



Mapping L-phenylalanine binding to the regulatory domains of human tryptophan hydroxylase isoforms

Vedel, Ida Marie

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Vedel, I. M. (2022). *Mapping L-phenylalanine binding to the regulatory domains of human tryptophan hydroxylase isoforms*. DTU Chemistry.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Mapping L-phenylalanine binding to the regulatory domains of *human* tryptophan hydroxylase isoforms

Ph.D. Thesis
Ida Marie Vedel



Department of Chemistry
Technical University of Denmark
April 2022

PREFACE

This dissertation is submitted to the Technical University of Denmark in partial fulfilment of the requirements of the Ph.D. degree. The work was carried out at the Department of Chemistry, Technical University of Denmark, under the supervision of Professor Günther H. J. Peters and Professor Pernille Harris. In addition, nuclear magnetic resonance spectroscopy experiments were carried out at the Department of Biology, University of Copenhagen, in the research group of Professor Birthe B. Kragelund. The cryo-electron microscopy data presented in this thesis was acquired and analyzed by Ph.D. Zhenwei Zhang at the Department of Structural Dynamics, Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany, in the research group of Professor Holger Stark.

Ida Marie Vedel

Kongens Lyngby, April 2022

ACKNOWLEDGEMENTS

This project would not have been the same without the involvement of many people, and I would like to thank everyone who made the project possible. First of all, thanks to my supervisors Professor Günther H. J. Peters and Professor Pernille Harris for giving me the opportunity to work on this project, which I have truly enjoyed working on for the past three years. Thanks to Günther, for help, scientific guidance, support, and for letting me pursue the parts of the project I found interesting. Thanks to Pernille, for valuable input, perspectives, and scientific discussions.

The project would have been very different without the collaboration on NMR spectroscopy experiments with Professor Birthe B. Kragelund and NMR manager Andreas Prestel from the Structural Biology and NMR Laboratory at the Department of Biology, University of Copenhagen, and a special thanks therefore goes to them. To Birthe, for welcoming me in your lab, for many suggestions on experiments to try, scientific discussions, and your always enthusiastic approach to the project. Thanks to Andi, for introducing me to the world of NMR, for sharing your NMR excitement, and for countless data discussions. Thanks also to Signe Sjørup for practical help around the lab, to Yulian Gavrilov for extensive advice on analysis of NMR relaxation data, and thanks to everyone else in SBiNLab for excellent company and for making me feel welcome.

I am additionally grateful to Ph.D. Zhenwei Zhang and Professor Holger Stark from the Department of Structural Dynamics at the Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany, for collaboration on the cryo-EM structure of TPH2. Thanks to Holger Stark for hosting me in his lab, and thanks to Zhenwei for a good collaboration.

Previous work by Assoc. Professor Hans E. M. Christensen and his research group on TPH that this work could build upon is additionally acknowledged. Thanks also to Master students Beatrice F. Schultz, Gabriella Björkengren, Maja van Hees, and Marta Vidal del Santo, who I co-supervised, for contributing to the knowledge on TPH with the work on their projects. I would also like to acknowledge David F. Nielsen for technical assistance in the labs at DTU. Thanks for always being helpful with ordering chemicals and supplies, for help with preparative work, and for good teamwork during the laboratory renovation period. Thanks to Assoc. Professor Charlotte H. Gotfredsen for letting me use the NMR facility at DTU and for discussions on NMR data. Thanks to Mariusz Kubus for technical assistance and introduction to the SEC-MALS equipment.

A huge thanks to my present and former colleagues and office mates at DTU: Helena D. Tjørnelund, Suk Kyu Ko, Tine M. Frimann, Alina Kulakova, Christin Pohl, Sowmya Indrakumar, Line A. Ryberg, Ulf Molich, Sujata Mahapatra, Sarah Hedberg, Natalia T. Skawinska, Marta Vidal del Santo, Aliihsan Bagci, Michael J. J. Bartholdy, and Nina R. Hamburger. Thanks for good company in the office and in the lab, and thanks for many good occasions outside of DTU including dinners and board game nights. Thanks to Alina, for invaluable help and input on SAXS data acquisition and analysis, and to both Alina and Helena for reading parts of my thesis. Thanks to Christin for Sunday runs and informal scientific discussions. Also thanks to Johanne M. Nielsen and Maria B. Bang for contributing to a positive atmosphere on the second floor of 206.

I feel very privileged to have wonderful family and friends. Thanks to all of you, for your support and for all the moments we share. Thanks to my dad and brother, for taking my mind of protein chemistry on our summer bicycle-rides and thanks to my mum and dad for autumn mushroom-hunting. Special thanks to Stine, Benny, Caro, Nina, and Marie, for a valued friendship since the second grade, and thanks to Maja for pleasant (yet tough) running company. Thanks to Julie, for being there when it matters. Last, but definitely not least, thanks to Sylvester, for lots of input on my figures, but above all for infinite amounts of love and support.

ABSTRACT

Tryptophan hydroxylase (TPH) catalyzes the reaction of L-Tryptophan into L-5-hydroxytryptophan, which is the rate-limiting step in the biosynthesis of serotonin. In the peripheral tissues, L-Trp hydroxylation is catalyzed by tryptophan hydroxylase isoform 1 (TPH1), while tryptophan hydroxylase isoform 2 (TPH2) catalyzes the reaction in the brain. The homologous homo-tetrameric enzymes contain three domains: A regulatory (RD), a catalytic (CD), and a tetramerization domain (TD). As a hormone and neurotransmitter, serotonin influences a variety of body functions and behaviors e.g. liver function, gastrointestinal motility, appetite, aggression, and attention, and consequently dysregulation of the serotonin system is associated with several diseases. Therefore, to shed further light on serotonin regulation and serotonin related diseases, a detailed understanding of TPH's regulation is important. In this project, the *human* TPH isoforms have been investigated in two different studies.

The first study concerned a structural characterization of TPH2. L-Phe has previously been shown to increase stability and induce dimerization of TPH2 variants containing the RD, which has been connected to a potential binding site in the RD. In this study, the structure of an N-terminally truncated variant of the isolated RD in complex with L-Phe was obtained using nuclear magnetic resonance (NMR) spectroscopy. L-Phe was bound at the dimer interface in two identical binding pockets, as also seen for a homologous enzyme, phenylalanine hydroxylase. It was additionally observed using NMR spectroscopy that L-Phe was superior to the natural substrate, L-Trp, as a ligand of the isolated RD. Cryo-electron microscopy was used to study an N-terminally truncated variant of the complete tetrameric enzyme, which resulted in a low-resolution structure of the variant with dimerized RDs. 2D class averages additionally indicated that the RDs exist in a monomer-dimer equilibrium in tetrameric TPH2, where RD monomers move randomly relative to the CDs.

Assessing the residues involved in L-Phe binding to the RD of TPH2, and comparing the sequences of TPH1 and TPH2, showed that many residues from the binding pockets were conserved in TPH1. The second study therefore concerned ligand binding to the RD of TPH1. Size exclusion chromatography coupled to multi-angle light scattering was used to investigate the oligomeric state of different TPH1 variants without the TD. The results showed that variants containing the RD, formed dimers and that dimer formation was stabilized by the presence of L-Phe, but not L-Trp. Additionally, NMR spectroscopy showed that the RD dimer of TPH1 bound L-Phe.

DANSK RESUMÉ

Tryptophan hydroxylase katalyserer reaktionen fra L-tryptophan (L-Trp) til L-5-hydroxytryptophan der er det hastighedsbestemmende trin i serotonin biosyntesen. I det perifære væv bliver L-Trp hydroxyleret af tryptophan hydroxylase isoform 1 (TPH1), mens tryptophan hydroxylase isoform 2 katalyserer den samme reaktion i hjernen (TPH2). De homologe homo-tetrameriske enzymer består af tre domæner: Et regulatorisk, et katalytisk og et tetrameriseringsdomæne. Som hormon og neurotransmitter påvirker serotonin en række fysiologiske og mentale funktioner f.eks. leverfunktion, tarmfunktion, appetit, aggression og opmærksomhed. Forstyrrelser i serotoninbalancen er således forbundet med flere forskellige sygdomme. En detaljeret forståelse af TPHs regulering er derfor vigtig for at opnå en større forståelse af serotoninregulering og serotonin relaterede sygdomme. I dette projekt er de to *humane* TPH isoformer blevet undersøgt i to forskellige studier.

Det første studie omhandlede en strukturel karakterisering af TPH2. Det er tidligere vist at L-fenylalanin (L-Phe) øger stabiliteten af TPH2 varianter der inkluderer det regulatoriske domæne og forskubber monomer-dimer ligevægten af varianterne mod dimer, hvilket indikerer eksistensen af en bindingslomme i det regulatoriske domæne. I dette studie blev NMR (nuclear magnetic resonance) spektroskopi brugt til at bestemme strukturen af en N-terminal trunkeret variant af det isolerede regulatoriske domæne af TPH2 i kompleks med L-Phe. L-Phe var bundet i dimer grænsefladen i to identiske bindingslommer, hvilket også er observeret for det homologe enzym, fenylalanin hydroxylase. Derudover viste NMR spektroskopi, at L-Phe var en markant bedre ligand for det isolerede regulatoriske domæne end det naturlige substrat L-Trp. Kryo-elektron mikroskopi blev brugt til at undersøge en N-terminal trunkeret variant af det komplette tetrameriske enzym, hvilket resulterede i en struktur i lav opløsning af varianten med dimeriserede regulatoriske domæner. 2D klassificeringerne indikerede desuden at de regulatoriske domæner eksisterer i en monomer-dimer ligevægt i TPH2 tetrameren, hvor regulatorisk domæne monomerer kan bevæge sig tilfældigt relativt til de katalytiske domæner.

Ved at undersøge aminosyreresterne involveret i L-Phe binding i det regulatoriske domæne af TPH2 og sammenligne sekvenserne af TPH1 og TPH2, blev det observeret at mange aminosyrerester i bindingslommen var identiske i TPH1. Det andet studie omhandlede derfor binding af ligander til det regulatoriske domæne af TPH1. Størrelseskromatografi koblet til lysspredningsmålinger blev brugt til at undersøge oligomerisering af TPH1 varianter uden tetrameriserings domænet. Resultaterne viste at varianter der inkluderede det regulatoriske domæne var dimerer, og at dimerkonformationen blev stabiliseret af L-Phe men ikke L-Trp. Derudover viste NMR spektroskopi at dimeren af det regulatoriske domæne af TPH1 kunne binde L-Phe.

