UNIVERSITÄT
BAYREUTH

Characterization and engineering of ligand binding in the periplasmic binding protein PotF

DISSERTATION

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
an der Fakultät für Biologie, Chemie und Geowissenschaften
der Universität Bayreuth

vorgelegt von

Pascal Kröger (M. Sc.)

aus Hamburg

Bayreuth 2022

Die vorliegende Arbeit wurde in der Zeit von November 2018 bis September 2022 in Bayreuth am Lehrstuhl Biochemie III unter Betreuung von Frau Professorin Dr. Birte Höcker angefertigt.

Vollständiger Abdruck der von der Fakultät für Biologie, Chemie und Geowissenschaften der Universität Bayreuth genehmigten Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.).

Dissertation eingereicht am: 12.09.2022

Zulassung durch Promotionskommission: 21.09.2022

Wissenschaftliches Kolloquium: 10.02.2023

Amtierender Dekan: Prof. Dr. Benedikt Westermann

Prüfungsausschuss:

Prof. Dr. Birte Höcker (Gutachterin)
Prof. Dr. Claus-D. Kuhn (Gutachter)
Prof. Dr. Stefan Schuster (Vorsitz)
Prof. Dr. Dirk Schüler

1. Table of Contents

1. Table of Contents III
2. List of Abbreviations IV
3. Summary VI
4. Zusammenfassung VIII
5. Introduction 1
   - 5.1 Biomolecular recognition and periplasmic-binding proteins 1
   - 5.2 The periplasmic-binding protein PotF and putrescine uptake in E. coli 3
   - 5.3 Biogenic amines: synthesis, function, and metabolism 4
   - 5.4 Spermidine plays an important role in bacterial pathogenesis 6
   - 5.5 Agmatine: a neuronal biogenic amine and potential PotF ligand 7
   - 5.6 The green fluorescent protein and its application in biosensor design 8
   - 5.7 Fluorescent biosensors 10
6. Synopsis 14
7. Own Contribution 21
   - 7.1 Manuscript 1: A comprehensive binding study illustrates ligand recognition in the periplasmic binding protein PotF 21
   - 7.2 Manuscript 2: Fine-tuning spermidine binding modes in the putrescine binding protein PotF 21
   - 7.3 Draft for manuscript 3: A fluorescent biosensor for the visualization of Agmatine 21
8. Manuscript 1 23
9. Manuscript 2 51
10. Draft for Manuscript 3 74
11. List of Publications 106
12. References 107
13. Acknowledgements 114
14. Eidesstattliche Versicherung und Erklärung 115

2. List of Abbreviations

5-HT	5-hydroxy-tryptamine
A. victoria	*Aequorea victoria*
AAV	adeno-associated virus
ABC	ATP binding cassette
ATP	adenosine triphosphate
Abox	aromatic box
ADC	arginine decarbocylase
AGM	agmatine
CDV	cadaverine
CFP	cyan fluorescent protein
cp	circularly permuted
ΔF	change in fluorescence
Dist	distal
DNA	deoxyribonucleic acid
E. coli	*Escherichia coli*
F_0	basic fluorescence
Fld	flavodoxin
FP	fluorescent protein
FRET	Förster resonance energy transfer
GABA	γ-aminobutyric acid
GFP	green fluorescent protein
HEK	human embryonic kidney
HSQC	heteronuclear single quantum coherence
ITC	isothermal titration calorimetry
LAO-BP	lysine, arginine, and ornithine binding protein
max.	maximal
MBP	maltose binding protein
MD	molecular dynamic
NMDAR	N-methyl-D-aspartate receptor
NMR	nuclear magnetic resonance
ODC	ornithine decarboxylase
Orn	ornithine
P. aeroguinosa	*Pseudomonas aeroguinosa*
PBP	periplasmic binding protein
Prox	proximal
PUT	putrescine
PuuP	putrescine uptake protein
RBP	ribose binding protein
S. flexneri	*Shigella flexneri*
S. pneumoniae	*Streptococcus pneumoniae*
SAMamine	s-adenosylmethioninamine
SAT1	spermidine/spermine N^1-acetyltransferase 1
sf	superfolder

SPD	spermidine
SPM	spermine
T3SS	type III secretion systems
YFP	yellow fluorescent protein

3. Summary

Periplasmic binding proteins (PBPs) are a highly versatile superfamily of proteins that cover a diverse spectrum of ligands in bacteria and archaea. They display a “Venus-flytrap” architecture which consists of two β/α-lobes connected by a hinge region with the ligand binding site located at the interface of the two lobes. PBPs are predominantly open in solution and undergo a large conformational change (closure) upon ligand binding. PotF is *E. coli's* putrescine (PUT) PBP and recruits its ligand to the PotFGHI ABC-transporter system to facilitate uptake. In this thesis apo PotF as well as its binding mechanism and capabilities for other metabolically relevant biogenic amines has been investigated thermodynamically, *in silico,* and structurally. Results showed how PotF promiscuously binds agmatine (AGM), cadaverine (CDV), spermidine (SPD) and spermine (SPM) in a nanomolar to high micromolar range. SPD and AGM raised particular interest for further investigations.

SPD uptake in *E. coli* is facilitated by PotD, a homolog of PotF that in its binding pocket displays seven differences. Prior to this work a PotF variant carrying these seven mutations was generated (PotF/D) and demonstrated a switch in polyamine preference from PUT to SPD. Unfortunately, PotF/D did not display nanomolar affinity for SPD, but simply conserved the already established mediocre binding from PotF wild type in PotF/D. Additionally, the crystal structure of PotF/D displayed a slightly more open conformation after binding SPD. This was only once previously observed for PotF binding with the large ligand SPM, which seems to interfere with complete closure. In this work the sequence space for the seven residue differences was combinatorically analyzed in terms of polyamine preference for SPD and PUT in several PotF constructs. Furthermore, specific residues that were deemed important for closure before, and their surrounding were targeted by mutagenesis to enable a full conformational change in PotF/D as complete closure can improve ligand binding. Through systematic analysis of all the different variants, effects on PUT and SPD polyamine preference in PotF can be pinpointed to specific binding pocket residues. Moreover, it was possible to improve the SPD affinity of PotF/D and to generate a version of PotF, which has its original polyamine specificity (nanomolar for PUT and micromolar for SPD) reversed. This can have implications for medical research as SPD uptake is linked to pathogenesis in multidrug resistant gram-negative bacteria and has been targeted by treatment strategies. Having highly adaptable PBPs like PotF at their disposal might allow pathogens to use the PotFGHI system to enable SPD uptake and bypass treatment strategies.

The ligand AGM on the other hand is a neuromodulator and shows an influence on multiple neurotransmitter systems. Additionally, it is assumed to be a non-selective drug with multiple targets. AGMs specific modes of action and mechanisms are largely unknown and unexplored; therefore, this thesis presents the development of an AGM biosensor by inserting a circularly permuted GFP (cpGFP) into PotF. The gained insight from the PotF wildtype and the PotF/D study were combined to semi-

rationally engineer an AGM specific PotF receptor module as well as a non-binding control variant, which combined with cpGFP resulted in the AGM sensor AGMsen and a control sensor. The non-binding control sensor offers a control when setting up experiments. AGMsen allows fluorescence tracking of AGM in a non-invasive manner upon ligand binding inside and displayed on HEK cells, as well as displayed on rat hippocampal neurons. This proves functionality of the PotF-based sensor which paves the way for further engineering and is a first step towards improving the understanding of AGMs function in the brain.

4. Zusammenfassung

Die periplasmatischen Bindeproteine (PBPs) sind eine vielfältige Superfamilie von Proteinen, die ein breites Spektrum an Liganden in Bakterien und Archaeen binden. Sie bestehen aus zwei β/α-Domänen, die durch eine Scharnierregion verbunden sind. Diese Konstellation wird auch als "Venus-Fliegenfallen"-Architektur bezeichnet, wobei sich die Ligandenbindungsstelle zwischen den Grenzflächen der beiden Domänen befindet. In Lösung liegen PBPs vorwiegend in einem offenen Zustand vor und vollziehen bei Ligandenbindung eine große Konformationsänderung (Schließung). PotF ist *E. colis* Putrescin (PUT) PBP und rekrutiert seinen Liganden an das PotFGHI ABC-Transportersystem, um so die Aufnahme von PUT ins Zellinnere zu ermöglichen. In der hier vorgelegten Promotionsarbeit wurden apo PotF und sein Bindungsmechanismus für andere metabolisch relevante biogene Amine thermodynamisch, *in silico* und strukturell untersucht. Die Ergebnisse zeigen, dass PotF neben PUT auch Agmatin (AGM), Cadaverin (CDV), Spermidin (SPD) und Spermin (SPM) mit nanomolaren bis hoch mikromolare Affinitäten bindet. SPD und AGM sind hierbei von besonderem Interesse für weitere Untersuchungen.

Die Aufnahme von SPD erfolgt in *E. coli* durch PotD, ein homologes Protein zu PotF, welches in seiner Bindungstasche sieben Unterschiede aufweist. Vor dieser Arbeit wurde eine PotF-Variante mit diesen sieben Mutationen erzeugt (PotF/D), die eine Veränderung der Polyaminpräferenz von PUT zu SPD zeigte. Leider weist PotF/D keine nanomolare Affinität für SPD auf, sondern konserviert lediglich die bereits vorhandene SPD-Affinität vom PotF-Wildtyp. Des Weiteren zeigt die Kristallstruktur von PotF/D nach Bindung von SPD eine halbgeschlossene Konformation, die zuvor nur einmal bei der Bindung von PotF mit SPM beobachtet wurde, einem Liganden, der auf Grund seiner Größe die vollständige Schließung zu stören scheint. In dieser Arbeit wurde der Sequenzraum für die sieben Positionsunterschiede in Bezug auf die Polyaminpräferenz für SPD und PUT in mehreren PotF-Konstrukten kombinatorisch analysiert. Darüber hinaus wurden spezifische Reste, die zuvor als wichtig für die Schließung identifiziert wurden, sowie ihre Umgebung durch Mutagenese gezielt verändert, um eine vollständige Schließung in PotF/D zu ermöglichen und so die Ligandenaffinität zu verbessern. Durch eine systematische Analyse der verschiedenen Varianten konnten die Auswirkungen auf die PUT- und SPD-Polyaminpräferenz in PotF auf spezifische Reste in der Bindetasche zurückgeführt werden. Darüber hinaus war es möglich, die Affinität von PotF/D zu SPD zu verbessern und eine Variante von PotF zu erzeugen, bei der die ursprüngliche Polyaminpräferenz (nanomolar für PUT und mikromolar für SPD) umgekehrt ist. Dies sind interessante Aspekte für die medizinische Forschung, da die SPD-Aufnahme mit der Pathogenese multiresistenter gramnegativer Bakterien in Verbindung gebracht wird und Gegenstand von Behandlungsstrategien gegen diese ist. Pathogene, die über

hochgradig anpassungsfähige PBPs wie PotF verfügen, könnten das PotFGHI-System nutzen, um SPD-Aufnahme zu ermöglichen und so Behandlungsstrategien umgehen.

Im Gegensatz zu SPD ist AGM ein Neuromodulator und beeinflusst verschiedene Neurotransmittersysteme. Zudem wird angenommen, dass AGM ein nicht-selektiver Wirkstoff mit mehreren molekularen Zielen ist, der Behandlungsstrategien unterstützen kann. AGMs spezifische Wirkungsweise und Mechanismen sind weitgehend unbekannt und unerforscht. Daher geht es in dieser Promotionsarbeit auch um die Entwicklung eines AGM-Biosensors der durch Einbringung eines zirkulär permutierten GFP (cpGFP) in PotF konstruiert wurde. Die gewonnenen Erkenntnisse aus der PotF-Wildtyp und der PotF/D-Studie wurden kombiniert, um ein AGM-spezifisches PotF-Rezeptormodul sowie eine nicht-bindende Kontrollvariante zu entwickeln. Die Kombination dieser mit cpGFP resultierten im AGM Sensor AGMsen und einem Kontrollsensor. Der Kontrollsensor kann in Experimenten als Negativkontrolle fungieren und AGMsen ermöglicht die nicht-invasive Fluoreszenzverfolgung von AGM nach Ligandenbindung in und auf HEK-Zellen sowie auf hippocampalen Neuronen der Ratte. Dies bestätigt die Funktionalität des PotF-basierten Sensors, der somit einen ersten Schritt zum besseren Verständnis der Funktion von AGM im neuronalen System darstellt.

5. Introduction

5.1 Biomolecular recognition and periplasmic-binding proteins

Molecular interactions are tightly regulated and important in all aspects of life. They range from being specifically tailored towards unique partners to being very promiscuous. Highly specific receptor interactions are invaluable for crucial processes in life i.e., specific neuronal transmissions in the brain, whereby promiscuity is a feature that allows for the binding of chemically and structurally similar molecules (Schreiber and Keating, 2011). This enables quick adaptability and reduces the number of needed recognition systems to similar stimuli.

In gram-negative bacteria and archaea, the periplasmic-binding proteins (PBP) are the main contributor to recognizing different types of solutes in the periplasm e.g., carbohydrates, amino acids, vitamins, and ions (Borrok et al., 2009). Interestingly, gram-positive bacteria have homologous extra-cytoplasmic lipoproteins which take up the same role as PBPs (Tam and Saier, 1993). All these proteins work hand in hand with ABC-transporter systems: they bind small molecules and escort them to the transporter (Tam and Saier, 1993; Moussatova et al., 2008), where they can be taken up into cells and used for further downstream processes. These multi disperse processes range from metabolism to chemotaxis and even quorum sensing (Quiocho and Ledvina, 1996; Chen et al., 2002).

The superfamily of PBPs displays a large sequence diversity concomitant with all the different solutes they recognize. Structurally, PBPs are highly conserved and are divided into class I and class II PBP-like folds. Some famous representatives of class I type PBPs are the ribose binding protein (RBP) and the promiscuous lysine, arginine, and ornithine binding protein (LAO-BP). Nevertheless, the most famous and well-studied example of a PBP belongs to class II and is the maltose binding protein (MBP) which is responsible for the uptake of maltodextrins in *E. coli* (Mächtel et al., 2019). PBP class II displays a slightly more complex structure compared to class I, which is in line with the current evolutionary pathway that is proposed for PBPs. Class II is suggested to have evolved from class I by domain dislocation while class I is proposed to be the result of a gene duplication event of a flavodoxin-like folded protein with an ancestral dimer intermediate (Figure 1; Fukami-Kobayashi et al., 1999).

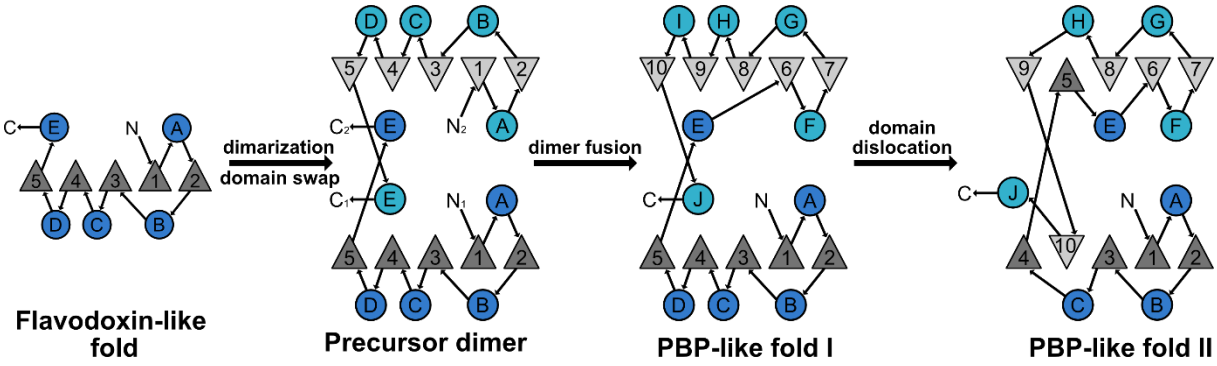


Figure 1: Proposed evolutionary pathway for PBPs. The flavodoxin (Fld)-like fold shows high similarity to the two lobes of PBPs, hence it was hypothesized that PBPs evolved via a gene duplication event of a Fld. The formation of a dimeric Fld-precursor enforced by helix swapping is implicated (E in schematic). Fusion of the duplicated gene copies resulted in PBP-like fold I in which the β-sheets arrange in the order 5, 4, 3, 1,2 in the lower and 10, 9,8,6,7 in the upper lobe, linking the swapped helix from the N-terminal half to β6. PBP-like fold II evolved from fold I by domain swapping, which is best portrayed in the rearrangement of the β-sheets. In fold II the β-sheets are ordered 4,10,3,1,2 in the lower and 9, 5, 8, 6, 7 in the upper lobe with 10 and 5 originating from the opposite lobes. Figure adapted from Fukuma-Kobayashi et al., 1999 and drawn with Affinity Designer (Serif Europe Ltd. 1987).

However, the most notable feature of PBPs is conserved: their large domain motion during ligand binding. The binding of small molecules takes place at the hinge region of the protein which is located between two β/α lobes. In solution PBPs are primarily in an open conformation, and upon ligand encounter the lobes close, burying the ligand inside the protein. This motion sequence combined with the overall structure is often described as a venus-flytrap architecture because of the clear similarities of ligand binding in PBPs and insect catching in the carnivorous plant *Dionaea muscipula*.

The specific mechanism of PBPs changing from open to closed conformation has been studied extensively. NMR experiments show that apo-MBP undergoes a rapid exchange between a major open (95%) and a minor closed form (5%) on a micro-to nanosecond timescale in solution (Tang et al., 2007). This raises the questions how ligand binding in PBPs can be modelled best. Bucher et al. 2011 used accelerated molecular dynamics simulations to analyze conformational selection, induced fit and a mixed model. Conformational selection is based on the presence of both conformations (open and closed) of the PBP in solution and the ligand choosing the closed as its preferred state. This was ruled out in favor of the induced fit model which is in line with the prior hypothesis where ligand encounter occurs in the open state leading to closure of the PBP. Interestingly, a mixed model based on a previously sampled semi-closed conformation of MBP (Bucher et al., 2011a) also seems plausible. In this model the protein would exchange between an open and a semi-closed form, with the latter being favored for ligand encounter, followed by a complete closure like in the induced fit model (Bucher et al., 2011b).

The large conformational change of the PBP structure allows for the formation of new interaction surfaces and is the driving force for ABC-transporter interaction (Hollenstein et al., 2007). The two distinct states enable exposure of the binding pocket to two highly different environments. Ligands

can be positioned at the solvent-excluded, low dielectric protein interior (closed) while still permitting binding site residues to be placed at the evolving sites of the protein surfaces (open), which could explain the adaptability and coverage of various ligands by PBPs (Dwyer and Hellinga, 2004). In addition to all of this, PBPs share an evolutionary relationship with the ligand binding domains of ionotropic glutamate and GABA (γ-aminobutyric acid) receptors (Felder et al., 1999; Lee et al., 2019) which makes them an interesting target for neuroscientific research as well.

5.2 The periplasmic-binding protein PotF and putrescine uptake in E. coli

A special representative of class II PBPs and the focus of this doctoral thesis is *Escherichia coli's (E. coli)* putrescine (PUT) binding protein PotF. It is 370 amino acids in length which corresponds to a size of around 41 kDa with the first 26 N-terminal residues making up the periplasmic export signal peptide. PotF constitutes the first element in the multicomponent PotFGHI uptake system for the cationic polyamine PUT. The operon for this transport system in *E. coli* was first mapped in 1993 (Pistocchi et al., 1993) and further analysis identified a nucleotide binding sequence in the membrane associated PotG and membrane spanning segments in PotH and PotI. Together these proteins constitute the PUT transporter, with PotH and PotI spanning the inner membrane and PotG working as the ATP-binding cassette (ABC; Figure 2C). The first crystal structure of PotF was solved by Vassylyev et al. in 1998, confirming the periplasmic binding protein architecture (Figure 2A). PotF displays an ellipsoid shape consisting of two globular lobes each composed of a five-stranded central β-sheet and six α-helices (bottom lobe, N-terminus) or eight α-helices (top lobe, C-terminus). In the center of the molecule these two lobes are linked by two antiparallel β-strands, referred to as hinge region (Vassylyev et al., 1998) with the PUT binding site being located at the interface of the two lobes.

Prior to PotF, its homologous protein PotD has been structurally characterized (Sugiyama et al., 1996). As is common for PBPs, PotD shows a low sequence identity (35%; Sugiyama et al., 1996) but a high structural similarity (Figure 2B) to PotF, which even made it possible to graft the binding pocket of PotD onto PotF (Scheib et al., 2014). Moreover, PotD preferably binds the polyamine spermidine (SPD), which is PUT extended by one propylamine moiety. Similarly, to PotF, PotD is the first component of the SPD uptake system PotABCD, with PotA being a homologue to the ABC PotG and PotH & I homologues to PotB & C, respectively with sequence identities ranging from 36% - 42%. Furthermore, it has been shown that PotD can bind PUT (Kashiwagi et al., 1993) and thus the PotABCD system can facilitate PUT uptake as well (Kashiwagi et al., 1991). Additionally, another PUT uptake protein PuuP has been reported (Terui et al., 2014). This protein imports PUT for its utilization as an energy source in the absence of glucose while the PotFGHI system is more involved in maintaining optimal concentrations of PUT for cell growth in the presence of glucose as an energy source.

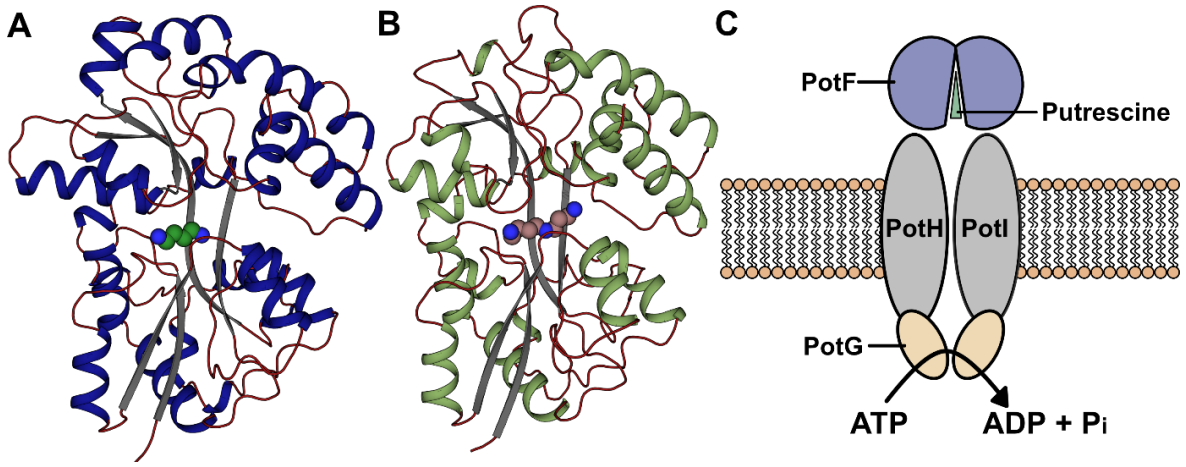


Figure 2: Crystal structure of PotF (A) and PotD (B) as well as the schematic representation of the PotFGHI PUT-uptake system in *E. coli*. PotF (**A**; PDB: 1A99) and PotD (**B**; PBD: 1POT) display a highly similar crystal structure, especially, the lower lobes do not deviate from each other besides in some loop regions. Both proteins are displayed as cartoon, β-sheets are shown in gray, loops in red and α-helices in blue (PotF) or green (PotD). The wildtype ligands PUT and SPD for both proteins are shown in the binding pocket and are represented as spheres in green and salmon, respectively. Protein structures were drawn with PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.). The PotFGHI PUT-uptake machinery (**C**) consists of the PBP PotF, which recognizes PUT in the periplasm and escorts it to the ABC transporter. PotH and PotI span the inner membrane and PotG functions as the ABC. Schematic was drawn with Affinity Designer (Serif Europe Ltd. 1987).

5.3 Biogenic amines: synthesis, function, and metabolism

PUT and SPD are the main polyamines in bacteria and over the years multiple homo- and orthologs to the biogenic amine uptakes systems of PotF and PotD have been reported in a multitude of different bacterial strains (Machius et al., 2007; Brandt et al., 2010). This illustrates the importance of biogenic amines as they play a central role in nucleic acid and protein biosynthesis as well as in proliferation and therefore, cell growth (Tabor and Tabor, 1984; Pegg, 1988; Igarashi and Kashiwagi, 2006). In *E. coli* PUT, cadaverine (CDV) and agmatine (AGM) can be generated by the decarboxylation of the amino acids ornithine (Orn), Lys and Arg, respectively (Figure 3). Besides its direct formation from Orn by the ornithine decarboxylase (ODC, *speC*), PUT can be generated from Arg in two steps. First, Arg is decarboxylated to AGM by arginine decarboxylase (ADC, *speA*) followed by the removal of urea by agmatinase (*speB*), producing PUT (Figure 3).

Two further exceptions are SPD and spermine (SPM), which are synthesized by the SPD synthase (*speE*). The SPD synthase converts PUT into SPD by transferring a propylamine group from S-adenosylmethioninamine (SAMamine) to PUT (Figure 3). This newly formed SPD can be further used as a substrate to which the transfer of another propylamine group from a SAMamine can occur to form SPM (Bowman et al., 1973). All these polycations can function as neutralizing charges on nucleic acids; they stabilize membranes or stimulate RNA-polymerases and some ribosomal proteins (Abraham, 1968; Tabor and Tabor, 1985; Huang et al., 1990). Furthermore, intracellular osmotic and ionic strength can be regulated and stabilized in accordance to changing conditions by excreting PUT rapidly from the cell (Munro et al., 1972).

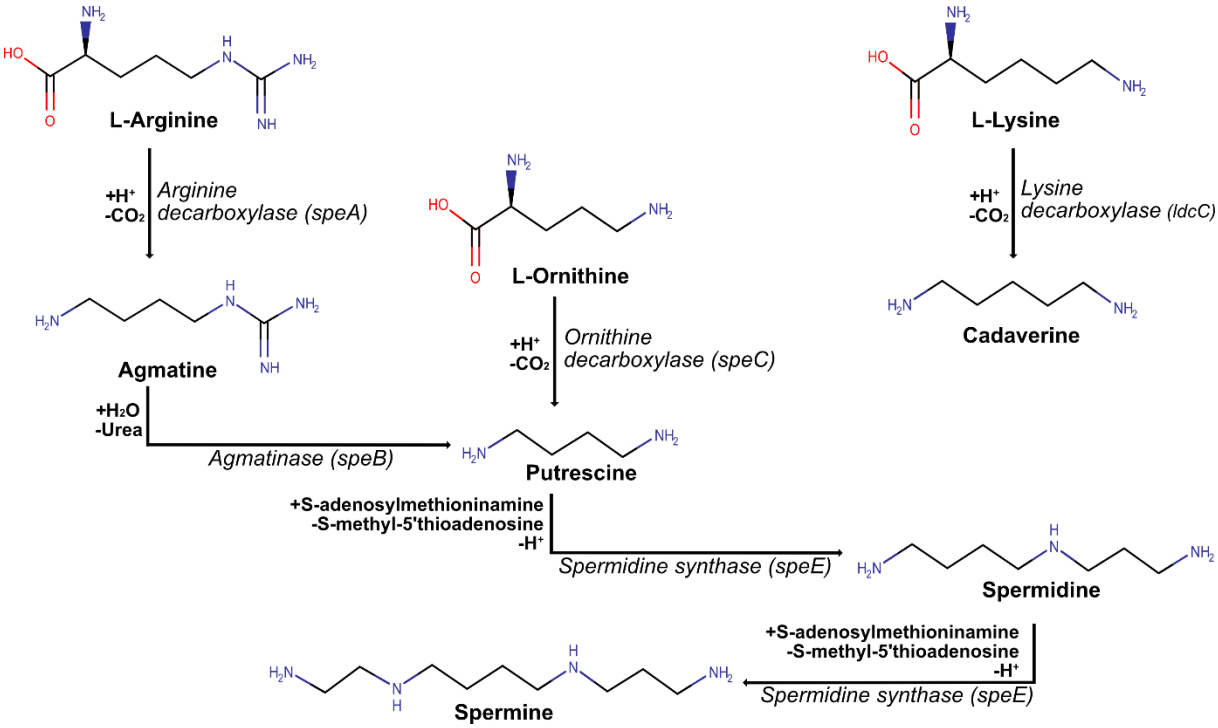


Figure 3: **Biogenic amine metabolism in E. coli.** AGM, PUT and CDV can be generated by the decarboxylation of their precursor amino acids Arg, Orn and Lys, by their respective decarboxylases. PUT can also be formed by agmatinase, which removes urea from AGM. SPD can be formed from PUT by transferring a propylamine group from S-adenosylmethioninamine onto it. In the same fashion spermine (SPM) can be generated from SPD. In *E coli* both reactions are catalyzed by the SPD synthase. MarvinSketch was used for drawing chemical structures, Marvin 19.27.0, 2019 (ChemAxon) and schematic was drawn with Affinity Designer (Serif Europe Ltd. 1987).

Biogenic polyamines do not only play an important role in bacteria but are involved in different processes in eukaryotes as well. The basic premise for function based on the polycationic properties, like nucleic acid or protein interaction, are the same in eukaryotes and hence polyamines influence replication, protein synthesis and regulate cell differentiation or apoptosis (Igarashi and Kashiwagi, 2010). Depletion of intracellular SPD and SPM levels by overexpression of the catabolic enzyme spermidine/spermine N^1-acetyltransferase 1 (SAT1) even leads to an arrest in protein synthesis (translation) and growth in mammalian cells (Mandal et al., 2013). Unfortunately, being involved in all these important regulatory processes makes polyamines a big contributor in tumorigenesis as well (Gerner and Meyskens, 2004): Their presence and metabolism are often up- and dysregulated in cancer cells (Soda, 2011; Casero et al., 2018). Therefore, polyamines have become the target of chemotherapy and chemoprevention (Rial et al., 2009) as well as biomarkers for tumor progression in specific cancer types (Casero et al., 2018).

5.4 Spermidine plays an important role in bacterial pathogenesis

PotF displays a 7.3 µM affinity for SPD, and with the mentioned possibility to graft PotDs binding pocket onto PotF (Scheib et al., 2014), considering SPD as a natural PotF ligand is highly important. Similar results regarding PUT and SPD binding have been observed for respective homologs SpuD (PotF, sequence identity: 57.8%) and SpuE (PotD, sequence identity: 34%) in *Pseudomonas aeroguinosa* (*P. aeroguinosa*; Wu et al., 2012). Wu et al. altered the exclusive SPD binding protein SpuE by exchanging Trp_{271} of the binding pocket to Phe, making space for an important water molecule, which seems to originally promote PUT binding in SpuD. This resulted in a SpuE variant, which exhibits a 1.12 µM affinity for PUT. In addition, the reverse mutation of Phe_{273} to Trp in SpuE lowered the PUT affinity by 250-fold compared to the wildtype. Unfortunately, changes in SPD affinity were not analyzed in this study. Nevertheless, these results show the importance of just one key aromatic binding pocket residue at position 271 (SpuE) and corresponding position 273 (SpuD) for polyamine specificity. These findings indicate that PotF and PotD as well as their homologs can fulfill their respective other's role which can be further enhanced by just single point mutations. This can be of great importance when looking at pathogenic bacteria. PotD was identified as a potential virulence factor in *Streptococcus pneumoniae* (*S. pneumoniae*) since a PotD deficient (*ΔpotD*) mouse-virulent strain showed a clear attenuation in virulence within models of systemic and pulmonary infection independent of the inoculation route or location (Ware et al., 2006). Furthermore, PotD was successfully applied as an immunization against *S. pneumoniae* infection in mice where the immunized animals showed a 91.7% survival rate upon a lethal pneumococcal challenge compared to a 100% mortality rate in the control group (Shah and Swiatlo, 2006). Nonetheless, the study on the *ΔpotD* strain showed wild type like growth upon exogenous PUT or SPD addition, which hints at another polyamine transport system for SPD in *S. pneumoniae* (Ware et al., 2006) and could suggest a possible co-usage of PotF to ensure adequate SPD levels inside bacterial cells.

Moreover, SPD uptake has been linked to the expression of type III secretion systems (T3SS; Zhou et al., 2007). T3SS are used by pathogens to deliver bacterial effector proteins into eukaryotic host cells (Ghosh, 2004) and, hence, play an essential role in in pathogenesis of multidrug-resistant gram-negative bacteria (Felise et al., 2008; Keyser et al., 2008). Zhou et al. showed that mutations in the major SPD uptake system *spuDEFGH* decrease the transcription of most T3SS genes in *P. aeroguinosa* significantly. It is not surprising that therefore, the design of a SpuE antibody to prevent SPD transport attenuates virulence and ultimately weakens *P. aeroguinosa* infections (Zhang et al., 2019). The potential of PotF and its relatives to be able to facilitate SPD besides PUT uptake can make PotF a reasonable target for treatments. In addition, PUT uptake has also been shown to influence virulence gene expression in the intestinal pathogen *Shigella flexneri* (*S. flexneri*; Durand and Björk, 2003).

5.5 Agmatine: a neuronal biogenic amine and potential PotF ligand

Biogenic amines do not only exert influence on different regulatory processes inside cells. One very famous class of biogenic amines are the neurotransmitters e.g., serotonin, histamine and the three catecholamines dopamine, epinephrine (adrenaline), and norepinephrine (noradrenaline). These neurotransmitters are also synthesized from amino acids: Serotonin is based on Trp, histamine on His and the other three originate from Tyr. Another molecule that is synthesized in the brain, stored in synaptic vesicles, accumulated by uptake, and released upon depolarization is AGM (Reis and Regunathan, 2000). Unfortunately, no AGM specific postsynaptic receptor or "agmatinergic" neuronal system has been identified by now, labeling it a neuromodulator and co-transmitter instead of a "classic" neurotransmitter. Still, AGM shows an influence on multiple molecular targets that include neurotransmitter systems like nicotinic, imidazoline I_1 and I_2, α_2-adrenergic, glutamate NMDAR, and serotonin 5-HT2A and 5HT-3 receptors (Piletz et al., 2013). Most common central nervous system disorders have a diverse origin and due to its high occurrence in the peripheral and central nervous system, AGM is hypothesized to be a "magical shotgun", a non-selective drug with multiple targets, that can lead to more effective treatments (Roth et al., 2004; Piletz et al., 2013). AGM has shown antidepressant-like effects (Zomkowski et al., 2002; Chen et al., 2018) and improves cognitive dysfunction in Alzheimer's disease (Song et al., 2014). Furthermore, it offers protection against schizophrenia by influencing the dopaminergic signaling in the brain (Kotagale et al., 2012) and has beneficial effects against multiple sclerosis (Ninkovic et al., 2015).

Compared to the other biogenic amine neurotransmitters, AGM is a precursor to PUT in metabolism (Figure 3). Therefore, their structures are highly similar: they share the same 1-aminobutane core, with AGM carrying a guanidine as a functional group at carbon atom four as opposed to the amino group of PUT (Figure 3). This similarity could allow PotF to bind AGM making it a possible natural ligand among the other described polyamines SPD, CDV and SPM. Moreover, the uniqueness of AGMs functional group could make an attractive target for engineering PotF towards AGM-specific binding. With AGM being a potential therapeutic as well as playing an important role in neuroscience this approach could be especially interesting for biosensor engineering. A fluorescent sensor could allow tracking of AGM in terms of spatiotemporal localization or molecular mechanisms *in vivo* in non-invasive experimental set-ups. Most of these biosensors consist of proteins that are either chemically modified or genetically coupled to a fluorescent reporter (Frommer et al., 2009; Tainaka et al., 2010). Another type of sensor are transcription factors that initiate the biosynthesis of downstream reporter genes upon ligand binding (Binder et al., 2012). Generally, a fluorescent signal is triggered by all these sensor classes after binding their respective target molecule. For PBPs like PotF, the sensor

construction focuses on the drastic conformational closure in the receptor module induced by ligand binding, with green fluorescent protein (GFP) being the most commonly used effector domain.

5.6 The green fluorescent protein and its application in biosensor design

The first discovery of the naturally fluorescent protein GFP dates back to 1962, where it was found in *Aequorea victoria* (*A. victoria*) by Osamu Shimomura (Shimomura et al., 1962; Shimomura, 2005). Its primary sequence was first isolated and sequenced in 1992 (Prasher et al., 1992) and it has been applied for laboratory use i.e., as a marker for gene expression, ever since (Chalfie et al., 1994). The crystal structure of GFP shows that the 238 amino acid long polypeptide folds into a 26.9 kDa β-barrel consisting of 11 strands that surround a central α-helix (Figure 5; Ormö et al., 1996) that includes the fluorophore formed by spontaneous cyclization and oxidation of the sequence Ser_{65}-Tyr_{66}-Gly_{67} (Figure 4).

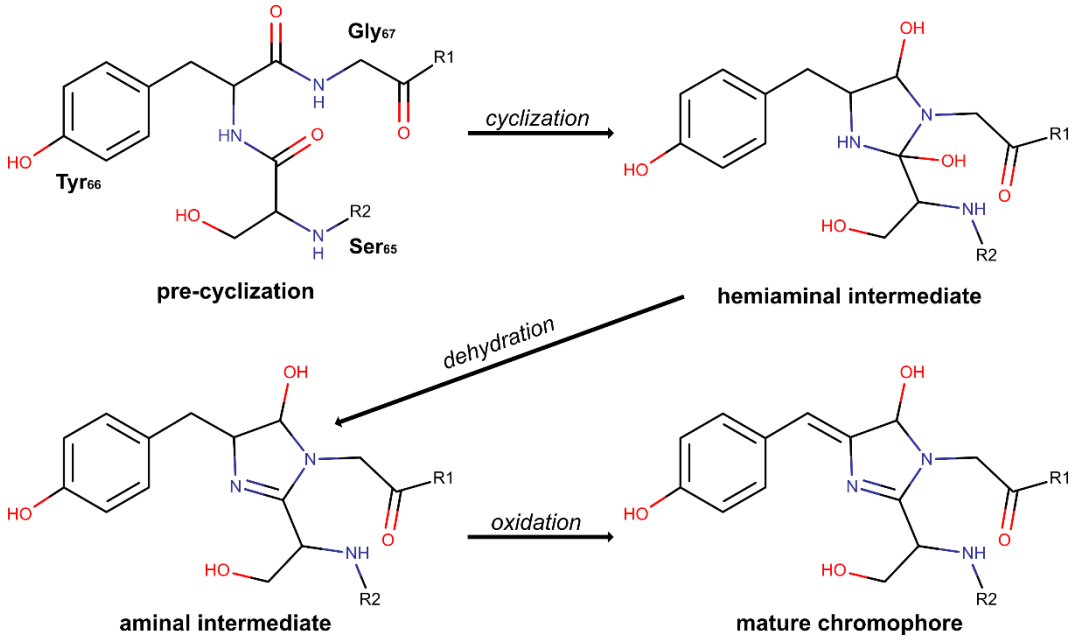


Figure 4: Chromophore formation in GFP. After correct folding of the protein, cyclization occurs due to a nucleophilic attack of the amid electron pair from Gly_{67} to the carbonyl group of Ser_{65}. This is followed by a dehydration and oxidation step to form a conjugated π-system which is linked to the functional phenol group of Tyr_{66}. This extended π-system is the mature chromophore and responsible for the fluorescence of GFP. MarvinSketch was used for drawing chemical structures, Marvin 19.27.0, 2019 (ChemAxon) and schematic was drawn with Affinity Designer (Serif Europe Ltd. 1987).

After folding of the protein, chromophore maturation occurs rapidly with residues Ser_{65} and Gly_{67} condensing into a five membered heterocycle by hemiaminal formation through nucleophilic attack of the amid electron pair from Gly_{67} to the carbonyl group of Ser_{65}. The hemiaminal undergoes dehydration and an aminal is formed in the now monosaturated five-ring. Through oxidation at c_α and c_β of Tyr_{66}, the newly formed double bond of the five ring is linked to the conjugated π-system of the functional phenol group. This extended π-system is also called the *p*-hydroxybenzylidene-imidazolinone chromophore and is responsible for the visible absorbance and fluorescence of GFP (Cubitt et al., 1995). The order in which dehydration and oxidation occur has not been fully clarified

yet, but it has been proposed that under highly aerobic *in vitro* conditions, the dominant reaction path follows a cyclization-oxidation-dehydration mechanism (Zhang et al., 2006) while in the originally proposed mechanism dehydration immediately occurs after ring closure as shown in figure 4 (Cubitt et al., 1995).

The fluorescence properties of GFP have been engineered since the start of its characterization. The first major improvement of GFP was the exchange of Ser_{65} to Thr, leading to enhanced fluorescence and photostability. Furthermore, this mutation shifted the major excitation peak to 488 nm, with the emission peak maintained at 509 nm (Heim et al., 1995) making it more convenient to use in a laboratory set-up than the wild type GFP which displays two excitation peaks at 395 and 475 nm. In addition to S65T, a mutation of Phe_{64} to Leu was discovered, improving folding stability at 37°C and thereby allowing application of GFP in mammalian cell culture (Cormack et al., 1996). Moreover, exchanging one residue adjacent to the chromophore (Thr_{203}) results in significantly red-shifted excitation and emission maxima (Ormö et al., 1996). Therefore, GFP is the progenitor of a whole range of fluorescent proteins (FPs) with different spectroscopic properties and a vast variety of hues of which most can be found in the fluorescent protein data base FPbase (Lambert, 2019). Still, all FPs share the same β-barrel architecture since the entire structure is essential for fluorescence development and maintenance (Remington, 2006; Kremers et al., 2011). In contrast to most other proteins, FPs have many charged residues in their core, which promote the binding of water molecules, locking them into specific conformations inside the protein (Kremers et al., 2011). This is important since the fluorescence is highly dependent on the specific chemical surroundings of the chromophore, with changes to the local environment resulting in massive changes in spectral characteristics, photostability, acid resistance and a variety of other physical properties (Kremers et al., 2011).

Another unique GFP is the circularly permuted version cpGFP (Figure 5; Baird et al., 1999). For circular permutation, the original termini of GFP were connected by a GGTGGS linker while new termini were created between residue positions 144 and 145 in the seventh β-strand of the original barrel. Baird et al. observed that this location within the cpGFP can tolerate the insertion of entire proteins, and conformational changes in the insert can have a strong influence on the fluorescence. They introduced calmodulin and zinc finger domains into a yellow fluorescing variant of cpGFP resulting in constructs that increase their fluorescence upon metal binding. In the case of the calmodulin construct this created a sensor that can monitor cytosolic Ca^{2+} content in mammalian cells (Baird et al., 1999).

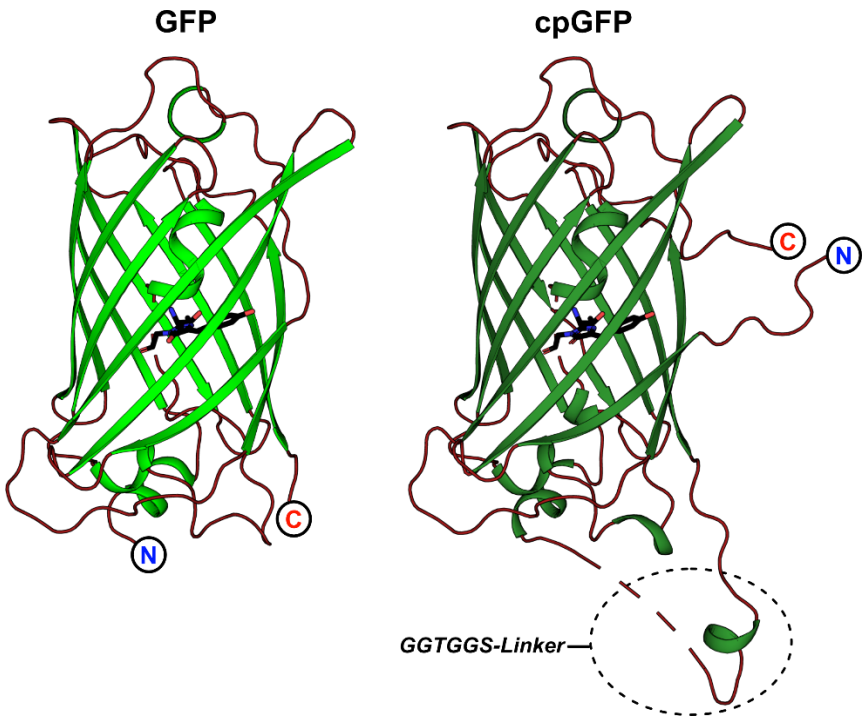


Figure 5: GFP from *A. Victoria* (left, PDB: 1EMA) and cpGFP (right, extracted from PDB: 3OSQ). In GFP N- and C-termini are in close proximity and located at the bottom of the barrel. For permutation these termini were linked by a GGTGGS linker in cpGFP as highlighted at the bottom of the structure. The linker seems to be highly flexible as seen in the dashed lines in the structure which represents missing electron density for five residues. The new termini were introduced at wild type positions 144 and 145 near the chromophore. Both proteins are displayed as cartoon, α-helices and β-sheets are shown in light green (GFP) or dark green (cpGFP), loops are shown in red and the chromophore in black sticks in the middle of the barrel. Protein structures were drawn with PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.) and modifications with Affinity Designer (Serif Europe Ltd. 1987).

To tackle the problem that wild type and other GFP variants used in protein fusion often misfold when expressed, Pédelacq and colleagues engineered superfolder GFP (sfGFP, Pédelacq et al., 2006), which shows improved tolerance of circular permutation, greater resistance to chemical denaturants and improved folding kinetics. Combining the properties of sf- and cpGFP, a superfolder circularly permuted GFP (sfcpGFP) offers perfect attributes to be used as an effector module in PBP-based biosensor engineering. The superfolder mutations will keep the overall construct stable and lead to robust folding, while the circular permutation will offer fluorescence signal sensitivity to the conformational change upon ligand binding of the PBP.

5.7 Fluorescent biosensors

The toolbox of FPs offers a wide range of different hues with unique photophysical properties spanning the whole spectrum of visible light by ranging from near-ultraviolet to near-infrared (Tomosugi et al., 2009; Shcherbakova and Verkhusha, 2013; Rodriguez et al., 2017). This was possible by discovering new variants in species other than *A. victoria* and further expanding them via protein engineering. Fluorescence is a rapid process on a nanosecond scale, all while being spatially precise with an emission wavelength smaller than cellular structures. Therefore, it makes an excellent tool for monitoring cellular processes with a combined temporal and spatial resolution not possible with traditional

biochemical methods (Greenwald et al., 2018). These features of an effector module joined with specific receptors that transmit ligand recognition to the effector leads to the creation of powerful sensing tools which can elevate many research fields to a new level. As stated before, proteins that undergo conformational changes make great receptor molecules because their global motion can be transferred to a FP resulting in a quantifiable signal. This has been used for the creation of two main classes of genetically encoded biosensors, which are either based on Förster resonance energy transfer (FRET) between two FPs or on allosteric modulation of fluorescence from a single FP (Nasu et al., 2021). In the case of FRET sensors, two FPs in which fluorescence spectra overlap, are chosen since the emitted light of the higher energy donor FP (blue shifted) needs to excite the lower energy acceptor FP (red shifted; Ciruela, 2008). This energy transfer is highly dependent on the donor and acceptor dipole moment and distance with an inverse 6^{th}-power law. This is especially important for biosensor engineering since fluorophores must be linked to the conformationally responsive receptor domain in a way that the chromophores get in a proximity of < 10 nm and that their dipoles are arranged in parallel upon ligand binding. The most famous FRET-FP pair are CFP (cyan fluorescent protein) and YFP (yellow fluorescent protein), where CFP gets excited by blue light (430 nm) and emits cyan fluorescence which in return can excite YFP, resulting in emission of yellow fluorescence. Energy transfer efficiency and FRET ratio can then be calculated to draw conclusions about binding and detection events. Recently, a FRET-biosensor for the visualization of the central plant hormone auxin has been engineered that allows the direct und dynamic measurement of auxin in different stages of the development of the plant at subcellular resolution, showing the robustness and potential of this approach for engineering sensors (Herud-Sikimić et al., 2021).

The usage of two FPs also brings disadvantages, the most obvious being the requirement to measure two emission colors, which can limit the opportunities of multicolor detections and multiparameter imaging (Carlson and Campbell, 2009). The FPs in a FRET pair can also show differences in photobleaching rates which can influence data analysis (Piston and Kremers, 2007). Additionally, different types of light are absorbed and scattered differently in tissue. This can influence the results of the observed emission ratios as a function of tissue depth (Boffi et al., 2018; Nasu et al., 2021). These problems are not present in single FP biosensors. Nasu et al., describe them as more user friendly but more difficult to engineer and less appropriate for quantitative imaging, however this does not affect neuroscientific sensing aspects much because of the all-or-none nature of neuronal signaling, which reduces the need for quantitative measurements. Furthermore, with high-throughput screening methods like microfluidics coupled to rapid fluorescence sorting systems being more easily accessible and computational prediction tools getting more powerful each year (Liu and Kuhlman, 2006; Leman et al., 2020; Jumper et al., 2021) the difficulties for engineering become less and less prominent. In recent years, sensors for neuronal compounds have been designed und successfully applied in

different experimental set-ups, thereby tracking some of the most important neurotransmitters e.g., Glu (Marvin et al., 2013), GABA (Marvin et al., 2019), serotonin (Unger et al., 2020), and dopamine (Patriarchi et al., 2018; Sun et al., 2018) among different species. In the case of dopamine, the sensor is based on a G-protein-coupled receptor (GPCR) sensing domain fused to cpGFP, whereas the others are based on PBP sensing domains. The GABA sensor iGABAsnFR can detect fluorescence increases *in vivo* in mice and zebrafish (Marvin et al., 2019), while Glu sensor iGlusnFR has been used for imaging in the mouse retina and *in vivo* Glu signaling in *Caenorhabditis elegans* (*C. elegans*) (Marvin et al., 2013). Furthermore, iGlusnFR has been extensively engineered in terms of affinity, stability, and differently colored emission profiles (Marvin et al., 2018). In addition to neurotransmitters, MBP has already been turned into a maltose sensor in 2011 proving the functionality of a PBP-based biosensor design (Marvin et al., 2011).

