

## RESEARCH ARTICLE



# A naturally occurring standalone TrpB enzyme provides insights into allosteric communication within tryptophan synthase

Thomas Kinader<sup>1</sup> | Lukas Drexler<sup>1</sup> | Cristina Duran<sup>2</sup> |  
Sílvia Osuna<sup>2,3</sup> | Reinhard Sterner<sup>1</sup>

<sup>1</sup>Institute of Biophysics and Physical Biochemistry, Regensburg Center for Biochemistry, University of Regensburg, Regensburg, Germany

<sup>2</sup>Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química, Universitat de Girona, Girona, Spain

<sup>3</sup>ICREA, Barcelona, Spain

## Correspondence

Sílvia Osuna, Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química, Universitat de Girona, Girona 17003, Spain.  
Email: [silvia.osuna@udg.edu](mailto:silvia.osuna@udg.edu)

Reinhard Sterner, Institute of Biophysics and Physical Biochemistry, Regensburg Center for Biochemistry, University of Regensburg, D-93040 Regensburg, Germany.  
Email: [reinhard.sterner@ur.de](mailto:reinhard.sterner@ur.de)

## Funding information

Generalitat de Catalunya for the consolidated group TCBioSys, Grant/Award Number: SGR 2021 00487; Spanish MICIN, Grant/Award Numbers: PDC2022-133950-I00, PID2021-129034NB-I00; European Research Council, Grant/Award Numbers: ERC-2015-StG-679001, ERC-2022-CoG-101088032, ERC-2022-POC-101112805, ERC-2023-POC-101158166; Spanish MINECO, Grant/Award Number: PRE2019-089147; Fonds der Chemischen Industrie

**Review Editor:** Lynn Kamerlin

## Abstract

Allosteric regulation of catalytic activity is a widespread property of multi-enzyme complexes. The tryptophan synthase is a prototypical allosteric enzyme where the constituting  $\alpha$  (TrpA) and  $\beta$  (TrpB) subunits mutually activate each other in a manner that is incompletely understood. Experimental and computational studies have shown that LBCA-TrpB from the last bacterial common ancestor contains six residues (Res<sub>6</sub>) distal from the active site that allow for high stand-alone catalytic activity in the absence of a TrpA subunit. In the present study, a database search revealed that Res<sub>6</sub> is also present in the extant *p*/TrpB from *Pelodictyon luteolum*. The *p*/TrpB enzyme showed a high stand-alone activity and only a moderate activation by *p*/TrpA. The replacement of LBCA-Res<sub>6</sub> in *p*/TrpB with the consensus residues from a multiple sequence alignment yielded *p*/TrpB-con, which showed a dramatically decreased stand-alone activity but was strongly stimulated by *p*/TrpA. These findings suggest that the effect of these six key allosteric residues is largely independent of the protein context within a specific TrpB enzyme. Analysis of the conformational landscapes of *p*/TrpB and *p*/TrpB-con revealed that *p*/TrpB in isolation displays efficient closure of both the active site and the communication (COMM) domain. In contrast, these catalytically competent states are destabilized in *p*/TrpB-con but can be recovered by the addition of *p*/TrpA. A correlation-based shortest path map (SPM) analysis reveals that the catalytically and allosterically relevant domains—specifically, the COMM domain in TrpB and loops 2 and 6 in TrpA—are tightly interconnected exclusively in *p*/TrpA:*p*/TrpB-con.

## KEYWORDS

allostery, molecular dynamics simulations, multi-enzyme complexes, shortest path map, stand-alone enzyme, tryptophan synthase

## 1 | INTRODUCTION

The term allostery describes the regulation of a biological macromolecule through the binding of a small

effector molecule to a position distant from the functional site (Hofmann, 2023; Monod et al., 1963). Allostery is a widespread phenomenon that can be observed, for example, in the context of cell signaling and receptor proteins (Cournia & Chatzigoulas, 2020), oligomeric proteins like hemoglobin that cooperatively

Thomas Kinader and Lukas Drexler contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Protein Science* published by Wiley Periodicals LLC on behalf of The Protein Society.

bind a ligand, or enzymes in which the catalytic activity is fine-tuned by an inhibiting or an activating molecule (Lisi & Loria, 2017).

Comprehending allostery at a molecular level is key to unraveling central questions at the heart of molecular biology. In the context of biocatalysis, a thorough understanding and the ability to tune allostery will unlock the catalytic potential of metabolic enzymes, paving the way to a sustainable and green bioproduction of fine chemicals (Buller et al., 2015; Murciano-Calles et al., 2016). The impact of mutations distal from the active site on the catalytic activity of enzymes indicates that a considerable amount of long-range allosteric effects takes place in many enzymatic systems (Gunasekaran et al., 2004; Osuna, 2021). Several models were put forward to explain allostery, assuming either the propagation of an input signal via interconnecting residues (allosteric network model) or an equilibrium between states of low and high activities (conformational selection and population shift model) (Liu & Nussinov, 2016). Despite considerable efforts, the mechanistic basis of allostery, in general, is still poorly understood (Hofmann, 2023), and additional studies are needed to further our understanding of this widespread molecular feature.

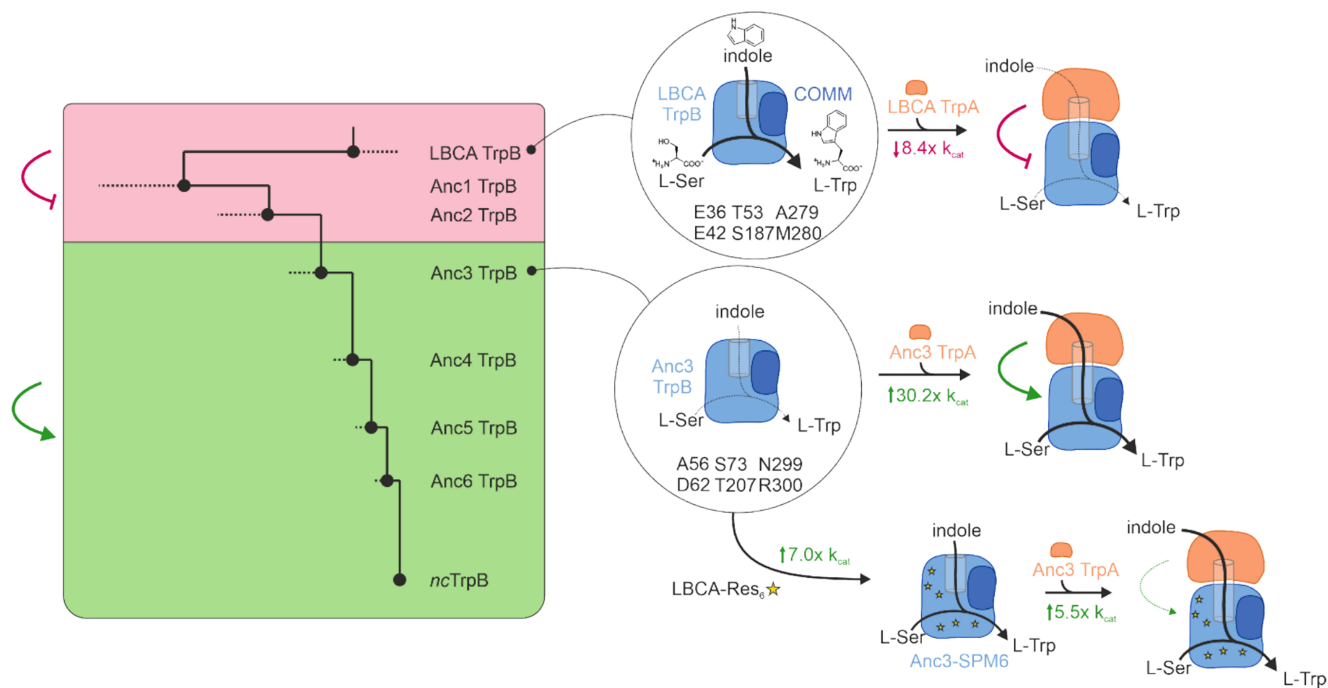
A popular model enzyme to study allostery is tryptophan synthase (TS) because it combines two enzymatic activities that tightly allosterically regulate each other to synchronize the overall reaction (Brzović et al., 1992; Dunn, 2012; Kirschner et al., 1991). In bacteria, TS occurs as  $\alpha\beta\beta\alpha$  heterotetramer consisting of two different subunits,  $\alpha$  (TrpA) and  $\beta$  (TrpB), whereby the  $\alpha\beta$  (TrpA:TrpB) dimer forms the functional unit that is responsible for the last two steps in tryptophan biosynthesis (Goldberg et al., 1966; Yanofsky & Crawford, 1972). Specifically, TrpA catalyzes the retro-aldol reaction of indole-3-glycerol-phosphate (IGP) to glyceraldehyde-3-phosphate (G-3-P) and indole (Lane & Kirschner, 1991; Yanofsky & Crawford, 1972), which is released into a hydrophobic intermolecular channel that connects the active site of TrpA to the active site of TrpB (Dunn et al., 1990; Hyde et al., 1988). Meanwhile, in the TrpB active site, serine reacts with the cofactor pyridoxal phosphate (PLP) to form an aminoacrylate (AA). Then, AA reacts with indole to form tryptophan while the catalytic cofactor PLP is recovered (Figure S1) (Raboni et al., 2009). To prevent futile reactions catalyzed either by TrpA or TrpB, several layers of allosteric regulation have evolved. First, both TrpA and TrpB are generally activated upon complex formation (Hatanaka et al., 1962; Yanofsky & Crawford, 1972). Second, the binding of IGP to the TrpA active site results in a lowered Michaelis constant for serine ( $K_M^{\text{Ser}}$ ) in TrpB (Dunn, 2012; Dunn et al., 1990), and third, the formation of the AA intermediate in TrpB leads to the catalytically activated conformation of TrpA, which enhances its turnover

number ( $k_{\text{cat}}$ ) for the TrpA reaction (Anderson et al., 1991; Banik et al., 1995).

While the current models are still insufficient to explain all aspects of the allosteric communication within the TS at a molecular level, it is known that the so-called COMM domain, which covers the active site in TrpB, plays an important role in allosteric signaling (Miles, 1991). Moreover, an (extended) open conformation of the COMM domain has been attributed to an inactive state of the TrpB monomer, whereas a closed conformation is associated with the catalytically competent state of TrpB in complex with TrpA (Brzović et al., 1992; Buller et al., 2018; Dunn et al., 2008; Maria-Solano et al., 2019).

