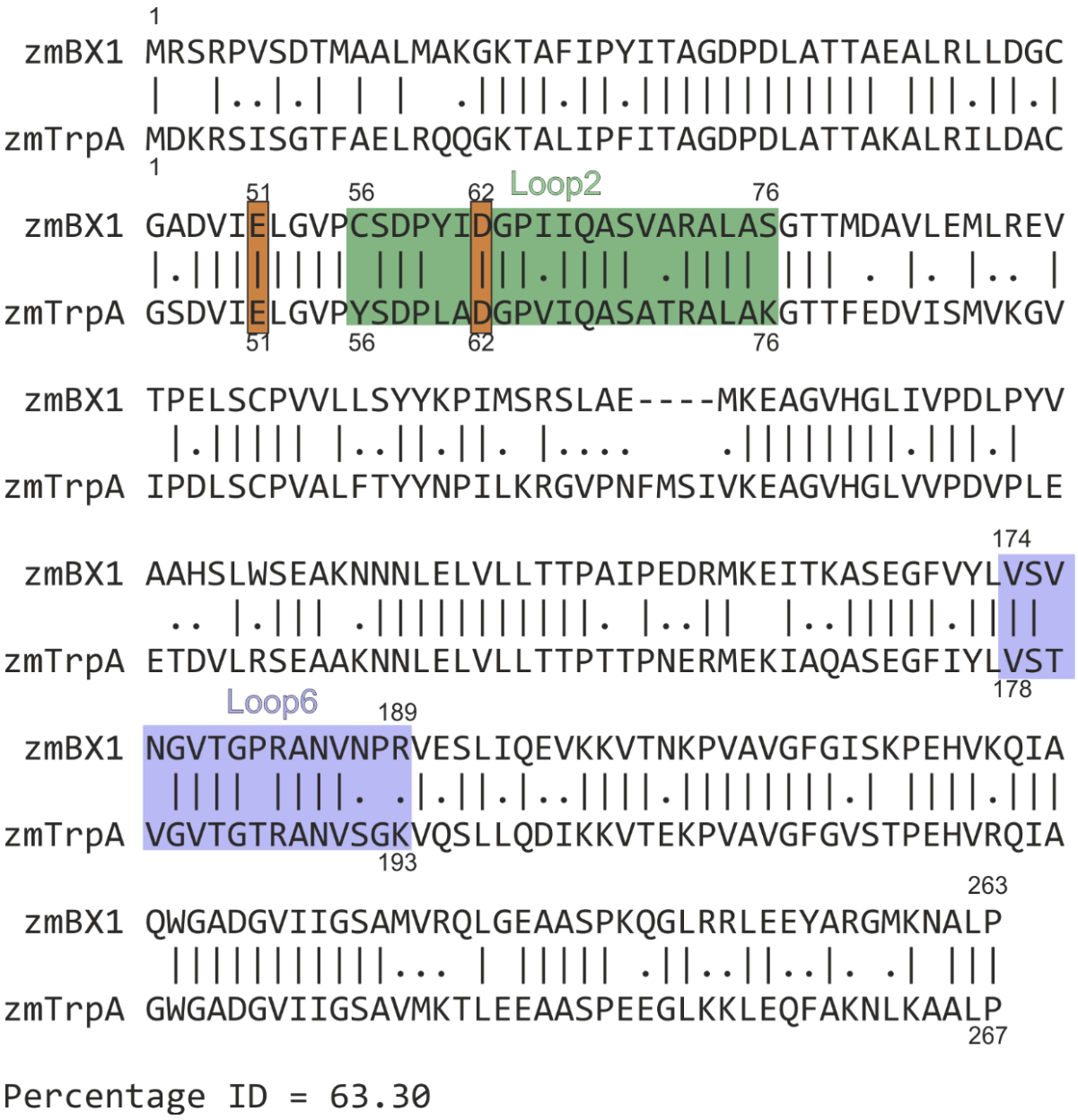


Figure 9: Pairwise sequence alignment of zmBX1 and zmTrpA

The alignment was calculated in Jalview (Waterhouse et al. 2009). The catalytic residues E51 and D62 are marked with orange boxes. Residues of Loop2 are highlighted in green, residues of Loop6 are highlighted in light blue.

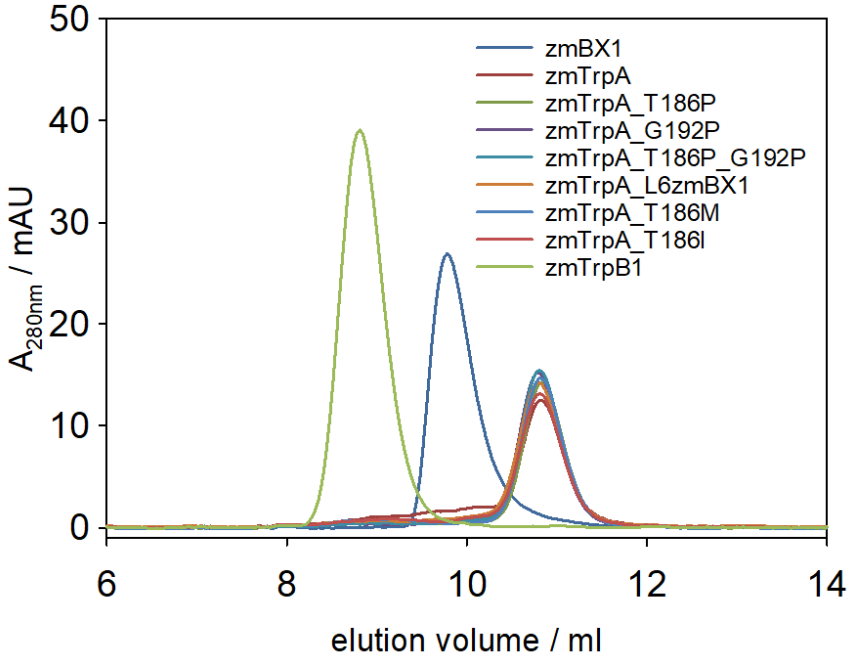
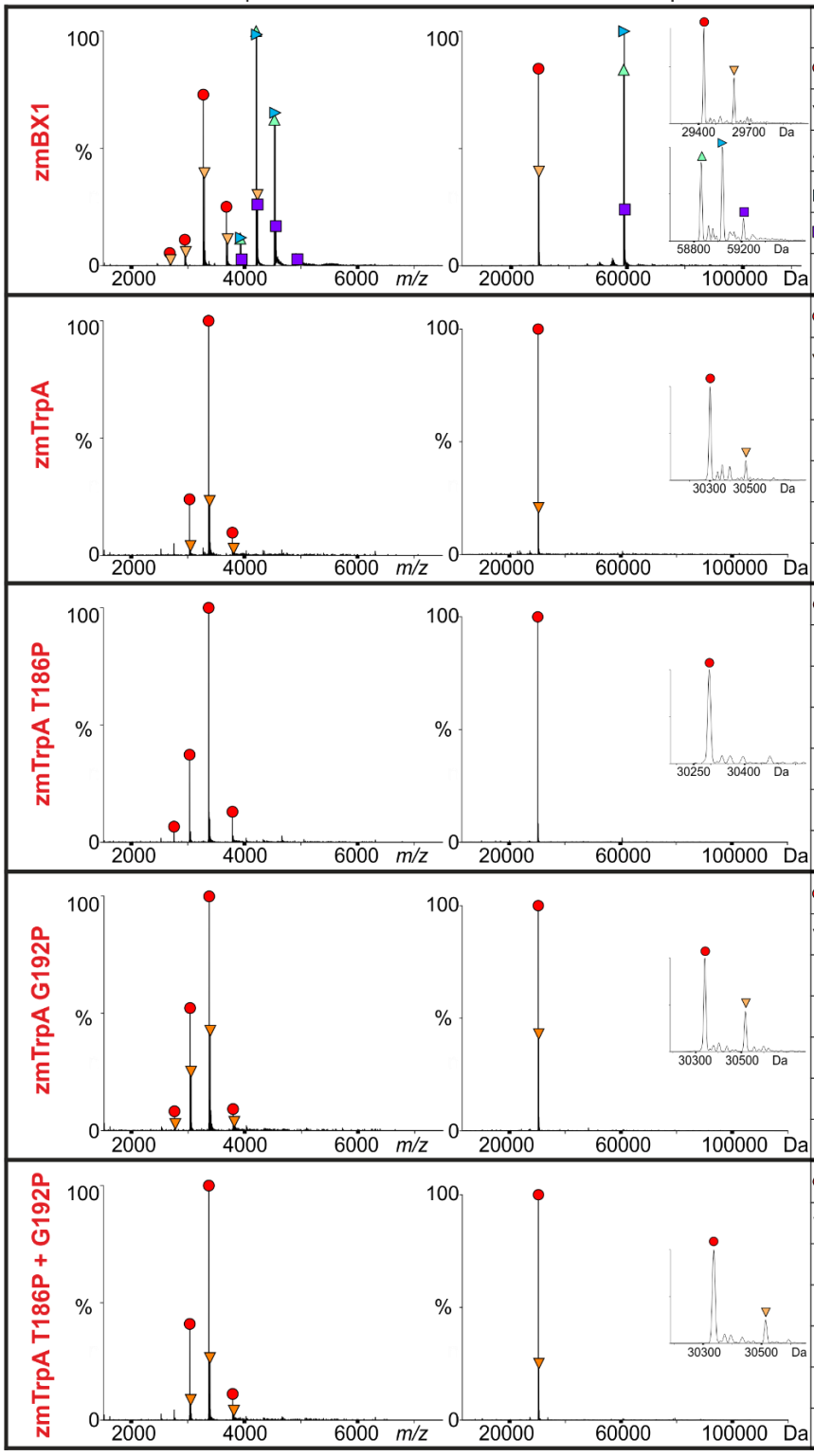


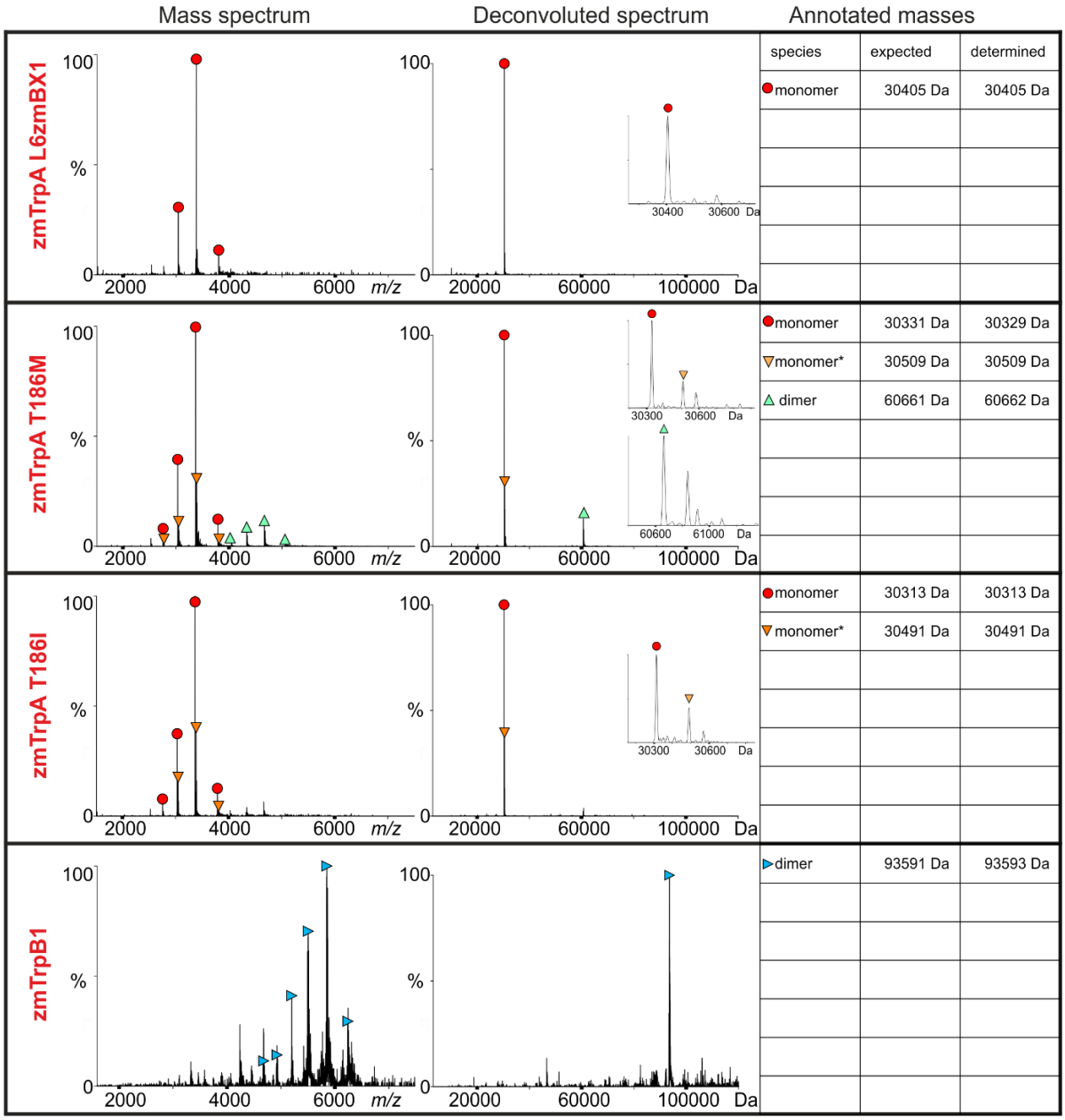
Figure 10: Analytical size-exclusion chromatograms

The experiments were performed on a calibrated Superdex S75 column, and the elution was followed at 280 nm by an absorbance detector. The applied protein (monomer) concentration was 30 µM for all shown samples.

Mass spectrum | Deconvoluted spectrum | Annotated masses



species	expected	determined
zmBX1		
monomer	29431 Da	29431 Da
monomer*	29609 Da	29609 Da
dimer	58862 Da	58860 Da
dimer*	59040 Da	59037 Da
dimer**	59218 Da	59215 Da
zmTrpA		
monomer	30301 Da	30301 Da
monomer*	30479 Da	30479 Da
zmTrpA T186P		
monomer	30297 Da	30296 Da
zmTrpA G192P		
monomer	30341 Da	30340 Da
monomer*	30519 Da	30518 Da
zmTrpA T186P + G192P		
monomer	30337 Da	30336 Da
monomer*	30515 Da	30514 Da



* α-6-gluconoylation (+178 Da)

for zmTrpB1: multiple salt-adducts

Figure 11: Purity and integrity of protein samples

Masses correspond to those expected based on protein sequence (see **Table 3**). Main reason for the heterogeneity of TrpA- and BX1-proteins is the α-6-gluconoylation due to Gly-Ser-Ser-$[His]_6$ at the N-terminus was 30 µM for all (Geoghegan et al. 1999). The difference between expected and determined mass of zmBX1 is likely due to (partial) cysteine oxidation.

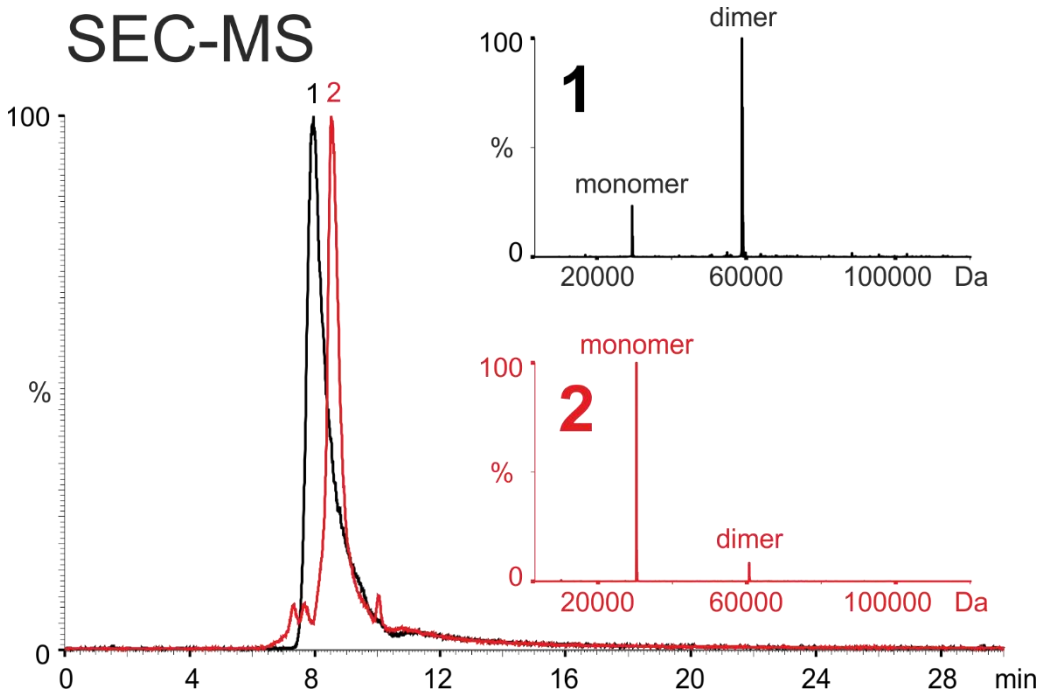
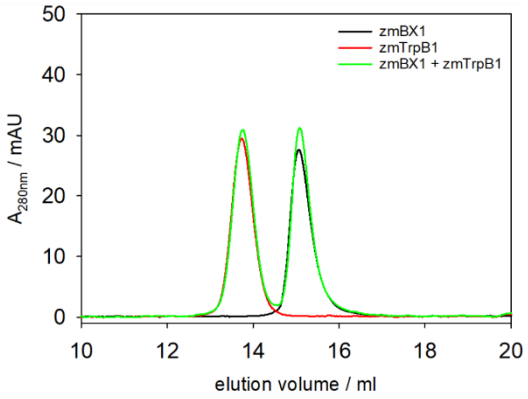
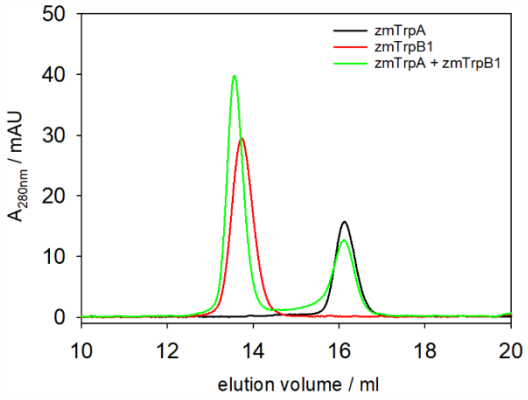
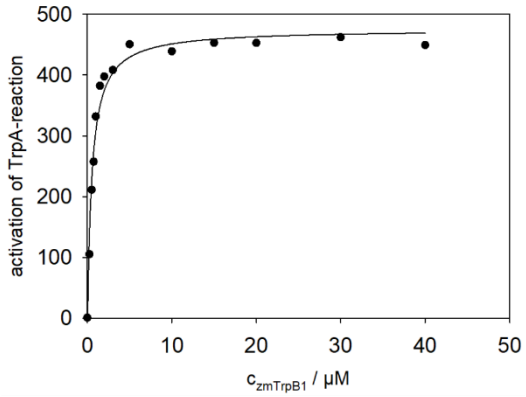
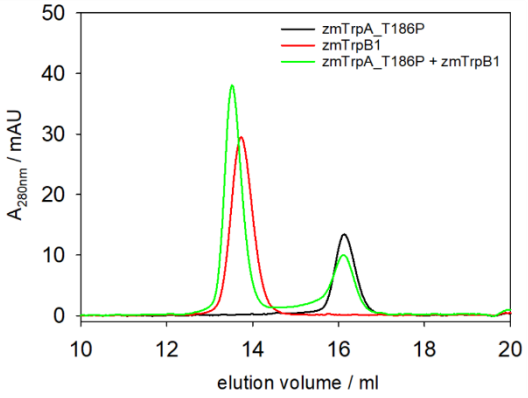
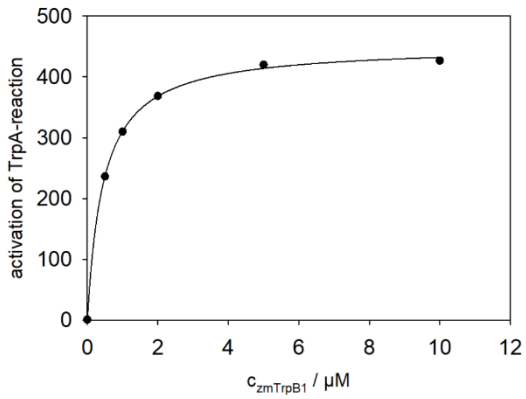


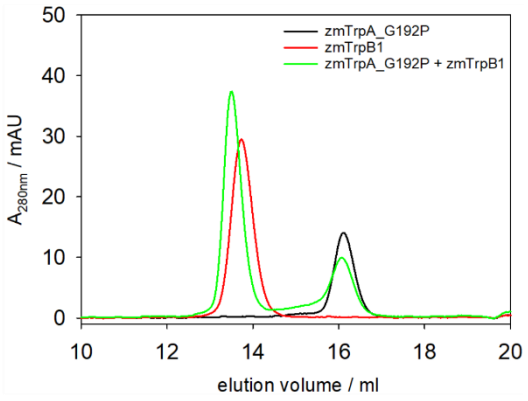
Figure 12: Oligomeric state determination of zmBX1 (1) and zmTrpA (2) by SEC-MS

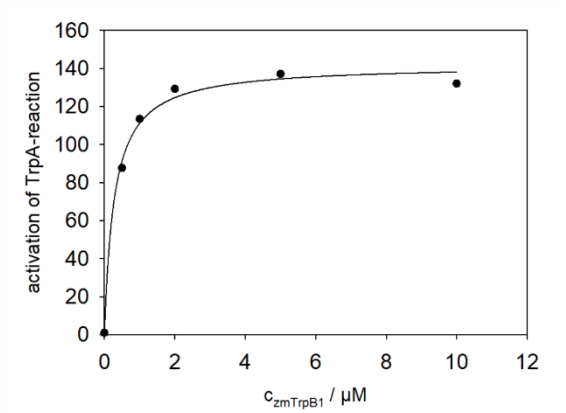
Deconvoluted spectra correspond to a single scan for the peak maximum of each eluting species. At the given protein concentration and conditions, zmBX1 appears to be primarily dimeric and zmTrpA appears to be primarily monomeric.