Taking the knowledge from all these studies it is conceivable to build a biosensor from PotF in a similar fashion as it has been done for other PBPs before. PotF offers a large variety of possible polyamine ligands that can provide interesting targets among multiple fields of research. PUT and CDV can occur in food as a product of fermentation or putrefaction of tissue (Silla Santos, 1996). As mentioned previously, polyamines are used as biomarkers for tumor progression in specific cancer types (Casero et al., 2018) and PotF could provide the perfect framework for engineering biosensors with desired properties to assist in clinical polyamine research. Finally, AGM a lesser known and more unexplored neuromodulator, makes an interesting biosensor target to solve unanswered questions regarding its neurological mechanisms. The PBPs motion upon closure could be transmitted to cpGFP, thus, altering the surroundings of the chromophore and resulting in a signal change (Figure 6) when binding the desired molecule (Lorimier et al., 2002; Nasu et al., 2021). This is possible because the permutation site in cpGFP is where the chromophore is closest to the β-barrel shell (Nasu et al., 2021) and opening the structure at this position exposes it to a new environment.

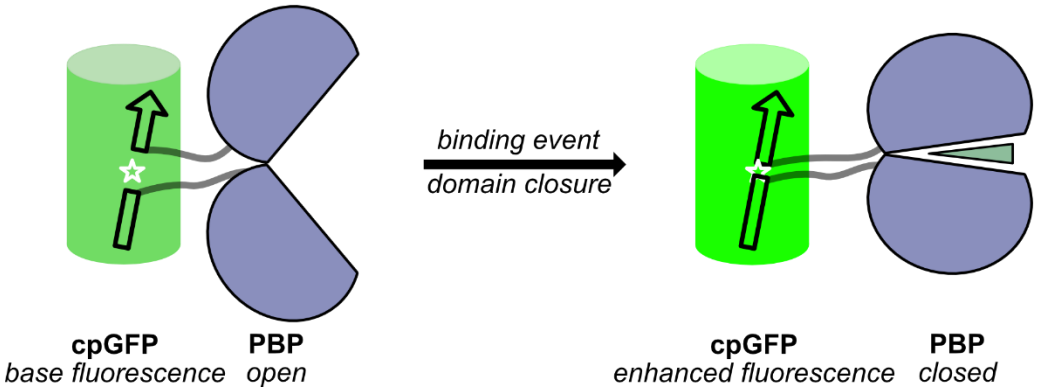


Figure 6: Depiction of a PBP-based cpGFP biosensor. Linking cpGFP to a conformationally responsive region of the PBP (e.g., PotF) can yield a functional biosensor construct. Upon ligand binding and domain closure the conformational change would be transmitted to cpGFP, affecting the chromophore (white star), and ideally result in a gain in fluorescence. Schematic was drawn with Affinity Designer (Serif Europe Ltd. 1987).

The closing of the PBP can then alter proximity between linkers and nearby cpGFP sidechains possibly resulting in a change in water structuring in the opening, leading to a shift in the chromophore protonation equilibrium as observed in the MBP sensor (Marvin et al., 2011). In detail this refers to the protonation of the functional phenol group of Tyr_{66} which is in equilibrium with its deprotonated phenolate. Opening the GFP through permutation and inserting a receptor domain at this specific position exposes the phenol(ate) moiety of Tyr_{66} enabling the pK_a to become more susceptible to its local surroundings (Nasu et al., 2021). In the case of GFP the phenol form absorbs shorter-wavelength light and is less fluorescent and the phenolate form absorbs longer-wavelength light and is typically the brighter fluorescent state which ideally presents the ligand bound state of the sensor (Figure 6).

6. Synopsis

The objective of this doctoral thesis is to gain insights into ligand recognition and discrimination using the putrescine (PUT) periplasmic binding protein (PBP) PotF and its interaction with a range of small molecule ligands as a model system. Furthermore, the aim is to investigate how to switch binding capabilities towards unique ligands in this protein to understand the underlying mechanisms between promiscuity and specificity. This knowledge can be used to improve design and engineering procedures for new receptor proteins as well as explain how these properties of a protein might impact medical research in the future. Lastly, the knowledge was applied to PotF in the framework of a fluorescent biosensor to create a tool that can visualize the neuromodulator agmatine (AGM) *in vivo*.

In manuscript one, we analysed the ligand promiscuity of PotF by isothermal titration calorimetry (ITC) for a subset of different biogenic amines and their respective amino acids. The wild type ligand PUT, AGM, cadaverine (CDV), spermidine (SPD) and spermine (SPM) showed substantial affinity to the protein in the nanomolar to micromolar range (Figure 7).

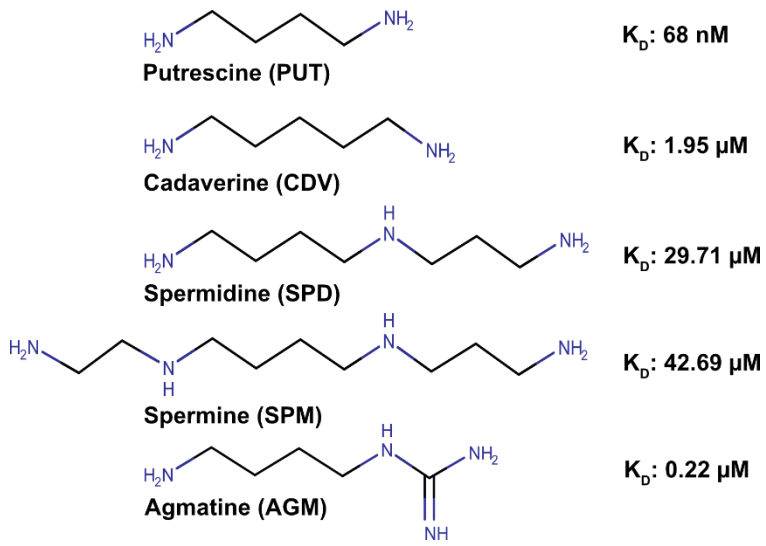


Figure 7: Structure and affinity as measured by ITC for all tested ligands. Affinity decreases with increasing size of the molecules. One exception is AGM, which displays the second highest affinity of the dataset. Further analysis of crystal structures showed that the reason for this is the perfect mimic of the PUT binding mode by AGM. MarvinSketch was used for drawing chemical structures, Marvin 19.27.0, 2019 (ChemAxon) and schematic was created with Affinity Designer (Serif Europe Ltd. 1987).

Generally, the affinity decreased with the size of the ligands from PUT to CDV, to SPD, to SPM; however, AGM showed the second highest affinity, ranking between PUT and CDV. This was surprising since AGM is longer than CDV and possesses a bulkier guanidine end group in comparison to all other ligands with only simple amines. To address this anomaly in our dataset we solved crystal structures of all PotF-ligand complexes to analyse differences in the binding modes of the ligands and identify important interaction residues. We observed that the primary amine binding sites in the pocket are crucial for ligand interactions. All molecules bind to the first, proximal primary amine binding site in the same fashion as PUT. Furthermore, the longer ligands CDV, SPD and SPM bend their methylene

backbone to fit their nitrogen atoms in the place of the secondary, distal primary amine binding site in the pocket, leading to somewhat strained conformations of the ligands. AGM on the other hand was able to perfectly place the intramolecular guanidino nitrogen in the distal amine pocket, with the rest of the guanidine group being additionally coordinated in the distal part of the pocket. This nearly perfect mimic of the PUT binding mode by AGM results in the second highest affinity of our data set. In ITC measurements we also observed a shift in thermodynamics from entropically unfavoured for PUT, CDV and AGM to entropically favoured for SPD and SPM. We suspect additional water release from the pocket upon binding SPD and SPM to be the reason for this shift. Since water molecules seemed to mediate an important role for ligand binding and thermodynamics, we solved two more crystal structures, an open and a closed apo form of PotF. The closed apo form revealed a tightly coordinated water network in the pocket, whose positions could be directly mapped to ligand molecules thereby acting as a blueprint for possible ligand placements. The open apo structure on the other hand illustrated the large motion between open and closed conformation of PBPs. Interestingly, when comparing all the structures we observed a semi-closed conformation for the PotF-SPM complex. Binding of the largest ligand, SPM, prevented complete closure of the protein by disrupting two salt bridges on either side of the binding pocket. To improve our understanding of the conformations PotF can adopt in the presence and absence of ligands we used the solved crystal structures as input for molecular dynamics (MD) simulations. The apo PotF simulations resulted in four main states which increase in order of their opening and exchanged at different rates towards the most populated ones (Figure 8).

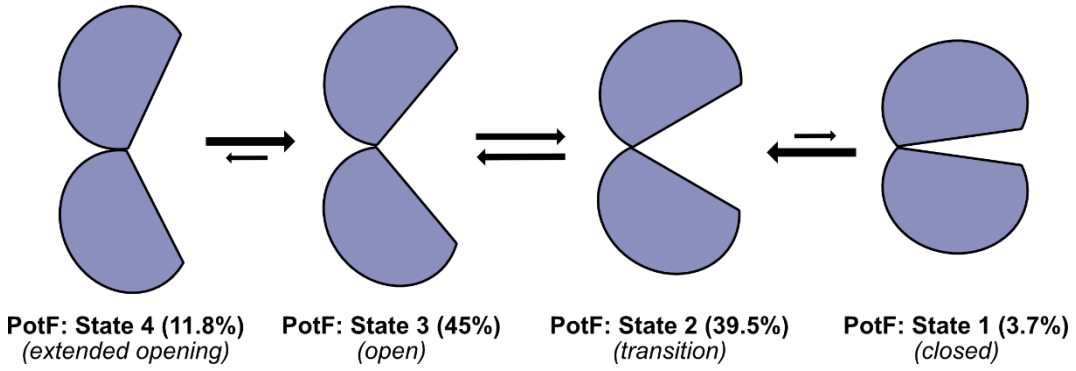


Figure 8: Schematic representation of the simulated behaviour of apo PotF in solution. As expected, the most populated state is state 3, which represents the open conformation. State 3 shows exchange to a transient more open state (4) with the equilibrium being favoured towards state 3. The exchange between state 3 and the slightly more closed and second most populated state 2 is more balanced. The lowest populated state 1 represents the fully closed conformation, which is normally reached by ligand binding. This state also seems unfavoured as it exchanges rapidly to state 2. Schematic was created with Affinity Designer (Serif Europe Ltd. 1987).

The highest populated state resembles the open apo structure the most. The least populated states were a transient more open state (State 4) and the fully closed state (State 1), which resembles the crystal structures in the closed form. The second most populated state (State 2) comprised of conformations between open and closed. Projection of the bound ensembles after ligand addition to

the simulation showed that the PUT, AGM and CDV states reflect the conformation of the crystal structures, whereby SPD and SPM showed slightly more open states. To investigate these discrepancies between the structures and calculations we applied nuclear magnetic resonance (NMR) spectroscopy to analyse the behaviour of PotF in solution. We measured initial ^{1}H-^{15}N HSQC spectra for apo PotF, to use as a reference before the addition of all ligands to saturating concentrations, where we observed clear signal changes. Although peaks moved along the same trajectory, most did not shift to an extent which would mimic PUT binding. CDV and AGM addition were the most similar, again, but SPD addition led to less pronounced shifts while SPM was even closer to the apo spectrum. These observations support the simulations and suggest that PotF adopts slightly different closed states in solution, depending on the type of ligand bound. These nuanced differences could have been overridden by crystal forces that prefer one particular state in the solved x-ray structures, with only large changes being shown, as it is the case for the more open PotF-SPM bound structure.

In the second manuscript we aimed to reveal determinants for polyamine preference for PUT and SPD in PotF. We followed up on a prior study by Scheib et al., where the different residues in the binding pocket of PotD, an SPD binding homolog of PotF, were grafted onto PotF. These seven mutations created PotF/D, a PotF variant that did not bind PUT anymore while conserving the moderate SPD affinity of the wild type. First, we looked at the different mutations and grouped them into proximal (Prox; T38S, E39D & D247S), distal (Dist; S87Y, A182D & L348Q) and aromatic box (Abox; F276W) residues of the pocket (Figure 9).

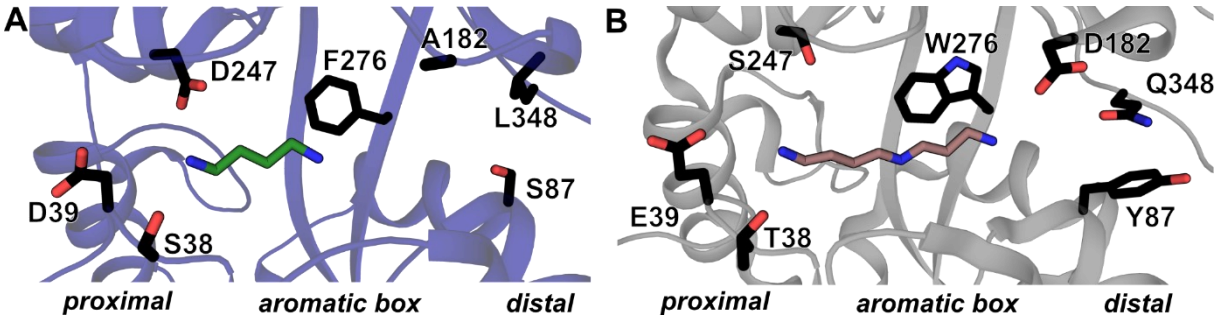


Figure 9: Representation of the binding pocket of PotF in complex with PUT (A; PDB 6YE0) and PotF/D in complex with SPD (B; PDB 7OYZ). The secondary structure of PotF and PotF/D is displayed in blue and grey cartoon, respectively. The ligands PUT (green) and SPD (salmon) as well as targeted residues (**A,** black) and their respective exchanges (**B,** black) are shown as sticks. The description under both binding pockets shows the location of the proximal and distal region as well as the aromatic box. Protein structures were drawn with PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.) and figure was created with Affinity Designer (Serif Europe Ltd. 1987).

We analysed the influence of the three groups and observed a big change in polyamine preference by only mutating Abox. Prox mutations led to a non-binding variant for PUT and SPD, while the Dist cluster displayed a low affinity for SPD only. Furthermore, we combined Abox with Prox and Dist to create Abox_Prox and Abox_Dist. The disadvantageous effect of the Prox mutations also led to no PUT and SPD binding in Abox_Prox, while Abox_Dist showed a 5-fold improvement for selective SPD binding in comparison to PotF/D. We combined all single Dist mutations with Abox, to characterize the specific

effects of the distal mutations. Abox_A182D showed the largest swap of polyamine preference for PUT (decrease 1200-fold) and SPD (increased 10-fold), while the L348Q carrying variant displays the same trend but less pronounced. Abox_S87Y was the only mutation displaying negative effects on both polyamine affinities when compared to the Abox mutation alone.

Another feature observed in PotF/D was a partially open conformation in its crystal structure. We only observed this feature one other time, when analysing SPM binding in PotF in our previous study. SPM seemed to interfere with closure by disrupting important salt bridges (Asp_{39}-Arg_{254} & Arg_{91}-Glu_{184}) flanking the binding pocket. Therefore, we closely analysed these salt bridges and their surroundings in PotF/D and tried to fix the closing mechanism. In the initial PotF/D Asp_{39} was exchanged to Glu, this still allowed for salt bridge formation, but we reverted this to the original Asp to introduce possible wild type-like closure. Furthermore, the distal mutation S87Y showed a flipped side chain outside the pocket and disrupted Arg_{91}-Glu_{184}. Consequently, we changed residues in the proximity to S87Y to allow it to flip back inside the pocket, in order to stay true to the initial design and enable PotD-like SPD coordination. We also reverted S87Y to allow salt bridge formation and introduced a Tyr in a nearby position to create a distal Tyr for SPD coordination. All constructs improved SPD affinity in PotF/D but to our surprise none showed a fully closed conformation in their respective crystal structures. So, we went back to the initial dataset we obtained from our combinatorial binding pocket analysis and the non-binding Prox constructs seemed to most interfere with binding and possibly closure. We already reverted D39E and deemed S38T as less of an influence since the general property of the residue was maintained. The most interesting exchange was D247S as we had identified this Asp as one of the first responders to ligand binding in PotF in our first manuscript. Therefore, it appears highly important for the binding mechanism and, hence, we reverted this residue in all our previously generated constructs and in PotF/D itself. The reversion alone already improved the SPD affinity comparable to the other constructs, but it did not lead to full closure. In our other variants, the change back to Asp combined with the re-establishment of the salt bridges led to the highest affinity SPD binders with residual PUT affinity in the medium to low micromolar range and to complete closure in the crystal structure. All in all, we were able to improve PotF/D and created a PotF construct that completely reverted its original polyamine preference for PUT (nanomolar to micromolar) and SPD (micromolar to nanomolar).

Overall, these two studies enable a more detailed understanding of the binding mechanisms in PotF. The results show that polyamine preference is mainly encoded in the aromatic box and the distal region of the binding pocket. These residues influence binding specificities by thermodynamically favouring or disfavouring the stabilization of water molecules. We pinpoint the prerequisites for complete closure to salt bridges and the key carboxyl-harbouring proximal residue D247, as fixing each component by itself did not lead to complete closure in PotF/D.

The combined insights of *in silico* analysis paired with robust biochemical and structural data can further serve as an evaluation benchmark in software development for ligand docking and protein design that might ultimately lead to improved design pipelines. An understanding of thermodynamic driving forces, including both solvent and molecular determinants, as well as of elements outside the binding pocket like salt bridges is crucial. These are important vantage points for protein engineering. Furthermore, the adaptability of polyamine transport proteins like PotF is important for medical research when looking at pathogenic bacteria. SPD uptake is linked to the expression of type III secretion systems (T3SS, Zhou et al., 2007), which play an important role in pathogenesis of multi-drug resistant bacteria (Felise et al., 2008; Keyser et al., 2008). We analysed the conservation rate of important binding pocket residues in 250 homo- and orthologs of PotF and found a much lower conservation throughout the distal residues, which are important to mediate polyamine preference. This shows an evolutionary trend towards more flexibility, a broader ligand spectrum and a possible higher promiscuity of polyamine uptake systems, which might allow pathogens to bypass treatments that only target one uptake system in the future.

In the draft for the third manuscript, we engineered PotF into a fluorescent biosensor for the neuromodulator AGM. Here, we used our knowledge of the binding mechanism from the two prior studies to change the binding preference of PotF towards AGM. We based our approach on the observation that AGM mimics the binding mode of PUT, which allowed for a high initial affinity, and exploited polyamine preference mediating properties of the distal residues in conjunction with the coordination of the unique AGM guanidino moiety in this region. To turn PotF into a fluorescent biosensor, we introduced a superfolder circular permuted GFP (sfcpGFP) into PotF, which facilitates quick screening of ligand specificity in plate reader assays upon engineering and the possibility to track AGM in a non-invasive manner *in vivo* in the future. For the initial construction, insertion sites were analysed and the linkers between PotF and sfcpGFP optimised for optimal fluorescence gain upon ligand binding. Additionally, the binding pocket was changed to lower the initial affinity for PUT, thus weakening carryover and prior saturation of the sensor binding pocket by endogenous ligands. The chosen mutations (S87Y & F276W) were known to lower affinity for PUT and SPD based on analysing PotF/D. In this initial construct, further mutations of the proximal and distal region were screened based on our prior PotF/D analysis.

This screening led to an AGM specific sensor construct (AGMsen) that carries the mutations S87Y and A182D. We also constructed a control sensor by introducing D247K. The Lys at position 247 masks the primary amine binding side in PotF and stops ligands from binding. Purified AGMsen showed an affinity for AGM of 38 µM and a dynamic range (max. $\Delta F/F_0$) of 3.0 in fluorescence-based dose-response assays. Other biogenic amines like SPD, PUT and CDV displayed K_D-values in the high micromolar to millimolar range which most probably will not interfere with specific AGM binding. The control sensor

showed insignificant residual interaction with AGM displaying a K_D of 2.4 mM. ITC measurements for the PotF receptor modules of both constructs confirmed no binding in the control sensor and selective AGM binding (K_D = 0.3 µM) for the final sensor. The difference between the determined K_D-values by ITC of the PotF receptor module and fluorescence-based dose-response assays of AGMsen are most likely attributable to sfcpGFP adding strain to a movement sensitive region of PotF in the sensor construct.

Furthermore, we introduced our sensors into the vectors pcDNA and pDisplay for intracellular expression and display on cells, respectively. The sensors showed robust expression in HEK cells for both systems and the expected membrane localization in the case of the displayed sensor. The measured affinities in HEK lysate matched the results of sensor purified from *E. coli* confirming no influence of eukaryotic specific factors e.g., glycosylation on sensor performance. The AGM affinity of HEK displayed AGMsen was lower (525 µM) compared to the lysate assay (35 µM). This is most probably attributable to elements like the transmembrane domain and secretion signal added by pDisplay, which might further influence the closing dynamics of the sensor. Still, we were able obtain a quantifiable signal and further tested our displayed sensor in primary rat hippocampal neuron culture. Neurons transduced with an adeno-associated virus (AAV) vector, which places the sensor gene under the control of a synapsin promotor, showed good expression and membrane localization after eight to ten days. We added AGM in different concentrations to neurons displaying AGMsen, which responded in the same regime (ca. 700 µM) as the HEK cells albeit with a low dynamic range. Nevertheless, these results confirmed functionality and will allow us set-up screening experiments which mimic neuronal performance, thereby enabling easier engineering of new sensor constructs in the future.

Overall, we confirmed that the distal binding pocket residues of PotF are the main driving force for AGM specificity and used our knowledge of polyamine binding at the proximal side of the pocket to design a control sensor. The control sensor can be used to set-up reliable controls and allows for the assessment of robustness of future experiments, since fluctuations in fluorescence not triggered by ligand binding will be easily traceable. We are aware, that AGMsen has room for improvement especially with respect to affinity and dynamic range, to detect weaker signals. New linker optimizations are a possibility to harvest the true potential of the high affinity AGM specific PotF receptor module identified by ITC. Additionally, mutations outside the binding pocket that destabilize the open conformation and drive the sensor towards a ligand bound form are good examples for application driven engineering without touching the already highly specific pocket. Our screening set-up is scalable and allows for rapid gene library screening in a high throughput fashion. Nevertheless, the current version of AGMsen shows robust expression and is functional in multiple experimental set-ups, which we hope to expand by working together with possible users to utilize their feedback and

improve the sensor. Providing this tool is a first step to better understand the neuronal modes of action of AGM in the future.

On a last note, all experimental features, and the functionality of the PotF-based sfcpGFP sensor are not AGM specific which allows us to alter the sensor towards other possible ligands by simply introducing point mutations in the binding pocket.

7. Own Contribution

7.1 Manuscript 1: A comprehensive binding study illustrates ligand recognition in the periplasmic binding protein PotF

For this manuscript I expressed and prepared the proteins and performed isothermal titration calorimetry (ITC) under the supervision of Sooruban Shanmugaratnam. I further set-up and screened crystallization conditions. Crystal preparation and data processing as well as structure building and deposition was done together with Sooruban Shanmugaratnam. The full *in silico* characterization by molecular dynamic simulations as well as the dihedral analysis was conducted by Noelia Ferruz. Kristian Schweimer measured and analysed all nuclear magnetic resonance spectra. All authors wrote their respective parts for the original draft. Finalization and submission of the manuscript was done by Birte Höcker.

7.2 Manuscript 2: Fine-tuning spermidine binding modes in the putrescine binding protein PotF

Ulrike Scheib and I performed mutagenesis to construct the different variants. Both of us expressed and purified different protein variants. Ulrike Scheib conducted initial ITC measurements on the combinatorial analysis of the binding pocket. I remeasured these in triplicates as well as performed and analysed all other ITCs. I conducted crystallization set-ups, while crystal preparation and data processing as well as structure building and deposition was done together with Sooruban Shanmugaratnam. Using these structures, we performed the dihedral analysis. I conducted the conservation analysis on binding pocket residues. Additionally, I wrote the initial draft and did the data curation as well as visualization together with Sooruban Shanmugaratnam. All authors finalized the manuscript and submission was done by Birte Höcker.

7.3 Draft for manuscript 3: A fluorescent biosensor for the visualization of Agmatine

Initial cloning and sensor construction was performed by André C. Stiel. I characterized the first and final sensor constructs. The set-up of the screening platform and mutagenesis for all tested variants was done by me. I expressed and purified all tested PotF receptor modules and analysed them via ITC. I set-up crystallization trials and performed crystal preparation with Sooruban Shanmugaratnam. I did the data processing as well as structure building for the sensor and control sensor structures. Final polishing of the structures and deposition was done together with Sooruban Shanmugaratnam. I cloned the sensor into the eukaryotic expression system and performed HEK cell experiments under supervision of Birthe Stüven in the laboratory of Dagmar Wachten. Birthe Stüven cloned the sensor

constructs into the adeno-associated virus (AAV) vector. AVV production as well as isolation and seeding of hippocampal neurons was performed in the group of Susanne Schoch. Transduction and experiments with neurons were conducted by Birthe Stüven. I wrote the initial draft, reviewing and editing was done by me, André C. Stiel, Dagmar Wachten and Birte Höcker.

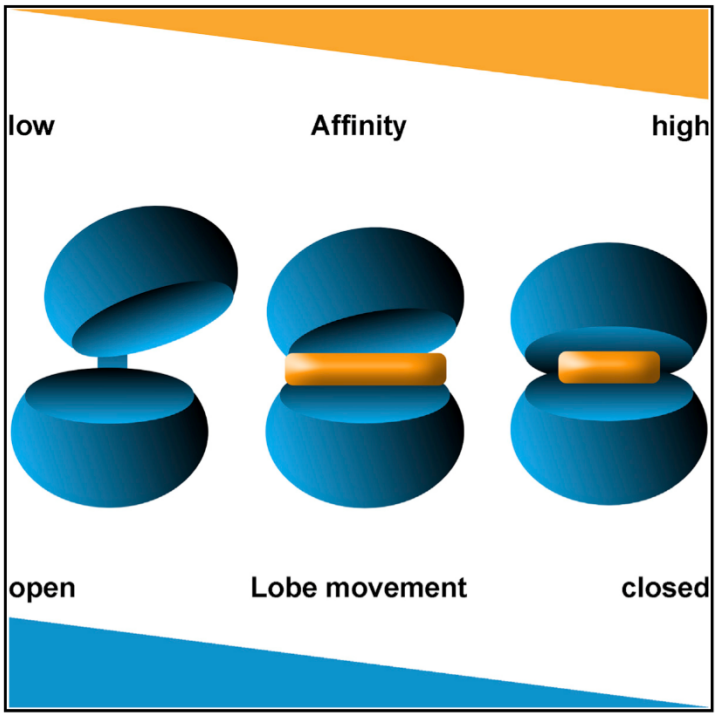
8. Manuscript 1

Article

Structure

A comprehensive binding study illustrates ligand recognition in the periplasmic binding protein PotF

Graphical Abstract



Authors

Pascal Kröger,
Sooruban Shanmugaratnam,
Noelia Ferruz, Kristian Schweimer,
Birte Höcker

Correspondence

birte.hoecker@uni-bayreuth.de

In Brief

Kröger et al. dissect PotF's binding capabilities for different ligands from polyamine biosynthesis utilizing orthogonal techniques. They combine extensive molecular dynamics simulations with in-solution NMR, crystal structures, and calorimetric data.

Highlights

- Binding analysis utilizing computational, biochemical, and structural techniques
- PotF binds several compounds from polyamine biosynthesis
- Crystal structures of closed and open apo states emphasize PotF's plasticity
- Binding pocket hydration pattern affects thermodynamic profile for the ligands

Kröger et al., 2021, Structure 29, 433–443
May 6, 2021
https://doi.org/10.1016/j.str.2020.12.005

CellPress

Structure

CellPress

Article

A comprehensive binding study illustrates ligand recognition in the periplasmic binding protein PotF

Pascal Kröger,[1,3] Sooruban Shanmugaratnam,[1,3] Noelia Ferruz,[1] Kristian Schweimer,[1,2] and Birte Höcker[1,4,*]
[1]Department of Biochemistry, University of Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany
[2]Northern Bavarian NMR Center, University of Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany
[3]These authors contributed equally
[4]Lead contact
*Correspondence: birte.hoecker@uni-bayreuth.de
https://doi.org/10.1016/j.str.2020.12.005

SUMMARY

Periplasmic binding proteins (PBPs) are ubiquitous receptors in gram-negative bacteria. They sense solutes and play key roles in nutrient uptake. *Escherichia coli's* putrescine receptor PotF has been reported to bind putrescine and spermidine. We reveal that several similar biogenic polyamines are recognized by PotF. Using isothermal titration calorimetry paired with X-ray crystallography of the different complexes, we unveil PotF's binding modes in detail. The binding site for PBPs is located between two lobes that undergo a large conformational change upon ligand recognition. Hence, analyzing the influence of ligands on complex formation is crucial. Therefore, we solved crystal structures of an open and closed apo state and used them as a basis for molecular dynamics simulations. In addition, we accessed structural behavior in solution for all complexes by ^{1}H-^{15}N HSQC NMR spectroscopy. This combined analysis provides a robust framework for understanding ligand binding for future developments in drug design and protein engineering.

INTRODUCTION

A myriad of protein-ligand interactions and their binding mechanisms are indispensable to most processes in life (Dunn, 2006) and are finely tuned for an organism's needs. For this purpose, receptor binding modes differ: some are specifically tailored toward unique binding partners, while others are more flexible and recognize a wide range of molecules. This promiscuity in ligand binding is a recurring feature, where molecules structurally and chemically associated with native ligands are usually bound with similar or lower affinities (Schreiber and Keating, 2011). A molecular understanding of the binding of different ligands provides insights into what determines and fine-tunes affinity and specificity. In addition, such receptors are promising candidates for changing specificity and thermodynamics in their promiscuous binding pockets. This approach paves the way for future rational protein engineering, ultimately leading to advanced therapeutics and diagnostics (Huang et al., 2016; Tiwari et al., 2012; Wilson, 2015).

In this work, we analyze the periplasmic binding protein (PBP) PotF from *E. coli*, which belongs to the PBP class II superfamily. It forms two α/β lobes, which are structurally similar and connected by a hinge region (Quiocho and Ledvina, 1996). The ligand-binding site is located at the interface between the two lobes adjacent to the hinge. Upon ligand recognition, PBPs bind with a twisting and closing motion often compared to a Venus flytrap (Felder et al., 1999). One representative of this ubiquitous protein family is the structurally and functionally well-characterized maltose-binding protein from the maltose/maltodextrin transport system of *E. coli* (Mächtel et al., 2019). Recent developments of biosensors for numerous solutes derived from a range of PBPs demonstrate the fold's versatility and applicability in life sciences (De Lorimier et al., 2009; Marvin et al., 2011, 2013, 2018). Especially, the evolutionary relationship to ionotropic glutamate receptor ligand-binding domains (Lee et al., 2019) illustrates the relevance of and opportunities for PBP-based tools in the neurosciences.

We previously showed that PotF has affinity toward its endogenous ligand putrescine (PUT) and additionally to spermidine (SPD), which consists of a PUT moiety extended with a propylamine (Scheib et al., 2014). To gain further knowledge of the molecular determinants of PotF's selectivity, we systematically characterized the binding event of several PUT analogs in polyamine biosynthesis. In detail, we performed isothermal titration calorimetry (ITC) experiments on PotF with a range of ligands, thereby gathering valuable insights into the forces that drive ligand binding (Keserü and Swinney, 2015). In addition, we solved high-resolution crystal structures for these PotF ligand complexes as well as for apo-PotF in a closed and open state, which provided details about the molecular mechanisms of ligand recognition and specificity as well as affinity.

CellPress

Structure
Article

Table 1. Thermodynamic signatures for binding of PotF to the five polyamines obtained by ITC measurements (n = 2)

Ligand	K_D (μM)	ΔH (kcal × mol^{-1})	−TΔS (kcal × mol^{-1})	ΔG (kcal × mol^{-1})	N
Putrescine	0.068 ± 0.040	−23.02 ± 0.12	13.28 ± 0.52	−9.74 ± 0.40	0.90 ± 0.02
Cadaverine	1.95 ± 0.16	−14.19 ± 0.02	6.53 ± 0.07	−7.66 ± 0.05	0.93 ± 0.02
Spermidine	29.71 ± 1.15	−3.65 ± 0.06	−2.42 ± 0.04	−6.07 ± 0.02	0.92 ± 0.00
Spermine	42.69 ± 2.13	−3.43 ± 0.01	−2.43 ± 0.02	−5.86 ± 0.03	0.97 ± 0.01
Agmatine	0.22 ± 0.08	−13.05 ± 0.21	4.09 ± 0.42	−8.96 ± 0.21	1.02 ± 0.04

The holo complexes reveal striking similarities. The ligands' amines are always coordinated in the same manner, sometimes even at the cost of adopting strained conformations. To study this unusual behavior, we performed molecular dynamics (MD) simulations, which revealed different preferred conformations in solution and showed how each ligand is recognized by PotF to different extents. We further probed this behavior using nuclear magnetic resonance (NMR) experiments. In accordance with the MD simulations, the degree of closing appears to correlate to affinity.

In summary, we show that PotF has extensive capabilities to bind different polyamines beyond what has previously been reported. We performed a highly comprehensive ligand-binding study, analyzing in detail crystallographic binding modes, thermodynamic properties, intrinsic dynamics, binding pathways, and equilibrium distributions of apo and holo systems. As such, we believe our study provides a robust platform to understand ligand recognition and lays the groundwork for future ligand and protein design work.

RESULTS

PotF binds several polyamines

We systematically characterized the binding of PUT and several of its analogs, such as SPD, cadaverine (CDV), spermine (SPM), and agmatine (AGM). Binding affinities and thermodynamics for all PotF-ligand interactions were analyzed using ITC (Table 1 and Figure S1). To prevent carryover of biogenic amines from expression in *E. coli*, the protein was unfolded and refolded during purification. Previous measurements showed that PotF binds PUT and SPD in the low nanomolar and micromolar range, respectively (Scheib et al., 2014). PotF has, as expected, a high affinity toward its endogenous ligand PUT (68 nM) and to CDV (1.95 μM), which is one methylene unit longer than PUT. Similar to prior studies (Scheib et al., 2014) PotF shows a moderate affinity toward SPD (29.71 μM), which is also the case for SPM (42.69 μM). Interestingly, PotF shows the second highest affinity for AGM (0.22 μM) in our test set, despite the molecule being bulkier because of its guanidino group. In addition, corresponding amino acids and other biogenic amines were tested, but did not show measurable affinities by ITC (Table S3). Although PotF binds many different biogenic amines, it seems that the carboxyl group of corresponding amino acids obstructs binding.

Another interesting property is a change in thermodynamics while binding different biogenic amines. For the wild-type ligand PUT, the thermodynamic contributions are enthalpy driven, as expected, whereas the thermodynamics shift together with decreasing affinities toward entropy-supported binding modes (Table 1).

Structural details from crystallographic studies

To understand observations made using ITC and to gain structural insights we crystallized apo-PotF in open (PDB: 6YED) and closed (PDB: 6YEB) conformations (Figures 1B and 2) as well as in complex with all five ligands (Figure 3). These seven structures present observable differences in their conformations, evidenced when projecting them onto a subspace defined by opening and twisting angles between the two lobes (see STAR Methods for definition of dihedral angles). Figure 1B summarizes these projections for the complexes of this work along with previously crystallized PotF structures for comparison, the PotF:PUT complex (Vassylyev et al., 1998) and the variant PotF_SPD (Scheib et al., 2014). For crystallographic details of all structures see Tables 2 and S2.

The structures adopt various conformations in which their twisting and opening angles almost linearly correlate. The most closed conformations have sharper twisting and opening angles and lie at the lower part of the diagram. These structures are the PotF complexes with PUT (PDB: 6YE0), SPD (PDB: 6YE8), AGM (PDB: 6YE6), and CDV (PDB: 6YE7). Interestingly, the apo-PotF closed conformation perfectly superposes with these structures, with angles of ∼119° and ∼33° (opening, twisting). Slightly more open is the PotF:SPM (PDB: 6YEC) conformation, in a similar range compared with previously crystallized PotF_SPD, defining a set of intermediate conformations between closed and open. The widest open conformation is the second apo-PotF structure, with angles of around 156° and 69°.

In addition to these similarities and differences from a global view of the structures, the binding pockets offer interesting insights as well. The closed apo structure shows a highly coordinated water network in the binding pocket (Figure 3A). Most water molecules in such internal pockets form at least three hydrogen bonds to be thermodynamically favorable (Schiebel et al., 2018). Ten water molecules appear coordinated in the pocket positions, all of them showing at least three hydrogen bonds. All water molecules in the proximal (section of the binding pocket formed by residues S38, D39, D247) and distal parts of the pocket (section formed by residues E185 and D278) can form four hydrogen bonds, while the waters in the aromatic box (formed by residues W37, Y40, W244, F276, and Y314) have three hydrogen bonds, with only one of them as a result of interactions with the protein. A characterization of the solvation pattern of the apo protein is crucial for understanding ligand binding, as the differences in the water pattern of unligated and ligated states gives an idea of which hydrogen bonds play an important role in complex formation. Interestingly, a superposition of all the complexes reveals that the water molecules overlap with the coordinates of the polyamine's nitrogens and some carbons (Figure S2D).

Structure
Article

CellPress

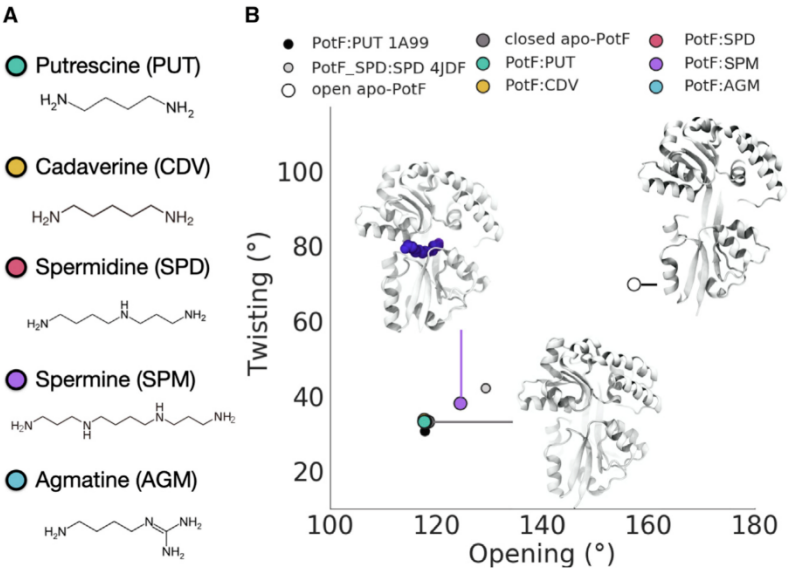


Figure 1. Complexes crystallized in this work and their general structural features
(A) Overview of chemical entities bound by PotF.
(B) Projection of the seven structures determined in this work along with two reference structures onto the opening and twisting angles. Closed apo-PotF, PotF:CDV, PotF:SPD, and PotF:AGM are congruent to PotF:PUT. Conformations of PotF:-PUT, PotF:SPM, and open apo-PotF are shown as references. A detailed description of the angles is shown in Tables S5 and S6. Protein representations were rendered using VMD (Humphrey et al., 1996).

The crystal structure of PotF in complex with PUT has been reported before (Vassylyev et al., 1998), and is the only available structure of a PotF wild-type ligand complex to date. Our structure aligns to the previously reported with a root-mean-square deviation (RMSD) of 0.35 Å for 341 Cα atoms. Furthermore, with 1.63 Å (Table 2), this structure is determined at roughly 0.6 Å higher resolution and shares the same space group with all our other complexes as well as the closed apo structure; hence, it allows better comparability throughout the dataset.

In complex with PotF, PUT is tightly coordinated at the proximal site of the binding pocket by several interactions. The protonated amine at N1 interacts via hydrogen bonds with the S38 and D39 backbone carbonyls, the Y314 hydroxyl, and three water molecules. It is also involved in a salt bridge with D247. PotF's central aromatic box forms strong cation-π interactions to N2 as well as C-H-π interactions to C4 of the ligand molecule, tightly enclosing PUT's aliphatic moiety. On the distal side of the pocket, the N2 amine forms a salt bridge with D278 (Figure 3B).

The formation of the amine salt bridges seems to be the main driving force for the binding of this series of compounds, as other ligands seem to take a strained conformation upon complex formation in order to coordinate its amines with D247 and D278. This is most visible in the case of CDV, because it is the only tested ligand that does not have a PUT-like element within its structure (Figure 1A). CDV forms its salt bridges to the downside by bending its methylene backbone and partially disrupting stacking interactions to fit its nitrogens into a PUT-like position inside the binding pocket (Figures 3C and S2A). Aside from these differences in conformation in the alkyl chain, CDV and PUT show remarkably similar binding modes, with water network and side-chain conformations superposing perfectly. Therefore, the additional strain put on CDV could account for a more than one order of magnitude lower binding affinity to PotF than PUT. This is expressed in the comparison of the dihedral angles between PUT C2-C3-C4-N2 (~162°) and CDV C2-C3-C4-C5 (~93°) (Figure S2A). A related case has already been described for the structurally similar compound n-butane (Woller and Garbisch, 1972).

SPD presents a structure similar to that of PUT, with the addition of a propylamine to its skeleton (Figure 3D). Its binding mode resembles that of PUT and CDV in the proximal side (Figure S2B), with its most proximal amine interacting with D247 and its secondary amine with D278. The extra propylamine extends in the distal side, displacing two of the well-coordinated water molecules (Figure S2D). The displacement of these thermodynamically favorable waters can explain the loss in enthalpy of SPD's binding profile. In addition, SPD binding is entropically driven as well, which could be the result of the more hydrophobic character of the interaction. SPM presents the largest molecule of the series, it is symmetric and is a polycation at physiological pH (Figure 1A). The two secondary amines are the ones that interact with D247 and D278, while the terminal amines extend in the proximal and distal sites, displacing the water networks in both cases. PotF:SPM's opening and twisting angles significantly deviate from those of the other complexes (Figure 1B). Due to the ligand's extended conformation, the two lobes are not able to close entirely, and thereby salt bridge formation between D39 and R254 at the proximal and R91 and E184 at the distal side is affected as well. The thermodynamics of binding are highly similar to those of SPD because SPM displaces the same water molecules at the distal side of the binding pocket.

More similar to PUT and CDV, AGM also coordinates very tightly in the PotF pocket. Despite AGM being slightly larger than PUT, its alkyl group does not adopt an unfavorable conformation to coordinate residues D247 and D278. In fact, the PUT and AGM structures superpose from the proximal nitrogen to guanidino NE, which, as PUT's N2, coordinates to D278. This is also reflected in almost the same dihedral angle (~160°) compared with PUT (~162°) (Figure S2C). The rest of the guanidino group nicely coordinates the distal water molecules via hydrogen bonds (Figure 3F).

In silico analysis using MD simulations

To gain a fundamental understanding of PotF's distribution of conformational states in the absence and presence of ligands, we performed all-atom, explicit-solvent MD simulations totaling

CellPress

Structure
Article

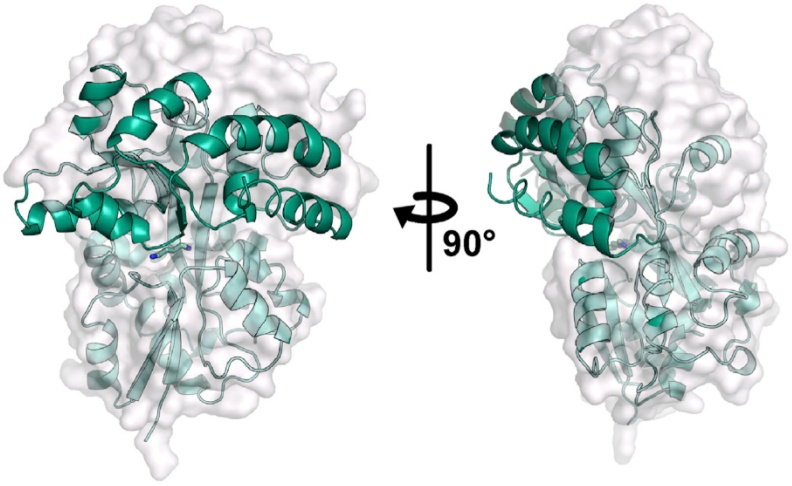


Figure 2. Comparison of PotF's open and closed conformations

Open conformation is shown as a light gray surface. As a representative of the closed conformation, PotF in complex with PUT is shown as a green cartoon with PUT in sticks. Alignment was done in PyMOL over the Cα atoms of residues 31–132 corresponding to the bottom lobe (RMSD 0.585 Å).

more than 595 μs of simulation time (Table S1). We ran six different simulation batches. First, we simulated apo-PotF at low salt concentration, where the initial protein input coordinates were taken from the open and closed apo-PotF X-ray structures (STAR Methods). Then, we performed five subsequent simulation ensembles with apo-PotF in the presence of the different ligands. In these cases, the input crystal structure was open apo-PotF and the ligands were initially placed at a considerable distance from the protein surface. All simulation batches were run in the absence of external forces, allowing protein and ligands to freely form a complex and stabilize different opening angles.

We analyzed these data by constructing Markov state models (MSMs) of each ensemble of trajectories. MSMs have been previously used to calculate slow processes from ensemble MD simulations and have proven very successful in ligand-binding studies (Ferruz et al., 2015, 2018; Plattner and Noé, 2015).

Exploring PotF's conformational space with MD

The simulation ensemble started from open and closed apo-PotF structures (Figure 1A) as the input conformations. Being interested in the dynamics of opening, we projected the trajectories into a space defined by the protein backbone dihedral angles and the global opening and closing angles for our analysis (STAR Methods). The conformational space was discretized into 1,000 clusters that we lumped into four states that contain similar conformations (Prinz et al., 2011) (Figure 4). The choice of the number of macrostates was based on inspecting the implied timescales (Figure S3) as suggested in previous studies (Bowman et al., 2009a; Prinz et al., 2011). These states describe PotF's most populated states in solution in the absence of ligands or salts at high concentration.

State 1 is the lowest populated state (3.7% ± 1.4%) and includes the structures with the narrowest average opening and twisting angles (120.4° and 29.2°). Conformations in this state are similar to those of most ligand complexes in this work (Figure 1A) and the closed apo-PotF structure (Figure 4A). State 1 converts to state 2, a slightly more open state (139.2°, 50.5°) and the second most populated in solution (39.5% ± 2.9%). The conversion occurs very rapidly, in 51.0 ± 3.3 ns. Reversion to state 1 occurs three orders of magnitude slower, in 11.5 ± 4.3 μs, leading to state 1's low population. State 2 converts rapidly to state 3 (in 188.7 ± 22.4 ns), the most populated state in solution (45.0% ± 2.9%), which in turn also reverts to state 2 on a similar timescale (223.4 ± 17.1 ns). State 3 presents a more open structure than the two previous states, and the ensemble structures are most similar to the open apo-PotF crystal structure (Figure 4A). State 4 is a transient state that converts very quickly to state 3 (68.1 ± 3.4 ns). The reverse process is two orders of magnitude slower and leads to its low population (11.8% ± 1.5%). State 4 presents the most open structure, with angles of 165.0° ± 0.2° and 87.2° ± 0.9°, respectively.

The slowest exchange process we identified in apo-PotF occurs in 11.8 ± 4.4 μs and describes the transition from the most open state (4) to the closest one (1). Previous work had already characterized an unligated closed state in proteins of the same fold. Specifically, Tang et al. showed that apo-MBP shows a rapid exchange in the nanosecond to microsecond timescale between a major open form (~95%) and a minor closed species (~5%) (Tang et al., 2007). More recently, Silva et al. found that the LAO protein is able to sample a closed state that presents the same structure as a rate-limiting encounter complex that forms with its ligand arginine before fully transitioning to the bound state (Silva et al., 2011).

Characterization of binding processes by MD

We ran simulation ensembles for the five ligands in this work totaling 0.5 ms of aggregate time (Table S1). The five systems present the same protein, ligand, and ion concentrations, thus allowing direct comparison. Projection of the raw data into the dihedral angles shows that the simulations explore slightly different regions of PotF's conformational space (Figure S4). Such differences, however, do not provide information on the most populated states in equilibrium. We thus constructed MSMs on each of the systems, which allowed us to identify the bound poses, intermediates, and binding pathways for each ligand (STAR Methods).