In the past, several studies reported on TrpB variants that were freed from their regulation by TrpA, including work using directed evolution (Buller et al., 2015; Murciano-Calles et al., 2016). In our own work, which was based on ancestral sequence reconstruction (ASR) (Schupfner et al., 2020), a remarkable functional switch was observed from the last common ancestor of all bacterial TrpB enzymes (LBCA TrpB) to an extant TrpB variant. Variants that were reasonably active on their own and were deactivated upon TrpA binding, that is, “stand-alone” TrpBs, were substituted along the evolutionary trajectory by TrpB variants that were poorly active on their own and were strongly activated by TrpA binding (Figure 1) (Schupfner et al., 2020). Subsequently, we applied a correlation-based computational method called shortest path map (SPM) (Casadevall et al., 2024; Osuna, 2021) to Anc3TrpB, that is, the first ancestral TrpB presenting allosteric activation, for enhancing its stand-alone activity, thus freeing it from dependence on TrpA activation. SPM identifies those residues within a protein that show correlated motions in Molecular Dynamics (MD) simulations and hence are thought to define the conformational equilibrium between different functional states of the enzyme (Casadevall et al., 2024; Maria-Solano et al., 2021; Osuna, 2021). The SPM conformationally relevant positions, together with sequence comparison between Anc3 and LBCA TrpB and thorough experimental characterization, identified six crucial residue positions ( $\text{Res}_6$ ), most of which are located distal from the active site (Figure S2). The identity of these six residues accounted for most of the observed functional changes between the stand-alone TrpBs and the allosterically controlled TrpBs, which is exemplarily shown for LBCA TrpB and Anc3 TrpB in Figure 1. Importantly, when we introduced the subset of these LBCA- $\text{Res}_6$  from LBCA TrpB into the context of Anc3 TrpB, we observed an increase in the stand-alone activity ( $k_{\text{cat}}$ ) by a factor of 7.0 and a reduction of the allosteric activation exerted by TrpA from 30.2-fold to only 5.5-fold (Figure 1).

Having identified the LBCA- $\text{Res}_6$ , we wondered whether the functional effect of this set of residues was



**FIGURE 1** Changes in the stand-alone activity of TrpB during evolution and based on computational predictions. Left panel: Ancestral sequence reconstruction of TrpB identified a switch from stand-alone TrpB variants that were highly active on their own but were deactivated upon complex formation with TrpA (highlighted in red) to variants that were poorly active on their own but were activated upon complex formation with TrpA (highlighted in green) (Schupfner et al., 2020). Upper right panel: The shortest path map (SPM) method previously identified six residues Res<sub>6</sub> in LBCA (E36, E42, T53, S187, A279, M280), which were assumed to be responsible for the 8.4-fold inactivation by TrpA (Maria-Solano et al., 2021). Middle right panel: The corresponding residues Res<sub>6</sub> in Anc3 TrpB (A56, D62, S73, T207, N299, R300) result in a 30.2-fold activation by TrpA. Lower right panel: The replacement of Res<sub>6</sub> in Anc3 TrpB by Res<sub>6</sub> from LBCA (LBCA-Res<sub>6</sub>) led to Anc3-SPM6, which shows a 7.0-fold improved stand-alone activity and only a 5.5-fold activation by TrpA (Maria-Solano et al., 2021).

dependent on the TrpB context. Therefore, we searched for extant TrpB sequences containing LBCA-Res<sub>6</sub> and investigated whether they also show high stand-alone activity and little dependence on their complementing TrpA partner. Our combined experimental and computational analysis indicates that the effect of these SPM-ASR mutations might be context-independent, as they exert a profound effect on the conformational landscape at different key reaction intermediates, thus impacting the long-range allosteric interaction between TrpA and TrpB subunits.

## 2 | RESULTS

In previous work, we identified six TrpB residues from LBCA, which are E36, E42 located at the surface, T53 at the  $\beta$ - $\beta$  interface, S187 at the active site, and A279, M280 at the  $\alpha$ - $\beta$  interface (Figure S1). The introduction of these LBCA-Res<sub>6</sub> into Anc3 TrpB (A56E + D62E + S73T + T207S + N299A + R300M) led to an increase in its stand-alone activity by 7.0-fold in terms of  $k_{\text{cat}}$  and a reduction of its allosteric activation by TrpA from 30.2-fold to 5.5-fold (Figure 1). To check whether LBCA-Res<sub>6</sub> were retained in any extant TrpB variant, we retrieved all 6373 sequences annotated as

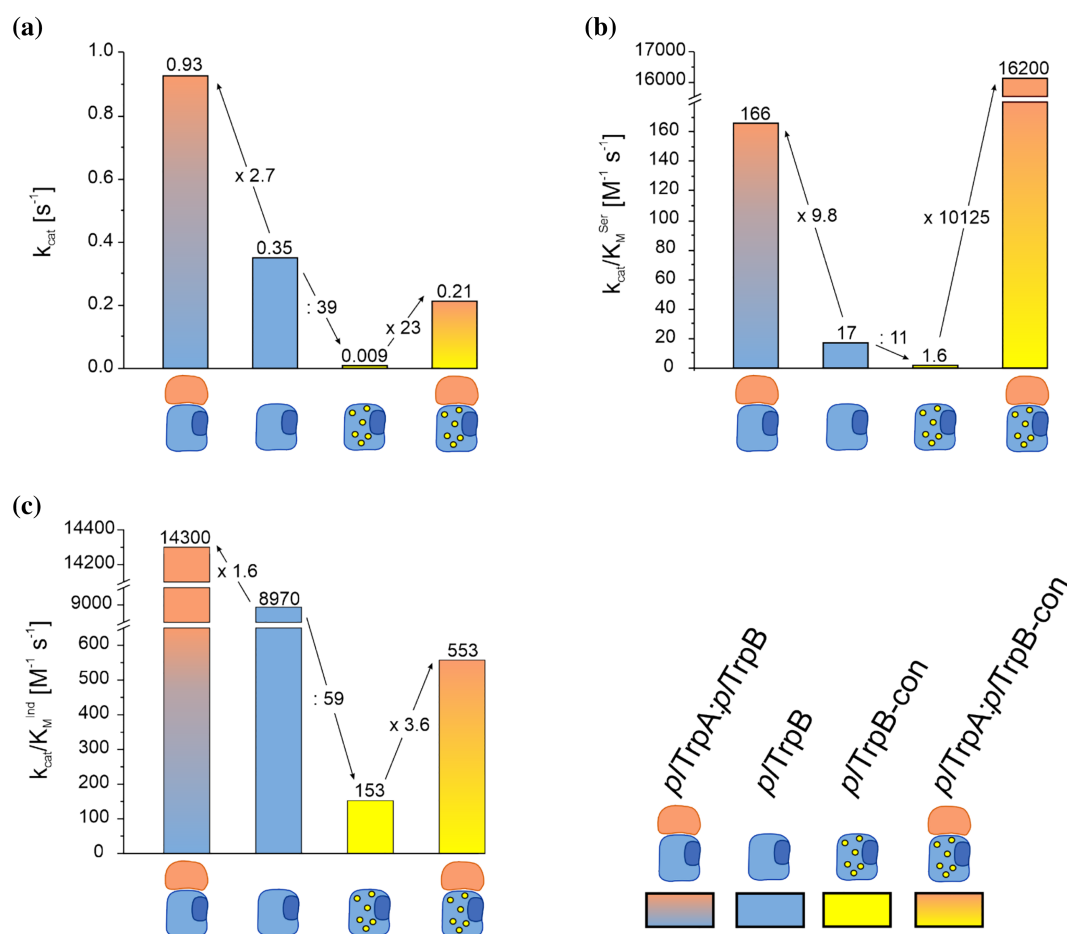
TrpB from the KEGG database (as of December 2021) and constructed an MSA. By filtering the dataset for the occurrence of one of these 6 amino acids at a time, we identified TrpB from *Pelodictyon luteolum* (*p*/TrpB) as the only sequence that possesses the whole Res<sub>6</sub> set.

To test the stand-alone activity of *p*/TrpB and its allosteric regulation by *p*/TrpA, the corresponding genes were codon-optimized for *Escherichia coli*, cloned into an appropriate vector, and expressed. Following the purification of recombinant *p*/TrpB and *p*/TrpA by affinity chromatography and size-exclusion chromatography, the purity of the enzymes was confirmed by SDS-PAGE (Figure S3) and their structural integrity was validated by far-UV CD spectroscopy (Figure S4).

Steady-state kinetic experiments were performed in which TrpB activity was assessed by a spectrophotometric assay directly observing the changes in absorption between indole and L-Trp. In the absence of *p*/TrpA, *p*/TrpB showed a  $k_{\text{cat}}$  value of  $0.35 \text{ s}^{-1}$  and a catalytic efficiency ( $k_{\text{cat}}/K_{\text{M}}$ ) of  $17 \text{ M}^{-1} \text{ s}^{-1}$  for serine and  $8970 \text{ M}^{-1} \text{ s}^{-1}$  for indole (Figure S5 and Table 1). While both  $k_{\text{cat}}$  and  $k_{\text{cat}}/K_{\text{M}}^{\text{Ind}}$  were comparatively high for an isolated TrpB enzyme,  $k_{\text{cat}}/K_{\text{M}}^{\text{Ser}}$  was surprisingly low, due to a remarkably high  $K_{\text{M}}^{\text{Ser}}$ . This prompted the question of whether L-Ser was the native

**TABLE 1** Steady-state enzyme kinetic parameters at 30°C of *p*/TrpB and *p*/TrpB-con in isolation and in complex with *p*/TrpA.

	$k_{\text{cat}}$ [ $\text{s}^{-1}$ ]	$K_{\text{M}}^{\text{Ser}}$ [mM]	$K_{\text{M}}^{\text{Ind}}$ [ $\mu\text{M}$ ]	$k_{\text{cat}}/K_{\text{M}}^{\text{Ser}}$ [ $\text{s}^{-1} \text{M}^{-1}$ ]	$k_{\text{cat}}/K_{\text{M}}^{\text{Ind}}$ [ $\text{s}^{-1} \text{M}^{-1}$ ]
<i>p</i> /TrpB	$0.35 \pm 0.02$	$21 \pm 3$	$39 \pm 7$	$17 \pm 2.6$	$8970 \pm 1720$
<i>p</i> /TrpA: <i>p</i> /TrpB	$0.93 \pm 0.21$	$5.6 \pm 0.4$	$65 \pm 7$	$166 \pm 39$	$14,300 \pm 3510$
<i>p</i> /TrpB-con	$0.009 \pm 0.002$	$5.8 \pm 0.7$	$59 \pm 3$	$1.6 \pm 0.39$	$153 \pm 34$
<i>p</i> /TrpA: <i>p</i> /TrpB-con	$0.21 \pm 0.06$	$0.013 \pm 0.001$	$380 \pm 76$	$16,200 \pm 4770$	$553 \pm 193$

**FIGURE 2** Steady-state kinetic parameters  $k_{\text{cat}}$  (a),  $k_{\text{cat}}/K_{\text{M}}^{\text{Ser}}$  (b), and  $k_{\text{cat}}/K_{\text{M}}^{\text{Ind}}$  (c) of *p*/TrpB and *p*/TrpB-con in isolation and in complex with *p*/TrpA. The data are taken from Table 1 and present the fold change in catalytic parameters resulting from either the addition of *p*/TrpA (indicated in orange) or the mutation of the Res<sub>6</sub> to the consensus (indicated as yellow spheres).

substrate of *p*/TrpB or, as in the case of TrpB2 enzymes, O-phospho-L-serine (Busch et al., 2014). To test this, we used an HPLC-based assay to investigate *p*/TrpB activity with O-phospho-L-serine and other possible substrates (D-serine, O-acetyl-L-serine, O-phospho-D-serine, O-phospho-L-threonine, L-threonine, and L-cysteine). However, only in the presence of L-serine complete turnover of indole to L-tryptophan was detected under the given experimental conditions (Figure S6).