LIST OF ABBREVIATIONS

30 Kan	30 μ M/mL kanamycin sulfate
3CP	<i>human</i> rhinovirus 3C protease
4a-hydroxy-BH ₄	4a-carbinolamine tetrahydrobiopterin
5-HT	5-hydroxytryptamine, serotonin
5-HTP	L-5-hydroxytryptophan, 5-hydroxy-L-Trp
AAAH	Aromatic amino acid hydroxylase
AC	Affinity chromatography
ACT	Aspartate kinase, chorismate mutase, and TyrA domain
AS	Active state
BBB	Blood-brain-barrier
BH ₂	7,8-dihydrobiopterin
BH ₄	5,6,7,8-tetrahydrobiopterin
CaMKII	Ca ²⁺ /Calmodulin-dependent protein kinase II
CD	Catalytic domain
cryo-EM	Cryo-electron microscopy
CSP	Chemical shift perturbation
CV	Column volume
dn/dc	Refractive index increment
dRI	Differential refractive index
DSF	Differential scanning fluorimetry
DTT	DL-dithiothreitol
FDA	U.S. Food & Drug Administration
FF	Force field
FID	Free induction decay
FT	Fourier transform, Fourier transformation, Fourier transformed
GF	Gel filtration
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
het-NOE	heteronuclear nuclear Overhauser effect
IPTG	isopropyl- β -D-thiogalactopyranoside
L-Dopa	L-3,4-dihydroxyphenylalanine
LNAA	large neutral amino acid
LS	Light scattering
MALS	Multi-angle light scattering
MBP	Maltose binding protein
Mw	Molecular weight
NMR	Nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
N-tail	N-terminal residues before the folded part of the AAAH's regulatory domains
OD ₆₀₀	Optical density at 600 nm
PAH	Phenylalanine hydroxylase

PDB	Protein data bank
PKA	Protein kinase A
PKU	phenylketonuria
PL	Protein-Ligand
PSI-BLAST	Position-Specific Iterated BLAST
RD	Regulatory domain
RF	Radio frequency
RI	Refractive index
RS	Resting state
SAXS	Small-angle X-ray scattering
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
TD	Tetramerization domain
TH	Tyrosine hydroxylase
TOCSY	Total correlation spectroscopy
TPH	Tryptophan hydroxylase
TPH1	Tryptophan hydroxylase isoform 1
TPH2	Tryptophan hydroxylase isoform 2
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
UV	Ultraviolet

TABLE OF CONTENTS

Preface	i
Acknowledgements	iii
Abstract	v
Dansk resumé	vii
List of abbreviations	ix
Introduction and outline of thesis	1
1 Tryptophan hydroxylase and the aromatic amino acid hydroxylases	3
1.1 The aromatic amino acid hydroxylases	3
1.2 Structure of the AAAHs	5
1.3 Regulatory mechanisms of the AAAH	13
2 Purification of <i>human</i> tryptophan hydroxylase variants	17
2.1 Introduction	17
2.2 Nomenclature and protein details	17
2.3 Experimental procedures	18
2.4 Results	22
2.5 Discussion	28
2.6 Conclusion	29
3 Methods	31
3.1 Nuclear magnetic resonance (NMR) spectroscopy	31
3.2 Small-angle X-ray scattering (SAXS)	37
3.3 Size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS)	38
3.4 Single particle cryo-electron microscopy (Cryo-EM)	39
4 Tryptophan hydroxylase 2 (Manuscript I)	41
5 Tryptophan hydroxylase 1 (Manuscript II)	79
6 Concluding remarks	97
7 Future perspectives	99
Bibliography	101
Appendix	111
A Sequence alignments	112
A1 Alignment of the <i>human</i> AAAHs	112
A2 Alignment of the catalytic domains of <i>human</i> TPH1, <i>human</i> TPH2 and <i>chicken</i> TPH1	113
B Purification of TPH variants	114

B1 Purification of TPH2.....	115
B2 Purification of NΔ47-TPH2	116
B3 Purification of NΔ47-TPH2-RC	117
B4 Purification of TPH2-CT.....	118
B5 Purification of TPH2-R.....	119
B6 Purification of NΔ47-TPH2-R.....	120
B7 Purification of TPH1-C.....	121
B8 Purification of TPH1-RC	122
C Sequences of protein variants.....	123
D ¹ H- ¹⁵ N HSQC spectra of NΔ47-TPH2-R	125
E A kinetic study of <i>human</i> TPH2 variants.....	128

INTRODUCTION AND OUTLINE OF THESIS

Serotonin is an important neurotransmitter and hormone implicated in a variety of diseases as e.g. depression, bipolar disorder, irritable bowel syndrome, and carcinoid syndrome¹⁻⁴. The first and rate-limiting step in the biosynthesis of serotonin is catalyzed by the enzymes tryptophan hydroxylase 1 (TPH1) in the peripheral tissues and tryptophan hydroxylase 2 (TPH2) in the brain^{5,6}, and consequently the regulation of TPH1 and TPH2 are relevant for serotonin regulation and serotonin related diseases. As TPH1 and TPH2 regulate each their spatially separated serotonin system, differences in the structure-function relationships of the two isoforms are additionally relevant for the possibility of therapeutic targeting of one isoform over the other⁷. However, TPH1 and TPH2 have proven difficult to characterize and structural information on especially the regulatory domains of the two enzymes remain sparse.

This work focuses on a structural characterization of the regulatory domain of *human* TPH2 and investigates binding of a ligand (L-Phe and L-Trp) to the domain. Nuclear magnetic resonance (NMR) spectroscopy and cryo-electron microscopy (cryo-EM) are used to shed light on TPH2, complemented by additional methods.

Sequential identities are observed between the regulatory domains of TPH1 and TPH2 with respect to the binding pocket of TPH2 determined in the first study. This leads to an investigation of the regulatory domain of *human* TPH1 using size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) and NMR spectroscopy.

Outline of thesis

The thesis contains seven chapters. Chapter 1 introduces the reader to the tryptophan hydroxylase isoforms and the aromatic amino acid hydroxylase enzyme family and presents the current knowledge on the enzymes relevant for this project. Chapter 2 concerns purification of TPH focusing on the isolated regulatory domain variants of TPH1 and TPH2. Chapter 3 provides a brief theoretical background on the main methods used in this project, which includes NMR spectroscopy, small-angle X-ray scattering, SEC-MALS, and cryo-EM. The most substantial part of the work conducted throughout this project is presented in Chapter 4, which concerns a structural study of ligand binding to the regulatory domain of *human* TPH2. This work is presented in the form of a manuscript (Under preparation). Chapter 5 investigates the oligomeric state of C-terminally truncated *human* TPH1 variants and ligand binding to the regulatory domain using SEC-MALS and NMR spectroscopy. This work is also presented in the form of a manuscript (Under preparation). Chapter 6 concludes on the project and chapter 7 offers future perspectives.

1

TRYPTOPHAN HYDROXYLASE AND THE AROMATIC AMINO ACID HYDROXYLASES

Tryptophan hydroxylase (TPH) belongs to the family of aromatic amino acid hydroxylases (AAAHs), which includes TPH, phenylalanine hydroxylase (PAH), and tyrosine hydroxylase (TH) ⁵. This chapter will briefly introduce the importance of TPH and the AAAHs and discuss the current knowledge on their structure and regulatory mechanisms relevant for this project with emphasis on TPH.

1.1 THE AROMATIC AMINO ACID HYDROXYLASES

TPH, PAH, and TH catalyze the hydroxylation of their respective amino acids, using non-heme Fe(II), molecular oxygen (O₂) and tetrahydrobiopterin (BH₄) as co-factor and co-substrates ⁵. TPH (EC 1.14.16.4) catalyzes the hydroxylation of L-Tryptophan into L-5-hydroxytryptophan (5-HTP). This reaction is the first and rate-limiting step in the two-step biosynthesis of serotonin, an important hormone and neurotransmitter ^{8,9} (Figure 1A). The second step is decarboxylation of 5-HTP by aromatic amino acid decarboxylase ¹⁰. BH₄ is in the reaction

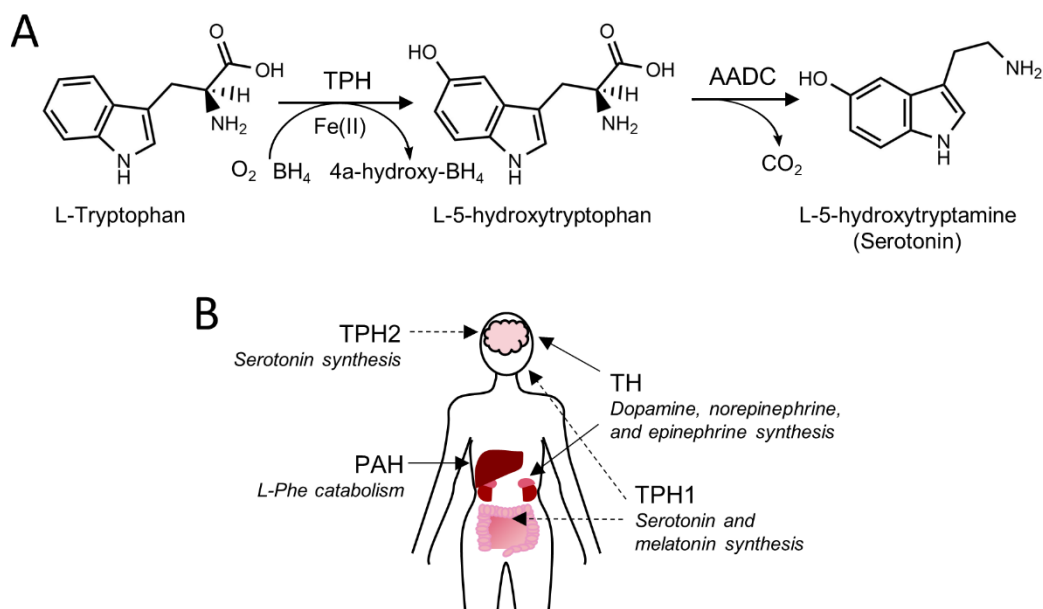


Figure 1 A) The two-step biosynthesis of serotonin (Figure made with ChemDraw 18.1). L-Trp is converted into L-5-hydroxytryptophan by TPH requiring Fe(II) as co-factor and O₂ and BH₄ as co-substrates. The second step is decarboxylation by aromatic amino acid decarboxylase (AADC) to yield serotonin ^{5,10}. B) Schematic illustration of the tissues where the AAAHs are most abundant. PAH is mainly found in the liver ¹², TH in the brain and in the adrenal gland ^{15,30}, TPH1 in the gut and in the pineal gland ⁶, and TPH2 in the brain ⁶.

converted to 4a-hydroxy-BH₄, which is subsequently regenerated back to BH₄ in two steps ¹¹. Serotonin is additionally a pre-cursor for the hormone melatonin predominantly produced in the pineal gland ⁶. There are two TPH isoforms, encoded by two different genes: TPH1, expressed mainly in enterochromaffin cells in the gut and in the pineal gland, and TPH2, expressed in the brain ^{6,9} (Figure 1B). Importantly, the second TPH isoform was discovered in 2003 ⁹, and earlier literature should therefore be read with attention to the source of TPH. As serotonin cannot permeate the blood-brain-barrier (BBB), TPH1 and TPH2 regulate each their spatially separated serotonin system ⁷. PAH (EC 1.14.16.1), located in the liver, is responsible for L-Phe catabolism converting L-Phe into L-Tyr ¹². TH converts L-Tyr into 3,4-dihydroxyphenylalanine (L-Dopa), a precursor in the primary synthesis pathway of the catecholamine neurotransmitters dopamine, norepinephrine, and epinephrine ¹³. TH is expressed in the brain and in the adrenal gland, and alternative splicing of the single TH gene leads to four very similar *human* isoforms only differing in the length of the N-terminal ^{14,15}.