In all our models we observe four major states (Figure S5). The first state, bulk, corresponds to the ligand-free form in solution. Encounter 1 and 2 states correspond to intermediate states that the ligand populates prior to reaching the final bound pose. Whereas encounter 1 is populated only in secondary pathways (i.e., the state is visited in less than 50% of the pathway fluxes), encounter 2 is a necessary step in the

Structure
Article

CellPress

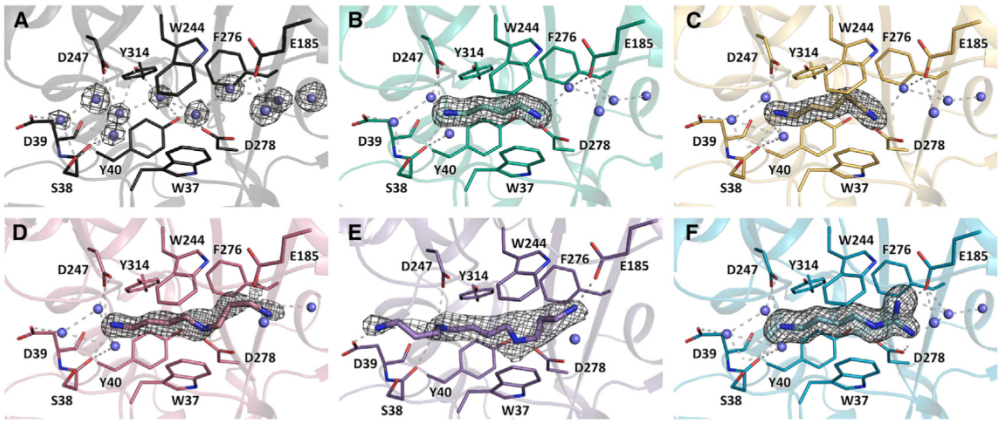


Figure 3. Comparison of binding pocket structures in PotF complexes

Binding pockets of PotF in the closed apo form (A; PDB: 6YEB) and in complex with PUT (B; PDB: 6YE0), CDV (C; PDB: 6YE6), SPD (D; PDB: 6YE8), SPM (E; PDB: 6YEC), and AGM (F; PDB: 6YE6). Ligand molecules and residues forming the binding pocket are shown as sticks. 2mFo-DFc maps exported as ccp4map files by phenix.mtz2map for ligands and water molecules in the closed apo structure are shown as mesh contoured at 1 σ using PyMOL. Water molecules are depicted as blue spheres and polar contacts as light gray dashed lines.

conversion from bulk to bound. Encounter 1 corresponds to a state that is semi-open, with opening and twisting angles ranging from 147° (SPD) to 157° (AGM) and 57° (SPD) to 70° (CDV). Encounter 2 corresponds to a state that is semi-closed, resembling state 2 in apo-PotF conformational space, with opening and twisting angles ranging from 127° (SPD) to 137° (PUT, AGM) and 37° (SPD) to 51° (PUT), respectively, and where the ligand has identified the binding cavity but has not yet adopted the most favorable conformation. In all cases, these transitions from bulk to bound states are dominated by very fast timescales, with on rates ranging from 1.56×10^8 (PUT) to 1.21×10^9 M^{-1} s^{-1} (SPM). Figure 5A shows a projection of the bound ensembles in comparison to the crystal structures. The bound poses for all ligands are within a 2 Å RMSD to their respective experimental structures, taking into account all protein α carbons and ligand heavy atoms for an ensemble of 1,000 bound structures. In detail, the RMSDs are 1.86 (PUT), 1.73 (CDV), 1.73 (SPD), 1.89 (SPM), and 1.50 Å (AGM). The X-ray structures present closed conformations with exact values of 118°, 33° (opening, twisting) for PUT, CDV, SPD, and AGM and 124°, 38° for SPM. The bound ensembles obtained from the MSMs are defined by a set of conformations that define the bound state, and therefore are presented by regions in space (Figure 5A). The regions sampled by PUT, CDV, and AGM cover their respective crystal structures. In particular, the average positions of these regions lie at values 119°, 27° (PUT); 119°, 28° (CDV); and 123°, 33° (AGM). SPD and SPM bound ensembles deviate more from the experimental structures. The SPD bound ensemble has average angles of 126° and 35°, sampling the region explored by the PotF:SPM and PotF_SPD:SPD X-ray structure. The SPM bound ensemble also populates a more open conformation than the experimental counterpart, with average angles of 131° and 42°.

The final poses resemble those observed in the X-ray structures (Figure 5B), with the amines being coordinated by D39, D247, D278, and E185 as described previously. The 12 closest binding residues in the crystal structures and MD bound ensembles are summarized in Table S4. Although the local binding pocket poses of the simulations and the crystal structures match, the global conformations of the protein differ. All experimental structures, except PotF:SPM, display the same twisting and opening angles, while the simulations show different conformations. To further investigate this observation, we applied NMR spectroscopy in order to gain a better understanding of the in-solution behavior upon ligand binding.

Solution NMR of PotF and its complexes

Thus, we measured ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) NMR spectra to gain further insights into the most populated states in solution. Addition of ligands to PotF causes chemical shift changes demonstrating binding of the ligands to PotF. For all used ligands the same resonances of PotF are affected, suggesting a similar binding mode (or binding in the same binding pocket) of the ligands in line with the previous results. A similar binding mode will generate similar peak shifts upon ligand recognition, and in our case, we assumed that completely closed states would show the same changes of chemical shifts. Due to the large conformational change induced by ligand binding, resonance changes are observable over the whole spectrum (Figure S6). Nevertheless, we see clear differences between the different saturated ligated complexes. Not surprisingly, the biggest effect in terms of chemical shift changes is detected after addition of PUT. The other complexes show very similar spectra and changes, but some peaks do not always shift to the same extent as is the case with PUT (Figure 6). Interestingly, the extent of the chemical shift changes correlates with the affinity measured with ITC. The peaks for PotF:CDV and PotF:AGM move close to or even overlap with the ones for PotF:PUT, whereas PotF:SPD and PotF:SPM show remarkable differences. The peaks do not shift

CellPress

Structure
Article

Table 2. Data collection and refinement statistics

	Apo PotF open	Apo PotF closed	PotF:PUT	PotF:CDV	PotF:SPD	PotF:SPM	PotF:AGM
PDB ID	6YED	6YEB	6YE0	6YE7	6YE8	6YEC	6YE6
Data collection							
Space group	P 1 2_1 1	P 3_2 2 1	P 3_2 2 1	P 3_2 2 1	P 3_2 2 1	P 3_2 2 1	P 3_2 2 1
Cell dimensions							
a, b, c (Å)	76.2, 53.5, 88.1	71.1, 71.1, 272.4	71.0, 71.0, 272.6	70.9, 70.9, 271.2	70.8, 70.8, 272.1	70.4, 70.4, 273.5	71.2, 71.2, 272.9
α, β, γ (°)	90, 112, 90	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Molecules per ASU	2	2	2	2	2	2	2
Resolution (Å)	45.90–2.18 (2.26–2.18)	45.68–1.97 (2.04–1.97)	45.64–1.63 (1.69–1.63)	45.50–1.60 (1.66–1.60)	40.71–1.50 (1.55–1.50)	45.58–2.09 (2.17–2.09)	45.73–1.56 (1.62–156)
R_{merge}	0.157 (0.970)	0.124 (3.324)	0.122 (2.482)	0.091 (2.547)	0.071 (2.663)	0.108 (3.334)	0.065 (2.922)
I/σ (I)	13.4 (4.7)	12.2 (0.7)	10.1 (0.7)	14.9 (0.7)	17.8 (0.8)	14.6 (0.7)	18.6 (0.7)
$CC_{1/2}$	0.997 (0.871)	0.999 (0.230)	0.998 (0.274)	0.999 (0.231)	0.999 (0.264)	0.999 (0.229)	0.999 (0.272)
Completeness (%)	99.9 (99.9)	99.5 (99.9)	99.9 (99.8)	99.2 (94.3)	100.0 (100.0)	100.0 (100.0)	99.8 (99.9)
Redundancy	10.9 (9.7)	10.9 (11.3)	10.7 (9.9)	10.8 (9.0)	11.5 (11.0)	11.0 (11.4)	11.0 (11.1)
No. total reflections	735,129 (33,409)	629,899 (64,561)	1,077,788 (98,256)	1,131,843 (88,835)	1,471,775 (138,718)	525,454 (53,487)	1,262,884 (125,863)
Refinement							
R_{work}/R_{free}	0.228/0.273 (0.326/0.352)	0.191/0.219 (0.323/0.305)	0.173/0.201 (0.324/0.325)	0.168/0.194 (0.328/0.378)	0.162/0.182 (0.334/0.345)	0.211/0.239 (0.344/0.362)	0.171/0.186 (0.338/0.359)
No. atoms							
Protein	5,381	5,373	5,411	5,377	5,510	5,351	5,393
Ligand	NA	NA	12	14	20	28	18
Water	307	261	588	612	606	141	494
B factors (Å^2)							
Protein	43.04	54.94	36.88	35.57	32.36	65.11	37.21
Ligand			23.80	28.59	39.01	66.19	29.65
Water	40.12	57.42	44.10	42.60	41.54	66.92	44.35
Root-mean-square deviations							
Bond length (Å)	0.002	0.009	0.012	0.009	0.009	0.002	0.010
Bond angles (°)	0.50	0.94	0.93	0.94	1.09	0.50	1.00

as much or, in the latter case, the spectrum for PotF:SPM looks almost like the one for apo-PotF.

Translational diffusion measurements by NMR spectroscopy

Ligand binding causes changes in the overall structure of PotF resulting in a more compact closed structure. Translational diffusion measurements were performed to detect changes in compactness. The translational diffusion coefficient is inversely proportional to the hydrodynamic radius of the protein. At 308 K for unbound PotF a translational diffusion coefficient of 1.013 ± 0.008 × 10^{-10} m^2/s was obtained. A saturating excess of PUT under otherwise identical conditions increased the diffusion coefficient to 1.126 ± 0.008 × 10^{-10} m^2/s, demonstrating a reduced hydrodynamic radius in line with a closed structure. In the case of SPM, a diffusion coefficient of 1.08 ± 0.01 × 10^{-10} m^2/s was observed, which is between the values for unbound and PUT-bound PotF (Figure S7). This supports a partially closed state of PotF in the presence of a weakly binding ligand in line with our previous observations.

DISCUSSION

PotF has two (α/β)-lobes connected by a hinge region that encloses the ligand-binding cavity. The protein undergoes large conformational rearrangements upon ligand binding, from an open to a closed state. It is assumed that the binding to PBPs occurs via an induced fit mechanism, where closure of the lobes is triggered upon ligand recognition. Our work provides detailed insights into how the PBP PotF from *E. coli* binds different polyamines on the structural level, supported by thermodynamic data as well as a detailed *in silico* analysis.

We could show with ITC experiments that PotF has the capability to bind even more polyamines than known so far. Apart from the reported affinity toward PUT and SPD, it also can bind related biogenic polyamines, namely CDV and SPM, perhaps not surprisingly, as they differ mainly in the size of their carbon chain. The measured affinities show a strong dependency on the length of the molecule. Basically, we can state that the longer the diamine, the weaker the affinity. When we

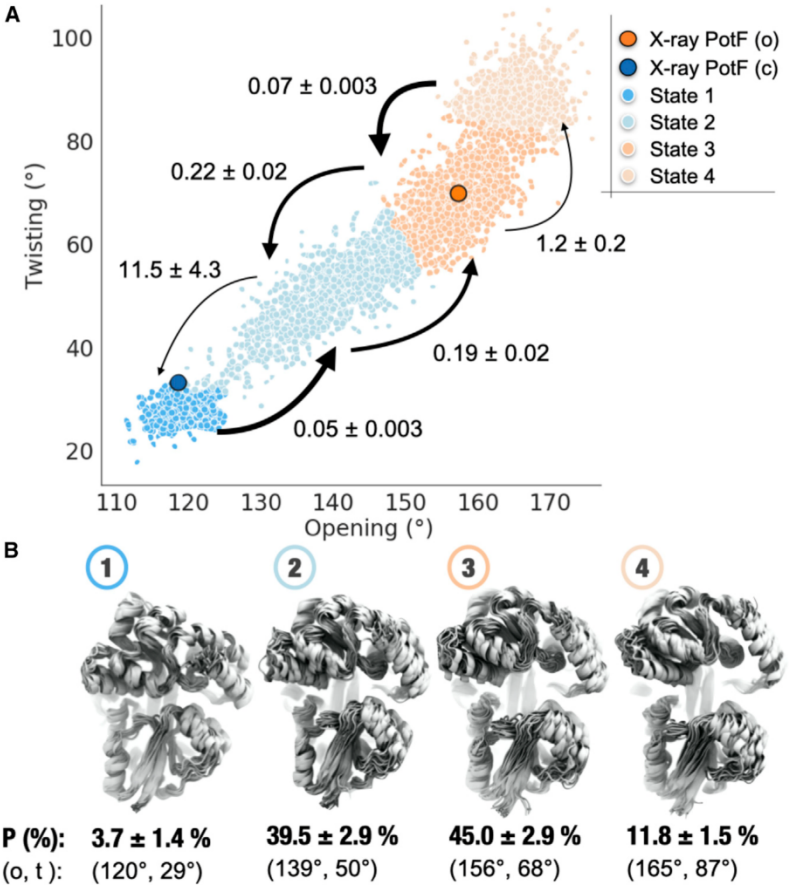


Figure 4. Populations, dynamics, and structural properties of apo-PotF states in solution

(A) Projection of the four states onto their opening and twisting angles. Each state is composed of an ensemble of structures that convert on a rapid timescale. Conversion kinetics between states are shown in microseconds. Open and closed PotF crystal structures are also shown for comparison. (B) Representative structures superimposed by using all protein α-carbons, population of conformations for each state, and mean opening and twisting angles for each ensemble in (A). Protein representations were rendered using VMD (Humphrey et al., 1996).

expanded the test set to the corresponding amino acids, no binding could be detected. Arguably, the carboxyl-group likely clashes with residues in the binding pocket or disrupts important interactions and thus prevents proper binding. Another scenario might be a missing interaction partner for the carboxyl group to accommodate displaced solvent molecules. Further exploration of ligands led us to AGM, a naturally occurring derivative of arginine, which acts as neurotransmitter and plays a role in many other metabolic and regulatory processes (Piletz et al., 2013; Raasch et al., 2001; Satriano, 2004). Surprisingly, ITC measurements with AGM showed an affinity in the regime between PUT and CDV, which raised the question of how PotF can accommodate AGM with its bulky guanidino group. To answer this, we applied crystallography and solved the structure of PotF in complex with AGM. A closer look into the binding pocket reveals that in addition to the coordinating residues for the PUT part, residues S85, D278, and E185 coordinate the guanidino moiety at the proximal side, explaining the initially unexpected high affinity.

Structure determination of PotF gave insights into the binding mode of AGM, so we crystallized PotF also in complex with the whole set of ligands to gain more detailed information about the molecular and structural determinants of binding. In addition to all ligand complex combinations we obtained apo structures in the closed and open unligated conformation. A structural superposition aligning on just one lobe provides the best overall view (Figure 2). The complexes with PUT, CDV, AGM, and SPD are strikingly similar. Beyond that, CDV and SPD adopt strained conformations in order to anchor the nitrogens to proximal and distal glutamates, exactly superposing with PUT (Figure 3). Even water molecules in the binding pocket seem to be highly conserved and serve as placeholders for bigger ligand molecules.

This fits well with our ITC data as we observe binding modes shifted to an entropically supported binding the larger the ligands get. The more water molecules are displaced in the binding pocket, the higher the solvation entropy will be, and this is reflected in the thermodynamics of the binding processes.

Nevertheless, the structure with SPM—the longest ligand in our test set—shows some interesting differences. Like the other ligands, SPM also adopts the nitrogen positions of PUT, but due to its size, it disrupts two salt bridges at both ends of the binding pocket. These missing interactions lead to a slight opening of the two lobes and might be the reason for the lowest affinity in our test set. It seems that these salt bridges act as two clamps stabilizing the closed state and, by doing so, enhance affinity.

By running high-throughput MD simulations the equilibrium distribution of apo-PotF in solution and the binding processes of the five ligands could be characterized. Apo-PotF in solution mostly populates two open states, resembling the most probable state in equilibrium distribution and the conformation observed in the open apo-PotF X-ray structure. The closed apo-PotF conformation is minimally sampled (3%), but definitely accessible under standard conditions in the microsecond timescale. Characterization of the binding events via MSMs led to five bound poses within 2 Å of the X-ray structures. Projection of the bound ensembles onto the dihedral subspace revealed that PUT-, CDV-, and AGM-bound poses cover the conformation

CellPress

Structure
Article

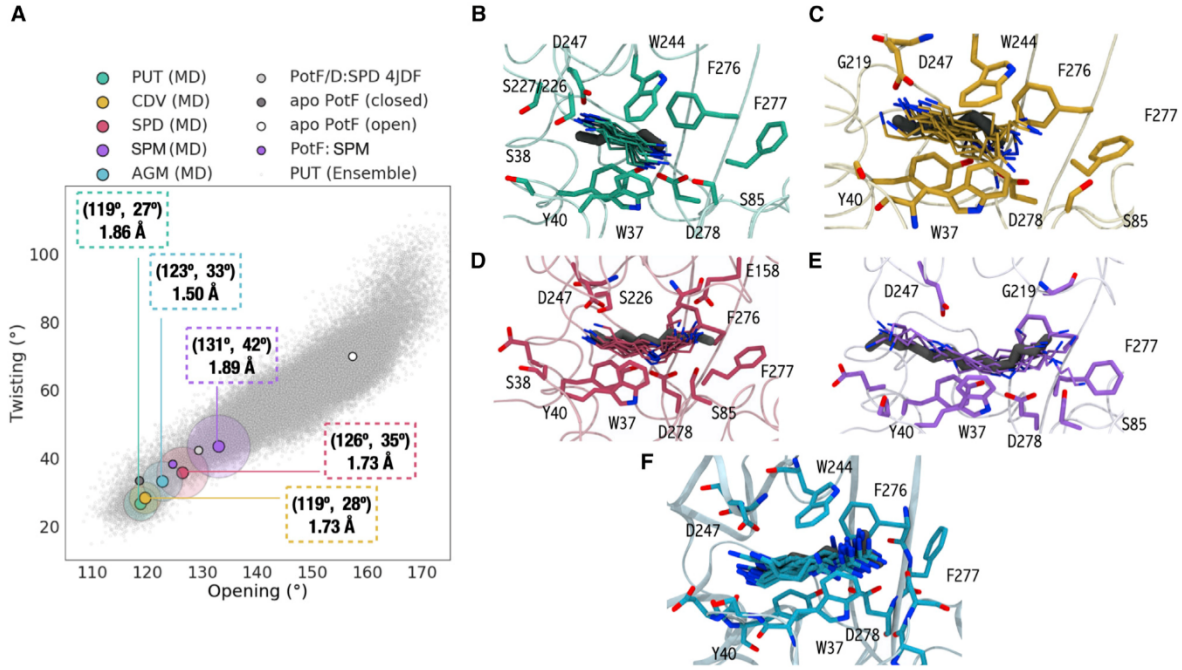


Figure 5. Binding poses for the five different ligands

(A) Projection onto the dihedral space for the five bound pose ensembles. Depicted for each pose are their average angles and their RMSD to the respective X-ray structure. Shown in the background are data points sampled by the PotF:PUT system in the simulations.

(B–F) Close-ups of the binding poses of PUT (B), CDV (C), SPD (D), SPM (E) and AGM (F) relative to the X-ray structures. Protein representations were rendered using VMD (Humphrey et al., 1996).

observed in the X-ray structures, but SPD and SPM populate a region with slightly larger opening and twisting angles than their experimental counterparts.

The discrepancy between the calculated models from the MD simulations and the crystal structures questions our hypothesis of complete closure upon ligand binding, regardless of the bulkiness of the ligand molecule. Consequently, we measured ^{1}H-^{15}N HSQC NMR spectra to test PotF's behavior in solution. Initial spectra of the apo protein provide the open conformation state before addition of any ligand. After supplying ligand in saturating concentrations, we could clearly observe various signal changes, and some of these showed a nuanced response to the different ligands. While moving along the same trajectory, peaks did not shift to the same extent as in the spectrum with PUT. Further NMR experiments will be necessary to analyze this behavior in solution. However, these findings support the MD simulations and suggest that PotF can indeed adopt slightly different closed states upon binding different ligands. Although X-ray crystallography provided nicely comparable complex structures, subtle differences in the global conformation as observed in MD and NMR might be overridden by crystal forces that select just one state. This highlights even more the necessity of integrated approaches to understand binding in the highly flexible PBP scaffold.

All our results taken together, we reveal not only that PotF binds more ligands than reported until now. We also show how PotF binds its ligands in a set of crystal structures together with thermodynamic data. Furthermore, we used NMR measurements to support theoretical data from MD simulations indicating that PotF can adopt different closed states in solution.

Ligand promiscuity has also been analyzed in other PBPs, such as HisJ and LAO-BP. HisJ is associated with high-affinity histidine binding and translocation, but also binds a range of naturally occurring related compounds as determined in ITC and NMR titrations (Paul et al., 2017). The protein's ability to access the closed conformation also in the apo form has been explored by MD simulations (Chu et al., 2014), but structures as we provide for PotF have not been otherwise determined so far. LAO-BP, on the other hand, is named for its ability to bind lysine, arginine, and ornithine and has been shown to also bind histidine with lower affinity. To study the binding mechanism MD studies were performed on LAO-BP showing the sampling of a closed state that resembles the rate-limiting encounter complex (Silva et al., 2011). In a recent study, then, an assessment of energetics regarding arginine and histidine binding using ITC was combined with structural analysis of protein-ligand complexes and revealed that affinity and selectivity in LAO-BP are shaped through the interplay of protein-ligand and solvent interactions (Vergara et al., 2020). In these aspects, there are commonalities in the findings of the LAO-BP study and our study. LAO-BP has also been compared with glutamine binding protein,

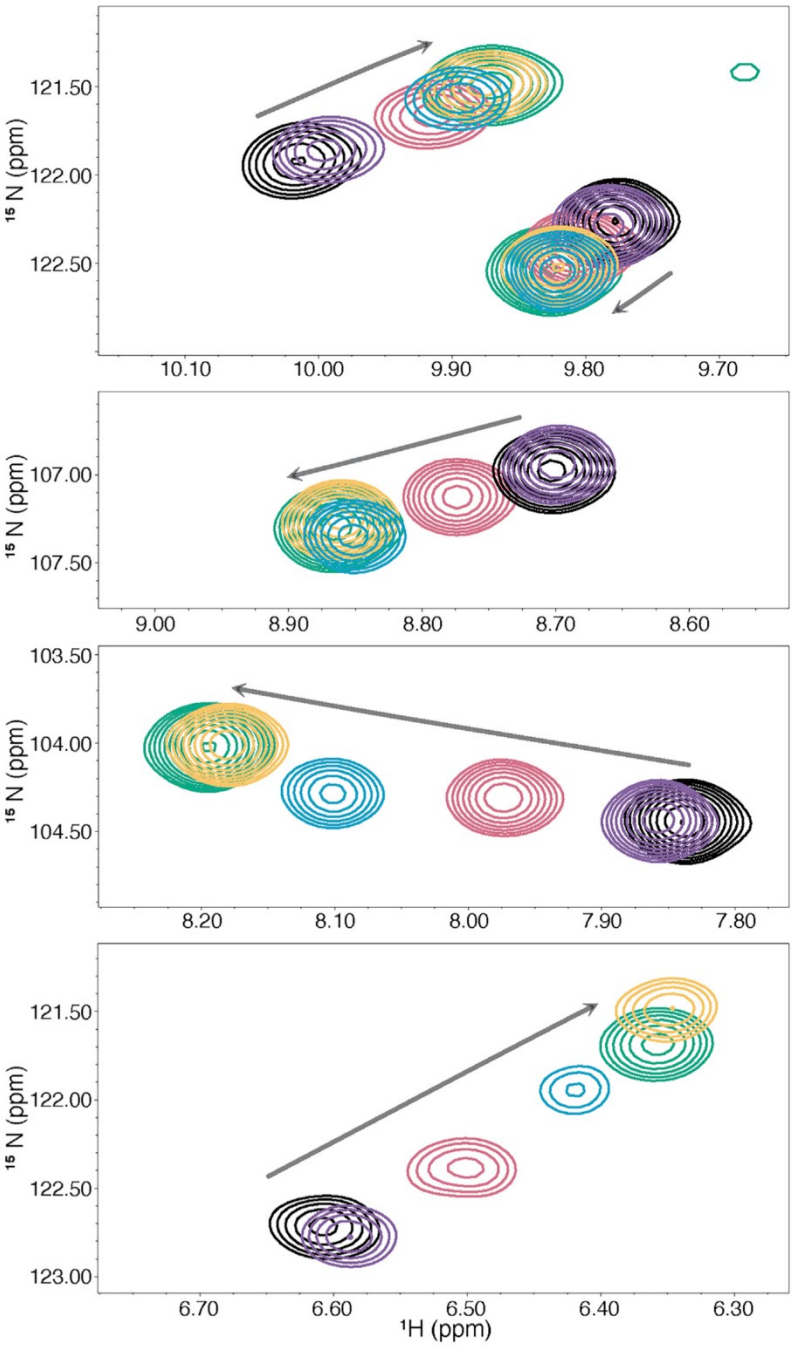


Figure 6. Chemical shift changes from ^{1}H-^{15}N HSQC spectra

Selection of four peaks that exemplify shift differences upon ligand binding. Peaks for apo-PotF are in black, PotF:PUT in green, PotF:CDV in yellow, PotF:AGM in blue, PotF:SPD in salmon, and PotF:SPM in purple. Arrows show the direction of shift.

which led to explorations of the evolution of binding specificity in these solute-binding proteins (Clifton and Jackson, 2016) as well as attempts to design new specificities (Banda-Vazquez et al., 2018). Dynamics have not been considered in such approaches, but are expected to play an important role and need to be considered.

Overall, our study enables a more detailed understanding of the binding mechanisms in PotF. These combined insights paired with robust biochemical and structural data are envisioned to further serve as an evaluation benchmark in software development for ligand docking and protein design. An understanding of thermodynamic driving forces, including

both solvent and molecular determinants, is an important vantage point for protein engineering that should ultimately lead to improved design pipelines in the future.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Bacterial strains
- METHOD DETAILS
  - Cloning, expression & purification
  - Isothermal titration calorimetry (ITC)
  - Crystallography
  - Molecular dynamics simulations
  - Markov state modeling analysis
  - NMR spectroscopy
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.str.2020.12.005.

ACKNOWLEDGMENTS

We thank the beamline staff at BESSY and Swiss Light Source (SLS) for support. We acknowledge financial support and allocation of synchrotron beamtime by Helmholtz-Zentrum-Berlin, as well as beamtime allocation by SLS. This work was supported by the Deutsche Forschungsgemeinschaft, grant HO4022/2-3.

AUTHOR CONTRIBUTIONS

P.K., S.S., N.F., and B.H. wrote the manuscript. P.K. and S.S. measured and analyzed ITC and crystallized, solved, and deposited the X-ray structures. N.F. set up the simulation systems, ran and analyzed the MD simulations, and performed the dihedral analyses. K.S. measured and analyzed the NMR spectra.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 7, 2020
Revised: October 28, 2020
Accepted: December 8, 2020
Published: January 5, 2021

REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Banda-Vazquez, J., Shanmugaratnam, S., Rodriguez-Sotres, R., Torres-Larios, A., Höcker, B., and Sosa-Peinado, A. (2018). Redesign of LAOBP to bind novel L-amino acid ligands. Protein Sci. *27*, 957–968.

Bowman, G.R., Huang, X., and Pande, V.S. (2009a). Using generalized ensemble simulations and Markov state models to identify conformational states. Methods *49*, 197–201.

Bowman, G.R., Beauchamp, K.A., Boxer, G., and Pande, V.S. (2009b). Progress and challenges in the automated construction of Markov state models for full protein systems. J. Chem. Phys. *131*, 124101.

Buch, I., Giorgino, T., and De Fabritiis, G. (2011). Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. Proc. Natl. Acad. Sci. U S A *108*, 10184–10189.

Chu, B.C.H., Chan, D.I., DeWolf, T., Periole, X., and Vogel, H.J. (2014). Molecular dynamics simulations reveal that apo-HisJ can sample a closed conformation. Proteins *82*, 386–398.

Clifton, B.E., and Jackson, C.J. (2016). Ancestral protein reconstruction yields insights into adaptive evolution of binding specificity in solute-binding proteins. Cell Chem. Biol. *23*, 236–245.

Doerr, S., and De Fabritiis, G. (2014). On-the-fly learning and sampling of ligand binding by high-throughput molecular simulations. J. Chem. Theory Comput. *10*, 2064–2069.

Doerr, S., Harvey, M.J., Noé, F., and De Fabritiis, G. (2016). HTMD: high-throughput molecular dynamics for molecular discovery. J. Chem. Theory Comput. *12*, 1845–1852.

Dunn, M.F. (2006). Protein-ligand Interactions: General Description (ELS).

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

Felder, C.B., Graul, R.C., Lee, A.Y., Merkle, H.P., and Sadee, W. (1999). The Venus flytrap of periplasmic binding proteins: an ancient protein module present in multiple drug receptors. AAPS Pharm. Sci. *1*, E2.

Ferruz, N., Harvey, M.J., Mestres, J., and De Fabritiis, G. (2015). Insights from fragment hit binding assays by molecular simulations. J. Chem. Inf. Model. *55*, 2200–2205.

Ferruz, N., Doerr, S., Vanase-Frawley, M.A., Zou, Y., Chen, X., Marr, E.S., Nelson, R.T., Kormos, B.L., Wager, T.T., Hou, X., et al. (2018). Dopamine D3 receptor antagonist reveals a cryptic pocket in aminergic GPCRs. Sci. Rep. *8*, 897.

Harvey, M.J., Giupponi, G., and DeFabritiis, G. (2009). ACEMD: accelerating biomolecular dynamics in the microsecond time scale. J. Chem. Theory Comput. *5*, 1632–1639.

Huang, P.S., Boyken, S.E., and Baker, D. (2016). The coming of age of de novo protein design. Nature *537*, 320–327.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. *14*, 33–38.

Joosten, R.P., Joosten, K., Murshudov, G.N., and Perrakis, A. (2012). PDB_REDO: constructive validation, more than just looking for errors. Acta Crystallogr. D Biol. Crystallogr. *68*, 484–496.

Keserü, G.M., and Swinney, D.C. (2015). Thermodynamics and Kinetics of Drug Binding (Wiley-VCH Verlag GmbH & Co. KGaA).

Lee, J.Y., Krieger, J., Herguedas, B., García-Nafría, J., Dutta, A., Shaikh, S.A., Greger, I.H., and Bahar, I. (2019). Druggability simulations and X-ray crystallography reveal a ligand-binding site in the GluA3 AMPA receptor N-terminal domain. Structure *27*, 241–252.

De Lorimier, R.M., Smith, J.J., Dwyer, M.A., Looger, L.L., Sali, K.M., Paavola, C.D., Rizk, S.S., Sadigov, S., Conrad, D.W., Loew, L., et al. (2009). Construction of a fluorescent biosensor family. Protein Sci. *11*, 2655–2675.

Mächtel, R., Narducci, A., Griffith, D.A., Cordes, T., and Orelle, C. (2019). An integrated transport mechanism of the maltose ABC importer. Res. Microbiol. *170*, 321–337.

Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., and Simmerling, C. (2015). ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J. Chem. Theory Comput. *11*, 3696–3713.

Marvin, J.S., Schreiter, E.R., Echevarría, I.M., and Looger, L.L. (2011). A genetically encoded, high-signal-to-noise maltose sensor. Proteins Struct. Funct. Bioinform. *79*, 3025–3036.

Marvin, J.S., Borghuis, B.G., Tian, L., Cichon, J., Harnett, M.T., Akerboom, J., Gordus, A., Renninger, S.L., Chen, T.W., Bargmann, C.I., et al. (2013). An optimized fluorescent probe for visualizing glutamate neurotransmission. Nat. Methods *10*, 162–170.

Marvin, J.S., Scholl, B., Wilson, D.E., Podgorski, K., Kazemipour, A., Müller, J.A., Schoch, S., Quiroz, F.J.U., Rebola, N., Bao, H., et al. (2018). Stability, affinity, and chromatic variants of the glutamate sensor iGluSnFR. Nat. Methods *15*, 936–939.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

Mueller, U., Förster, R., Hellmig, M., Huschmann, F.U., Kastner, A., Malecki, P., Pühringer, S., Röwer, M., Sparta, K., Steffien, M., et al. (2015). The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: current status and perspectives. Eur. Phys. J. Plus *130*, 141–150.

Paul, S., Banerjee, S., and Vogel, H.J. (2017). Ligand binding specificity of the *Escherichia coli* periplasmic histidine binding protein, HisJ. Protein Sci. *26*, 268–279.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: machine learning in python. J. Mach. Learn. Res. *12*, 2825–2830.

Pérez-Hernández, G., Paul, F., Giorgino, T., and De Fabritiis, G. (2013). Identification of slow molecular order parameters for Markov model construction. J. Chem. Phys. *139*, 015102.

Piletz, J.E., Aricioglu, F., Cheng, J.T., Fairbanks, C.A., Gilad, V.H., Haenisch, B., Halaris, A., Hong, S., Lee, J.E., Li, J., et al. (2013). Agmatine: Clinical applications after 100 years in translation. Drug Discov. Today *18*, 880–893.

Plattner, N., and Noé, F. (2015). Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models. Nat. Commun. *6*, 7653.

Prinz, J.-H., Wu, H., Sarich, M., Keller, B., Senne, M., Held, M., Chodera, J.D., Schütte, C., and Noé, F. (2011). Markov models of molecular kinetics: generation and validation. J. Chem. Phys. *134*, 174105.

Quiocho, F.A., and Ledvina, P.S. (1996). Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: Variation of common themes. Mol. Microbiol. *20*, 17–25.

Raasch, W., Schäfer, U., Chun, J., and Dominiak, P. (2001). Biological significance of agmatine, an endogenous ligand at imidazoline binding sites. Br. J. Pharmacol. *133*, 755–780.

Satriano, J. (2004). Arginine pathways and the inflammatory response: Interregulation of nitric oxide and polyamines: Review article. Amino Acids *26*, 321–329.

Scheib, U., Shanmugaratnam, S., Farías-Rico, J.A., and Höcker, B. (2014). Change in protein-ligand specificity through binding pocket grafting. J. Struct. Biol. *185*, 186–192.

Schiebel, J., Gaspari, R., Wulsdorf, T., Ngo, K., Sohn, C., Schrader, T.E., Cavalli, A., Ostermann, A., Heine, A., and Klebe, G. (2018). Intriguing role of water in protein-ligand binding studied by neutron crystallography on trypsin complexes. Nat. Commun. *9*, 3559.

Schreiber, G., and Keating, A.E. (2011). Protein binding specificity versus promiscuity. Curr. Opin. Struct. Biol. *21*, 50–61.

Silva, D.-A., Bowman, G.R., Sosa-Peinado, A., and Huang, X. (2011). A role for both conformational selection and induced fit in ligand binding by the Lao protein. PLoS Comput. Biol. *7*, e1002054.

Sparta, K.M., Krug, M., Heinemann, U., Mueller, U., and Weiss, M.S. (2016). Xdsapp2.0. J. Appl. Cryst. *49*, 1085–1092.

Stejskal, E.O., and Tanner, J.E. (1965). Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. J. Chem. Phys. *42*, 288–292.

Tang, C., Schwieters, C.D., and Clore, G.M. (2007). Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. Nature *449*, 1078–1082.

Tiwari, M.K., Singh, R., Singh, R.K., Kim, I.-W., and Lee, J.-K. (2012). Computational approaches for rational design of proteins with novel functionalities. Comput. Struct. Biotechnol. J. *2*, e201209002.

Vanommeslaeghe, K., and MacKerell, A.D., Jr. (2012). Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. J. Chem. Inf. Model. *52*, 3144–3154.

Vassylyev, D.G., Tomitori, H., Kashiwagi, K., Morikawa, K., and Igarashi, K. (1998). Crystal structure and mutational analysis of the Escherichia coli putrescine receptor. Structural basis for substrate specificity. J. Biol. Chem. *273*, 17604–17609.

Vergara, R., Romero-Romero, S., Velazquez-Lopez, I., Espinoza-Perez, G., Rodrieguez-Hernandez, A., Pulido, N.O., Sosa-Peinado, A., and Fernandez-Velasco, D.A. (2020). The interplay of protein-ligand and water-mediated interactions shape affinity and selectivity in LAO binding protein. FEBS J. *287*, 763–782.

Weber, M., and Kube, S. (2005). Robust Perron Cluster Analysis for Various Applications in Computational Life Science (Berlin, Heidelberg: Springer), pp. 57–66.

Wilson, C.J. (2015). Rational protein design: Developing next-generation biological therapeutics and nanobiotechnological tools. Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol. *7*, 330–341.

Woller, P.B., and Garbisch, E.W. (1972). The conformational analysis of n-Butane. J. Am. Chem. Soc. *94*, 5310–5314.

Wu, D.H., Chen, A., and Johnson, C.S. (1995). An improved diffusion-ordered spectroscopy experiment incorporating bipolar-gradient pulses. J. Magn. Reson. Ser. A *115*, 260–264.

STAR★METHODS

KEY RESOURCES TABLE

Reagent or Resource	Source	Identifier
Bacterial and virus strains		
E. coli BL21 (DE3)	Invitrogen	n/a
E. coli Top10	Novagen	n/a
Chemicals, peptides, and recombinant proteins		
Putrescine dihydrochloride	Sigma-Aldrich	CAS: 333-93-7
Cadaverine dihydrochloride	Sigma-Aldrich	CAS: 1476-39-7
Spermidine trihydrochloride	Sigma-Aldrich	CAS: 334-50-9
Spermine	Sigma-Aldrich	CAS: 71-44-3
Agmatine sulfate	Alfa Aesar	CAS: 2482-00-0
Bicine	Alfa Aesar	CAS: 150-25-4
Ammonium sulfate	Sigma-Aldrich	CAS: 7783-20-2
Jeffamine M600	Fluka	Order No.: 09303
PEG 3350	Sigma-Aldrich	CAS: 25322-68-3
Ammonium ^{15}N chloride	Sigma-Aldrich	CAS: 3946-62-1
PotF	This paper	UniProt: P31133
Deposited data		
PotF apo open	This paper	PDB: 6YED
PotF apo closed	This paper	PDB: 6YEB
PotF-Putrescine-complex	Vassylyev et al., 1998	PDB: 1A99
PotF-Putrescine-complex	This paper	PDB: 6YE0
PotF-Cadaverine-complex	This paper	PDB: 6YE7
PotF-Spermidine-complex	This paper	PDB: 6YE8
PotF-Spermine-complex	This paper	PDB: 6YEC
PotF-Agmatine-complex	This paper	PDB: 6YE6
Recombinant DNA		
pET21b(+) - PotF	Scheib et al., 2014	n/a
Oligonucleotides		
PotF_forward_primer TTTCATATGGCTGAACAAAAAA CACTCCAC	Sigma-Aldrich	n/a
PotF_reverse_primer AAACTCGAGTTTTCCGCTCTTCACTT TGGT	Sigma-Aldrich	n/a
Software and algorithms		
NanoAnalyze 3.10.0	TA Instruments	https://www.tainstruments.com/
XDSAPP 2.0	HZB, Helmholtz-Center Berlin	Sparta et al., 2016
XSCALE 20180126	XDS package	http://xds.mpimf-heidelberg.mpg.de/
Phenix package 1.13-2998	Adams et al., 2010	http://www.phenix-online.org/
Coot 0.89	Emsley et al., 2010	https://www2.mrc-lmb.cam.ac.uk/
PyMOL 2.1.0	Schrödinger	https://www.pymol.org/
PDB_REDO 6.00	Joosten et al., 2012	https://pdb-redo.org/
NMRViewJ 9.2.0	n/a	https://nmrfx.org/nmrfx/nmrviewj
ACEMD	Harvey et al., 2009	https://pubs.acs.org/doi/10.1021/ct9000685
VMD	Humphrey et al., 1996	http://www.ks.uiuc.edu/Research/vmd/

Structure
Article

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Birte Höcker (birte.hoecker@uni-bayreuth.de).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The datasets generated during this study are available at the PDB under the accession codes 6YED (PotF apo open), 6YEB (PotF apo closed), 6YE0 (PotF:PUT), 6YE7 (PotF:CDV), 6YE8 (PotF:SPD), 6YEC (PotF:SPM), 6YE6 (PotF:AGM).

Complete simulation data and code for analysis are available upon request due to size limitations.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial strains

- Escherichia Coli Top10: mcrA, Δ(mrr-hsdRMS-mcrBC), Phi80lacZ(del)M15, ΔlacX74, deoR, recA1, araD139, Δ(ara-leu)7697, galU, galK, rpsL(SmR), endA1, nupG (Invitrogen)
- Escherichia Coli BL21(DE3): E. coli B dcm ompT hsdS($r_B^-m_B^-$) gal (Novagen)

The strains were grown in LB media at 37 °C in shaker flasks.

METHOD DETAILS

Cloning, expression & purification

The gene for PotF (UniProt ID P31133) missing the N-terminal signal peptide was amplified by PCR using genomic DNA of *E. coli* K-12 and primers, which introduced flanking restriction sites for NdeI and XhoI. All cloning work was performed in Top10 cells. After digestion the DNA fragment was ligated into a pET21b(+)-vector, which added a C-terminal His6-tag. BL21 (DE3) cells were transformed with the plasmid and plated out on LB agar plates containing 100 μg/ml ampicillin as selection marker.

For protein expression 2 L LB were inoculated with 20 mL of an overnight culture containing 100 μg/ml ampicillin and incubated at 37 °C until OD_{600} reached a value of ~0.7. Overexpression was induced by adding isopropyl-ß-thiogalactoside (IPTG) to a final concentration of 1 mM and further incubation for 4 h at 37 °C. Cells were harvested by centrifugation (Beckman Coulter, JLA 8.1000 4000 g, 20 min, 4°C) and pellets were washed with 30 mL buffer (50 mM Tris pH 8, 300 mM NaCl, 20 mM Imidazole).

Resuspended cells were lysed by sonication (Branson 6.3 mm tip, 2 x 3 min, 40% duty cycle, output power 4), followed by a centrifugation step (Beckman Coulter, JA25.50, 40.000 g, 60 min) to separate the soluble from the insoluble fraction and cell debris. The supernatant was loaded onto an equilibrated (50 mM Tris pH 8, 300 mM NaCl, 20 mM Imidazole) HisTrap HP 5 mL column (GE Healthcare) using a peristaltic pump (Pump P-1, GE Healthcare). After washing for 10 column volumes with lysis buffer, unfolding of the protein to wash off endogenous ligands, was achieved by washing and incubating for an hour with 6 M Guanidinium Hydrochloride (GdHCl) and subsequently washing with 10 CV of lysis buffer for refolding. The protein was eluted by a stepwise increase of the Imidazole concentration to 260 mM with an ÄKTA system. Fractions containing the protein were pooled and concentrated with a centrifugal concentrator to a maximum volume of 13 mL and applied to an equilibrated (50 mM Tris pH 8, 300 mM NaCl) preparative size exclusion column (HiLoad Superdex 75 26/60, GE Healthcare). Fractions with monomeric protein were pooled and concentrated for ITC measurements or crystallization setups. Protein concentration was checked photometrically using the absorption at 280 nm. Expression and purification were verified by SDS-PAGE.

For ^{15}N-labelled samples expression was performed in M9 minimal media supplemented with ^{15}N-Ammoniumchloride. Besides that, expression and purification protocol was kept the same as for non-labelled protein. Final buffer for NMR measurements was 20 mM Tris pH 8.0, 150 mM NaCl.

Isothermal titration calorimetry (ITC)

Ligands were always freshly prepared in the exact same buffer as was used for the size exclusion run. Samples were degassed and temperature equilibrated using a degassing station (TA Instruments). 300 μL protein sample of different concentrations depending on affinity, was transferred into the sample cell of a nanoITC (TA Instruments) and 50 μL of an 8-10-fold concentrated ligand solution was loaded into the injection needle. Multiple injection measurements were carried out at 293 K, 300 rpm stirring rate and 300 s spacings between each 2 μL injection. The heat quantity past injection was determined by integration of the measured peaks. Every protein-ligand combination was measured in duplicates. Subtraction of heat of dilution measurements, peak integration and one site binding fits were done with NanoAnalyze (TA Instruments). Reported errors are the errors between fits of the measured duplicates.

CellPress

Structure
Article

Crystallography

All crystallization experiments were set up as sitting drop vapour diffusion experiments in either 3-well Intelliplates (Art Robbins Instruments) or MRC Maxi plates (Swissci) using a protein concentration of 40 mg/mL or 15 mg/mL and a 20-fold molar excess of ligand. Protein-ligand mixtures were equilibrated at 293 K for several hours before crystallization setups.

First crystals were obtained after an initial sparse matrix screen using the commercially available JCSG CORE I-IV screens (QIAGEN). These hits were further optimized by applying the Additive Screen (Hampton Research) and then using a grid screen to improve promising conditions.

This resulted in conditions containing 2.4 M Ammonium sulfate, 0.1 M Bicine pH 8.8 and either 5% PEG 3550 or 7% Jeffamine M-600 or 4.5% Jeffamine M-600 as additive, except for the open apo structure. Here we obtained crystals in 0.1 M Tris pH 7.3, 0.2 M NaCl and 37.5% PEG 3000. Detailed information can be found in Table S2.

Crystals were mounted using CryoLoops and transferred into a cryogenic solution made of reservoir solution and either 50% Glycerol or 1.7 M Malonate pH 8.8 (Table S2) and then cooled down in liquid nitrogen. Data collection at 100 K was done at the beamlines BL 14.1 at the synchrotron BESSY II, Helmholtz-Zentrum Berlin (Mueller et al., 2015) and at PXIII at the Swiss Light Source, Villigen. Diffraction data was processed using XDSAPP (Sparta et al., 2016) and decision of resolution cut-off was made according to $CC_{1/2}$ around 0.25, I/Sigma > 0.5, and completeness in the outer shell > 75%. For the open apo structure two datasets from crystals of the same condition were merged using XSCALE to obtain better density maps at the expense of resolution. Molecular replacement with both lobes independently and missing hinge region residues of PotF (PDB: 1A99 (Vassylyev et al., 1998)) as search model, to account for different opening angles, was done using Phaser-MR (McCoy et al., 2007). Data quality was assessed utilizing phenix.xtriage. Refinements were done with Phenix.refine (Adams et al., 2010). Manual model building was performed in Coot (Emsley et al., 2010). Final models were evaluated by PDB_REDO 6.00 (Joosten et al., 2012).

Molecular dynamics simulations

System building, equilibration, and production runs

Simulation systems apo-PotF, PotF:PUT, and PotF:SPD were built and run with the software HTMD (Doerr et al., 2016). In all cases systems comprised solvated all-atom cubic boxes containing <67000 atoms. Simulation boxes consisted of a protein centered at the origin of coordinates, and a ligand –when present- positioned at least 15 Å away from the protein surface. Explicit solvent molecules and neutralizing NaCl ions were added on each box. The apo-PotF batch consisted of two sub-batches, in one closed apo-PotF served as the basis for the protein initial coordinates, while in the other open apo-PotF. All PotF:ligand complex systems started from open apo-PotF as coordinates for the protein. The ligands were parameterized with the paramchem tool (Vanommeslaeghe and MacKerell, 2012).

Ten replicas were built per simulation batch, which in the case of the PotF:ligand systems led therefore to ten different initial positions for the ligand around the protein. The final ligand concentration was 2.7 mM. All systems were minimized, equilibrated and run using ACEMD (Harvey et al., 2009) and amberff14SB as forcefield (Maier et al., 2015). Standard simulation parameters were used as described in previous work (Ferruz et al., 2015).

For the MD production, we ran an intelligent adaptive sampling scheme that performs the simulations in successive epochs by analyzing them with Markov state models (MSMs) (Doerr and De Fabritiis, 2014). The metric used during the adaptive runs for the MSMs analysis were the protein dihedrals (see below) for the apo-PotF batch and the ligand heavy atoms versus the protein alpha carbons for the PotF:ligand batches.

Markov state modeling analysis

Markov state modeling proceeds from the discretization of the conformational space and the description of the dynamics of the system of interest as a sequence of transitions between these discrete states. A properly discretized MSM shows converging timescales with high probability of transition among kinetically similar states, and lower probability between kinetically separated states. From this model, the pathways and kinetic rates between distinct conformations may be derived (Bowman et al., 2009b)

For the apo system, we run 1149 75-ns long trajectories totaling 86.2 μs of aggregated simulation time, from where 25.8 μs came from the simulations starting in the closed conformation, and 60.4 μs from the open one. The trajectories were projected onto the protein dihedral angles that represent opening and twisting, defined by the following two four-atom sets: opening: Cα of residues 27, 249, 109, and 193; and twisting: Cα of residues 36, 249, 286, and 334 (simulation numbering, for other dihedral definitions see Table S5). We projected the multidimensional data onto the slow order parameters using TICA (Pérez-Hernández et al., 2013) (time-lagged independent component analysis). After projecting the data into 2 TICA dimensions, 1000 clusters were computed using the mini batch k-means algorithm (Pedregosa et al., 2011). The clusters were lumped together into 4 macrostates by the PCCA algorithm (Weber and Kube, 2005), using a lag time of 30 ns.

For PotF:ligand complexes the trajectories were projected onto a space defined by opening and twisting dihedral angles, contacts between ligand nitrogen and protein alpha carbons, and a binary metric that identifies when the ligand is close to any protein residue by at least 10 Å. We projected the contact dimensions onto 2 TICA dimensions leading to 5 total dimensions. Dimensions were clustered onto 1500 microstates and MSM was with 5, 7, 5, 4, and 5 macrostates was built, respectively. The binary contact metric was used to define a bulk state. Lag times were 40, 40, 40, 30, and 30ns, respectively. The four macrostates visited in 90% of the pathway fluxes were selected for Figure S5.