Complex formation with zmTrpB1

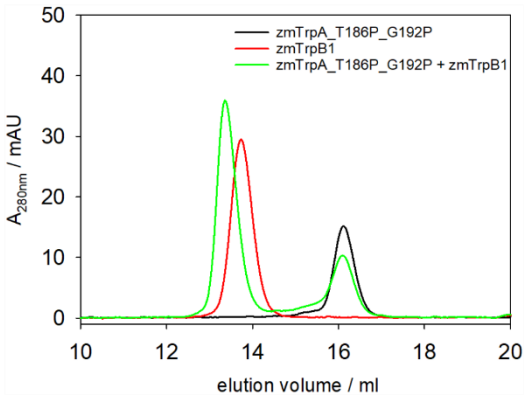
	structural / SEC	functional / activity titration
zmBX1		n.d.
zmTrpA		
zmTrpA_T186P		

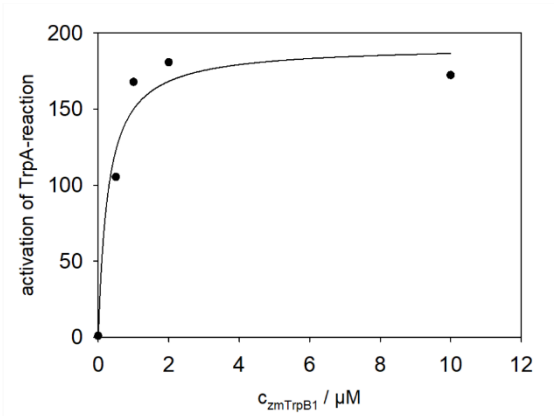
zmTrpA_G192P



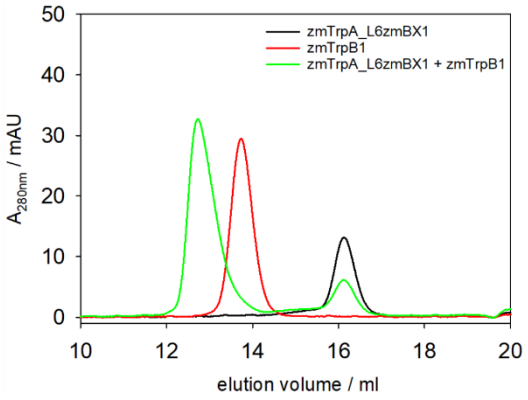


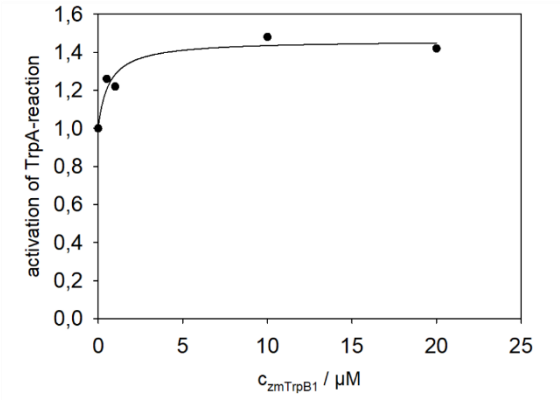
zmTrpA_T186P_G192P



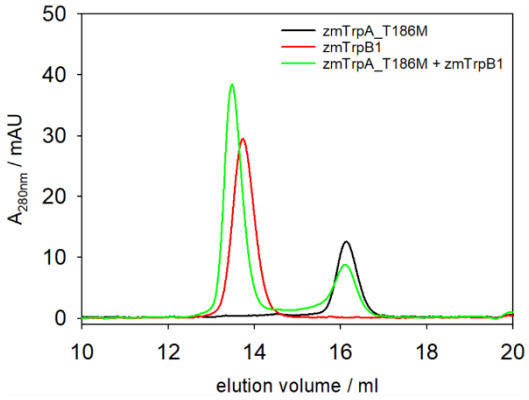


zmTrpA_L6zmBX1



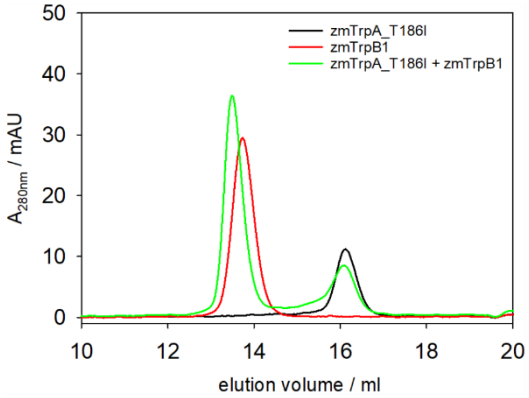


zmTrpA_T186M



n.d.

zmTrpA_T186I



n.d.

Figure 13: Structural (S200 SEC) and functional (activity titrations) interaction studies

Left columns: Analytical size-exclusion chromatograms of zmBX1, wild-type zmTrpA and its variants, in absence and presence of zmTrpB1. The experiments were performed on a calibrated Superdex S200 column. The elution was followed at 280 nm by an absorbance detector. The applied protein (monomer) concentration was 30 µM for all samples. **Right columns**: Activity titration experiments for determination of apparent K_D-values for the binding of zmTrpA and its variants to zmTrpB1. n.d.= not determined.

2.5 Supplementary Tables

Table 2: Primer used for site-directed mutagenesis

Target construct	Template	Primer name	Primer sequence 5'→3'	Method
pET28a_zmTrpA _L6zmBX1	pET28a_ zmTrpA	5'zmA_L6zmBX1 3'zmA_L6zmBX1	**CCTCGCGCAAACGTGAACCCACGA**GTGCAATCTCTTCTTCAGGA **ACCTGTCACTCCGTTCACGCTCAC**AAGATAAATAAATCCTTCTG	PNK
pET28a_zmTrpA _T186P	pET28a_ zmTrpA	5'zmA_T186P_PNK 3'zmA_T186P_PNK	**CCG**CGTGCAAATGTAAGC ACCAGTAACTCCAAC	PNK
pET28a_zmTrpA _G192P	pET28a_ zmTrpA	5'zmA_G192P_PNK 3'zmA_ G192P_PNK	**CCG**AAGGTGCAATCTCTT CGTGCAAATGTAAGC	PNK
pET28a_zmTrpA _T186P_G192P	pET28a_ zmTrpA	5'zmA_T186P_G192P 3'zmA_T186P_G192P	**CCG**CGTGCAAATGTAAGC**CCG**AAGGTGCAATCTCTT ACCAGTAACTCCAAC	PNK
pET28a_zmTrpA _T186A	pET28a_ zmTrpA	5'zmA_T186A_QCM 3'zmA_T186A_QCM	GAGTTACTGGT**GCG**CGTGCAAATG CATTTGCACG**CGC**ACCAGTAACTC	QCM
pET28a_zmTrpA _T186M	pET28a_ zmTrpA	5'zmA_T186M_QCM 3'zmA_T186M_QCM	GAGTTACTGGT**ATG**CGTGCAAATG CATTTGCACG**CAT**ACCAGTAACTC	QCM
pET28a_zmTrpA _T186E	pET28a_ zmTrpA	5'zmA_T186E_QCM 3'zmA_T186E_QCM	GAGTTACTGGT**GAA**CGTGCAAATG CATTTGCACG**TTC**ACCAGTAACTC	QCM
pET28a_zmTrpA _T186H	pET28a_ zmTrpA	5'zmA_T186H_QCM 3'zmA_T186H_QCM	GAGTTACTGGT**CAT**CGTGCAAATG CATTTGCACG**ATG**ACCAGTAACTC	QCM
pET28a_zmTrpA _T186Y	pET28a_ zmTrpA	5'zmA_T186Y_QCM 3'zmA_T186Y_QCM	GAGTTACTGGT**TAT**CGTGCAAATG CATTTGCACG**ATA**ACCAGTAACTC	QCM
pET28a_zmTrpA _T186I	pET28a_ zmTrpA	5'zmA_T186I_QCM 3'zmA_T186I_QCM	GAGTTACTGGT**ATT**CGTGCAAATG CATTTGCACG**AAT**ACCAGTAACTC	QCM
pET28a_zmTrpA _T186R	pET28a_ zmTrpA	5'zmA_T186R_QCM 3'zmA_T186R_QCM	GAGTTACTGGT**CGT**CGTGCAAATG CATTTGCACG**ACG**ACCAGTAACTC	QCM
pET28a_zmTrpA _T186G	pET28a_ zmTrpA	5'zmA_T186G_QCM 3'zmA_T186G_QCM	GAGTTACTGGT**GGC**CGTGCAAATG CATTTGCACG**GCC**ACCAGTAACTC	QCM
pET28a_zmTrpA _ G192A	pET28a_ zmTrpA	5'zmA_G192A_QCM 3'zmA_G192A_QCM	CAAATGTAAGC**GCC**AAGGTGCAAT ATTGCACCTT**GGC**GCTTACATTTG	QCM
pET28a_zmTrpA _ G192M	pET28a_ zmTrpA	5'zmA_ G192M_QCM 3'zmA_ G192M_QCM	CAAATGTAAGC**ATG**AAGGTGCAAT ATTGCACCTT**CAT**GCTTACATTTG	QCM
pET28a_zmTrpA _ G192E	pET28a_ zmTrpA	5'zmA_ G192E_QCM 3'zmA_ G192E_QCM	CAAATGTAAGC**GAA**AAGGTGCAAT ATTGCACCTT**TTC**GCTTACATTTG	QCM
pET28a_zmTrpA _ G192H	pET28a_ zmTrpA	5'zmA_ G192H_QCM 3'zmA_ G192H_QCM	CAAATGTAAGC**CAT**AAGGTGCAAT ATTGCACCTT**ATG**GCTTACATTTG	QCM
pET28a_zmTrpA _ G192Y	pET28a_ zmTrpA	5'zmA_ G192Y_QCM 3'zmA_ G192Y_QCM	CAAATGTAAGC**TAT**AAGGTGCAAT ATTGCACCTT**ATA**GCTTACATTTG	QCM
pET28a_zmTrpA _ G192I	pET28a_ zmTrpA	5'zmA_ G192I_QCM 3'zmA_ G192I_QCM	CAAATGTAAGC**ATT**AAGGTGCAAT ATTGCACCTT**AAT**GCTTACATTTG	QCM
pET28a_zmTrpA _ G192R	pET28a_ zmTrpA	5'zmA_ G192R_QCM 3'zmA_ G192R_QCM	CAAATGTAAGC**CGT**AAGGTGCAAT ATTGCACCTT**ACG**GCTTACATTTG	QCM
pET28a_zmTrpA _ G192T	pET28a_ zmTrpA	5'zmA_ G192T _QCM 3'zmA_ G192T _QCM	CAAATGTAAGC**ACC**AAGGTGCAAT ATTGCACCTT**GGT**GCTTACATTTG	QCM

Bases in **bold** indicate mutated positions. The column "Method" indicates the method that was used for site-directed mutagenesis.

Table 3: Amino acid sequences of experimentally characterized proteins

Construct	Protein	Amino acid sequence	*average $mass_{calc}$/ Da	**average $mass_{calc}$/ Da
pUR28a_ zmBX1	zmBX1	MGSSHHHHHHLEMRSRPVSDTMAALMAKGKTAFIPYITAGDPDLATTAEALRLLD GCGADVIELGVPCSDPYIDGPIIQASVARALASGTTMDAVLEMLREVTPELSCPVV LLSYYKPIMSRSLAEMKEAGVHGLIVPDLPYVAAHSLWSEAKNNNLELVLLTTPAIP EDRMKEITKASEGFVYLVSVNGVTGPRANVNPRVESLIQEVKKVTNKPVAVGFGI SKPEHVKQIAQWGADGVIIGSAMVRQLGEAASPKQGLRRLEEYARGMKNALP	29431	29609
PET28a_ zmTrpA	zmTrpA	MGSSHHHHHHSSGLVPRGSHMDKRSISGTFAELRQQGKTALIPFITAGDPDLATT AKALRILDACGSDVIELGVPYSDPLADGPVIQASATRALAKGTTFEDVISMVKGVIP DLSCPVALFTYYNPILKRGVPNFMSIVKEAGVHGLVVPDVPLEETDVLRSEAAKNN LELVLLTTPTTPNERMEKIAQASEGFIYLVSTVGVTGTRANVSGKVQSLLQDIKKVT EKPVAVGFGVSTPEHVRQIAGWGADGVIIGSAVMKTLEEAASPEEGLKKLEQFAK NLKAALP	30301	30479
pET28a_ zmTrpA _T186P	zmTrpA_ T186P	MGSSHHHHHHSSGLVPRGSHMDKRSISGTFAELRQQGKTALIPFITAGDPDLATT AKALRILDACGSDVIELGVPYSDPLADGPVIQASATRALAKGTTFEDVISMVKGVIP DLSCPVALFTYYNPILKRGVPNFMSIVKEAGVHGLVVPDVPLEETDVLRSEAAKNN LELVLLTTPTTPNERMEKIAQASEGFIYLVSTVGVTG**P**RANVSGKVQSLLQDIKKVT EKPVAVGFGVSTPEHVRQIAGWGADGVIIGSAVMKTLEEAASPEEGLKKLEQFAK NLKAALP	30297	30475
pET28a_ zmTrpA _G192P	zmTrpA_ G192P	MGSSHHHHHHSSGLVPRGSHMDKRSISGTFAELRQQGKTALIPFITAGDPDLATT AKALRILDACGSDVIELGVPYSDPLADGPVIQASATRALAKGTTFEDVISMVKGVIP DLSCPVALFTYYNPILKRGVPNFMSIVKEAGVHGLVVPDVPLEETDVLRSEAAKNN LELVLLTTPTTPNERMEKIAQASEGFIYLVSTVGVTGTRANVS**P**KVQSLLQDIKKVT EKPVAVGFGVSTPEHVRQIAGWGADGVIIGSAVMKTLEEAASPEEGLKKLEQFAK NLKAALP	30341	30519
pET28a_ zmTrpA _T186P_G 192P	zmTrpA_ T186P_ G192P	MGSSHHHHHHSSGLVPRGSHMDKRSISGTFAELRQQGKTALIPFITAGDPDLATT AKALRILDACGSDVIELGVPYSDPLADGPVIQASATRALAKGTTFEDVISMVKGVIP DLSCPVALFTYYNPILKRGVPNFMSIVKEAGVHGLVVPDVPLEETDVLRSEAAKNN LELVLLTTPTTPNERMEKIAQASEGFIYLVSTVGVTG**P**RANVS**P**KVQSLLQDIKKVT EKPVAVGFGVSTPEHVRQIAGWGADGVIIGSAVMKTLEEAASPEEGLKKLEQFAK NLKAALP	30337	30515
pET28a_ zmTrpA _L6zmBX1	zmTrpA_ L6zmBX 1	MGSSHHHHHHSSGLVPRGSHMDKRSISGTFAELRQQGKTALIPFITAGDPDLATT AKALRILDACGSDVIELGVPYSDPLADGPVIQASATRALAKGTTFEDVISMVKGVIP DLSCPVALFTYYNPILKRGVPNFMSIVKEAGVHGLVVPDVPLEETDVLRSEAAKNN LELVLLTTPTTPNERMEKIAQASEGFIYLVS**VN**GVTG**P**RANV**NPR**VQSLLQDIKKV TEKPVAVGFGVSTPEHVRQIAGWGADGVIIGSAVMKTLEEAASPEEGLKKLEQFA KNLKAALP	30405	30583
pET28a_ zmTrpA _T186M	zmTrpA_ T186M	MGSSHHHHHHSSGLVPRGSHMDKRSISGTFAELRQQGKTALIPFITAGDPDLATT AKALRILDACGSDVIELGVPYSDPLADGPVIQASATRALAKGTTFEDVISMVKGVIP DLSCPVALFTYYNPILKRGVPNFMSIVKEAGVHGLVVPDVPLEETDVLRSEAAKNN LELVLLTTPTTPNERMEKIAQASEGFIYLVSTVGVTG**M**RANVSGKVQSLLQDIKKV TEKPVAVGFGVSTPEHVRQIAGWGADGVIIGSAVMKTLEEAASPEEGLKKLEQFA KNLKAALP	30331	30509
pET28a_ zmTrpA _T186I	zmTrpA_ T186I	MGSSHHHHHHSSGLVPRGSHMDKRSISGTFAELRQQGKTALIPFITAGDPDLATT AKALRILDACGSDVIELGVPYSDPLADGPVIQASATRALAKGTTFEDVISMVKGVIP DLSCPVALFTYYNPILKRGVPNFMSIVKEAGVHGLVVPDVPLEETDVLRSEAAKNN LELVLLTTPTTPNERMEKIAQASEGFIYLVSTVGVTG**I**RANVSGKVQSLLQDIKKVT EKPVAVGFGVSTPEHVRQIAGWGADGVIIGSAVMKTLEEAASPEEGLKKLEQFAK NLKAALP	30313	30491
Construct	Protein	Amino acid sequence	*average $mass_{calc}$/ Da	***average $mass_{calc}$/ Da
pUR28a_ zmTrpB1	zmTrpB1	MHHHHHHLEMAASPAAALEMGNGAAVAGLQRPDAMGRFGRFGGKYVPETLMH ALTELENAFHALATDDEFQKELDGILKDYVGRESPLYFAERLTEHYKRADGTGPLI YLKREDLNHTGAHKINNAVAQALLAKRLGKQRIIAETGAGQHGVATATVCARFGL QCIIYMGAQDMERQALNVFRMKLLGAEVRAVHSGTATLKDATSEAIRDWVTNVET THYILGSVAGPHPYPMMVREFHKVIGKETRRQAMHKWGGKPDVLVACVGGGSN AMGLFHEFVEDQDVRLIGVEAAGHGVDTDKHAATLTKGQVGVLHGSMSYLLQDD DGQVIEPHSISAGLDYPGVGPEHSFLKDIGRAEYDSVTDQEALDAFKRVSRLEGII PALETSHALAYLEKLCPTLPDGVRVVLNCSGRGDKDVHTASKYLDV	46474	46796

Sequence colored in red: His_6-tag for IMAC purification. In case of pET28a constructs the His_6-tag sequence also contains a thrombin cleavage site (LVPR|GS). Sequence colored in blue: Initial methionine (M1). Sequence in **bold**: Residues that were mutated compared to the respective wild-type sequence. *Organic source corrected average masses for monomers were calculated with NIST Mass and Fragment Calculator v1.32 (Kilpatrick et al. 2012). **Organic source corrected average masses for monomers including PTM= alpha-N-6-phosphogluconoylation (Geoghegan et al. 1999), *** Organic source corrected average masses for monomer including internal aldimine with cofactor PLP+4Na+.

Table 4: Overview of structural and functional complex formation

Proteins	Structural complex formation as detected by SEC	Functional complex formation as detected by activity titration
		K_D^{app} / µM
zmBX1 + zmTrpB1	no	n.d.
zmTrpA + zmTrpB1	yes	0.5
zmTrpA_T186P + zmTrpB1	yes	0.5
zmTrpA_G192P + zmTrpB1	yes	0.3
zmTrpA_T186P_G192P + zmTrpB1	yes	0.3
zmTrpA_L6zmBX1 + zmTrpB1	yes	0.7
zmTrpA_T186M + zmTrpB1	yes	n.d.
zmTrpA_T186I + zmTrpB1	yes	n.d.

n.d.= not determined.

3 Analysis of Allosteric Communication in a Multi Enzyme Complex by Ancestral Sequence Reconstruction

Michael Schupfner, Kristina Straub, Rainer Merkl, and Reinhard Sterner

This chapter contains unpublished data.

Key words: Allostery, ancestral sequence reconstruction, vertical approach, tryptophan synthase

3.1 Abstract

Tryptophan synthase (TS) is a hetero-tetrameric $\alpha\beta\beta\alpha$ complex. It is characterized by the channeling of the reaction intermediate indole and the mutual activation of the α-subunit TrpA and the β-subunit TrpB via a complex allosteric network. We have analyzed this allosteric network by means of ancestral sequence reconstruction (ASR), which is an *in silico* method to resurrect extinct ancestors of modern proteins. Previously, the sequences of TrpA and TrpB from the last bacterial common ancestor (LBCA) have been computed by means of ASR and characterized. LBCA-TS is similar to modern TS by forming a $\alpha\beta\beta\alpha$ complex with indole channeling taking place. However, LBCA-TrpA allosterically decreases the activity of LBCA-TrpB by a factor of 5, whereas, for example, the modern ncTrpA from *Neptuniibacter caesariensis* allosterically increases the activity of ncTrpB by a factor of 30. To identify amino acid residues that are responsible for this inversion of the allosteric effect, all six evolutionary TrpA and TrpB intermediates that stepwise link LBCA-TS with ncTS were characterized. Remarkably, the switching from TrpB-inhibition to TrpB-activation by TrpA occurred between two successive TS intermediates. Sequence comparison of these two intermediates and iterative rounds of site-directed mutagenesis allowed us to identify four out of 413 residues from TrpB that are necessary and sufficient for its allosteric activation by TrpA. The effect of our mutational studies were rationalized by a community analysis based on molecular dynamics simulations. Our findings demonstrate that ancestral sequence reconstruction can efficiently identify residues essential for allosteric signal propagation in multi-enzyme complexes.

3.2 Introduction

Allostery describes the phenomenon that protein function can be altered by the binding of an effector ligand at a site that is topographically distinct from the active site (Lu et al. 2019). Allosteric communication is central for the regulation of the activity of many enzyme complexes and has been intensely studied with various approaches over the past decades. Nevertheless, due to the complexity of the studied systems, there are still plenty of questions open with respect to the structural basis of allostery (Wodak et al. 2019). Previous studies have indicated that allostery is based on dynamic effects of the protein backbone and of side chains that can be understood by a combination of computational and experimental approaches (Lisi and Loria 2017, Liang et al. 2019). However, allosteric communication between distant sites is likely mediated via multiple paths, which makes it difficult to identify the role of individual structural elements for signal propagation (Wodak et al. 2019). One possibility to identify such elements is the analysis of specific mutations that disturb communication between the allosteric site to the active site (Lu et al. 2019). It has been shown, for instance, that allosteric signals in protein kinase A (PKA) (Lisi and Loria 2017), in imidazole glycerol phosphate synthase (ImGPS) (List et al. 2012), and in aminodeoxychorismate synthase (ADCS) (Semmelmann et al. 2019) can be disrupted with single point mutations.