1.1.1 Physiological aromatic amino acid levels

The amino acids L-Phe and L-Trp are not metabolized in the body and are therefore only obtained through the diet, while L-Tyr is also synthesized by PAH ¹⁶. In healthy adults, plasma L-Phe levels have been reported to ~35-85 µM and fluctuating less than 50% during the day ^{16,17}. Plasma L-Tyr levels have been reported to ~50-80 µM and plasma L-Trp levels to ~55-65 µM ¹⁶. In the brain, the levels of the aromatic amino acids depend on transport across the BBB by the LAT1 transporter ^{16,18}. LAT1 transports the large neutral amino acids (LNAAs), L-Tyr, L-Phe, L-Trp, L-Leu, L-Ile, L-Met, L-His, and L-Thr, and is nearly saturated at physiological LNAA levels, making LAT1 transport a competitive process for the LNAAs ^{16,18}. L-Phe binds strongest to LAT1 of the LNAAs ¹⁸. In rats, a study reported brain L-Phe levels to ~13.4 µg/g (80 µmol/kg), L-Tyr levels to ~12 µg/g (70 µmol/kg), and L-Trp levels to ~3 µg/g (15 µmol/kg) ¹⁹.

1.1.2 Serotonin related diseases and TPH as a therapeutic target

Serotonin is a neurotransmitter and hormone affecting a variety of body functions and behavioral processes through interactions with serotonin receptors (5-HT receptors) ²⁰. Central effects include among others aggression, appetite, memory, sexuality, and attention ^{20,21}. Irregular TPH2 function, affecting the central serotonin system, has been connected to multiple neurological disorders, such as depression, schizophrenia, and bipolar disorder ^{1,2,22}. Peripheral serotonin affects among others liver function, hemostasis, gastrointestinal motility, and vasoconstriction ²⁰. Consequently, dysfunction of the peripheral serotonin system is related to diseases as e.g. ulcerative colitis and irritable bowel syndrome ³ and carcinoid syndrome ⁴. The complexity of the serotonergic systems, makes it challenging to design specific drugs for serotonin related diseases without unwanted side-effects. For example, selective serotonin reuptake inhibitors are used as antidepressants (to increase central serotonin levels), however, these are associated with a range of side-effects in the peripheral system including e.g. cardiovascular problems, weight gain, and gastro intestinal distress ²³. A strategy to selectively target either the peripheral or central serotonin system is isoform specific TPH targeting. Currently, one TPH1 inhibitor, telotristat ethyl, has FDA approval for treatment of carcinoid syndrome ²⁴. The inhibitor is not isoform specific, but as it cannot permeate the BBB it only influences TPH1 and the peripheral serotonin levels ²⁴. While new compounds are being designed for TPH1 specific inhibition ^{25,26}, isoform specific targeting

of TPH2 remains more challenging for at least two reasons: i) To target TPH2 the drug has to pass the BBB. ii) In the central system, the therapeutic goal is most frequently an increased serotonin level, as opposed to a decreased serotonin level in the peripheral system ⁷. Thus, the goal is to develop TPH2 specific activators, which is currently greatly hindered by the limited knowledge on TPH2 (and TPH1) regulation and structure-function relationship. Therefore, improved knowledge on both TPH isoforms is important for the possibility of TPH2 directed drug targeting.

1.1.3 Physiological relevance of PAH and TH

Mutations in the PAH gene can lead to dysfunctional PAH variants resulting in reduced or absent PAH activity and thus elevated L-Phe levels, a condition known as phenylketonuria (PKU) ²⁷. In classical PKU, plasma L-Phe levels are increased to above 1200 μM ²⁸. High L-Phe levels are neurotoxic and left untreated it can lead to e.g. mental retardation, epilepsy, and behavioral problems ²⁹. The catecholamine neurotransmitters and hormones dopamine, norepinephrine, and epinephrine, the downstream products after conversion of L-Tyr to L-Dopa by TH, influences brain functions e.g. attention, memory, and cognition ^{30,31}, and reduced levels of dopamine are found in relation to Parkinson's disease ³². Considering the importance of L-Phe catabolism, catecholamine synthesis, and serotonin synthesis, it is natural that extensive research has gone into mapping the structure-function relationships of the AAAHs which are still not fully elucidated.

1.2 STRUCTURE OF THE AAAHs

The homologous AAAHs consist of three different domains: An N-terminal regulatory domain (RD), a catalytic domain (CD), and a C-terminal tetramerization domain (TD, Figure 2A-B) ^{33–36}. Of these domains, the CDs have the highest sequence identity and the RDs the lowest (Figure 2A). Being dynamic multi-domain proteins, the structures of the AAAHs have proven difficult to study. PAH is the best structurally characterized of the enzymes, with structures available of both the full enzyme and multiple truncated variants ^{34,37–42}. A PAH monomer is illustrated in Figure 2B, showing the fold of the individual domains. The AAAHs mainly present as tetramers, formed through coiled-coil interactions of the TDs ^{33–36} (Figure 2C). PAH can additionally form dimers ⁴³, and the presence of an octameric species of TH has been reported in one study ⁴⁴. The conformation of the CDs and TDs are similar in the tetramers of PAH, TH, and TPH2 ^{34,36,45}, and though no tetrameric structure is available of TPH1 the high degree of homology suggests a similar conformation. The positions of the RDs in the tetramers are found to vary among the AAAHs as illustrated in Figure 2D. In PAH, the RDs can be docked in between two CDs as illustrated in Figure 2C or form RD dimers above and below the CD tetramer ^{34,38,41} (Figure 2D). The latter conformation is favored in the presence of L-Phe ³⁸. In the TH tetramer, the RDs form dimers ^{36,46}. The location of the RDs in the tetramers of TPH1 and TPH2 have not been fully determined yet. In the presence of L-Phe (0.6 mM to 9 mM), SAXS data support a TPH2 tetramer model with dimerized RDs ⁴⁷, but currently no data is available on the RDs position in the absence of L-Phe or on the positions of the RDs in the TPH1 tetramer. In the following sections, the structure and function of the CDs and RDs of the AAAHs will be presented in more detail.

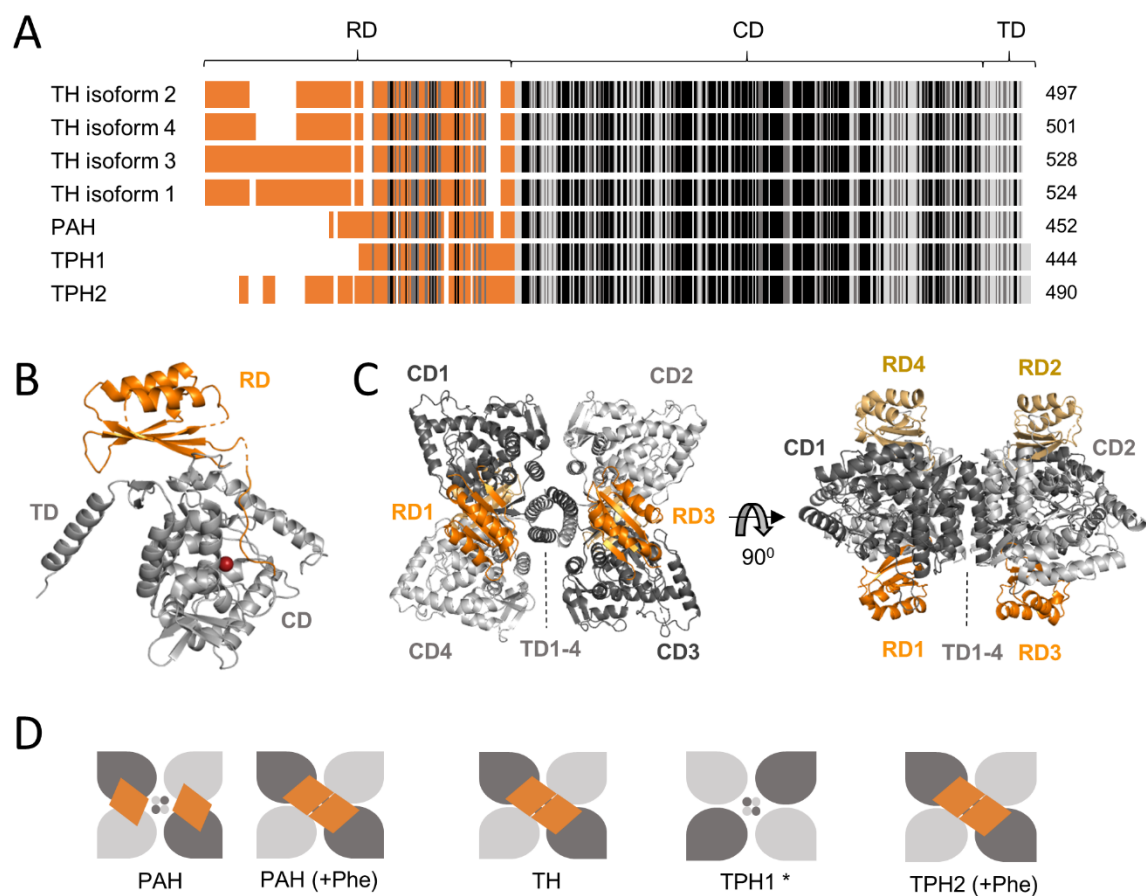


Figure 2 Schematic sequence alignment of the human AAAH isoforms. The RD is colored in orange and the CD and TD are colored in light grey. White spaces indicate gaps, dark grey indicates residues with strongly similar properties, and black indicates identical residues. The alignment was performed with Clustal Omega¹⁵¹ and the figure made with inspiration from¹⁵². The full alignment is included in Appendix A1. **B)** Cartoon illustration of a PAH monomer (PDB ID: 6N1K³⁸). The RD is colored in orange and the CD and TD in light grey. **C)** Cartoon illustration of a PAH tetramer (PDB ID: 5DEN³⁴). The RDs are colored in orange and the CDs and TDs are colored in grey. A dark orange RD is from the same monomer as a dark grey CD. **D)** Schematic illustration of the AAAH tetramers (adapted from Chapter 4). The CDs and TDs are colored in grey, and the RDs are colored in orange. The structure of PAH with L-Phe has been confirmed by X-ray crystallography^{34,38}, and SAXS data supports the conformation with L-Phe^{38,41}. The TH structure has been determined by cryo-EM³⁶. The RDs are not illustrated for TPH1 as there is no structure available of the full-length enzyme. * Additionally there is no structure available of the CD tetramer. The dimeric conformation of the RDs in TPH2 (+L-Phe) is supported by SAXS data⁴⁷.