Thermodynamics and kinetics were computed as follows. After TICA discretization, the master equation is then built as:

$$P_i(t) = \sum_{j=1}^{N} [k_{ij}P_j(t) - k_{ji}P_i(t)] = K_{ij}P_j(t),$$

where $P_i(t)$ is the probability of state i at time t, and k_{ij} are the transition rates from j to i, and K = (K_{ij}) is the rate matrix with elements K_{ij} = k_{ij} for i ≠ j and $K_{ii} = -\sum_{j \neq i} k_{ji}$. The master equation d**P**/dt = K P has solution with initial condition P(0) given by P(t) = T(t) P(0), where the transition probability matrix is defined as: $T_{ij}(t) = (\exp[\mathbf{K}t])_{ij} = p(i,t|j,0)$, i.e. the probability of being in state i at time t, given that the system was in state j at time 0. In practical terms, $p_{ij}(\Delta t)$ is estimated from the simulation trajectories for a given lag time Δt using a maximum likelihood estimator compatible with detailed balance. The eigenvector π with eigenvalue 1 of the matrix $_{yu0poy78ihj,}$ corresponds therefore to the stationary, equilibrium probability, and thus the thermodynamics of the process. On the other hand, first-order kinetics are derived from mean-first passage times (MFPT) (Buch et al., 2011): Where k_{on}, measured in units of $M^{-1}s^{-1}$, is inversely dependent on the ligand concentration, in this case 0.0027 M. These analyses were performed with HTMD (Doerr et al., 2016).

NMR spectroscopy

Purified ^{15}N-labelled protein samples were concentrated to 200-400 μM final concentration. After addition of 10 % D_2O for frequency lock, samples were transferred into a 5 mm NMR tube. ^{1}H-^{15}N-HSQC spectra were recorded at 293 K on a 700MHz Bruker Avance III HD spectrometer equipped with a cryogenically cooled probe. Stock solutions of the ligands were prepared in the same gel filtration buffer the protein was purified. Ligands were added after an initial apo measurement in saturating concentration according to the determined affinities by ITC.

Translational diffusion was measured on a 700 MHz Bruker Avance III HD NMR spectrometer. Gradient strength was calibrated using a doped water sample (2mM $CuSO_4$, 1 % H_2O in D_2O) assuming a diffusion coefficient of 1.90 × 10^{-9} m^2/s at a calibrated temperature of 298 K using a 1D 1H pulse gradient stimulated echo with bipolar gradients (Wu et al., 1995). For protein measurements a calibrated temperature of 308 K was used and the 3-9-19 watergate water suppression was applied in the diffusion experiment. The samples consisted of 200 μM unlabeled PotF in 20 mM Tris, pH 7.5, 150 mM NaCl. In presence of ligands their concentration was adjusted to achieve full saturation of the protein, i.e. 3-fold excess for PUT and 20-fold excess SPM. The integrated signal intensity of the methyl group region (-0.5 – 1.25 ppm) was observed as a function of increased gradient strength, and analysed according to $S(Q) = S(0) * \exp\{-DQ\}$ with $Q = \gamma^2 g^2 \delta^2 (\Delta - \delta/3 - \tau/2)$ (γ = gyromatic ratio, g = gradient strength, δ = gradient pulse length, Δ = diffusion delay, τ = gradient recovery delay (Stejskal and Tanner, 1965). For each sample three independent measurements were performed and the corresponding intensities averaged before data fitting. NMR spectra were visualized using NMRViewJ.

QUANTIFICATION AND STATISTICAL ANALYSIS

ITC experiments were analyzed using NanoAnalyze (TA Instruments). Binding curves were derived using the built-in one-site-binding fit. Each system was measured twice and the fitted parameters and their standard deviation are given in Table 2.

Error estimation for MD equilibrium populations and kinetic conversions was done as follows:

We estimated errors for all properties using a bootstrapping technique. We performed 10 independent runs in which 20% of the trajectories were randomly eliminated and a new MSM was built after re-clustering.

Supplementary Information for

A comprehensive binding study illustrates ligand recognition in the periplasmic binding protein PotF

by Pascal Kröger, Sooruban Shanmugaratnam, Noelia Ferruz, Kristian Schweimer, and Birte Höcker

FIGURES

Figure S1: ITC measurements, Related to Table 1

Figure S2: Overlay of ligand and water molecules in the binding site to illustrate coordination, Related to Figure 3.

Figure S3: Implied timescales for the MSM of the apo simulations, Related to Figure 4.

Figure S4: Distribution of the simulation ensembles in the dihedral space, Related to STAR Methods.

Figure S5: Summary of the simulation results for the five ligands of this work, Related to Figure 5.

Figure S6: 1H-15N-HSQC spectra, Related to Figure 6.

Figure S7: Translational diffusion measurements by pulsed field gradient NMR, Related to Figure 6.

TABLES

Table S1. Summary of the simulation batches analyzed in this work, Related to STAR Methods.

Table S2: Crystallization conditions, cryo information and protein concentrations, Related to Table 2 and STAR Methods.

Table S3: Tested ligands with ITC, Related to Table 1.

Table S4: Comparison of binding residues between X-ray and MD simulations, Realted to Figure 3 and 5.

Table S5: Atom sets used for the definition of dihedral angles for each PDB in this work,

Table S6: Opening and twisting angles for the PDB structures and simulations of this work, Related to Figure 1 and 4.

FIGURES

Figure S1: ITC measurements, Related to Table 1 with corrected heat rate in upper panels and calculated enthalpy (data points) with fit (solid line) in lower panel for PotF with PUT (a), CDV (b), SPD (c), SPM (d) and AGM (e)

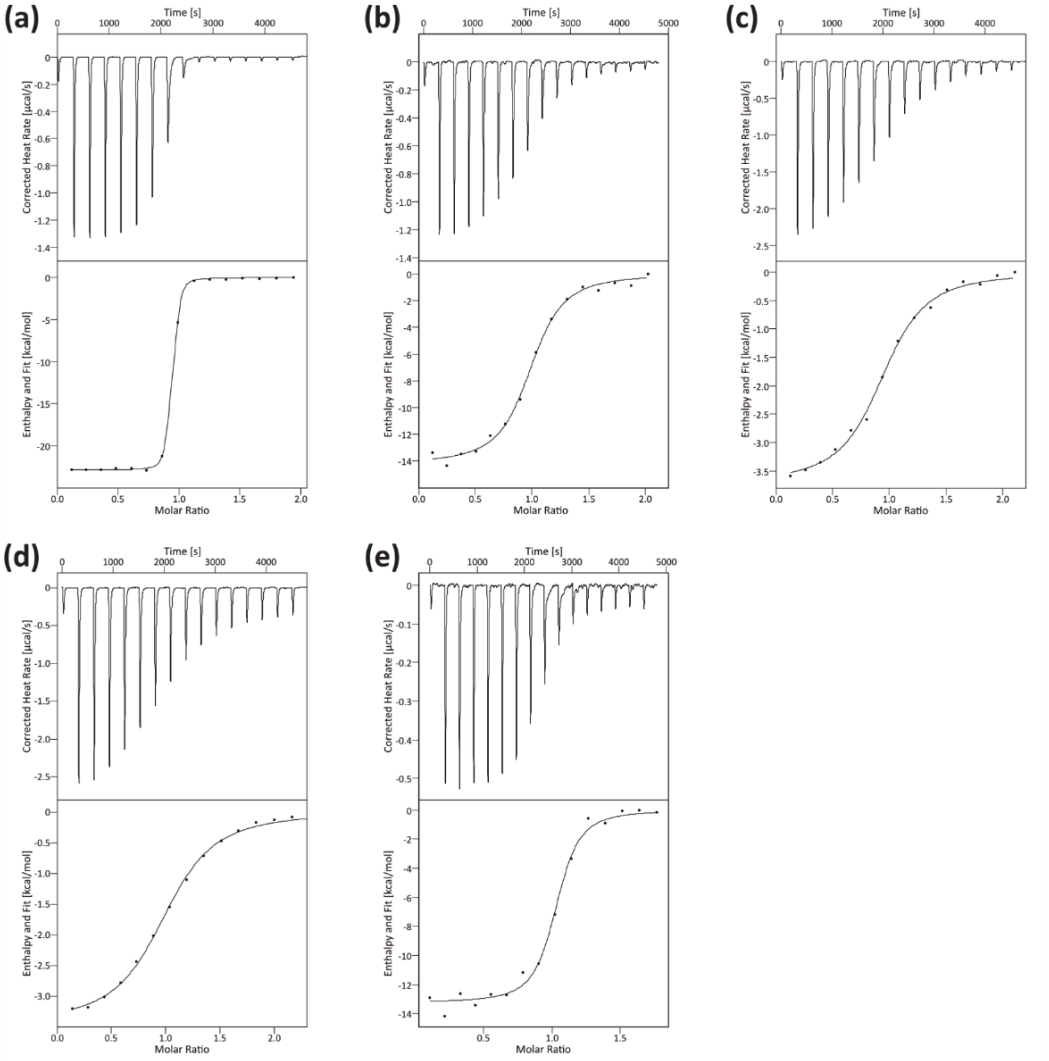


Figure S2: Overlay of ligand and water molecules in the binding site to illustrate coordination, Related to Figure 3. (a) PUT (green) and CDV (yellow) and arrows indicating the bonds used for dihedral angle determination; (b) PUT (green) and SPD (salmon); (c) PUT (green) and AGM (blue); and (d) PUT (green), CDV (yellow), SPD (salmon), AGM (blue) and water (spheres). Red arrows indicate the bonds used for dihedral angle determination.

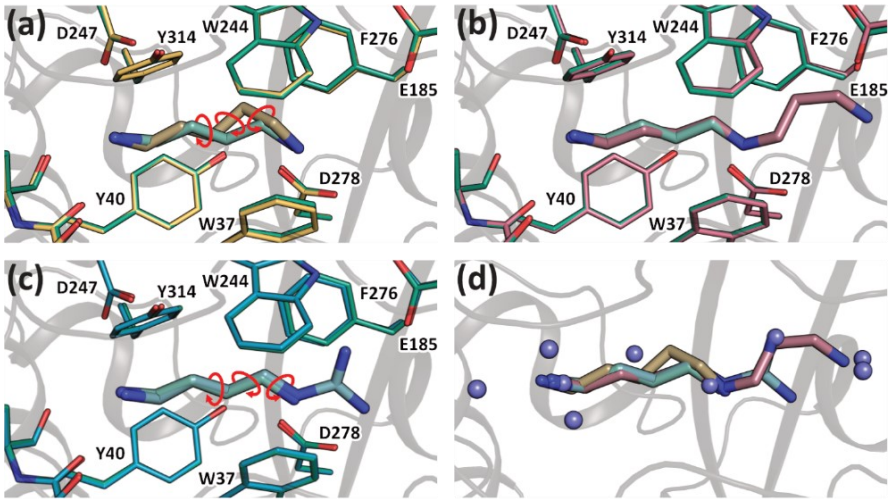


Figure S3: Implied timescales for the MSM of the apo simulations, Related to Figure 4. Lag time was chosen at 30 ns.

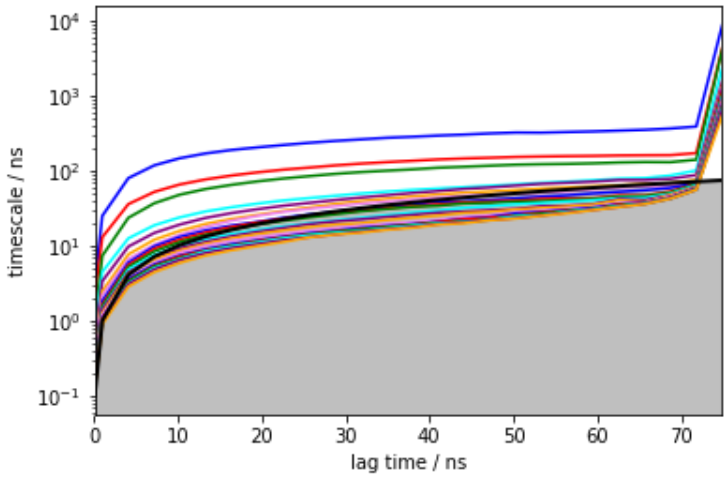


Figure S4: Distribution of the simulation ensembles in the dihedral space, Related to STAR Methods. All the ensembles contain the same amount of data (70 μs). The plots show raw data and as such are no indicative of equilibrium states.

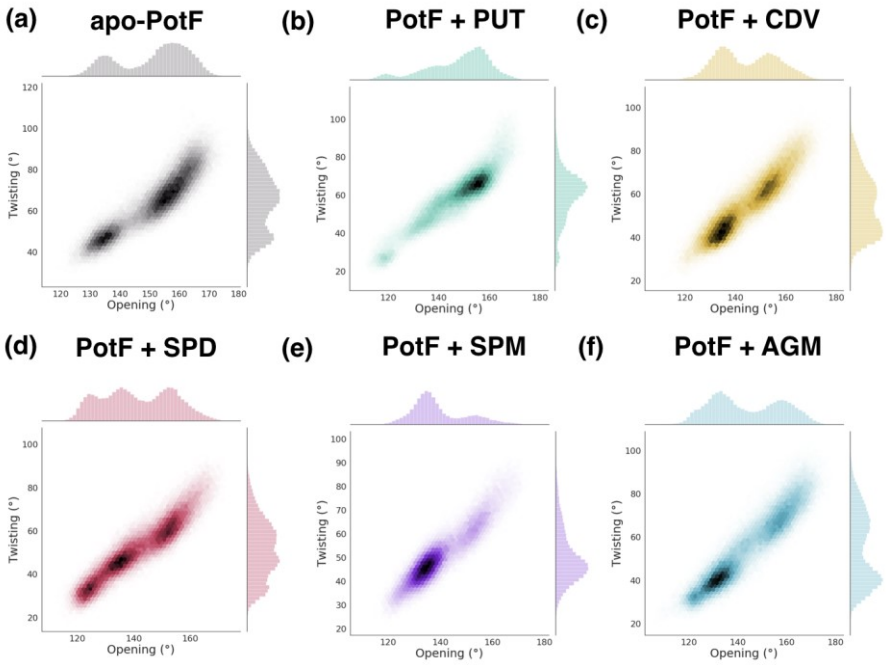


Figure S5: Summary of the simulation results for the five ligands of this work, Related to Figure 5. For each binding process (PUT (a), CDV (b), SPD (c), SPM (d), and AGM (e)) we show the projection of the macrostates onto the dihedral subspace, along with the mean opening and twisting angles for bound and encounter 1 and 2; the RMSDs for each state; the conformations of the four states (bulk, encounter 1 and 2, and bound) and the major 2 pathways of binding.

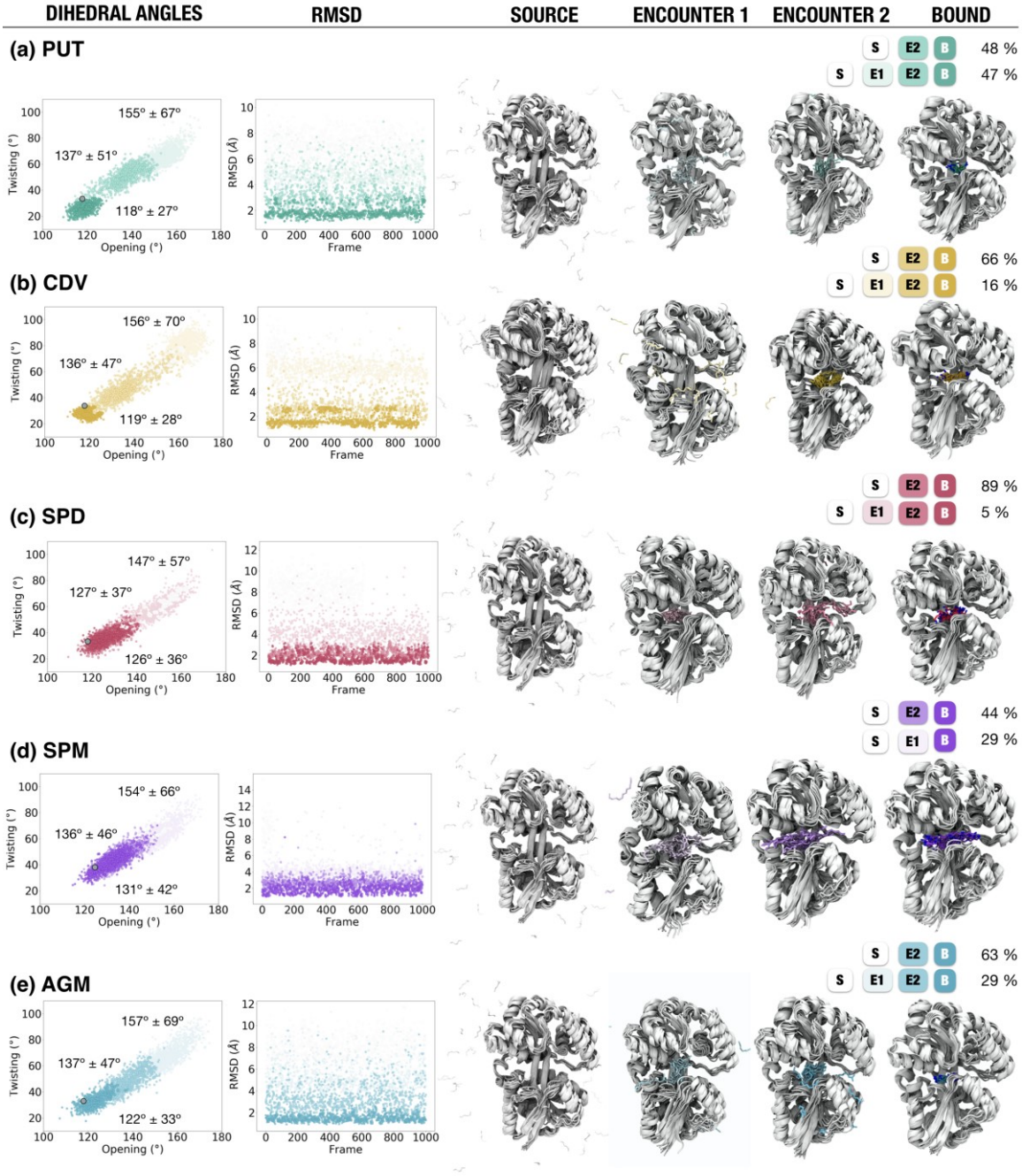


Figure S6: 1H-15N-HSQC spectra, Related to Figure 6. apo PotF is shown in black in all panels, PotF:PUT in green (a), PotF:CDV in yellow (b), PotF:AGM in blue (c), PotF:SPD in salmon (d), and PotF:SPM in purple (e).

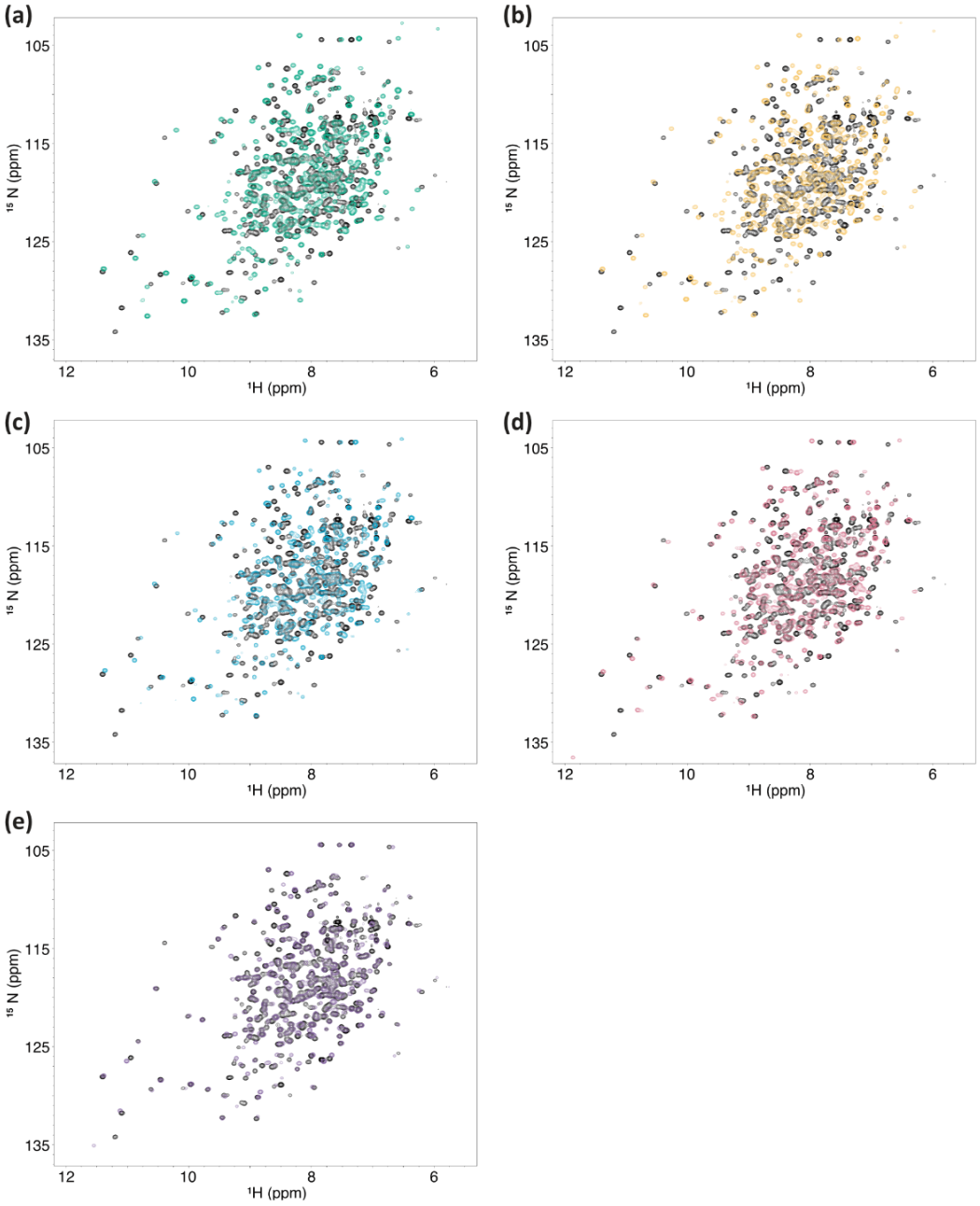
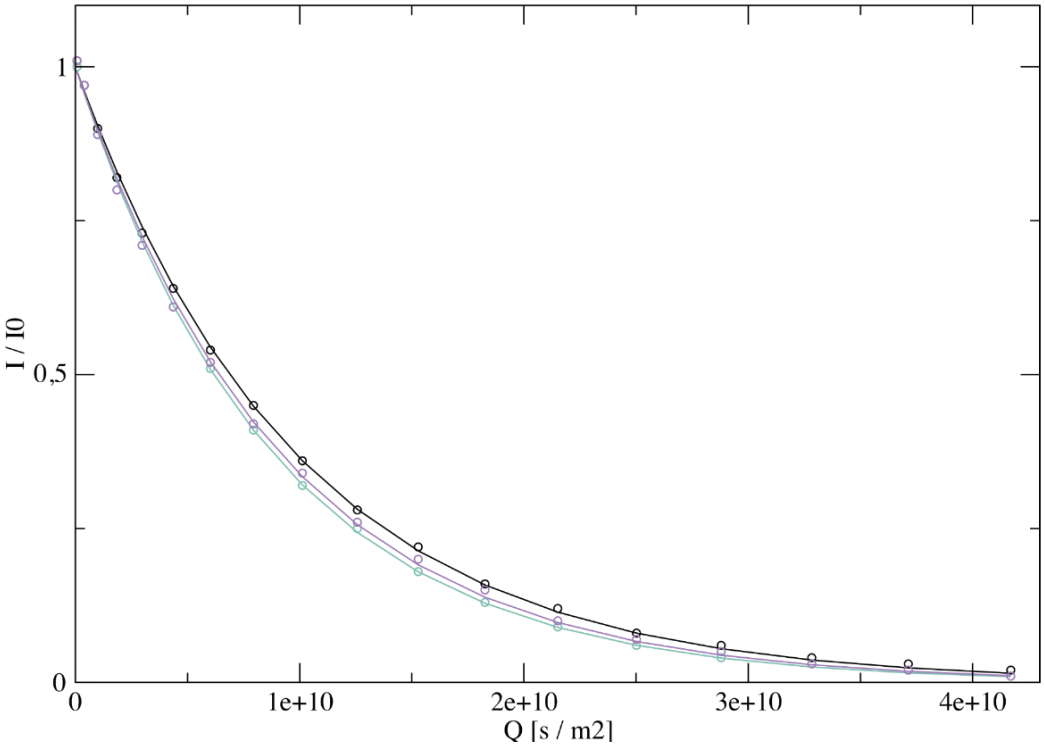


Figure S7: Translational diffusion measurements by pulsed field gradient NMR, Related to Figure 6. PotF in absence (black) and presence of PUT (green) or SPM (purple). The normalized integrated signal intensity of the methyl region is shown as function of Q = $\gamma^2 g^2 \delta^2 (\Delta - \delta/3 - \tau/2)$ (γ = gyromatic ratio, g = gradient strength, δ = gradient pulse length, Δ = diffusion delay, τ = gradient recovery delay).



Tables

Table S1. Summary of the simulation batches analyzed in this work, Related to STAR Methods.

System	Simulation time (μs)
apo-PotF	85.7
PotF:PUT	86.9
PotF:CDV	108.4
PotF:SPD	153.3
PotF:SPM	72.9
PotF:AGM	88.5
Total	**595.7**

Table S2: Crystallization conditions, cryo information and protein concentrations, Related to Table 2 and STAR Methods.

Protein concentration	Conformation or Ligand	Condition	Additive	Cryo
15 mg/mL	“open”	0.1 M Tris pH 7.3, 0.2 M NaCl, 37.5% PEG 3000	-	-
40 mg/mL	“closed”	2.4 M $(NH_4)_2SO_4$, 0.1M Bicine pH 8.8	5% PEG 3550	25% Glycerol
15 mg/mL	Putrescine	2.4 M $(NH_4)_2SO_4$, 0.1M Bicine pH 8.8	4.5% Jeffamine M600 pH 7	50% 3.4 M Malonate pH 8.8
40 mg/mL	Agmatine	2.4 M $(NH_4)_2SO_4$, 0.1M Bicine pH 8.8	7% Jeffamine M600 pH 7	25% Glycerol
15 mg/mL	Spermidine	2.4 M $(NH_4)_2SO_4$, 0.1M Bicine pH 8.9	10% Jeffamine M600 pH 7	50% 3.4 M Malonate pH 8.8
40 mg/mL	Spermine	2.4 M $(NH_4)_2SO_4$, 0.1M Bicine pH 8.8	5% PEG 3550	50% Glycerol
40 mg/mL	Cadaverine	2.4 M $(NH_4)_2SO_4$, 0.1M Bicine pH 8.8	5% PEG 3550	50% 3.4 M Malonate pH 8.8

Table S3: Tested ligands with ITC, Related to Table 1.

Ligand	Abbreviation	Molecule type	Binding capability
Lysine	LYS	Amino acid, carboxylated cadaverine	-
Agmatine	AGM	Biogenic amine, neuromodulator	High
Spermidine	SPD	Biogenic amine related to putrescine, PotD ligand	Moderate
Putrescine	PUT	Biogenic amine, $PotF^{wt}$-ligand	Very high
Serotonin	SERO	Biogenic amine, neurotransmitter	-
Cadaverine	CDV	Biogenic amine, very similar to putrescine	High
γ-Aminobutyric acid	GABA	Non-proteinogenic amino acid	-
Arginine	ARG	Amino acid, carboxylated agmatine	-
Ornithine	ORN	Amino acid, carboxylated putrescine	-
Tyramine	TYRA	Biogenic amine	-
Spermine	SPM	Biogenic amine related to spermidine and putrescine, longest ligand tested	Moderate

Table S4: Comparison of binding residues between X-ray and MD simulations, Realted to Figure 3 and 5. X-ray column shows the 12 closest residues to the ligand. Bound MD ensemble shows the 12 residues that more often are within 8 Å of any ligand heavy atom.

	X-Ray	Bound MD Ensemble
PUT	Trp37, Ser38, Asp39, Tyr40, Asn65, Ser85, Trp244, Asp247, Gln250, Phe276, Asp278, Tyr314	Trp37, Ser38, Tyr40, Pro84, Ser85, Ser226, Ser227, Trp244, Asp247, Phe276, Phe277, Asp278
CDV	Trp37, Ser38, Asp39, Tyr40, Asn65, Ser85, Trp244, Asp247, Gln250, Phe276, Asp278, Tyr314	Trp37, Ser38, Tyr40, Pro84, Ser85, Trp135, Trp244, Gly245, Asp247, Phe276, Phe277, Asp278
SPD	Trp37, Ser38, Tyr40, Asn65, Ser85, Glu185, Trp244, Asp247, Met274, Phe276, Asp278, Tyr314	Trp37, Ser38, Asp39, Tyr40, Ser85, Ala182, Glu185, Gly245, Asp247, Phe276, Phe277, Asp278
SPM	Trp37, Ser38, Asp39, Asn65, Ser85, Glu185, Ile230, Trp244, Asp247, Phe276, Asp278, Tyr314	Trp37, Ser38, Asp39, Tyr40, Asp63, Ser85, Ala182, Glu185, Asp247, Phe276, Phe277, Asp278
AGM	Trp37, Ser38, Asp39, Tyr40, Asn65, Ser85, Glu185, Trp244, Asp247, Phe276, Asp278, Tyr314	Trp37, Ser38, Asp39, Tyr40, Pro84, Ser85, Trp244, Gly245, Asp247, Phe276, Phe277, Asp278

Table S5: Atom sets used for the definition of dihedral angles for each PDB in this work, Related to STAR Methods and Figure 1. Numbers refer to alpha carbon atoms. Numbering remains identical for PDBs where chain B is also present. Figure 1 shows dihedrals for chain A.

PDB	Opening	Twisting
1A99	54 276 136 220	63 276 313 361
4JDF	54 276 136 220	63 276 313 361
1POT	51 255 132 205	60 255 292 339
6YED (PotF apo open)	54 276 136 220	63 276 313 361
6YEB (PotF apo closed)	54 276 136 220	63 276 313 361
6YE0 (PotF:PUT)	54 276 136 220	63 276 313 361
6YE7 (PotF:CDV)	54 276 136 220	63 276 313 361
6YE8 (PotF:SPD)	54 276 136 220	63 276 313 361
6YEC (PotF:SPM)	54 276 136 220	63 276 313 361
6YE6 (PotF:AGM)	54 276 136 220	63 276 313 361

Table S6: Opening and twisting angles for the PDB structures and simulations of this work, Related to Figure 1 and 4. Chain A and B are considered. Atoms for the computation of angles are listed in **Table S5**.

PDB	Opening/Twisting Angles chain A	Opening/Twisting Angles chain B	Opening/Twisting Angles MSMs
6YED (PotF apo open)	157.2º 69.9º	155.4º 67.1º	see Fig. 3 and text
6YEB (PotF apo closed)	118.4º 33.3º	119.8º 32.9º	see Fig. 3 and text
6YE0 (PotF:PUT)	117.7º 33.4º	118.8º 32.7º	118.6 ± 2.9 26.6 ± 3.2
6YE7 (PotF:CDV)	117.9º 33.3º	119.5º 32.8º	119.46 ± 2.7 28.2 ± 2.9
6YE8 (PotF:SPD)	117.8º 33.1º	119.0º 33.2º	126.46 ± 5.0 35.69 ± 5.2
6YEC (PotF:SPM)	124.6º 38.1º	126.6º 38.7º	132.85 ± 5.8 43.38 ± 6.6
6YE6 (PotF:AGM)	117.7º 33.9º	118.9º 32.8º	122.62 ± 4.3 33.07 ± 4.6

9. Manuscript 2

JBC RESEARCH ARTICLE

Fine-tuning spermidine binding modes in the putrescine binding protein PotF

Received for publication, September 9, 2021, and in revised form, November 10, 2021 Published, Papers in Press, November 19, 2021,
https://doi.org/10.1016/j.jbc.2021.101419

Pascal Kröger[1], **Sooruban Shanmugaratnam**[1,2], **Ulrike Scheib**[2], **and Birte Höcker**[1,2,*]

From the [1]*Department for Biochemistry, University of Bayreuth, Bayreuth, Germany;* [2]*Max Planck Institute for Developmental Biology, Tübingen, Germany*

Edited by Wolfgang Peti

A profound understanding of the molecular interactions between receptors and ligands is important throughout diverse research, such as protein design, drug discovery, or neuroscience. What determines specificity and how do proteins discriminate against similar ligands? In this study, we analyzed factors that determine binding in two homologs belonging to the well-known superfamily of periplasmic binding proteins, PotF and PotD. Building on a previously designed construct, modes of polyamine binding were swapped. This change of specificity was approached by analyzing local differences in the binding pocket as well as overall conformational changes in the protein. Throughout the study, protein variants were generated and characterized structurally and thermodynamically, leading to a specificity swap and improvement in affinity. This dataset not only enriches our knowledge applicable to rational protein design but also our results can further lay groundwork for engineering of specific biosensors as well as help to explain the adaptability of pathogenic bacteria.

Biomolecular recognition and discrimination are crucial for many biological functions. Binding interactions ranging from highly specific to promiscuous determine the regulation and functioning of a multitude of parallel cellular processes (1, 2). Periplasmic binding proteins (PBPs) are versatile bacterial nonenzymatic receptors that sense a range of different solutes, like carbohydrates, amino acids, vitamins, and ions (3). This large ligand diversity is reflected in a high sequence diversity within the superfamily, although the general structure of PBPs is conserved (4). They consist of two β/α-lobes connected *via* a hinge region with the ligand-binding site located at the interface of the lobes. PBPs are predominantly open in solution and close by undergoing a large conformational change upon ligand recognition, which is often compared to a Venus flytrap (5). These receptors work hand in hand with bacterial prokaryotic-type ABC transporters, since they require additional solute-binding proteins to recruit substrates (6).

The binding-induced change of the overall structure of PBPs might be the prerequisite for the adaptability and large coverage of different ligands by PBPs. The two conformations enable the positioning of ligands at the solvent-excluded and low dielectric protein interior (closed) while still allowing binding site residues to be placed at the evolving sites of protein surfaces (open) (4).

To gain insight into the evolution and designability of ligand binding and specificity, we analyzed the homologous PBPs, PotF and PotD, which are *Escherichia coli* putrescine (PUT)-binding protein and spermidine (SPD)-binding protein, respectively. They constitute the first elements of two separate multicomponent uptake systems (PotFGHI and PotABCD) to transport the cationic polyamines PUT and SPD across the cellular membrane (7, 8). PotF and PotD only share 35% overall sequence identity, but their binding pockets as well as their respective ligands encompass a high structural similarity. Still, the protein-binding modes differ: PotF shows affinity for PUT and SPD, whereas PotD exclusively binds SPD. Prior to this work, we grafted the seven differences in amino acids from PotD binding pocket onto PotF (9); this resulted in PotF_SPD, which will be referred to as PotF/D throughout this work to keep naming for mutants more concise. PotF/D solely binds SPD, whereas PUT affinity is abolished. Nonetheless, this variant did not show an affinity for SPD as high as PotD. A structural intriguing feature of PotF/D (Protein Data Bank [PDB]: 7OYZ) is the semiclosed conformation of the ligated crystal structure. We only observed PotF to adopt a similar semiclosed conformation when binding the large polyamine spermine (PDB: 6YEC). The size of this ligand disrupts two salt bridges flanking the binding pocket (D_{39}–R_{254} and R_{91}–E_{184}), which were deemed important for tight binding and ligand affinity (10). Upon closer inspection of the seven mutations in PotF/Ds crystal structure, two exchanges (D39E and S87Y) were identified to influence the wildtype-like salt bridge formation of PotFs between the lobes.

In our previous study on PotF (10), we deduced that binding events in PBPs are not narrowed down to two fixed conformations and a final ligand pose but are more dynamic and must be analyzed considering the overall protein dynamic as well. The observation that SPD affinity was solely maintained and PotF/Ds crystal structure adopted a semiclosed state led us to the approaches presented in this work. We conducted a combinatorial test of the sequence space between PotF-binding and PotF/D-binding pockets by dividing the pocket into three parts (proximal, distal, and aromatic box) to determine disadvantageous and beneficial residue changes. In

* For correspondence: Birte Höcker, birte.hoecker@uni-bayreuth.de.

addition, we approached the effects of more global changes on affinity by a stepwise reintroduction and reestablishment of the salt bridges. For both approaches, we utilized isothermal titration calorimetry (ITC) and X-ray crystallography as methods of analysis. Furthermore, we illustrate how a promiscuous PBP can quickly adapt high specificity and discuss the impact this phenomenon could have on pathogens like multidrug-resistant gram-negative bacteria.

Results

Polyamine-binding pockets in PotF and PotD

Despite great similarities among their binding pockets, PotF and PotD differ in their polyamine-binding profiles: PotF is promiscuous for PUT ($K_D \approx 68$ nM) and SPD ($K_D \approx 30$ μM) (10), whereas PotD binds exclusively SPD ($K_D \approx 6$ nM) (9). We showed previously that this specificity can be swapped by exchanging respective residues in the binding pocket of PotF against their PotD counterparts (9). This resulted in PotF/D, which solely binds SPD ($K_D \approx 37$ μM; Table 1). The mentioned PotF/D:SPD–K_D differs somewhat from our previously published one because of a changed setup (different ITC device and buffer conditions). Therefore, we remeasured all binding constants to keep comparability extremely high during the whole dataset of this study and to be able to interpret also slight differences in affinity and thermodynamics. Nonetheless, PotF/D does not reveal the specific interactions that direct the ligand-binding profiles of PotF and PotD for the different polyamines. To understand why this is the case, we analyzed the contributions to specificity from first shell mutations in the binding pocket using targeted mutagenesis and ITC. Comparison of the binding pockets of PotF and PotD suggests three distinct groups based on their location and ligand interaction (Table S1). Residues of the primary amine-binding site interacting with N1 of PUT and SPD in PotF were assigned to the proximal group (S38, D39, and D247; residues and numbering according to the PotF sequence without periplasmic signal peptide). Aromatic residues, which anchor the methylene backbone and N2 of the ligands *via* van der Waals, CH–π, and cation–π interactions, are assigned to the central aromatic box (W37, W244, and F276). Finally, residues forming direct interactions with N2 of PUT and N3 of SPD in PotF as well as equivalent residues forming interactions with N3 of SPD in PotD are assigned to the distal group (S85, S87, A182, E185, D278, and L348; Fig. 1). Aromatic box residues are conserved in both PBPs, the only difference is F276 in PotF that is substituted by tryptophan (W255) in PotD. Similarly, both proteins tightly coordinate the proximal primary amine of their ligands. The main differences between both proteins are found at the distal side. In PotD, polar residues (S83, Y85, D168, E171, and Q327) form ionic and hydrophilic contacts to the cationic N3 of SPD, whereas the distal binding pocket of PotF is slightly more hydrophobic in its characteristics (S85, S87, A182, E185, and L348) and shows a tightly coordinated water network (10). Three variants were generated based on the clustering of the active site: PotF_Prox (S38T, D39E, and D247S), PotF_Dist (S87Y, A182D, and L348Q), and PotF_Abox (F276W) (Fig. 1*B*).

Residue influence on polyamine specificity in PotF mutants

PUT and SPD binding profiles of the generated variants were analyzed by ITC. The proximal and distal cluster mutagenesis (PotF_Prox and PotF_Dist) abolishes association of PUT, whereas only PotF_Dist maintained a very low SPD affinity (Table 1). The single mutation F276W in the aromatic box increases SPD binding about threefold ($K_D \approx 11$ μM), whereas PUT affinity decreases threefold ($K_D \approx 0.21$ μM) compared with wildtype PotF. The single mutation F276W does not result in exclusive SPD binding as it is observed for

Table 1
PUT and SPD affinities as determined by ITC for PotF and PotD as well as all variants constructed for the combinatorial mutation analysis between the sequences of PotF and PotF/D

Protein	Ligand	K_D (μM)	n	ΔG (kcal × mol^{-1})	ΔH (kcal × mol^{-1})	$-T\Delta S$ (kcal × mol^{-1})
PotF (10)	PUT	0.07 ± 0.04	0.90 ± 0.02	−9.74 ± 0.40	−23.02 ± 0.12	13.28 ± 0.52
	SPD	29.71 ± 1.15	0.92 ± 0.00	−6.07 ± 0.02	−3.65 ± 0.06	−2.42 ± 0.04
PotD (9)	PUT	N/D	—	—	—	—
	SPD	5.80 ± 1.2	0.86 ± 0.20	−11.13 ± 0.12	−16.40 ± 1.4	5.24 ± 1.39
PotF/D	PUT	N/D	—	—	—	—
	SPD	37.32 ± 2.4	0.94 ± 0.02	−5.94 ± 0.04	−9.30 ± 0.20	3.36 ± 0.24
PotF_Prox	PUT	N/D	—	—	—	—
	SPD	N/D	—	—	—	—
PotF_Abox	PUT	0.21 ± 0.04	0.90 ± 0.03	−8.97 ± 0.10	−22.25 ± 0.77	13.28 ± 0.75
	SPD	11.05 ± 0.88	0.96 ± 0.06	−6.65 ± 0.04	−6.25 ± 0.16	−0.40 ± 0.21
PotF_Dist	PUT	N/D				
	SPD	88.19 ± 33.10	0.91 ± 0.02	−5.48 ± 0.20	−1.83 ± 0.08	−3.65 ± 0.28
PotF_Abox_Prox	PUT	N/D	—	—	—	—
	SPD	N/D	—	—	—	—
PotF_Abox_Dist	PUT	N/D	—	—	—	—
	SPD	7.77 ± 0.32	0.94 ± 0.00	−6.85 ± 0.02	−5.28 ± 0.07	−1.57 ± 0.09
PotF_Abox-S87Y	PUT	14.92 ± 0.54	0.92 ± 0.03	−6.47 ± 0.02	−9.97 ± 0.20	3.50 ± 0.18
	SPD	19.95 ± 0.70	0.97 ± 0.02	−6.30 ± 0.02	−5.66 ± 0.12	−0.65 ± 0.12
PotF_Abox-A182D	PUT	83.22 ± 2.08	0.91 ± 0.01	−5.45 ± 0.02	−9.79 ± 0.09	4.33 ± 0.08
	SPD	3.13 ± 0.62	0.91 ± 0.02	−7.39 ± 0.11	−4.37 ± 0.20	−3.02 ± 0.15
PotF_Abox-L348Q	PUT	2.66 ± 0.08	0.94 ± 0.03	−7.48 ± 0.02	−14.22 ± 0.75	6.74 ± 0.74
	SPD	3.26 ± 0.14	0.98 ± 0.01	−7.83 ± 0.68	−7.72 ± 0.03	−0.11 ± 0.69

If the K_D values are not cited, all are measured in biological triplicates. The error is the standard deviation between the three measurements.
Abbreviation: N/D, not determinable.

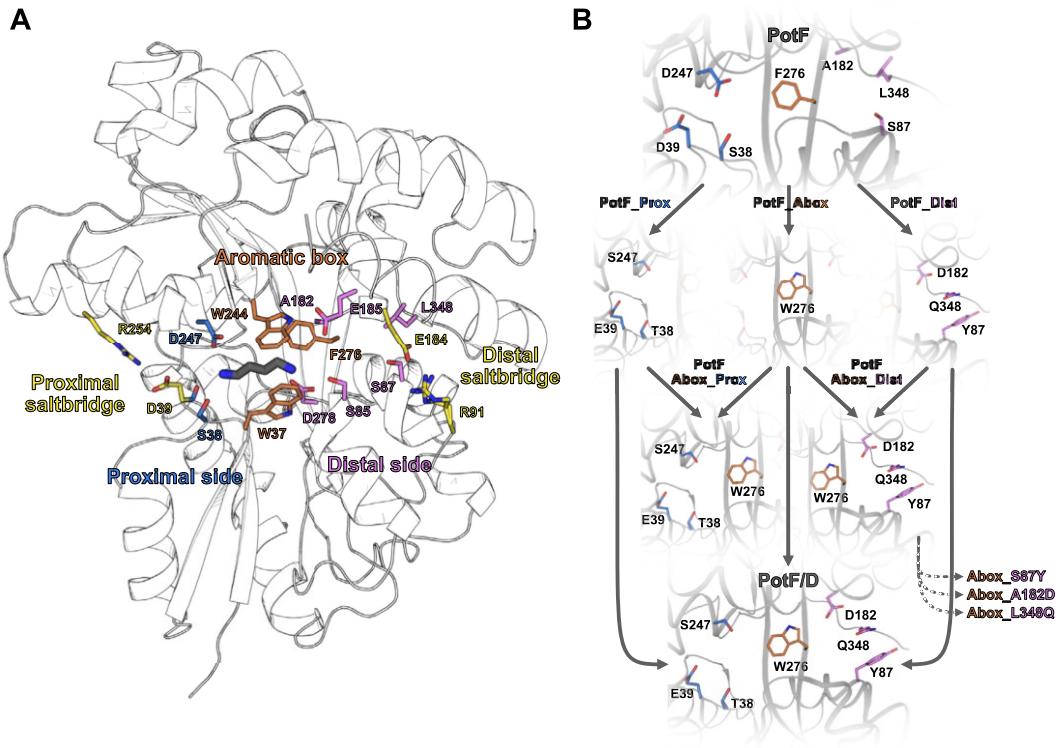


Figure 1. PotF binding pocket and mutational steps towards PotF/D. *A*, binding pocket of PotF in complex with PUT (*gray sticks*). Proximal side chains are shown as *blue sticks*, aromatic box residues as *orange sticks*, and distal side chains as *pink sticks*. Salt bridges flanking the binding pocket are depicted as *yellow sticks*. Protein backbone is shown as *white cartoon with black outline*. *B*, schematic flowchart of the combinatorial sequence space analysis between PotF and PotF/D. In the first round, residues were grouped into proximal (PotF_Prox), aromatic box (PotF_Abox), and distal (PotF_Dist). Coloring of variants and residues was kept in relation with the figure. Binding pocket for PotF is illustrated from PotF in complex with PUT (Protein Data Bank ID: 6YE0), binding pocket of PotF/D and single groups is shown using PotF/D in complex with SPD (PDB ID: 7OYZ). Following the single-group analysis, proximal and distal residues were combined with the aromatic box substitution (PotF_Abox_Prox and PotF_Abox_Dist). Since the distal region seems to exert influence on ligand specificity, single mutants from the distal group in combination with the aromatic box were created and analyzed as well (PotF_Abox_S87Y, PotF_Abox_A182D, and PotF_Abox_L348Q). Protein structures were visualized using PyMOL (The PyMOL Molecular Graphics System; version 2.3; Schrödinger, LLC). PUT, putrescine; SPD, spermidine.

PotF/D (Table 1). The affinity-modulating effect of an equivalent position was also observed for the respective homologs of PotD and PotF, SpuD (PotF, sequence identity: 57.8%) and SpuE (PotD, sequence identity: 34%), in *Pseudomonas aeruginosa* (11).

Because of the key role of F276W in affinity modulation, F276W was combined with the proximal and distal mutations (Fig. 1*B*), respectively. The PotF_Abox_Prox variant (S38T, D39E, D247S, and F276W) showed no binding of PUT similar to PotF_Prox but restored a very low SPD affinity (Table 1). This illustrates the power of F276W in relation to SPD binding and the importance of the proximal residues that form interactions with the N1 primary amines of the respective ligands for polyamine binding in general. This is in accordance with previous mutational studies (12, 13) as well as molecular dynamic simulations, where the proximal side was the first responding region upon ligand encounter (10). The combination of the distal substitutions (S87Y, A182D, and L348Q) with F276W in the aromatic box not only maintains SPD binding but also improves it almost twofold (K_D(SPD) ≈ 7.8 μM) compared with PotF_Abox and roughly fourfold to fivefold compared with PotF and PotF/D, respectively. In addition, PUT affinity was not detectable. This proves that the distal residues in combination with F267W are the main switches to alter the specificity of PotF toward SPD. This is in line with the study of Machius *et al.* (14), who analyzed a PotD homolog from *Treponema pallidum* (TpPotD) and declared the architecture of the distal part of the binding pocket as one principle to explain polyamine specificity and in particular SPD preference. Conversely, the authors link PUT preference to a tighter anchoring of N1. However, this second principle seems questionable since SPD is also bound by wildtype PotF, and the exchanges of the proximal residues abolished the affinity for both polyamines.

We combined each single distal mutation (S87Y, A182D, or L348Q) with F276W (Abox, Fig. 1*B*) to evaluate the specific contributions of these residues further. All three variants show polyamine promiscuity, albeit with different apparent affinities (Table 1). The most prominent swap of the original preference for PUT over SPD and the highest affinity for SPD along this mutational approach is achieved through the combination of A182D with F276W. It exhibits a roughly

tenfold increase in affinity for SPD (K_D(SPD) ≈ 3.1 μM) with a concomitant decrease of PUT affinity by around 1200-fold (K_D(PUT) ≈ 83 μM) compared with wildtype PotF. The variant based on L348Q shows the same tendency but less pronounced (K_D(SPD) ≈ 2.6 μM, K_D(PUT) ≈ 3.3 μM), whereas S87Y has a negative effect on both polyamines compared with the single mutation F276W. In summary, we can achieve high SPD affinity by only switching two amino acids in the binding pocket of PotF, of which F276W is important to maintain affinity in general and A182D or L348Q to improve affinity. For SPD specificity, the combinatorial synergy of multiple discriminatory mutations is important.