The shown data are the average and standard deviation of values obtained by triplicate measurements. Experimental conditions included 50 mM potassium

phosphate (pH 7.5), 180 mM KCl, 40  $\mu\text{M}$  PLP, saturating concentrations of one substrate (indole/L-serine) and varying concentrations of the other substrate (L-serine/indole). The individual substrate saturation curves and the values determined from their fitting to the Michaelis–Menten equation are shown in Figure S5.

Next, we investigated the effect of complex formation with *p*/TrpA on the catalytic activity of *p*/TrpB. In the presence of a 1.5-fold molar excess of *p*/TrpA over *p*/TrpB, we recorded a  $k_{\text{cat}}$  of  $0.93 \text{ s}^{-1}$ , and catalytic efficiencies of  $166 \text{ M}^{-1} \text{ s}^{-1}$  for serine and  $14,300 \text{ M}^{-1} \text{ s}^{-1}$  for indole (Figure S5 and Table 1). This



translates to a moderate activation with respect to  $k_{\text{cat}}$  by a factor of 2.7 (Figure 2a), for  $k_{\text{cat}}/K_{\text{M}}^{\text{Ser}}$  by a factor of 9.8 (Figure 2b), and for  $k_{\text{cat}}/K_{\text{M}}^{\text{Ind}}$  by a factor of 1.6 (Figure 2c). Taken together, these results suggest that—except for the low apparent affinity to L-serine— $p\text{TrpB}$  exploits a substantial fraction of its catalytic potential already in the absence of  $p\text{TrpA}$ , supporting that the LBCA-Res<sub>6</sub> conveys stand-alone activity independent of the protein context.

To further substantiate this hypothesis, we exchanged the LBCA-Res<sub>6</sub> of  $p\text{TrpB}$  with the consensus amino acid found at each position as determined from the MSA of all TrpB1 enzymes. The resulting  $p\text{TrpB-con}$  variant contained the substitutions E46A + E52D + T63S + S197T + A289G + M290N.  $p\text{TrpB-con}$  was produced in the same way as wild-type  $p\text{TrpB}$ , and its purity and structural integrity were again experimentally confirmed by means of SDS-PAGE and CD spectroscopy (Figures S3 and S4).

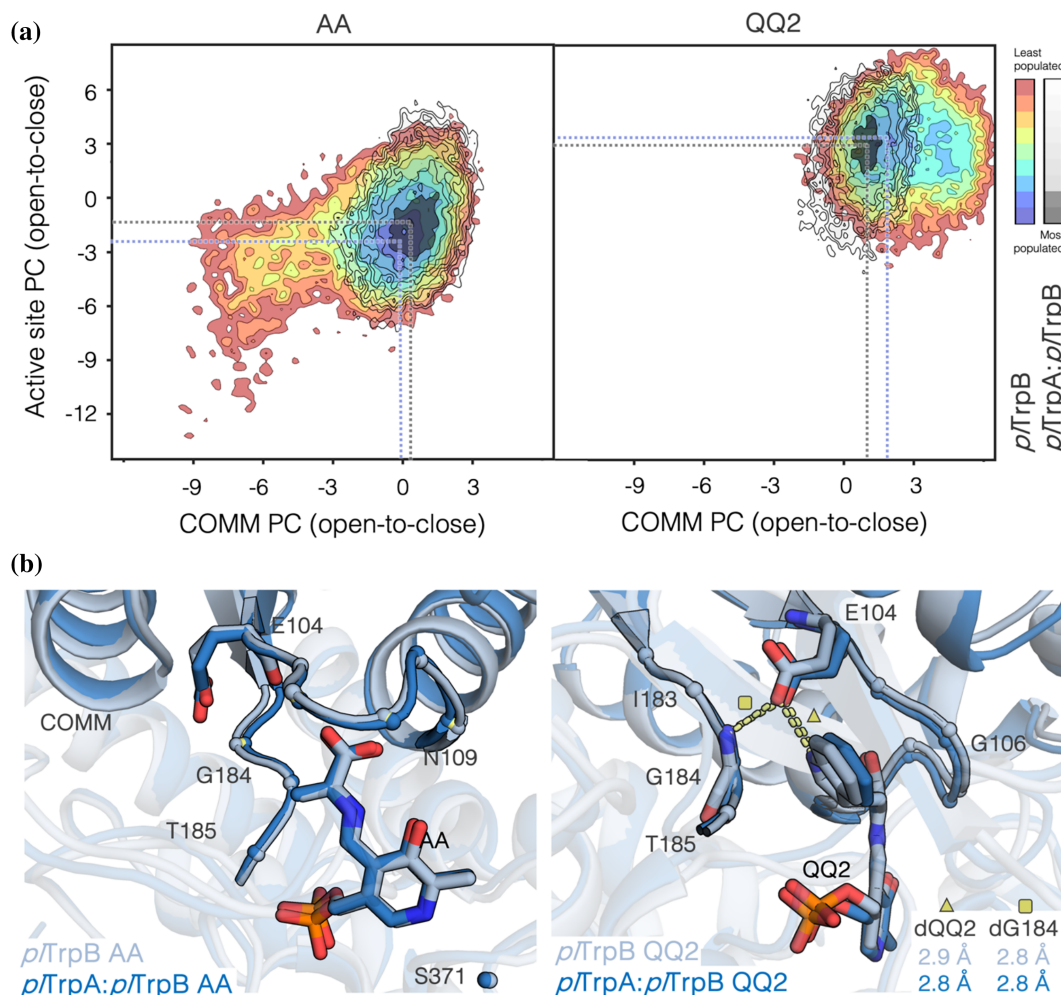
We next performed steady-state kinetic experiments with  $p\text{TrpB-con}$  in isolation, which yielded a  $k_{\text{cat}}$  of  $0.009 \text{ s}^{-1}$  and catalytic efficiencies of  $1.6 \text{ M}^{-1} \text{ s}^{-1}$  for serine and  $153 \text{ M}^{-1} \text{ s}^{-1}$  for indole (Figure S5 and Table 1). Compared to  $p\text{TrpB}$ , these values correspond to a large decrease of  $k_{\text{cat}}$  by a factor of 39 (Figure 2a), of  $k_{\text{cat}}/K_{\text{M}}^{\text{Ser}}$  by a factor of 11 (Figure 2b), and of  $k_{\text{cat}}/K_{\text{M}}^{\text{Ind}}$  by a factor of 59 (Figure 2c). Obviously, these reductions could be either caused by a significantly reduced catalytic potential of  $p\text{TrpB-con}$  compared to  $p\text{TrpB}$  or a much stronger dependency of  $p\text{TrpB-con}$  on allosteric activation by  $p\text{TrpA}$ . To discriminate between these two alternatives, we measured the activity of  $p\text{TrpB}$  in the presence of a 1.5-fold excess of  $p\text{TrpA}$ . Remarkably, the determined  $k_{\text{cat}}$  value was  $0.21 \text{ s}^{-1}$ , while the catalytic efficiencies were  $16,200 \text{ M}^{-1} \text{ s}^{-1}$  for serine and  $553 \text{ M}^{-1} \text{ s}^{-1}$  for indole, which corresponds to activations of  $k_{\text{cat}}$  by a factor of 23 (Figure 2a), of  $k_{\text{cat}}/K_{\text{M}}^{\text{Ser}}$  by a factor of more than 10,000 (Figure 2b), and of  $k_{\text{cat}}/K_{\text{M}}^{\text{Ind}}$  by a factor of 3.6 (Figure 2c). Interestingly, the  $k_{\text{cat}}/K_{\text{M}}^{\text{Ser}}$  of the  $p\text{TrpA}:p\text{TrpB-con}$  complex surpassed the value for the  $p\text{TrpA}:p\text{TrpB}$  complex by approximately two orders of magnitude.

To exclude that the observed increased  $p\text{TrpA}$ -induced allosteric activation of  $p\text{TrpB-con}$  compared to  $p\text{TrpB}$  is caused by higher subunit binding affinities in the  $p\text{TrpA}:p\text{TrpB-con}$  than in the  $p\text{TrpA}:p\text{TrpB}$  complex, we determined apparent  $K_{\text{d}}^{\text{app}}$  values by activity titrations. For this, we added varying concentrations of  $p\text{TrpA}$  to a constant concentration of  $p\text{TrpB}$  or  $p\text{TrpB-con}$  and measured the reaction rate at saturating substrate concentrations. The determined apparent dissociation constants  $K_{\text{d}}^{\text{app}}$  of  $8.5 \mu\text{M}$  for  $p\text{TrpA}:p\text{TrpB}$  (Figure S7) and  $9.9 \mu\text{M}$  for  $p\text{TrpA}:p\text{TrpB-con}$  (Figure S8) indicate that the fractions of  $p\text{TrpB}$  and  $p\text{TrpB-con}$  complexed by  $p\text{TrpA}$  as present in our steady-state kinetic measurements were basically

identical. This finding was confirmed by analytical size exclusion chromatography. The determined elution volumes showed that  $p\text{TrpA}$  in isolation forms  $\alpha$ -monomers, whereas for both  $p\text{TrpB}$  and  $p\text{TrpB-con}$  a  $\beta\beta$ -dimer was detected. Mixing of  $p\text{TrpA}$  with  $p\text{TrpB}$  or  $p\text{TrpB-con}$  yielded similar fractions of isolated  $\alpha$ , isolated  $\beta\beta$ , as well as  $\alpha\beta\beta$  and  $\alpha\beta\beta\alpha$  complexes for  $p\text{TrpA}:p\text{TrpB}$  and  $p\text{TrpA}:p\text{TrpB-con}$ , respectively (Figure S9). To estimate the theoretical allosteric activation for complete  $\alpha\beta\beta\alpha$  complex formation, we performed a simple calculation based on the law of mass action, according to which the actual  $k_{\text{cat}}$  values of  $p\text{TrpB}$  and  $p\text{TrpB-con}$  in the respective complexes may have been underestimated by factors of 2.5 for  $p\text{TrpA}:p\text{TrpB}$  and 3.6 for  $p\text{TrpA}:p\text{TrpB-con}$ , respectively (Calculation S1). Hence, the allosteric activation by  $p\text{TrpA}$  is even more pronounced for  $p\text{TrpB-con}$  compared to  $p\text{TrpB}$  than has been estimated from the results of the steady-state measurements listed in Table 1.