1.2.1 The catalytic domain

The CDs are as illustrated in Figure 2A the most similar of the domains among the AAAs, and their large degree of homology is also evident from both sequential and structural alignments of the CDs (Figure 3). The CDs are mainly helical and form a cavity where the active site is located ^{45,48–50}. The isolated CDs of TPH1 and TPH2 are monomeric species and are fully active without the RD and TD ^{51–53}.

The iron-atom, positioned around 13 Å from the protein surface, separates the cavity into two binding pockets, one for BH₄ and one for the amino acid substrate as illustrated in Figure 4A based on structures of TPH1 ^{49,54}. The iron atom is bound in a 2-histidine-1-carboxylate facial triad (coordinated to two histidines and one glutamate) and is additionally coordinating to three water molecules in a hexagonal complex, as illustrated in Figure 4B ⁴⁹. Upon binding of both BH₄ and amino acid substrate, it is suggested that Fe(II) forms a bidentate coordination to glutamate and excludes two water molecules, leaving space for O₂ to bind ⁵⁵.

The binding pocket of BH₄ can be visualized from a structure of *human* TPH1 with BH₄ analogue BH₂ bound ⁴⁹ (Figure 4C). Crystal structures of PAH have shown similar binding of BH₄ and BH₂, allowing extrapolation of the BH₂ binding pocket ³⁷. In the TPH1 structure, BH₂ is kept in close proximity of the iron by hydrogen bonding to two of the iron-coordinated water molecules, and a water bridge to E273 (Figure 4C). BH₂ is additionally forming ionic interactions with the backbone of G234 and L236, and is π -stacking with Y235 and F241. All of these residues are conserved in *human* TPH2, see Appendix A2.

The binding pocket of L-Trp was determined from a structure of *chicken* TPH1 (91% sequence identity to *human* TPH1), revealing that L-Trp is surrounded by residues Y235, R257, T265, P268, H272, F313, F318, S336, and I366 and that L-Trp forms polar interactions with R257, T265, S336, and I366 ⁵⁴ (Figure 4D). All of these residues are conserved between *human* TPH1, *chicken* TPH1 and *human* TPH2, see Appendix A2.

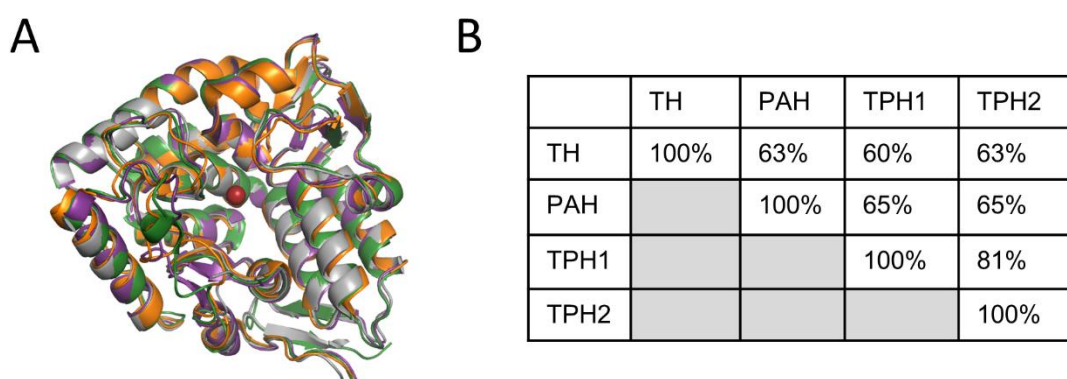


Figure 3 A) Structural alignment of the CDs of the four human AAAs. PAH (PDB ID: 6HPO ⁴⁸, green), TPH2 (PDB ID: 4V06 ⁴⁵, purple), TPH1 (PDB ID: 1MLW ⁴⁹, orange), TH (PDB ID: 2XSN ⁵⁰, grey). The iron atom located in the active site is shown as a red sphere. B) Sequence identities of the CDs of the human AAAs from alignment with clustal Omega ¹⁵¹. The following residues were included in the alignment: TH isoform 3 residues 189–502, PAH residues 113–426, TPH1 residues 100–413, TPH2 residues 146–459. Only TH isoform 3 was included in the alignment as the CDs of the four human TH isoforms are identical.

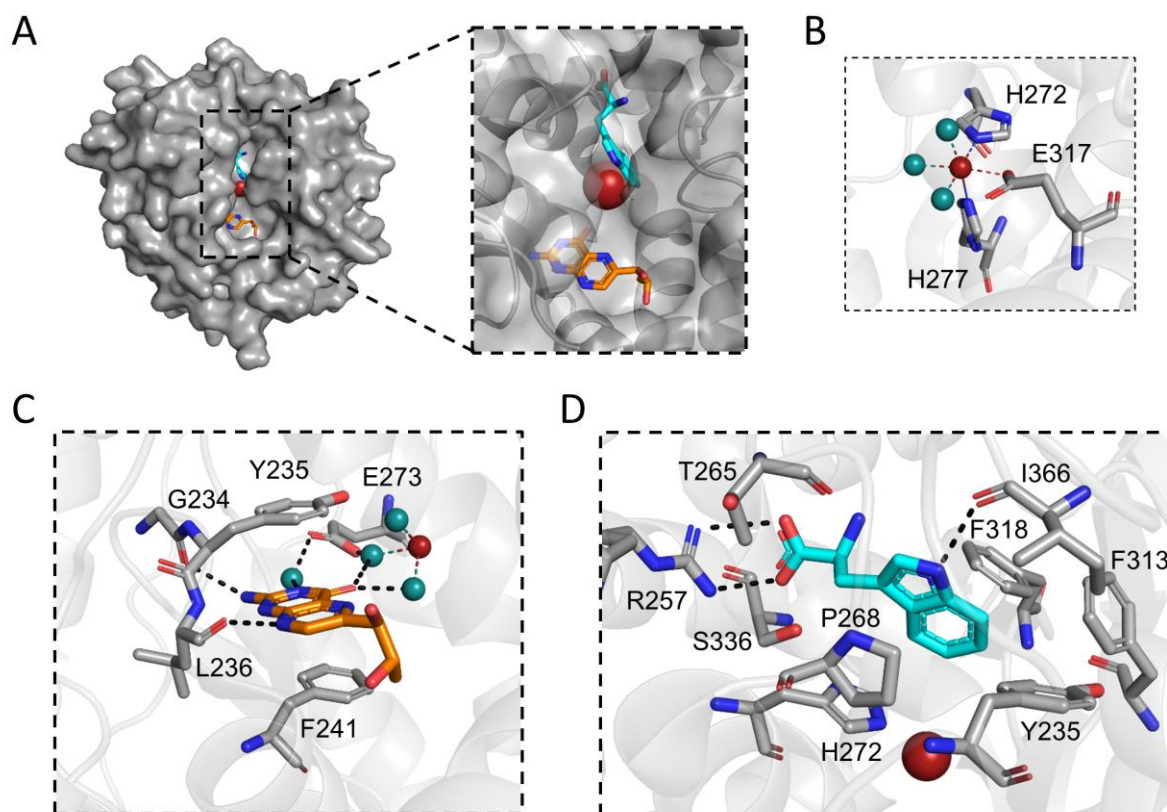


Figure 4 A) Surface representation of the CD of human TPH1 (PDB ID: 1MLW⁴⁹), with co-substrate analogue BH₂ shown as an orange stick, Fe(III) as a red sphere, and L-Trp (superimposed from the crystal structure of chicken TPH1 (PDB ID: 3E2T⁵⁴) shown as cyan stick. A zoom in on the active site cleft is shown to the right with a transparent surface rendering. B) Illustration of the coordination sphere of iron in the crystal structure of human TPH1 (PDB ID: 1MLW⁴⁹). Fe(III) is shown as a red sphere and water molecules as blue spheres. C) Location of the BH₄ binding pocket in the structure of human TPH1 (PDB ID: 1MLW⁴⁹). Co-substrate analogue BH₂ is shown as orange stick, Fe(III) as a red sphere, and water molecules as blue spheres. D) Binding pocket of L-Trp in chicken TPH1 (PDB ID: 3E2T⁵⁴) L-Trp is shown as a cyan stick and Fe(III) as a red sphere. The indicated residue numbers are those of the conserved residues in human TPH1, all differing by one integer. I.e. I366 in human TPH1 is I367 in chicken TPH1.

1.2.1.1 Kinetic parameters and substrate specificity of TPH1 and TPH2

Despite a large degree of identity between the BH₄ and L-Trp binding pockets of TPH1 and TPH2, differences are observed in the kinetic parameters of the two isoforms (Summarized in Table 1): $k_m(\text{BH}_4)$ is found to be larger for TPH1 than for TPH2, while the V_{max} for TPH1 has been reported to be both larger and smaller than that of TPH2^{51,56,57}. The isolated CD of TPH1 is found to be substrate inhibited by L-Trp^{51,56}, while substrate inhibition is generally not observed for the isolated CD of TPH2^{33,51,56,58}. However, there are a few reports of substrate inhibition of full-length TPH2 with large K_i constants (420-970 μM)^{57,59}. The difference in substrate inhibition has been attributed to a loop lining the active site, as a study indicated that the loop specifically in TPH1 closes upon L-Trp binding and blocks the entrance to the BH₄ binding pocket⁵⁶. Differences are also observed in the amino acid substrate specificity of TPH1 and TPH2. Due to the high degree of homology in the CDs of the AAHs, they are not fully specific towards their natural amino acid substrates. For MBP-tagged *human* TPH1 and TPH2 variants, McKinney et al. found that TPH1 had a K_m (Trp) of $10 \pm 12 \mu\text{M}$ and a K_m (Phe) of $34 \pm 8 \mu\text{M}$ with twice the catalytic efficiency with L-Trp as substrate compared to L-Phe⁵⁷.

For TPH2 they determined a K_m (Trp) of $43 \pm 4 \mu\text{M}$, a K_m (Phe) of $284 \pm 38 \mu\text{M}$, and four times the catalytic efficiency with L-Trp as substrate compared to PAH⁵⁷. Thus indicating that TPH2 is more selective than TPH1 with respect to L-Trp⁵⁷. A similar degree of specificity for L-Trp compared to L-Phe was observed by Moran et al. for the isolated CD of *rabbit* TPH1⁵³. L-Tyr has been reported as a poor substrate for TPH1, and thus likely also for TPH2, and is converted at least 5000 times slower than L-Trp and L-Phe⁵³. The residue F313 (TPH1) in the tryptophan binding pocket has been shown to be important for substrate specificity, and mutation to tryptophan, the corresponding residue in PAH, increases the preference for L-Phe as substrate⁶⁰. Mice lacking the ability to produce both TPH1 and TPH2, have been shown to still have 10 % of the normal serotonin level in the blood, which has been attributed to the residual ability of PAH to hydroxylate L-Trp⁶¹.

Table 1 Kinetic parameters of TPH1 and TPH2 from a selected subset of studies. Only studies where both the parameters of TPH1 and TPH2 were determined has been included, to enable easier comparison between the isoforms.