Tryptophan synthase (TS) is a hetero-tetrameric $\alpha\beta\beta\alpha$ complex where a central dimer of two β-subunits (TrpB) is surrounded by two α-subunits (TrpA) (Hyde et al. 1988). These subunits catalyze the last two steps in the biosynthesis of the essential amino acid tryptophan: TrpA cleaves indole-3-glycerol phosphate (IGP) into glyceraldehyde-3-phosphate and the reaction intermediate indole, which is subsequently channeled to the active site of TrpB where it reacts in a PLP-mediated reaction with L-serine to L-tryptophan (**Figure 14**). In order to coordinate these reactions, the TrpA and TrpB subunits mutually activate each other in an allosteric manner (Miles 2001). These properties and a long history of research make the TS an excellent model system for the analysis of allosteric communication within multi-enzyme complexes (Dunn 2012, Buller et al. 2015, Buller et al. 2018). However, many aspects of allostery in TS are still not understood and only few residues involved in signal propagation from the active site of TrpA to the active site of TrpB (and vice versa) have been identified.

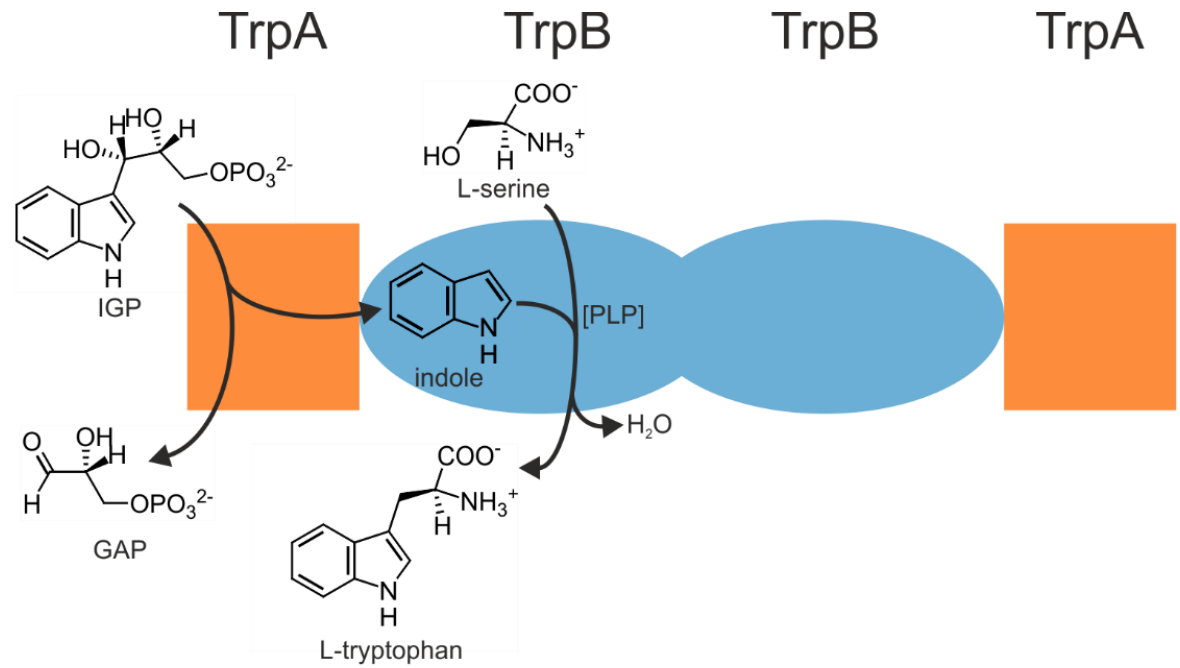


Figure 14: Structure and function of tryptophan synthase (TS)

The αββα TS complex is constituted of a central dimer of β-subunits (TrpB; colored light blue) with two adjacent α-subunits (TrpA, colored orange). TrpA catalyzes the cleavage of indole-3-glycerol-phosphate (IGP) into glyceraldehyde-3-phosphate (GAP) and indole. The latter is channeled to the active site of TrpB where it condenses in a PLP-dependent manner with L-serine, yielding L-tryptophan.

Ancestral sequence reconstruction (ASR) is an *in silico* method that allows one to calculate the most probable sequences of proteins from extinct species with the help of a multiple sequence alignment of extant species and a phylogenetic model (Merkl and Sterner 2016, Hochberg and Thornton 2017). We have previously used ASR to determine the amino acid sequences of TrpA and TrpB from the Last Bacterial Common Ancestor (LBCA-TrpA and LBCA-TrpB) (Busch et al. 2016). Both LBCA-TrpA and LBCA-TrpB were catalytically active and assembled to a αββα complex within which indole was channeled with similar efficiency as in modern TS complexes. However, although IGP cleavage of LBCA-TrpA was elevated by two orders of magnitude in the presence of LBCA-TrpB similar as in modern TS, the condensation of L-serine and indole catalyzed by LBCA-TrpB was slowed down fivefold by LBCA-TrpA. This unexpected allosteric inhibition might be an ancient feature of primordial TS complexes or an artifact of the reconstruction process. In any case it offers the opportunity to use a “vertical” approach for the identification of structural elements and specific amino acids that are crucial for the communication between TrpA and TrpB. In contrast to the conventional “horizontal” approach in which extant homologous proteins are compared, in a vertical approach reconstructed proteins representing extinct species of the phylogenetic tree are compared with modern ones homologues. Vertical approaches are much more

straightforward for the analysis of sequence-function relationship of proteins than horizontal approaches, especially because functional differences can be traced back to much fewer differences in sequence (Hochberg and Thornton 2017).

Specifically in the case of TS, starting from the root of the tree (identical to LBCA-TS), all internal TS nodes leading to a specific leaf (a modern TS) are available and sequence differences can be correlated with functional differences. Hence, we experimentally characterized six phylogenetically intermediate TS sequences (Anc1nc-TS - Anc6nc-TS) that link LBCA-TS with the modern TS from the γ-proteobacterium *Neptuniibacter caesariensis* (ncTS), in which ncTrpB is stimulated 30-fold by ncTrpA. We found that on the way from LBCA-TS to ncTS the transition from allosteric inhibition to allosteric activation occurs between the adjacent Anc2nc-TS and Anc3nc-TS and is entirely based on sequences differences in TrpB. Subsequent site-directed mutagenesis proved that only four (out of 413) TrpB residues are crucial for the inversion of allostery from inhibitory to activating, and the effect of these residues could be rationalized by a computational community analysis (Sethi et al. 2009) that is based on MD simulations.

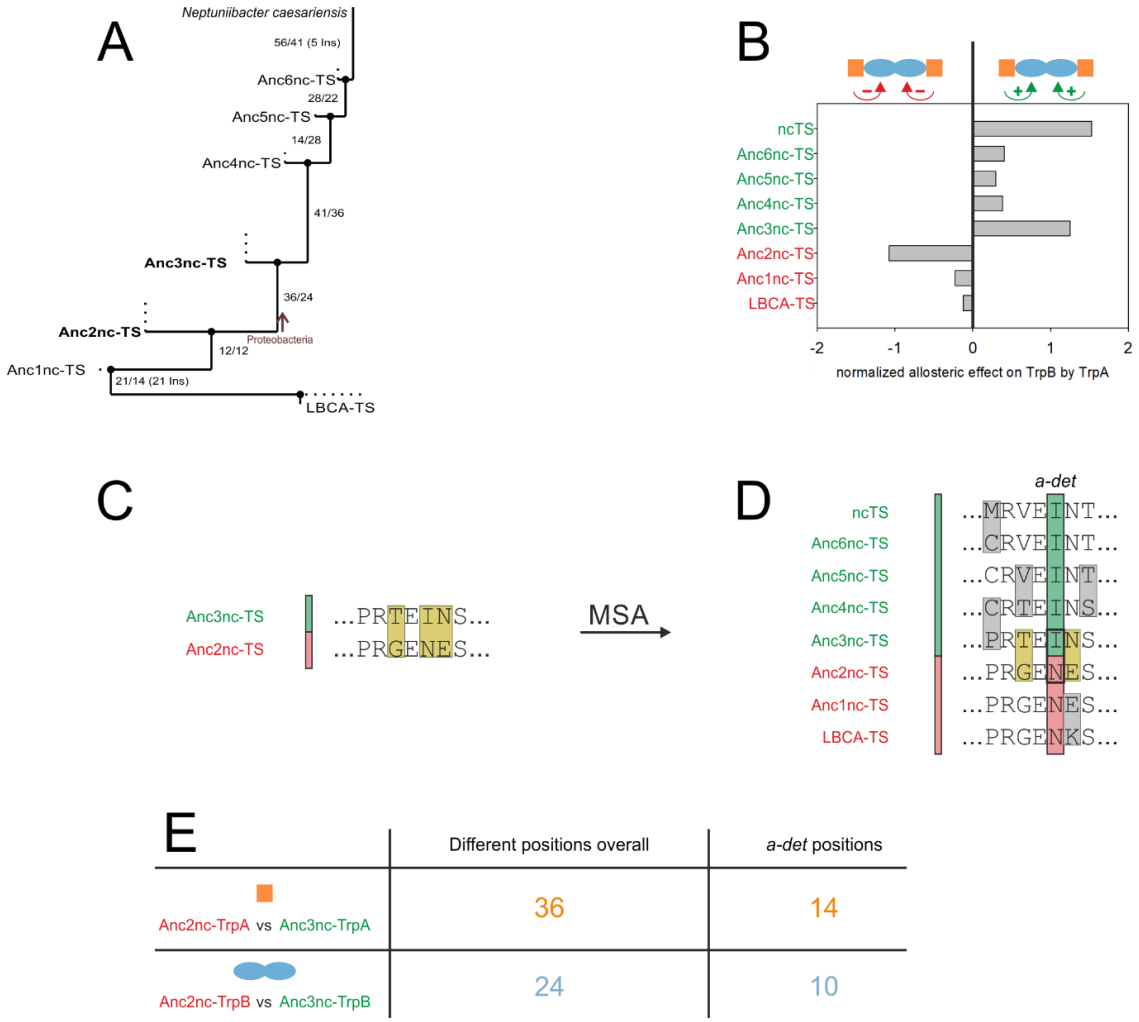
3.3 Results

3.3.1 A Vertical Approach Identifies Ten TrpB Positions That Are Essential for the Allosteric Effect

The reconstructed LBCA-TS displays similar features as the well-studied TS complexes from *Salmonella typhimurium* or *Escherichia coli* (stTS/ecTS) except for the “inversed” allosteric effect of TrpA on TrpB. While in all characterized extant TS the TrpB subunit is activated by the respective TrpA subunit (Dunn 2012, Miles 2013), in LBCA-TS the LBCA-TrpB subunit is deactivated five-fold by LBCA-TrpA (Busch et al. 2016). In order to identify residues that are responsible for these different allosteric effects of LCBA-TS and extant TS complexes, a vertical approach was applied. In the first step, a path within the phylogenetic tree connecting LBCA-TS (root of the tree) with the extant ncTS (one leaf of the tree), which shows an activating allosteric effect (see below), was selected. LBCA-TrpA and ncTrpA share a sequence identity of 48 %, while LBCA-TrpB and ncTrpB share a sequence identity of 72 %. This means that LBCA-TrpA and ncTrpA differ in 137 residues, while LBCA-TrpB and ncTrpB differ in 111 residues; moreover, compared to LBCA-TrpB, the ncTrpB sequence contains 13 additional residues. These drastic sequence variations make it impossible to test experimentally which residues are responsible for the different allosteric effects. Therefore, in a second step, all internal

nodes of the selected path linking LBCA-TS and ncTS (Anc1nc-TS - Anc6nc-TS) were analyzed (**Figure 15A**). For this purpose, the genes coding for the studied TrpA and TrpB proteins were cloned (see **Table 6** for primers), and the recombinant proteins (see **Table 7** for sequences) were generated in *E. coli* and purified. An initial characterization revealed that not only LBCA-TS but also Anc1nc-TS and Anc2nc-TS display an inhibitory allosteric effect, whereas Anc3nc-TS - Anc6-ncTS and ncTS show an activating allosteric effect (**Figure 15B**). These findings show that a discrete transition from the deactivating to the activating allosteric effect is occurring between Anc2nc-TS and Anc3nc-TS.

Pairwise sequence comparisons identified 36 residue differences between Anc2nc-TrpA and Anc3nc-TrpA and 24 residue differences between Anc2nc-TrpB and Anc3nc-TrpB (**Figure 15C** and **Figure 15E**). The detailed experimental assessment of these 60 residue differences would still not be possible with reasonable effort. To further narrow down the number of essential residues, a multiple sequence alignment (MSA) for the concatenated TrpA and TrpB sequences was created, which contained all the sequences of the selected path from LBCA-TS up to ncTS. Within the MSA, the sequences were grouped according to the allosteric effect of TrpA on TrpB, whether activating or deactivating. Positions occupied with residues that differ between the two groups but are conserved within each group were designated as allostery-determining (*a-det*) positions (**Figure 15D**). Within the TrpA sequences 14 *a-det* positions and within the TrpB sequences 10 *a-det* positions were identified (**Figure 15E**).



E

	Different positions overall	*a-det* positions
Anc2nc-TrpA vs Anc3nc-TrpA	36	14
Anc2nc-TrpB vs Anc3nc-TrpB	24	10

Figure 15: Identification of *a-det* positions in TrpA and TrpB by a vertical approach

(**A**) Path within the phylogenetic tree of TS connecting LBCA-TS with TS from *Neptuniibacter caesariensis* (ncTS) via the intermediate nodes Anc1nc-TS - Anc6nc-TS. The numbers of differing amino acids between the connected sequences are indicated. (Number of differences in TrpA sequence/number of differences in TrpB sequence.) Ins.: insertions. (**B**) Allosteric effect of TrpA subunits on their respective TrpB subunits. The TrpB reactions were performed at 30 °C in presence of 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP, 45 mM L-serine, 300 µM indole, 0.5 µM TrpB (monomer) and 1 µM TrpA (monomer). The resulting reaction mixture was analyzed by reversed phase HPLC. The normalized allosteric effect is defined as the decadic logarithm of the TrpB activity in presence of TrpA divided by the TrpB activity in absence of TrpA. Raw data are given in **Table 9**. TS complexes with allosteric activation of TrpB by TrpA are colored green, TS complexes with allosteric inhibition are colored red. (**C**) Scheme of the pairwise sequence comparison between Anc2nc-TS and Anc3nc-TS. The shown sequences are randomly selected examples with differing amino acids framed. (**D**) Scheme of the multiple sequence alignment (MSA) of the concatenated TS amino acid sequences. The sequences were grouped according to the allosteric effect on the TrpB subunit by the TrpA subunit. Sequences, where the TrpB subunit is activated are colored green. Sequences, where the TrpB subunit is deactivated are colored red. Positions occupied with residues that differ between the two groups, but are conserved within each group were designated as “allostery-determining” (*a-det*) positions. The shown sequences are random examples. (**E**) Table listing the number of “different overall positions” identified in (**C**) and of *a-det* positions identified in (**D**).

We next asked whether the identity of the TrpA subunit is at all significant for the activation or inactivation of the TrpB subunit. To this end, cross-wise activity titrations between Anc2nc-TrpB and Anc2nc-TrpA (**Figure 16A**) or Anc3nc-TrpA (**Figure 16B**) were performed, as well as between Anc3nc-TrpB and Anc2nc-TrpA (**Figure 16C**) or Anc3nc-TrpA (**Figure 16D**). The results revealed that Anc2nc-TrpB is deactivated by both Anc2nc-TrpA and Anc3nc-TrpA, while Anc3nc-TrpB is activated by both Anc2nc-TrpA and Anc3nc-TrpA. This led to the conclusion that the nature of the TrpA subunit is unimportant for the allosteric effect on the TrpB subunit. We concluded that the 10 *a-det* positions identified for the TrpB subunit are sufficient for the inversion of the allosteric effect of TrpA on TrpB.

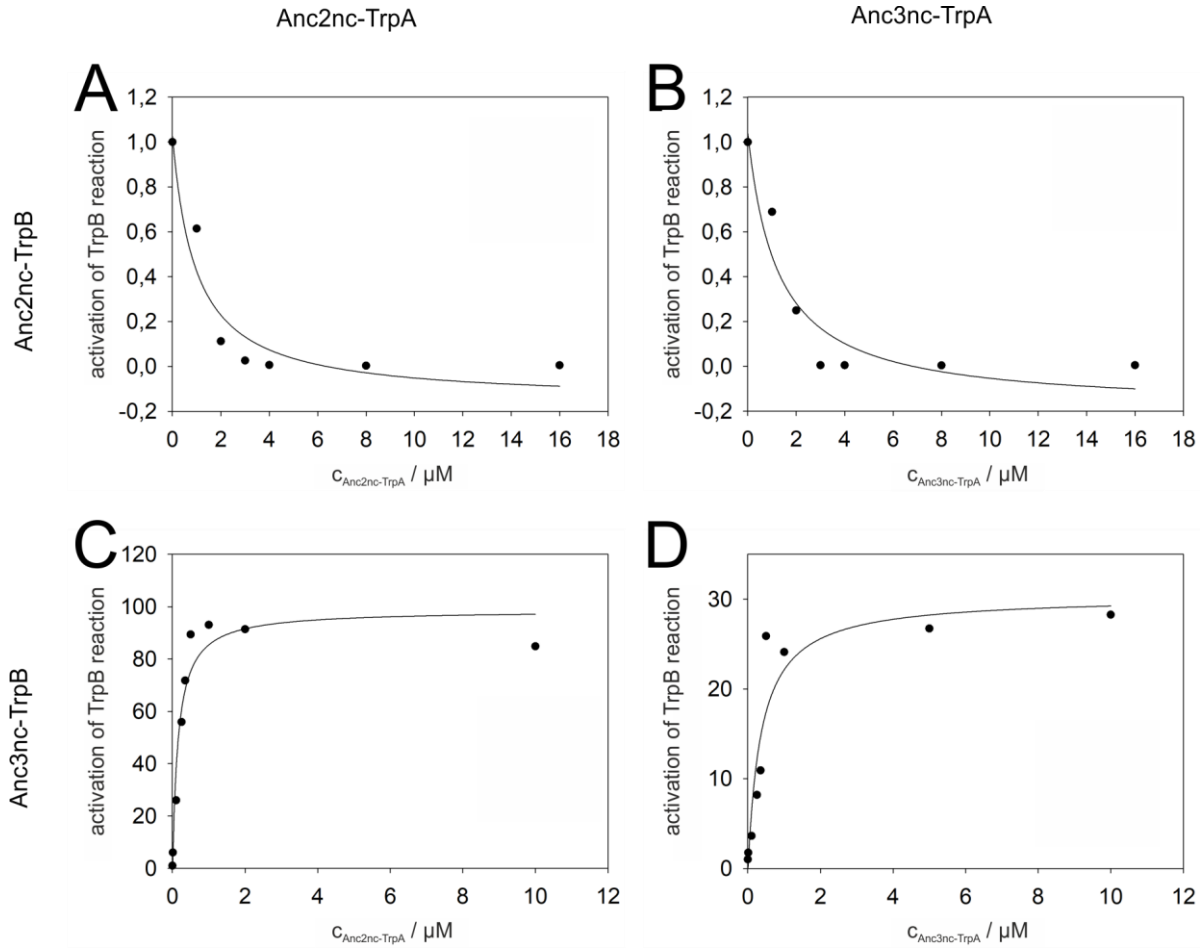


Figure 16: Activity titrations of TrpB subunits with TrpA subunits of Anc2nc-TS and Anc3nc-TS

Anc2nc-TrpB (2 µM, monomer concentration) was titrated with varying concentrations of Anc2nc-TrpA (**A**) and Anc3nc-TrpA (**B**). Anc3nc-TrpB (0.5 µM, monomer concentration) was titrated with varying concentrations of Anc2nc-TrpA (**C**) and Anc3nc-TrpA (**D**). The titrations were performed at 30 °C in presence of 100 mM KP pH 7.5, 180 mM KCl, 40 µM PLP, 200 µM indole, and 50 mM L-serine for Anc2nc-TrpB or 250 mM L-serine for Anc3nc-TrpB. TrpB activities were followed by monitoring the condensation of indole and L-serine to L-tryptophan at a wavelength of 290 nm.

To confirm this conclusion, two variants were created. In variant Anc2nc-TrpB_AM10, the 10 *a-det* residues as present in Anc2nc-TrpB were mutated towards the Anc3nc-TrpB residues. Inversely, in the variant Anc3nc-TrpB_AM10, the 10 *a-det* residues as present in Anc3nc-TrpB were mutated towards the Anc2nc-TrpB residues. These variants were assayed for the allosteric effect of Anc2nc-TrpA (for Anc2nc-TrpB_AM10) or Anc3nc-TrpA (for Anc3nc-TrpB_AM10) on the TrpB subunit. The results obtained with both variants confirmed that the exchange of the 10 *a-det* residues are sufficient to inverse the allosteric effect of TrpA on TrpB in both ways, from a deactivated TrpB to an activated TrpB (Anc2nc-TrpB vs. Anc2nc-TrpB_AM10) and from an activated TrpB to a deactivated TrpB (Anc3nc-TrpB vs. Anc3nc-TrpB_AM10) (**Figure 17A**).

3.3.2 Site-Directed Mutagenesis Identifies Four TrpB Positions That Are Essential for the Allosteric Effect

We next wanted to identify among the 10 *a-det* positions in Anc2nc-TrpB_AM10 those that are essential for the conversion of the allosteric effect from deactivating to activating. The 10 *a-det* positions are distributed all over the 3D structural model of Anc2nc-TrpB (**Figure 17B**), but can be assigned to one of four spatial clusters (A - D; **Figure 21A**). Cluster A comprises two residues close to the TrpA - TrpB interface. Cluster B contains four positions in the COMM domain, which has been shown to be significant for the allosteric communication between TrpA and TrpB (Schneider et al. 1998). The two positions of cluster C are located close to the active site, and the two positions of cluster D are located next to the TrpB homo-dimer interface. Starting from Anc2nc-TrpB_AM10, first residues of clusters A, C, and D were individually mutated back to their identities in Anc2nc-TrpB (**Table 8**). The protein variants with back-mutated cluster A (Anc2nc-TrpB_AM8-A) and cluster D (Anc2nc-TrpB_AM8-D) residues retained the activating allosteric effect, whereas the variant with back mutated cluster C (Anc2nc-TrpB_AM8-C) residues showed a deactivating allosteric effect (**Figure 21B**). These results indicated that cluster A and D residues are not relevant for the conversion of the allosteric effect from deactivating to activating. A variant with the combined back mutations of cluster A and D (Anc2nc-TrpB_AM6) residues confirmed this conclusion. In contrast, the residues from cluster C seem to contribute to the inversion of the allosteric effect. In a next step, the four residues of cluster B were individually back-mutated. The resulting variants Anc2nc-TrpB_AM5_M187I and Anc2nc-TrpB_AM5_D199E retained the activating allosteric effect as observed in Anc2-ncTrpB_AM10, whereas Anc2nc-

TrpB_AM5_R142M and Anc2nc-TrpB_AM5_V156I showed a deactivating allosteric effect.