In addition to determining the apparent binding constants, ITC unravels the underlying thermodynamic contributions. We previously described highly conserved and coordinated water molecules in the PotF-binding pocket that are thermodynamically favorable (10, 15). They serve as placeholders for ligand molecules and can be displaced upon binding. Two water molecules are removed when accommodating the propylamine extension of SPD compared with PUT, which results in higher solvation entropy during the binding process (Fig. S1). In addition, hydrophobicity of the binding pocket increases, if more water molecules (solvent) are released, which is also linked to more entropically driven interactions (16). This is in line with all tested PotF-binding pocket mutants in which the thermodynamic properties for SPD remain enthalpically and entropically favored, albeit to varying degrees. For example, the gain of PotF_Abox in enthalpic shares ($\Delta\Delta H$ = −2.65 kcal/mol) for binding SPD is accompanied by a large loss in the favorable entropic contribution ($\Delta(-T\Delta S)$ = +2.02 kcal/mol) compared with the wildtype (Table 1), which could be explained through tighter stacking interactions of tryptophan with the longer ligand SPD. In contrast, PUT binding is exclusively driven by an enthalpic term that compensates for unfavorable entropies. In PotF_Abox, PUT binding solely shows a loss of enthalpy contributions ($\Delta\Delta H$ = +0.77 kcal/mol) compared with the wildtype, whereas the unfavorable entropy term stays remarkably similar ($\Delta(-T\Delta S)$ = ±0 kcal/mol; Table 1). Other interesting variants from a thermodynamic point of view are PotF_Abox_L348Q and PotF_Abox_S87Y. These constructs roughly display the same affinity for both polyamines and consequently a similar Gibbs free energy but show completely different thermodynamic profiles. This appears to be a textbook example for entropy–enthalpy compensation (17). The changes in enthalpy and entropy upon polyamine binding in PotF_Abox_L348Q or S87Y are too big compared with the neglectable change in Gibbs energy to be due to only conformational changes, which is the conventional explanation for entropy–enthalpy compensation. Hence, they must partially result from variations in the amounts of water immobilized or released upon complex formation (18). This underlines the importance of water molecules in the binding pocket of PotF and their role in polyamine recognition and specificity.

Taken together, the ITC measurements show that the polyamine specificity is predominantly encoded within the aromatic box residue 276 and the distal positions 182 and 348 of the PotF binding pocket. Furthermore, the distal residues seem to exert their influence on binding specificity *via* thermodynamically favoring or disfavoring the stabilization of water molecules instead of N3 of SPD.

Approaching polyamine specificity from the conformational perspective of the proteins

Beyond the contribution of the individual positions in the binding pocket, the opening of the structure is essential; it influences all residues lining the binding pocket and the accessibility for mediating water molecules. The formation of salt bridges and the resulting stabilization of the closed state upon ligand binding seem to play a major role for ligand affinity in PotF, highlighting the influence of other important protein elements besides the binding pocket residues (10). PotF/D showed a more open conformation in the crystal structure (PDB: 7OYZ) when binding SPD (Fig. 2). A similar phenomenon, albeit less pronounced, was observed in PotF:spermine (PDB: 6YEC; Fig. 2*A*), in which interlobe salt bridges flanking the binding pocket D_{39}–R_{254} and R_{91}–E_{184} are disrupted by interactions with the ligand itself. Thereby, complete closure is hindered in these structures.

Closer examination of PotF/D revealed that two mutations interfere with wildtype-like salt bridge formation. Although D39E would increase the length of the one residue involved in the proximal salt bridge and partially open the structure, the exchange from Asp to Glu still allows salt bridge formation in general. However, the second mutation S87Y seemed to have a much bigger impact on the ligand recognition of PotF as shown in the prior analysis. In PotD, this Tyr faces inward into the binding pocket to coordinate N3 of SPD (Fig. 3*A*). In contrast, when introduced in PotF/D, this Tyr turns outward and disrupts the salt bridge between E184 and R91 (Fig. 3*C*). This observation explains the negative influence of S87Y on polyamine binding in general as well as the lack of affinity improvement in PotF/D, as the Tyr does not only disrupt the salt bridge and thereby obstructs complete closure but also is not able to coordinate N3 of SPD. Nevertheless, the unfavorable position of Y87 seemed irregular as, in principle, there should be space to accommodate this residue inside the binding pocket. This shifted our attention to the close surroundings of position Y87, in particular F88. In PotF (Fig. 3*B*), this Phe points inside the binding pocket, but in PotF/D, its rotamer is vertically flipped by ∼180°. It appears that F88 and Y87 repel each other in PotF/D, driving both residues away from the binding pocket and eliminating PotD-like positioning of Y87 (Fig. 3*C*). All prior observations combined led to the following three constructs: the basic double salt bridge reintroduction E39D–Y87S as well as E39D–F88A and E39D–Y87S–F88Y (Fig. 3*D*). The F88A containing variant was constructed to stay with the original plan of grafting the SPD-binding mode of PotD onto PotF. The removal of this Phe should allow Tyr to flip inside the binding pocket to accommodate PotD-like positioning. We also introduced F88Y in the basic double salt bridge mutant (E39D + Y87S), thereby hijacking a bulky aromatic residue that already points toward the binding pocket to introduce a distal PotD-like Tyr.

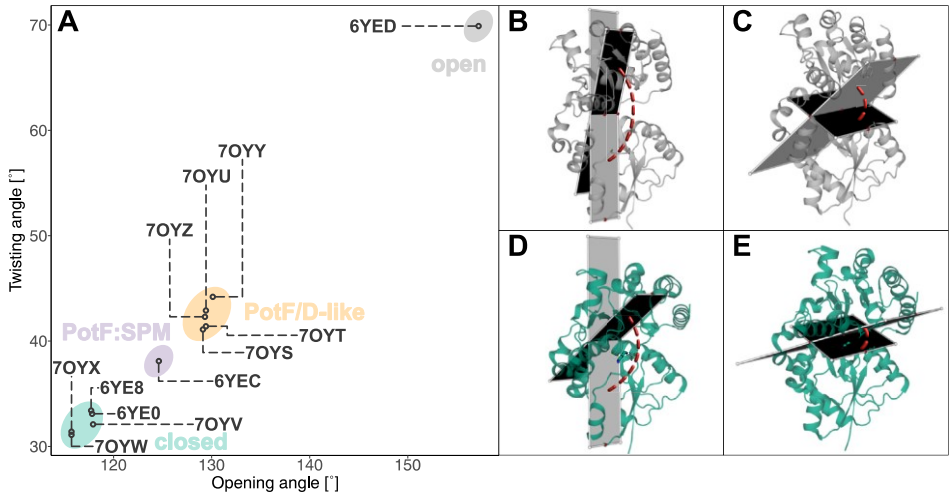


Figure 2. Opening and twisting angles of crystallized variants. *A*, plot of opening and twisting angles of all variants, where we solved an X-ray structure. Each *circle* corresponds to the Protein Data Bank (PDB) identifier it is linked to. *B–E*, illustration of how the opening and twisting angles were determined. Opening angle for (*B*) PotF apo open (PDB ID: 6YED) and (*D*) PotF:PUT (PDB ID: 6YE0) and, respectively, twisting angle for PotF apo (*C*) and PotF:PUT (*E*). Planes in *gray* depict the dihedral angle between the Cα atoms of the four chosen residues. Twisting and opening angles were calculated according to Kröger *et al.* between residues 55, 136, 220, and 276 and 63, 276, 313, and 361, respectively. Protein structures were visualized using PyMOL (The PyMOL Molecular Graphics System, version 2.3; Schrödinger, LLC). PUT, putrescine.

Upon analysis by ITC, an increased SPD affinity was observed for each of the constructed variants. A rudimentary but insignificant PUT affinity was observable in these variants compared with the original PotF/D and thereby still deemed as not determinable (Table 2) in our measurements. The F88A containing mutant showed a similar affinity for SPD as the double salt bridge construct but a different thermodynamic profile. Binding modes of constructs containing Y87S are again similar to the SPD binding of PotFs, revealing beneficial enthalpic and entropic shares (Table 2). On the other hand, the F88A carrying mutant shows a substantial increase in enthalpic and an unfavorable entropic contribution, in this case, more like PotF binding PUT (Table S2). The improvement of the enthalpic contribution by the F88A substitution could suggest a rearrangement of Y87 forming hydrophilic interactions or hydrogen bonds. The highest affinity was measured for the double salt bridge mutant containing the additional F88Y (K_D(SPD) ≈ 5.3 μM). The introduction of F88Y changed the themodynamic profile of the double salt-bridge construct toward the F88A variant albeit less pronounced (Table 2).

Structural analysis of salt bridge constructs

In order to further investigate the mutant's contributions, we tried crystallization of the PotF/D constructs E39D–Y87S, E39D–Y87S–F88Y, and E39D–F88A. Unfortunately, we were not able to crystallize F88A after numerous attempts, while high-resolution structures (Table S3) for both other variants were acquired. It seemed the removal of a whole benzene ring in F88A and keeping only a methyl group destabilized the protein by creating a large unoccupied space. To counteract the aforementioned difficulties, we constructed a variant bearing F88L instead of F88A. This construct displays a slightly lower SPD affinity (K_D(SPD) ≈ 9.9 μM; Table 2) while maintaining the general thermodynamic profile. It acts as a substitute to evaluate whether the structural positioning of Y87 can be altered by removing F88. We obtained high-resolution datasets for this variant in a similar fashion as for the other constructs (Table S3). To our surprise, all solved structures adopt a PotF/D-like semiclosed conformation (Fig. 2*A*).

Also, SPD adopts similar positions to the one in PotF/D in all structures. In E39D–Y87S, no specific differences to PotF/D are observable. Reverting S87 to Y did not result in complete closure or F88 flipping back inside the binding pocket (Fig. 4*A*). The same behavior is observed in the F88L variant where Leu tries to mimic Phe positioning (Fig. 4*C*). It seems that in these constructs, occupation of the space around residue 88 is important; this puts further emphasis on the problems regarding the F88A exchange. In F88L, Y87 remains flipped outside the distal region of the binding pocket. Tyr fills the unoccupied space between the salt bridge residues, interacts with R91 of the salt bridge, and stacks with Q348. Salt bridge residue E184 interacts with Q348 and Y87 as well (Fig. 4*C*).

Furthermore, R91 can interact with D356. In the E39D–Y87S structure (Fig. 4*A*), R91 shows an additional alternate conformation that can interact with E69. This interaction is possible in the completely open "apo" structure (PDB: 6YED) of wildtype PotF as well, hinting at an Arg "hand-off" between E69 or D356 and E184 upon closure.

At the proximal side, D39 is part of the water network involved in coordinating N1 of the ligand. The respective R254 can partially interact with D39 and N231. This time, R254 adopts an alternative conformation in E39D–Y87S and E39D–F88L, favoring one of the possible interactions in each (Fig. 4, *B* and *D*), respectively. The interaction with N231 supports the

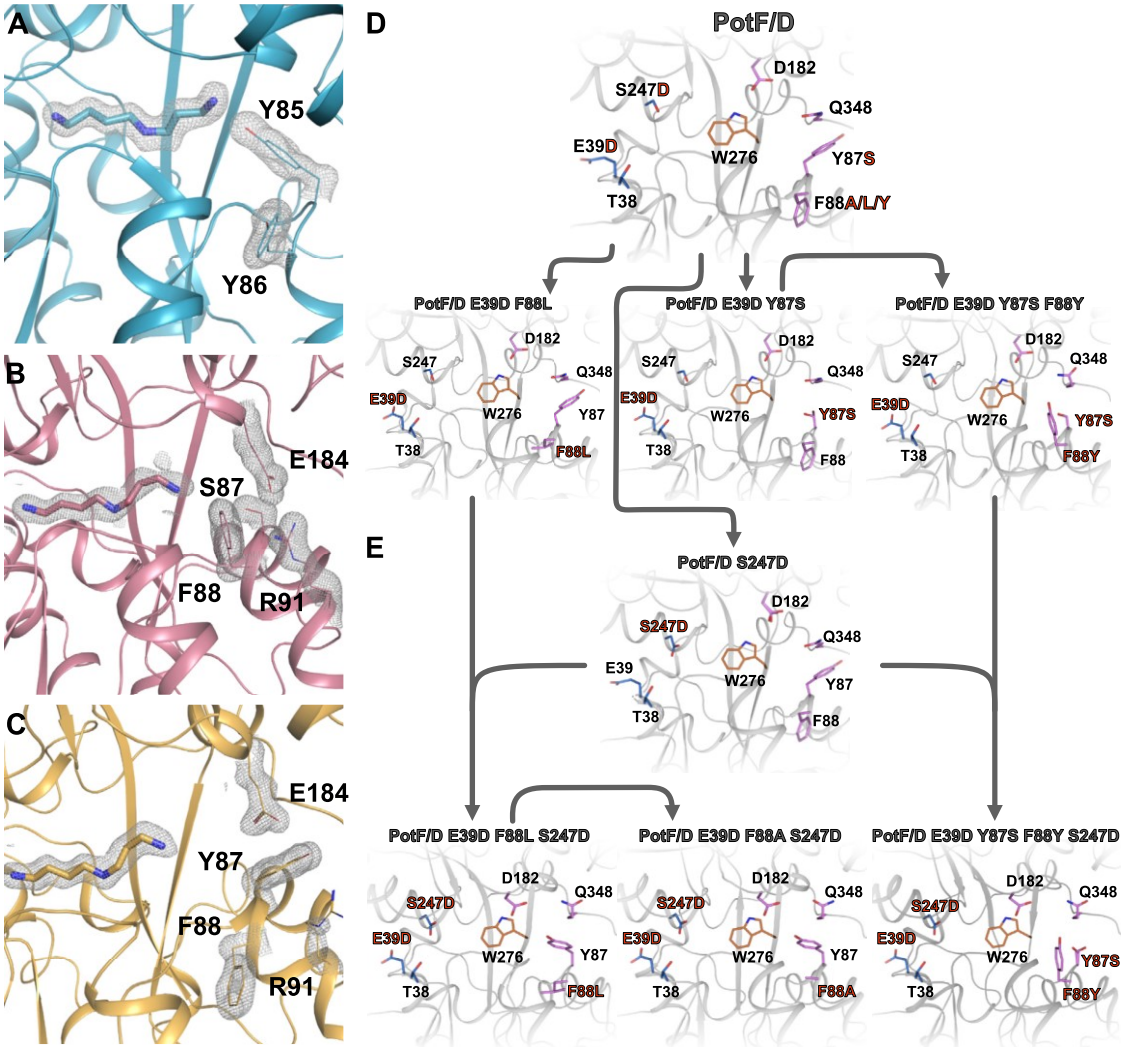


Figure 3. Structural analysis of distal tyrosine and mutational steps on PotF/D. Design puzzles of PotF/D and the interference for closure (*A–C*) as well as the design pathway for improvement (*D* and *E*). In PotD (*A*; Protein Data Bank [PDB] ID: 1POT), Y85 faces inward, directly coordinating SPDs N3. In wildtype PotF in complex with SPD (*B*; PDB ID: 6YE8), the corresponding position is occupied by S87 facing toward the binding pocket but not directly reaching SPD. F88 is in close proximity to S87. In the designed PotF/D (*C*; PDB ID: 7OYZ), the newly introduced Y87 to improve SPD binding faces away from the pocket and thereby disrupts the distal salt bridge (E_{184}–R_{91}). In addition, Y87 and F88 seem to repulse each other. *D*, shows the first steps of reintroducing wildtype-like salt bridge residues and changes made at position 88 to allow the initial design of Y87 to fit inside the pocket and allow for salt bridge formation in addition. The consecutive final design step following *D* to allow complete closure in our final constructs is shown in *E*. In *A–C*, SPD and important side chain residues are shown as sticks with their $2F_o$–F_c densities contoured at 1σ as *gray mesh*. Protein structures were visualized using PyMOL (The PyMOL Molecular Graphics System, version 2.3; Schrödinger, LLC). SPD, spermidine.

conformation of R254 in the semiclosed state of these structures and is observable in the open apo state of PotF wildtype again. This might hint at a second "hand-off" mechanism for salt bridge formation upon closure similar to the distal side.

In the highest affinity F88Y mutant, the newly introduced Tyr slightly rotates inside the binding pocket but does not directly interact with the ligand. It is involved in cation–π interactions with R91 and can stack with N65. The second salt bridge residue E184 interacts with D182 (Fig. 4*E*). Whether all these constructs can completely close in solution remains elusive. It is always possible that, with these specific mutations and the constraints they exercise, the semiclosed conformation is more favored for crystallization of these variants, since the main interplaying distal residues Y87, D182, and Q348, are all newly introduced in PotF/D. Nonetheless, more factors than just the two salt bridges need to influence the preference for complete closure.

Joining the two approaches

To solve the conundrum of a completely closed PotF wildtype-like conformation for our designs, we looked at the dataset as a whole. The importance of the proximal side of the binding pocket and its role for polyamine binding in general

Table 2
PUT and SPD affinities as determined by ITC for PotF/D constructs

Protein	Ligand	K_D (μM)	n	ΔG (kcal × mol^{-1})	ΔH (kcal × mol^{-1})	$-T\Delta S$ (kcal × mol^{-1})
PotF/D–E39D–Y87S	PUT	N/D	—	—	—	—
	SPD	9.44 ± 0.91	0.98 ± 0.01	−6.74 ± 0.06	−6.56 ± 0.13	−0.18 ± 0.19
PotF/D–E39D–F88A	PUT	N/D	—	—	—	—
	SPD	9.60 ± 0.54	0.98 ± 0.03	−6.73 ± 0.03	−12.22 ± 0.35	5.49 ± 0.36
PotF/D–E39D–F88L	PUT	N/D	—	—	—	—
	SPD	9.90 ± 0.45	1.03 ± 0.01	−6.71 ± 0.03	−12.31 ± 0.11	5.60 ± 0.13
PotF/D–E39D–Y87S–F88Y	PUT	N/D	—	—	—	—
	SPD	5.32 ± 1.43	0.99 ± 0.02	−7.09 ± 0.15	−7.81 ± 0.21	0.71 ± 0.32
PotF/D–S247D	PUT	N/D	—	—	—	—
	SPD	9.43 ± 3.07	0.94 ± 0.01	−6.77 ± 0.18	−6.76 ± 0.08	−0.01 ± 0.25
PotF/D–E39D–F88A–S247D	PUT	125.90 ± 5.29	0.92 ± 0.04	−5.23 ± 0.02	−5.21 ± 0.14	−0.02 ± 0.15
	SPD	1.48 ± 0.14	0.91 ± 0.01	−7.82 ± 0.06	−4.75 ± 0.09	−3.07 ± 0.14
PotF/D–E39D–F88L–S247D	PUT	202.70 ± 43.13	0.90 ± 0.01	−4.97 ± 0.13	−5.54 ± 0.73	0.57 ± 0.86
	SPD	3.33 ± 0.53	0.93 ± 0.01	−7.35 ± 0.09	−4.19 ± 0.06	−3.17 ± 0.10
PotF/D–E39D–Y87S–F88Y–S247D	PUT	77.92 ± 3.17	0.90 ± 0.01	−5.51 ± 0.02	−4.15 ± 0.05	−1.36 ± 0.05
	SPD	0.84 ± 0.27	0.91 ± 0.02	−8.17 ± 0.17	−5.46 ± 0.08	−2.71 ± 0.16

K_D values are measured as biological triplicates. The reported error is the error between the three measurements.
Abbreviation: N/D, not determinable.

stood out. We already addressed the minor influence of D39E and its reverse counterpart. The T38S exchange was deemed to have less of an influence since the general properties of the residue stayed constant with just an addition of a methyl group. Hence, the focus shifted toward D247S, the wildtype Asp bears one of two major carboxyl groups important in primary amine coordination upon ligand binding and recognition in PotF. In addition, there must still be a direct influence on closure at the proximal side even after reverting D39 to E.

We introduced the S247D exchange into PotF/D and into our high-affinity constructs E39D–Y87S–F88Y and E39D–F88A/L (Fig. 3*E*). PotF/D–S247D showed an improved SPD affinity in the regime of the first-round salt bridge construct (K_D(SPD) ≈ 9.4 μM; Table 2). The combination of S247D with the other constructs led to an affinity improvement with which the nanomolar range was reached for the first time, while keeping the order of affinities in the different mutants the same compared with the prior analysis. All binding modes shifted toward being enthalpically and entropically favored, reflecting the wildtype binding mode of PotF for SPD (Table 2). Nonetheless, all constructs regained PUT affinity in the medium to low micromolar range (Table 2). This was expected as we reintroduced the wildtype-like proximal primary amine–binding site in these variants.

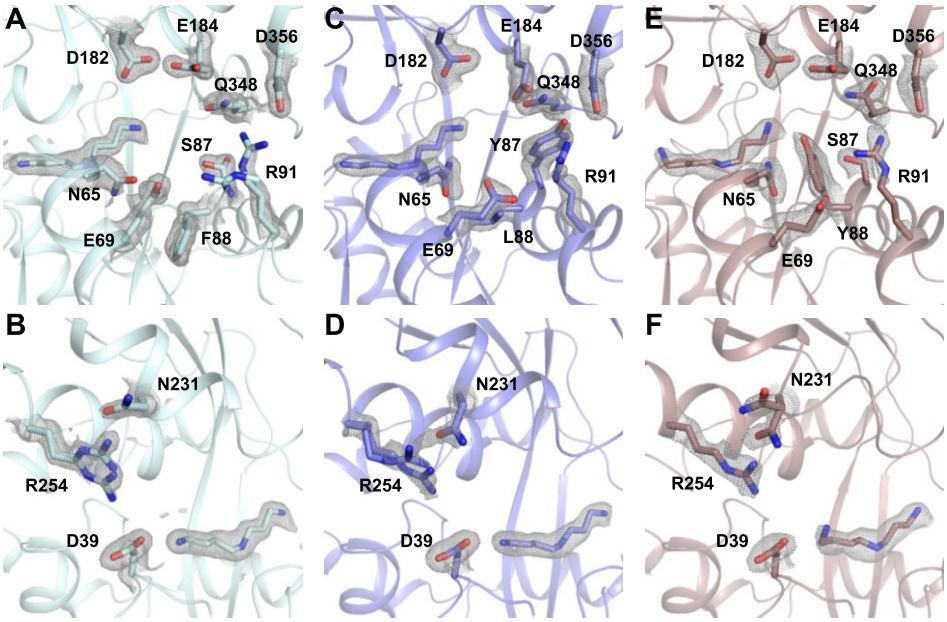


Figure 4. Structural details of the first-round salt bridge mutants related to Figure 3*D*. *A*, *C*, and *E*, distal side of the binding pocket of PotF/D–E39D–Y87S (*A*, Protein Data Bank [PDB] ID: 7OYS), PotF/D–E39D–F88L (*C*, PDB ID: 7OYT), and PotF/D–E39D–Y87S–F88Y (*E*, PDB ID: 7OYU). Proximal salt bridge side of the binding pocket of PotF/D–E39D–Y87S (*B*), PotF/D–E39D–F88L (*D*), and PotF/D–E39D–Y87S–F88Y (*F*). SPD and important side chain residues are shown as *sticks* with their $2F_o$–F_c densities contoured at 1σ as *gray mesh*. Protein structures were visualized using PyMOL (The PyMOL Molecular Graphics System, version 2.3; Schrödinger, LLC). SPD, spermidine.

Tuning specificity in PotF

We successfully obtained X-ray data for all newly generated variants including the F88A carrying construct (Table S3). All structures but PotF/D–S247D adopt a fully closed wildtype conformation (Fig. 2). This backs up the theory that major elements on either side of the binding pocket play a crucial role in modulating closure, excluding the proximal salt bridge since it never truly was disrupted but rather extended. In the F88L/A variants, Y87 is finally able to rotate toward the binding pocket and allows distal salt bridge formation. The residue, however, does not coordinate the primary amine at N3 of SPD directly but takes part in a hydrogen bonding network with D182, N65, and a water molecule, which is a direct interaction partner to N3 of SPD (Fig. 5, *A* and *B*).

The orientation of the residues in both mutants is identical, confirming the F88L containing variant as a structural example for the noncrystallizable F88A construct in the first part of this study. An interesting feature of the F88A containing construct is the orientation of the N2 of SPD, as it faces toward the aromatic box residues to form stacking interaction in contrast to all other constructs, in which interaction with D278 is always preferred. The F88A containing structure is less resolved than the others, and the ligand density could allow for both conformations of N2 in the density, but multiple refinement cycles ended up preferring the built one. This puts more emphasis on the remaining question regarding the slightly lower affinity of F88L. As previously described, Leu at position 88 rotates downward, similar to Phe in the original PotF/D, thereby α-helix 86 to 95 slides back and loses typical α-helical properties as determined by DSSP (Fig. 5*C*; (19)). This slight structural change is propagated over two loops and α-helix 114 to 123 with an RMSD (over all Cαs) of 0.610 Å for this specific region compared with the residues before (29–86, 0.272 Å), after (124–369, 0.335 Å), as well as the complete chain A RMSD of 0.481 Å. Chain B as well as the A–B chain comparison behaves in a comparable manner (Table S2). This deviation puts extra strain on the structure of F88L and especially the distal salt bridge since R91 is in the affected region. In general, the noise in the F_o–F_c map of this structure seems to indicate many flexible side chains. The "flexible" F88L structure is also the first in which we ever observed SPD taking alternate conformations in the binding pocket of which one does not mimic PUT-like positioning. PotF/D–E39D–Y87S–F88Y–S247D is the construct with the highest SPD affinity (840 nM; Fig. 5*D* and Table 2). In this structure, Y88 is stacking with N65 and the salt bridge residue R91, stabilizing the latter and thereby adopting the fully closed conformation. Furthermore, the hydroxyl group of Y88 is coordinated by E66, whereas R91 is interacting with S87 in two alternate conformations (Fig. 5*E*). Taken together, the interplay of the mentioned residues forms a tight pocket surrounding Y88, thereby stabilizing the distal region and supporting the completely closed conformation. In all constructs that carry S247D, SPD adopts a "relaxed" conformation and does not show unexpected bending of its backbone to fit a specific position as it does when being bound by PotF wildtype or PotF/D (Fig. 5, *A*, *B*, and *D*).

All in all, we determined salt bridges and the key carboxyl-harboring proximal residue D247 as the basic prerequisites for

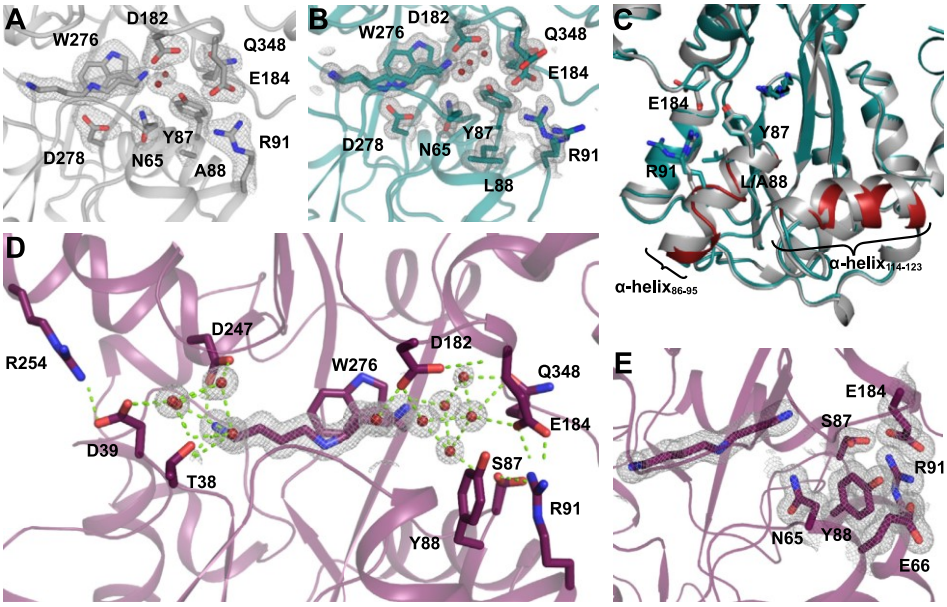


Figure 5. Structural details of second-round salt bridge mutants related to Figure 3*E*. *A*, PotF/D–E39D–F88A–S247D (Protein Data Bank [PDB] ID: 7OYV). *B*, PotF/D–E39D–F88L–S247D (PDB ID: 7OYW). *C*, PotF/D–E39D–F88L–S247D in *green* and PotF/D–E39D–F88A–S247D in *gray*. *Red* highlights the differences in the helices of PotF/D–E39D–F88L–S247D (*green*) to PotF/D–E39D–F88A–S247D (*gray*). *D*, PotF/D–E39D–Y87S–F88Y–S247D (PDB ID: 7OYX) with waters in the binding pocket shown as *red spheres* and the polar contacts highlighted as *light green dashed lines*. *E*, distal side of the binding pocket of PotF/D–E39D–Y87S–F88Y–S247D highlighting the tight and pocket-like coordination of Tyr88. SPD and important side chain residues are shown as *sticks*. If densities are shown, $2F_o$–F_c maps contoured at 1σ are depicted as *gray mesh*. Protein structures were visualized using PyMOL (The PyMOL Molecular Graphics System, version 2.3; Schrödinger, LLC). SPD, spermidine.

a complete closure of PotF constructs. Fixing each component by itself was not sufficient. Only the combination of both allows full closure, deeming it an intricately regulated process.

Discussion

Design aspects and interpretation

We were able to pinpoint SPD specificity to the aromatic box and the interplay of several distal residues. Nonetheless, higher affinity SPD binders always came at a partial loss of specificity by reintroducing a slight PUT affinity. In most cases, PUT affinity remained marginal and at best in high micromolar ranges. This affinity is detectable in *in vitro* measurements, whereas in an *in vivo* situation our final variants would almost always prefer SPD binding over PUT if given the choice. On the one hand, we were able to develop an improved specific SPD-binding PotF variant in PotF_Abox_Dist (7.8 μM; Table 1) as well as a whole variety of enhanced PotF/D constructs (E39D–Y87S, E39D–F88A, E39D–F88L, E39D–Y87S–F88Y, and S247D) with good SPD affinity (9.4, 9.6, 9.9, 5.3, & 9.4 μM, respectively) and an array of different thermodynamic binding profiles (Table 2). On the other hand, we designed a PotF/D construct that completely switched its ligand preference compared with the wildtype by exhibiting nanomolar affinity for SPD (840 nM) and medium micromolar affinity for PUT (88 μM). It appears that shaping affinity comes always at the risk of losing specificity and vice versa.

Implications for evolutionary adaptation of polyamine uptake in pathogenic bacteria

Promiscuity of a receptor allows recognition of a range of structurally and chemically similar molecules; this comes with the advantage of adaptability for the host organism as well, hence it is less susceptible to changes exerting selective pressure. Nonetheless, this adaptability bears negative aspects, especially looking at pathogens. In the case of PotF and PotD, numerous orthologs with good sequence conservation, especially in the binding pocket, have been identified. Highly similar receptors have been characterized for polyamine specificity in different bacterial species (14, 20). PotD has been identified as a potential virulence factor in *Streptococcus pneumoniae* (21) and successfully applied as an immunization against systemic infection in mice (22). Multidrug-resistant gram-negative bacteria have a significant impact on public health, and SPD uptake has been linked to the expression of type III secretion (T3SS) system genes (23), which are an essential part in their pathogenesis (24, 25). Therefore, a SpuE antibody was designed to prevent SPD transport and ultimately weakening *Pseudomonas aeruginosa* infection (26). Bacterial pathogens have proven their adaptability numerous times, and by having promiscuous PBPs like PotF (or its homologs) at their disposal, they might just be able to hijack another system to facilitate SPD uptake and bypass treatment methods in the long run. Two mutations in PotF_Abox_A182D are enough to reverse polyamine specificity in PotF to favor SPD. This does not take into consideration the already mediocre micromolar (30 μM) affinity of PotF for SPD, which could explain why a previous study on a *ΔpotD S. pneumoniae* strain (21) already hinted at the existence of an alternative SPD uptake system. It is conceivable that this might have been the co-usage of the PotFGHI system by both polyamines. This is not uncommon in nature, since two other highly similar PBPs (70% sequence identity (27)), LAOBP and HisJ, both of which show promiscuous binding capabilities for Lys, Arg, Orn, and His but in a different order (28–30), even share the same inner membrane ATPase-permease complex ($HisQMP_2$; (27)). Major findings of our study compared with others on the PBPs of the polyamine uptake system are in mutual agreement, thus it can be feasible to extrapolate these results onto other bacterial species in which homologs or orthologs of these proteins can be found. This might suggest looking at PotD and PotF as well as their respective relatives in more of a joint manner from a medical perspective.

Conservation and evolution of polyamine transporter systems

To put further emphasis on this point, we utilized a protein BLAST on the UniProt entry of PotF (UniProt ID: P31133), which resulted in 250 unbiased hits of polyamine transporter systems from numerous different organisms (Material S1). We analyzed conservation and mutational frequency of important binding pocket and salt bridge residues. This resulted in a median conservation of ∼75% for all 15 residues. The highest conservation is present for aromatic box residue W37 (100%) and key carboxyl harboring residues D247, D278, and E185 as well as proximal salt bridge residues D39 and R254 (all ≥96%). Interestingly, in a third of the sequences, F276W is present, and with generally less conservation at the distal side of the binding pocket (S87: 45%, A182: 60%, and L348: 43%), an evolution toward the aforementioned co-usage of PotF by multiple polyamines is conceivable. This is further supported considering PotF sequences being annotated as PUT/SPD transporter systems for several of the different organisms. The lowest conservation percentage shows distal salt bridge residue E184 with 26%. In 10% of the cases, E184D would still allow for salt bridge formation, but E184T (37%) seems to be the preferred evolutionary trajectory for this residue. Nonetheless, the opposing salt bridge residue R91 shows good conservation with 59% and R91K being the nearest possible exchange with 37%, therefore keeping the possibility of distal salt bridge formation (E/D_{184}–R/K_{91}) in over a third of the cases. The results of the conservation analysis are visualized in Figure 6 as a WebLogo.

The lower conservation rate throughout many of the distal residues and the potential loss of salt bridge formation enable

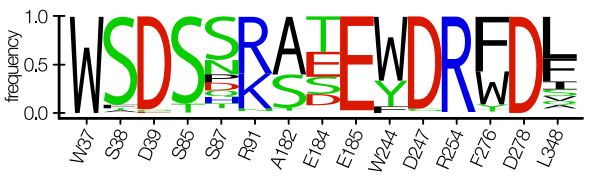


Figure 6. WebLogo representation of the conservation analysis of PotF homologs.

the distal part of the binding pocket to be less structurally fixed and, thus, allows for more flexibility and thereby adaptability for a broader ligand spectrum. We know already that PotF can recognize multiple polyamines on top of PUT and SPD (10), and the aforementioned findings suggest the same for different orthologs, which might have evolved toward even increased promiscuity by being exposed to various external stimuli throughout evolution.

Further medical relevance

Polyamines are not only important for bacterial regulatory pathways but are involved in a multitude of processes in eukaryotes (31–33). Their presence and metabolism are commonly upregulated and dysregulated in cancer cells (34, 35). Therefore, polyamines have become the target of multiple treatment strategies as well as biomarkers for tumor progression in specific cancer types (34). In recent years, the usage of PBPs as receptor modules fused to circularly permuted fluorescent proteins as biosensors became a well-established approach to track small molecules (36–38). Our studies on PotF provide the perfect framework for engineering similar biosensors with desired affinities and specificities to assist clinical polyamine research.

Experimental procedures

Cloning of PotF and PotF/D variants

The gene for PotF (UniProt ID: P31133) without the N-terminal signal peptide was amplified by PCR using genomic DNA of *E. coli* K-12 and primers, which introduced flanking restriction sites for NdeI and XhoI. After digestion, the DNA fragment was ligated into a pET21b(+)-vector, thereby adding a C-terminal His6-tag. PotF/D (previously called PotF_SPD) was constructed prior to this study (9). Mutations for the construction of different PotF and PotF/D variants were introduced by a modified QuickChange PCR utilizing KAPA polymerase (Roche) followed by an additional ligation step using T4 DNA ligase (New England Biolabs) according to the manufacturer's protocol. All used oligonucleotides can be found in Table S4. Top10 cells were transformed with reaction mixture *via* heat shock and plated on LB agar containing ampicillin (100 μg/ml) as selection marker. Overnight cultures were grown in LB medium containing ampicillin (100 μg/ml), and DNA was isolated using NucleoSpin Plasmid EasyPure-Kit (Machery & Nagel) according to the manufacturer's protocol. To confirm correct implementation of mutations, plasmids were sequenced (Eurofins Genomics) using standard T7 primers. For expression of proteins, BL21 (DE3) cells were transformed with the plasmid and plated out on LB agar plates containing 100 μg/ml ampicillin as selection marker to obtain single colonies, which were further used to inoculate overnight cultures.

Protein expression

For protein expression, 2 l LB were inoculated with 20 ml of an overnight culture containing 100 μg/ml ampicillin and incubated at 37 °C until an absorbance reached a value of 0.6 to 0.8 at 600 nm. Overexpression was induced by adding IPTG to a final concentration of 1 mM and further incubation for 4 h at 37 °C. Cells were harvested by centrifugation (Beckman Coulter; JLA 8.1000; 4000*g*, 20 min, 4 °C), and pellets were resuspended and washed with 30 ml buffer (50 mM Tris [pH 8], 300 mM NaCl, and 20 mM Imidazole). After recentrifugation (Eppendorf 5920R; S-4x750; 4000*g*, 1 h, 4 °C), the washed pellets were stored at −20 °C until further use.

Protein purification

Cells were resuspended (50 mM Tris [pH 8], 300 mM NaCl, and 20 mM imidazole) and lysed by sonication (Branson; 6.3 mm tip, 3 × 3 min, 40% duty cycle, output power 4), followed by a centrifugation step (Beckman Coulter; JA25.50; 40,000*g*, 60 min) to separate the soluble from the insoluble fraction and cell debris. The supernatant was loaded onto an equilibrated (50 mM Tris [pH 8], 300 mM NaCl, and 20 mM imidazole) HisTrap HP 5 ml column (GE Healthcare) using a peristaltic pump (Pump P-1; GE Healthcare). After washing for ten column volumes with lysis buffer, unfolding of the protein to wash off endogenous ligands was achieved by washing and incubating for an hour with 6 M guanidinium hydrochloride and subsequently washing with ten column volumes of lysis buffer for refolding. The protein was eluted by a stepwise increase of the imidazole concentration to 260 mM with an ÄKTA system. Fractions containing the protein were pooled and concentrated with a centrifugal concentrator (Amicon; 10 kDa molecular weight cutoff) to a maximum volume of 12 ml and applied to an equilibrated (50 mM Tris [pH 8] and 300 mM NaCl) preparative size exclusion column (HiLoad Superdex 75 26/60; GE Healthcare). Fractions with monomeric protein were pooled and concentrated for ITC measurements or crystallization setups (Amicon; 10 kDa molecular weight cutoff). Protein concentration was determined photometrically using the absorption at 280 nm. Expression and purification were verified by SDS-PAGE.

ITC

Ligands were always freshly prepared in the exact same buffer as was used for the size exclusion run. Samples were degassed, and temperature was equilibrated using a degassing station (TA Instruments). About 400 μl protein sample of different concentrations (Table S5) depending on affinity was transferred into the sample cell of a nanoITC (TA Instruments), and 50 μl of about tenfold concentrated ligand solution was loaded into the injection needle. Multiple injection measurements were carried out at 293 K, 300 rpm stirring rate, and 250 s spacings between each 2 μl injection. The heat quantity past injection was determined by integration of the measured peaks. Every protein–ligand combination was measured as biological triplicate. Subtraction of heat of dilution measurements, peak integration, and one site binding fits

were done with NanoAnalyze (TA Instruments). Reported errors are the errors between fits of the measured triplicates.

Crystallography

All crystallization experiments were set up as sitting drop vapor diffusion experiments in either 3-well Intelliplates (Art Robbins Instruments) or MRC Maxi plates (Swissci) using a protein concentration of 40, 30, or 15 mg/ml and a 20-fold molar excess of ligand. Protein–ligand mixtures were equilibrated at 293 K for several hours before crystallization setups.

First crystals were obtained after an initial sparse matrix screen using the commercially available JCSG CORE I-IV screens (QIAGEN). If needed, these hits were further optimized by using a grid screen to improve promising conditions and applying the Additive Screen (Hampton Research).

This resulted in conditions containing 2.4 M ammonium sulfate, 0.1 M bicine, pH 8.3 or 9.0, and either 4.5% or 5% Jeffamine M-600 as additive. We also obtained crystals in 0.1 M sodium acetate (pH 4.6), 0.2 M ammonium acetate, and 30% PEG 4000; 0.085 M sodium acetate [pH 4.7], 0.17 M ammonium acetate with 30% PEG 4000 and 15% glycerol; and 0.1 M MES (pH 5) with 30% PEG 6000. Detailed information regarding each crystal can be found in Table S6.

Crystals were mounted using CryoLoops and transferred into a cryogenic solution made of reservoir solution and either 25% glycerol or 1.7 M malonate matching the pH of the condition (Table S6) and then cooled down in liquid nitrogen. Data collection at 100 K was done at the beamlines BL 14.1 and 14.2 at the synchrotron BESSY II, Helmholtz-Zentrum Berlin (39). Diffraction data were processed using XDSAPP (40), and decision of resolution cutoff was made according to $CC_{1/2}$ around 0.25, I/Sigma >0.5, and completeness in the outer shell >75%. Molecular replacement was done using Phaser-MR (14) with both lobes independently and missing hinge region residues of PotF (PDB: 1A99 (13)) as search model, to account for different opening angles. Data quality was assessed utilizing phenix.xtriage. Refinements were done with Phenix.refine (41). Manual model building was performed in Coot (42). Final models were evaluated by PDB_REDO 6.00 (43). Refinement statistics and crystallographic data are shown in Table S3.

Bioinformatics analysis of PotF and related sequences

BLAST was used on the UniProt ((44); 2021) entry of PotF (POTF_ECOLI; P31133) with the settings listed in Table 3. The resulting hits were analyzed on the mutational distribution of each residue of interest for this study. A detailed list of all analyzed sequences and their respective organisms as well as their score, identity, and E-value with regard to PotF can be found in Material S1.

Data availability

X-ray coordinates of all solved structures have been deposited at the PDB (http://www.rcsb.org) with accession codes: 7OYS, 7OYT, 7OYU, 7OYV, 7OYW, 7OYX, 7OYY, and 7OYZ.

Table 3
pBLAST parameters used for the conservational analysis of PotF and PotF-related structures

Program	BLASTP (BLASTP 2.9.0+)
Database	uniprotkb_refprotswissprot (Protein) generated for BLAST on December 2, 2020
Sequences	57,391,823 sequences consisting of 21,838,191,652 letters
Matrix	blosum62
Threshold	10
Filtered	False
Gapped	True
Maximum no. of hits reported	250

Supporting information—This article contains supporting information.

Acknowledgments—We acknowledge financial support and allocation of synchrotron beamtime by HZB and thank the beamline staff at BESSY for support. This work was supported by Deutsche Forschungsgemeinschaft grant HO4022/2-3.

Author contributions—B. H. conceptualization; P. K., S. S., and U. S. investigation; P. K. and S. S. data curation; P. K., S. S., and B. H. writing–original draft; P. K., S. S., U. S., and B. H. writing–reviewing & editing; B. H. funding acquisition.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: ITC, isothermal titration calorimetry; PBP, periplasmic binding protein; PDB, Protein Data Bank; PUT, putrescine; SPD, spermidine.

References

1. Nobeli, I., Favia, A. D., and Thornton, J. M. (2009) Protein promiscuity and its implications for biotechnology. *Nat. Biotechnol.* **27**, 157–167
2. Schreiber, G., and Keating, A. E. (2011) Protein binding specificity versus promiscuity. *Curr. Opin. Struct. Biol.* **21**, 50–61
3. Borrok, M. J., Zhu, Y., Forest, K. T., and Kiessling, L. L. (2009) Structure-based design of a periplasmic binding protein antagonist that prevents domain closure. *ACS Chem. Biol.* **4**, 447–456
4. Dwyer, M. A., and Hellinga, H. W. (2004) Periplasmic binding proteins: A versatile superfamily for protein engineering. *Curr. Opin. Struct. Biol.* **14**, 495–504
5. Felder, C. B., Graul, R. C., Lee, A. Y., Merkle, H.-P., and Sadee, W. (1999) The venus flytrap of periplasmic binding proteins: An ancient protein module present in multiple drug receptors. *AAPS PharmSci.* **1**, 7–26
6. Moussatova, A., Kandt, C., O'Mara, M. L., and Tieleman, D. P. (2008) ATP-binding cassette transporters in Escherichia coli. *Biochim. Biophys. Acta* **1778**, 1757–1771
7. Furuchi, T., Kashiwagi, K., Kobayashi, H., and Igarashi, K. (1991) Characteristics of the gene for a spermidine and putrescine transport system that maps at 15 min on the Escherichia coli chromosome. *J. Biol. Chem.* **266**, 20928–20933
8. Pistocchi, R., Kashiwagi, K., Miyamoto, S., Nukui, E., Sadakata, Y., Kobayashi, H., and Igarashi, K. (1993) Characteristics of the operon for a putrescine transport system that maps at 19 minutes on the Escherichia coli chromosome. *J. Biol. Chem.* **268**, 146–152

9. Scheib, U., Shanmugaratnam, S., Farías-Rico, J. A., and Höcker, B. (2014) Change in protein-ligand specificity through binding pocket grafting. *J. Struct. Biol.* **185**, 186–192
10. Kröger, P., Shanmugaratnam, S., Ferruz, N., Schweimer, K., and Höcker, B. (2021) A comprehensive binding study illustrates ligand recognition in the periplasmic binding protein PotF. *Structure* **29**, 433–443.e4
11. Wu, D., Lim, S. C., Dong, Y., Wu, J., Tao, F., Zhou, L., Zhang, L.-H., and Song, H. (2012) Structural basis of substrate binding specificity revealed by the crystal structures of polyamine receptors SpuD and SpuE from Pseudomonas aeruginosa. *J. Mol. Biol.* **416**, 697–712
12. Kashiwagi, K., Pistocchi, R., Shibuya, S., Sugiyama, S., Morikawa, K., and Igarashi, K. (1996) Spermidine-preferential uptake system in Escherichia coli. Identification of amino acids involved in polyamine binding in PotD protein. *J. Biol. Chem.* **271**, 12205–12208
13. Vassylyev, D. G., Tomitori, H., Kashiwagi, K., Morikawa, K., and Igarashi, K. (1998) Crystal structure and mutational analysis of the Escherichia coli putrescine receptor. Structural basis for substrate specificity. *J. Biol. Chem.* **273**, 17604–17609
14. Machius, M., Brautigam, C. A., Tomchick, D. R., Ward, P., Otwinowski, Z., Blevins, J. S., Deka, R. K., and Norgard, M. V. (2007) Structural and biochemical basis for polyamine binding to the Tp0655 lipoprotein of Treponema pallidum: Putative role for Tp0655 (TpPotD) as a polyamine receptor. *J. Mol. Biol.* **373**, 681–694
15. Schiebel, J., Gaspari, R., Wulsdorf, T., Ngo, K., Sohn, C., Schrader, T. E., Cavalli, A., Ostermann, A., Heine, A., and Klebe, G. (2018) Intriguing role of water in protein-ligand binding studied by neutron crystallography on trypsin complexes. *Nat. Commun.* **9**, 3559
16. Du, X., Li, Y., Xia, Y.-L., Ai, S.-M., Liang, J., Sang, P., Ji, X.-L., and Liu, S.-Q. (2016) Insights into protein-ligand interactions: Mechanisms, models, and methods. *Int. J. Mol. Sci.* **17**, 144
17. Chodera, J. D., and Mobley, D. L. (2013) Entropy-enthalpy compensation: Role and ramifications in biomolecular ligand recognition and design. *Annu. Rev. Biophys.* **42**, 121–142
18. Dragan, A. I., Read, C. M., and Crane-Robinson, C. (2017) Enthalpy-entropy compensation: The role of solvation. *Eur. Biophys. J.* **46**, 301–308
19. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637
20. Brandt, A.-M., Raksajit, W., Yodsang, P., Mulo, P., Incharoensakdi, A., Salminen, T. A., and Mäenpää, P. (2010) Characterization of the substrate-binding PotD subunit in Synechocystis sp. strain PCC 6803. *Arch. Microbiol.* **192**, 791–801
21. Ware, D., Jiang, Y., Lin, W., and Swiatlo, E. (2006) Involvement of potD in Streptococcus pneumoniae polyamine transport and pathogenesis. *Infect. Immun.* **74**, 352–361
22. Shah, P., and Swiatlo, E. (2006) Immunization with polyamine transport protein PotD protects mice against systemic infection with Streptococcus pneumoniae. *Infect. Immun.* **74**, 5888–5892
23. Zhou, L., Wang, J., and Zhang, L.-H. (2007) Modulation of bacterial type III secretion system by a spermidine transporter dependent signaling pathway. *PLoS One* **2**, e1291
24. Felise, H. B., Nguyen, H. V., Pfuetzner, R. A., Barry, K. C., Jackson, S. R., Blanc, M.-P., Bronstein, P. A., Kline, T., and Miller, S. I. (2008) An inhibitor of gram-negative bacterial virulence protein secretion. *Cell Host Microbe* **4**, 325–336
25. Keyser, P., Elofsson, M., Rosell, S., and Wolf-Watz, H. (2008) Virulence blockers as alternatives to antibiotics: Type III secretion inhibitors against gram-negative bacteria. *J. Intern. Med.* **264**, 17–29
26. Zhang, Y., Sun, X., Qian, Y., Yi, H., Song, K., Zhu, H., Zonta, F., Chen, W., Ji, Q., Miersch, S., Sidhu, S. S., and Wu, D. (2019) A potent anti-SpuE antibody allosterically inhibits type III secretion system and attenuates virulence of Pseudomonas aeruginosa. *J. Mol. Biol.* **431**, 4882–4896
27. Higgins, C. F., and Ames, G. F. (1981) Two periplasmic transport proteins which interact with a common membrane receptor show extensive homology: Complete nucleotide sequences. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 6038–6042
28. Nikaido, K., and Ames, G. F. (1992) Purification and characterization of the periplasmic lysine-, arginine-, ornithine-binding protein (LAO) from Salmonella typhimurium. *J. Biol. Chem.* **267**, 20706–20712
29. Paul, S., Banerjee, S., and Vogel, H. J. (2017) Ligand binding specificity of the Escherichia coli periplasmic histidine binding protein, HisJ. *Protein Sci.* **26**, 268–279
30. Pulido, N. O., Silva, D.-A., Tellez, L. A., Pérez-Hernández, G., García-Hernández, E., Sosa-Peinado, A., and Fernández-Velasco, D. A. (2015) On the molecular basis of the high affinity binding of basic amino acids to LAOBP, a periplasmic binding protein from Salmonella typhimurium. *J. Mol. Recognit.* **28**, 108–116
31. Igarashi, K., and Kashiwagi, K. (2010) Modulation of cellular function by polyamines. *Int. J. Biochem. Cell Biol.* **42**, 39–51
32. Mandal, S., Mandal, A., Johansson, H. E., Orjalo, A. V., and Park, M. H. (2013) Depletion of cellular polyamines, spermidine and spermine, causes a total arrest in translation and growth in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 2169–2174
33. Pegg, A. E., and Casero, R. A. (2011) Current status of the polyamine research field. *Methods Mol. Biol.* **720**, 3–35
34. Casero, R. A., Murray Stewart, T., and Pegg, A. E. (2018) Polyamine metabolism and cancer: Treatments, challenges and opportunities. *Nat. Rev. Cancer* **18**, 681–695
35. Soda, K. (2011) The mechanisms by which polyamines accelerate tumor spread. *J. Exp. Clin. Cancer Res.* **30**, 95
36. Marvin, J. S., Borghuis, B. G., Tian, L., Cichon, J., Harnett, M. T., Akerboom, J., Gordus, A., Renninger, S. L., Chen, T.-W., Bargmann, C. I., Orger, M. B., Schreiter, E. R., Demb, J. B., Gan, W. B., Hires, S. A., *et al.* (2013) An optimized fluorescent probe for visualizing glutamate neurotransmission. *Nat. Methods* **10**, 162–170
37. Marvin, J. S., Schreiter, E. R., Echevarría, I. M., and Looger, L. L. (2011) A genetically encoded, high-signal-to-noise maltose sensor. *Proteins* **79**, 3025–3036
38. Marvin, J. S., Shimoda, Y., Magloire, V., Leite, M., Kawashima, T., Jensen, T. P., Kolb, I., Knott, E. L., Novak, O., Podgorski, K., Leidenheimer, N. J., Rusakov, D. A., Ahrens, M. B., Kullmann, D. M., and Looger, L. L. (2019) A genetically encoded fluorescent sensor for in vivo imaging of GABA. *Nat. Methods* **16**, 763–770
39. Mueller, U., Förster, R., Hellmig, M., Huschmann, F. U., Kastner, A., Malecki, P., Pühringer, S., Röwer, M., Sparta, K., Steffien, M., Ühlein, M., Wilk, P., and Weiss, M. S. (2015) The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives. *Eur. Phys. J. Plus* **130**, 662
40. Sparta, K. M., Krug, M., Heinemann, U., Mueller, U., and Weiss, M. S. (2016) XDSAPP2.0. *J. Appl. Crystallogr.* **49**, 1085–1092
41. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., *et al.* (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221
42. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
43. Joosten, R. P., Joosten, K., Murshudov, G. N., and Perrakis, A. (2012) PDB_REDO: Constructive validation, more than just looking for errors. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 484–496
44. UniProt Consortium (2021) UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489

Supplementary information for

Fine-tuning spermidine binding modes in the putrescine binding protein PotF

Pascal Kröger[1], Sooruban Shanmugaratnam[1,2], Ulrike Scheib[2], Birte Höcker[1,2*]

[1]Department for Biochemistry, University of Bayreuth, Bayreuth, Germany

[2] Max-Planck-Institute for Developmental Biology, Tübingen, Germany

*Corresponding author

Table of Content

Figure S1: Binding pocket of PotFwt:PUT and PotFwt:SPD with waters

Figure S2a: ITC measurements and analysis related to Table 1.