Study		TPH1	TPH2
Windahl et al., isolated CDs of <i>human</i> TPH1 and TPH2 ⁵¹	K_m (Trp) K_m (BH ₄) V_{max} (Trp) K_i (Trp)	$22.8 \pm 0.9 \mu\text{M}$ $315 \pm 37 \mu\text{M}$ $201 \pm 11 \mu\text{M}/\text{min}$ $72 \pm 7 \mu\text{M}$	$15 \pm 0.4 \mu\text{M}$ $26.5 \pm 1.3 \mu\text{M}$ $151 \pm 1 \mu\text{M}/\text{min}$ N/A
Tidemand et al., isolated CDs of <i>human</i> TPH1 and TPH2 ⁵⁶	K_m (Trp) K_m (BH ₄) k_{cat} (Trp) K_i (Trp)	$12.7 \pm 1.4 \mu\text{M}$ $285 \pm 15 \mu\text{M}$ $108 \pm 10 \text{ min}^{-1}$ $165 \pm 18 \mu\text{M}$	$13.1 \pm 1.3 \mu\text{M}$ $26.7 \pm 0.9 \mu\text{M}$ $258 \pm 20 \text{ min}^{-1}$ N/A
McKinney et al., MBP-tagged full-length <i>human</i> TPH1 and TPH2 ⁵⁷	K_m (Trp) K_m (BH ₄) V_{max} (Trp) K_m (Phe) V_{max} (Phe)	$10.6 \pm 1.2 \mu\text{M}$ $27 \pm 3 \mu\text{M}$ 651 ± 26 $34.4 \pm 7.8 \mu\text{M}$ $1214 \pm 107 \text{ nmol}/\text{min}/\text{mg}$	$43.4 \pm 3.6 \mu\text{M}$ $12.8 \pm 1.1 \mu\text{M}$ $202 \pm 8.9 \text{ nmol}/\text{min}/\text{mg}$ $284 \pm 38 \mu\text{M}$ $316 \pm 22 \text{ nmol}/\text{min}/\text{mg}$
McKinney et al., full-length <i>human</i> TPH1 and TPH2 ⁵⁷	K_m (Trp) K_m (BH ₄) V_{max} (Trp) K_i (Trp)	$22.8 \pm 3.9 \mu\text{M}$ $39 \pm 8 \mu\text{M}$ $646 \pm 43 \text{ nmol}/\text{min}/\text{mg}$ $385 \pm 70 \mu\text{M}$	$40.3 \pm 3.2 \mu\text{M}$ $20.2 \pm 1.5 \mu\text{M}$ $194 \pm 7.2 \text{ nmol}/\text{min}/\text{mg}$ $970 \pm 328 \mu\text{M}$

1.2.2 The regulatory domain

The RD is the least studied domain of both TPH1 and TPH2, which is linked to low expression levels and instability of TPH variants containing the domain ^{62,63}. Currently, no structures are available in the protein data bank of the RD of either TPH1 or TPH2. One crystal structure is available of the isolated RD of *human* PAH (PDB ID: 5FII ³⁹), while an NMR structure is available of the isolated RD of *rat* TH (PDB ID: 2MDA ⁶⁴). The RD of PAH and TH both consist of an ordered core of around 80-90 residues, preceded by a mainly flexible N-terminal tail (N-tail) of different lengths, of which the N-tail is the most divergent ^{34,36,39,64} (Figure 5). The RD core of PAH is an ACT domain (Aspartate kinase, chorismate mutase, Tyr A domain), a domain family originally identified by use of PSI-BLAST ⁶⁵. ACT domains are characterized by a ferredoxin like fold ($\alpha\beta$ - sandwich), with $\beta\alpha\beta\alpha\beta$ topology, and the ability to bind a ligand ⁶⁵. Multiple enzymes containing the ACT domain are involved in amino acid metabolism and are regulated by their ACT domain ligand ⁶⁶. This includes PAH, which is allosterically regulated by L-Phe binding to the ACT dimer ^{39,41}. Another example is, *E. Coli* D-3-Phosphoglycerate dehydrogenase which is involved in the synthesis of L-Ser, and is allosterically inhibited by the amino acid binding to its ACT domain dimer ⁶⁷. The symmetric RD dimer of PAH forms an 8-stranded β -sheet and additionally forms two identical binding pockets at the dimer interface, where L-Phe binds (Figure 6A) ³⁹. In addition to the side-by-side dimer association of the ACT domain as seen in PAH, ACT domains in other enzymes are observed to form face-to-face ACT dimers ⁶⁸, ACT trimers ⁶⁹, and tetramers ^{70,71}. Though the RD of TH forms a dimer and resembles the ACT domain, there are notable differences (Figure 6B) ⁶⁶. First, the topology is slightly different with a $\beta\beta\alpha\beta\alpha$ fold ⁶⁴, and secondly, the RD of TH is found not be thermally stabilized by L-Tyr indicating that the domain does not bind the amino acid ³⁹. Therefore, the RD of TH is only reported as a “putative ACT domain” ⁶⁶. The sequential identities of the RD cores of TPH1 and TPH2 compared to PAH and TH provided in Figure 5B, indicate that both TPH1 and TPH2 are mostly resembling PAH, and sequence analysis has proposed that TPH1 and TPH2 contain the ACT domain ^{65,72}. Importantly, the N-tail of TPH2 is 46 residues longer than that of TPH1 which is expected to contain ~17 residues (Figure 5A).

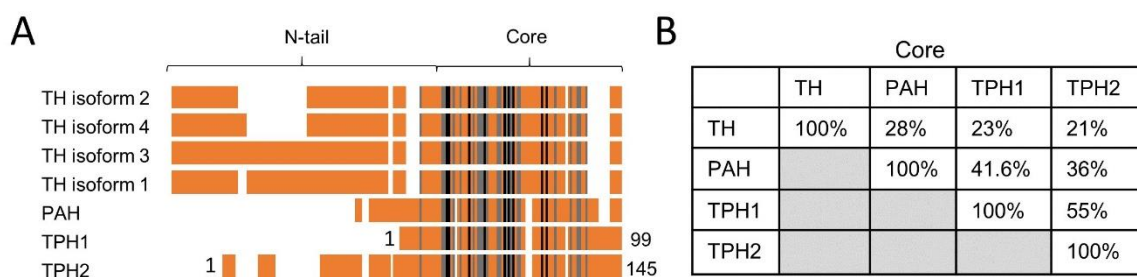


Figure 5 A) Schematic sequence alignment of the RDs of the AAAH. White spaces indicate gaps, dark grey indicates residues with strongly similar properties, and black indicates identical residues. The alignment was performed with clustal Omega ¹⁵¹. B) Sequence identities among the core of the RDs of the AAAHs, determined from an alignment with clustal Omega ¹⁵¹. The following sequences were used: human TPH2 (residues 64-145), human TPH1 (residues 18-99), human PAH (residues 35-112), human TH isoform 3 (residues 113-188). The TH isoforms only differ in the length of the N-tail, and therefore only one TH isoform was used in the alignment of the RD cores.

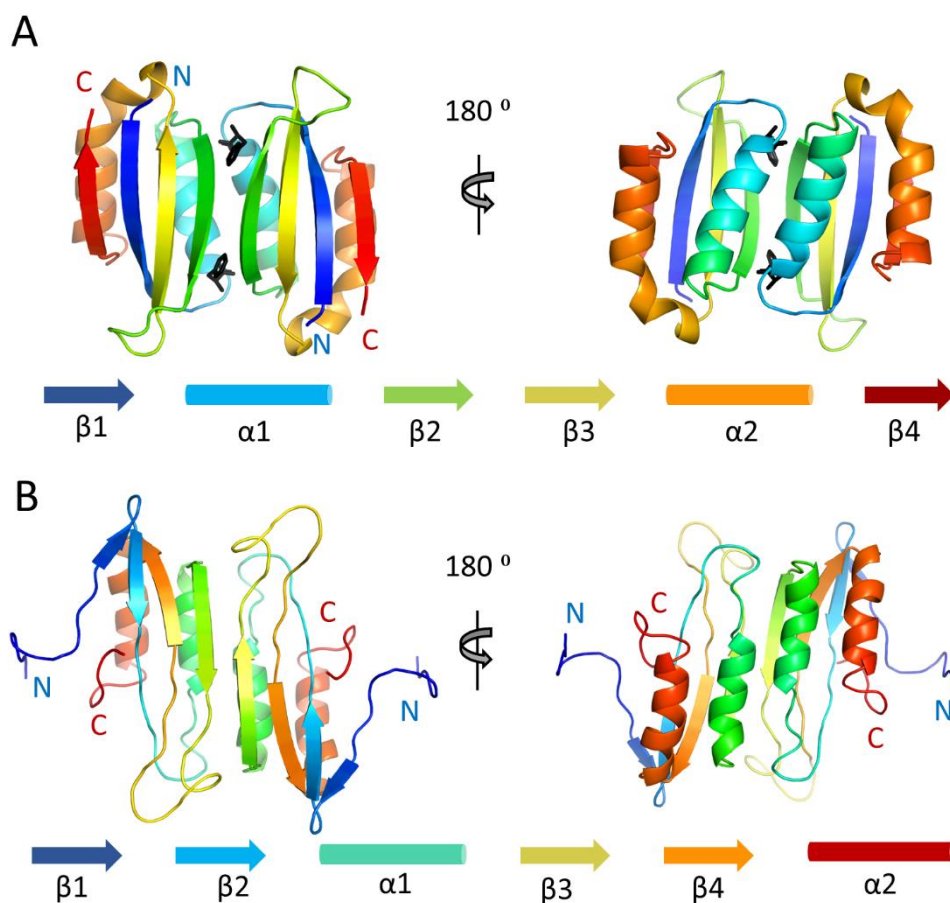
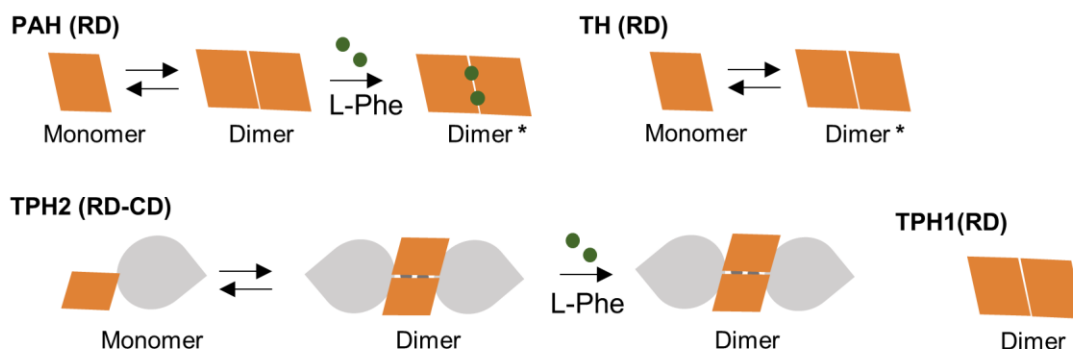


Figure 6 A) Crystal structure of the RD core dimer of human PAH with L-Phe (PDB ID: 5FII³⁹). The two L-Phe molecules (black sticks) are bound in identical binding pockets at the symmetric RD dimer interface. The topology of the domain is illustrated below the structure. B) NMR conformer of the RD core dimer of rat TH (PDB ID: 2MDA⁶⁴). The topology of the domain is illustrated below the structure.