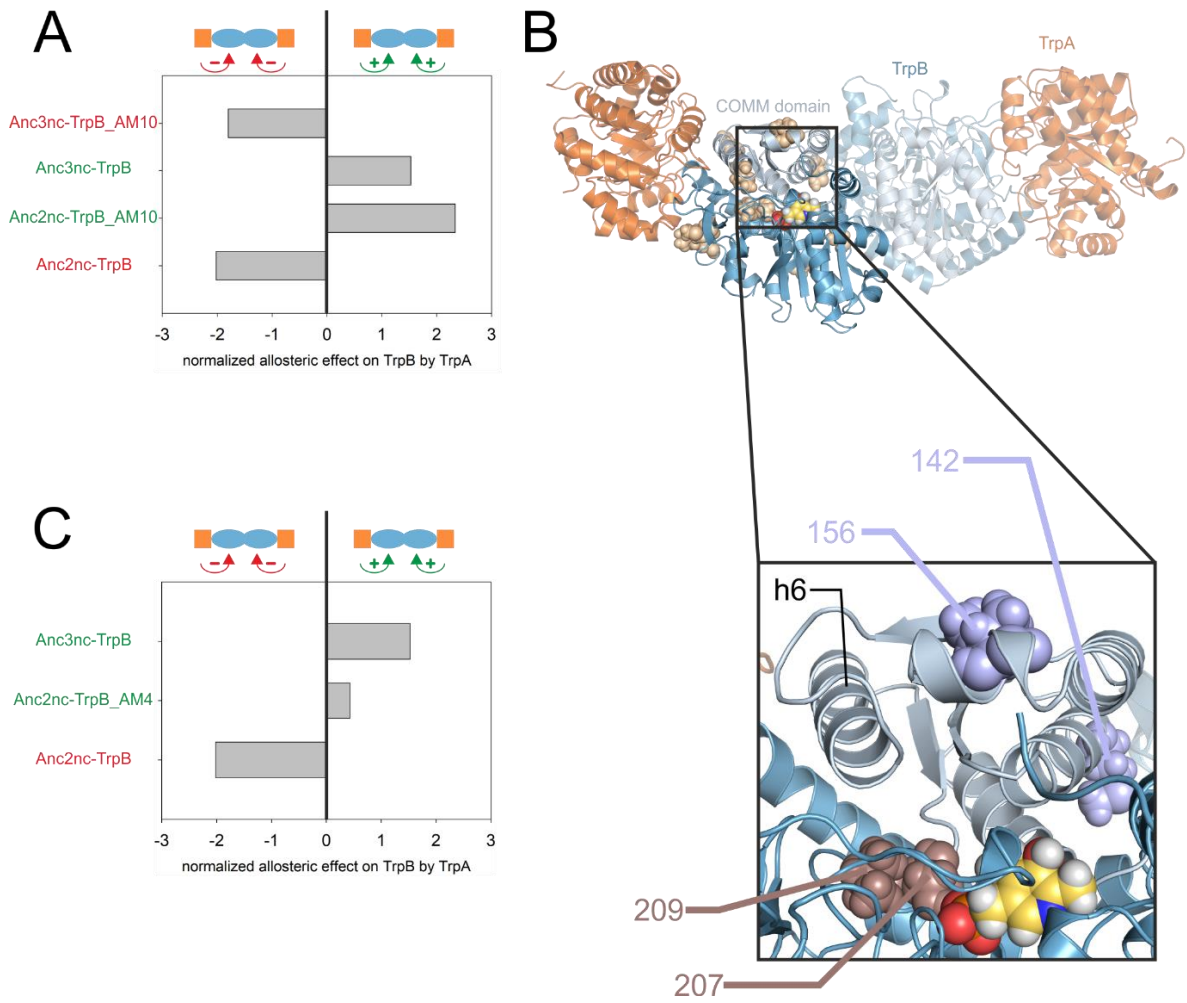


Figure 17: Mutational exchange of *a-det* and *a-ess* residues between Anc2nc-TrpB and Anc3nc-TrpB (**A**) The transfer of the ten *a-det* positions from Anc2nc-TrpB into Anc3nc-TrpB leads to a deactivating allosteric effect in Anc3nc-TrpB_AM10. Vice versa, the transfer of the ten *a-det* positions from Anc3nc-TrpB into Anc2nc-TrpB leads to an activating allosteric effect in Anc2nc-TrpB_AM10. The normalized allosteric effect is defined as the decadic logarithm of the TrpB activity in presence of TrpA divided by the TrpB activity in absence of TrpA. Raw data are given in **Table 10**. (**B**) Homology model of Anc2nc-TS. The ten *a-det* positions are shown as beige spheres. The active site is indicated by the spherical representation of the PLP cofactor colored yellow. Magnification: Detail of the homology model of Anc2nc-TrpB. The four *a-ess* positions responsible for the inversion of the allosteric effect are numbered and shown as spheres. The two positions from cluster B (cf. **Figure 21**) located in the COMM domain are colored in light blue, and the two positions from cluster C close to the active site (indicated by the PLP cofactor) are colored in brown. α-Helix 6 (h6) in TrpB is also marked. (**C**) The transfer of the four *a-ess* residues from Anc3nc-TrpB into Anc2nc-TrpB leads to an activating allosteric effect in Anc2nc-TrpB_AM4. The *a-ess* residues were identified by sequential back-mutation experiments (see the text for details). The normalized allosteric effect is defined as the decadic logarithm of the TrpB activity in presence of TrpA divided by the TrpB activity in absence of TrpA. Raw data are given in **Table 12**.

The four residues of Anc2nc-TrpB_AM10 that were shown to be essential for the inversion of the allosteric effect from deactivating to activating (*a-ess*) were then combined in the variant Anc2nc-TrpB_AM4. The activating residues R142 and V156 from cluster B are located within the COMM domain, while residues S207 and V209 from cluster C are close to the active site (**Figure 17B zoom**). As expected, Anc2nc-TrpB_AM4 showed an activating allosteric effect (**Figure 17C**; **Figure 22**). Further studies with Anc2nc-TrpB_AM3 variants where one of the four positions was back-mutated confirmed that the mutations in Anc2nc-TrpB_AM4 are necessary and sufficient for an activating allosteric effect.

3.3.3 The Effect of the Four Identified Residues is Independent of a Specific TrpB Background

To test whether the four identified residues are also effective in TrpB backgrounds other than Anc2nc-TrpB, variants of LBCA-TrpB and of Anc3nc-TrpB were created. In LBCA-TrpB_AM4 the four *a-ess* residues were mutated to their respective Anc3nc-TrpB identities. In Anc3nc-TrpB_AM4, the four *a-ess* residues were mutated to their respective Anc2nc-TrpB identities. For both variants, we could show the inversion of the allosteric effect by the introduced mutations. While LBCA-TrpB is deactivated by LBCA-TrpA, the variant LBCA-TrpB_AM4 is activated upon binding of LBCA-TrpA. On the other hand, Anc3nc-TrpB is activated by Anc3nc-TrpA, but the variant Anc3nc-TrpB_AM4 is deactivated upon binding of Anc3nc-TrpA (**Figure 18**). These results show that the inversion of the allosteric effect of TrpA on TrpB is also possible in other TrpB subunits than Anc2nc-TrpB, meaning that the role of the four positions is independent of a specific TrpB background.

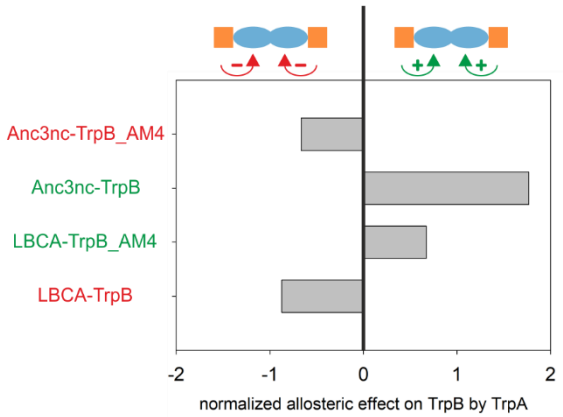


Figure 18: Mutational exchange of *a-ess* residues between LBCA-TrpB and Anc3nc-TrpB

The transfer of the four *a-ess* residues from Anc3nc-TrpB into LBCA-TrpB leads to an activating allosteric effect in LBCA-TrpB_AM4. Vice versa, the transfer of the four *a-ess* residues from LBCA-TrpB into Anc3nc-TrpB leads to a deactivating allosteric effect in Anc3nc-TrpB_AM4. The normalized allosteric effect is defined as the decadic logarithm of the TrpB activity in presence of TrpA divided by the TrpB activity in absence of TrpA. Raw data are given in **Table 11**.

3.3.4 Biochemical Characterization of the *a-ess* TrpB Positions

To unravel the mechanism by which the residues in Anc2nc-TrpB_AM4 inverse the allosteric effect compared to Anc2nc-TrpB, comparative experimental and computational studies were performed. Analytical size-exclusion chromatography showed that the four residue differences between Anc2nc-TrpB and Anc2nc-TrpB_AM4 do not influence their ability to form a stable complex with Anc2nc-TrpA (**Figure 19**).

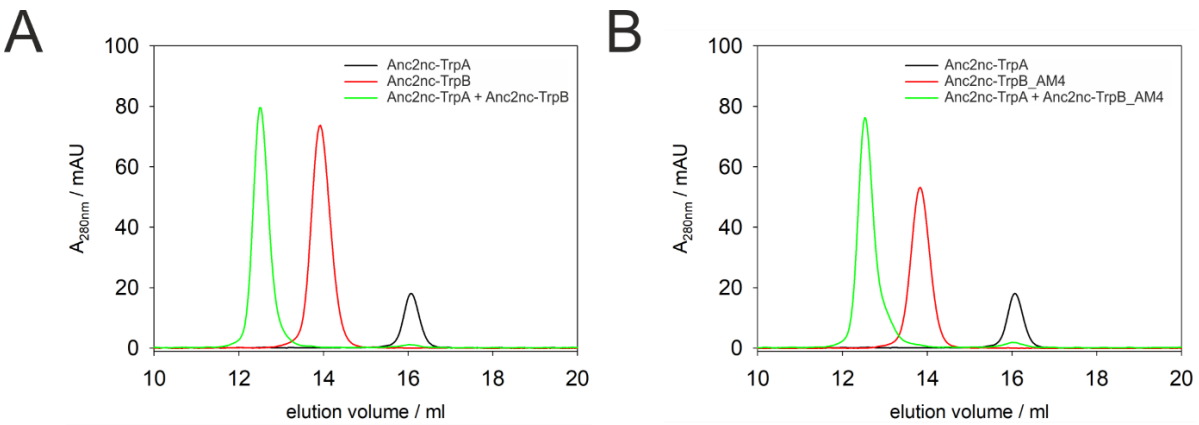


Figure 19: Structural complex formation of Anc2nc-TrpA with Anc2nc-TrpB and Anc2nc-TrpB_AM4 as observed by analytical gel filtration chromatography

(**A**) Elution profiles of Anc2nc-TrpA, Anc2nc-TrpB, and a mixture of the two proteins. (**B**) Elution profiles of Anc2nc-TrpA, Anc2nc-TrpB_AM4, and a mixture of the two proteins. The proteins (30 µM monomer concentration for each sample) were loaded on a Superdex S200 column that was equilibrated with 50 mM potassium phosphate (pH 7.5) and 300 mM KCl and eluted at a flow rate of 0.3 ml/min. Elution was followed by following the absorbance at 280 nm.

For a more detailed study of the allosteric effect, steady-state kinetic data of Anc2nc-TrpB and Anc2nc-TrpB_AM4 in absence and presence of Anc2nc-TrpA were recorded (**Table 5**). Anc2nc-TrpA reduces the turnover number (k_{cat}-value) of Anc2nc-TrpB about 35 - 65-fold but does not change significantly the Michaelis constants $K_m^{L\text{-ser}}$ and K_m^{indole}. In contrast, Anc2nc-TrpA increases the k_{cat}-value of Anc2nc-TrpB_AM4 about 2-fold whereas $K_m^{L\text{-ser}}$ and K_m^{indole} also do not change much.

Table 5: Steady-state kinetic parameters for the TrpB reactions of Anc2nc-TrpB and Anc2nc-TrpB_AM4, in absence and presence of Anc2nc-TrpA

	L-serine dependent			Indole dependent		
Protein(s)	k_{cat} / s^{-1}	$K_m^{L\text{-}ser}$ / mM	$k_{cat}/K_m^{L\text{-}ser}$ / $M^{-1}s^{-1}$	k_{cat} / s^{-1}	K_m^{Indole} / µM	k_{cat}/K_m^{Indole} / $M^{-1}s^{-1}$
Anc2nc-TrpB	0.3533 ± 0.0053	0.73 ± 0.06	$4.82*10^2 \pm 4.40*10^1$	0.1928 ± 0.0079	2.52 ± 0.76	$7.65*10^1 \pm 2.62*10^1$
Anc2nc-TrpB + Anc2nc-TrpA	0.0055 ± 0.0003	1.30 ± 0.50	$4.24*10^0 \pm 1.87*10^0$	0.0055 ± 0.0004	6.14 ± 2.20	$8.91*10^{-1} \pm 3.81*10^{-1}$
Anc2nc-TrpB_AM4	0.2464 ± 0.0155	3.99 ± 1.03	$6.18*10^1 \pm 1.97*10^1$	0.2372 ± 0.0195	16.63 ± 4.16	$1.43*10^1 \pm 4.74*10^0$
Anc2nc-TrpB_AM4 + Anc2nc-TrpA	0.4744 ± 0.0216	3.79 ± 0.85	$1.25*10^2 \pm 3.38*10^1$	0.5012 ± 0.0317	26.00 ± 5.20	$1.93*10^1 \pm 5.08*10^0$

Experimental conditions: For L-serine dependent kinetics: 100 mM potassium phosphate buffer pH 7.5, 180 mM KCl, 40 µM PLP, 200 µM indole and varying concentrations of L-serine in presence of 4 µM Anc2nc-TrpB (and if indicated 10 µM Anc2nc-TrpA), or 0.25 µM Anc2nc-TrpB_AM4 (and if indicated) 2.5 µM Anc2nc-TrpA. For indole dependent kinetics: 100 mM potassium phosphate buffer pH 7.5, 180 mM KCl, 40 µM PLP, 50 mM L-serine and varying concentrations of indole in presence of 4 µM Anc2nc-TrpB (and if indicated 10 µM Anc2nc-TrpA), or 0.25 µM Anc2nc-TrpB_AM4 (and if indicated) 2.5 µM Anc2nc-TrpA. All enzyme concentrations are monomer concentrations. The reaction was followed at 30 °C by monitoring the condensation of L-serine and indole to L-tryptophan at 290 nm. The mean and the standard errors were calculated from at least three independent measurements.

3.3.5 *In silico* analysis of the *a-ess* mutations

The dynamic network model (Sethi et al. 2009) assumes that the propagation of allosteric signals in a protein complex depends on the communication between “communities”. These structures of sub-domain size can be identified by a correlation analysis of the motions observed during a molecular dynamics (MD) simulation. We were interested to find out, how the *a-ess* mutations affect the organization of TrpB communities and consequently the propagation of the allosteric signal coming from TrpA. Thus, we performed MD simulations of Anc2nc-TS_AM4 and Anc2nc-TS, which we considered representative for the allosterically activated and the deactivated state, respectively.

Due to the lack of crystal structures, we derived a homology model for the full complex Anc2nc-TS by means of Yasara (Krieger et al. 2009) and introduced the corresponding mutations by means of PyMOL (Schrodinger 2015) to create a model of the Anc2nc-TS_AM4 complex. For both models, a MD simulation of 3 µs duration was performed, which consisted of 100 ns runs computed for 30 replicates. We then utilized a standardized protocol (VanWart et al. 2012) to identify for the two TrpB variants all communities consisting of more than two residues.

Generally, networks that are well suited for topology analysis contain 2 - 10 communities (Yao et al. 2019). This requirement is met by our MD simulations, because the dynamics of the activated state (Anc2nc-TS_AM4) gave rise to seven dominant communities named *asc*_1 to *asc*_7 in the following (**Figure 20A, right panel**). Interestingly, *asc*_5 contains nearly all residues of the COMM domain (residues 110 – 206 in Anc2nc-TS_AM4), which is a critical element of TrpB activation (Schneider et al. 1998). Thus, the residue composition of *asc*_5 confirmed that this analysis of our MD simulations is biochemically plausible. The comparison of the communities *dsc*_1 to *dsc*_7 of the deactivated state (Anc2nc-TS) (**Figure 20A, left panel**) with *asc*_1 to *asc*_7 indicated a high degree of overlap in the residue composition of the activation-specific communities. However, we also noticed the existence of some smaller communities whose residue compositions were specific for the activated or the deactivated state (**Figure 20B**).

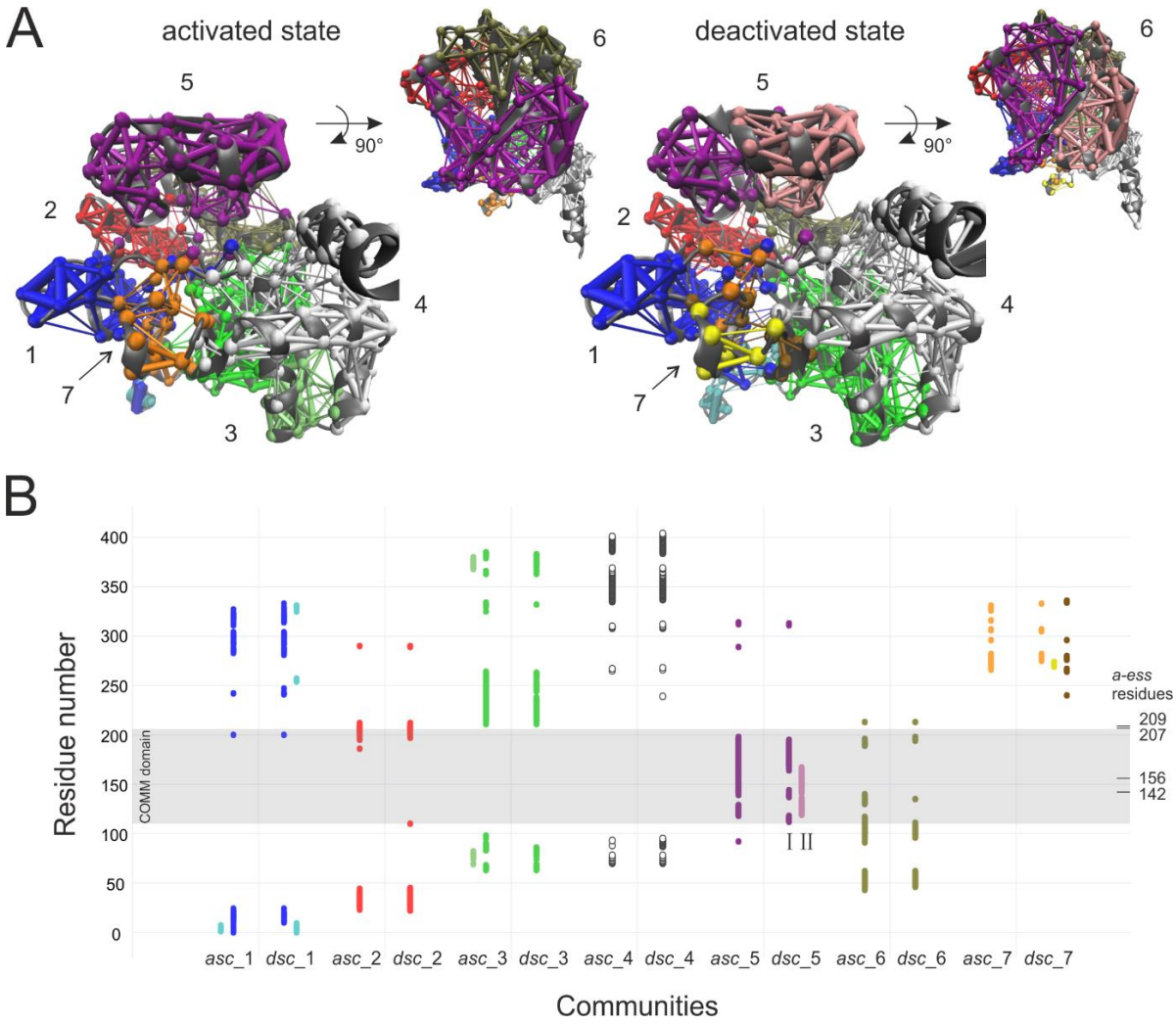


Figure 20: Identification of TrpB communities induced by correlated motions in MD simulations

(**A**) Communities identified by analyzing MD simulations of the activated (Anc2nc-TS_AM4) and the deactivated (Anc2nc-TS) state. Communities, whose residue composition overlap to a great extent are shown in the same color. The diameter of edges interconnecting residue positions *i* and *j* corresponds to the value -log|c_{ij}|, where c_{ij} is the correlation between residue *i* and *j*. (**B**) Residue-specific representation of community compositions. Communities observed in the activated state (*asc*) and the deactivated state (*dsc*) are color coded as in (**A**). Note that the disintegration of communities gave rise to smaller communities that are positioned next to the major ones. The localization of the COMM domain and the *a-ess* residues is indicated. The COMM domain residues constitute in the activated state the community *asc*_5 and disintegrate in the deactivated state into the two smaller communities *dsc*_5-I and *dsc*_5-II.

One of the most pronounced differences is the separation of *asc*_5 into two smaller communities in the deactivated state *dsc*_5 (named I and II in **Figure 20B**). This finding suggests that the four *a-ess* mutations turn the intrinsically rigid but moveable COMM domain (Kulik et al. 2002) into two, less coordinated structural elements. Most plausibly, this decoupling impedes the propagation of the allosteric signal from helix h6 of the COMM domain in TrpB to its catalytic center (**Figure 17B zoom**).

3.4 Discussion

3.4.1 Vertical Versus Horizontal Approaches for the Analysis of Sequence-Function Relationships

One of the central objectives in biochemistry is to uncover the relationship between amino acid sequence and protein function. A conventional approach to achieve this goal is to align the sequences of homologous proteins and to use site-directed mutagenesis to test the effect of individual residue differences for functional divergence. However, such an approach based on the horizontal comparison of extant proteins is facing two major problems: First, due to neutral drift, even closely related proteins often share sequence identities of less than 50 %, which corresponds to 100 different positions in a small protein of 200 amino acids. Second, the effect of certain amino acids depends on the context within the protein, a property that is called epistasis (Starr and Thornton 2016). Both neutral drift and epistasis significantly impair the identification of essential amino acids by site-directed mutagenesis. This is especially true for complex functional properties such as allostery as observed in metabolic enzyme complexes, which is most probably realized by the interplay of networks of interconnected amino acids. An elegant alternative to such cumbersome classical horizontal approaches is a vertical comparison of sequences that represent the nodes of a phylogenetic tree of a given protein. These sequences are determined with the help of ASR (Merkl and Sterner 2016, Hochberg and Thornton 2017), followed by the expression of the corresponding synthetic genes and the characterization of the recombinant proteins. Since the sequences of adjacent nodes in a phylogenetic tree are generally more similar to each other than the sequences of extant proteins, the identification of residues crucial for functional differences is more straightforward (Holinski et al. 2017, Siddiq et al. 2017).