We next asked why  $p\text{TrpB}$  is much more active than  $p\text{TrpB-con}$  and why  $p\text{TrpB-con}$  is much more efficiently allosterically activated by  $p\text{TrpA}$ , even though most of the differences between the two proteins are conservative amino acid exchanges. To address these questions, we resorted to a computational analysis of the two  $p\text{TrpB}$  variants in isolation and in complex with  $p\text{TrpA}$ , considering two key reaction intermediates: the aminoacrylate (AA) Schiff base and the quinonoid II (QQ2) (Figure S1). In a previous study, we found that the evaluation of the conformational landscape of  $p\text{TrpB}$  from *Pyrococcus furiosus* and some stand-alone variants at these two reaction intermediates was crucial for rationalizing the changes in stand-alone activity and allosteric regulation (Maria-Solano et al., 2019). Since the catalytic proficiency of TrpB depends on efficient closure of the active site and allosteric communication mediated by the COMM domain (Dunn et al., 2008), we decided to reconstruct the conformational landscapes based on Principal Component analysis (PCA), focusing on these two parameters in our analysis (Section 4, Figure 3 and Figure S10).

The conformational landscape of  $p\text{TrpB}$  alone and in the  $p\text{TrpA}:p\text{TrpB}$  complex reveals a higher flexibility of the COMM domain (x-axis) in the case of  $p\text{TrpB}$ , especially at the AA intermediate (Figure 3a), thus indicating that  $p\text{TrpB}$  in the absence of  $p\text{TrpA}$  can adopt both open and closed conformations. This can be interpreted as a restriction of the conformational space that is accessible for  $p\text{TrpB}$  exerted by its binding partner  $p\text{TrpA}$ . However, only minor differences are observed in the most populated conformation, as both  $p\text{TrpB}$  and  $p\text{TrpA}:p\text{TrpB}$  share the same minima at closed conformations of the COMM and active site (minima are at ca. 0, -2 for AA and at ca. 1.5, 3 at QQ2, Figure 3a). The overlay of a representative structure of the most stable conformation for  $p\text{TrpB}$  and  $p\text{TrpA}$ :



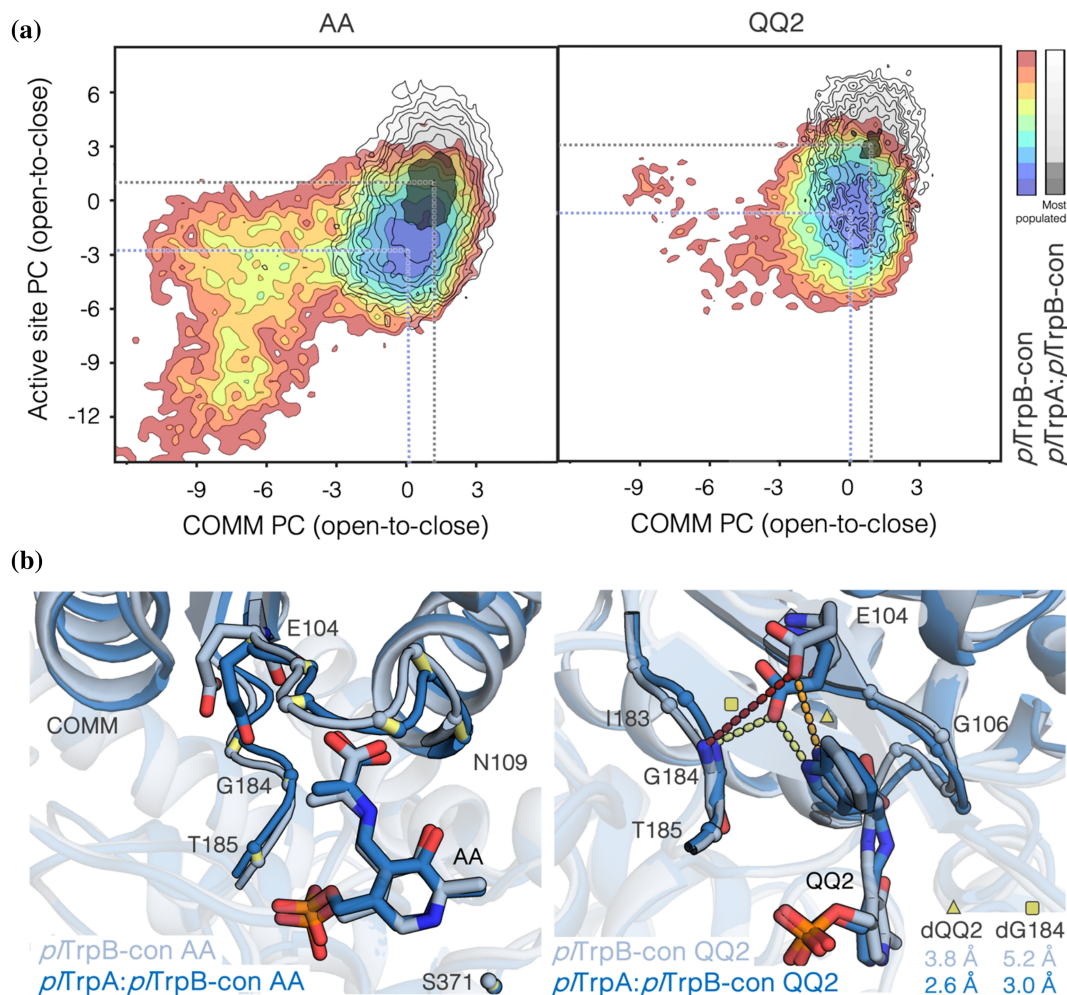
**FIGURE 3** Reconstructed conformational landscapes of  $p\text{TrpB}$  in isolation and in the  $p\text{TrpA}:p\text{TrpB}$  complex at the aminoacrylate AA (left) and quinonoid QQ2 (right) reaction intermediates. (a) Overlay of the PCA-generated conformational landscape considering distances between C $\alpha$ -atoms of the residues included in the COMM domain (x-axis) and active site (y-axis). Negative values of PC correspond to open states, whereas positive values correspond to closed states. The conformational landscape of  $p\text{TrpB}$  is represented using a red-to-blue coloring scheme (red for the least populated conformation and blue for the most populated ones), whereas the landscape of the  $p\text{TrpA}:p\text{TrpB}$  complex is shown with black lines (the most stable minimum is colored in black). (b) Overlay of two representative structures extracted from the most stable minima for  $p\text{TrpB}$  (light blue) and  $p\text{TrpA}:p\text{TrpB}$  (dark blue) at the AA (left panel) and QQ2 (right) intermediates. The mean distance between E104 and G184 or indole nitrogen at the QQ2 state is represented in Å.

$p\text{TrpB}$  indeed shows minor structural differences at both the AA and QQ2 intermediates (Figure 3b), hence indicating that  $p\text{TrpB}$  in the absence of  $p\text{TrpA}$  can adopt the catalytically productive closed states of the COMM domain and active site. The small conformational deviations observed when comparing isolated  $p\text{TrpB}$  with  $p\text{TrpA}:p\text{TrpB}$  are in line with the high stand-alone activity ( $k_{\text{cat}}$ ) of  $p\text{TrpB}$  (Table 1). In the case of the  $p\text{TrpA}:p\text{TrpB}$  complex, the COMM domain flexibility is restricted as the minimum appears to be overall more narrow at both the AA and QQ2 intermediates, which could explain the slight increase in the  $k_{\text{cat}}$  value upon complex formation. Furthermore, taking into account that  $p\text{TrpB}$  and  $p\text{TrpA}:p\text{TrpB}$  represent the two best catalysts with respect to  $k_{\text{cat}}$ , the shared minimum found in the two conformational landscapes

presenting the COMM domain closed with a proper positioning of E104 to stabilize the positive charge of the nitrogen of the indole at QQ2 represents the catalytically productive conformation, which ideally stabilizes the transition state and leads to the highest turnover numbers.

Although in our MD simulations L-Ser is already coupled to pyridoxal phosphate (AA intermediate), the higher conformational flexibility of the COMM domain observed in the case of  $p\text{TrpB}$ , as well as a more open active site tunnel (Figure S11) might explain the high Michaelis constant  $K_{\text{M}}^{\text{Ser}}$  observed experimentally (Table 1).

Next, we analyzed the conformational landscapes of  $p\text{TrpB}$ -con in isolation and in the  $p\text{TrpA}:p\text{TrpB}$ -con complex. Intriguingly, large differences are observed

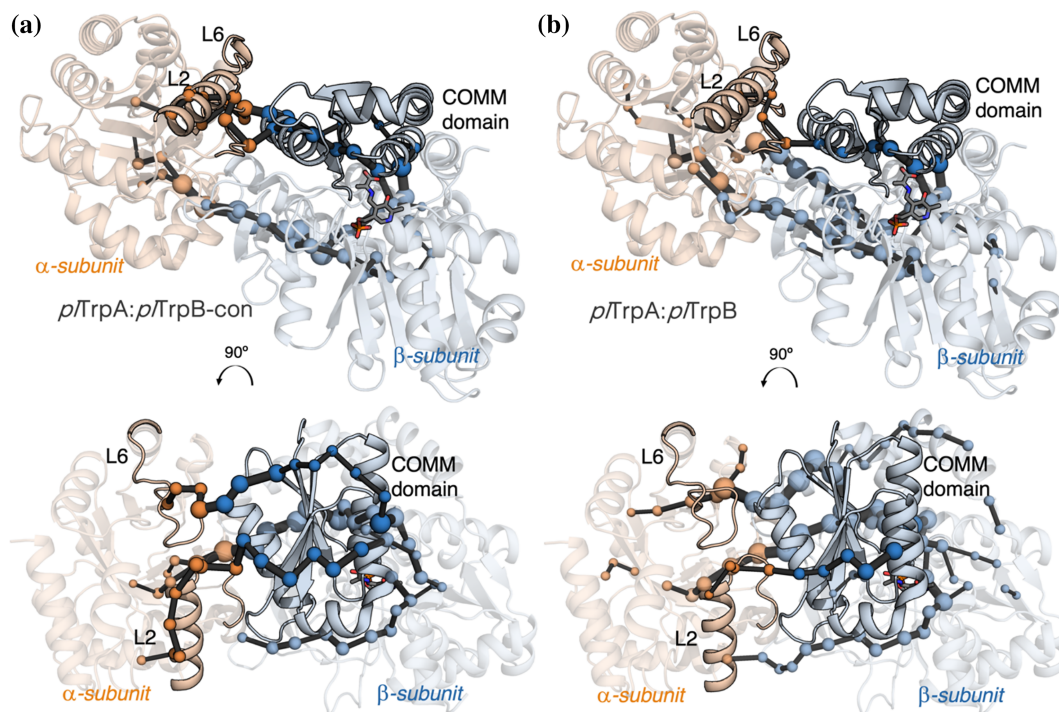


**FIGURE 4** Reconstructed conformational landscapes of isolated *pTrpB-con* and complexed *pTrpB-con* in *pTrpA:pTrpB-con* at the aminoacrylate AA (left) and quinonoid QQ2 (right) reaction intermediates. (a) Overlay of the PCA-generated conformational landscape considering distances between C $\alpha$ -atoms of the residues included in the COMM domain (x-axis) and active site (y-axis). Negative values of PC correspond to open states, whereas positive ones correspond to closed states. The conformational landscape of *pTrpB-con* is represented using a red-to-blue coloring scheme (red for the least populated conformation, blue for the most populated ones), whereas the landscape of the *pTrpA:pTrpB-con* complex is shown with black lines (the most stable minimum is colored in black). (b) Overlay of two representative structures extracted from the most stable minima for *pTrpB-con* (light blue) and *pTrpA:pTrpB-con* (dark blue) at the AA (left panel) and QQ2 (right) intermediates. The mean distance between E104 and G184 or indole nitrogen at the QQ2 state is represented in Å.