The oligomeric states of the isolated RDs are similar among the AAAHs, while it varies if the RDs have been observed to bind a ligand as mentioned for PAH and TH above. The isolated RD of PAH has been reported to exist in a monomer-dimer equilibrium (with a $K_d \sim 46 \mu\text{M}$), while addition of L-Phe induces dimer formation^{39,73} (Figure 7). The isolated RD of TH has been reported to be dimeric at concentrations down to $5 \mu\text{M}$ ⁶⁴. Additionally a study indicated the presence of a monomer at lower protein concentrations, suggesting the existence of a monomer-dimer equilibrium⁷⁴. Furthermore, the same study reported that truncating the first 64 residues of the TH N-tail resulted in a monomer⁷⁴. The isolated RD of TPH1 is reported to be dimeric independent of L-Phe (investigated in the range of $2.7\text{-}35 \mu\text{M}$ TPH1)⁷⁵. Additionally Patel et al. observed using DSF that L-Trp did not affect the stability of the RD of TPH1³⁹. A study on a truncated variant of TPH2 containing both the RD and CD, showed that the variant exists in a monomer-dimer equilibrium with the monomeric species mainly being observed in the protein concentration range of $2\text{-}10 \mu\text{M}$ ⁵². Addition of 3 mM L-Phe shifted this equilibrium towards the dimeric form⁵². Additionally, L-Phe and L-Trp (10 mM) were shown to increase thermal stability of the variant with $\sim 3^\circ\text{C}$ and $\sim 6^\circ\text{C}$, respectively, while D-Phe neither affected the stability or induced dimerization of the variant⁵². Additionally, dimerization and stabilization by L-Phe and L-Trp were not observed for the isolated CD suggesting the presence of an RD binding site in TPH2⁵². A hydrogen/deuterium exchange mass spectrometry study of the TPH2 variant (combined RD and CD) reported decreased deuterium uptake values for the majority of the peptides in the RD in the presence of L-Phe, supporting an RD binding site⁴⁷.



*Figure 7 Schematic illustration of the oligomeric states of the isolated RDs of PAH^{39,73}, TH^{64,74}, TPH1⁷⁵, and the N-terminally truncated variant of TPH2 containing the RD and CD⁵². * A crystal structure is available of the RD dimer of PAH with L-Phe³⁹, and an NMR structure is available of the RD dimer of TH⁶⁴. There are no structures available of the TPH1 and TPH2 variants in the protein data bank, and they are therefore illustrated based on their expected conformations extrapolating from PAH and TH. SAXS data on the TPH2 variant with L-Phe has indicated a flexible dimeric conformation⁴⁷. The structures of the PAH RD monomer and dimer without L-Phe and the TH RD monomer have additionally not been determined.*

1.3 REGULATORY MECHANISMS OF THE AAAH

While the regulatory mechanisms of TPH1 and TPH2 are not completely mapped, the regulatory mechanisms of PAH and TH are better structurally characterized. Importantly, the differences in the N-tail between the two enzymes and in their ability to bind an RD ligand are reflected in the specific mechanisms regulating their activity, which will be briefly outlined in the following.

1.3.1 Regulatory mechanism of PAH

PAH is known to be allosterically activated by L-Phe⁷⁶. However, earlier studies debated the origin of the cooperative activation and the location of the allosteric L-Phe binding site. L-Phe activation was previously proposed to result from cooperative binding in the active site^{77,78}, or by binding in an allosteric site at the interface between the RD and CD^{40,79}. Currently, substantial structural studies support that the location of the allosteric site(s) is at the interface of two RD dimers^{38,39,41,80}, as in the structure in Figure 6A. In this model, the allosteric activation by L-Phe is accompanied by large structural changes in the tetrameric enzyme⁴¹ (Figure 8): In the absence of L-Phe, PAH is mostly populating a resting state (RS), where the RDs are docked to the CD of an adjacent monomer⁴¹. In this state, the N-tail of PAH (residues 20-25), lie across the opening to the active site hindering substrate access³⁴. To form the allosteric binding site, it is suggested that the RDs undock from the CDs and dimerize with the RD of a diagonal monomer, creating the dimer interface where 2 L-Phe molecules bind and stabilize the active state (AS)^{41,80}. A molecular dynamics simulation study on the RD PAH dimer has indicated a conformational selection mechanism for L-Phe binding⁸¹. Once the RDs are dimerized, the N-tail no longer occlude the active site, facilitating substrate access. Another study suggested that formation of the AS tetramer first requires the dissociation of the RS PAH tetramer into a dimer, before re-association to form the AS⁸². Removal of the N-terminal 29 amino acids, results in a variant that is no longer allosterically activated by L-Phe, supporting the role of the N-terminal in PAH regulation⁸³. Notably, a similar allosteric mechanism involving large structural changes upon dimerization of the ACT domain has been observed in DAH7PS, another enzyme of the ACT domain family⁶⁸. The AS-RS equilibrium of PAH is additionally found to be affected by phosphorylation of Ser16 in the N-tail⁸⁴, a site phosphorylated by cAMP-dependent protein kinase A (PKA)⁸⁵. In agreement with crystal structures of non-phosphorylated and phosphorylated PAH⁴⁰, the phosphorylation is

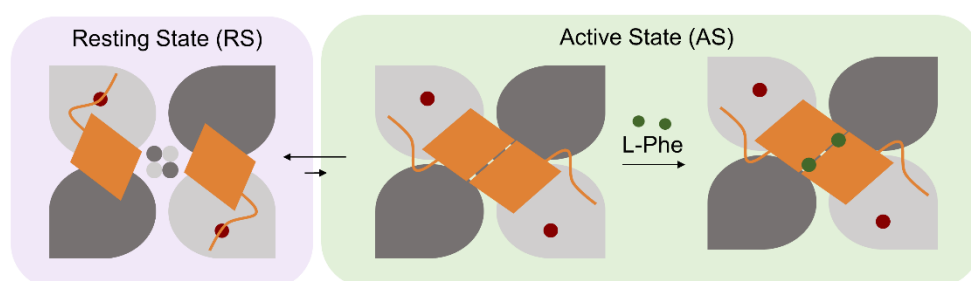


Figure 8 Allosteric mechanism of PAH as described by Meisburger et al.⁴¹. In absence of L-Phe, the RDs (orange) are mainly docked to the CDs (grey) allowing the N-tail to cover the active site (red circle). Upon dimerization of the RDs the N-tail consequently move away from the active site, resulting in an active state stabilized by L-Phe binding at the RD dimer interface⁴¹. Made with inspiration from⁴¹ and adapted from Chapter 4.

proposed to cause very local structural effects in the N-tail, that weakens the interaction of the N-tail with the CD, facilitating the conformational change of the RDs ⁸⁴. Thus, phosphorylation is suggested to increase the population of the active state in absence of L-Phe, thereby lowering the amount of L-Phe needed to activate PAH ⁸⁴. A cation- π interaction involving residues Phe80 from the RD and Arg123 and Arg420 from the CD has been shown to be important for formation of the RS, and mutations of Phe80 alters the equilibrium between RS, AS, and the intermediate “undocked states” where the RDs are neither docked to the CDs or forming dimers ⁸⁰. The study indicated that undocked states were equally active to the AS ⁸⁰.

1.3.1.1 Regulatory mechanism of TH

TH is feedback-inhibited by catecholamines (dopamine, epinephrine, and norepinephrine) which compete with binding of BH₄ in the active site ^{86,87}. This inhibition is correlated with phosphorylation of the N-tail, which is longer compared to the N-tail of PAH, and therefore can interact with the active site regardless of the dimerized RDs in TH ³⁶. In unphosphorylated form, the N-tail interacts with the active site if a catecholamine is bound, strengthening the inactive TH-catecholamine complex ^{36,87} (Figure 9). Upon phosphorylation of Ser40 (*human* TH isoform 2 /TH type 1 notation), it is suggested that the affinity of the N-tail with the active site decreases, thereby detaching the N-tail and lowering the K_d of the TH-catecholamine complex and thus minimizing the inhibitory effect ^{36,87,88}. While studies agree on the lower binding affinity of catecholamines upon Ser40 phosphorylation, the magnitude of the effect has been described differently with reported decreases in binding affinity for dopamine of 10-300 times that of unphosphorylated TH ^{87,89}. Bueno-Carrasco et al. provided the full-length structure of unphosphorylated *human* TH isoform 2 with and without dopamine bound, revealing that a helical part of the N-tail (residues 39-58) forms interactions with both dopamine and active site residues ³⁶. No density was observed for the N-tail without dopamine bound indicating that the N-tail is flexible and not forming interactions with the CDs without a catecholamine bound in the active site ³⁶. Besides phosphorylation at Ser40 by multiple kinases including PKA or Ca²⁺/Calmodulin-dependent protein kinase II (CaMKII) ³⁰, TH isoform 2 is also phosphorylated at Ser31 ⁹⁰, and at Ser19 which has been found to increase the rate of phosphorylation at Ser40 ⁹¹. Additionally, phosphorylation of Ser19 increases binding to 14-3-3 protein ⁹². 14-3-3 proteins are a family of highly conserved proteins expressed in several different tissues including the brain ⁹³. The proteins are around 30 kDa in size, are dimeric,

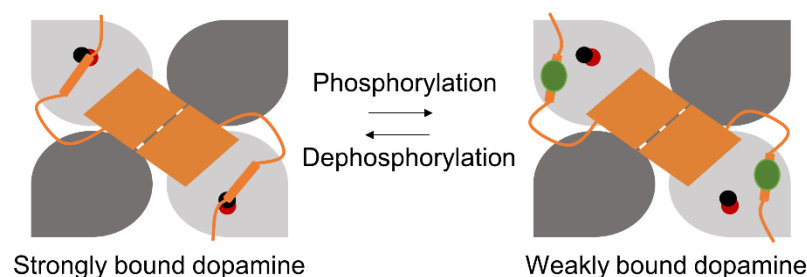


Figure 9 Schematic illustration of the effect of Ser40 phosphorylation in TH, as described by ³⁶ and adapted from Chapter 4. In unphosphorylated form, the N-tail is docked to the active site (red circle) when dopamine (black circle) is bound. Upon phosphorylation of Ser40 (green oval), the tail is released and the affinity of dopamine with TH decreases ³⁶.

and are known to bind to many different, mainly phosphorylated, proteins ^{93,94}. A crystal structure is available of the N-terminal residues of TH (1-43) in complex with 14-3-3 γ ⁹⁵.