In our case of TS, the vertical approach proved to be of great value for the identification of residues that are essential for the allosteric regulation of TrpB by TrpA. We started with two variants in which the allosteric effect was either deactivating (LBCA-TS) or activating (ncTS) and faced the practically insolvable problem to assign this functional disparity to one or more out of 261 amino acid differences (137 differences in the TrpA sequences / 111 differences and 13 insertions/deletions in TrpB sequences). Moreover, since known positions within TrpB that affect allostery are conserved (for example: R141 in stTrpB / R153 in ncTrpB / R138 in LBCA-TrpB or D305 in stTrpB / D317 in ncTrpB / D302 in LBCA-TrpB) (Dunn 2012), there was no straightforward way to identify crucial positions responsible for the inversion of the allosteric effect of TrpA on TrpB. However,

our vertical approach allowed us to narrow down the number of amino acids responsible for this functional difference to only four residues, all of which were located in TrpB. It is important to note that we cannot exclude that the deactivating allosteric effect observed in LBCA-TS, Anc1nc-TS, and Anc2nc-TS is an artifact of the ASR. However, this uncertainty does not compromise the utility of these proteins for the analysis of the structure-function relationship of TS described in this study.

3.4.2 Function of the Four *a-ess* Residues That Are Necessary and Sufficient for the Inversion of the Allosteric Effect

A major factor in allosteric activation of the TrpB subunit by TrpA is the shift of the equilibrium from a rather inactive open conformation towards the active closed conformation. Moreover, it has been shown that certain mutations in the TrpB subunit (e.g. R141A or D305A in stTrpB) are able to disrupt its allosteric activation by TrpA and modify the effect of monovalent cations catalytic activity (Ferrari et al. 2003). We have now identified four positions (142, 156, 207, and 209) within the TrpB subunit that exert a drastic effect on the allosteric manipulation by TrpA but have not been analyzed so far. In order to elucidate the molecular function of these *a-ess* positions, analytical gel filtration, steady-state enzyme kinetics and *in silico* analysis were performed. In terms of structural complex formation the four *a-ess* positions seem to be neutral as Anc2nc-TrpB (deactivating allosteric effect) and Anc2nc-TrpB_AM4 (activating allosteric effect) both assemble with Anc2nc-TrpA to the regular $\alpha\beta\beta\alpha$ TS (**Figure 18**). Steady-state enzyme kinetics confirmed that the k_{cat}-value of Anc2nc-TrpB is decreased 35-65 fold by Anc2nc-TrpA, while the k_{cat}-value of Anc2nc-TrpB_AM4 is enhanced 2-fold in presence of Anc2nc-TrpA. In contrast, the $K_m^{L\text{-}Ser}$ and K_m^{Indole} values are not influenced significantly by Anc2nc-TrpA, neither for Anc2nc-TrpB nor for Anc2nc-TrpB_AM4 (**Table 5**). These results show that the residues at the four *a-ess* positions do not change the way by which substrate binding by TrpB is modulated by TrpA. However, these residues drastically change the way by which catalysis by TrpB is influenced by TrpA.

Interestingly, all four positions are located within the COMM domain or close the active site. In agreement with their localization, our *in silico* studies further strengthened the role of COMM domain conformations for the activation of the TrpB subunit. Most important is the “closed” conformation that induces the highly active state of the TrpB subunit. The existence of two communities I and II in the deactivated state that splits the COMM domain, suggest a less coordinated movement of this domain in Anc2nc-TrpB.

Taken together, the vertical approach used in this study provided novel and unexpected insights into the allosteric communication within TS. Specifically, four residues in TrpB could be identified that had escaped attention in spite of intense research on this fascinating enzyme complex during the past decades. We trust that vertical approaches will prove to be highly valuable for the analyses of sequence-function relationship of other proteins in the future.

Acknowledgments

We thank Christiane Endres and Sonja Fuchs for excellent technical assistance.

3.5 Methods

3.5.1 Cloning

The genes LBCA-*trp*A und LBCA-*trp*B were previously cloned into pET21a(+) using *Nde*I/*Xho*I restriction sites for expression with a C-terminal His_6-tag (Busch et al. 2016). The genes *anc1nc-trp*A, *anc1nc-trp*B, *anc2nc-trp*A, *anc2nc-trp*B, *anc2nc-trp*B_AM10, *anc3nc-trp*A, *anc3nc-trp*B, *anc3nc-trp*B_AM10, *anc4nc-trp*A, *anc4nc-trp*B, *anc5nc-trp*A, *anc5nc-trp*B, *anc6nc-trp*A, *anc6nc-trp*B, *nctrp*A (GenBank: EAR61762.1) and *nctrp*B (GenBank: EAR61761.1) were codon-optimized for expression in *E.coli* and synthesized by Thermo Fisher Scientific (GeneArt Strings DNA Fragments). *Anc1nc-trp*A, *anc2nc-trp*A, *anc3nc-trp*A, *anc4nc-trp*A and *nctrp*B were cloned into pUR28_BsaI using *Bsa*I restriction sites (Rohweder et al. 2018) for expression with a N-terminal His_6-tag. *Anc1nc-trp*B, *anc2nc-trp*B, *anc2nc-trp*B_AM10, *anc3nc-trp*B, *anc3nc-trp*B_AM10, *anc4nc-trp*B, *anc5nc-trp*B, *anc6nc-trp*B, *anc3nc-trp*B_AM4 and LBCA-*trp*B_AM4 were cloned into pET21a_BsaI using *Bsa*I restriction sites for expression with a C-terminal His_6-tag. *Anc5nc-trp*A, *anc6nc-trp*A and *nctrp*A were cloned into pMAL-c5T_BsaI using *Bsa*I restriction sites for expression with a N-terminal His_6-maltose binding protein/MBP-tag.

3.5.2 Site-Directed Mutagenesis

The variants studied in this work were generated starting from *anc2nc-trp*B_AM10 using QuikChange™ site-directed mutagenesis (QCM) (Wang and Malcolm 1999) or the method according to the NEB Q5® site-directed mutagenesis kit (PNK). The templates, primer pairs, and the methods for the generation of all variants are listed in **Table 6**. For all constructs Sanger sequencing was applied to confirm the desired sequence.

3.5.3 Gene Expression and Protein Purification

E.coli BL21 (DE3) Gold cells were transformed with respective plasmids. The cells were grown in lysogeny broth medium supplemented with 150mg/ml ampicillin in case of pET21 and pMAL-c5T plasmids or 75 mg/ml kanamycin in case of pUR28 plasmids; 20 mM potassium phosphate buffer (pH 7.5) was added in case of *trp*A constructs; 20 µM PLP was added in case of *trp*B constructs. At a cell density of OD_{600} = 0.6, protein expression was induced by addition of 0.5 mM isopropyl-β-thiogalactopyranoside. After growth overnight at 20 °C, the cells were harvested by centrifugation (Beckman Coulter Avanti J-26S XP, JLA-8.1, 15 min, 4000 rpm, 4 °C). Cell pellets were suspended in 100 mM potassium phosphate buffer (pH 7.5), 300 mM KCl and 10 mM imidazole for

purification of TrpA proteins or 50 mM potassium phosphate buffer (pH 7.5), 300 mM KCl, 10 mM imidazole and 40 µM PLP for purification of TrpB proteins. The cells were disrupted by sonification (Branson Sonifier W-250D, amplitude 50%, 2 x 2 min, 2 s pulse/2 s pause). The insoluble cell fraction was removed by centrifugation (Beckman Coulter Avanti J-26S XP, JA-25.50, 45 min, 14000 rpm, 4 °C). The soluble fractions of Anc1nc-TrpA, Anc1nc-TrpB, Anc4nc-TrpA and Anc4nc-TrpB underwent heat treatment (75 °C for 20 min), denatured proteins were removed by centrifugation (Beckman Coulter Avanti J-26S XP, JA-25.50, 45 min, 14000 rpm, 4 °C). The supernatants were subjected to purification by immobilized metal ion affinity chromatography (IMAC) (GE Healthcare, HisTrap FF Crude). The proteins of interest were eluted in 100 mM potassium phosphate buffer (pH 7.5) in case of TrpA proteins or 50 mM potassium phosphate buffer (pH 7.5) in case of TrpB proteins, with 300 mM KCl and a linear gradient of 10 to 1000 mM imidazole. The fractions with enriched protein of interest were pooled and subjected to preparative size-exclusion chromatography (SEC) (GE Healthcare, HiLoad Superdex 75 pg column for TrpA proteins or GE Healthcare, HiLoad Superdex 200 pg column for TrpB proteins). The proteins were eluted in 100 mM potassium phosphate (pH 7.5) in case of TrpA proteins or in 50 mM potassium phosphate (pH 7.5) in case of TrpB proteins, with 300 mM KCl. The soluble fractions of MBP-Anc5nc-TrpA, MBP-Anc6nc-TrpA and MBP-ncTrpA were subjected to purification by IMAC (GE Healthcare, HisTrap FF Crude). The fusion proteins consisting of MBP and the proteins of interest were eluted in 100 mM potassium phosphate buffer (pH 7.5) with 300 mM KCl and a linear gradient of 10 to 500 mM imidazole. The fractions with enriched fusion protein were pooled and 5 U/ml thrombin was added. For protease cleavage, the solution was subjected to dialysis against 100 mM potassium phosphate buffer (pH 7.5) with 300 mM KCl over night at 4 °C. After dialysis the protein solution was subjected to IMAC (GE Healthcare, HisTrap FF Crude). The flow-through fraction contained the proteins of interest, was collected and subjected to preparative SEC (GE Healthcare, HiLoad Superdex 75 pg column). The proteins were eluted in 100 mM potassium phosphate (pH 7.5). The fractions containing purified protein were pooled and used for further experiments. Protein concentrations in case of TrpA proteins were determined by absorbance spectroscopy at 280 nm (Thermo Scientific, NanoDrop One) using molar extinction coefficients that were calculated from the amino acid sequence (Gasteiger E. 2005), and by using a commercial Bradford reagent (Bio-Rad, Bradford protein assay) for TrpB proteins.

3.5.4 HPLC Assay for Recording of TrpB Activity

The conversion of L-serine and indole to L-tryptophan by TrpB in presence or absence of TrpA was observed by a discontinuous assay. The reactions contained 0.5 µM TrpB and 0 or 1 µM TrpA, 45 mM L-serine, 100 mM EPPS/KOH (pH 7.5), 180 mM KCl and 40 µM PLP. The reactions were initiated by addition of 300 µM indole and incubated for 2 min at 30 °C. Reactions were stopped by addition of 0.5 volume of potassium hydroxide solution and 1 volume of methanol. The reaction products were subsequently analyzed by reversed-phase HPLC using an Agilent system (1200 series) with a C18 column (Phenomenex, Luna 5 µm C18(2) 100Å, 150 x 3 mm). The separation was performed according to (Busch et al. 2016). The elution time of L-tryptophan was 16.86 min. L-Tryptophan was quantified according to a calibration curve with different concentrations of L-tryptophan detected by fluorescence (λ_{ex} = 278 nm, λ_{em} = 350 nm).

3.5.5 Steady-State Enzyme Kinetics of TrpB Enzymes and Activity Titrations by Detection of TrpB Activity

The conversion of L-serine and indole to L-tryptophan by TrpB was directly observed at 30 °C by absorbance spectroscopy (V-650 spectrophotometer, JASCO) using $\Delta\varepsilon_{290}$(tryptophan-indole) = 1.89 $mM^{-1}cm^{-1}$. (Faeder and Hammes 1970) Initial velocities were recorded in 100 mM potassium phosphate (pH 7.5), 180 mM KCl and 40 µM PLP with saturating concentrations of L-serine or indole. For steady-state kinetics, the initial velocities were plotted as a function of the L-serine or indole concentration and fitting a hyperbolic function to the obtained data points (SigmaPlot 12.5). For activity titration experiments, the initial velocities were plotted as a function of the added TrpA concentration the resulting data points were fitted with a hyperbolic function with an offset for the TrpB activity in absence of TrpA.

3.5.6 Analytical Size-Exclusion Chromatography

Analytical size-exclusion chromatography was performed at room temperature with a chromatography system (GE Healthcare, ÄKTAmicro) coupled with an autosampler (Spark Holland, Alias GE Bio Cool) providing a cooled sample compartment at 4 °C. Proteins were eluted from an analytical size exclusion column (GE Healthcare , Superdex S200, 10/300 GL) in 50 mM potassium phosphate (pH 7.5) and 300 mM KCl at a flow rate of 0.3 ml/min. Elution was followed by an absorbance detector at 280 nm.

3.5.7 Homology Modelling

To compute a 3D structure of Anc2nc-TS, the homology modeling module of YASARA (version 17.4.17) (Krieger et al. 2009) was used. In brief, the corresponding protocol consisted of the following steps: To begin with, two templates were added manually, namely LBCA-TS (PDB-ID 5ey5) and stTS (PDB-ID 2j9x). Additional three templates were identified by PSI-BLASTing (Altschul et al. 1997) the target sequence against UniRef90. The resulting position-specific scoring matrices were then used to search the PDB, which resulted in the three templates with PDB-IDs 5ey5, 5kin, and 5tch. All five templates were ranked based on the alignment score and their structural quality assessed by means of WHAT_CHECK (Hooft et al. 1996). Analogously, a profile was deduced. Target and template proteins were aligned based on the corresponding profiles and additional structural information contained in the templates. Then, initial side-chain rotamer configurations were determined by using dead-end elimination in the context of a repulsive function (Canutescu et al. 2003). YASARA optimized loops by testing a large number of different conformations and by fine-tuning the orientation of side-chains based on electrostatic and knowledge-based interactions. Subsequently, for each of the models, the hydrogen bonding network was optimized and a high-resolution energy minimization was run to confirm the stability of each structure. Finally, YASARA determined model quality based on scores determined for each residue; the structural model had an acceptable Z-score of -0.253.

The homology model for Anc2nc-TS_AM4 was created by introducing the four Anc2nc-TS_AM4 specific mutations into Anc2nc-TS via PyMOL (Schrodinger 2015). For each substitution, the rotamer was chosen that induced the fewest steric clashes.

3.5.8 MD Simulation

MD simulations with the homology models of Anc2nc-TS and Anc2nc-TS_ AM4 were performed with GROMACS (Berendsen et al. 1995) (version 2016.4), and the AMBER03 force field. The complex was placed in a rectangular water box; the simulation cell was at least 5 Å larger than the solute along each axis and the system was neutralized by adding NaCl ions. To prepare a production run, the solvated system was at first energy minimized by a maximum of 5000 steps of steepest descent MD minimization followed by a two-part equilibration phase (each lasted 50 ps) that began with an NVT simulation followed by an NPT simulation. Finally, the production run (NVT conditions, time step 1 fs) was started at 298 K and ran for 100 ns, generating a snapshot every 0.1 ns. During

the simulation, the temperature was kept constant at 298 K by using a Berendsen thermostat; for detail see GROMACS documentation. 30 replicates were generated for each complex resulting in a total of 3µs simulation time each.

3.5.9 Community Analysis

All replicates from the MD simulation were pooled and a cross correlation matrix for Anc2nc-TS and Anc2nc-TS_AM4 based on the Cα trajectories was generated with the help of the VMD (Humphrey et al. 1996) plugin NetworkView as described in (Sethi et al. 2009). The content of this matrix specifies a network of nodes and edges and contains substructures (communities of nodes) that are more densely interconnected to each other than to the other nodes of the network. Per definition, residues belonging to the same community can communicate with each other through multiple routes. The structure of the communities was determined by means of the Girwan-Newman (Girvan and Newman 2002) algorithm implemented in NetworkView. Communities with less than 3 residues were discarded.

3.6 Supplementary Figures

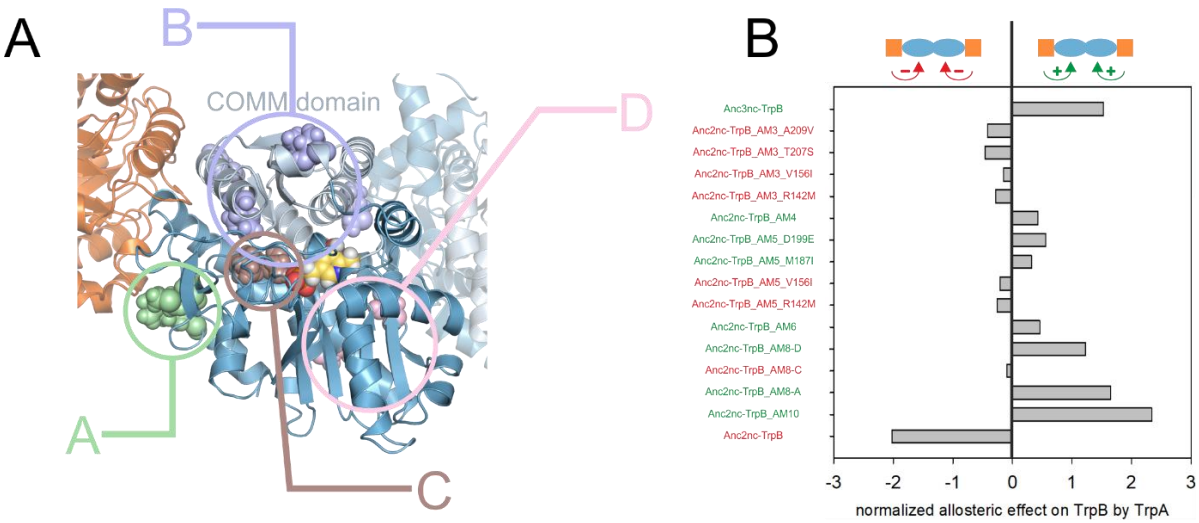


Figure 21: Structural clustering of 10 *a-det* positions in TrpB and subsequent identification of four *a-ess* positions

(**A**) Homology model of Anc2nc-TS with TrpB colored in blue and TrpA colored in orange. The 10 *a-det* positions are represented as spheres, colored according to the clusters they belong to. Cluster A (colored in light green) is located next to the TrpA-TrpB interface. Cluster B (colored in light blue) is positioned within the COMM domain. Cluster C (colored in brown) is located close to the active site, which is indicated by the spherical representation of the PLP cofactor (colored in yellow). Cluster D (colored in pink) lies close to the TrpB-TrpB homodimer interface. (**B**) Assay for the allosteric effect of TrpA on TrpB. All protein variants that contributed to the identification of the four relevant positions *a-ess* are shown. The normalized allosteric effect is defined as the decadic logarithm of the TrpB activity in presence of TrpA divided by the TrpB activity in absence of TrpA. Raw data are given in **Table 12**.

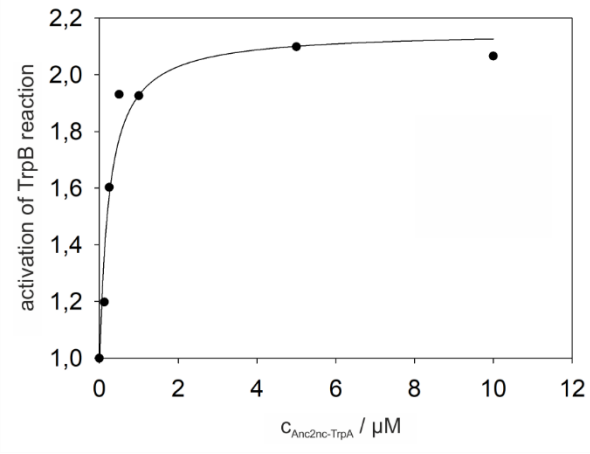


Figure 22: Activity titration of Anc2nc-TrpB_AM4 with Anc2nc-TrpA

TrpB_AM4 (0.5 µM, monomer concentration) was titrated with varying concentrations of Anc2nc-TrpA. The experiment was performed at 30 °C in presence of 100 mM KP pH 7.5, 180 mM KCl, 40 µM PLP, 200 µM indole, 50 mM L-serine. TrpB activity was followed by monitoring the condensation of indole and L-serine to L-tryptophan at a wavelength of 290 nm.