between *pTrpB-con* and *pTrpA:pTrpB-con*: in the absence of the binding partner, *pTrpB-con* exhibits a high conformational heterogeneity for both the active site and COMM domain and, more importantly, the most stable conformation is substantially shifted when in complex with *pTrpA* (Figure 4a). Specifically, the active site and the COMM domain exhibit a slightly more open conformation at both the AA and QQ2 intermediates in the case of *pTrpB-con*. The overlay of a representative structure of the most stable minima of *pTrpB-con* and *pTrpA:pTrpB-con* (Figure 4b) indeed indicates suboptimal stabilization of QQ2 in *pTrpB-con* (i.e., a larger distance between E104 and the nitrogen of indole, Figure S12), which is in line with the poor  $k_{cat}$  value found for this variant. Some deviations of the COMM domain loop containing E104, G106, and

the 183–185 loop are observed between isolated *pTrpB-con* and *pTrpB-con* in complex (Figure 4b). Interestingly, the conformational landscape of the *pTrpA:pTrpB-con* complex resembles the one found for *pTrpA:pTrpB*; however, a shift in the location of the minimum towards a more closed active site conformation is observed for the former, especially at the AA intermediate. This means that *pTrpA* seems to alter the conformational equilibrium, pushing it towards a more closed active site, which improves the Michaelis constant  $K_M$  for L-Ser, while hampering indole access, thus yielding a higher  $K_M$  for indole in *pTrpA:pTrpB-con* (Table 1). The analysis of the tunnels with respect to the COMM domain closure also evidences a narrower (more closed) tunnel in the case of *pTrpA:pTrpB-con*, again explaining the elevated  $K_M$  for indole found





**FIGURE 5** SPM-identified allosteric communication pathway for *pTrpA:pTrpB-con* (a) and *pTrpA:pTrpB* (b). The key structural elements described to be allosterically connected are labeled: loop 6 (L6), loop 2 (L2) in *pTrpA*, and the COMM domain in *pTrpB/pTrpB-con*. In both cases, the *pTrpA* subunit is represented in orange, whereas *pTrpB/pTrpB-con* is colored in blue.

experimentally (Table 1 and Figure S11). Our MD simulations therefore provide an explanation for the reversal of catalytic efficiency between indole and serine observed when comparing both *pTrpA:pTrpB* and *pTrpA:pTrpB-con* complexes: The improved  $K_M^{\text{Ser}}$  in *pTrpA:pTrpB-con* is explained by the more closed TrpB active site pocket, which helps retain L-Ser in a productive pose for the external aldimine 1 (Aex1) and AA formation. Such a closed conformation of the TrpB active site induces the closing of the TrpA–TrpB tunnel, thus hampering indole diffusion and affecting its  $K_M^{\text{Ind}}$ . The opposite effect is observed in *pTrpA:pTrpB*, which presents a more open TrpB active site pocket detrimental for the productive coupling of L-Ser to PLP, but also a more open TrpA–TrpB tunnel helping indole channeling and thus leading to a lower  $K_M^{\text{Ind}}$  value compared to *pTrpA:pTrpB-con*.

According to our computational and experimental results, standalone activity is obtained if the isolated subunit can efficiently stabilize the catalytically productive closed conformation of the COMM domain and active site, similar to what is observed in the presence of the allosteric binding partner. This is in line with our previous studies focused on comparing the allosterically driven conformational ensemble with that of isolated subunits of *pTrpB* (Casadevall et al., 2022; Maria-Solano et al., 2019). The kinetic characterization of isolated *pTrpB* and *pTrpB-con* in the two corresponding complexes indicates that efficient allosteric

activation is found if (i) the allosteric effector—in our case TrpA—restricts the conformational space of the allosterically controlled enzyme and (ii) catalytically productive states with closed conformations of the COMM domain and active site are stabilized over catalytically unproductive ones.

This leads to the final question of how the mutations influence the communication between the two proteins. To address this problem, we evaluated the allosteric communication between *pTrpA* and *pTrpB/pTrpB-con* using our correlation-based SPM tool and considering the AA intermediate in which the allosteric signal is hypothesized to be maximal (Dunn, 2012; Ito et al., 2022; Ito et al., 2023). The formation of the AA intermediate in TrpB stabilizes the catalytically activated closed conformation of TrpA in which both L6 and L2 cover the active site and promote the IGP retroaldol cleavage (Duran et al., 2024; Kulik et al., 2005). At the same time, the catalytically activated closed state of TrpA favors the closing of the COMM domain for retaining indole and promoting its coupling with AA. Interestingly, we observe such a sophisticated allosteric communication only in the case of *pTrpB-con*: the computed SPM shows an intertwined pathway connecting both active site loops 6 and 2 (L6 and L2) of *pTrpA* with the COMM domain and the *pTrpB-con* active site where PLP is located (Figure 5a). In contrast, *pTrpB*, which is substantially less activated by the presence of *pTrpA*, preserves the communication



between L2 and the COMM domain but lacks the direct connection with L6 (Figure 5b). Therefore, reverting LBCA-Res<sub>6</sub> to the consensus residues has led to a significant change in the dynamic properties and allosteric communication, which impacts activity and substrate affinity.

### 3 | DISCUSSION

In comparison to previously studied extant TrpB enzymes, *p*TrpB exhibits several unusual properties: a superior standalone activity, weak allosteric activation by *p*TrpA, and poor binding affinity for its substrate L-serine. The first two properties can be attributed to the presence of the LBCA-Res<sub>6</sub> residues (Figure 1), which obviously have persisted over time in this extant TrpB enzyme. Consequently, the same residues are responsible for a substantial enhancement of the stand-alone properties of both an ancestral (LBCA TrpB) and an extant (*p*TrpB) enzyme, and this effect is also observed in a previously computationally designed variant (Anc3TrpB-SPM6) (Maria-Solano et al., 2021). It is intriguing to note that this particular set of conformationally relevant residues seems to function independently of the specific protein context, as the comparison of the SPM networks between ancestral and multiple extant TrpBs from different organisms shows similar SPM pathways, all identifying these six positions as relevant (Figure S13). Of note is that this Res<sub>6</sub> subset differs from previously identified sets of residues as deduced from MSAs that can promote standalone activity (Buller et al., 2015; Schupfner et al., 2020). Hence, there exist different sets of activating residues, which could be transferable to other TrpB enzymes, providing a potential strategy for enhancing their activity. In fact, in a recent publication, more than 500,000 variants of TrpB from *Thermotoga maritima* were reported using a continuous evolution strategy, in which key mutations at the enzyme surface, but also at the  $\alpha$ - $\beta$  interface, were found to be important for improving the stand-alone function of *Tm*TrpB (Rix et al., 2024). Actually, one of these key positions was I271 (P271 in LBCA), which is located right next to two of the Res<sub>6</sub> positions studied here, that is, A279 and M280 at the  $\alpha$ - $\beta$  interface.

These characteristics raise fundamental questions about the physiological roles of *p*TrpB as well as *p*TrpA within their host organism, *P. luteolum*. Based on the exclusive substrate specificity of *p*TrpB for L-serine, its sequence homology to other TrpB1 enzymes, and the presence of the characteristic catalytic Asp residue in the active site, this enzyme can be confidently classified as a member of the TrpB1 group (Table S1) (Fleming et al., 2018; Merkl, 2007; Xie et al., 2002). However, analysis of the genomic neighborhood of *p*TrpB (Figure S14) reveals two striking differences compared to other TrpB1 enzymes: First, *pltrpA* and

*pltrpB* are separated by nine interjacent genes. Second, both genes are encoded on different DNA strands and are thus not co-transcribed. Also, neither *pltrpA* nor *pltrpB* is embedded in a *trp* operon. Based on the genomic proximity, further nucleophiles were tested as substrates that might replace indole; however, none of them showed any product formation (Figure S15a). As suggested by the stand-alone properties of *p*TrpB, the moderate affinity for *p*TrpA, as shown by the relatively high  $K_d^{\text{app}}$  value of 8.5  $\mu$ M (Figure S7), and the isolated genomic location of *pltrpA* and *pltrpB*, the organism does not appear to be overly concerned with losing nascent indole from the *p*TrpA:*p*TrpB complex. One could speculate that *P. luteolum* possesses other pathways for indole biosynthesis and scavenges the metabolite by another enzyme. The existence of a TrpB2 protein, a class of TrpB homologues that have a high affinity for indole (Hettwer & Sterner, 2002) supports this hypothesis. Alternatively, uncoupling of *p*TrpA and *p*TrpB may allow the organism to utilize indole for processes other than tryptophan biosynthesis, suggesting that indole may serve as a precursor for other metabolites. Moreover, the stand-alone nature of *p*TrpB led us to hypothesize that it might play a role in L-tryptophan degradation. We therefore analyzed the reverse reaction from L-tryptophan to indole and L-serine using an HPLC-based enzyme assay. However, we found no evidence of indole production (Figure S15b). Further studies are needed to elucidate the precise role of this enzyme within the organism and to understand the underlying molecular mechanisms that account for its unique properties.