Figure S2b: ITC measurements and analysis related to Table 1.

Figure S3: ITC measurements and analysis related to Table 2.

Table S1: Residue groups in PotF and their respective counterparts in PotD

Table S2: RMS calculated by aligning the stated residues from PotF/D-E39D-F88A-S247D onto PotF/D-E39D-F88L-S247

Table S3: Crystallographic data and refinement statistics for all solved crystal structures

Table S4: Oligonucleotides used for the generation of PotF and PotF/D variants

Table S5: Concentrations of protein and ligand solutions used for the triplicate ITC measurements

Table S6: Full crystallization conditions, protein concentrations and cryogenic solutions used for the structure determination of all constructs

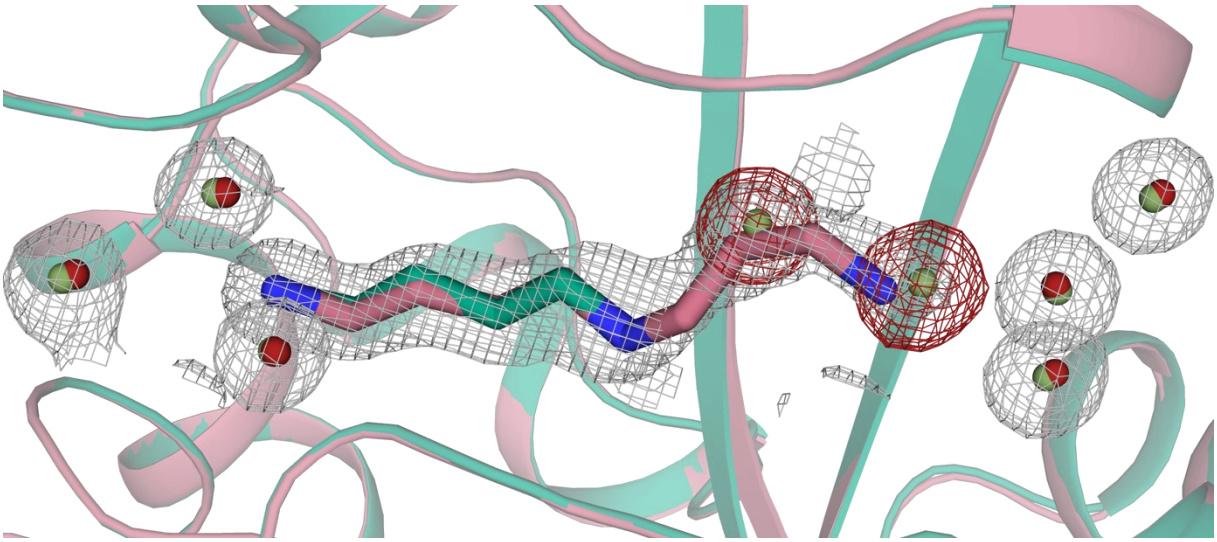


Figure S1: Binding pocket of PotFwt:PUT (green) and PotFwt:SPD (salmon) with waters (green and red spheres, respectively). Ligand molecules are depicted as sticks in the matching color. Maps for SPD and waters are shown as 2Fo-Fc-densities contoured at 1σ as gray mesh. The waters highlighted with a red mesh are displaced by SPD in PotFwt:SPD. Protein structures were visualized using PyMOL (The PyMOL Molecular Graphics System, Version 2.3 Schrödinger, LLC).

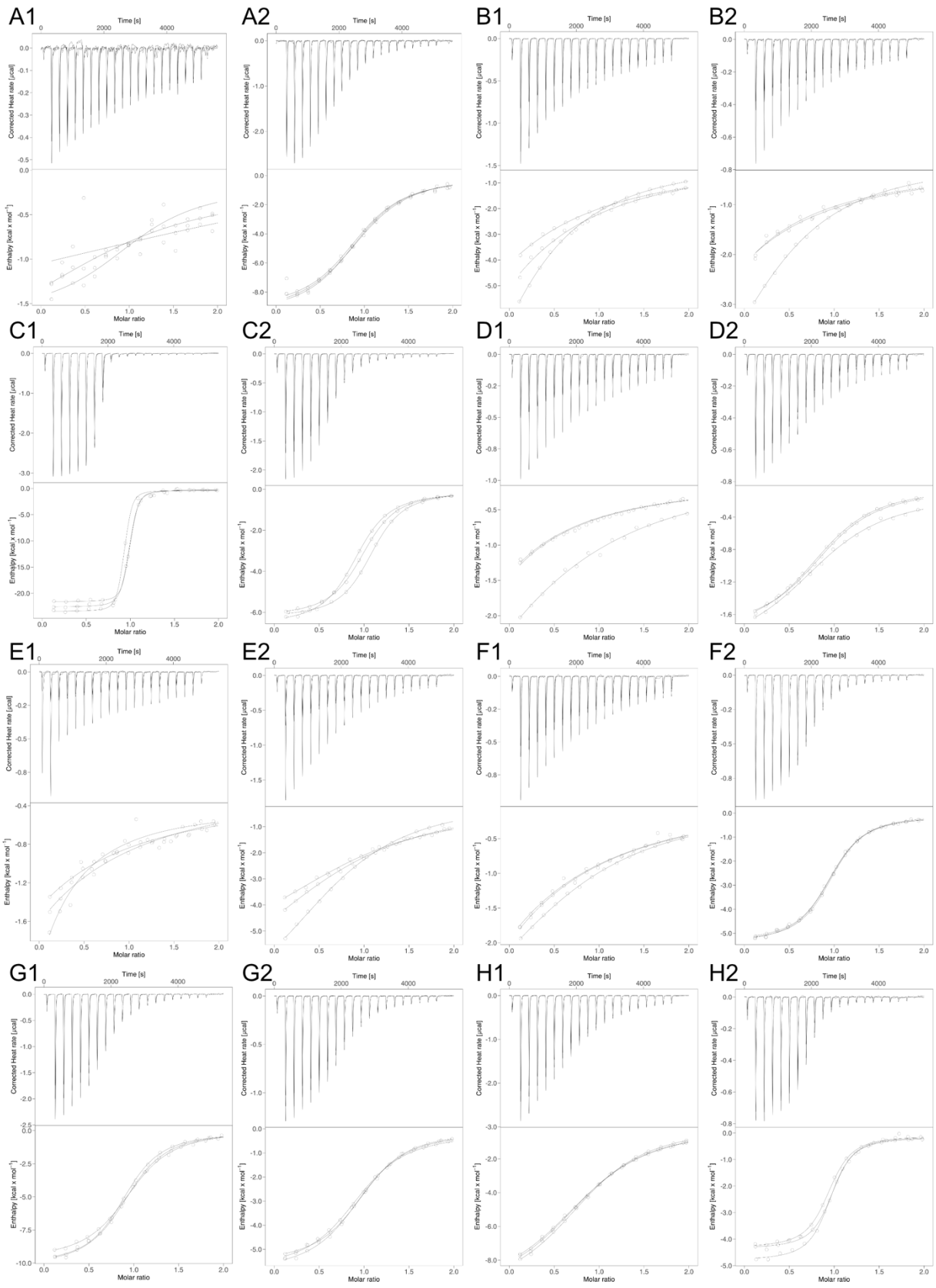


Figure S2a: ITC measurements and analysis related to Table 1. Measurements with PUT are labelled with 1 and measurements with SPD are labelled with 2. A: PotF/D, B: PotF_Prox, C: PotF_Abox, D: PotF_Dist, E: PotF_Abox_Prox, F: PotF_Abox_Dist, G: PotF_Abox-S87Y, H: PotF_Abox-A182D. Plots contain data from 3 biological replicates.

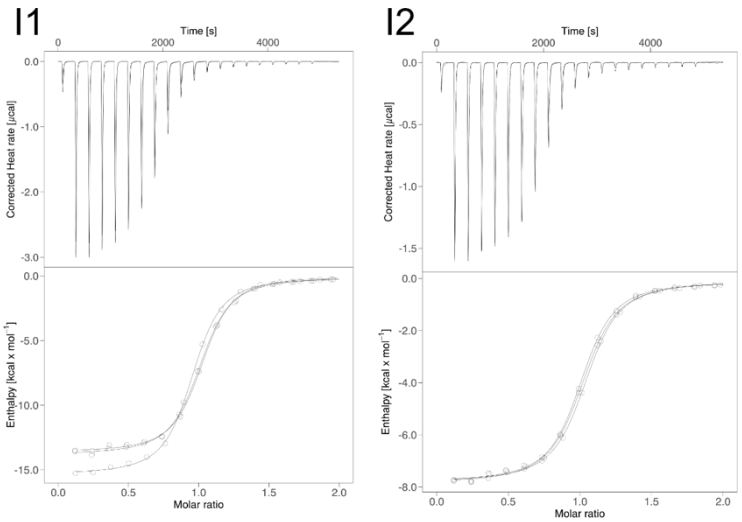


Figure S2b: ITC measurements and analysis related to Table 1. Measurements with PUT are labelled with 1 and measurements with SPD are labelled with 2. I: PotF_Abox-L348Q. Plots contain data from 3 biological replicates.

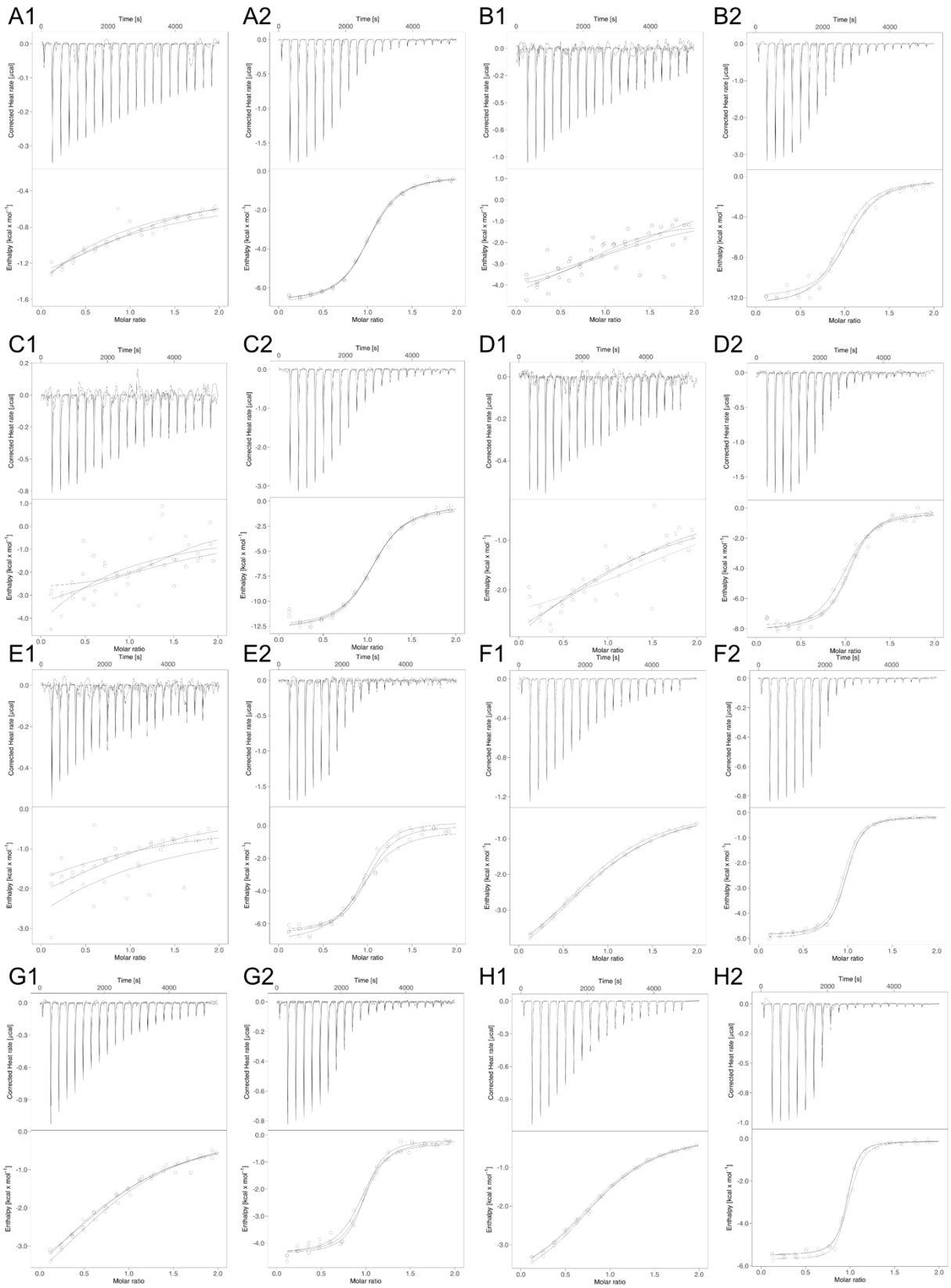


Figure S3: ITC measurements and analysis related to Table 2. Measurements with PUT are labelled with 1 and measurements with SPD are labelled with 2. A: PotF/D-E39D-Y87S, B: PotF/D-E39D-F88A, C: PotF/D-E39D-F88L, D: PotF/D-E39D-Y87S-F88Y, E: PotF/D-S247D, F: PotF/D-E39D-F88A-S247D, G: PotF/D-E39D-F88L-S247D, H: PotF/D-E39D-Y87S-F88Y-S247D. Plots contain data from 3 biological replicates

Table S1: Residue groups in PotF and their respective counterparts in PotD. All resulting mutations in PotF/D are listed as well.

	PotF	PotD	Mutations in PotF/D
Proximal (Prox)	S38, D39 & D247	T35, E39 & S232	S38T, D39E & D247S
Aromatic box (Abox)	W37, W244 & F276	W34, W229 & W255	F276W
Distal (Dist)	S85, S87, A182, E285, D278 & L348	S83, Y85, D168, E171, D257 & Q327	S87Y, A182D & L348Q

Table S2: RMS calculated by aligning the stated residues from PotF/D-E39D-F88A-S247D onto PotF/D-E39D-F88L-S247 via PyMol without outlier rejection cycles.

Residues ***Align F88A onto L***	**RMS over Cα Chain A/A**	**RMS over Cα Chain A/B**	**RMS over Cα Chain B/A**	**RMS over Cα Chain B/B**
29-369	0.481	0.478	0.370	0.409
29-85	0.272	0.259	0.179	0.190
86-123	0.610	0.544	0.518	0.472
124-369	0.335	0.376	0.321	0.371

Table S3: Crystallographic data and refinement statistics for all solved crystal structures.

	PotF/D	**E39D-F88L**	**E39D-Y87S**	**E39D-Y87S-F88Y**	**S247D**	**E39D-F88L-S247D**	**E39D-F88A-S247D**	**E39D-Y87S-F88Y-S247D**
PDB ID	7OYZ	7OYT	7OYS	7OYU	7OYY	7OYW	7OYV	7OYX
Data collection								
Wavelength [Å]	1.0370	0.9184	0.9184	0.9184	0.9184	0.9184	0.9184	0.9184
Resolution range [Å]	46.14 - 1.49 (1.54 - 1.49)	39.01 - 1.60 (1.68 - 1.60)	39.29 - 1.57 (1.63 - 1.57)	46.36 - 1.95 (2.02 - 1.95)	35.15 - 1.36 (1.41 - 1.36)	49.39 - 1.28 (1.33 - 1.28)	47.58 - 1.90 (1.97 - 1.90)	40.70 - 1.37 (1.41 - 1.37)
Space group	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 2_1 2_1 2	C 2 2 2_1	P 2_1 2_1 2
Cell parameter								
a, b, c, [Å]	37.20 81.85 111.74	37.14 78.03 118.44	37.26 78.58 114.44	37.14 78.61 114.83	37.06 82.15 111.07	116.88 71.29 92.42	73.61 122.58 190.32	117.16 71.49 92.66
α,β,γ [°]	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Total reflections	720880 (67643)	292090 (26997)	260238 (25374)	138481 (14023)	538447 (52254)	1306722 (114299)	594419 (40596)	1212027 (117504)
Unique reflections	56881 (5388)	46390 (4363)	47693 (4677)	25021 (2455)	73425 (7123)	198363 (19073)	62244 (5363)	163742 (16029)
Multiplicity	12.7 (12.5)	6.5 (6.2)	5.5 (5.4)	5.5 (5.7)	7.3 (7.3)	6.6 (6.0)	9.5 (7.6)	7.4 (7.3)
Completeness [%]	99.6 (95.7)	99.8 (99.7)	99.6 (99.5)	98.9 (99.2)	99.4 (96.2)	99.3 (96.4)	90.9 (79.5)	99.8 (98.9)
Mean I/sigma [I]	11.11 (0.55)	9.03 (0.66)	10.35 (0.73)	6.44 (0.79)	7.58 (0.61)	8.80 (0.59)	13.74 (1.00)	7.21 (0.46)
Wilson B-factor	23.7	23.3	21.8	30.1	14.5	16.3	33.1	16.6
No. of molecules per a.u.	1	1	1	1	1	2	2	2
Matthews coefficient	2.16	2.26	2.13	2.14	2.15	2.53	2.82	2.55
R_{merge}	0.136 (4.220)	0.128 (2.421)	0.098 (2.027)	0.206 (2.231)	0.171 (2.530)	0.104 (2.316)	0.121 (2.003)	0.154 (3.290)
R_{meas}	0.142 (4.397)	0.139 (2.645)	0.108 (2.242)	0.227 (2.451)	0.184 (2.721)	0.114 (2.536)	0.128 (2.139)	0.166 (3.540)
R_{pim}	0.040 (1.218)	0.054 (1.051)	0.045 (0.943)	0.093 (0.996)	0.068 (0.988)	0.044 (1.015)	0.039 (0.712)	0.061 (1.295)
$CC_{1/2}$	0.999 (0.281)	0.998 (0.273)	0.999 (0.256)	0.994 (0.262)	0.997 (0.280)	0.997 (0.233)	0.998 (0.375)	0.997 (0.240)
CC*	1.000 (0.662)	0.999 (0.655)	1.000 (0.639)	0.998 (0.644)	0.999 (0.661)	0.999 (0.615)	1.000 (0.739)	0.999 (0.622)

Table S3 (continued): Crystallographic data and refinement statistics for all solved crystal structures.

Refinement								
Reflections used in refinement	56879 (5388)	46390 (4363)	47683 (4677)	25007 (2581)	73415 (4555)	198355 (19073)	62236 (5363)	163728 (16025)
Reflections used for R_{free}	2101 (199)	2100 (197)	2098 (206)	1251 (135)	2099 (134)	2098 (202)	2101 (182)	2099 (206)
R_{work}	0.177 (0.425)	0.178 (0.362)	0.174 (0.340)	0.201 (0.320)	0.156 (0.334)	0.132 (0.332)	0.227 (0.403)	0.168 (0.416)
R_{free}	0.216 (0.446)	0.209 (0.380)	0.209 (0.345)	0.242 (0.355)	0.186 (0.367)	0.161 (0.346)	0.276 (0.459)	0.194 (0.461)
CC_{work}	0.958 (0.614)	0.965 (0.627)	0.966 (0.587)	0.963 (0.606)	0.974 (0.635)	0.978 (0.577)	0.946 (0.578)	0.978 (0.559)
CC_{free}	0.967 (0.495)	0.961 (0.646)	0.955 (0.635)	0.950 (0.493)	0.975 (0.699)	0.970 (0.438)	0.901 (0.630)	0.976 (0.522)
Number of non-hydrogen atoms	3106	3078	3155	2964	3444	6970	5749	6639
macromolecules	2750	2726	2764	2729	2855	5994	5432	5718
solvent	341	218	301	180	545	878	230	814
Protein residues	341	341	342	341	342	696	682	694
RMS bond lengths [Å]	0.002	0.011	0.005	0.005	0.007	0.008	0.007	0.017
RMS bond angles [°]	0.52	1.05	0.75	0.69	0.91	1.05	0.86	1.46
Ramachandran favored [%]	98.2	97.9	98.2	97.6	97.9	97.8	97.5	98.1
Ramachandran allowed [%]	1.8	2.1	1.8	2.4	2.1	2.2	2.5	1.9
Ramachandran outliers [%]	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Rotamer outliers [%]	0.7	0.7	1.0	0.7	1.3	1.8	1.0	1.0
Clashscore	1.27	5.03	2.61	4.85	2.39	6.09	6.49	3.54
Average B-factor	31.14	31.4	28.3	35.9	19.7	21.8	37.4	23.5
macromolecules	29.7	29.6	26.8	35.3	16.9	19.2	37.3	21.7
solvent	42.1	40.9	36.8	40.6	35.0	37.8	36.8	33.4
Number of TLS groups		1	1	1			2	2

Table S4: Oligonucleotides used for the generation of PotF and PotF/D variants via QuickChange during this study

Mutation	Orientation	Sequence (5‘-3‘)
S38T & D39E	forward	TTTATAACTGGACCGAATATATCGCCCCG
	reverse	CGGGGCGATATATTCGGTCCAGTTATAAA
E39D	forward	CATTTATAACTGGACTGATTATATCGCCCCGGACACG
	reverse	CGTGTCCGGGGCGATATAATCAGTCCAGTTATAA
S87Y	forward	GGTTCCATCTGCCTACTTTCTGGAGCGCC
	reverse	GGCGCTCCAGAAAGTAGGCAGATGGAACC
Y87S	forward	GGTGGTTCCATCTGCCAGCTTTCTGGAGCGCCAG
	reverse	CTGGCGCTCCAGAAAGCTGGCAGATGGAACCACC
F88A	forward	GTGGTTCCATCTGCCTACGCGCTGGAGCGCCAGTTGACT
	reverse	AGTCAACTGGCGCTCCAGCGCGTAGGCAGATGGAACCAC
F88L	forward	GGTTCCATCTGCCTACTTACTGGAGCGCC
	reverse	GGCGCTCCAGTAAGTAGGCAGATGGAACC
F88Y	forward	TTCCATCTGCCAGCTATCTGGAGCGCCAG
	reverse	CTGGCGCTCCAGATAGCTGGCAGATGGAA
A182D	forward	CTCTTTCCTGGATGATCCAGAAGAAGTTT
	reverse	AAACTTCTTCTGGATCATCCAGGAAAGAG
D247S	forward	CGGCTGGGCAGGTTCTGTCTGGCAGGCGT
	reverse	ACGCCTGCCAGACAGAACCTGCCCAGCCG
S247D	forward	ATCGGCTGGGCAGGTGATGTCTGGCAGGCG
	reverse	CGCCTGCCAGACATCACCTGCCCAGCCGAT
F276W	forward	AGGGGCGATGGCGTGGTTTGATGTATTCG
	reverse	CGAATACATCAAACCACGCCATCGCCCCT
L348Q	forward	AAGCTGTTCACTCAGAAAGTGCAGGATCC
	reverse	GGATCCTGCACTTTCTGAGTGAACAGCTT

Table S5: Concentrations of protein and ligand solutions used for the triplicate ITC measurements during this study

Protein Variant	Protein concentrations [µM]	Ligand	Ligand concentrations [mM]
PotF_Prox	***1)*** 348 ***2)*** 360 ***3)*** 261	PUT	***1)*** 3.28 ***2)*** 3.64 ***3)*** 2.70
	1) 348 ***2)*** 235 ***3)*** 261	SPD	***1)*** 3.28 ***2)*** 2.40 ***3)*** 2.70
PotF_Abox	***1)*** 146 ***2)*** 147 ***3)*** 140	PUT	***1)*** 1.16 ***2)*** 1.66 ***3)*** 1.53
	1) 328 ***2)*** 298 ***3)*** 283	SPD	***1)*** 4.31 ***2)*** 3.27 ***3)*** 3.09
PotF_Dist	***1)*** 588 ***2)*** 502 ***3)*** 472	PUT	***1)*** 6.00 ***2)*** 5.00 ***3)*** 5.00
	1) 602 ***2)*** 502 ***3)*** 470	SPD	***1)*** 6.00 ***2)*** 5.00 ***3)*** 5.00
PotF_Abox_Prox	***1)*** 425 ***2)*** 208 ***3)*** 209	PUT	***1)*** 4.25 ***2)*** 2.00 ***3)*** 2.10
	1) 409 ***2)*** 205 ***3)*** 209	SPD	***1)*** 4.25 ***2)*** 2.00 ***3)*** 2.10
PotF_Abox_Dist	***1)*** 563 ***2)*** 382 ***3)*** 425	PUT	***1)*** 6.00 ***2)*** 3.80 ***3)*** 4.25
	1) 194 ***2)*** 200 ***3)*** 209	SPD	***1)*** 2.00 ***2)*** 2.00 ***3)*** 2.09
PotF_Abox_S87Y	***1)*** 235 ***2)*** 250 ***3)*** 291	PUT	***1)*** 2.35 ***2)*** 2.55 ***3)*** 3.00
	1) 235 ***2)*** 250 ***3)*** 291	SPD	***1)*** 2.35 ***2)*** 2.55 ***3)*** 3.00
PotF_Abox_A182D	***1)*** 381 ***2)*** 433 ***3)*** 345	PUT	***1)*** 3.80 ***2)*** 4.40 ***3)*** 3.50
	1) 193 ***2)*** 217.5 ***3)*** 161	SPD	***1)*** 2.00 ***2)*** 2.20 ***3)*** 1.77
PotF_Abox_L348Q	***1)*** 240 ***2)*** 238 ***3)*** 227	PUT	***1)*** 2.40 ***2)*** 2.40 ***3)*** 2.35
	1) 240 ***2)*** 235 ***3)*** 245	SPD	***1)*** 2.40 ***2)*** 2.40 ***3)*** 2.45
PotF/D	***1)*** 470 ***2)*** 370 ***3)*** 360	PUT	***1)*** 4.70 ***2)*** 3.70 ***3)*** 3.60
	1) 395 ***2)*** 370 ***3)*** 360	SPD	***1)*** 4.00 ***2)*** 3.70 ***3)*** 3.60
PotF/D-E39D-Y87S	***1)*** 319 ***2)*** 308 ***3)*** 312	PUT	***1)*** 3.20 ***2)*** 3.10 ***3)*** 3.10
	1) 319 ***2)*** 308 ***3)*** 312	SPD	***1)*** 3.20 ***2)*** 3.10 ***3)*** 3.10
PotF/D-E39D-F88A	***1)*** 308 ***2)*** 317 ***3)*** 310	PUT	***1)*** 3.00 ***2)*** 3.10 ***3)*** 3.10
	1) 308 ***2)*** 317 ***3)*** 310	SPD	***1)*** 3.00 ***2)*** 3.10 ***3)*** 3.10
PotF/D-E39D-F88L	***1)*** 294 ***2)*** 296 ***3)*** 299	PUT	***1)*** 2.90 ***2)*** 2.90 ***3)*** 3.00
	1) 294 ***2)*** 300 ***3)*** 299	SPD	***1)*** 2.90 ***2)*** 3.00 ***3)*** 3.00
PotF/D-E39D-Y87S-F88Y	***1)*** 260 ***2)*** 265 ***3)*** 248	PUT	***1)*** 2.60 ***2)*** 2.60 ***3)*** 2.50
	1) 260 ***2)*** 265 ***3)*** 248	SPD	***1)*** 2.60 ***2)*** 2.60 ***3)*** 2.50
PotF/D-S247D	***1)*** 310 ***2)*** 305 ***3)*** 305	PUT	***1)*** 3.00 ***2)*** 3.00 ***3)*** 3.00
	1) 310 ***2)***305 ***3)*** 305	SPD	***1)*** 3.00 ***2)*** 3.00 ***3)*** 3.00
PotF/D-E39D-F88A-S247D	***1)*** 361 ***2)*** 378 ***3)*** 374	PUT	***1)*** 4.00 ***2)*** 4.00 ***3)*** 4.00
	1) 183 ***2)*** 193 ***3)*** 189	SPD	***1)*** 2.00 ***2)*** 2.00 ***3)*** 2.00
PotF/D-E39D-F88L- S247D	***1)*** 391 ***2)*** 360 ***3)*** 370	PUT	***1)*** 4.00 ***2)*** 3.60 ***3)*** 3.70
	1) 205 ***2)*** 180 ***3)*** 220	SPD	***1)*** 2.00 ***2)*** 1.80 ***3)*** 2.20
PotF/D-E39D-Y87S-F88Y- S247D	***1)*** 393 ***2)*** 388 ***3)*** 379	PUT	***1)*** 4.00 ***2)*** 4.00 ***3)*** 4.00
	1) 202 ***2)*** 186 ***3)*** 189	SPD	***1)*** 2.10 ***2)*** 2.10 ***3)*** 2.00

Table S6: Full crystallization conditions, protein concentrations and cryogenic solutions used for the structure determination of all constructs throughout this study.

Construct	Concentration	Condition	Cryogenic Solution
PotF/D-E39D-Y87S	15 mg/ml	0.085 M Sodium acetate pH 4.7, 0.17M Ammonium acetate, 32.5% PEG 4000, 15% Glycerol	-
PotF/D-E39D-F88L	30 mg/ml	0.1 M Sodium acetate pH 4.6, 0.2 M Ammonium acetate, 30% PEG 4000	25% Glycerol + SPD
PotF/D-E39D-Y87S-F88Y	40 mg/ml	0.1 M Sodium acetate pH 4.6, 0.2 M Ammonium acetate, 30% PEG 4000	25% Glycerol + SPD
PotF/D-E39D-F88A-S247D	30 mg/ml	2.4 M AmSO4, 0.1 M Bicine pH 8.3, 4.5% Jeffamine M600	50% 3.4 M Malonate pH 8.3 + SPD
PotF/D-E39D-F88L-S247D	40 mg/ml	2.4 M AmSO4, 0.1 M Bicine pH 9.0, 5% Jeffamine M600	50% 3.4 M Malonate pH 8.8 + SPD
PotF/D-E39D-Y87S-F88Y-S247D	30 mg/ml	2.4 M AmSO4, 0.1 M Bicine pH 8.3, 4.5% Jeffamine M600	50% 3.4 M Malonate pH 8.3 + SPD
PotF/D-S247D	40 mg/ml	0.1 M MES pH 5, 30% PEG 6000	25% Glycerol + SPD

10. Draft for Manuscript 3

A fluorescent biosensor for the visualization of Agmatine

Pascal Kröger[1], Andre C. Stiel[2,3], Birthe Stüven[4], Sooruban Shanmugaratnam[1,2], Susanne Schoch[5], Dagmar Wachten[4], Birte Höcker[1,2*]

Affiliations:
[1] Department of Biochemistry, University of Bayreuth, Bayreuth, Germany.
[2] Max Planck Institute for Developmental Biology, Tübingen.
[3] Helmholtz Center Munich, Institute of Biological and Medical Imaging, Munich, Germany.
[4] Institute of Innate Immunity, Medical Faculty, University of Bonn, Bonn, Germany.
[5] Institute of Neuropathology, Medical Faculty University of Bonn, Bonn, Germany.

[*] Correspondence: birte.hoecker@uni-bayreuth.de

Abstract:

Agmatine regulates multiple neurotransmitter systems. However, the underlying molecular mechanisms are not fully understood. To visualize agmatine dynamics in cellular networks and thereby unravel its physiological function, we generated a genetically encoded fluorescent agmatine biosensor (AGMsen) based on the periplasmic putrescine binding protein PotF. We first characterized the agmatine binding properties of the PotF receptor module based on its crystal structure and mutated the binding pocket to favor agmatine binding over other biogenic amines. We then applied the biosensor in different cell types to test its functionality and response to agmatine. Our results show that AGMsen allows visualization of agmatine in live cells, in particular in primary neuronal cultures. Thus, the sensor can be used as a tool to contribute to our understanding of agmatine distribution and dynamics, and its effect on neuronal functions *in vivo*.

Introduction:

Neurotransmitters are key regulators of brain function, and their dysregulation leads to a number of neuropathologies[1–3]. Optical imaging reveals spatio-temporal neurotransmitter distributions and as such largely contributes to understanding neuronal function and the development of new concepts for therapies to treat neuropathologies. Prime examples are genetically encoded biosensors that allow to visualize neurotransmitters like dopamine, glutamate, GABA, and serotonin in different organisms[4–8]. A neurotransmitter that is yet at the advent of its research is agmatine (AGM), the decarboxylated form of the amino acid arginine. AGM can be taken up by axon terminals and is localized in synaptic vesicles, from which it can be released in a calcium-dependent manner[9,10]. This is in line with the proposed neurotransmitter-like function of AGM since it shows an influence on multiple molecular

targets that include neurotransmitter systems like nicotinic, imidazoline I_1 and I_2, α_2-adrenergic, glutamate NMDAR, and serotonin 5-HT2A and 5HT-3 receptors[11]. Even with this multitude of different targets, agmatine is still referred to as a neuromodulator or co-transmitter because no agmatine-specific, postsynaptic receptor or agmatinergic system has been identified yet. Still, the most common central nervous system disorders seem to have a polygenic origin and due to its large presence in the peripheral and central nervous system, agmatine is conjectured to be a "magical shotgun", a non-selective drug with multiple targets, which can lead to more effective treatments[11,12]. AGM function in the central nervous system includes antidepressant-like effects[13,14], protection against schizophrenia[15], improvement of cognitive dysfunction in Alzheimer's disease[16], and improvement of multiple sclerosis[17].

However, the molecular mechanisms underlying its function in the brain are unknown. This is also due to a lack of knowledge on the spatio-temporal effector function of agmatine *in vivo*. Hence, the development of a genetically encoded biosensor that allows to visualize agmatine localization and dynamics in a non-invasive manner is key to understand the role of agmatine in neuro(patho)physiology. Here, we developed a fluorescence-based, genetically encoded AGM biosensor, consisting of an AGM binding-domain derived from the putrescine periplasmic binding-protein (PBP) PotF from *E. coli* fused to a super-folder circularly permuted GFP (sfcpGFP[18,19]). The large conformational change of the PBP upon AGM binding can be transmitted to the GFP, thus, altering the chromophore environment and increasing GFP fluorescence [20,21]. We chose PotF due to its promiscuous binding of biogenic amines as characterized in prior studies[22, 23]. Our engineering employed a semi-rational approach combined with a medium throughput fluorescence screening method to optimize linker positions and improve AGM specificity. We applied the final AGM biosensor, AGMsen, in different cell types *in vitro*, i.e., in primary neuronal cultures. Our results show that the agmatine sensor AGMsen can be used as the first of its kind to track agmatine in various experimental set-ups, which will allow to shed light on AGM function not only *in vitro* but also *in vivo*.

Results:

Scaffold selection and initial sensor construction

In a previous study, we discovered AGM as a natural ligand (K_D = 0.22 µM) for the periplasmic binding-protein PotF. The high affinity of PotF towards AGM can be explained by the perfect mimic of the binding mode of putrescine (PUT), as determined by crystallography (FIGURE 2a & b)[22]. To facilitate the engineering of high agmatine affinity and specificity by reengineering the binding pocket, we first established a sensor for the native ligand putrescine. We inserted the circular permuted green fluorescent protein (cpGFP, carrying mutation K12R) in the putrescine binding-protein PotF of *E. coli*. Entry sites were chosen based on assumed major backbone displacement in the hinge region, as

elucidated on the bound structure of *E. coli* PotF (1A99[24], Figure S1). Most entry sites gave rise to a sensor that responded to addition of putrescine (PUT; Figure S2) with only position 332 showing no response. Most of the sites resulted in a sensor that decreased its fluorescence upon PUT addition. Only the sensor 328/332 showed a fluorescence increase of $\Delta F/F_0$ ~0.3 upon the addition of 50 µM PUT. As we aimed for a sensor that increased its fluorescence upon binding, we continued with the sensor 328/332. We next explored the effect of different linkers between N-PotF and cpGFP as well as cpGFP and PotF-C. We limited ourselves to changes of the two residues exiting (N-linker) and entering PotF (C-linker), respectively. From this screen, A327L/E328I of the N-linker emerged as the best variant with the C-linker remaining unchanged (N332/P333). Since the screen was based on measuring fluorescence changes after PUT addition to crude extract of *E. coli*, we introduced mutations (S87Y, F276W) as they reduce initial PUT affinity from the native nanomolar to micromolar range[23] to prevent confounding effects of intracellular ligand. Next, we replaced the initially used cpGFP by the circular permuted variant of superfolder GFP[19] (sfcpGFP). We took this preliminary sensor (PUTsen) as a template for the engineering of a sensor with high sensitivity and specificity for agmatine.

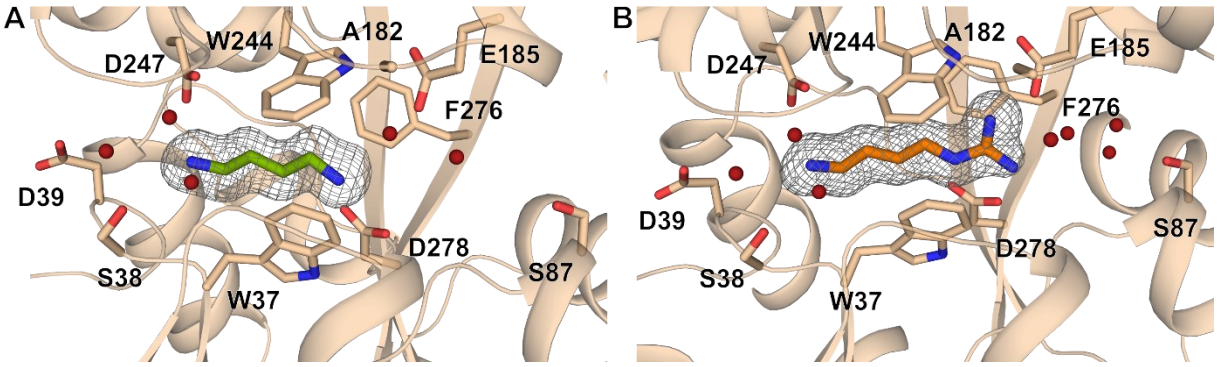


Figure 2: Overview of the binding pocket of PotF in complex with putrescine (A) and agmatine (B). Agmatine binds to wild-type PotF with high affinity as it mimics the putrescine binding mode. Ligand molecules and residues forming the binding pocket are shown as sticks. 2mFo-DFc maps for ligands are shown as gray mesh contoured at 1 σ using PyMOL.

Tuning affinity towards agmatine

The PotF receptor module needs to be engineered to specifically bind AGM so that sensor performance can be assessed in cell culture experiments and possibly *in vivo* in the future. The natural biogenic amine ligands of wild-type PotF play an essential role in cell proliferation, they will be present in experimental set-ups, and can even be included in media[25–28]. An unspecific sensor would recognize these other molecules, thus leading to high background levels or false positive results.

A high initial affinity to the target molecule is desirable for engineering, since shaping ligand specificity in a promiscuous binder like PotF often comes at the cost of losing affinity[23]. As a starting point for possible mutations, we first investigated PotF/D, a PotF variant into which the binding pocket of the homologous PotD was grafted[29], and were surprised to already find high AGM affinity in isothermal

titration calorimetry (ITC) measurements (K_D = 4 µM, Figure 3a). Additionally, PotF/D does no longer bind PUT and recognizes spermidine (SPD) with an affinity of 37 µM, making it a first construct with improved AGM specificity compared to PotF. Building on a thorough analysis of the differences between the binding pockets of PotF and PotF/D[23], we randomly combined mutations (Table S1) from this study and started to screen biosensor constructs for affinity and specificity to AGM in *E. coli* lysate based on fluorescence changes upon ligand addition. The most interesting variants showed a high fluorescence increase in response to AGM and low or no response to other PotF ligands like PUT, SPD, and cadaverine (CDV). These variants were further purified by affinity chromatography to limit the influence of lysate compounds and were re-screened for biogenic amine binding. The results of our lysate and protein assay screening round that yielded the final variant is shown in Figure S3.

This final variant from our screening carried the mutations S87Y and A182D. Additionally, we aimed to generate a non-binding control sensor to confirm that changes in fluorescence are triggered by ligand binding only. Therefore, we introduced the mutation D247K, which is located in the first responding region of the protein to ligand encounter[22]. By introducing a lysine at this position, we can mask the primary amine binding-site in PotF and generate a sensor that cannot bind any of the tested polyamines. We verified the effects of these mutations on AGM specificity by measuring binding affinities with ITC for the PotF receptor modules without fusion to a fluorescence protein: PotF-S87Y-A182D displayed an affinity of 0.3 µM for AGM and no apparent K_D for the other tested ligands PUT, SPD, and CDV (Figure 3b). The control sensor (PotF-D247K) showed no binding for any of the tested ligands (Figure 3c).

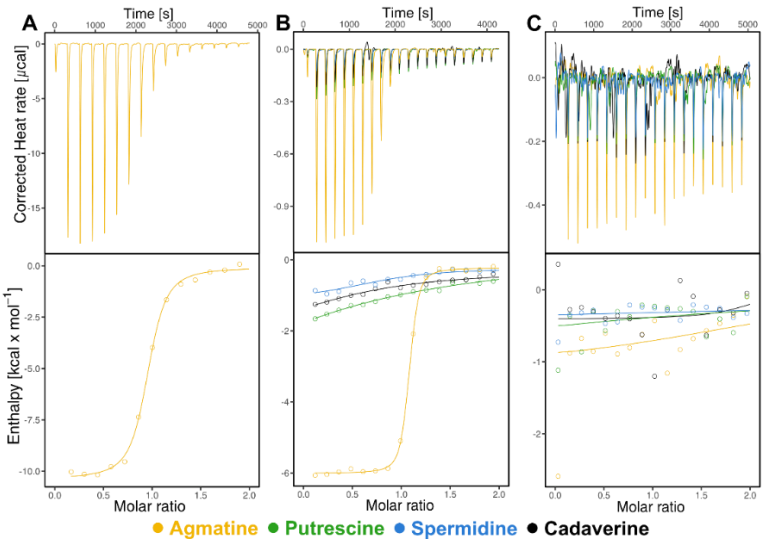


Figure 3: ITC measurements for PotF/D (A), PotF-S87Y-A182D (B) and PotF-D247K (C). PotF/D displays a K_D of 4 µM for AGM. The PotF-S87Y-A182D receptor molecules displays an affinity of 0.3 µM (**B**). Any binding of other polyamines was not detectable in ITC. The designed control receptor PotF-D247K (**C**) shows no binding of any tested ligand. Measured thermodynamics and stoichiometry of all measurements can be found in Table S2.

To examine and understand the binding of AGM further, we solved the crystal structure of PotF-S87Y-A182D in complex with AGM and compared it to wild-type PotF. Interestingly, the positioning of the two lobes in PotF is slightly displaced and more open in the S87Y-A182D variant (Figure 4a). We analyzed the opening and twisting between the two lobes in this structure as described before[22] and found 10.9° wider opening and 7.1° wider twisting angles compared to the closure observed in wild type. Still, most AGM interacting residues in the binding pocket do not differ from the wild type (Figure 4b & 2b). Since the lobes are slightly further apart, distances between interacting residues and important binding pocket waters can change. These distances are compensated by the bulkier ligand AGM but are most probably detrimental for high affinity binding of slim ligands like PUT and SPD.

The control receptor module PotF-D247K showed no apparent binding to all tested polyamines in ITC (Figure 3c). Furthermore, the crystal structure confirmed that the amine of the newly introduced lysine 247 occupies the primary amine binding-site of the PotF pocket and, thereby, blocks ligand binding as expected (Figure 4c).

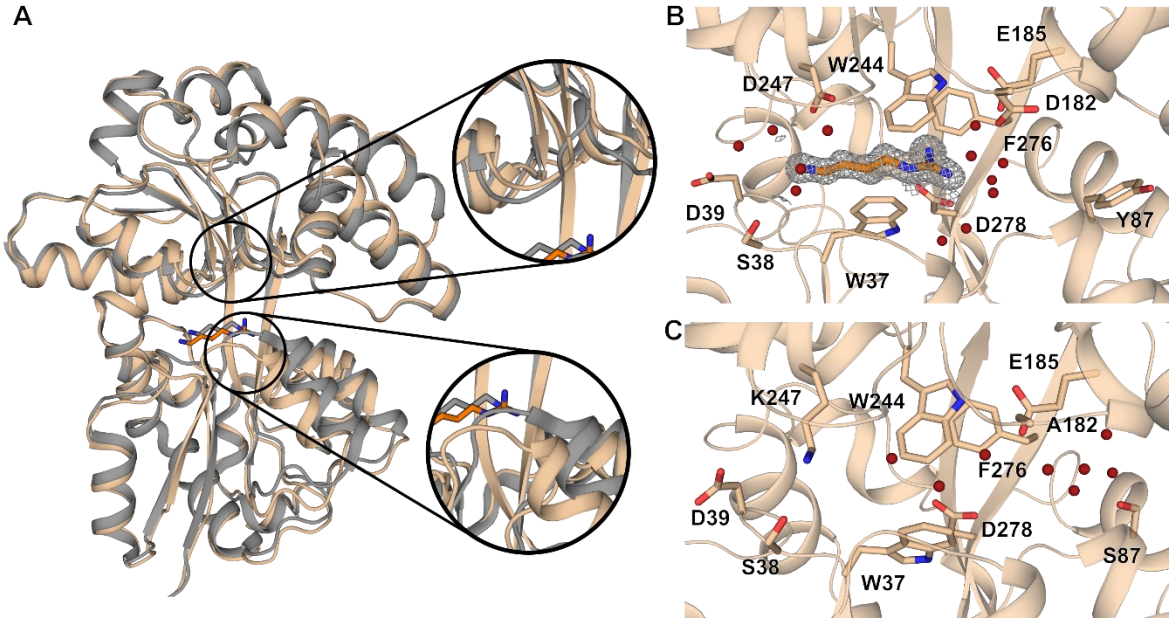


Figure 4: Alignment of the PotF receptor module (S87Y-A182D) of AGMsen in complex with AGM (pale orange) and PotF in complex with PUT (grey; A). Overview of the binding pocket of the PotF receptor module of AGMsen (B) and the control sensor (C). The structure of the AGMsen receptor module is slightly more open than wild type PotF. The positioning and interactions of agmatine in the pocket remain mainly unchanged compared to in the wild type PotF (Figure 2b). In the control sensor residue K247 occupies the primary amine binding site and, thereby, prevents ligand binding. Ligand molecules and residues forming the binding pocket are shown as sticks. 2mFo-DFc maps for ligands are shown as gray mesh contoured at 1 σ using PyMOL. Structure statistics can be found in table S4.