3.7 Supplementary Tables

Table 6: Primer used for site-directed mutagenesis

Target construct	Template	Introduced mutation	forward/ reverse	Primer sequence 5'→3'	Method
pET21a_Anc2nc-TrpB_AM10_H24R	pET21a_Anc2nc-TrpB_AM10	H24R	forward reverse	**CGT**TTTGGTCCGTATGGT ACCACGTGCATCCGG	PNK
pET21a_ Anc2nc-TrpB_AM8-A	pET21a_ Anc2nc-TrpB_AM10_H24R	R300M	forward reverse	**ATG**ACCTATCTGCTGC GCTACCATGCAGAAC	PNK
pET21a_ Anc2nc-TrpB_AM10_T207S	pET21a_ Anc2nc-TrpB_AM10	T207S	forward reverse	ACATTATTGGC**TCT**GTTGCCGG CCGGCAAC**AGA**GCCAATAATGT	QCM
pET21a_ Anc2nc-TrpB_AM8-C	pET21a_ Anc2nc-TrpB_AM10_T207S	A209V	forward reverse	TGGCTCTGTT**GTT**GGTCCGCAT ATGCGGACC**AAC**AACAGAGCCA	QCM
pET21a_ Anc2nc-TrpB_AM10_ S73T	pET21a_ Anc2nc-TrpB_AM10	S73T	forward reverse	GTCGTCCG**ACC**CCGCTGTATT AATACAGCGG**GGT**CGGACGAC	QCM
pET21a_ Anc2nc-TrpB_AM8-D	pET21a_ Anc2nc-TrpB_AM10_ S73T	R227E	forward reverse	GCGTTATTGGT**GAA**GAAGCACGTC GACGTGCTTC**TTC**ACCAATAACGC	QCM
pET21a_ Anc2nc-TrpB_AM7-A_ S73T	pET21a_ Anc2nc-TrpB_AM8-A	S73T	forward reverse	GTCGTCCG**ACC**CCGCTGTATT AATACAGCGG**GGT**CGGACGAC	QCM
pET21a_ Anc2nc-TrpB_AM6	pET21a_ Anc2nc-TrpB_AM7-A_ S73T	R227E	forward reverse	GCGTTATTGGT**GAA**GAAGCACGTC GACGTGCTTC**TTC**ACCAATAACGC	QCM
pET21a_ Anc2nc-TrpB_AM5_R142M	pET21a_ Anc2nc-TrpB_AM6	R142M	forward reverse	CGTTGCAGCA**ATG**TTTGGTCTGGAAT ATTCCAGACCAAA**CAT**TGCTGCAACG	QCM
pET21a_ Anc2nc-TrpB_AM5_V156I	pET21a_ Anc2nc-TrpB_AM6	V156I	forward reverse	GTGCCGAAG**ATA**TTGAACGTCAG CTGACGTTCAA**TAT**CTTCGGCAC	QCM
pET21a_ Anc2nc-TrpB_AM5_M187I	pET21a_ Anc2nc-TrpB_AM6	M187I	forward reverse	TGAAAGATGCA**ATT**AATGAAGCAATG CATTGCTTCATT**AAT**TGCATCTTTCA	QCM
pET21a_ Anc2nc-TrpB_AM5_D199E	pET21a_ Anc2nc-TrpB_AM6	D199E	forward reverse	TACCAATGTG**GAA**GATACCTATTAC GTAATAGGTATC**TTC**CACATTGGTA	QCM
pET21a_ Anc2nc-TrpB_AM4	pET21a_ Anc2nc-TrpB_AM5_M187I	D199E	forward reverse	TACCAATGTG**GAA**GATACCTATTAC GTAATAGGTATC**TTC**CACATTGGTA	QCM
pET21a_ Anc2nc-TrpB_AM3_T207S	pET21a_ Anc2nc-TrpB_AM4	T207S	forward reverse	ACATTATTGGC**TCT**GTTGCCGG CCGGCAAC**AGA**GCCAATAATGT	QCM
pET21a_ Anc2nc-TrpB_AM3_A209V	pET21a_ Anc2nc-TrpB_AM4	A209V	forward reverse	GCACCGTT**GTG**GGTCCGC GCGGACC**CAC**AACGGTGC	QCM
pET21a_ Anc2nc-TrpB_AM3_ R142M	pET21a_ Anc2nc-TrpB_AM4	R142M	forward reverse	CGTTGCAGCA**ATG**TTTGGTCTGGAAT ATTCCAGACCAAA**CAT**TGCTGCAACG	QCM
pET21a_ Anc2nc-TrpB_AM3_ V156I	pET21a_ Anc2nc-TrpB_AM4	V156I	forward reverse	GTGCCGAAG**ATA**TTGAACGTCAG CTGACGTTCAA**TAT**CTTCGGCAC	QCM

Bases in **bold** and underlined indicate mutated positions.

Table 7: Amino acid sequences of experimentally characterized proteins

Construct	Protein	Studied amino acid sequence
pET21a_LBCA-TrpA (Busch et al. 2016)	LBCA-TrpA	MNRIAEAFEELKKKGEKALIPFITAGDPDLETTLELVRALVEAGADIIELGIPFSDPLADGPTIQRASQRALA SGTTLDKVFEMVRELREKNTDVPIVFLTYYNPIFRYGIERFVKECAEAGVDGLIVPDLPPEEAADLAAAAE KYGVDLIFLVAPTSTDERIKMIAKHASGFVYCVSVTGVTGARSEIAADLAELVSRIRKHTDLPIAVGFGIST PEQAAEVAQVADGVIVGSAIVKRIEENQDEEDIVEEVREFVRELREAVKLEHHHHHH
pET21a_LBCA-TrpB (Busch et al. 2016)	LBCA-TrpB	MIGRFGKYGGQYVPETLMPALEELEEAYERAKNDPEFQAELEYYLRDYVGRPTPLYFAENLTKDLGGA KIYLKREDLNHTGAHKINNALGQALLAKRMGKKRVIAETGAGQHGVATATVAAMFGLECVVYMGAEDIE RQALNVFRMKLLGAKVRPVTSGSRTLKDAINEAMRDWVTNVEDTFYIIGSVVGPHPYPMMVRDFQSVI GEEARQQILEKEGRLPDAIVACVGGGSNAMGIFHPFIDDESVRLIGVEAAGKGIETGKHAATLSAGRPGV LHGAMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHAYLKDTGRAEYVSVTDDEALEAFQLLSRTEGIIPAL ESSHAVAYAMKLAPELSKDQIIVVNLSGRGDKDVNTVARYLLGVELDLEHHHHHH
pUR28a_Anc1nc-TrpA	Anc1nc-TrpA	MHHHHHHLEMNRIAEAFAKLKAEGEKALIPYITAGDPDLETTLELVRALVEAGADIIELGIPFSDPLADGPT IQRASQRALASGTTLAKVLEMVRELREKNTDVPIVFMTYYNPIYSYGLERFVKECAEAGVDGLIVPDLPP EEAADLAAAARKHGLDLIFLLAPTSTDERIKLIAKHASGFVYCVSVTGVTGARSELAADLAELVSRIRKHT DLPIAVGFGISTPEQAAEVAQVADGVVVGSAIVKRIEENQDEEDLVEEVAAFVRELREAVK
pET21a_Anc1nc-TrpB	Anc1nc-TrpB	MNVSTNVMATTTKSKAALPDARGRFGKYGGRYVPETLMPALEELEEAYERAKRDPEFQAELEYYLKDY VGRPTPLYFAENLTEHLGGAKIYLKREDLNHTGAHKINNALGQALLAKRMGKKRVIAETGAGQHGVATA TVAAMFGLECVVYMGAEDIERQALNVFRMKLLGAEVRPVTSGSRTLKDAINEAMRDWVTNVEDTYYIIG SVVGPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDAIVACVGGGSNAMGIFHPFIDDESVRLIGVEAA GEGIETGKHAASLSAGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYVSV TDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELDLE HHHHHH
pUR28a_Anc2nc-TrpA	Anc2nc-TrpA	MHHHHHHHLEMNRIAEAFAKLKAEGKKALIPYITAGDPDLETTLELVRALVEAGADIIELGVPFSDPLADGP VIQRASQRALASGTTLRKVLEMVRELREKNTEVPIVLMTYYNPIYSYGLERFVKEAAEAGVDGLIVPDLP PEEAADLAAAAREHGLDLIFLLAPTSTDERIKLIAKHASGFVYYVSVTGVTGARSELAADLAEKVARIRKH TDLPIAVGFGISTPEQAAEVAQVADGVVVGSAIVKRIEENQDEEDLVAEVAAFVRELREAVK
pET21a_Anc2nc-TrpB	Anc2nc-TrpB	MNVSTNVMATTPKSKAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAAMFGLECVVYMGAEDIERQALNVFRMKLLGAEVRPVTSGSRTLKDAINEAMRDWVTNVEDTYYIIGS VVGPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVEAG GHGIETGKHAASLSGGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYVSV TDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELDLE HHHHHH
pUR28a_Anc3nc-TrpA	Anc3nc-TrpA	MHHHHHHHLEMNRIAAAFAKLKAEGKKALIPYITAGDPDLETTLELMHALVEAGADIIELGVPFSDPMADG PVIQRASQRALASGTTLRDVLEMVAEFRETDTETPIVLMGYANPIYSMGYERFAKAAAEAGVDGLIVVDL PPEEAAELAAALREHGIDLIFLLAPTSPDERIKLIAEHASGFVYYVSVTGVTGARSADAADVAAKVARIRK HTDLPIAVGFGIKTPEQAAEVAKVADGVVVGSAIVNEIEEAQDEENLTAAVAALVRELRAAVK
pET21a_Anc3nc-TrpB	Anc3nc-TrpB	MNASTNVSATTPKSYAALPDARGHFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKHY VGRPSPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATA TVAARFGLECVVYMGAEDVERQALNVFRMKLLGAEVRPVTSGSRTLKDAMNEAMRDWVTNVDDTFYII GTVAGPHPYPMMVRDFQSVIGREARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVE AGGHGIETGKHAASLSGGRPGVLHGNRTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSWLKDIGRAEYV SVTDDEALAAFQLLSRLEGIIPALESSHALAYAAKLAPTMSKDQIIVVNLSGRGDKDVNTVARYLLGVDLD LEHHHHHH
pUR28a_Anc4nc-TrpA	Anc4nc-TrpA	MHHHHHHHLEMNRIAACFAKLKAEGKKALIPYITAGDPDPDTTVDLMHALVEAGADIIELGVPFSDPMADG PVIQRASERALAHNTSLRDVLEMVAEFRETDTETPVVLMGYANPIEAMGYERFAKAAAEAGVDGVLTVD LPPEEAAELAAALKEHGIDPIFLLAPTTPEQRVKLIAEHASGYVYYVSVKGVTGAGNLDVDDVAAKLARIR QHTDLPIGVGFGIKDGETAASVAEVADGVVVGSALVNKIEEAQDEDNLKAAVAALVAELRAAVD
pET21a_Anc4nc-TrpB	Anc4nc-TrpB	MNASTNVSAISLKSYAQLPDARGHFGPYGGRFVAETLMAALDELEEAYERAKRDPAFQAEFDRDLKHY VGRPSPLYFAERWTEHLGGAKIYLKREDLNHTGAHKVNNTIGQALLAKRMGKKRVIAETGAGQHGVAS ATVAARFGLECVVYMGAEDVERQALNVYRMKLLGAEVVPVTSGSRTLKDAMNEAMRDWVTNVDDTFY IIGTVAGPHPYPMMVRDFQSVIGREARQQMLEQEGRLPDALVACVGGGSNAIGLFHPFIEDEGVRMIGV EAGGHGIETGKHAAPLSAGRPGVLHGNRTYLMQDEDGQIIETHSVSAGLDYPGVGPEHSWLKDIGRAE YVAVTDDEALAAFHALTRIEGIMPALESSHALAYAAKLAPTMSKDQIIVVNLSGRGDKDINTVAQYLSGIN LDLEHHHHHH
pMAL-c5T_Anc5nc-TrpA	Anc5nc-TrpA	GSHMNRIDACFAKLKAEGKKALIPYITAGDPDPDVTVDLMHALVEAGADIIELGVPFSDPMADGPVIQLA CERALAHNTSLRDVLEMVAEFRETDTETPIVLMGYANPIEAMGYERFAKAAAEAGVDGVLTVDLPPEEA AELNAALKEHGIDTIFLLAPTTPEQRVKLIVEHASGYVYYVSVKGVTGAGNLDVDDVAAKLARIRQHTDLP IGVGFGIKDGESAASVAEVADGVVVGSALVNKIGELQDEDNIKAAVAALVAEIRSAVD
pET21a_Anc5nc-TrpB	Anc5nc-TrpB	MNASTNVSAIDLKAYAQLPDANGHFGPYGGRFVSETLMAALDDLEEMYERLKRDPAFQAEFDKDLAHY VGRPSPLYFAERWTEKVGGAKIYLKREDLNHTGAHKVNNTIGQALLAKYMGKKRVIAETGAGQHGVAS ATVAARLGLECQVYMGAEDVERQALNVYRMKLLGAEVVPVTSGTRTLKDAMNEAMRDWVTNVDDTFY IIGTVAGPHPYPMLVRDFQSVIGREARQQCLEQEGRLPDALVACVGGGSNAIGLFHPFIEDEGVAMYGV EAGGHGIETGKHAAPLSAGRPGVLHGNRTYLMQDEDGQIIETHSVSAGLDYPGVGPEHSWLKDIGRVE YVAATDDEALAAFHALTRVEGIMPALESSHALAYAAKLAATMSKDQIIVVNLSGRGDKDIHTVAEYLDGIN IFLEHHHHHH
pMAL-c5T_Anc6nc-TrpA	Anc6nc-TrpA	GSHMNRIDACFAKLKAEGKKALIPYITAGDPEPDVTVDLMHALVEAGADIIELGVPFSDPMADGPVIQLA CERALAHNTSLRDVLEMVAEFRETDTETPIVLMGYLNPIEVMGYERFAKAAQEAGVDGVLTVDLPPEES AELNQILKEHGIDTIYLLAPTTPEQRVKLICEHGSGYLYYVSVKGVTGSASLDVDDVANKLERIRKHTDLP LGVGFGIKDGESAASVAKVADGVIVGSALVNKIAELQDDDNIKSEVAALITEIRSAMD
pET21a_Anc6nc-TrpB	Anc6nc-TrpB	MNASTNVSAIDLKALSQLPDANGHFGPYGGRFVSETLMAALDDLEEMYERLKRDPAFQAEFDRDLAHY VGRPSPLYFAERLTEKVGGAKIYLKREDLNHTGAHKVNNTIGQALLAKYTGKPRVIAETGAGQHGVASA TVAARLGLECQVYMGAEDVQRQALNVYRMKLLGAEVVPVTSGTRTLKDAMNEAMRDWVTNVDDTFYII

		GTVAGPHPYPKLVRDFQSVIGREARQQCLEQTGRLPDALVACVGGGSNAIGLFHPFIEDEGVAMYGVE AGGHGVETGQHAAPLSAGKPGVLHGNRTYLMQDENGQIMGTHSVSAGLDYPGVGPEHSYLKDIGRVN YVAATDEEALDAFHALTRVEGIMPALESSHAVAYAMKLAATMDKDQIIVVNLSGRGDKDIHTVAEYLDGI NIFLEHHHHHH
pMAL-c5T_ncTrpA	ncTrpA	GSHMSRISSSFTKLKTEGRKALIPYVTAGDPAPGVTVGLLHAMVEAGADVIELGVPFSDPMADGPVIQLA CERALMHNTRLTDVLEMVAEFRQKDDVTPIVLMGYLNPIEILGYERFAQEAQKAGVDGVLTVDLPPEES QEFNQIMREHDIDTIYLLAPTTEESRVKYVCENGSGYLYYVSVKGVTGSASLDVQSVANKLEVVRKYTD LPLGVGFGIKDAESATSVSKVAEGVIVGSVLVNKIAELVKDQDNIAPEVAAIINDMRTAMDS
pUR28a_ncTrpB	ncTrpB	MHHHHHHLEMPTNIDLKALSQLPDSNGHFGPYGGRFVSETLMAALKQLEETYEKLWRDPAFQEEFDR DLAHYVGRPSPLYFAKHLTEKVGGAKIFLKREDLNHTGAHKVNNTIGQALLAKYTGKPRVIAETGAGQH GVASATVAARLGLECQVYMGAEDVQRQALNVYRMKLLGAEVISVESGTRTLKDAMNEAMRDWVTNVD NTFYIIGTAAGPHPYPKLVRDFQCVIGREARQQCLDQTGRLPDALVACVGGGSNAIGLFHPFIEDADVK MYGVEAGGYGVETGQHAAPLSAGSPGVLHGNRTYLMSDEGGQILGTHSVSAGLDYPGVGPEHAYLKD TGRATYVDATDEEAMDAFRALTRVEGIMPAIESSHAVAYAMKLAKEMDKDQTIVVNLSGRGDKDIHTVA QIDGINI
pET21a_Anc2nc-TrpB_AM10	Anc2nc-TrpB_AM10	MNVSTNVMATTPKSKAALPDARGHFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPSPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATA TVAARFGLECVVYMGAEDVERQALNVFRMKLLGAEVRPVTSGSRTLKDAMNEAMRDWVTNVDDTYYII GTVAGPHPYPMMVRDFQSVIGREARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVE AGGHGIETGKHAASLSGGRPGVLHGSRTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYV SVTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELD LEHHHHHH
pET21a_Anc3nc-TrpB_AM10	Anc3nc-TrpB_AM10	MNASTNVSATTPKSYAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKHY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAAMFGLECVVYMGAEDIERQALNVFRMKLLGAEVRPVTSGSRTLKDAINEAMRDWVTNVEDTFYIIGS VVGPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVEAG GHGIETGKHAASLSGGRPGVLHGNMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSWLKDIGRAEYVSV TDDEALAAFQLLSRLEGIIPALESSHALAYAAKLAPTMSKDQIIVVNLSGRGDKDVNTVARYLLGVDLDLE LEHHHHHH
pET21a_Anc2nc-TrpB_AM8-A	Anc2nc-TrpB_AM8-A	MNVSTNVMATTPKSKAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPSPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATA TVAARFGLECVVYMGAEDVERQALNVFRMKLLGAEVRPVTSGSRTLKDAMNEAMRDWVTNVDDTYYII GTVAGPHPYPMMVRDFQSVIGREARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVE AGGHGIETGKHAASLSGGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYV SVTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELD LEHHHHHH
pET21a_Anc2nc-TrpB_AM8-C	Anc2nc-TrpB_AM8-C	MNVSTNVMATTPKSKAALPDARGHFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPSPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATA TVAARFGLECVVYMGAEDVERQALNVFRMKLLGAEVRPVTSGSRTLKDAMNEAMRDWVTNVDDTYYII GSVVGPHPYPMMVRDFQSVIGREARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVE AGGHGIETGKHAASLSGGRPGVLHGSRTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYV SVTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELD LEHHHHHH
pET21a_Anc2nc-TrpB_AM8-D	Anc2nc-TrpB_AM8-D	MNVSTNVMATTPKSKAALPDARGHFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAARFGLECVVYMGAEDVERQALNVFRMKLLGAEVRPVTSGSRTLKDAMNEAMRDWVTNVDDTYYII GTVAGPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVE AGGHGIETGKHAASLSGGRPGVLHGSRTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYV SVTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELD LEHHHHHH
pET21a_Anc2nc-TrpB_AM6	Anc2nc-TrpB_AM6	MNVSTNVMATTPKSKAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAARFGLECVVYMGAEDVERQALNVFRMKLLGAEVRPVTSGSRTLKDAMNEAMRDWVTNVDDTYYII GTVAGPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVE AGGHGIETGKHAASLSGGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYV SVTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELD LEHHHHHH
pET21a_Anc2nc-TrpB_AM5_R142M	Anc2nc-TrpB_AM5_R142M	MNVSTNVMATTPKSKAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAAMFGLECVVYMGAEDVERQALNVFRMKLLGAEVRPVTSGSRTLKDAMNEAMRDWVTNVDDTYYII GTVAGPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVE AGGHGIETGKHAASLSGGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYV SVTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELD LEHHHHHH
pET21a_Anc2nc-TrpB_AM5_V156I	Anc2nc-TrpB_AM5_V156I	MNVSTNVMATTPKSKAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAARFGLECVVYMGAEDIERQALNVFRMKLLGAEVRPVTSGSRTLKDAMNEAMRDWVTNVDDTYYIIG TVAGPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVEA GGHGIETGKHAASLSGGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYVS VTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELDL EHHHHHH
pET21a_Anc2nc-TrpB_AM5_M187I	Anc2nc-TrpB_AM5_M187I	MNVSTNVMATTPKSKAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAARFGLECVVYMGAEDVERQALNVFRMKLLGAEVRPVTSGSRTLKDAINEAMRDWVTNVDDTYYIIG

		TV**A**GPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVEA GGHGIETGKHAASLSGGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYVS VTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELDL EHHHHHH
pET21a_Anc2nc -TrpB_AM5_ D199E	Anc2nc-TrpB_AM5_ D199E	MNVSTNVMATTPKSKAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAA**R**FGLECVVYMGAED**V**ERQALNVFRMKLLGAEVRPVTSGSRTLKDA**M**NEAMRDWVTNVEDTYYIIG **T**V**A**GPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVEA GGHGIETGKHAASLSGGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYVS VTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELDL EHHHHHH
pET21a_Anc2nc -TrpB_AM4	Anc2nc-TrpB_AM4	MNVSTNVMATTPKSKAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAA**R**FGLECVVYMGAED**V**ERQALNVFRMKLLGAEVRPVTSGSRTLKDAINEAMRDWVTNVEDTYYIIG **T**V**A**GPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVEA GGHGIETGKHAASLSGGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYVS VTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELDL EHHHHHH
pET21a_Anc2nc -TrpB_AM3_ R142M	Anc2nc-TrpB_AM3_ R142M	MNVSTNVMATTPKSKAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAAMFGLECVVYMGAED**V**ERQALNVFRMKLLGAEVRPVTSGSRTLKDAINEAMRDWVTNVEDTYYIIG **T**V**A**GPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVEA GGHGIETGKHAASLSGGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYVS VTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELDL EHHHHHH
pET21a_Anc2nc -TrpB_AM3_ V156I	Anc2nc-TrpB_AM3_ V156I	MNVSTNVMATTPKSKAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAA**R**FGLECVVYMGAEDIERQALNVFRMKLLGAEVRPVTSGSRTLKDAINEAMRDWVTNVEDTYYIIG**T** V**A**GPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVEAG GHGIETGKHAASLSGGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYVSV TDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELDLE HHHHHH
pET21a_Anc2nc -TrpB_AM3_ T207S	Anc2nc-TrpB_AM3_ T207S	MNVSTNVMATTPKSKAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAA**R**FGLECVVYMGAED**V**ERQALNVFRMKLLGAEVRPVTSGSRTLKDAINEAMRDWVTNVEDTYYIIG SV**A**GPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVEA GGHGIETGKHAASLSGGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYVS VTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELDL EHHHHHH
pET21a_Anc2nc -TrpB_AM3_ A209V	Anc2nc-TrpB_AM3_ A209V	MNVSTNVMATTPKSKAALPDARGRFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKEY VGRPTPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATAT VAA**R**FGLECVVYMGAED**V**ERQALNVFRMKLLGAEVRPVTSGSRTLKDAINEAMRDWVTNVEDTYYIIG **T**VVGPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVEA GGHGIETGKHAASLSGGRPGVLHGSMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSYLKDTGRAEYVS VTDDEALEAFQLLSRTEGIIPALESSHAIAYAVKLAPEMSKDQIIVVNLSGRGDKDVNTVARYLLGVELDL EHHHHHH
pET21a_LBCA-TrpB_AM4	LBCA-TrpB_AM4	MIGRFGKYGGQYVPETLMPALEELEEAYERAKNDPEFQAELEYYLRDYVGRPTPLYFAENLTKDLGGA KIYLKREDLNHTGAHKINNALGQALLAKRMGKKRVIAETGAGQHGVATATVAA**R**FGLECVVYMGAED**V**E RQALNVFRMKLLGAKVRPVTSGSRTLKDAINEAMRDWVTNVEDTFYIIG**T**V**A**GPHPYPMMVRDFQSVI GEEARQQILEKEGRLPDAIVACVGGGSNAMGIFHPFIDDESVRLIGVEAAGKGIETGKHAATLSAGRPGV LHGAMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHAYLKDTGRAEYVSVTDDEALEAFQLLSRTEGIIPAL ESSHAVAYAMKLAPELSKDQIIVVNLSGRGDKDVNTVARYLLGVELDLEHHHHHH
pET21a_Anc3nc TrpB_AM4	Anc3nc-TrpB_AM4	MNASTNVSATTPKSYAALPDARGHFGPYGGRYVPETLMPALEELEEAYERAKRDPAFQAELDYYLKHY VGRPSPLYFAERLTEHLGGAKIYLKREDLNHTGAHKINNAIGQALLAKRMGKKRVIAETGAGQHGVATA TVAA**M**FGLECVVYMGAED**I**ERQALNVFRMKLLGAEVRPVTSGSRTLKDAMNEAMRDWVTNVDDTFYII G**S**V**V**GPHPYPMMVRDFQSVIGREARQQILEKEGRLPDALVACVGGGSNAMGLFHPFIDDEGVRMIGVE AGGHGIETGKHAASLSGGRPGVLHGNRTYLLQDEDGQIIEAHSISAGLDYPGVGPEHSWLKDIGRAEYV SVTDDEALAAFQLLSRLEGIIPALESSHALAYAAKLAPTMSKDQIIVVNLSGRGDKDVNTVARYLLGVDLD LEHHHHHH

Sequence colored in red: His_6-tag for IMAC purification. In case of pMAL-c5T constructs, the His_6-MBP-tag sequence contains a thrombin cleavage site (LVPR|GS) which was exploited to remove the His_6-MBP-tag from the desired amino acid sequence. The residual GSH sequence is colored in green. Sequence colored in blue: Initial methionine (M1). Sequence in **bold** and underlined: Residues that were mutated compared to the respective wild-type sequence.