The observed decrease in both  $k_{\text{cat}}$  (39-fold) and  $k_{\text{cat}}/K_M^{\text{Ser}}$  (11-fold) when transitioning from wildtype – *p*TrpB to *p*TrpB-con (Figure 2a,b) is striking. However, the subsequent dramatic increase upon the addition of *p*TrpA (24-fold increase for  $k_{\text{cat}}$ , 10.125-fold increase for  $k_{\text{cat}}/K_M^{\text{Ser}}$ ) is equally intriguing. Here, the introduction of mutations initially led to a significant decrease in enzymatic activity, which can, however, be restored upon the addition of an interaction partner. Moreover, in terms of  $k_{\text{cat}}/K_M^{\text{Ser}}$  wildtype levels are even surpassed by two orders of magnitude. Our findings provide a conclusive example of sign epistasis, a phenomenon where mutations are detrimental or negative in one context but beneficial or positive in another context (Weinreich et al., 2005). This is particularly interesting from an evolutionary perspective because it means that mutations that may initially harm an enzyme can become beneficial through the acquisition of an interaction partner, resulting in a dependence on this partner. Such an effect has indeed been observed for the evolution of ribulose-1,5-bisphosphate carboxylase/oxygenase (Schulz et al., 2022). However, during the evolution of stand-alone *p*TrpB from stand-alone LBCA TrpB, either a strong dependence on *p*TrpA has never existed or the enzyme managed at some point during

its evolution to free itself from this sign epistasis by acquiring residues that provided enhanced stand-alone activity. The evaluation of the allosteric regulation pathway operating between subunits indeed indicates that the catalytically relevant COMM domain of *p*/TrpB is not optimally connected to the two active site loops of *p*/TrpA, thus explaining the low activation of *p*/TrpB in the presence of the allosteric binding partner *p*/TrpA. When analyzing the different enzyme variants and enzyme complexes computationally, we found a strong correlation between the activity of each variant and the shift and stabilization of the catalytically competent closed states at the reconstructed conformational landscapes. The shift that is observed in the energetic minima of *p*/TrpB-con upon complex formation can be interpreted as the stabilization of the catalytically competent state of TrpB. In essence, our results can best be explained by the conformation selection and population shift model (Liu & Nussinov, 2016), as *p*/TrpA changes the distribution of different *p*/TrpB-con conformations toward productive ones. This is also supported by the observation that the Res<sub>6</sub> are scattered all over the protein and do not form a coherent network, which would be more in line with an allosteric network model.

## 4 | MATERIALS AND METHODS

### 4.1 | Bacterial strains and chemicals

All proteins that were analyzed in this study were expressed in the *E. coli* strain BL21 Gold (DE3), purchased from Agilent Technologies. All chemicals used in this study were purchased from commercial sources and were of analytical grade or higher.

### 4.2 | Cloning

The genes for *p*/TrpA, *p*/TrpB, and *p*/TrpB-con were purchased from GeneArt (Thermo Fisher Scientific). They were codon-optimized for recombinant gene expression in *E. coli* and equipped with flanking *Bsa*I restriction sites. Then, the genes were cloned into pET21a-*Bsa*I expression vectors (Rohweder et al., 2018) using a coupled digestion/ligation reaction with *Bsa*I and T4-DNA ligase. The resulting constructs allow for IPTG-inducible expression with a C-terminal His<sub>6</sub>-tag. The deduced amino acid sequences of *p*/TrpA, *p*/TrpB, and *p*/TrpB-con are shown in Table S2.

### 4.3 | Gene expression and protein purification

The *E. coli* expression strain BL21 Gold (DE3) was transformed with the expression plasmids coding for

*p*/TrpA, *p*/TrpB, and *p*/TrpB-con, respectively. The cells were grown at 37°C in lysogenic broth (LB) medium supplemented with 150 mg/mL ampicillin to an OD<sub>600</sub> of 0.6. Next, gene expression was induced by the addition of 0.5 mM IPTG, and the cultures were further incubated overnight at 20°C. Cells were then harvested by centrifugation (4000 g, 20 min) and resuspended in 50 mM KP (pH 7.5), 300 mM KCl, and 10 mM imidazole. Afterward, cells were disrupted by sonication (Branson Sonifier W-250D, 60% amplitude, 2.5 min, 2 s pulse, 2 s pause) and cell debris and insoluble aggregates were removed by centrifugation (14,000 g, 45 min). All proteins analyzed in this study were expressed with a C-terminal His<sub>6</sub>-tag. The His<sub>6</sub>-tagged proteins were purified from the supernatant by affinity chromatography using an ÄKTA-purifier system with a HisTrap excel column (CV 5 mL, GE Healthcare) applying a linear imidazole gradient (10–500 mM over 15 CV). The proteins were further purified by preparative size exclusion chromatography using an ÄKTA-purifier system with a Hi Load 16/600 Superdex 75 pg. column (CV 330 mL, GE Healthcare). The proteins were eluted in 50 mM potassium phosphate (pH 7.5) and 300 mM KCl. Fractions containing the purified protein were identified by SDS-PAGE and pooled. In the case of *p*/TrpA, the protein concentration was determined by absorbance spectroscopy at 280 nm (Thermo Fisher Scientific, NanoDrop One) using a molar extinction coefficient of 26,930 (M<sup>-1</sup> cm<sup>-1</sup>) (Wilkins et al., 1999). In the case of *p*/TrpB and *p*/TrpB-con, due to the additional absorption of the PLP cofactor, the protein concentration was determined with the Bradford assay using a commercial Bradford reagent (Bio-Rad, Bradford protein assay). The purified proteins were dripped into liquid nitrogen and stored at –70°C.

### 4.4 | HPLC-based substrate screening

An HPLC-based substrate screening was performed to identify potential additional alpha-amino acid substrates of *p*/TrpB that might be preferred as substrates over L-serine. All enzymatic assays contained 500 μM indole, 100 mM potassium phosphate (pH 7.5), 180 mM KCl, 40 μM PLP, 2 mM of the respective alpha-amino acid, and 5 μM *p*/TrpB. As a reference, standard samples were used, which contained 100 mM potassium phosphate (pH 7.5), 180 mM KCl, and 500 μM indole or 2 mM L-tryptophan, respectively. Further HPLC-based enzyme assays were performed to unravel alternative functions of *p*/TrpB and *p*/TrpA: To identify possible additional nucleophiles that might be preferred over indole, the same assay was performed in the absence of indole and in the presence of 500 μM of the tested compound. Either 10 μM *p*/TrpB, 10 μM *p*/TrpA, or 10 μM *p*/TrpA:*p*/TrpB were used to start the reaction. In a third approach, the reversed reaction starting with

L-tryptophan was assessed, again using the same assay including 10  $\mu\text{M}$  *p*/TrpB but no L-serine and no indole. Following incubation at 30°C and 500 rpm for 60 min, all reactions were stopped by centrifugation using a filter tube with a pore size of 10 kDa to remove any enzymes. The reaction products were subsequently analyzed by reversed-phase HPLC using an Agilent system (1100 series) with an Eclipse XDB-C18 ( $4.6 \times 150$ ) column. The separation was performed at 20°C with a flow rate of 0.25 mL/min using 0.1% formic acid in water as buffer A and 0.1% formic acid in acetonitrile as buffer B (gradient: 5%–100% buffer B).

#### 4.5 | Steady-state kinetic measurements of TrpB enzymes

To determine TrpB activity, the difference in absorbance between indole and L-tryptophan was used ( $\Delta\epsilon_{290} = 1890 \text{ M}^{-1} \text{ cm}^{-1}$ ) (Faeder & Hammes, 1970). Reactions were performed in triplicates at 30°C, and changes in absorbance were monitored using a spectrophotometer (JASCO V-750). The experimental conditions included 50 mM potassium phosphate (pH 7.5), 180 mM KCl, 40  $\mu\text{M}$  PLP, saturating concentrations of one substrate (indole/L-serine) and varying concentrations of the other substrate (L-serine/indole). When a constant baseline absorbance was reached, reactions were initiated by the addition of 0.5–7  $\mu\text{M}$  of *p*/TrpB or *p*/TrpB-con. For the TrpB reactions in the presence of *p*/TrpA, complex formation was induced by the addition of a 1.5-fold molar excess of *p*/TrpA to 0.5–3  $\mu\text{M}$  of *p*/TrpB or *p*/TrpB-con prior to the start of the reaction. Initial velocities ( $v_i$ ) were calculated from the initial linear part of the resulting curve via division by  $\Delta\epsilon_{290}$ . The determined reaction velocities were then normalized to the applied enzyme concentration ( $v_i/E_0$ ) and plotted against the substrate concentration. The Michaelis constant  $K_M$  and the turnover number  $k_{\text{cat}}$  were obtained by fitting the data to the Michaelis–Menten equation using Origin 2022 (© OriginLab Corporation).

#### 4.6 | Determination of apparent $K_d^{\text{app}}$ values by activity titrations

To determine apparent  $K_d^{\text{app}}$  values for the *p*/TrpA: *p*/TrpB and *p*/TrpA: *p*/TrpB-con complexes, activity titrations were performed where TrpB activity was monitored as described above. Reactions were performed in duplicates at 30°C. The experimental conditions included 2  $\mu\text{M}$  *p*/TrpB or *p*/TrpB-con, 50 mM potassium phosphate (pH 7.5), 180 mM KCl, 40  $\mu\text{M}$  PLP, saturating concentrations of both substrates (75 mM L-serine, 1 mM indole) and varying concentrations of *p*/TrpA. The initial velocities measured for *p*/TrpB or *p*/TrpB-con in the absence of *p*/TrpA were used for normalization

and subtracted from all data points recorded in the presence of *p*/TrpA. The initial velocities were plotted as a function of the added *p*/TrpA concentration. To determine the apparent  $K_d^{\text{app}}$  the resulting data points were fitted to the data with a hyperbolic saturation curve using Origin 2022 (© OriginLab Corporation).

#### 4.7 | Determination of complex formation by analytical size-exclusion chromatography

Analytical size exclusion chromatography was performed with *p*/TrpA (75  $\mu\text{M}$ ), *p*/TrpB (50  $\mu\text{M}$ ), and *p*/TrpB-con (50  $\mu\text{M}$ ) that were applied individually or as a mixture between *p*/TrpA and *p*/TrpB or *p*/TrpB-con, respectively, to a Superdex200 increase 10/300 GL column (GE Healthcare) operated on an ÄKTAmicro system (GE Healthcare). The column was equilibrated with 50 mM potassium phosphate (pH 7.5), 300 mM KCl, and 75 mM L-serine at 25°C. Protein elution was performed at a flow rate of 0.3 mL/min and was followed by absorbance measurements at 280 nm. Calibration was performed with the Cytiva LMW and HMW calibration kits.