Purified sensor performance

After the promising ITC results for the isolated receptor modules, we assessed the binding capabilities in the framework of the biosensor with the added sfcpGFP. We named the sensor containing S87Y and A182D AGMsen and conducted fluorescence-based dose-response measurements with AGMsen purified from *E coli* upon addition of AGM, PUT, SPD, and CDV at concentrations ranging from 0.1 µM to 10 mM. Here, AGMsen showed a dynamic range ($\Delta F_{max}/F_0$) of 3.0 and an AGM affinity of 38 µM (Figure 5a, Table S3). For the control sensor, only a small increase in fluorescence intensity at very high AGM concentrations was observable in the dose-response measurements, which corresponds to a K_D of 3.6 mM (Figure 5b, Table S3).

Overall, PotF receptor variants showed good performance in ITC and fluorescence-based dose-response measurements as sensors, albeit with differences in affinity between the two techniques. This is most likely due to the influence of the inserted GFP on the PotF receptor modules. We introduced the sensors into pcDNA3.1. and pDisplay for eukaryotic expression to further characterize them in cell culture experiments. With pDisplay, we introduced an N-terminal signal peptide to guide the sensor to the secretory pathway and a C-terminal transmembrane domain to anchor the sensor on top of the cells.

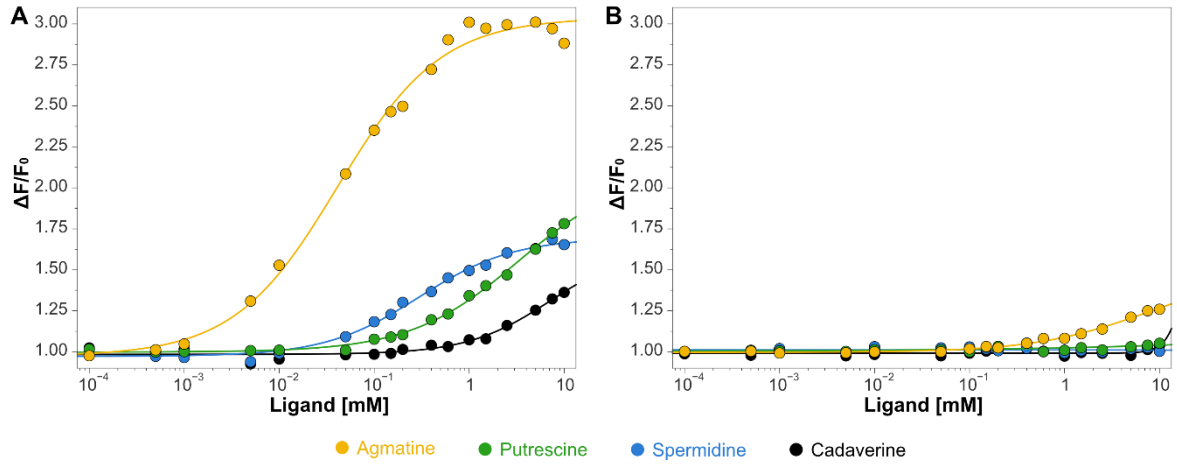


Figure 5: Dose-response curves in response to different biogenic amine ligands for purified AGMsen (A) and purified control sensor (B) from *E. coli*. (**A**) AGMsen that carries the mutations S87Y and A182D shows a dynamic of 3.0 and an affinity of 38 µM for AGM. Additionally, AGMsen shows low affinities for SPD, PUT and CDV with K_D's of 244 µM, 1.9 mM, and 3.4 mM, respectively. (**B**) The control sensor displays residual AGM affinity (3.6 mM) accompanied by a low dynamic range ($\Delta F_{max}/F_0$ 1.3). Data points represent the mean of triplicates and the fit was done with the Hill equation using the fit-o-mat[30]. All K_D values can be found in Table S3.

Sensor characterization in cells

First, we performed immunocytochemistry to visualize the expression of AGMsen in the pcDNA3.1 and the pDisplay backbone in HEK293 cells. To analyze anchoring of the sensor into the outer cell membrane, we performed labeling with and without cell permeabilization (Figure 6a). The expression of the sensor in both constructs was verified in permeabilized cells but membrane anchoring was only observed in the non-permeabilized cells expressing pDisplay-AGMsen. Thus, the sensor is secreted and incorporated into the outer cell membrane.

We next characterized the sensor performance in different cell-based set-ups. Therefore, we expressed AGMsen and the control sensor in HEK293 and performed dose-response measurements in cell lysate, which we compared to the prior *E. coli* assays to ensure the constructs behave similarly and are not influenced by specific eukaryotic features, e.g., posttranslational modifications. No notable differences between the assay results from *E. coli* proteins and HEK lysate for affinities of all ligands were observed (Figure 6b & c and Figure 5a &b, Table S3), thereby confirming our bacterial set-up as a robust screening platform that can be used to predict sensor function in eukaryotic cells. The only observable difference is a 20% drop in dynamic range upon AGM addition, which is expected as some lysate components might quench the fluorescent signal (Figure 6b). We then measured dose-response curves for HEK293 live-cell suspensions in a PTI spectrofluorometer for pDisplay-AGMsen (Figure 6d). Here, the secreted sensor displayed a lower K_D of 525 µM and dynamic range of 1.6 compared to the lysate from intracellularly expressed sensor (K_D = 35 µM, $\Delta F_{max}/F_0$ = 2.4, Figure 5c). Still, the sensors response is likely enough to visualize release events as neurotransmitters reach high concentrations in the synaptic cleft (as discussed below).

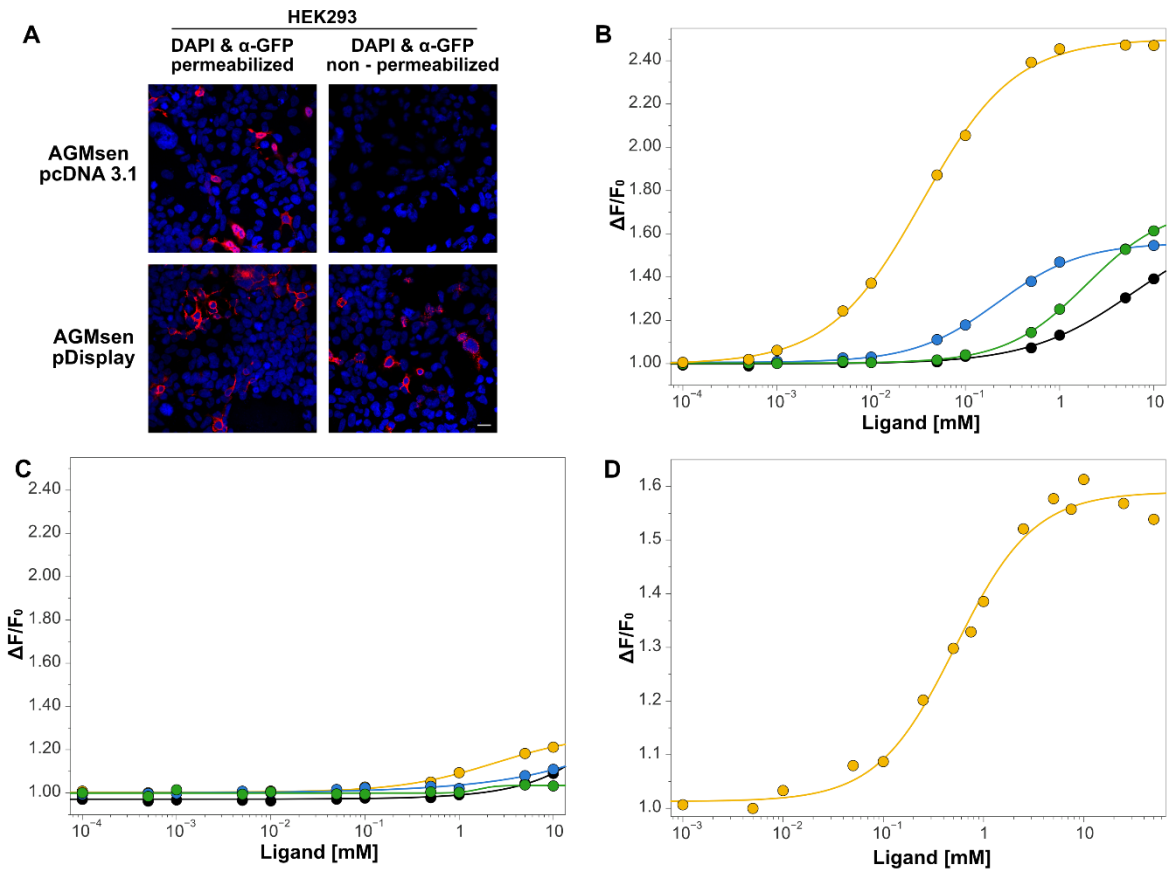


Figure 6: Immunocytochemistry of HEK293 cells expressing AGMsen in pcDNA3.1 or pDisplay, with or without permeabilization (A). Dose-response of AGMsen (B) and the control sensor (D) in HEK293 cell lysate. AGMsen was expressed in HEK293 cells using either the pcDNA3.1 or the pDisplay backbone (**A**). In permeabilized cells, GFP expression was detected in both cases, while in non-permeabilized cells, GFP was only observed for the pDisplay variant, confirming the correct incorporation of the sensor after secretion into the outer cell membrane. Cells are counterstained with DAPI. Scale bar = 20 µm. In HEK cell lysate the affinities for AGMsen (**C**) are similar measured affinities for purified sensor from *E. coli* with the dynamic range being slightly lower (AGM K_D = 35 µM, $\Delta F_{max}/F_0$ = 2.4). The control sensor also performs the same displaying residual AGM affinity of 2.4 mM accompanied by a low dynamic range of 1.2. The response of AGMsen displayed on top of HEK293 cells (**D**) to AGM shows a reduced K_D (525 µM) and dynamic range (1.6) compared to the non-displayed variant. Data fits were done with the Hill equation using the fit-o-mat[30]. All K_D values can be found in Table S3.

Therefore, we subcloned pDisplay-AGMsen and the control sensor into a pAAV backbone with a synapsin promotor for neuronal expression and generated Adeno-associated virus containing the sensors. We next transduced primary rat hippocampal neurons with the pDisplay-AGMsen and control sensor and observed expression after 8 – 10 days (Figure S3). To quantify the AGM response, we analyzed neurons for each added AGM concentration using ImageJ as described in Figure S4. Results for different cells in the same experiment do scatter (Figure S5), nevertheless, clear trends are observable (Figure 7 & S5). Overall, fluorescence signal changes in neurons were much lower compared to HEK but the resulting K_D of ca. 700 µM is in the same regime as for the sensor displayed on HEK cells. This confirms functionality for AGMsen displayed on rat hippocampal neurons.

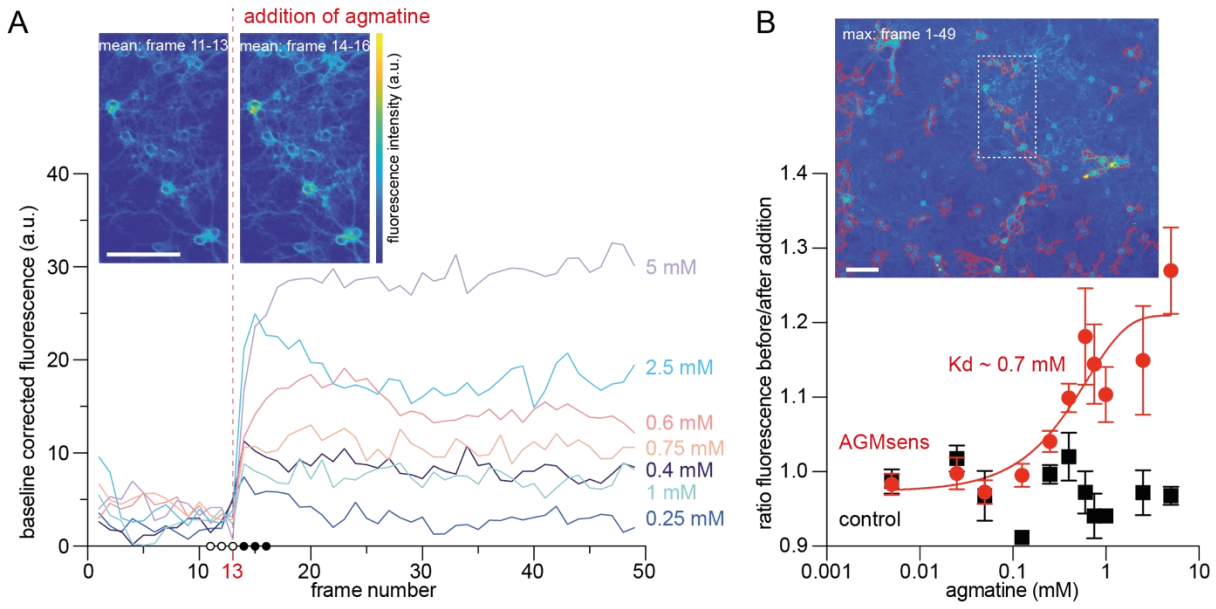


Figure 7: Response of AGMsen in neurons. AGMsen was expressed in rat hippocampal neurons and imaged using a Zeiss Observer.Z1 widefield microscope before and after the addition of agmatine (dashed red line in a). **A**: Exemplary curves of single ROIs after the addition of AGM at concentrations between 5 and 0.25 mM. Top left insets show magnifications of the 5 mM addition data (dashed white rectangle in **B**) with the mean of three frames before and after agmatine addition shown (hollow and filled circles on x-axis). **B:** Mean fluorescence change ratio of AGMsen and control sensor at different AGM concentrations. Shown is the mean pixel value of a minimum of 10 ROIs per datapoint. Error bars indicate the standard deviation. The AGMsen data can be fitted with a sigmoidal yielding a K_D of ~0.7 mM (quality of fit r-squared 0.83). The inset shows a maximum intensity projection of the data for the 5 mM addition. The automatically chosen ROIs are indicated in red. All scale bars are 100 µm. All color bars as in inset in **A**. Curves for AGMsen and control sensor for all concentrations and ROIs can be found in Figure S5 and a detailed description of the analysis procedure can be found in Figure S4.

Discussion:

Design aspects

In the engineering of the sensor the functional guanidino group proved to be a major player to achieve selectivity for an otherwise indiscriminate polyamine binding protein. The unconventional pairing of the high similarity of the binding mode with the uniqueness of a different functional group was the key ingredient in the sensor design process. In the sensor crystal structure, the positioning of AGM does not differ from the one in wild type suggesting that the other amines still recognize the primary amine binding site with rudimentary affinity. The main changes of the structure are in the distal site of the pocket (S87Y & A182D, Figure 4) where in the case of AGM, the guanidino group is located. It seems that this large and bulkier group withstands the mutational changes and maintains high affinity while the smaller primary amine groups of the other ligands loose most of their affinity. This is most probably caused by slightly incomplete closure of the PotF-S87Y-A182D receptor module of AGMsen, which results in worse coordination for the less bulky ligands like PUT and SPD. This also has consequences

for the sensor, whose functionality depends on closing of PotF, hence, it might display a lower dynamic range than possible with optimal closure.

We observed discrepancies between the actual sensor affinities and the affinity of the receptor module by itself determined by ITC. On the one hand, the sensor is purified only by IMAC and mostly screened in lysate which can affect the performance of the receptor module. On the other hand, GFP with a size of 28 kDa is inserted into a movement sensitive region of PotF and, therefore, can influence PotF by exerting forces or resistance on the closing mechanism. It is notable that AGMsen does not completely omit binding of the other amines, especially for the longest comparable ligand SPD. Still, fluorescence signal and the affinity in purified sensor are much lower for SPD (245 µM), hence, this will most likely not influence the results since the sensor would always prefer AGM in an *in vivo* setting.

In addition, our knowledge of the PotF pocket and dynamics allowed us to construct a functional control sensor. D247K occupies the primary amine binding site, where typically all characterized ligands of PotF would bind in the same fashion. It is notable that the crystal structure displays a nearly fully closed conformation which could explain the good signal of the control sensor. This allows for reliable controls and the assessment of robustness for future experimental set-ups since drifts or other issues in fluorescence signal will be easily traceable with the control sensor.

Sensor performance

As there are no significant differences comparing the assay results between *E. coli* and HEK cells (Figure 5a & b and & 6b & c), it confirms the robustness of our screening system and dispelled any concerns that codon usage would impact results since the sensor genes were not optimized for HEK cell expression. The HEK displayed AGMsen performed worse than the non-displayed construct, showing a lower apparent affinity and dynamic range. This suggests that the dynamics of the sensor are further influenced by the addition of the secretion signal, linkers, and transmembrane domain. This is not unexpected, as like the insertion of sfcpGFP, these modifications can influence the dynamics of PotF and its closing mechanism. We know from ITC measurements that the potential K_D of AGMsen's PotF receptor variant is 0.3 µM. Therefore, the sensor might be further optimized by improving the different linkers to obtain better signals and to provide less influence on the affinity, especially at lower ligand concentrations. Moreover, the epitope tags from the construct can be removed as in pMinDisplay, a version of pDisplay missing the HA tag[4]. Another factor that is important to address are reaction conditions: the displayed sensor ideally works in an extracellular environment and the other sensor intracellularly, meaning a different experimental set-up is required i.e., changing salt content and concentrations. Ligand binding in the PotF pocket is highly mediated by coulomb forces between ligand amines and carboxyl harboring sidechains[22], therefore the ionic strength of the buffer could potentially impact affinities. Nonetheless, the highest potential affinity (0.3 µM, Figure 3b, Table S2)

for the PotF receptor module was determined by ITC in a high salt buffer containing 300 mM NaCl, hence we believe the impact of the different extensions on PotF play a bigger role than slight differences in buffer conditions.

The importance of improving dynamics further becomes apparent when assessing the neuron data. Due to generally lower signal increase, background noise can have a strong influence on the results and neuron culture is overall more inconsistent. This could be counteracted by having a stronger, more prominent signal. The response of AGMsen displayed on neurons was slightly lower but in a similar regime as observed for the HEK displayed AGMsen. One reason could be the difficulties with signal intensity and high fluctuations between cells in the same measurements, resulting in high errors. Additionally, neurons were prepared from rat, a different organism, which could also lead to small changes in performance. Nevertheless, HEK cell and rat neuron displayed AGMsen are comparable and newly optimized constructs could be easily characterized and compared in HEK cells first before moving to other organisms and experimental set-ups to save time and resources.

Further engineering approaches

There are further possibilities for improvements in case linker engineering only increases the dynamic range but not the affinity. Recent studies revealed a new possible strategy to engineer binding in PBPs. Van den Noort et al.[31] discovered that distant mutations that tweak free energies of available conformations can influence affinity in the maltose-binding protein (MBP). Using this method, we do not have to reengineer the already optimized pocket for binding again but target other areas of the receptor module. A final remark considering the neuroscientific research is that we only investigated sensor functionality and response to external AGM addition. Analyzing specific effects of AGM by tracking of natural release upon stimuli in the synaptic cleft will be essential. A way to target the post-synapse and improve spatial representation is by adding neuroligin (Nlgn) as a localization tag to the displayed construct as it has been done for similar sensors like iSeroSnFR[8].

Cellular Agmatine mechanisms

Besides its neuronal effects AGM has a strong influence on polyamine metabolism in general. It can inhibit ornithine decarboxylase (ODC), the rate limiting enzyme of polyamine biosynthesis[32]. As mentioned before, polyamines are important for proliferation and it has been shown that ODC levels are elevated in tumors, making it a proto-oncogene[33,34]. Therefore, AGM is an interesting antagonist to proliferation because through ODC inhibition, intracellular polyamine levels are reduced, and growth is hindered. AGM is taken up by mammalian cells by using the polyamine transport system, which is positively correlated with the rate of cell proliferation[35]. To quantify uptake in different cell types, radioactively ([^{3}H]) labeled AGM or the determined ODC activity was used[32,36–41]. Our

intracellularly expressed sensor is a promising tool that could provide a solution to measuring AGM import without the need for labeled compounds due to the fluorescent readout.

Finally, it is worth mentioning that we can alter the polyamine preference in PotF by only a few point mutations making engineering towards other ligands straightforward. Our screening in *E. coli* is fluorescence based and can be scaled up to screen libraries for potential candidates. Combining these features with the presented results allows to create ligand independent tests and apply this knowledge to develop further PotF-based sensors.

Conclusion:

Using our understanding and detailed prior studies on PotF, we were able to engineer an AGM sensor, which is functional in HEK cells and rat hippocampal neurons. The sensor is available in multiple vectors for expression inside cells (pcDNA 3.1) or for display on cells (pDisplay and pAAV-pDisplay). We also provide a non-binding control sensor that can be used to directly confirm the authenticity of signal changes in AGMsen being induced by AGM binding and no other components of the experimental set-up. We are aware that there is room for improvement regarding dynamic range and affinities in the current sensor framework to track weaker AGM signals. As outlined above these imperfections can be addressed with further engineering in the future. Still, for AGM to fulfill its neuromodulatory functions, like other transmitters, high concentrations in the synaptic cleft are needed. It has been shown that AGM is colocalized with glutamate, the major excitatory neurotransmitter of the CNS, in the hippocampus[42]. The concentration of glutamate at cultured hippocampal synapses peaks at 1.1 mM[43] and it has been described that concentrations surpass 1 mM after stimulation by an action potential[44]. It is conceivable that AGM is capable to reach concentrations in this range which is reliably detected by our sensor. In addition, the high µM affinity allows the sensor to not be influenced by small fluctuations in AGM concentrations and, hence, it should only signal genuine release events. Therefore, AGMsen is a first important step towards better understanding AGM-specific mechanisms in a non-invasive manner and we hope by providing this tool we can work closely together with users to further tune it towards different needs.

Methods:

Initial construction in pET21b(+)

Initial sensor construction was conducted by a two-step splicing overlap extension PCR with Phusion polymerase (New England Biolabs) using standard protocol with an adjusted elongation time of 50 s. Used oligonucleotides can be found in Table S5. Resulting fragments were purified after agarose gel electrophoresis. After subsequent restriction hydrolysis with NdeI and XhoI (both from Thermo Fisher Scientific) the fragment was ligated into equivalently linearized pET21b(+).Top10 cells were transformed with the reaction mixture via heat shock. DNA of overnight cultures was isolated using NucleoSpin® Plasmid EasyPure-Kit (Machery & Nagel) according to the manufacturers protocol and correct implementation of mutations was checked by sequencing at Eurofins Genomics using standard primers.

Targeted mutagenesis

Mutations for the construction of different AGMsen and PotF variants were introduced by a modified QuickChange PCR utilizing KAPA® polymerase (Roche) followed by an additional ligation step using T4 DNA ligase (New England Biolabs) according to the manufacturers protocol. All used oligonucleotides for mutations presented in this work can be found in table S5**.** Transformation, DNA isolation and construct verification was conducted as described above.

Cloning into pDisplay

Sensor constructs were amplified with Q5 polymerase (New England Biolabs) using oligonucleotides (Table S6) that introduced SfiI and PstI (both from Thermo Fisher Scientific) restriction sites. After restriction hydrolysis using standard protocol for the respective enzymes and subsequent purification, fragments were ligated into equivalently linearized pDisplay. Transformation, DNA isolation and construct verification was conducted as described above.

Cloning into pcDNA

Final and control sensor fragments were cloned into pcDNA using Gibson assembly. Fragments were amplified using a standard Q5® High-fidelity DNA-Polymerase protocol (New England Biolabs). Primers were designed using NEBuilder® and are shown in table S6. The PCR added a Kozak sequence to facilitate translation in eukaryotic cells. Amplified vector and insert fragments were purified from 1% (w/v) agarose gels using NucleoSpin ® Gel and PCR Clean-up kit (Macherey & Nagel) according to manufacturer's protocol. Subsequently, fragments were mixed in a 1:1 ratio (n/n) to a final DNA quantity of 200 ng in 5µl reaction volume. DNA mix was added to Gibson Assembly reaction mix and

incubated for 1 h at 50°C. Transformation, DNA isolation and construct verification was conducted as described above.

Cloning of sensor genes into an adeno associated virus (AAV) vector

For expression on the surface of neurons, the genes for the displayed AGMsen and the control sensor were placed under the control of the human synapsin I promoter. Therefore, both sensors were excised from the respective pDisplay-plasmids including the N-terminal Igk leader and HA, as well as C-terminal Myc and PDGFR domain sequences using XbaI and HindIII. Afterwards the genes were subcloned into an AAV vector (rAAV-Syn1-MCS; serotype 2/1) and used for recombinant AAV production. rAAV-Syn1-MCS was kindly provided by Martin Schwarz (Bonn University Medical School).

Protein expression and purification

Receptor modules were expressed and purified as described in Kröger et al 2021a &b.[22,23]

Sensor variants were overexpressed in autoinduction medium ZYM-5052[45] for 18 h at 30°C and 180 rpm using *E. coli* BL21 (DE3) cells. After expression, cultures were pelleted (3500 g, 20 min, 4°C) and resuspended in 20 mL buffer S1 (20 mM Tris pH 8, 100 mM NaCl, 20 mM imidazole) per pellet of 500 mL culture. Cells were disrupted by sonication (Branson medium tip, 2 x 3 min, Duty 40%, Output: 4). After centrifugation (40.000 g, 1 h, 4°C) supernatant containing C-terminally His_6-tagged target protein was loaded onto a 1 mL HisTrap™ FF column equilibrated with 10 CV buffer S1. The column was washed with 10 CV buffer S1 followed by elution with buffer S2 (20 mM Tris pH 8, 100 mM NaCl, 500 mM Imidazole). Elution of the sensor was tracked by the greenish hue. Afterwards, buffer was exchanged to the desired assay buffer using NAP 25 columns. Purified sensor was stored at 4°C if not used immediately.

Isothermal Titration Calorimetry (ITC) of receptor modules

ITC was performed as described in Kröger et al 2021a & b[22,23]. Used protein and ligand concentrations can be found in table S7

Crystallography of receptor modules

All crystallization experiments were set up as sitting drop vapour diffusion experiments in 3-well Intelli plates (Art Robbins Instruments) using a protein concentration of 40 mg/mL (PotF-S87Y-A182D) and 30 mg/mL PotF-D247K and a 20-fold molar excess of ligand if added. Protein-ligand mixtures were equilibrated at 293 K for several hours before crystallization setups.

Crystals for PotF-S87Y-A182D were obtained in 0.085 M sodium acetate pH 4.6, 0.17M ammonium acetate, 30% PEG 4000 and 15% glycerol and for PotF-D247K in 2.4 M ammonium sulfate, 0.1 M Bicine pH 9.0 and 10% Jeffamine M600.

Crystals were mounted using CryoLoops and then flash-cooled in liquid nitrogen. In case of PotF-D247K crystals were first transferred into a cryogenic solution made of reservoir solution and 1.7 M malonate matching the pH of the condition. Data collection at 100 K was done at the beamline BL 14.1 at the synchrotron BESSY II, Helmholtz-Zentrum Berlin[46]. Diffraction data was processed as described in Kröger et al 2021a & b.

Lysate-based fluorescence assay

Sensor variants were overexpressed in 10 mL autoinduction medium for 18 h at 30°C and 180 rpm. Afterwards, bacteria were pelleted by centrifugation (20 min, 3500 g, 4°C) and resuspended in 2 mL of assay buffer (20 mM Tris pH8, 100 mM NaCl). The suspension was transferred to two 1.2 mL tubes of an eight's-cluster-strip (biolab products) and 180 - 220 µg of 0.1 mm glass beads (biolab products) were added. Cells were disrupted by using a bead mill (Bead ruptor, Omni®) 24 at 6 m/s for ten cycles of 30 s with 1 min breaks between each cycle. Debris and beads were pelleted by centrifugation (20 min, 4800 g, 4°C). 180 µL of supernatant was transferred to a Nunclon F96 MicroWell™ black plate and a first fluorescence intensity measurement (λ_{ex} = 470/10 nm, λ_{em} = 515/10 nm) was conducted to adjust the gain for the tested constructs to a fluorescence of roughly 10000 - 20000 units. The detection method of the reader (Tecan Spark) was set to fluorescence top reading, the gain according to the initial scan, the excitation and emission as stated before, and the Z-Position to 16800 µm. Solutions of different ligands were prepared in assay buffer. For each well the initial fluorescence intensity was measured, followed by 10 x 5 µL injections of 100 mM ligand solution using injector pumps. After each injection, a 2 mm and 150 rpm double orbital shake was conducted for 5 s prior to the measurement. As a control, buffer was added in an additional well in 10 x 5 µL injections. The measured values were corrected by the fold-change in fluorescence intensity of the buffer control and the resulting fold increase in fluorescence intensity was plotted against the ligand concentration to evaluate ligand binding.

Protein-based fluorescence assay

Plate reader setting and data evaluation were kept consistent with the lysate assay. First, a dilution series (1:10 - 1:10.000) of the sensor was measured in the respective assay buffer (20 mM Tris pH8, 100 mM NaCl) in a total reaction volume of 180 µL to adjust fluorescence emission to 10000 –20000 fluorescence counts. Fluorescence intensity was measured as triplicates. For each well the initial fluorescence intensity was measured, followed by 5 x 5 µL injections of 20 mM ligand solution and

9 x 5 µL injections of 100 mM ligand solution using injector pumps. After each injection, a 2 mm and 150 rpm double orbital shake was conducted for 5 s prior to the measurement. As a control, buffer was added in an additional well in 14 x 5 µL injections.

Dose-response assay in E. coli

Plate reader settings and dilutions were kept consistent with the other fluorescence assay. To measure binding affinities of the final sensor variant, 180 µl of adequate sensor dilution was combined with 20 µl of ligand solutions ranging from 0.01 mM to 500 mM. Reaction mix was incubated for 5 - 10 min at RT to ensure endpoint measurements. A buffer was chosen that represents HEK cell lysate assay buffer (10 mM NaCl, 5 mM KCl, 2 mM $MgCl_2$*$6H_2O$, 2 mM $CaCl_2$*$2H_2O$, 10 mM glucose &10 mM HEPES pH 7.4). Data was normalized to the buffer measurement and fluorescence increase was plotted against ligand concentration. Data was fitted sigmoidal for a 1:1 binding ratio of sensor and ligand using the Hill-equation.

Preparation of primary hippocampal cultures

Primary hippocampal neurons were cultured from C57BL/6 mice as described before[47]. Hippocampi were dissected from mice at embryonic day 15-19 (animal license protocol AZ 81-02.04.2020.A100, LANUV, NRW, Germany), washed with HBSS (Life Technologies), and incubated with 0.025 g/ml trypsin (Life Technologies) for 20 min at 37°C. After washing with HBSS, DNA was digested using 0.001 g/ml DNase I (Roche) and the tissue was further dissociated using filter tips. The dissociated cells were seeded on coverslips coated with Poly-D-lysine (Sigma-Aldrich) in a 24-well plate at a density of 40.000 cells per well. The cells were incubated in Neurobasal Medium supplemented with 2% B-27 and 1 mM L-glutamine or in Basal Medium Eagle supplemented with 2% B-27, 0.5% glucose, 0.5 mM L-glutamine, and 1% fetal bovine serum (Thermo Fisher Scientific) at 37°C with 5% CO_2 until further use.

Recombinant adeno-associated virus (rAAV) production

rAAV of serotype 2/1 were produced as described before[48]. Briefly, HEK293T cells were transfected using the calcium-phosphate method with pAAV-Syn-Display-AGMsen and pAAV-Syn-Display-control sensor together with the adenoviral helper plasmid pFΔ6, pRV1, and pH21, the latter encoding rep and cap genes. Cells were harvested 48 hours after transfection and lysed in 0.5% sodium deoxycholate (Sigma-Aldrich) and 50 units/ml Benzonase endonuclease (Millipore). Viral particles were purified by HiTrap™ heparin columns (GE Healthcare) and concentrated with Amicon Ultra Centrifugal Filters (Millipore) to a final volume of 400 µl. An aliquot of the virus was validated on an SDS-polyacrylamide gel by Coomassie Blue staining.

Transfection of HEK cells and transduction of hippocampal neurons

HEK293 (CRL-1573) cells were obtained from American Type Culture Collection (ATCC) and grown in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10 % FCS to 90 - 95% confluency. Cells were transfected with 9.5 µg sensor plasmid per 9 cm petri dish and 0.5 µg per 4 well plate. Therefore, DNA and polyethyleneimine (PEI, Sigma Aldrich) were mixed in OptiMEM (Gibco) to a final concentration of 10 ng/µl and 20 ng/µl, respectively. Reaction mix was incubated for 10 min at RT before adding it to the culture. Prior to that, old media was removed from the culture and replaced with 4x the volume to the reaction mix of fresh medium containing only 2% serum. Cells were left for 2 days at 37°C and 5% CO_2 before further experiments to ensure good sensor expression.

For transduction 0.5 - 1 µl of AAV was added directly into BME medium 4 days after isolation and incubated for 8 - 10 days at 37°C and 5% CO_2.

Dose-response assay in HEK lysate

For harvest media was removed, and cells were washed with 1x PBS. Afterwards 1 ml of 1xPBS was added to the dish, cells were scraped and transferred to 1.5 ml tube. After centrifugation (500g, 5 min, RT) PBS was removed and replaced with assay buffer (10 mM NaCl, 5 mM KCl, 2 mM $MgCl_2*6H_2O$, 2 mM $CaCl_2*2H_2O$, 10 mM glucose & 10 mM HEPES pH 7.4) for hypotonic lysis. Cells were incubated for 15 min on ice and sonicated in a water bath for 1x 30 s and 2x 10 s to assist lysis. Afterwards cells were incubated on ice for 5 min, followed by vortexing thoroughly and another 5 min incubation on ice. Finally, debris was pelleted by centrifugation (1000 g, 10 min, 4°C) and supernatant was diluted 1:10 for the lysate assay. 180 µl of lysate was transferred to black walled fluorescence plates (Grainer) and 20 µl ligand solutions with concentrations ranging from 0.001 mM to 100 mM were added. Mixture was incubated for 5 min to ensure endpoint measurements. Plate was transferred to a FluoStar Omega plate reader (BMG Labtech) and whole well fluorescence was determined as follows: λ_{ex} = 485, λ_{em} = 520, Flashes 20, top optics, Gain set to 90% fluorescence in the well with highest AGM concentration. Data was normalized as described before.

Dose-response assay for HEK displayed sensor

Fist, media was removed, and cells were washed with 1xPBS followed by incubation with 1 ml 1xPBS/EDTA for 15 - 30 min (37°C, 5%CO_2) or until cells start to detach. Cells were flushed from the plate and spun down (500g, 5 min, RT). After discarding the PBS/EDTA solution cells were resuspended in 1x HBSS (Gibco). Cells were diluted 1:20 and 1.8 ml of the solution was transferred to 2 ml reaction tube to which either 200 µl buffer or ligand solution ranging from 0.01 mM to 500 mM were added. Mixture was incubated at RT for 5 min to ensure endpoint measurements. Afterwards it was transferred to a fluorescence cuvette. Each cell solution was measured in a PTI QuantaMaster

spectrofluorometer (Horiba) at 20°C and λ_{ex} = 488/7.5 nm, λ_{em} = 512/5 nm for 5 seconds under constant stirring to ensure homogeneity during the measurement. For each concentration the measurements over 5 s were averaged and data was normalized as described before.

Immunocytochemistry with and without permeabilization

For staining, HEK cells and hippocampal neurons were grown on coverslips in four well or 24 well plates, respectively. Without permeabilization primary antibody (ab6556 Anti-GFP, Abcam, 1:500) was added to the medium (DMEM, 10% FCS) and cells were incubated for 30 min at 37°C and 5% CO_2 afterwards cells were washed with PBS followed by fixation with 4% paraformaldehyde (PFA, Alfa Aesar) and another wash. Staining was blocked for 30 min with CT buffer (0.5% Triton X-100, 5% ChemiBLOCKER in 0.1M NaP pH 7, Sigma Aldrich & Merck) at RT. Secondary antibody (Goat anti-Rabbit A647 IgG, A21245 Invitrogen 1:500 or 111-604-144 Jackson Immuno 1:400) was added in CT buffer simultaneously with DAPI (4', 6-Diamidino-2-Phenyl-indole dihydrochloride, 1:10000) as a DNA counterstain and cells were incubated for 45 min in the dark. Cells were washed with PBS, and coverslips were placed face down on microscope slides with a drop Aqua-Poly/mount (Tebu-Bio) mounting medium being left to harden for 24 h at 4°C. For staining with permeabilization the primary antibody is added after fixation and blocking with an additional PBS washing step between the antibodies. Confocal images were recorded with a Leica Sp5 microscope using the 63x magnification HCX PL APO lambda blue objective. DAPI and AF647 were imaged at 30% laser power with a 488 nm argon laser and a 633 nm HeNe laser, respectively. All imaging was performed at the Microscopy facility of the University Hospital Bonn.

Live cell imaging of neurons expressing the sensor in response to Agmatine

After transduction and expression, the activity of the sensor was analysed using a Zeiss Observer.Z1 widefield microscope equipped with an Axiocam 506 trans imaging device. GFP fluorescence (λ_{ex}: 450-490 nm, λ_{em}: 500-550 nm) was imaged using the 10x magnification LD EC PlanNeoFluor objective. For the measurements, medium was replaced with 150 µl ES buffer, followed by recording a GFP baseline for 2 min in 10 s. intervals. Different concentrations of agmatine were added to final concentrations ranging from (5 µM – 5 mM) in a 1:2 dilution and measurements were continued further for 6 min.

Data Availability:

X-ray coordinates of all solved structures have been deposited at the protein databank with accession codes: 8ASZ, 8AT0

Acknowledgment:

We thank Susanne Gillig and Simon Prisner for exploring early sensor constructs and Cornelius Fischer for help in testing agmatine sensor constructs. We acknowledge financial support and allocation of synchrotron beamtime by HZB and thank the beamline staff at BESSY for support. This work was supported by the Deutsche Forschungsgemeinschaft grant HO4022/2-3.

Declaration of interests:

The authors declare that they have no conflict of interest.

References:

1. Blier, P. & El Mansari, M. Serotonin and beyond: therapeutics for major depression. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **368,** 20120536; 10.1098/rstb.2012.0536 (2013).
2. Cowen, P. J. & Browning, M. What has serotonin to do with depression? *World psychiatry : official journal of the World Psychiatric Association (WPA)* **14,** 158–160; 10.1002/wps.20229 (2015).
3. Dunlop, B. W. & Nemeroff, C. B. The role of dopamine in the pathophysiology of depression. *Archives of general psychiatry* **64,** 327–337; 10.1001/archpsyc.64.3.327 (2007).
4. Marvin, J. S. *et al.* An optimized fluorescent probe for visualizing glutamate neurotransmission. *Nature methods* **10,** 162–170; 10.1038/nmeth.2333 (2013).
5. Marvin, J. S. *et al.* A genetically encoded fluorescent sensor for in vivo imaging of GABA. *Nature methods* **16,** 763–770; 10.1038/s41592-019-0471-2 (2019).
6. Patriarchi, T. *et al.* Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science (New York, N.Y.)* **360**; 10.1126/science.aat4422 (2018).
7. Sun, F. *et al.* A Genetically Encoded Fluorescent Sensor Enables Rapid and Specific Detection of Dopamine in Flies, Fish, and Mice. *Cell* **174,** 481-496.e19; 10.1016/j.cell.2018.06.042 (2018).
8. Unger, E. K. *et al.* Directed Evolution of a Selective and Sensitive Serotonin Sensor via Machine Learning. *Cell* **183,** 1986-2002.e26; 10.1016/j.cell.2020.11.040 (2020).
9. Goracke-Postle, C. J., Nguyen, H. O. X., Stone, L. S. & Fairbanks, C. A. Release of tritiated agmatine from spinal synaptosomes. *Neuroreport* **17,** 13–17; 10.1097/01.wnr.0000192739.38653.aa (2006).
10. Goracke-Postle, C. J., Overland, A. C., Riedl, M. S., Stone, L. S. & Fairbanks, C. A. Potassium- and capsaicin-induced release of agmatine from spinal nerve terminals. *Journal of neurochemistry* **102,** 1738–1748; 10.1111/j.1471-4159.2007.04647.x (2007).
11. Piletz, J. E. *et al.* Agmatine: clinical applications after 100 years in translation. *Drug discovery today* **18,** 880–893; 10.1016/j.drudis.2013.05.017 (2013).
12. Roth, B. L., Sheffler, D. J. & Kroeze, W. K. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nature reviews. Drug discovery* **3,** 353–359; 10.1038/nrd1346 (2004).
13. Zomkowski, A. D. E. *et al.* Agmatine produces antidepressant-like effects in two models of depression in mice. *Neuroreport* **13,** 387–391; 10.1097/00001756-200203250-00005 (2002).
14. Chen, Z.-D. *et al.* Antidepressant-like action of agmatine in the acute and sub-acute mouse models of depression: a receptor mechanism study. *Metabolic brain disease* **33,** 1721–1731; 10.1007/s11011-018-0280-9 (2018).
15. Kotagale, N. R. *et al.* Psychopharmacological study of agmatine in behavioral tests of schizophrenia in rodents. *Pharmacology, biochemistry, and behavior* **100,** 398–403; 10.1016/j.pbb.2011.09.006 (2012).

16. Song, J. *et al.* Agmatine Improves Cognitive Dysfunction and Prevents Cell Death in a Streptozotocin-Induced Alzheimer Rat Model. *Yonsei Medical Journal* **55,** 689–699; 10.3349/ymj.2014.55.3.689 (2014).

17. Ninkovic, M. *et al.* The Use of Agmatine Provides the New Insight in an Experimental Model of Multiple Sclerosis. *Neurochemical research* **40,** 1719–1727; 10.1007/s11064-015-1655-5 (2015).

18. Baird, G. S., Zacharias, D. A. & Tsien, R. Y. Circular permutation and receptor insertion within green fluorescent proteins. *Proceedings of the National Academy of Sciences of the United States of America* **96,** 11241–11246; 10.1073/pnas.96.20.11241 (1999).

19. Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. *Nature biotechnology* **24,** 79–88; 10.1038/nbt1172 (2006).

20. Lorimier, R. M. de *et al.* Construction of a fluorescent biosensor family. *Protein science : a publication of the Protein Society* **11,** 2655–2675; 10.1110/ps.021860 (2002).

21. Nasu, Y., Shen, Y., Kramer, L. & Campbell, R. E. Structure- and mechanism-guided design of single fluorescent protein-based biosensors. *Nature chemical biology* **17,** 509–518; 10.1038/s41589-020-00718-x (2021).

22. Kröger, P., Shanmugaratnam, S., Ferruz, N., Schweimer, K. & Höcker, B. A comprehensive binding study illustrates ligand recognition in the periplasmic binding protein PotF. *Structure (London, England : 1993)* **29,** 433-443.e4; 10.1016/j.str.2020.12.005 (2021).

23. Kröger, P., Shanmugaratnam, S., Scheib, U. & Höcker, B. Fine-tuning spermidine binding modes in the putrescine binding protein PotF. *The Journal of biological chemistry* **297,** 101419; 10.1016/j.jbc.2021.101419 (2021).

24. Vassylyev, D. G., Tomitori, H., Kashiwagi, K., Morikawa, K. & Igarashi, K. Crystal structure and mutational analysis of the Escherichia coli putrescine receptor. Structural basis for substrate specificity. *The Journal of biological chemistry* **273,** 17604–17609; 10.1074/jbc.273.28.17604 (1998).

25. Igarashi, K. & Kashiwagi, K. Modulation of cellular function by polyamines. *The international journal of biochemistry & cell biology* **42,** 39–51; 10.1016/j.biocel.2009.07.009 (2010).

26. Landau, G., Bercovich, Z., Park, M. H. & Kahana, C. The role of polyamines in supporting growth of mammalian cells is mediated through their requirement for translation initiation and elongation. *The Journal of biological chemistry* **285,** 12474–12481; 10.1074/jbc.M110.106419 (2010).

27. Mandal, S., Mandal, A., Johansson, H. E., Orjalo, A. V. & Park, M. H. Depletion of cellular polyamines, spermidine and spermine, causes a total arrest in translation and growth in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **110,** 2169–2174; 10.1073/pnas.1219002110 (2013).

28. Pegg, A. E. & McCann, P. P. Polyamine metabolism and function. *The American journal of physiology* **243,** C212-21; 10.1152/ajpcell.1982.243.5.C212 (1982).

29. Scheib, U., Shanmugaratnam, S., Farías-Rico, J. A. & Höcker, B. Change in protein-ligand specificity through binding pocket grafting. *Journal of structural biology* **185,** 186–192; 10.1016/j.jsb.2013.06.002 (2014).

30. Möglich, A. An Open-Source, Cross-Platform Resource for Nonlinear Least-Squares Curve Fitting. *J. Chem. Educ.* **95,** 2273–2278; 10.1021/acs.jchemed.8b00649 (2018).

31. van den Noort, M., Boer, M. de & Poolman, B. Stability of Ligand-induced Protein Conformation Influences Affinity in Maltose-binding Protein. *Journal of molecular biology* **433,** 167036; 10.1016/j.jmb.2021.167036 (2021).

32. Arndt, M. A. *et al.* The arginine metabolite agmatine protects mitochondrial function and confers resistance to cellular apoptosis. *American journal of physiology. Cell physiology* **296,** C1411-9; 10.1152/ajpcell.00529.2008 (2009).

33. Scalabrino, G. & Ferioli, M. E. Polyamines in Mammalian Tumors Part I (Elsevier1981), Vol. 35, pp. 151–268.

34. Scalabrino, G. & Ferioli, M. E. Polyamines in Mammalian Tumors Part II (Elsevier1982), Vol. 36, pp. 1–102.

35. Isome, M. *et al.* The antiproliferative effects of agmatine correlate with the rate of cellular proliferation. *American journal of physiology. Cell physiology* **293,** C705-11; 10.1152/ajpcell.00084.2007 (2007).

36. Cabella, C. *et al.* Transport and metabolism of agmatine in rat hepatocyte cultures. *European journal of biochemistry* **268,** 940–947; 10.1046/j.1432-1327.2001.01955.x (2001).

37. del Valle, A. E., Paz, J. C., Sánchez-Jiménez, F. & Medina, M. A. Agmatine uptake by cultured hamster kidney cells. *Biochemical and biophysical research communications* **280,** 307–311; 10.1006/bbrc.2000.4101 (2001).

38. Molderings, G. J., Bönisch, H., Göthert, M. & Brüss, M. Agmatine and putrescine uptake in the human glioma cell line SK-MG-1. *Naunyn-Schmiedeberg's archives of pharmacology* **363,** 671–679; 10.1007/s002100100418 (2001).

39. Sastre, M., Regunathan, S. & Reis, D. J. Uptake of agmatine into rat brain synaptosomes: possible role of cation channels. *Journal of neurochemistry* **69,** 2421–2426; 10.1046/j.1471-4159.1997.69062421.x (1997).

40. Satriano, J., Isome, M., Casero, R. A., Thomson, S. C. & Blantz, R. C. Polyamine transport system mediates agmatine transport in mammalian cells. *American journal of physiology. Cell physiology* **281,** C329-34; 10.1152/ajpcell.2001.281.1.C329 (2001).

41. Satriano, J. *et al.* Agmatine suppresses proliferation by frameshift induction of antizyme and attenuation of cellular polyamine levels. *The Journal of biological chemistry* **273,** 15313–15316; 10.1074/jbc.273.25.15313 (1998).

42. Seo, S., Liu, P. & Leitch, B. Spatial learning-induced accumulation of agmatine and glutamate at hippocampal CA1 synaptic terminals. *Neuroscience* **192,** 28–36; 10.1016/j.neuroscience.2011.07.007 (2011).

43. Clements, J. D., Lester, R. A., Tong, G., Jahr, C. E. & Westbrook, G. L. The time course of glutamate in the synaptic cleft. *Science (New York, N.Y.)* **258,** 1498–1501; 10.1126/science.1359647 (1992).

44. Dzubay, J. A. & Jahr, C. E. The Concentration of Synaptically Released Glutamate Outside of the Climbing Fiber–Purkinje Cell Synaptic Cleft. *J. Neurosci.* **19,** 5265–5274; 10.1523/JNEUROSCI.19-13-05265.1999 (1999).

45. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein expression and purification* **41,** 207–234; 10.1016/j.pep.2005.01.016 (2005).

46. Mueller, U. *et al.* The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives. *Eur. Phys. J. Plus* **130,** 662; 10.1140/epjp/i2015-15141-2 (2015).

47. Woitecki, A. M. H. *et al.* Identification of Synaptotagmin 10 as Effector of NPAS4-Mediated Protection from Excitotoxic Neurodegeneration. *J. Neurosci.* **36,** 2561–2570; 10.1523/JNEUROSCI.2027-15.2016 (2016).