1.3.2 Regulatory mechanisms of TPH1 and TPH2

Both TPH1 and TPH2 are known to be phosphorylated by PKA and CaMKII ^{96,97}. TPH1 is phosphorylated at Ser58 (located in the ACT domain) which has been suggested to increase activity in presence of 14-3-3 proteins ⁹⁷. Additionally, PKA and CaMKII has been shown to phosphorylate TPH1 at Ser260 (located in the CD) ⁹⁸. TPH2 is phosphorylated at Ser19 (isoform specific site located in the N-tail) by PKA and CaMKII, and at Ser104 (corresponding to Ser58 in TPH1) by PKA ^{57,96,99}. Winge et al. found that phosphorylation at Ser104 likely increases the rate of phosphorylation at Ser19 ⁹⁶. Ser19 phosphorylation is reported to have modest effects on the catalytic turnover of TPH2, but greatly increases the binding of TPH2 to 14-3-3 proteins ⁹⁶. 14-3-3 binding to TPH2 both increases activity and stability of the enzyme ⁹⁶, but whether the activity increase is due to structural effects, as in the case of TH and PAH, or mainly a result of increased stability remains unclear. Truncation of the N-terminal 44 residues in TPH2 results in a variant with increased activity ¹⁰⁰. Importantly, the limited information available on the regulatory domains of TPH1 and TPH2 and their structure-function relationships exposes the need for further studies to unravel regulation of the enzymes.

2 PURIFICATION OF *HUMAN* TRYPTOPHAN HYDROXYLASE VARIANTS

2.1 INTRODUCTION

As introduced in Chapter 1 the RDs of TPH1 and TPH2 are the least characterized of the domains, which has been connected to low stability of the RDs. Protein characterization requires sufficient amounts of pure protein and as such an efficient purification strategy is required. The research group led by Assoc. Professor Hans E. M. Christensen, has previously developed successful purification strategies for different TPH1 and TPH2 variants ^{51,101–104}. One strategy, used in this project, relies on using MBP-tags and a two-step purification procedure comprising an affinity chromatography (AC) step with on-column cleavage by MBP-tagged *human* rhinovirus 3C protease (MBP-3CP) and a subsequent gel filtration (GF) step. The method allows the use of high salt concentration in the buffers, which has been found to increase stability of TPH variants containing the RD, and which is not possible when using e.g. ion exchange chromatography ¹⁰³. Additionally, the MBP-tag has been found to increase expression levels and solubility of full-length TPH1 and TPH2 ^{57,62}. Including L-Phe in the purification buffer has been found to increase the yield of TPH2 variants containing the RD ⁵². The development of the procedure has been described in more detail elsewhere ^{47,104}.

In this chapter, the nomenclature and details of the studied TPH variants are presented, the purification procedure is introduced, and the purifications of the isolated RDs of TPH2 and TPH1 are outlined and discussed. As purification of TPH variants containing the CD have been described in detail previously, it will not be presented in detail here, but examples of chromatograms and SDS-PAGES of all the variants studied throughout the project are provided in Appendix B. In total, more than 50 purifications were performed during the project. Laboratory technician David F. Nielsen kindly assisted in part of the protein expressions and in purifying TPH1-C and Master student Maja van Hees assisted in part of the TPH1-R and TPH1-RC purifications.

2.2 NOMENCLATURE AND PROTEIN DETAILS

Throughout this dissertation, abbreviations will be used when referring to the different TPH variants. All the studied variants in the project are *human* TPH, and the organism

will therefore not be included in the notation. Variants with N-terminal truncations, will be noted as “NΔx-TPH” where x is the number of removed residues. In case of variants only containing some of the domains, the included domains will be indicated after TPH with the letters “R”, “C”, and “T”, referring to the RD, CD, and TD, respectively. E.g. TPH2-CT consists of the CD and TD. When referring to the MBP-tagged protein (before cleavage) “MBP” will be added before the TPH variant, e.g. MBP-TPH2. Upon cleavage of the tag by 3CP, 2 residues from the recognition sequence (LEVLFQGP) are left on the N-terminal (GP). Details of the studied TPH variants (included residues, molecular weight, extinction coefficient, and abbreviations) can be found in Table 2. Additionally, the full sequences of each variant are provided in Appendix C. All variants in Table 2 have been used for experiments during the project, however, not all variants are included in the results presented in this thesis.

Table 2 Protein details of the studied TPH variants and their abbreviation used in the project. ϵ_{280} values were estimated based on the protein sequence using ProtParam from the ExPASy server ¹⁰⁵.

Abbreviation	Domains	Isoform	N-terminal residues from affinity tag	Residues	Mw [Da]	ϵ_{280} [M ⁻¹ cm ⁻¹]
TPH2	RD+CD+TD	TPH2	GP	2-490	56080	53290
NΔ47-TPH2	RD+CD+TD	TPH2	GP	48-490	50940	46300
TPH2-CT	CD+TD	TPH2	GP	146-490	39691	40800
TPH2-R	RD	TPH2	GP	2-145	16407	12490
NΔ47-TPH2-R	RD	TPH2	GP	48-145	11267	5500
NΔ47-TPH2-RC	RD+CD	TPH2	GP	48-459	47434	44810
TPH1-R	RD	TPH1	GP	2-99	11623	0
TPH1-RC	RD+CD	TPH1	GPGS	2-413	47816	33350
TPH1-C	CD	TPH1	GP	100-413	36220	33350

2.3 EXPERIMENTAL PROCEDURES

2.3.1 Materials and equipment

Buffers and solutions were prepared using water from an 18.2 MΩcm Sartorius Arium® Pro VF Ultrapure Water system (Sartorius), and filtered through a 0.22 μm filter after preparation. Purifications were performed using an Akta start (GE Healthcare) or Akta Explorer (GE Healthcare) chromatographic system.

2.3.2 Expression of MBP-tagged protein

An LB agar plate with 30 µg/mL kanamycin (30 Kan) was streaked with a frozen glycerol stock of *E. Coli* BL21(DE3) cells containing the relevant expression plasmid (pET26 vector) and incubated overnight at 37 °C. A single colony from the plate was added to 50 mL LB 30 Kan and transferred to a 250 mL shaking flask. The preculture was incubated at 37 °C and 140 rpm shaking until an optical density at 600 nm (OD_{600}) of 0.6-1.0. The cells were then pelleted by centrifugation (3000 rpm, 10 min, 4 °C) and resuspended in 50 mL fresh LB 30 Kan. 650 mL LB 30 Kan in a 2 L shaking flask was then inoculated with 6.5 mL of the resuspended cells and incubated at 30 °C and 140 rpm. For expression of MBP-3CP, isopropyl- β -thiogalactopyranoside (IPTG) was added to a final concentration of 0.1 mM at an OD_{600} of 0.5-0.6, and the incubation continued at 30 °C for 16-18 h. For MBP-TPH variants, the temperature was set to 20 °C and the shaking to 240 rpm at an OD_{600} of 0.3. The incubation was continued until an OD_{600} of 0.5-0.6 and expression was then induced with IPTG and the incubation continued for 16-18h (20 °C, 240 rpm). In case the TPH variant contained the CD, $(NH_4)_2Fe(II)SO_4$ was added to a final concentration of 0.2 mM Fe(II) at the time of IPTG induction. The cells were harvested by centrifugation (4000 rpm, 15 min, 4 °C), re-suspended in 25 mL purification buffer (Table 3), and transferred to a 50 mL falcontube. The cells were pelleted (4000 rpm, 15 min, 4 °C), the supernatant discarded, and the cells stored at -80 °C until purification.

2.3.3 Purification of TPH variants

Frozen cells from 650 mL culture were thawed and resuspended in specific binding buffer (Table 3) and lysed by sonication (80% amplitude, 1 cycle, Labsonic P sonicator, Sartorius) for 3x30 s on ice water with at least 1 min cooling in between. After centrifugation (2x20 min, 16000 g, 4 °C), the supernatant was recovered and filtered through a 0.45 µm syringe filter (Whatman™). The sample was loaded onto a pre-equilibrated 16/11 Dextrin Sepharose affinity column (GE Healthcare) with a flow rate of 5 mL/min. The flow through was eluted with 5-6 column volumes (CVs) of binding buffer, and 30 mL of MBP-3CP (diluted from a stock to 2.2 µM in binding buffer) was then loaded onto the column before the system was paused for 1 h to ensure sufficient proteolytic cleavage. The TPH variant was then eluted with binding buffer, and ~8 mL protein was collected. MBP, MBP-3CP, and residual MBP-TPH were eluted from the column with elution buffer (binding buffer + 1 mM maltose). For variants containing the CD, the protein concentration was measured as described in section 2.3.5 and the sample incubated with 1.5x molar equivalent of $(NH_4)_2Fe(SO_4)_2$ for 15 min. The protein was then filtered through a 0.45 µm syringe filter (PALL) and loaded onto to a pre-equilibrated HiLoad 26/600 Superdex 75 prep grade column (or Superdex 200 dependent on the variant, GE Healthcare) with binding buffer and a flow rate of 2.6 mL/min. The protein was collected from the peak eluting at the expected elution volume, which was verified by SDS-PAGE analysis. The concentration was measured, and if needed the sample volume was reduced by centrifugation (4000 g, 4 °C) using a Vivaspin Turbo centrifugal concentrator (Sartorius) with either a 3, 10, or 30 kDa membrane dependent on the protein variant. The protein was frozen in liquid N₂ in relevant sized aliquots and stored at -80 °C.

Table 3 The most frequently used buffers for purification of TPH variants in this project. Despite the previously observed improve in yield when including L-Phe in TPH2 buffers, L-Phe was omitted from the buffer in most cases to ensure it did not affect later experiments when studying the difference of TPH2 variants +/- L-Phe. DTT: dithiothreitol.

TPH variant	Binding buffer(s)
TPH2	20 mM HEPES, 300 mM (NH ₄) ₂ SO ₄ , pH 7.0 (Adjusted with NH ₄ OH)
NΔ47-TPH2	20 mM HEPES, 300 mM (NH ₄) ₂ SO ₄ , pH 7.0 (Adjusted with NH ₄ OH)
TPH2-CT	20 mM HEPES, 100 mM (NH ₄) ₂ SO ₄ , pH 7.0 (Adjusted with NH ₄ OH)
TPH2-R	12 mM Na ₂ HPO ₄ /8 mM NaH ₂ PO ₄ , 100 mM (NH ₄) ₂ SO ₄ , pH 7.0 (Adjusted with NH ₄ OH)
NΔ47-TPH2-R	12 mM Na ₂ HPO ₄ /8 mM NaH ₂ PO ₄ , 100 mM (NH ₄) ₂ SO ₄ , pH 7.0 (Adjusted with NH ₄ OH) 12 mM Na ₂ HPO ₄ /8 mM NaH ₂ PO ₄ , 100 mM (NH ₄) ₂ SO ₄ , 10 mM L-Phe, pH 7.0 (Adjusted with NH ₄ OH)
NΔ47-TPH2-RC	20 mM HEPES, 300 mM (NH ₄) ₂ SO ₄ , pH 7.0 (Adjusted with NH ₄ OH)
TPH1-R	20 mM Tris, 100 mM (NH ₄) ₂ SO ₄ , 5 mM DTT, pH 8.0 (Adjusted with HCl)
TPH1-RC	20 mM Tris, 300 mM (NH ₄) ₂ SO ₄ , 5 mM DTT, pH 8.0 (Adjusted with HCl)
TPH1-C	20 mM Tris, 100 mM (NH ₄) ₂ SO ₄ , pH 8.0 (Adjusted with HCl)

2.3.4 Purification of MBP-3CP

For purification of MBP-3CP, the sample preparation was as described in section 2.3.3, except that the binding buffer was 20 mM Tris, 70mM (NH₄)₂SO₄, pH 8.0 and cells from 2x650 mL cultures were used. The sample was loaded onto a pre-equilibrated Dextrin Sepharose affinity column (GE Healthcare) with a flow rate of 5 mL/min. After removing unbound proteins with 5 CV binding buffer, MBP-3CP was eluted with elution buffer (binding buffer + 0.1 mM maltose). The buffer was exchanged to storage buffer (50mM Tris, 50mM (NH₄)₂SO₄, 1mM DTT, 10mM EDTA, 20v/v% glycerol, pH 8.0) by ultrafiltration (>500x dilution of the elution buffer) using a 200 mL Amicon stirred cell fitted with a PLGC NMWL 10 kDa membrane (EMD Millipore) and pressurized with N₂. The protein was stored at -20 °C in fractions containing 80 nmol MBP-3CP.