Table 8: Strategy for the identification of positions relevant for the allosteric effect of TrpA on TrpB

Step	Protein	Mutated positions	Allosteric effect on TrpB	Conclusion
#1 "wildtype" Anc2nc-TrpB	Anc2nc-TrpB	none	−	none
#2 Anc2nc-TrpB with 10 unique positions mutated towards Anc3nc identity	Anc2nc-TrpB_AM10	R24H M300R M142R I156V I187M E199D S207T V209A T73S E227R	+	10 mutated positions are sufficient to inverse the allosteric effect on TrpB
#3 Anc2nc-TrpB_AM10 Cluster A, C and D seperatly backmutated to Anc2nc identity	Anc2nc-TrpB_AM8-A	M142R I156V I187M E199D S207T V209A T73S E227R	+	R24H M300R are not relevant for inversion of the allosteric effect on TrpB
	Anc2nc-TrpB_AM8-C	R24H M300R M142R I156V I187M E199D T73S E227R	−	S207T V209A are relevant for inversion of the allosteric effect on TrpB
	Anc2nc-TrpB_AM8-D	R24H M300R M142R I156V I187M E199D S207T V209A	+	T73S E227R are not relevant for inversion of the allosteric effect on TrpB
#4 Combinational backmutation of the irrelevant positions from Cluster A and D	Anc2nc-TrpB_AM6	M142R I156V I187M E199D S207T V209A	+	the 6 mutated positions are sufficient to inverse the allosteric effect on TrpB

Step	Protein	Mutated positions	Allosteric effect on TrpB	Conclusion
#5 Anc2nc-TrpB_AM6 single backmutation of the positions in Cluster B	Anc2nc-TrpB_AM5_R142M	I156V I187M E199D S207T V209A	−	M142R is relevant for inversion of the allosteric effect on TrpB
	Anc2nc-TrpB_AM5_V156I	M142R I187M E199D S207T V209A	−	I156V is relevant for inversion of the allosteric effect on TrpB
	Anc2nc-TrpB_AM5_M187I	M142R I156V E199D S207T V209A	+	I187M is not relevant for inversion of the allosteric effect on TrpB
	Anc2nc-TrpB_AM5_D199E	M142R I156V I187M S207T V209A	+	E199D is not relevant for inversion of the allosteric effect on TrpB
#6 Combinational backmutation of the irrelevant positions identified in #5	Anc2nc-TrpB_AM4	M142R I156V S207T V209A	+	the 4 mutated positions are sufficient to inverse the allosteric effect on TrpB
#7 Anc2nc-TrpB_AM4 single backmutation of the last 4 positions	Anc2nc-TrpB_AM3_R142M	I156V S207T V209A	−	M142R is relevant for inversion of the allosteric effect on TrpB
	Anc2nc-TrpB_AM3_V156I	M142R S207T V209A	−	I156V is relevant for inversion of the allosteric effect on TrpB
	Anc2nc-TrpB_AM3_T207S	M142R I156V V209A	−	S207T is relevant for inversion of the allosteric effect on TrpB
	Anc2nc-TrpB_AM3_A209V	M142R I156V S207T	−	V209A is relevant for inversion of the allosteric effect on TrpB

The color code is as in **Figure 21A**: Mutations colored in light green are located in Cluster A. Mutations colored in light blue are located in Cluster B. Mutations colored in brown are located in Cluster C. Mutations colored in pink are located in Cluster D. The activation or deactivation column is based on experimental results shown in **Figure 21B**.

Table 9: Tryptophan formation by TrpB-subunits in absence and presence of the respective TrpA-subunits followed by an HPLC-based assay

Protein	Initial rate of tryptophan formation / nM/s
LBCA-TrpB	625,0 ± 106,2
LBCA-TrpA+ LBCA-TrpB	478,2 ± 138,4
Anc1nc-TrpB	595,0 ± 36,7
Anc1nc-TrpA+ Anc1nc-TrpB	364,8 ± 66,6
Anc2nc-TrpB	872,3 ± 169,8
Anc2nc-TrpA+ Anc2nc-TrpB	102,2 ± 38,2
Anc3nc-TrpB	62,2 ± 37,4
Anc3nc-TrpA+ Anc3nc-TrpB	577,4 ± 88,6
Anc4nc-TrpB	374,2 ± 217,0
Anc4nc-TrpA+ Anc4nc-TrpB	860,4 ± 144,9
Anc5nc-TrpB	445,6 ± 75,2
Anc5nc-TrpA+ Anc5nc-TrpB	852,7 ± 87,3
Anc6nc-TrpB	236,1 ± 67,2
Anc6nc-TrpA+ Anc6nc-TrpBB	552,1 ± 95,3
ncTrpB	41,0 ± 33,4
ncTrpA+ ncTrpB	352,5 ± 43,1

Experimental conditions: 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP, 45 mM L-serine, 300 µM indole, 0.5 µM TrpB (monomer) and 1 µM TrpA (monomer) at 30 °C. The mean and the standard deviations of two independent measurements are given.

Table 10: Tryptophan formation of TrpB-subunits in absence and presence of the respective TrpA-subunits followed by absorbance spectroscopy at 290 nm

Protein	Initial rate of tryptophan formation / nM/s
Anc2nc-TrpB	150,3 ± 0,4
Anc2nc-TrpB +Anc2nc-TrpA	1,5 ± 0,0
Anc2nc-TrpB_AM10	0,6 ± 2,0
Anc2nc-TrpB_AM10 + Anc2nc-TrpA	122,4 ± 4,4
Anc3nc-TrpB	14,2 ± 0,7
Anc3nc-TrpB + Anc3nc-TrpA	484,9 ± 24,6
Anc3nc-TrpB_AM10	604,4 ± 31,3
Anc3nc-TrpB_AM10 + Anc3nc-TrpA	9,6 ± 1,2

Experimental conditions: 100 mM potassium phosphate pH 7.5, 180 mM KCl, 40 µM PLP, 50 mM L-serine, 200 µM indole, 0.5 µM TrpB (monomer) and 10 µM TrpA (monomer) at 30 °C. The mean and the standard deviations of two independent measurements are given.

Table 11: Tryptophan formation of TrpB-subunits in absence and presence of the respective TrpA-subunits followed by absorbance spectroscopy at 290 nm

Protein	Initial rate of tryptophan formation / nM/s
Anc3nc-TrpB	7,1 ± 0,2
Anc3nc-TrpB + Anc3nc-TrpA	417,0 ± 7,7
Anc3nc-TrpB_AM4	118,9 ± 5,6
Anc3nc-TrpB_AM4 + Anc3nc-TrpA	25,6 ± 4,4
LBCA-TrpB	293,6 ± 16,1
LBCA-TrpB + LBCA-TrpA	39,2 ± 3,3
LBCA-TrpB_AM4	35,9 ± 2,6
LBCA-TrpB_AM4 + LBCA-TrpA	168,6 ± 1,6

Experimental conditions: 100 mM potassium phosphate pH 7.5, 180 mM KCl, 40 µM PLP, 50 mM L-serine, 200 µM indole, 0.5 µM TrpB (monomer) and 10 µM TrpA (monomer) at 30 °C. The mean and the standard deviations of two independent measurements are given.

Table 12: Tryptophan formation of TrpB-subunits in absence and presence of the respective TrpA-subunits followed by absorbance spectroscopy at 290 nm

Protein	Initial velocity of tryptophan formation / nM/s
Anc2nc-TrpB	150,3 ± 0,4
Anc2nc-TrpB + Anc2nc-TrpA	1,5 ± 0,0
Anc2nc-TrpB_AM10	0,6 ± 2,0
Anc2nc-TrpB_AM10 + Anc2nc-TrpA	122,4 ± 4,4
Anc2nc-TrpB_AM8-A	6,3 ± 1,4
Anc2nc-TrpB_AM8-A + Anc2nc-TrpA	280,1 ± 12,0
Anc2nc-TrpB_AM8-C	5,2 ± 0,5
Anc2nc-TrpB_AM8-C + Anc2nc-TrpA	4,2 ± 0,1
Anc2nc-TrpB_AM8-D	18,0 ± 2,8
Anc2nc-TrpB_AM8-D + Anc2nc-TrpA	303,2 ± 19,8
Anc2nc-TrpB_AM6	118,5 ± 0,2
Anc2nc-TrpB_AM6 + Anc2nc-TrpA	341,1 ± 37,8
Anc2nc-TrpB_AM5_R142M	364,5 ± 18,6
Anc2nc-TrpB_AM5_R142M + Anc2nc-TrpA	201,0 ± 7,9
Anc2nc-TrpB_AM5_V156I	291,5 ± 3,6
Anc2nc-TrpB_AM5_V156I + Anc2nc-TrpA	178,1 ± 9,4
Anc2nc-TrpB_AM5_M187I	105,1 ± 4,2
Anc2nc-TrpB_AM5_M187I + Anc2nc-TrpA	218,3 ± 1,9
Anc2nc-TrpB_AM5_D199E	51,0 ± 2,3
Anc2nc-TrpB_AM5_D199E + Anc2nc-TrpA	186,8 ± 3,8
Anc2nc-TrpB_AM4	101,7 ± 1,2
Anc2nc-TrpB_AM4 + Anc2nc-TrpA	271,3 ± 2,3
Anc2nc-TrpB_AM3_R142M	317,2 ± 31,6
Anc2nc-TrpB_AM3_R142M + Anc2nc-TrpA	167,6 ± 8,5
Anc2nc-TrpB_AM3_V156I	185,6 ± 6,1
Anc2nc-TrpB_AM3_V156I + Anc2nc-TrpA	131,8 ± 7,0
Anc2nc-TrpB_AM3_T207S	177,5 ± 0,2
Anc2nc-TrpB_AM3_T207S + Anc2nc-TrpA	62,1 ± 0,1
Anc2nc-TrpB_AM3_A209V	91,2 ± 0,8
Anc2nc-TrpB_AM3_A209V + Anc2nc-TrpA	34,8 ± 1,1

Experimental conditions: 100 mM potassium phosphate pH 7.5, 180 mM KCl, 40 µM PLP, 50 mM L-serine, 200 µM indole, 0.5 µM TrpB (monomer) and 10 µM TrpA (monomer) at 30 °C. The mean and the standard deviations of two independent measurements are given.

4 Comprehensive Summary and Outlook

4.1 Comprehensive Summary

In this work, both subunits of the TS were studied with respect to the nature of the mutual allosteric communication that takes place between TrpA and TrpB during catalysis. In both manuscripts, residues that play a crucial role in allosteric activation of the respective subunit were identified by two different approaches for sequence comparisons. Subsequently the significance of the identified residues for allostery was confirmed by mutational studies.

In manuscript A, a horizontal approach was employed to perform evolution guided rational design experiments that provided insight into allosteric activation of the TrpA subunit within the TS complex. In addition, it was shown that the transition from an enzyme of primary metabolism to an enzyme of secondary metabolism can be accomplished by the mutation of just one amino acid residue. In general, the canonical TrpA subunit is activated by its TrpB interaction partner (Dunn 2012). A homologue of TrpA acting in *Zea mays* secondary metabolism, named zmBX1, catalyzes the same IGP lyase reaction as TrpA, but is not dependent on allosteric activation (Kriechbaumer et al. 2008). Two loop regions, Loop2 and Loop6, which contribute to the active site of zmTrpA and zmBX1 were compared regarding their amino acid compositions (Axe et al. 2015). This horizontal sequence comparison revealed two positions within Loop6 that differ between zmTrpA and zmBX1. At both positions in zmBX1 proline residues are present, while zmTrpA contains a threonine and a glycine residue, respectively. In order to assess the effect of the identified proline positions, multiple variants of zmTrpA were created and assayed for their IGP lyase activity. It turned out that if one of the two positions carries a proline, an increase of the turnover number (k_{cat}-value) by at least an order of magnitude is observed. Further enhancement of the k_{cat}-value could be obtained by the combined incorporation of proline at both identified positions, and the maximum activation was achieved by replacing the complete Loop6 of zmTrpA with the corresponding loop from zmBX1. That means the TrpA variants are decoupled from the allosteric activation by TrpB. At the same time, all variants displayed a dramatic increase in the K_m^{IGP}-value. The results of manuscript A show that single mutations can alter the allosteric regulation mechanism of an enzyme in primary metabolism. This suggests an evolutionary trajectory, which allowed a copy of TrpA to evolve into zmBX1 and act in

secondary metabolism. These findings provide another example of a primary metabolism enzyme that gave rise to a secondary metabolism enzyme by a very low number of mutations (Vining 1992).

In manuscript B, a vertical approach was pursued to identify the molecular cause for the unusual deactivation of a TrpB subunit by a TrpA subunit. While in well-known extant TS complexes the TrpB subunit is activated by the TrpA subunit, a recent study reported that LBCA-TrpB from the last common ancestor of all bacteria that is deactivated by LBCA-TrpA (Busch et al. 2016). Both LBCA-TrpA and LBCA-TrpB were “resurrected” by an *in silico* method termed ancestral sequence reconstruction. The vertical approach, in combination with site-directed mutagenesis, enabled the identification of four *a-ess* residues in TrpB that are sufficient for the inversion of the allosteric activation. Size-exclusion chromatography showed that these four positions did not alter TS complex formation. Moreover, steady-state enzyme kinetics revealed that the four residues influence only the k_{cat}-value of the TrpB subunit in presence of TrpA, but not binding of the substrates L-serine and indole. The COMM domain within TrpB is crucial for the allosteric communication (Schneider et al. 1998). Molecular dynamics simulations were performed in order to identify communities of adjacent residues that move in a concerted manner. Interestingly, these *in silico* studies indicated that there are two separated communities within the COMM domain of the deactivated variant of TrpB, while for the activated variant only one concerted community within the COMM domain was identified. These findings lead to the hypothesis that the disintegration of the communities within the COMM domain of TrpB is responsible for the deactivation of TrpB by TrpA. Such conclusion is in agreement with a previously postulated hypothesis according to which activation of TrpB by TrpA is coupled to a coordinated movement of the COMM domain (Kulik et al. 2002).

4.2 Outlook

Within manuscript A, it remained unanswered how the incorporation of proline residues in Loop6 of zmTrpA affects its biochemical properties and catalysis. Limited trypsin proteolysis was performed to obtain insights into Loop6 flexibility (Miles and Higgins 1978, Leopoldseder 2005) but no differences between activated zmTrpA variants (zmTrpA_T186P, zmTrpA_G192P, zmTrpA_T186P_G192P, zmTrpA_L6zmBX1) and wild-type zmTrpA were observed. Subsequently, molecular dynamics simulations were conducted in cooperation with Leonhard Heizinger (group of Prof. Dr. Rainer Merkl, University of Regensburg). Protein dynamics on a 30 ns timescale, however, did not reveal any differences between a zmTrpA homology model and the homology models of its activated variants. Other possible approaches to study loop dynamics that have recently been followed by our group include FRET measurements (in cooperation with Prof. Dina Grohmann, University of Regensburg) or PRE-NMR studies (in cooperation with Prof. Remco Sprangers, University of Regensburg) (Ruisinger 2019).

A way to uncover the effect of activating mutations in Loop6 for catalysis would be the kinetic analysis of the separate mechanistic steps within the α-reaction. This could be performed by applying chemical quench-flow methods (Anderson et al. 1991). Probably the most urgent question left open is the molecular basis for the increased K_m^{IGP}-value of the activated zmTrpA variants, which indicates that the mutations in Loop6 hamper the binding of IGP to the active site. The drastically decreased K_m^{IGP}-value in presence of zmTrpB1 demonstrates that this effect can be reversed by an interaction partner, which might eliminate an unfavorable Loop6 conformation induced by the introduced mutations. Presumably, only X-ray structures of wild-type zmTrpA and its activated Loop6 variants in presence of an IGP analogue like indole propanol phosphate (IPP) (Ehrmann et al. 2010) could bring light into this matter. However, crystallization attempts have not been successful so far.

An interesting question is whether the activating effect of Loop6 mutations observed for zmTrpA can be transferred to other TrpA enzymes. The TrpA enzymes from several species were analyzed for the compositions of Loop6. It turned out that – in contrast to zmTrpA – most TrpA enzymes contain one proline residue in Loop6 close to the proline 188 of zmBX1. A second proline residue at the corresponding position of proline 182 in zmBX1 is missing in every other inspected TrpA enzyme. Various TrpA variants (stTrpA, ecTrpA, sfTrpA, LBCA-TrpA, ttTrpA, pfTrpA and tmTrpA) with a second proline corresponding to proline 182 of zmBX1 were generated, but none of these showed an enhanced IGP lyase activity.

In the course of manuscript B, four novel *a-ess* residues that are essential for the allosteric activation/deactivation of TrpB by TrpA subunit were identified and studied. Steady-state enzyme kinetics showed that the cause for the altered normalized allosteric effect induced by the four residue differences between Anc2nc-TrpB and Anc2nc-TrpB_AM4 exclusively lead to altered turnover numbers (k_{cat}-values). In contrast, the $K_m^{L\text{-}ser}$ - and K_m^{indole} – values of TrpB remain practically constant. Based on these results, a deeper insight in the molecular role of the four *a-ess* positions could be obtained by stopped-flow methods that allow one to determine the microscopic rate constants for the separate steps of catalysis. Such measurements were previously conducted for TS from *S. typhimurium*, where the rate constants of the main steps within the catalytic cycle of TrpB were determined by exploiting the change in absorbance of the PLP cofactor (Anderson et al. 1991). Preliminary stopped-flow experiments with Anc2nc-TrpB and Anc2nc-TrpB_AM4 in presence and absence of Anc2nc-TrpA were performed in collaboration with Dr. Sandra Schlee, but failed due to a missing spectral signal in the presence of TrpA.

In addition to such functional studies, the elucidation of dynamical and structural properties of Anc2nc-TrpB and Anc2nc-TrpB_AM4 would be of high interest. Community analyses indicated that the concerted dynamics of the intact COMM domain as realized in Anc2nc-TrpB_AM4 was interrupted in Anc2nc-TrpB. An experimental confirmation of this hypothesis is of urgent need. A NMR-based method that employs ^{13}C labeled methyl groups as probes has been reported for the detection of changes in dynamics of protein complexes even larger than 1000 kDa (Wiesner and Sprangers 2015). This method could provide first evidence whether the *a-ess* residues indeed determine the coordination of COMM domain movement. Another approach for elucidation of changes in dynamics of the COMM domain by NMR could be the employment of a ^{13}C labeled pyridoxal 5'-phosphate (PLP) cofactor (O'Leary and Payne 1976). Since the two of the *a-ess* positions are located at the active site and the allosteric communication affects mainly the k_{cat}-value, the PLP cofactor could prove as a valuable probe. Complementary to NMR studies, solving crystal structures of Anc2nc-TrpB and Anc2nc-TrpB_AM4 both in presence and absence of Anc2nc-TrpA could provide additional evidence about the conformations adapted by the COMM domain.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic acids research* 25(17): 3389-3402.

Anderson, K. S., Miles, E. W. and Johnson, K. A. (1991). Serine Modulates Substrate Channeling in Tryptophan Synthase. A Novel Intersubunit Triggering Mechanism. *The Journal of biological chemistry* 266(13): 8020-8033.

Axe, J. M., O'Rourke, K. F., Kerstetter, N. E., Yezdimer, E. M., Chan, Y. M., Chasin, A. and Boehr, D. D. (2015). Severing of a Hydrogen Bond Disrupts Amino Acid Networks in the Catalytically Active State of the Alpha Subunit of Tryptophan Synthase. *Protein science* 24(4): 484-494.

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 : a European journal of chemical biology* 9(7): 1024-1028.

Berendsen, H. J. C., Vanderspoel, D. and Vandrunen, R. (1995). Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation. *Computer Physics Communications* 91(1-3): 43-56.

Blei, F., Baldeweg, F., Fricke, J. and Hoffmeister, D. (2018). Biocatalytic Production of Psilocybin and Derivatives in Tryptophan Synthase-Enhanced Reactions. *Chemistry* 24(40): 10028-10031.

Boville, C. E., Romney, D. K., Almhjell, P. J., Sieben, M. and Arnold, F. H. (2018). Improved Synthesis of 4-Cyanotryptophan and Other Tryptophan Analogs in Aqueous Solvent Using Variants of TrpB from *Thermotoga maritima*. *The Journal of organic chemistry* 83(14): 7447-7452.

Brzovic, P. S., Sawa, Y., Hyde, C. C., Miles, E. W. and Dunn, M. F. (1992). Evidence That Mutations in a Loop Region of the α-Subunit Inhibit the Transition from an Open to a Closed Conformation in the Tryptophan Synthase Bienzyme Complex. *The Journal of biological chemistry* 267(18): 13028-13038.

Buller, A. R., Brinkmann-Chen, S., Romney, D. K., Herger, M., Murciano-Calles, J. and Arnold, F. H. (2015). Directed Evolution of the Tryptophan Synthase β-Subunit for Stand-Alone Function Recapitulates Allosteric Activation. *PNAS* 112 (47): 14599-14604.

Buller, A. R., van Roye, P., Cahn, J. K. B., Scheele, R. A., Herger, M. and Arnold, F. H. (2018). Directed Evolution Mimics Allosteric Activation by Stepwise Tuning of the Conformational Ensemble. *Journal of the American Chemical Society* 140(23): 7256-7266.