48. van Loo, K. M. J. *et al.* Transcriptional regulation of T-type calcium channel CaV3.2: bi-directionality by early growth response 1 (Egr1) and repressor element 1 (RE-1) protein-silencing transcription factor (REST). *The Journal of biological chemistry* **287,** 15489–15501; 10.1074/jbc.M111.310763 (2012).

Supplementary Information:

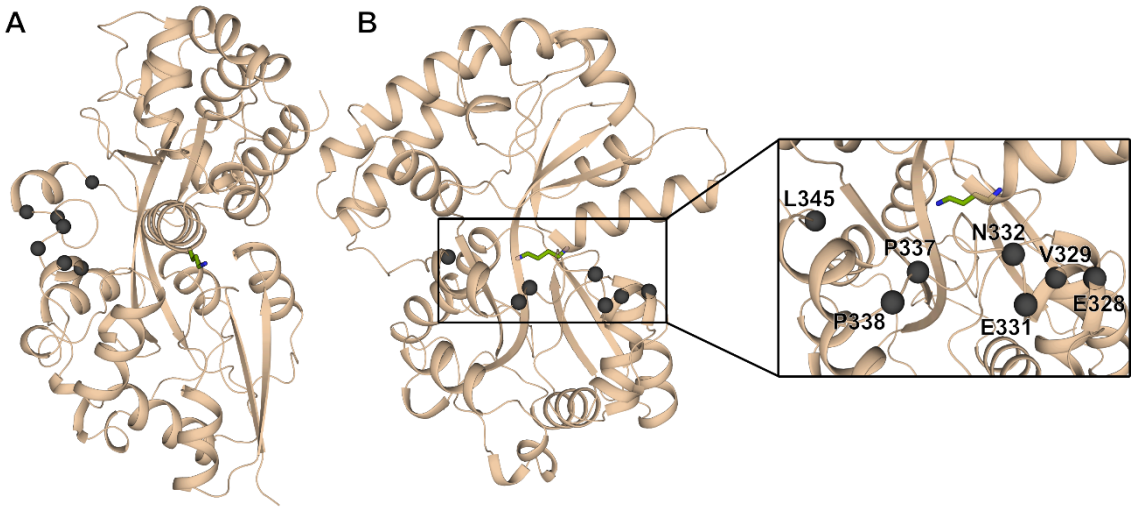


Figure S1: Side (A) and back view (B) of PotF to highlight the insertion site of cpGFP. The insertion site is located on the back side of the binding pocket close to the hinge region. Thereby, an insertion of cpGFP in this region should result in a motion-sensitive sensor. The Cα of the target residues for splicing overlap extension PCR are shown as dark grey spheres. Putrescine is shown as green sticks to indicate the position of the binding pocket in relation to the cpGFP insertion region.

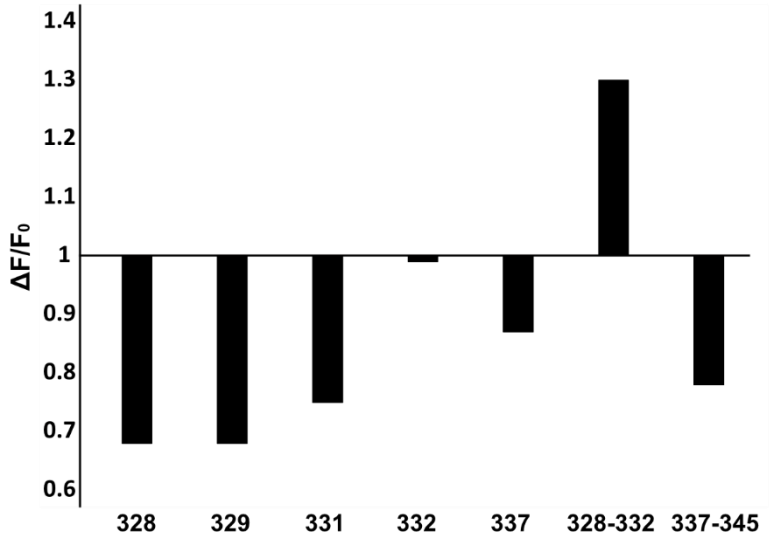


Figure S2: Fluorescence signal upon putrescine addition at different cpGFP insertion sites. All single insertion sites result in a negative fluorescence signal, with position 332 not responding in general. Only the combination of positions 328 and 332 resulted in the desired gain in fluorescence upon adding 50 – 100 µM of putrescine.

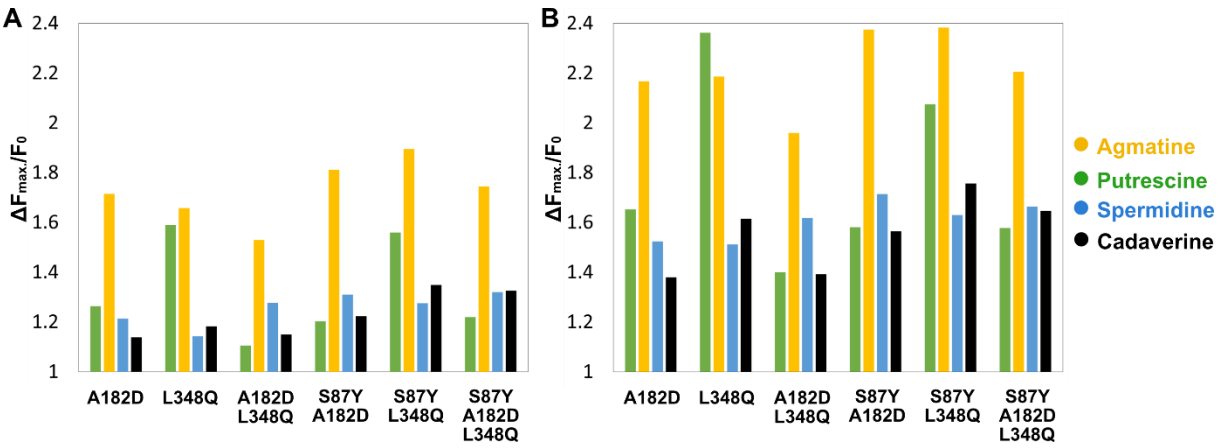


Figure S3: Example of lysate (A) and purified protein (B) screening assays for the same AGM sensor constructs. Overall, the fluorescence signal is similar albeit lower in the lysate screening. This representation only shows the maximum fluorescence gain after the last ligand addition. Protein assays are conducted to confirm lysate screening results and to lower the influence of possible endogenous biogenic amines from *E. coli* and other lysate components. S87Y-A182 is the variant that ultimately became AGMsen.

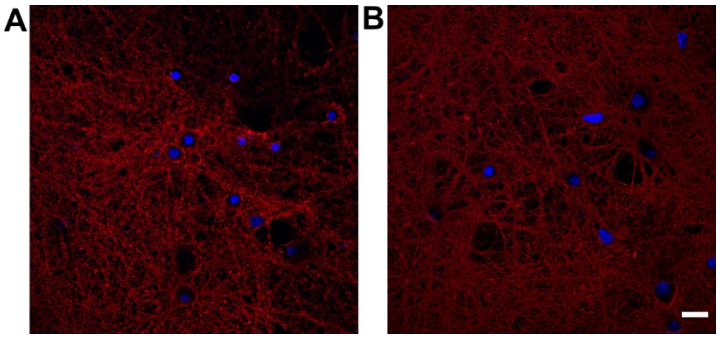


Figure S3: Antibody staining of primary rat hippocampal neurons transduced with pAAV-Syn-Display-AGMsen (A) and the control sensor (B). Both sensors express in neurons and are located in the outer membrane as confirmed by α-GFP antibody staining without permeabilization. Nucleus is counterstained with DAPI. Scale bar corresponds to 20 µm.

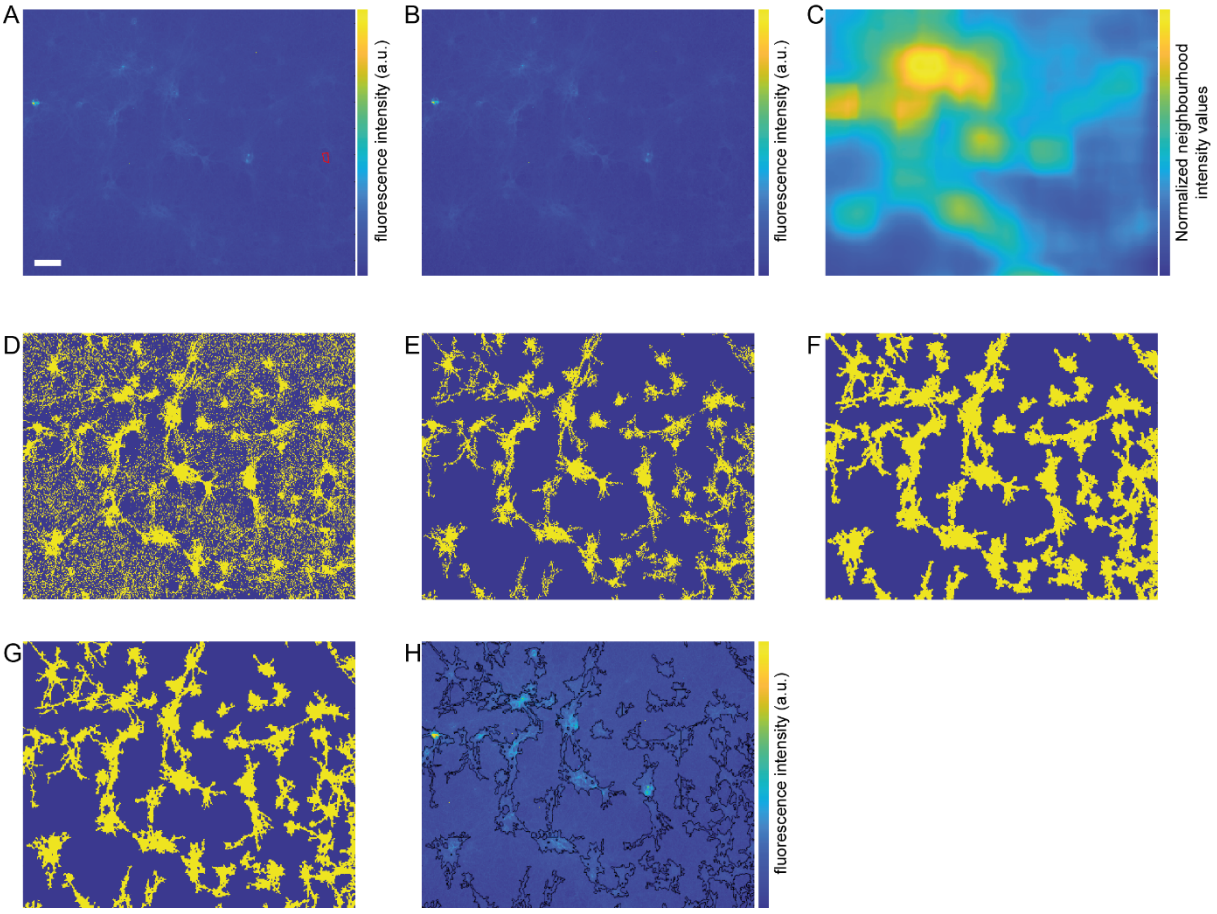


Figure S4: Analysis routine for neuronal imaging. All images shown are exemplary outputs of the described steps. Based on a maximum intensity projection of the tiff stack (**A**) a region with apparent homogenous minimal signal was chosen manually (red ROI). For each frame the mean pixel values of the background ROI were calculated and subtracted yielding a background corrected tiff-stack. On a maximum intensity projection of this stack (**B**) an adaptive threshold algorithm was used (Matlab R2022a, "adaptthresh", threshold setting 0.5, if not indicated differently all functions were used with settings as per default of R2022a) yielding a normalized map of neighborhood intensities (**C**). This map was used to binarize the maximum intensity projection of the background corrected stack yielding initial ROIs for the neuronal cells ("imbinarize", **D**). Subsequently, small regions below 5000 pixels were removed ("bwareaopen", **E**), the remaining ROIs were dilated ("dilate", **F**) and subsequently eroded ("erode", **G**) with a disk as morphological structuring element with a radius of 4 pixels (approximated with 6 lines, "strel") yielding the final ROIs (**G**). In **H** the final ROIs are shown overlaying the maximum intensity projection of the background corrected tiff-stack. All images with scalebars as in **A** of 100 µm. The ROIs were used individually on each tiff-stack to extract the mean pixel values. Concatenation of mean pixel values yielded the time-trace for a given ROI. For each experimental condition 10-20 ROIs were analyzed. For each trace the change in fluorescence was calculated by dividing the mean of three frames after addition (14-16) by the three frames before addition (11-13).

Figure S5: Time traces of individual ROIs. Type of experiment indicated on top. X axis is frame number. Y axis is fluorescence intensity (a.u.).

Table S1: Differences between PotF and PotF/D

Position	PotF Residue	PotF/D Residue
38	Ser	Thr
39	Asp	Glu
87	Ser	Tyr
182	Ala	Asp
247	Asp	Ser
276	Phe	Trp
348	Leu	Glu

Table S2: ITC data for all measurements. N/D = Not detectable

Variant	Ligand	K_D [µM]	ΔG [kcal/mol]	ΔH [kcal/mol]	-TΔS [kcal/mol]	n
PotF/D	AGM	4.0	-7.2	-10.3	3.1	0.90
PotF-S87Y-A182D	AGM	0.3	-8.7	-5.8	-2.9	1.0
	PUT	N/D	-	-	-	-
	SPD	N/D	-	-	-	-
	CDV	N/D	-	-	-	-
PotF-D247K	AGM	N/D	-	-	-	-
	PUT	N/D	-	-	-	-
	SPD	N/D	-	-	-	-
	CDV	N/D	-	-	-	-

Table S3: Affinity from dose-response measurements. Data was fitted by Hill-equation using the fit-o-mat[30]. N/D = Not detectable.

	Agmatine	Putrescine	Spermidine	Cadaverine
AGMsen *E. coli* purified	41 ± 5 µM	3.3 ± 0.8 mM	297 ± 38 µM	7.2 ± 7.3 mM
Control sensor *E. coli* purified	8.0 ± 5.6 mM	N/D	N/D	N/D
AGMsen HEK lysate	35 ± 2 µM	1.9 ± 0.1 mM	220 ± 10 µM	6.2 ± 2.8 mM
Control sensor HEK lysate	2.4 ± 0.8 mM	N/D	N/D	N/D
AGMsen HEK Display	525 ± 60 µM	-	-	-

Table S4: Crystallographic data and refinement statistics for all solved crystal structures

	PotF-S87Y-A182D	PotF-D247K		PotF-S87Y-A182D	PotF-D247K
PDB ID	8ASZ	8AT0	**Refinement**		
Data collection			Reflections used in refinement	86267 (8426)	54717 (5385)
Wavelength [Å]	0.9184		Reflections used for R_{free}	2099 (205)	2098 (206)
Resolution range [Å]	37.33 - 1.28 (1.33 - 1.28)	45.55 - 2.00 (2.07 - 2.00)	R_{work}	0.179 (0.391)	0.201 (0.306)
Space group	P 21 21 21	P 32 2 1	R_{free}	0.209 (0.380)	0.243 (0.359)
Cell parameter			CC_{work}	0.970 (0.604)	0.968 (0.587)
a, b, c, [Å]	37.3 79.1 113.3	70.7 70.7 272.6	CC_{free}	0.970 (0.676)	0.953 (0.460)
α,β,γ [°]	90 90 90	90 90 120	Number of non-hydrogen atoms	3305	5840
Total reflections	628916 (62355)	602533 (61804)	macromolecules	2851	5444
Unique reflections	86290 (8427)	54733 (5385)	solvent	378	308
Multiplicity	7.3 (7.4)	11.0 (11.5)	Protein residues	343	682
Completeness [%]	99.0 (98.0)	99.9 (100.0)	RMS bond lengths [Å]	0.016	0.003
Mean I/sigma [I]	14.6 (0.6)	11.2 (0.7)	RMS bond angles [°]	1.04	0.57
Wilson B-factor	17.9	39.34	Ramachandran favored [%]	97.1	97.4
No. of molecules per a.u.	1	2	Ramachandran allowed [%]	2.9	2.7
Matthews coefficient			Ramachandran outliers [%]	0.0	0.0
R_{merge}	0.065 (2.904)	0.181 (2.980)	Rotamer outliers [%]	1.28	0.7
R_{meas}	0.070 (3.118)	0.190 (3.120)	Clashscore	4.08	4.63
R_{pim}	0.026 (1.122)	0.057 (0.918)	Average B-factor	30.53	45.95
$CC_{1/2}$	1.000 (0.280)	0.999 (0.256)	macromolecules	29.23	45.53
CC*	1.000 (0.661)	1.000 (0.639)	solvent	37.71	47.96
			Number of TLS groups	1	2

Table S5: Oligonucleotides used for initial construction of the sensor

Type	Position / Orientation	Sequence (5‘-3‘)
Initial construct design	cpGFP-PotF327-fwd	GCTGGAGTACAACTTTAACGAAGTCCGTGAGAACCC
	cpGFP-PotF327-rev	GGGTTCTCACGGACTTCGTTAAAGTTGTACTCCAGC
	PotF-328-cpGFP-fwd	ACGCCGCTGGTGAGTGCGGAAAGCCACAACGTCTATATCATG
	PotF-328-cpGFP-rev	CATGATATAGACGTTGTGGCTTTCCGCACTCACCAGCGGCGT
	cpGFP-PotF-328-fwd	AAGCTGGAGTACAACTTTAACGTCCGTGAGAACCCAGGTATT
	cpGFP-328-PotF-rev	AATACCTGGGTTCTCACGGACGTTAAAGTTGTACTCCAGCTT
	PotF329-cpGFP-fwd	CTGGTGAGTGCGGAAGTCAGCCACAACGTCTATATCATG
	PotF329-cpGFP-rev	CATGATATAGACGTTGTGGCTGACTTCCGCACTCACCAG
	cpGFP-PotF329-fwd	TGGAGTACAACTTTAACCGTGAGAACCCAGGTATTTAT
	cpGFP-PotF329-rev	ATAAATACCTGGGTTCTCACGGTTAAAGTTGTACTCCA
	PotF331-cpGFP-fwd	GAAGTCCGTGAGAGCCACAACGTCTATATC
	PotF331-cpGFP-rev	GATATAGACGTTGTGGCTCTCACGGACTTC
	cpGFP-PotF331-fwd	TGGAGTACAACTTTAACAACCCAGGTATTTATCCGCC
	cpGFP-PotF331-rev	GGCGGATAAATACCTGGGTTGTTAAAGTTGTACTCC
	PotF-332-cpGFP-fwd	AGTGCGGAAGTCCGTGACAACAGCCACAACGTCTATATCATG
	PotF-332-cpGFP-rev	CATGATATAGACGTTGTGGCTGTTGTCACGGACTTCCGCACT
	cpGFP-PotF-332-fwd	AAGCTGGAGTACAACTTTAACCCAGGTATTTATCCGCCTGCG
	cpGFP-332-PotF-rev	CGCAGGCGGATAAATACCTGGGTTAAAGTTGTACTCCAGCTT
	PotF-337-cpGFP-fwd	GAGAACCCAGGTATTTATCCGAGCCACAACGTCTATATCATG
	PotF-337-cpGFP-rev	CATGATATAGACGTTGTGGCTCGGATAAATACCTGGGTTCTC
	cpGFP-337-potF-fwd	AAGCTGGAGTACAACTTTAACCCTGCGGATGTTCGTGCGAAG
	cpGFP-337-potF-rev	CTTCGCACGAACATCCGCAGGGTTAAAGTTGTACTCCAGCTT
	potF345cpGFP_fwd	CGTGCGAAGCTGAGCCACAACGTCTATATCATGGC
	cpGFPpotF345_rev	CGTTGTGGCTCAGCTTCGCACGAACATCCGCAGGCGG

Table S6: Oligonucleotides used for cloning and the generation of sensor variants via QuickChange

Type	Position / Orientation	Sequence (5‘-3‘)
Mutation: S87Y	forward	GGTTCCATCTGCCTACTTTCTGGAGCGCC
	reverse	GGCGCTCCAGAAAGTAGGCAGATGGAACC
Mutation: A182D	forward	CTCTTTCCTGGATGATCCAGAAGAAGTTT
	reverse	AAACTTCTTCTGGATCATCCAGGAAAGAG
Mutation: D247K	forward	CGGCTGGGCAGGTAAGGTCTGGCAGGCGT
	reverse	ACGCCTGCCAGACCTTACCTGCCCAGCCG
Gibson pcDNA	Insert_fw	CTGGGCCACCATGGCTGAACAAAAAACAC
	Insert_rv	GCCCTCTAGATTATTTTCCGCTCTTCAC
	Vector_fw	CGGAAAATAATCTAGAGGGCCCGTTTAAACCC
	Vector_rv	GTTCAGCCATGGTGGCCCAGCTTGGGTC
pDisplay	Insert_fw	GGGGCCCAGCCGGCCAGATCTGCTGAACAAAAAACACTCCAC
	Insert_rv	GACCTGCAGTTTTCCGCTCTTCACTTTGGTCC

Table S7: Concentrations of protein and ligand solutions used for ITC measurements

Protein Variant	Protein concentration [µM]	Ligand	Ligand concentration [mM]
PotF-S87Y-A182D	210	AGM	2.10
	210	PUT	2.10
	210	SPD	2.10
	210	CDV	2.10
PotF-D247K	190	AGM	1.90
	190	PUT	1.90
	190	SPD	1.90
	190	CDV	1.90
PotF/D	442	AGM	4.95

11. List of Publications

1. Kröger, P., Shanmugaratnam, S., Ferruz, N., Schweimer, K. & Höcker, B. **A comprehensive binding study illustrates ligand recognition in the periplasmic binding protein PotF**. *Structure (London, England: 1993)* **29,** 433-443.e4; 10.1016/j.str.2020.12.005 (2021).

2. Kröger, P., Shanmugaratnam, S., Scheib, U. & Höcker, B. **Fine-tuning spermidine binding modes in the putrescine binding protein PotF.** *The Journal of biological chemistry* **297,** 101419; 10.1016/j.jbc.2021.101419 (2021).

3. Draft for:
   Kröger, P., Stiel, A.C., Stüven, B., Shanmugaratnam, S., Schoch, S., Wachten, D. & Höcker, B. **A fluorescent biosensor for the visualization of Agmatine.** *To be submitted*

12.References

Abraham, K.A. (1968). Studies on DNA-dependent RNA polymerase from Escherichia coli. 1. The mechanism of polyamine induced stimulation of enzyme activity. European journal of biochemistry *5*, 143-146. https://doi.org/10.1111/j.1432-1033.1968.tb00348.x.

Baird, G.S., Zacharias, D.A., and Tsien, R.Y. (1999). Circular permutation and receptor insertion within green fluorescent proteins. Proceedings of the National Academy of Sciences of the United States of America *96*, 11241-11246. https://doi.org/10.1073/pnas.96.20.11241.

Binder, S., Schendzielorz, G., Stäbler, N., Krumbach, K., Hoffmann, K., Bott, M., and Eggeling, L. (2012). A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. Genome biology *13*, R40. https://doi.org/10.1186/gb-2012-13-5-r40.

Boffi, J.C., Knabbe, J., Kaiser, M., and Kuner, T. (2018). KCC2-dependent Steady-state Intracellular Chloride Concentration and pH in Cortical Layer 2/3 Neurons of Anesthetized and Awake Mice. Frontiers in cellular neuroscience *12*, 7. https://doi.org/10.3389/fncel.2018.00007.

Borrok, M.J., Zhu, Y., Forest, K.T., and Kiessling, L.L. (2009). Structure-based design of a periplasmic binding protein antagonist that prevents domain closure. ACS chemical biology *4*, 447-456. https://doi.org/10.1021/cb900021q.

Bowman, W.H., Tabor, C.W., and Tabor, H. (1973). Spermidine biosynthesis. Purification and properties of propylamine transferase from Escherichia coli. Journal of Biological Chemistry *248*, 2480-2486.

Brandt, A.-M., Raksajit, W., Yodsang, P., Mulo, P., Incharoensakdi, A., Salminen, T.A., and Mäenpää, P. (2010). Characterization of the substrate-binding PotD subunit in Synechocystis sp. strain PCC 6803. Archives of microbiology *192*, 791-801. https://doi.org/10.1007/s00203-010-0607-3.

Bucher, D., Grant, B.J., Markwick, P.R., and McCammon, J.A. (2011a). Accessing a hidden conformation of the maltose binding protein using accelerated molecular dynamics. PLoS computational biology *7*, e1002034. https://doi.org/10.1371/journal.pcbi.1002034.

Bucher, D., Grant, B.J., and McCammon, J.A. (2011b). Induced fit or conformational selection? The role of the semi-closed state in the maltose binding protein. Biochemistry *50*, 10530-10539. https://doi.org/10.1021/bi201481a.

Carlson, H.J., and Campbell, R.E. (2009). Genetically encoded FRET-based biosensors for multiparameter fluorescence imaging. Current opinion in biotechnology *20*, 19-27. https://doi.org/10.1016/j.copbio.2009.01.003.

Casero, R.A., Murray Stewart, T., and Pegg, A.E. (2018). Polyamine metabolism and cancer: treatments, challenges and opportunities. Nature reviews. Cancer *18*, 681-695. https://doi.org/10.1038/s41568-018-0050-3.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. (1994). Green fluorescent protein as a marker for gene expression. Science (New York, N.Y.) *263*, 802-805. https://doi.org/10.1126/science.8303295.

Chen, X., Schauder, S., Potier, N., van Dorsselaer, A., Pelczer, I., Bassler, B.L., and Hughson, F.M. (2002). Structural identification of a bacterial quorum-sensing signal containing boron. Nature *415*, 545-549. https://doi.org/10.1038/415545a.

Chen, Z.-D., Chen, W.-Q., Wang, Z.-Y., Cao, D.-N., Wu, N., and Li, J. (2018). Antidepressant-like action of agmatine in the acute and sub-acute mouse models of depression: a receptor mechanism study. Metabolic brain disease *33*, 1721-1731. https://doi.org/10.1007/s11011-018-0280-9.

Ciruela, F. (2008). Fluorescence-based methods in the study of protein-protein interactions in living cells. Current opinion in biotechnology *19*, 338-343. https://doi.org/10.1016/j.copbio.2008.06.003.

Cormack, B.P., Valdivia, R.H., and Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). Gene *173*, 33-38. https://doi.org/10.1016/0378-1119(95)00685-0.

Cubitt, A.B., Heim, R., Adams, S.R., Boyd, A.E., Gross, L.A., and Tsien, R.Y. (1995). Understanding, improving and using green fluorescent proteins. Trends in biochemical sciences *20*, 448-455. https://doi.org/10.1016/s0968-0004(00)89099-4.

Durand, J.M.B., and Björk, G.R. (2003). Putrescine or a combination of methionine and arginine restores virulence gene expression in a tRNA modification-deficient mutant of Shigella flexneri: a possible role in adaptation of virulence. Molecular microbiology *47*, 519-527. https://doi.org/10.1046/j.1365-2958.2003.03314.x.

Dwyer, M.A., and Hellinga, H.W. (2004). Periplasmic binding proteins: a versatile superfamily for protein engineering. Current opinion in structural biology *14*, 495-504. https://doi.org/10.1016/j.sbi.2004.07.004.

Felder, C.B., Graul, R.C., Lee, A.Y., Merkle, H.P., and Sadee, W. (1999). The Venus flytrap of periplasmic binding proteins: an ancient protein module present in multiple drug receptors. AAPS pharmSci *1*, E2. https://doi.org/10.1208/ps010202.

Felise, H.B., Nguyen, H.V., Pfuetzner, R.A., Barry, K.C., Jackson, S.R., Blanc, M.-P., Bronstein, P.A., Kline, T., and Miller, S.I. (2008). An inhibitor of gram-negative bacterial virulence protein secretion. Cell host & microbe *4*, 325-336. https://doi.org/10.1016/j.chom.2008.08.001.

Frommer, W.B., Davidson, M.W., and Campbell, R.E. (2009). Genetically encoded biosensors based on engineered fluorescent proteins. Chemical Society reviews *38*, 2833-2841. https://doi.org/10.1039/b907749a.

Fukami-Kobayashi, K., Tateno, Y., and Nishikawa, K. (1999). Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. Journal of molecular biology *286*, 279-290. https://doi.org/10.1006/jmbi.1998.2454.

Gerner, E.W., and Meyskens, F.L. (2004). Polyamines and cancer: old molecules, new understanding. Nature reviews. Cancer *4*, 781-792. https://doi.org/10.1038/nrc1454.

Ghosh, P. (2004). Process of protein transport by the type III secretion system. Microbiology and molecular biology reviews : MMBR *68*, 771-795. https://doi.org/10.1128/MMBR.68.4.771-795.2004.

Greenwald, E.C., Mehta, S., and Zhang, J. (2018). Genetically Encoded Fluorescent Biosensors Illuminate the Spatiotemporal Regulation of Signaling Networks. Chemical reviews *118*, 11707-11794. https://doi.org/10.1021/acs.chemrev.8b00333.

Heim, R., Cubitt, A.B., and Tsien, R.Y. (1995). Improved green fluorescence. Nature *373*, 663-664. https://doi.org/10.1038/373663b0.

Herud-Sikimić, O., Stiel, A.C., Kolb, M., Shanmugaratnam, S., Berendzen, K.W., Feldhaus, C., Höcker, B., and Jürgens, G. (2021). A biosensor for the direct visualization of auxin. Nature *592*, 768-772. https://doi.org/10.1038/s41586-021-03425-2.

Hollenstein, K., Frei, D.C., and Locher, K.P. (2007). Structure of an ABC transporter in complex with its binding protein. Nature *446*, 213-216. https://doi.org/10.1038/nature05626.

Huang, S.C., Panagiotidis, C.A., and Canellakis, E.S. (1990). Transcriptional effects of polyamines on ribosomal proteins and on polyamine-synthesizing enzymes in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America *87*, 3464-3468. https://doi.org/10.1073/pnas.87.9.3464.

Igarashi, K., and Kashiwagi, K. (2006). Polyamine Modulon in Escherichia coli: genes involved in the stimulation of cell growth by polyamines. Journal of biochemistry *139*, 11-16. https://doi.org/10.1093/jb/mvj020.

Igarashi, K., and Kashiwagi, K. (2010). Modulation of cellular function by polyamines. The international journal of biochemistry & cell biology *42*, 39-51. https://doi.org/10.1016/j.biocel.2009.07.009.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., and Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589. https://doi.org/10.1038/s41586-021-03819-2.

Kashiwagi, K., Miyamoto, S., Nukui, E., Kobayashi, H., and Igarashi, K. (1993). Functions of potA and potD proteins in spermidine-preferential uptake system in Escherichia coli. Journal of Biological Chemistry *268*, 19358-19363.

Kashiwagi, K., Suzuki, T., Suzuki, F., Furuchi, T., Kobayashi, H., and Igarashi, K. (1991). Coexistence of the genes for putrescine transport protein and ornithine decarboxylase at 16 min on Escherichia coli chromosome. Journal of Biological Chemistry *266*, 20922-20927. https://doi.org/10.1016/S0021-9258(18)54798-0.

Keyser, P., Elofsson, M., Rosell, S., and Wolf-Watz, H. (2008). Virulence blockers as alternatives to antibiotics: type III secretion inhibitors against Gram-negative bacteria. Journal of internal medicine *264*, 17-29. https://doi.org/10.1111/j.1365-2796.2008.01941.x.

Kotagale, N.R., Taksande, B.G., Wadhwani, P.J., Palhade, M.W., Mendhi, S.M., Gawande, D.Y., Hadole, P.N., and Chopde, C.T. (2012). Psychopharmacological study of agmatine in behavioral tests of schizophrenia in rodents. Pharmacology, biochemistry, and behavior *100*, 398-403. https://doi.org/10.1016/j.pbb.2011.09.006.

Kremers, G.-J., Gilbert, S.G., Cranfill, P.J., Davidson, M.W., and Piston, D.W. (2011). Fluorescent proteins at a glance. Journal of cell science *124*, 157-160. https://doi.org/10.1242/jcs.072744.

Lambert, T.J. (2019). FPbase: a community-editable fluorescent protein database. Nature methods *16*, 277-278. https://doi.org/10.1038/s41592-019-0352-8.

Lee, J.Y., Krieger, J., Herguedas, B., García-Nafría, J., Dutta, A., Shaikh, S.A., Greger, I.H., and Bahar, I. (2019). Druggability Simulations and X-Ray Crystallography Reveal a Ligand-Binding Site in the GluA3 AMPA Receptor N-Terminal Domain. Structure (London, England : 1993) *27*, 241-252.e3. https://doi.org/10.1016/j.str.2018.10.017.

Leman, J.K., Weitzner, B.D., Lewis, S.M., Adolf-Bryfogle, J., Alam, N., Alford, R.F., Aprahamian, M., Baker, D., Barlow, K.A., and Barth, P., et al. (2020). Macromolecular modeling and design in Rosetta: recent methods and frameworks. Nature methods *17*, 665-680. https://doi.org/10.1038/s41592-020-0848-2.

Liu, Y., and Kuhlman, B. (2006). RosettaDesign server for protein design. Nucleic acids research *34*, W235-8. https://doi.org/10.1093/nar/gkl163.

Lorimier, R.M. de, Smith, J.J., Dwyer, M.A., Looger, L.L., Sali, K.M., Paavola, C.D., Rizk, S.S., Sadigov, S., Conrad, D.W., and Loew, L., et al. (2002). Construction of a fluorescent biosensor family. Protein science : a publication of the Protein Society *11*, 2655-2675. https://doi.org/10.1110/ps.021860.

Machius, M., Brautigam, C.A., Tomchick, D.R., Ward, P., Otwinowski, Z., Blevins, J.S., Deka, R.K., and Norgard, M.V. (2007). Structural and biochemical basis for polyamine binding to the Tp0655 lipoprotein of Treponema pallidum: putative role for Tp0655 (TpPotD) as a polyamine receptor. Journal of molecular biology *373*, 681-694. https://doi.org/10.1016/j.jmb.2007.08.018.

Mächtel, R., Narducci, A., Griffith, D.A., Cordes, T., and Orelle, C. (2019). An integrated transport mechanism of the maltose ABC importer. Research in microbiology *170*, 321-337. https://doi.org/10.1016/j.resmic.2019.09.004.

Mandal, S., Mandal, A., Johansson, H.E., Orjalo, A.V., and Park, M.H. (2013). Depletion of cellular polyamines, spermidine and spermine, causes a total arrest in translation and growth in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America *110*, 2169-2174. https://doi.org/10.1073/pnas.1219002110.

Marvin, J.S., Borghuis, B.G., Tian, L., Cichon, J., Harnett, M.T., Akerboom, J., Gordus, A., Renninger, S.L., Chen, T.-W., and Bargmann, C.I., et al. (2013). An optimized fluorescent probe for visualizing glutamate neurotransmission. Nature methods *10*, 162-170. https://doi.org/10.1038/nmeth.2333.

Marvin, J.S., Scholl, B., Wilson, D.E., Podgorski, K., Kazemipour, A., Müller, J.A., Schoch, S., Quiroz, F.J.U., Rebola, N., and Bao, H., et al. (2018). Stability, affinity, and chromatic variants of the glutamate sensor iGluSnFR. Nature methods *15*, 936-939. https://doi.org/10.1038/s41592-018-0171-3.

Marvin, J.S., Schreiter, E.R., Echevarría, I.M., and Looger, L.L. (2011). A genetically encoded, high-signal-to-noise maltose sensor. Proteins *79*, 3025-3036. https://doi.org/10.1002/prot.23118.

Marvin, J.S., Shimoda, Y., Magloire, V., Leite, M., Kawashima, T., Jensen, T.P., Kolb, I., Knott, E.L., Novak, O., and Podgorski, K., et al. (2019). A genetically encoded fluorescent sensor for in vivo imaging of GABA. Nature methods *16*, 763-770. https://doi.org/10.1038/s41592-019-0471-2.

Moussatova, A., Kandt, C., O'Mara, M.L., and Tieleman, D.P. (2008). ATP-binding cassette transporters in Escherichia coli. Biochimica et biophysica acta *1778*, 1757-1771. https://doi.org/10.1016/j.bbamem.2008.06.009.

Munro, G.F., Hercules, K., Morgan, J., and Sauerbier, W. (1972). Dependence of the Putrescine Content of Escherichia coli on the Osmotic Strength of the Medium. Journal of Biological Chemistry *247*, 1272-1280. https://doi.org/10.1016/S0021-9258(19)45643-3.

Nasu, Y., Shen, Y., Kramer, L., and Campbell, R.E. (2021). Structure- and mechanism-guided design of single fluorescent protein-based biosensors. Nature chemical biology *17*, 509-518. https://doi.org/10.1038/s41589-020-00718-x.

Ninkovic, M., Stevanovic, I., Stojanovic, I., Ljubisavljevic, S., Basic, J., and Peric, P. (2015). The Use of Agmatine Provides the New Insight in an Experimental Model of Multiple Sclerosis. Neurochemical research *40*, 1719-1727. https://doi.org/10.1007/s11064-015-1655-5.

Ormö, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., and Remington, S.J. (1996). Crystal structure of the Aequorea victoria green fluorescent protein. Science (New York, N.Y.) *273*, 1392-1395. https://doi.org/10.1126/science.273.5280.1392.

Patriarchi, T., Cho, J.R., Merten, K., Howe, M.W., Marley, A., Xiong, W.-H., Folk, R.W., Broussard, G.J., Liang, R., and Jang, M.J., et al. (2018). Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. Science (New York, N.Y.) *360*. https://doi.org/10.1126/science.aat4422.

Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T.C., and Waldo, G.S. (2006). Engineering and characterization of a superfolder green fluorescent protein. Nature biotechnology *24*, 79-88. https://doi.org/10.1038/nbt1172.

Pegg, A.E. (1988). Polyamine metabolism and its importance in neoplastic growth and a target for chemotherapy. Cancer research *48*, 759-774.

Piletz, J.E., Aricioglu, F., Cheng, J.-T., Fairbanks, C.A., Gilad, V.H., Haenisch, B., Halaris, A., Hong, S., Lee, J.E., and Li, J., et al. (2013). Agmatine: clinical applications after 100 years in translation. Drug discovery today *18*, 880-893. https://doi.org/10.1016/j.drudis.2013.05.017.

Pistocchi, R., Kashiwagi, K., Miyamoto, S., Nukui, E., Sadakata, Y., Kobayashi, H., and Igarashi, K. (1993). Characteristics of the operon for a putrescine transport system that maps at 19 minutes on the Escherichia coli chromosome. Journal of Biological Chemistry *268*, 146-152. https://doi.org/10.1016/S0021-9258(18)54126-0.

Piston, D.W., and Kremers, G.-J. (2007). Fluorescent protein FRET: the good, the bad and the ugly. Trends in biochemical sciences *32*, 407-414. https://doi.org/10.1016/j.tibs.2007.08.003.

Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G., and Cormier, M.J. (1992). Primary structure of the Aequorea victoria green-fluorescent protein. Gene *111*, 229-233. https://doi.org/10.1016/0378-1119(92)90691-h.

Quiocho, F.A., and Ledvina, P.S. (1996). Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. Molecular microbiology *20*, 17-25. https://doi.org/10.1111/j.1365-2958.1996.tb02484.x.

Reis, D.J., and Regunathan, S. (2000). Is agmatine a novel neurotransmitter in brain? Trends in Pharmacological Sciences *21*, 187-193. https://doi.org/10.1016/S0165-6147(00)01460-7.

Remington, S.J. (2006). Fluorescent proteins: maturation, photochemistry and photophysics. Current opinion in structural biology *16*, 714-721. https://doi.org/10.1016/j.sbi.2006.10.001.

Rial, N.S., Meyskens, F.L., and Gerner, E.W. (2009). Polyamines as mediators of APC-dependent intestinal carcinogenesis and cancer chemoprevention. Essays in biochemistry *46*, 111-124. https://doi.org/10.1042/bse0460008.

Rodriguez, E.A., Campbell, R.E., Lin, J.Y., Lin, M.Z., Miyawaki, A., Palmer, A.E., Shu, X., Zhang, J., and Tsien, R.Y. (2017). The Growing and Glowing Toolbox of Fluorescent and Photoactive Proteins. Trends in biochemical sciences *42*, 111-129. https://doi.org/10.1016/j.tibs.2016.09.010.

Roth, B.L., Sheffler, D.J., and Kroeze, W.K. (2004). Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. Nature reviews. Drug discovery *3*, 353-359. https://doi.org/10.1038/nrd1346.

Scheib, U., Shanmugaratnam, S., Farías-Rico, J.A., and Höcker, B. (2014). Change in protein-ligand specificity through binding pocket grafting. Journal of structural biology *185*, 186-192. https://doi.org/10.1016/j.jsb.2013.06.002.

Schreiber, G., and Keating, A.E. (2011). Protein binding specificity versus promiscuity. Current opinion in structural biology *21*, 50-61. https://doi.org/10.1016/j.sbi.2010.10.002.

Shah, P., and Swiatlo, E. (2006). Immunization with polyamine transport protein PotD protects mice against systemic infection with Streptococcus pneumoniae. Infection and immunity *74*, 5888-5892. https://doi.org/10.1128/IAI.00553-06.

Shcherbakova, D.M., and Verkhusha, V.V. (2013). Near-infrared fluorescent proteins for multicolor in vivo imaging. Nature methods *10*, 751-754. https://doi.org/10.1038/nmeth.2521.

Shimomura, O. (2005). The discovery of aequorin and green fluorescent protein. Journal of microscopy *217*, 1-15. https://doi.org/10.1111/j.0022-2720.2005.01441.x.

Shimomura, O., Johnson, F.H., and Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. Journal of cellular and comparative physiology *59*, 223-239. https://doi.org/10.1002/jcp.1030590302.

Silla Santos, M.H. (1996). Biogenic amines: their importance in foods. International journal of food microbiology *29*, 213-231. https://doi.org/10.1016/0168-1605(95)00032-1.

Soda, K. (2011). The mechanisms by which polyamines accelerate tumor spread. Journal of experimental & clinical cancer research : CR *30*, 95. https://doi.org/10.1186/1756-9966-30-95.

Song, J., Hur, B.E., Bokara, K.K., Yang, W., Cho, H.J., Park, K.A., Lee, W.T., Lee, K.M., and Lee, J.E. (2014). Agmatine improves cognitive dysfunction and prevents cell death in a streptozotocin-induced Alzheimer rat model. Yonsei medical journal *55*, 689-699. https://doi.org/10.3349/ymj.2014.55.3.689.

Sugiyama, S., Matsuo, Y., Maenaka, K., Vassylyev, D.G., Matsushima, M., Kashiwagi, K., Igarashi, K., and Morikawa, K. (1996). The 1.8-A X-ray structure of the Escherichia coli PotD protein complexed with spermidine and the mechanism of polyamine binding. Protein science : a publication of the Protein Society *5*, 1984-1990. https://doi.org/10.1002/pro.5560051004.

Sun, F., Zeng, J., Jing, M., Zhou, J., Feng, J., Owen, S.F., Luo, Y., Li, F., Wang, H., and Yamaguchi, T., et al. (2018). A Genetically Encoded Fluorescent Sensor Enables Rapid and Specific Detection of Dopamine in Flies, Fish, and Mice. Cell *174*, 481-496.e19. https://doi.org/10.1016/j.cell.2018.06.042.

Tabor, C.W., and Tabor, H. (1984). Polyamines. Annual review of biochemistry *53*, 749-790. https://doi.org/10.1146/annurev.bi.53.070184.003533.

Tabor, C.W., and Tabor, H. (1985). Polyamines in microorganisms. Microbiological Reviews *49*, 81-99. https://doi.org/10.1128/mr.49.1.81-99.1985.

Tainaka, K., Sakaguchi, R., Hayashi, H., Nakano, S., Liew, F.F., and Morii, T. (2010). Design strategies of fluorescent biosensors based on biological macromolecular receptors. Sensors (Basel, Switzerland) *10*, 1355-1376. https://doi.org/10.3390/s100201355.

Tam, R., and Saier, M.H. (1993). Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. Microbiological Reviews *57*, 320-346.

Tang, C., Schwieters, C.D., and Clore, G.M. (2007). Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. Nature *449*, 1078-1082. https://doi.org/10.1038/nature06232.

Terui, Y., Saroj, S.D., Sakamoto, A., Yoshida, T., Higashi, K., Kurihara, S., Suzuki, H., Toida, T., Kashiwagi, K., and Igarashi, K. (2014). Properties of putrescine uptake by PotFGHI and PuuP and their physiological significance in Escherichia coli. Amino acids *46*, 661-670. https://doi.org/10.1007/s00726-013-1517-x.

Tomosugi, W., Matsuda, T., Tani, T., Nemoto, T., Kotera, I., Saito, K., Horikawa, K., and Nagai, T. (2009). An ultramarine fluorescent protein with increased photostability and pH insensitivity. Nature methods *6*, 351-353. https://doi.org/10.1038/nmeth.1317.

Unger, E.K., Keller, J.P., Altermatt, M., Liang, R., Matsui, A., Dong, C., Hon, O.J., Yao, Z., Sun, J., and Banala, S., et al. (2020). Directed Evolution of a Selective and Sensitive Serotonin Sensor via Machine Learning. Cell *183*, 1986-2002.e26. https://doi.org/10.1016/j.cell.2020.11.040.

Vassylyev, D.G., Tomitori, H., Kashiwagi, K., Morikawa, K., and Igarashi, K. (1998). Crystal structure and mutational analysis of the Escherichia coli putrescine receptor. Structural basis for substrate specificity. Journal of Biological Chemistry *273*, 17604-17609. https://doi.org/10.1074/jbc.273.28.17604.

Ware, D., Jiang, Y., Lin, W., and Swiatlo, E. (2006). Involvement of potD in Streptococcus pneumoniae polyamine transport and pathogenesis. Infection and immunity *74*, 352-361. https://doi.org/10.1128/IAI.74.1.352-361.2006.

Wu, D., Lim, S.C., Dong, Y., Wu, J., Tao, F., Zhou, L., Zhang, L.-H., and Song, H. (2012). Structural basis of substrate binding specificity revealed by the crystal structures of polyamine receptors SpuD and SpuE from Pseudomonas aeruginosa. Journal of molecular biology *416*, 697-712. https://doi.org/10.1016/j.jmb.2012.01.010.

Zhang, L., Patel, H.N., Lappe, J.W., and Wachter, R.M. (2006). Reaction progress of chromophore biogenesis in green fluorescent protein. Journal of the American Chemical Society *128*, 4766-4772. https://doi.org/10.1021/ja0580439.

Zhang, Y., Sun, X., Qian, Y., Yi, H., Song, K., Zhu, H., Zonta, F., Chen, W., Ji, Q., and Miersch, S., et al. (2019). A Potent Anti-SpuE Antibody Allosterically Inhibits Type III Secretion System and Attenuates Virulence of Pseudomonas Aeruginosa. Journal of molecular biology *431*, 4882-4896. https://doi.org/10.1016/j.jmb.2019.10.026.

Zhou, L., Wang, J., and Zhang, L.-H. (2007). Modulation of bacterial Type III secretion system by a spermidine transporter dependent signaling pathway. PloS one *2*, e1291. https://doi.org/10.1371/journal.pone.0001291.

Zomkowski, A.D.E., Hammes, L., Lin, J., Calixto, J.B., Santos, A.R.S., and Rodrigues, A.L.S. (2002). Agmatine produces antidepressant-like effects in two models of depression in mice. Neuroreport *13*, 387-391. https://doi.org/10.1097/00001756-200203250-00005.

13.Acknowledgements

I would like to thank…

… Birte for the opportunity to do my Ph.D. under her guidance and for all her advice since I joined her group.

… Sooruban, who contributed a lot to the whole project. Starting out as my supervisor and colleague, he quickly turned into to a friend I was able to depend on throughout this time.

… past and present colleagues for the great working atmosphere.

… my friends, who accompanied me during my Ph.D. journey.

… Holly, for putting up with me during the last months and bringing my English skills for this thesis to the next level.

And finally, I would like to thank my family for the support throughout my whole time in Bayreuth, starting with my Bachelor and finishing with this Ph.D. It would not have been possible without you.

14. Eidesstattliche Versicherung und Erklärung

(§ 8 Satz 2 Nr. 3 PromO Fakultät)

Hiermit versichere ich eidesstattlich, dass ich die Arbeit selbstständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe (vgl. Art. 64 Abs. 1 Satz 6 BayHSchG).

(§ 8 Satz 2 Nr. 3 PromO Fakultät)

Hiermit erkläre ich, dass ich die Dissertation nicht bereits zur Erlangung eines akademischen Grades eingereicht habe und dass ich nicht bereits diese oder eine gleichartige Doktorprüfung endgültig nicht bestanden habe.

(§ 8 Satz 2 Nr. 4 PromO Fakultät)

Hiermit erkläre ich, dass ich Hilfe von gewerblichen Promotionsberatern bzw. –vermittlern oder ähnlichen Dienstleistern weder bisher in Anspruch genommen habe noch künftig in Anspruch nehmen werde.

(§ 8 Satz 2 Nr. 7 PromO Fakultät)

Hiermit erkläre ich mein Einverständnis, dass die elektronische Fassung der Dissertation unter Wahrung meiner Urheberrechte und des Datenschutzes einer gesonderten Überprüfung unterzogen werden kann.

(§ 8 Satz 2 Nr. 8 PromO Fakultät)

Hiermit erkläre ich mein Einverständnis, dass bei Verdacht wissenschaftlichen Fehlverhaltens Ermittlungen durch universitätsinterne Organe der wissenschaftlichen Selbstkontrolle stattfinden können.

Bayreuth, 21.02.2023