#### 4.8 | Circular dichroism (CD) spectroscopy

To assess the structural integrity of the proteins, far-UV CD spectroscopy was used. Spectra were recorded with a CD spectrometer (J-815, JASCO) between 280 and 190 nm using a quartz cuvette (0.2 mm). Measurements were conducted at 25°C in five replicas. All spectra were corrected for buffer absorption (50 mM KP, pH 7.5; 300 mM KCl) and smoothed using the Savitzky–Golay algorithm (Savitzky & Golay, 1964) (convolution width 7) implemented in the Spectra Analysis software provided by JASCO. The mean molar ellipticity per residue  $\theta_{\text{MRW}}$  ( $\text{deg cm}^2 \text{ dmol}^{-1}$ ) was calculated from the observed ellipticity  $\theta_{\text{obs}}$  (mdeg), the pathlength of the cuvette  $d$  (cm), the protein concentration  $c$  ( $\mu\text{M}$ ), and the number of residues  $N_A$ , according to the following equation:

$$\theta_{\text{MRW}} = \frac{\theta_{\text{obs}} \times 10^5}{c \times d \times N_A}.$$

#### 4.9 | Analysis of the genomic neighborhood of *p*/TrpA and *p*/TrpB

Genome neighborhood diagrams were retrieved from the Enzyme-Function-Initiative Genome-Neighborhood-Tool (EFI-GNT) (Oberg et al., 2023;

Zallot et al., 2019) using the amino acid sequence of either *pTrpA* or *pTrpB* as input. The genes flanking *pTrpA* and *pTrpB* were analyzed regarding their functional annotation.

#### 4.10 | Multiple sequence alignment (MSA) and identification of *pTrpB*

All 6373 sequences that are annotated as TrpB1 enzymes in the KEGG (Kanehisa & Goto, 2000) database were retrieved and used to generate a multiple sequence alignment with MAFFT (Kato & Standley, 2013). To identify proteins which contain the LBCA-Res<sub>6</sub>, the MSA was iteratively filtered by discarding any proteins which did not contain one of the six residues. This shrunk the dataset to only one protein variant that contained all six LBCA-Res<sub>6</sub> residues. For mutational planning of the Res<sub>6</sub> residues in *pTrpB*, we analyzed the MSA in JalView (Waterhouse et al., 2009) and exchanged the LBCA-Res<sub>6</sub> to the consensus at each position (E46A, E52D, T63S, S197T, A289G, M290N) giving rise to *pTrpB*-con. In the case of E46, the Res<sub>6</sub> residue of *pTrpB* wild-type represented the most commonly found amino acid. The corresponding position was occupied by an Ala in Anc3TrpB, which is highly dependent on TrpA. Therefore, and because Ala is the second most commonly found amino acid at this position, we replaced Glu with Ala.

#### 4.11 | Molecular modeling system preparation

The starting structures for the two systems (the *pTrpA*:*pTrpB* and *pTrpA*:*pTrpB*-con) were generated with the multimer version of the AlphaFold2 (AF2) (Jumper et al., 2021) neural network. The AF2 predicted structures are available at <https://doi.org/10.5281/zenodo.14872182>. The AF2 models simulated had a predicted LDDT-C $\alpha$  score (pLDDT) higher than 92. To generate the TrpB homodimer enzyme, the TrpA subunits of the predicted structures were removed. The aminoacrylate (AA) Schiff base and quinonoid (QQ2) intermediates were placed in the TrpB subunits through superposition to the external aldimine 2 (Aex2) intermediate of the engineered TrpB crystal structure with PDB accession code 6AM8. Also, to avoid clashes with the QQ2 intermediate, the  $\chi_2$ ,  $\chi_3$ , and  $\chi_4$  torsion angles of the catalytic lysine (i.e., Lys84) were switched to the ones in the 6AM8 x-ray structure.

The water molecules added to each homodimer were selected from the DBSCAN clusterization (Ester et al., 1996; Jukić et al., 2017) algorithm implemented in the scikit-learn Python library (Pedregosa et al., 2011), of different x-ray TrpB monomers. Additionally, three conserved sodium ions in the crystal

structures were added to all structures located at the dimer interface and in each monomer close to the active site.

The MD parameters for QQ2 and AA intermediates were generated with the antechamber and parmchk2 modules of AMBER22 (Case et al., 2022) using the second generation of the general amber force field (GAFF2) (Case et al., 2022; Wang et al., 2004). The intermediates were optimized at the B3LYP/6-31G(d) level of theory, including Grimme's dispersion correction with Becke–Johnson Damping (D3-BJ) and the polarizable conductor model (PCM) (dichloromethane,  $\epsilon = 8.9$ ) as an estimation of the dielectric permittivity in the enzyme active site (Schutz & Warshel, 2001). The partial charges (RESP model) (Bayly et al., 1993) were set to fit the electrostatic potential generated at the HF/6-31G(d) level of theory. The charges were calculated according to the Merz–Singh–Kollman (Besler et al., 1990; Singh & Kollman, 1984) scheme using the Gaussian16 software package (Frisch et al., 2016). The protonation states were predicted using PROPKA (Olsson et al., 2011; Søndergaard et al., 2011). However, the protonation state of the catalytic residue Lys84 was neutral (i.e., LYN84), as is described in the mechanism at the QQ2 intermediate. The enzyme structures were solvated in a pre-equilibrated system using the OPC water model and neutralized by the addition of explicit counterions (i.e., Na<sup>+</sup>) using the AMBER22 leap module. All MD simulations were performed using the ff19SB force field.

#### 4.12 | MD simulation details

MD equilibration phase was done following the protocol described by Roe and Brooks with small differences fine-tuned to our systems (Roe & Brooks, 2020). The bonds involving hydrogen are constrained by the SHAKE algorithm during the non-minimization steps. Long-range electrostatic effects were modeled using the particle mesh-Ewald method (Darden et al., 1993). For Lennard-Jones and electrostatic interactions, a 10 Å cut-off was applied. The MD protocol starts with the minimization phase of 1500 steps of the steepest descent method followed by 3500 steps of the conjugate gradient method with a positional restraint (i.e., a force constant of 5.0 kcal mol<sup>-1</sup> Å<sup>-2</sup>) to the protein heavy atoms. In the following heating phase, a temperature increment from 25 to 300 K during 20 ps of MD simulation time, a Langevin thermostat with a collision frequency of 5 ps<sup>-1</sup>, and a positional restraint (i.e., a force constant of 5.0 kcal mol<sup>-1</sup> Å<sup>-2</sup>) to the protein heavy atoms is performed. A minimization and heating of all atoms in the system is the following step. This starts with two minimization stages of 1000 steps of the steepest descent method followed by 1500 steps of the conjugate gradient method, each with a



positional restraint (i.e., force constant of  $2.0 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  in the first minimization and  $0.1 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  in the second) to the protein heavy atoms. Following this, a third minimization phase of 1500 steps of the steepest descent method followed by 3500 steps of the conjugate gradient method without any positional restraint is performed. The system is then heated in accordance with the previously established procedure. Finally, a five-round equilibration phase at the NPT ensemble with a constant pressure of 1 atm is performed. The first four rounds were done with the Berendsen barostat, whereas the fifth one was done with a Monte Carlo barostat. For all equilibration rounds, a Langevin thermostat with a collision frequency of  $1 \text{ ps}^{-1}$  was used. A positional restraint to the protein heavy atoms with a force constant of 1.0 and  $0.5 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  was applied to the first and second equilibration rounds, respectively. In the third round of 10 ps equilibration, a positional restraint to the backbone heavy atoms with a force constant of  $0.5 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  was used. The fourth and fifth equilibration rounds of 10 ps and 1 ns, respectively, were performed without any restraint. The production runs were performed at the NVT ensemble with the Langevin thermostat with a collision frequency of  $1 \text{ ps}^{-1}$  during 50 ns for all systems. A total of 20 replicas of equilibration and production runs were performed, reaching a total simulation time of  $1 \mu\text{s}/\text{system}$  ( $20 \text{ replicas} \times 50 \text{ ns}$ ) for systems *p*/TrpB, *p*/TrpA: *p*/TrpB, *p*/TrpB-con, and *p*/TrpA:*p*/TrpB-con, with QQ2 or AA intermediates, leading to a final number of 8 systems. The MD trajectories were analyzed using the Python packages MDTraj (McGibbon et al., 2015), pytraj (Nguyen et al., 2016) which is part of the cpptraj package (Roe & Brooks, 2020), MDAnalysis (Gowers et al., 2016), and PyEMMA (Scherer et al., 2015).

## AUTHOR CONTRIBUTIONS

**Thomas Kinatader:** Conceptualization; writing – original draft; methodology; investigation. **Lukas Drexler:** Conceptualization; investigation; writing – original draft; methodology. **Cristina Duran:** Investigation; methodology; writing – review and editing; software. **Silvia Osuna:** Conceptualization; funding acquisition; supervision; writing – review and editing; software; project administration. **Reinhard Sterner:** Conceptualization; funding acquisition; writing – review and editing; supervision; project administration.

## ACKNOWLEDGMENTS

This work was supported by the Generalitat de Catalunya for the consolidated group TCBioSys (SGR 2021 00487), Spanish MICIU for grant projects PID2021-129034NB-I00 and PDC2022-133950-I00. S. O. is grateful for the funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (ERC-2015-

StG-679001, ERC-2022-POC-101112805, ERC-2022-CoG-101088032, and ERC-2023-POC-101158166). C. D. was supported by the Spanish MINECO for a PhD fellowship (PRE2019-089147) and ERC-2023-POC-101158166. L. D. was supported by a PhD fellowship from the Fonds der Chemischen Industrie. The authors are grateful to Dr. Frieda Sorgenfrei for the first filtering of the Res<sub>6</sub> positions in the generated MSA and to Sonja Fuchs and Sabine Laberer for excellent technical assistance. Open Access funding enabled and organized by Projekt DEAL.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.


## ORCID

Thomas Kinatader  <https://orcid.org/0000-0001-5739-044X>

Lukas Drexler  <https://orcid.org/0000-0001-8193-3256>

Cristina Duran  <https://orcid.org/0000-0003-3094-8823>

Silvia Osuna  <https://orcid.org/0000-0003-3657-6469>

Reinhard Sterner  <https://orcid.org/0000-0001-8177-8460>

## REFERENCES

- Anderson KS, Miles EW, Johnson KA. Serine modulates substrate channeling in tryptophan synthase. A novel intersubunit triggering mechanism. *J Biol Chem*. 1991;266:8020–33.
- Banik U, Zhu DM, Chock PB, Miles EW. The tryptophan synthase alpha 2 beta 2 complex: kinetic studies with a mutant enzyme (beta K87T) to provide evidence for allosteric activation by an aminoacrylate intermediate. *Biochemistry*. 1995;34:12704–11.
- Bayly CI, Cieplak P, Cornell W, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem*. 1993;97:10269–80.
- Besler BH, Merz KM, Kollman PA. Atomic charges derived from semi-empirical methods. *J Comput Chem*. 1990;11:431–9.
- Brzović PS, Ngo K, Dunn MF. Allosteric interactions coordinate catalytic activity between successive metabolic enzymes in the tryptophan synthase holoenzyme complex. *Biochemistry*. 1992;31:3831–9.
- Buller AR, Brinkmann-Chen S, Romney DK, Herger M, Murciano-Calles J, Arnold FH. Directed evolution of the tryptophan synthase  $\beta$ -subunit for stand-alone function recapitulates allosteric activation. *Proc Natl Acad Sci U S A*. 2015;112:14599–604.
- Buller AR, Roye P, Cahn JK, Scheele RA, Herger M, Cahn JKB, Arnold FH. Directed evolution mimics allosteric activation by stepwise tuning of the conformational ensemble. *J Am Chem Soc*. 2018;140:7256–66.
- Busch F, Rajendran C, Mayans O, Löffler P, Merkl R, Sterner R. TrpB2 enzymes are O-phospho-L-serine dependent tryptophan synthases. *Biochemistry*. 2014;53:6078–83.
- Casadevall G, Casadevall J, Duran C, Osuna S. The shortest path method (SPM) webserver for computational enzyme design. *Protein Eng Des Sel*. 2024;37:gzae005.
- Casadevall G, Duran C, Estévez-Gay M, Osuna S. Estimating conformational heterogeneity of tryptophan synthase with a template-based AlphaFold2 approach. *Protein Sci*. 2022;31:e4426.