Busch, F., Rajendran, C., Heyn, K., Schlee, S., Merkl, R. and Sterner, R. (2016). Ancestral Tryptophan Synthase Reveals Functional Sophistication of Primordial Enzyme Complexes. *Cell chemical biology* 23(6): 709-715.

Busch, F. M. (2015). Tryptophan Synthase Complexes: Evolution, Substrate Specificity, and Quaternary Structure. PhD Thesis. University of Regensburg.

Canutescu, A. A., Shelenkov, A. A. and Dunbrack, R. L. (2003). A Graph-Theory Algorithm for Rapid Protein Side-Chain Prediction. *Protein Science* 12(9): 2001-2014.

Crawford, I. P. (1975). Gene Rearrangements in the Evolution of the Tryptophan Synthetic Pathway. *Bacteriol Rev* 39(2): 87-120.

Creighton, T. E. (1970). A Steady-State Kinetic Investigation of the Reaction Mechanism of the Tryptophan Synthetase of *Escherichia coli*. *European Journal of Biochemistry* 13(1): 1-10.

Desvergne, B., Michalik, L. and Wahli, W. (2006). Transcriptional Regulation of Metabolism. *Physiol Rev* 86(2): 465-514.

Dever, T. E., Kinzy, T. G. and Pavitt, G. D. (2016). Mechanism and Regulation of Protein Synthesis in *Saccharomyces cerevisiae*. *Genetics* 203(1): 65-107.

Dierkers, A. T., Niks, D., Schlichting, I. and Dunn, M. F. (2009). Tryptophan Synthase: Structure and Function of the Monovalent Cation Site. *Biochemistry* 48(46): 10997-11010.

Dunn, M. F. (2012). Allosteric Regulation of Substrate Channeling and Catalysis in the Tryptophan Synthase Bienzyme Complex. *Archives of Biochemistry and Biophysics* 519(2): 154-166.

Dunn, M. F., Niks, D., Ngo, H., Barends, T. R. and Schlichting, I. (2008). Tryptophan Synthase: The Workings of a Channeling Nanomachine. *Trends in Biochemical Sciences* 33(6): 254-264.

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 αββ Subunit Stoichiometry. *Biochemistry* 49(51): 10842-10853.

Faeder, E. J. and Hammes, G. G. (1970). Kinetic Studies of Tryptophan Synthetase - Interaction of Substrates with B Subunit. *Biochemistry* 9(21): 4043-4049.

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(25): 7807-7818.

Ferrari, D., Yang, L. H., Miles, E. W. and Dunn, M. F. (2001). βD305A Mutant of Tryptophan Synthase Shows Strongly Perturbed Allosteric Regulation and Substrate Specificity. *Biochemistry* 40(25): 7421-7432.

Firn, R. D. and Jones, C. G. (2000). The Evolution of Secondary Metabolism - A Unifying Model. *Mol Microbiol* 37(5): 989-994.

Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grun, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R. B., Briggs, S. P., Simcox, K. and Gierl, A. (1997). Analysis of a Chemical Plant Defense Mechanism in Grasses. *Science* 277(5326): 696-699.

Frey, M., Schullehner, K., Dick, R., Fiesselmann, A. and Gierl, A. (2009). Benzoxazinoid Biosynthesis, a Model for Evolution of Secondary Metabolic Pathways in Plants. *Phytochemistry* 70(15-16): 1645-1651.

Frey, M., Stettner, C., Pare, P. W., Schmelz, E. A., Tumlinson, J. H. and Gierl, A. (2000). An Herbivore Elicitor Activates the Gene for Indole Emission in Maize. *PNAS* 97(26): 14801-14806.

Gasteiger E., H. C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. The Proteomics Handbook. J. M. Walker, Humana Press**:** 571-607.

Geoghegan, K. F., Dixon, H. B., Rosner, P. J., Hoth, L. R., Lanzetti, A. J., Borzilleri, K. A., Marr, E. S., Pezzullo, L. H., Martin, L. B., LeMotte, P. K., McColl, A. S., Kamath, A. V. and Stroh, J. G. (1999). Spontaneous α-*N*-6-Phosphogluconoylation of a "His Tag" in *Escherichia coli*: The Cause of Extra Mass of 258 or 178 Da in Fusion Proteins. *Analytical biochemistry* 267(1): 169-184.

Girvan, M. and Newman, M. E. (2002). Community Structure in Social and Biological Networks. *PNAS* 99(12): 7821-7826.

Herger, M., van Roye, P., Romney, D. K., Brinkmann-Chen, S., Buller, A. R. and Arnold, F. H. (2016). Synthesis of β-Branched Tryptophan Analogues Using an Engineered Subunit of Tryptophan Synthase. *Journal of the American Chemical Society* 138(27): 8388-8391.

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. *The Journal of biological chemistry* 277(10): 8194-8201.

Hochberg, G. K. A. and Thornton, J. W. (2017). Reconstructing Ancient Proteins to Understand the Causes of Structure and Function. *Annual review of biophysics* 46: 247-269.

Holinski, A., Heyn, K., Merkl, R. and Sterner, R. (2017). Combining Ancestral Sequence Reconstruction with Protein Design to Identify an Interface Hotspot in a Key Metabolic Enzyme Complex. *Proteins* 85(2): 312-321.

Hooft, R. W., Vriend, G., Sander, C. and Abola, E. E. (1996). Errors in Protein Structures. *Nature* 381(6580): 272.

Humphrey, W., Dalke, A. and Schulten, K. (1996). VMD: Visual Molecular Dynamics. *J Mol Graph* 14(1): 33-38, 27-38.

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*. *The Journal of biological chemistry* 263(33): 17857-17871.

Kilpatrick, E. L., Liao, W. L., Camara, J. E., Turko, I. V. and Bunk, D. M. (2012). Expression and Characterization of ^{15}N-Labeled Human C-Reactive Protein in

Escherichia coli and *Pichia pastoris* for Use in Isotope-Dilution Mass Spectrometry. *Protein expression and purification* 85(1): 94-99.

Kriechbaumer, V., Weigang, L., Fiesselmann, A., Letzel, T., Frey, M., Gierl, A. and Glawischnig, E. (2008). Characterisation of the Tryptophan Synthase Alpha Subunit in Maize. *BMC plant biology* 8: 44.

Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D. and Karplus, K. (2009). Improving Physical Realism, Stereochemistry, and Side-Chain Accuracy in Homology Modeling: Four Approaches that Performed Well in CASP8. *Proteins* 77 Suppl 9: 114-122.

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 Alpha Subunit of Tryptophan Synthase from *Salmonella typhimurium* and BX1 from Maize, Two Evolutionarily Related Enzymes. *Journal of molecular biology* 352(3): 608-620.

Kulik, V., Weyand, M., Seidel, R., Niks, D., Arac, D., Dunn, M. F. and Schlichting, I. (2002). On the Role of αThr183 in the Allosteric Regulation and Catalytic Mechanism of Tryptophan Synthase. *Journal of molecular biology* 324(4): 677-690.

Lane, A. N. and Fan, T. W. (2015). Regulation of Mammalian Nucleotide Metabolism and Biosynthesis. *Nucleic acids research* 43(4): 2466-2485.

Leopoldseder, S. (2005). Neuartige Tryptophan Synthasen aus Hyperthermophilen: Charakterisierung der Enzyme aus *Sulfolobus solfataricus*. PhD Thesis. University of Regensburg.

Liang, Z., Verkhivker, G. M. and Hu, G. (2019). Integration of Network Models and Evolutionary Analysis Into High-Throughput Modeling of Protein Dynamics and Allosteric Regulation: Theory, Tools and Applications. *Briefings in Bioinformatics*.

Lisi, G. P. and Loria, J. P. (2016). Solution NMR Spectroscopy for the Study of Enzyme Allostery. *Chemical reviews* 116(11): 6323-6369.

Lisi, G. P. and Loria, J. P. (2017). Allostery in Enzyme Catalysis. *Current opinion in structural biology* 47: 123-130.

List, F., Vega, M. C., Razeto, A., Hager, M. C., Sterner, R. and Wilmanns, M. (2012). Catalysis Uncoupling in a Glutamine Amidotransferase Bienzyme by Unblocking the Glutaminase Active Site. *Chemistry & Biology* 19(12): 1589-1599.

Lu, S., Shen, Q. and Zhang, J. (2019). Allosteric Methods and Their Applications: Facilitating the Discovery of Allosteric Drugs and the Investigation of Allosteric Mechanisms. *Acc Chem Res* 52(2): 492-500.

Lucero, H. A. and Kagan, H. M. (2006). Lysyl Oxidase: An Oxidative Enzyme and Effector of Cell Function. *Cellular and molecular life sciences : CMLS* 63(19-20): 2304-2316.

Marty, M. T., Baldwin, A. J., Marklund, E. G., Hochberg, G. K., Benesch, J. L. and Robinson, C. V. (2015). Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Anal Chem* 87(8): 4370-4376.

Merkl, R. (2007). Modelling the Evolution of the Archeal Tryptophan Synthase. *BMC evolutionary biology* 7: 59.

Merkl, R. and Sterner, R. (2016). Ancestral Protein Reconstruction: Techniques and Applications. *Biological Chemistry* 397(1): 1-21.

Miles, E. W. (1991). Structural Basis for Catalysis by Tryptophan Synthase. *Advances in enzymology and related areas of molecular biology* 64: 93-172.

Miles, E. W. (2001). Tryptophan Synthase: A Multienzyme Complex with an Intramolecular Tunnel. *The Chemical Record* 1(2): 140-151.

Miles, E. W. (2013). The Tryptophan Synthase $\alpha_2\beta_2$ Complex: A Model for Substrate Channeling, Allosteric Communication, and Pyridoxal Phosphate Catalysis. *Journal of Biological Chemistry* 288(14): 10084-10091.

Miles, E. W. and Higgins, W. (1978). An Active $\alpha'_2\beta_2$ Derivative of Tryptophan Synthase Formed by Limited Proteolysis. *The Journal of biological chemistry* 253(17): 6266-6269.

Miles, E. W., Rhee, S. and Davies, D. R. (1999). The Molecular Basis of Substrate Channeling. *The Journal of biological chemistry* 274(18): 12193-12196.

Monod, J., Changeux, J. P. and Jacob, F. (1963). Allosteric Proteins and Cellular Control Systems. *Journal of molecular biology* 6: 306-329.

Murciano-Calles, J., Romney, D. K., Brinkmann-Chen, S., Buller, A. R. and Arnold, F. H. (2016). A Panel of TrpB Biocatalysts Derived from Tryptophan Synthase through the Transfer of Mutations that Mimic Allosteric Activation. *Angewandte Chemie* 55(38): 11577-11581.

Musnier, A., Blanchot, B., Reiter, E. and Crepieux, P. (2010). GPCR Signalling to the Translation Machinery. *Cell Signal* 22(5): 707-716.

Nagata, S., Hyde, C. C. and Miles, E. W. (1989). The α Subunit of Tryptophan Synthase. Evidence That Aspartic Acid 60 is a Catalytic Residue and That the Double Alteration of Residues 175 and 211 in a Second-Site Revertant Restores the Proper Geometry of the Substrate Binding Site. *The Journal of biological chemistry* 264(11): 6288-6296.

Ngo, H., Kimmich, N., Harris, R., Niks, D., Blumenstein, L., Kulik, V., Barends, T. R., Schlichting, I. and Dunn, M. F. (2007). Allosteric Regulation of Substrate Channeling in Tryptophan Synthase: Modulation of the L-Serine Reaction in Stage I of the β-Reaction by α-Site Ligands. *Biochemistry* 46(26): 7740-7753.

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(37): 6396-6411.

Nussinov, R. and Tsai, C. J. (2013). Allostery in Disease and in Drug Discovery. *Cell* 153(2): 293-305.

O'Leary, M. H. and Payne, J. R. (1976). ^{13}C NMR Spectroscopy of Labeled Pyridoxal 5'-Phosphate. Model Studies, D-Serine Dehydratase, and L-Glutamate Decarboxylase. *The Journal of biological chemistry* 251(8): 2248-2254.

O'Rourke, K. F., Axe, J. M., D'Amico, R. N., Sahu, D. and Boehr, D. D. (2018). Millisecond Timescale Motions Connect Amino Acid Interaction Networks in Alpha Tryptophan Synthase. *Frontiers in molecular biosciences* 5: 92.

O'Rourke, K. F., Sahu, D., Bosken, Y. K., D'Amico, R. N., Chang, C. A. and Boehr, D. D. (2019). Coordinated Network Changes across the Catalytic Cycle of Alpha Tryptophan Synthase. *Structure*.

Plach, M. G., Loffler, P., Merkl, R. and Sterner, R. (2015). Conversion of Anthranilate Synthase into Isochorismate Synthase: Implications for the Evolution of Chorismate-Utilizing Enzymes. *Angewandte Chemie-International Edition* 54(38): 11270-11274.

Pommie, C., Levadoux, S., Sabatier, R., Lefranc, G. and Lefranc, M. P. (2004). IMGT Standardized Criteria for Statistical Analysis of Immunoglobulin V-REGION Amino Acid Properties. *Journal of Molecular Recognition* 17(1): 17-32.

Raboni, S., Bettati, S. and Mozzarelli, A. (2009). Tryptophan Synthase: A Mine for Enzymologists. *Cellular and molecular life sciences : CMLS* 66(14): 2391-2403.

Reyes-Lamothe, R. and Sherratt, D. J. (2019). The Bacterial Cell Cycle, Chromosome Inheritance and Cell Growth. *Nature Reviews Microbioliology*.

Rohweder, B., Semmelmann, F., Endres, C. and Sterner, R. (2018). Standardized Cloning Vectors for Protein Production and Generation of Large Gene Libraries in *Escherichia coli*. *BioTechniques* 64(1): 24-26.

Romney, D. K., Murciano-Calles, J., Wehrmuller, J. E. and Arnold, F. H. (2017). Unlocking Reactivity of TrpB: A General Biocatalytic Platform for Synthesis of Tryptophan Analogues. *Journal of the American Chemical Society* 139(31): 10769-10776.

Ruisinger, A. (2019). Identification of Active Site Loop Confromations and Analysis of their Significance for Enzymatic Function in Imidazole Glycerol Phosphate Synthase. Master Thesis. University of Regensburg.

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. *The Journal of biological chemistry* 279(28): 29572-29582.

Schneider, T. R., Gerhardt, E., Lee, M., Liang, P. H., Anderson, K. S. and Schlichting, I. (1998). Loop Closure and Intersubunit Communication in Tryptophan Synthase. *Biochemistry* 37(16): 5394-5406.

Schrodinger, LLC (2015). The PyMOL Molecular Graphics System, Version 1.8.

Semmelmann, F., Straub, K., Nazet, J., Rajendran, C., Merkl, R. and Sterner, R. (2019). Mapping the Allosteric Communication Network of Aminodeoxychorismate Synthase. *Journal of molecular biology* 431(15): 2718-2728.

Sethi, A., Eargle, J., Black, A. A. and Luthey-Schulten, Z. (2009). Dynamical Networks in tRNA:Protein Complexes. *PNAS* 106(16): 6620-6625.

Siddiq, M. A., Hochberg, G. K. and Thornton, J. W. (2017). Evolution of Protein Specificity: Insights from Ancestral Protein Reconstruction. *Current opinion in structural biology* 47: 113-122.

Starr, T. N. and Thornton, J. W. (2016). Epistasis in Protein Evolution. *Protein science* 25(7): 1204-1218.

Sterner, R. and Höcker, B. (2005). Catalytic Versatility, Stability, and Evolution of the $(\beta\alpha)_8$-Barrel Enzyme Fold. *Chemical reviews* 105(11): 4038-4055.

Straub, K. and Merkl, R. (2019). Ancestral Sequence Reconstruction as a Tool for the Elucidation of a Stepwise Evolutionary Adaptation. *Methods Mol Biol* 1851: 171-182.

VanAernum, Z. L., Gilbert, J. D., Belov, M. E., Makarov, A. A., Horning, S. R. and Wysocki, V. H. (2019). Surface-Induced Dissociation of Noncovalent Protein Complexes in an Extended Mass Range Orbitrap Mass Spectrometer. *Anal Chem* 91(5): 3611-3618.

VanWart, A. T., Eargle, J., Luthey-Schulten, Z. and Amaro, R. E. (2012). Exploring Residue Component Contributions to Dynamical Network Models of Allostery. *Journal of Chemical Theory and Computation* 8(8): 2949-2961.

Vining, L. C. (1992). Secondary Metabolism, Inventive Evolution and Biochemical Diversity--a Review. *Gene* 115(1-2): 135-140.

Waitt, G. M., Xu, R., Wisely, G. B. and Williams, J. D. (2008). Automated In-Line Gel Filtration for Native State Mass Spectrometry. *J Am Soc Mass Spectrom* 19(2): 239-245.

Wang, W. Y. and Malcolm, B. A. (1999). Two-Stage PCR Protocol Allowing Introduction of Multiple Mutations, Deletions and Insertions Using QuikChange™ Site-Directed Mutagenesis. *BioTechniques* 26(4): 680-682.

Washburn, R. S. and Gottesman, M. E. (2015). Regulation of Transcription Elongation and Termination. *Biomolecules* 5(2): 1063-1078.

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(9): 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. *European Journal of Biochemistry* 65(2): 375-385.

Wellington, S., Nag, P. P., Michalska, K., Johnston, S. E., Jedrzejczak, R. P., Kaushik, V. K., Clatworthy, A. E., Siddiqi, N., McCarren, P., Bajrami, B., Maltseva, N. I., Combs, S., Fisher, S. L., Joachimiak, A., Schreiber, S. L. and Hung, D. T. (2017). A Small-Molecule Allosteric Inhibitor of *Mycobacterium tuberculosis* Tryptophan Synthase. *Nature chemical biology*.

Wiesner, S. and Sprangers, R. (2015). Methyl Groups as NMR Probes for Biomolecular Interactions. *Current opinion in structural biology* 35: 60-67.

Wodak, S. J., Paci, E., Dokholyan, N. V., Berezovsky, I. N., Horovitz, A., Li, J., Hilser, V. J., Bahar, I., Karanicolas, J., Stock, G., Hamm, P., Stote, R. H., Eberhardt, J., Chebaro, Y., Dejaegere, A., Cecchini, M., Changeux, J. P., Bolhuis, P. G., Vreede, J., Faccioli, P., Orioli, S., Ravasio, R., Yan, L., Brito, C., Wyart, M., Gkeka, P., Rivalta, I., Palermo, G., McCammon, J. A., Panecka-Hofman, J., Wade, R. C., Di Pizio, A., Niv, M. Y., Nussinov, R., Tsai, C. J., Jang, H., Padhorny, D., Kozakov, D. and McLeish, T. (2019). Allostery in Its Many Disguises: From Theory to Applications. *Structure* 27(4): 566-578.

Yang, X. J. and Miles, E. W. (1992). Threonine 183 and Adjacent Flexible Loop Residues in the Tryptophan Synthase α Subunit Have Critical Roles in Modulating the Enzymatic Activities of the β Subunit in the $\alpha_2\beta_2$ Complex. *The Journal of biological chemistry* 267(11): 7520-7528.

Yanofsky, C. and Rachmeler, M. (1958). The Exclusion of Free Indole as an Intermediate in the Biosynthesis of Tryptophan in *Neurospora crassa*. *Biochimica et biophysica acta* 28(3): 640-641.

Yao, X.-Q., Momin, M. and Hamelberg, D. (2019). Establishing a Framework of Using Residue–Residue Interactions in Protein Difference Network Analysis. *Journal of Chemical Information and Modeling*.

Yutani, K., Ogasahara, K., Tsujita, T., Kanemoto, K., Matsumoto, M., Tanaka, S., Miyashita, T., Matsushiro, A., Sugino, Y. and Miles, E. W. (1987). Tryptophan Synthase α Subunit Glutamic Acid 49 is Essential for Activity. Studies with 19 Mutants at Position 49. *The Journal of biological chemistry* 262(28): 13429-13433.

Acknowledgment

This fascinating journey comes to an end and there are a lot of people, who contributed and to whom I want to express my gratitude.

Foremost, I would like to thank my PhD advisor Prof. Dr. Reinhard Sterner for his guidance and constant support throughout this thesis. I am grateful for the opportunity to gain expertise in the course of this project and to develop on a personal and scientific level. In addition, I want to thank you for the opportunities to attend conferences that contributed substantially to my development.

Furthermore, I would like to thank Prof. Dr. Rainer Merkl for his mentorship during this thesis, support in the field of bioinformatics and his contribution on the evolution of figures.

I want to thank Prof. Dr. Harald Kolmar for his mentorship in context of the Regensburg International Graduate School auf Life Sciences.

I am very grateful to Christiane Endres, Sonja Fuchs, Sabine Laberer and Jeannette Ueckert for their invaluable technical support and terrific teamwork.

Thank you to Claudia Pauer for administrative support.

I would like to thank Dr. Chitra Rajendran and Klaus-Jürgen Tiefenbach for their support regarding crystallography and IT infrastructure, respectively. In addition, I want to thank them for interesting and enriching discussions during lunch breaks.

Extraordinary thanks to Dr. Sandra Schlee for scientific support and scientific as well as personal discussions.

I gratefully thank my collaboration partners Dr. Florian Busch for performing native mass spectrometry in course of the TrpA project; Prof. Dr. Vicki Wysocki for making available state-of-the-art mass spectrometry instrumentation; Dr. Kristina Straub and Prof. Dr. Rainer Merkl for the informative and fruitful cooperation on the TrpB project.

I thank my rotation students Thomas Kinateder, Manuel Saller, Maria Deichner and Moritz Ertelt for their interest and enthusiasm. I hope you learned as much as I did.

I would also like to thank all former members of the Sterner group I spend time with during my bachelor and master thesis.

Cordial thanks to my current colleagues in the Sterner group for their valuable support and almost always scientific discussions: PD Dr. Patrick Babinger, Markus Busch,

Leonhard Heizinger, Enrico Hupfeld, Thomas Klein, Dr. Andrea Kneuttinger, Cosimo Kropp, Julian Nazet and Florian Semmelmann. I appreciate the friendly atmosphere during work as well as on great excursions, barbecue and Christmas parties.

Especially I want to thank my former supervisor Dr. Florian Busch and my former colleagues Dr. Alexandra Holinski, Dr. Wolfgang Kaiser and Dr. Mona Linde, who became dear friends. Thank you for your personal support, numerous coffees and many joyful hours to come.

Finally, I want to express my utmost gratitude to my parents and grandparents, all of them contributed in an incredible way to who I am today. I am also deeply grateful to my whole family and my friends for their encouragement and constant support. Especially, I want to thank Julia for her constant understanding, confidence, unlimited support and everything that goes beyond this work.