- Case DA, Darden TA, Cheatham TE, Simmerling CL, Wang J, et al. AMBER22. University of California, San Francisco; 2022.
- Cournia Z, Chatzigoulas A. Allosteric in membrane proteins. *Curr Opin Struct Biol*. 2020;62:197–204.
- Darden T, York D, Pedersen L. Particle mesh Ewald: an  $N \log(N)$  method for Ewald sums in large systems. *J Chem Phys*. 1993;98:10089–92.
- Dunn MF. Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase holoenzyme complex. *Arch Biochem Biophys*. 2012;519:154–66.
- Dunn MF, Aguilar V, Brzović P, Drewe WF, Houben KF, Leja CA, et al. The tryptophan synthase holoenzyme complex transfers indole between the  $\alpha$ - and  $\beta$ -sites via a 25–30 Å long tunnel. *Biochemistry*. 1990;29:8598–607.
- Dunn MF, Niks D, Ngo H, Barends TRM, Schlichting I. Tryptophan synthase: the workings of a channeling nanomachine. *Trends Biochem Sci*. 2008;33:254–64.
- Duran C, Kinatader T, Hiefinger C, Sterner R, Osuna S. Altering active-site loop dynamics enhances standalone activity of the tryptophan synthase  $\alpha$  subunit. *ACS Catal*. 2024;14:16986–95.
- Ester M, Kriegel H-P, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. 1996.
- Faeder EJ, Hammes GG. Kinetic studies of tryptophan synthetase. Interaction of substrates with the B subunit. *Biochemistry*. 1970;9:4043–9.
- Fleming JR, Schupfner M, Busch F, Baslé A, Ehrmann A, Sterner R, et al. Evolutionary morphing of tryptophan synthase: functional mechanisms for the enzymatic channeling of indole. *J Mol Biol*. 2018;430:5066–79.
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al. Gaussian 16. Wallingford, CT: Gaussian, Inc.; 2016.
- Goldberg ME, Creighton TE, Baldwin RL, Yanofsky C. Subunit structure of the tryptophan synthetase of *Escherichia coli*. *J Mol Biol*. 1966;21:71–82.
- Gowers R, Linke M, Barnoud J, Reddy T, Melo M, et al. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. In: Proceedings of the 15th Python in Science Conference. SciPy, pp 98–105. 2016.
- Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? *Proteins*. 2004;57:433–43.
- Hatanaka M, White EA, Horibata K, Crawford IP. A study of the catalytic properties of *Escherichia coli* tryptophan synthetase, a two-component enzyme. *Arch Biochem Biophys*. 1962;97:596–606.
- Hettwer S, Sterner R. A novel tryptophan synthase  $\beta$ -subunit from the hyperthermophile *Thermotoga maritima*. Quaternary structure, steady-state kinetics, and putative physiological role. *J Biol Chem*. 2002;277:8194–201.
- Hofmann H. All over or overall - do we understand allostery? *Curr Opin Struct Biol*. 2023;83:102724.
- Hyde CC, Ahmed SA, Padlan EA, Miles EW, Davies DR. Three-dimensional structure of the tryptophan synthase  $\alpha$  2  $\beta$  2 multienzyme complex from *Salmonella typhimurium*. *J Biol Chem*. 1988;263:17857–71.
- Ito S, Yagi K, Sugita Y. Computational analysis on the allostery of tryptophan synthase: relationship between  $\alpha/\beta$ -ligand binding and distal domain closure. *J Phys Chem B*. 2022;126:3300–8.
- Ito S, Yagi K, Sugita Y. Allosteric regulation of  $\beta$ -reaction stage I in tryptophan synthase upon the  $\alpha$ -ligand binding. *J Chem Phys*. 2023;158:115101.
- Jukić M, Konc J, Gobec S, Janežič D. Identification of conserved water sites in protein structures for drug design. *J Chem Inf Model*. 2017;57:3094–103.
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596:583–9.
- Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28:27–30.
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30:772–80.
- Kirschner K, Lane AN, Strasser AW. Reciprocal communication between the lyase and synthase active sites of the tryptophan synthase holoenzyme complex. *Biochemistry*. 1991;30:472–8.
- Kulik V, Hartmann E, Weyand M, Frey M, Gierl A, Niks D, et al. 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. *J Mol Biol*. 2005;352:608–20.
- Lane AN, Kirschner K. Mechanism of the physiological reaction catalyzed by tryptophan synthase from *Escherichia coli*. *Biochemistry*. 1991;30:479–84.
- Lisi GP, Loria JP. Allostery in enzyme catalysis. *Curr Opin Struct Biol*. 2017;47:123–30.
- Liu J, Nussinov R. Allostery: an overview of its history, concepts, methods, and applications. *PLoS Comput Biol*. 2016;12:e1004966.
- Maria-Solano MA, Iglesias-Fernández J, Osuna S. Deciphering the allosterically driven conformational ensemble in tryptophan synthase evolution. *J Am Chem Soc*. 2019;141:13049–56.
- Maria-Solano MA, Kinatader T, Iglesias-Fernández J, Sterner R, Osuna S. In silico identification and experimental validation of distal activity-enhancing mutations in tryptophan synthase. *ACS Catal*. 2021;11:13733–43.
- McGibbon RT, Beauchamp KA, Harrigan MP, Klein C, Swails JM, Hernández CX, et al. MDTraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophys J*. 2015;109:1528–32.
- Merkel R. Modelling the evolution of the archeal tryptophan synthase. *BMC Evol Biol*. 2007;7:59.
- Miles EW. The tryptophan synthase  $\alpha$  2  $\beta$  2 complex. Cleavage of a flexible loop in the  $\alpha$  subunit alters allosteric properties. *J Biol Chem*. 1991;266:10715–8.
- Monod J, Changeux JP, Jacob F. Allosteric proteins and cellular control systems. *J Mol Biol*. 1963;6:306–29.
- Murciano-Calles J, Romney DK, Brinkmann-Chen S, Buller AR, Arnold FH. A panel of TrpB biocatalysts derived from tryptophan synthase through the transfer of mutations that mimic allosteric activation. *Angew Chem Int Ed Eng*. 2016;55:11577–81.
- Nguyen H, Roe DR, Swails J, Case DA. PYTRAJ v1.0.0.dev1: Interactive data analysis for molecular dynamics simulations. Zenodo; 2016.
- Oberg N, Zallot R, Gerlt JA. EFI-EST, EFI-GNT, and EFI-CGFP: enzyme function initiative (EFI) web resource for genomic enzymology tools. *J Mol Biol*. 2023;435:168018.
- Olsson MHM, Søndergaard CR, Rostkowski M, Jensen JH. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J Chem Theory Comput*. 2011;7:525–37.
- Osuna S. The challenge of predicting distal active site mutations in computational enzyme design. *WIREs Comput Mol Sci*. 2021;11:11.
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, et al. Scikit-learn: machine learning in {Python}. *J Mach Learn Res*. 2011;12:2825–30.
- Raboni S, Bettati S, Mozzarelli A. Tryptophan synthase: a mine for enzymologists. *Cell Mol Life Sci*. 2009;66:2391–403.
- Rix G, Williams RL, Hu VJ, Spinner A, Pisera AO, et al. Continuous evolution of user-defined genes at 1 million times the genomic mutation rate. *Science*. 2024;386:eadm9073.
- Roe DR, Brooks BR. A protocol for preparing explicitly solvated systems for stable molecular dynamics simulations. *J Chem Phys*. 2020;153:054123.

- Rohweder B, Semmelmann F, Endres C, Sterner R. Standardized cloning vectors for protein production and generation of large gene libraries in *Escherichia coli*. *BioTechniques*. 2018;64:24–6.
- Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem*. 1964;36:1627–39.
- Scherer MK, Trendelkamp-Schroer B, Paul F, Pérez-Hernández G, Hoffmann M, Plattner N, et al. PyEMMA 2: a software package for estimation, validation, and analysis of Markov models. *J Chem Theory Comput*. 2015;11:5525–42.
- Schulz L, Guo Z, Zarzycki J, Steinchen W, Schuller JM, Heimerl T, et al. Evolution of increased complexity and specificity at the dawn of form I Rubiscos. *Science*. 2022;378:155–60.
- Schupfner M, Straub K, Busch F, Merkl R, Sterner R. Analysis of allosteric communication in a multienzyme complex by ancestral sequence reconstruction. *Proc Natl Acad Sci U S A*. 2020;117:346–54.
- Schutz CN, Warshel A. What are the dielectric “constants” of proteins and how to validate electrostatic models? *Proteins*. 2001;44:400–17.
- Singh UC, Kollman PA. An approach to computing electrostatic charges for molecules. *J Comput Chem*. 1984;5:129–45.
- Søndergaard CR, Olsson MHM, Rostkowski M, Jensen JH. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. *J Chem Theory Comput*. 2011;7:2284–95.
- Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *J Comput Chem*. 2004;25:1157–74.
- Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. 2009;25:1189–91.
- Weinreich DM, Watson RA, Chao L. Perspective: sign epistasis and genetic constraint on evolutionary trajectories. *Evolution*. 2005;59:1165–74.
- Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, et al. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol*. 1999;112:531–52.
- Xie G, Forst C, Bonner C, Jensen RA. Significance of two distinct types of tryptophan synthase beta chain in bacteria, archaea and higher plants. *Genome Biol*. 2002;3:RESEARCH0004.
- Yanofsky C, Crawford IP. 1 Tryptophan Synthetase. *The Enzymes*. Vol 7. Amsterdam: Elsevier; 1972. p. 1–31.
- Zallot R, Oberg N, Gerlt JA. The EFI web resource for genomic enzymology tools: leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. *Biochemistry*. 2019;58:4169–82.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Kinatader T, Drexler L, Duran C, Osuna S, Sterner R. A naturally occurring standalone TrpB enzyme provides insights into allosteric communication within tryptophan synthase. *Protein Science*. 2025; 34(4):e70103. <https://doi.org/10.1002/pro.70103>