2.3.5 Protein concentration determination

Protein concentration was determined by UV absorbance at 280 nm using an ND-1000 NanoDrop Spectrophotometer (Saveen Werner). The extinction coefficients were estimated using the online tool ProtParam from the ExPASy server¹⁰⁵. For TPH1-R, with an ε₂₈₀ of 0, the concentration was determined using an Abbematt550 refractometer (Anton Paar) and dn/dc=0.185^{106,107}.

2.3.6 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were diluted with water if needed to avoid overloading the lanes, and 15 μ L sample was mixed with 5 μ L 4xLeammli sample buffer (Bio-Rad, with 10% V/V β -mercaptoethanol added), and heated for 10 min at 80 $^{\circ}$ C prior to loading on either an AnyKd or 7.5 % Mini-PROTEAN[®] TGX Stain-Free[™] gel (Bio-Rad) dependent on the protein variant. To estimate the molecular weight, Precision Plus Protein Standard (product no. 161-0363, Bio-Rad) was used. The gels were run at 250 V, and visualized using a Gel Doc[™] EZ System (Bio-Rad) and Image Lab software (Bio-Rad). In case of TPH1-R, the gels were visualized by staining with Coomassie Blue G-250, as the variant does not contain any Trp residues and is thus undetectable using the Gel Doc[™] EZ System.

2.3.7 Nano differential scanning fluorimetry (nanoDSF)

To investigate the stability of N Δ 47-TPH2-R in different buffers, nanoDSF was used. A frozen batch of purified N Δ 47-TPH2-R was thawed on ice, filtered through a 0.22 μ m centrifuge tube filter (Corning), and the concentration was determined as described in section 2.3.5. The protein had been purified in either 12 mM Na₂HPO₄/8 mM NaH₂PO₄, 100 mM (NH₄)₂SO₄, pH 7.0 or 20 mM Hepes, 100 mM (NH₄)₂SO₄, pH 7.0 depending on which buffer system was used in the screen. When investigating the effect of different salts on the melting temperature, the buffer was additionally exchanged to a “no salt buffer” after purification (20 mM Hepes, pH 7.0) to remove (NH₄)₂SO₄ using a Vivaspin Turbo centrifugal concentrator (Sartorius) with a 3 kDa membrane ensuring >500x dilution of the original buffer. To investigate the effect of pH on the melting temperature, the protein was exchanged into a buffer containing the highest and lowest tested pH value, e.g. 20 mM Hepes, 100 mM (NH₄)₂SO₄, pH 8.0 or pH 6.5, using Vivaspin Turbo centrifugal concentrators (Sartorius) with a 3 kDa membrane. The intermediate pH values were then obtained by mixing the pH 7.0 buffer (with protein) with a calculated amount of the high or low pH value buffer (the resulting pH was measured on buffers mixed from the calculated ratios). Each final sample contained 100 μ M N Δ 47-TPH2-R and was loaded manually into standard grade capillaries (NanoTemper Technologies). The experiments were performed with the Prometheus NT.48 (NanoTemper Technologies) with a linear thermal ramp of 0.5 $^{\circ}$ C/min from 15 to 80 $^{\circ}$ C and using 30 % gain. The final melting temperature was determined as the inflection point of the 350/330 nm intrinsic fluorescence ratio analyzed using the software PR.ThermControl (NanoTemper Technologies).

2.4 RESULTS

Chromatograms and SDS-PAGES of TPH2, NΔ47-TPH2, NΔ47-TPH2-RC, TPH2-CT, TPH1-RC, and TPH1-C can be found in Appendix B. Purifications were performed according to the method described in section 2.3.3

2.4.1 Purification of NΔ47-TPH2-R

An example from a purification of NΔ47-TPH2-R using 12 mM Na₂HPO₄/8 mM NaH₂PO₄, 100 mM (NH₄)₂SO₄, 10 mM L-Phe, pH 7.0 is shown in Figure 10. The purification profile of purifications without L-Phe is very similar (Appendix B.6). The SDS-PAGE (Figure 10C) shows that MBP-NΔ47-TPH2-R is well expressed (lane R) and that the protein is mainly soluble (low amount of protein in lane P relative to S). Additionally, the flow through (FT) during affinity chromatography only contains small amounts of MBP-NΔ47-TPH2-R indicating efficient binding to the Dextrin Sepharose column. After cleaving off the MBP-tag with MBP-3CP, the target peak (Tp) was collected, and as seen in the SDS-PAGE, the protein is already very pure at this stage, as only a small amount of other proteins are visible. This was also evident during the GF where only one main peak eluted containing pure NΔ47-TPH2-R. The “plateau” region preceding the Gp peak (Figure 10B) likely contains the few impurities present in the Tp (Figure 10C).

The oligomeric state of NΔ47-TPH2-R was found to be dimeric (discussed in Chapter 4). TPH1-R is able to form intermolecular disulfide bonds between the single cysteine residue (Cys66), resulting in a covalently bonded dimer⁷⁵. Cys66 is conserved in TPH2 (Cys112), and TPH2-R contains an additional cysteine residue at position 114. To ensure that the dimer formation of NΔ47-TPH2-R was not a result of disulfide bonds, an SDS-PAGE of purified NΔ47-TPH2-R (purified in 12 mM Na₂HPO₄/8 mM NaH₂PO₄, 100 mM (NH₄)₂SO₄, pH 7.0) was run with and without reduction agent (β-mercaptoethanol). In both cases, only one band was visible on the gel (Figure 10D) at a size corresponding to a monomer, suggesting that NΔ47-TPH2-R does not form dimers through disulfide bonds. Thus, reduction agents can be omitted during purifications and experiments.

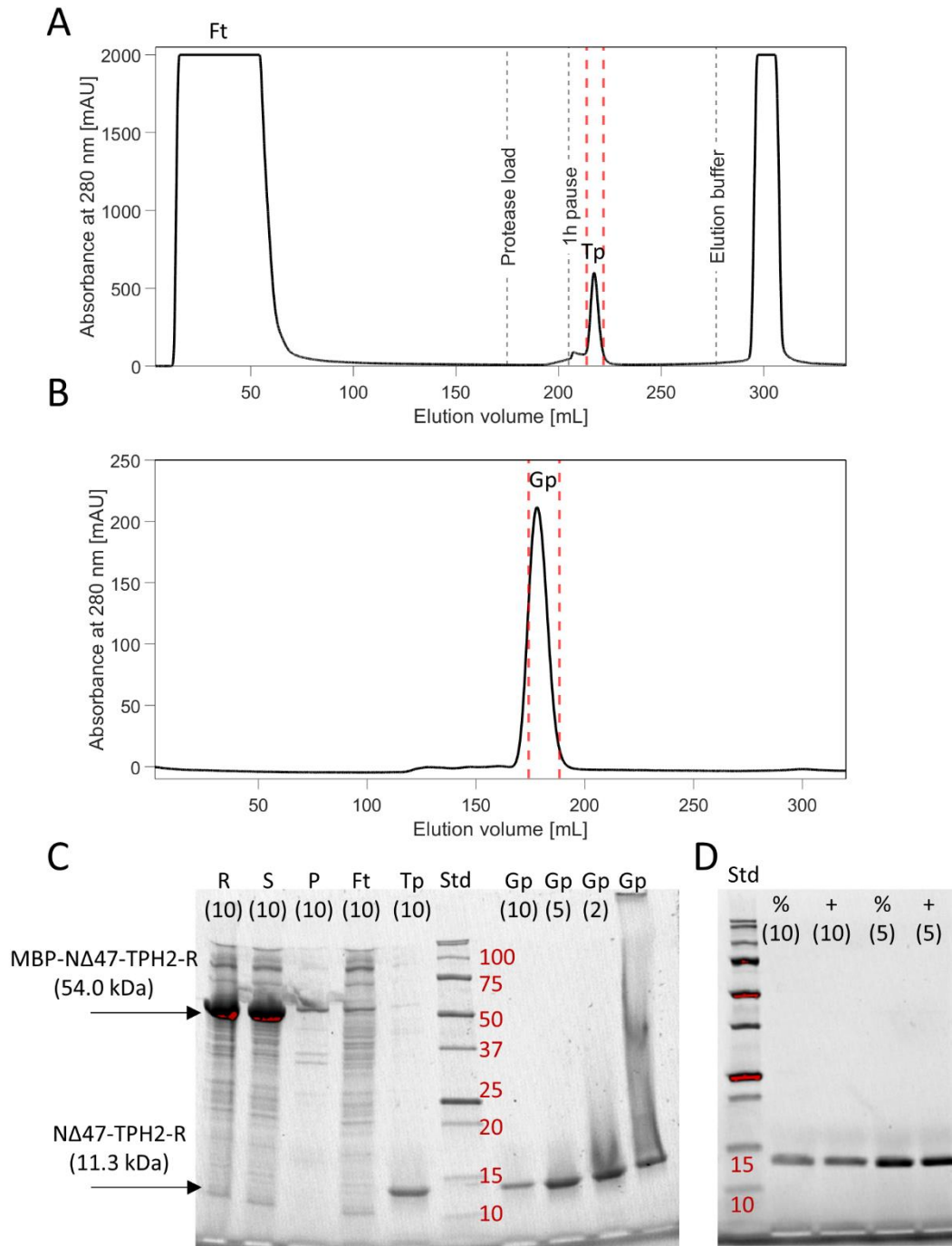


Figure 10 Purification of NΔ47-TPH2-R in 12 mM Na₂HPO₄/8 mM NaH₂PO₄, 100 mM (NH₄)₂SO₄, 10 mM L-Phe, pH 7.0. A) Affinity chromatogram, B) GF chromatogram, and C) SDS-PAGE analysis. Abbreviations: From sample preparation [R – lysate, S – Supernatant from centrifugation, P – pellet from centrifugation,]. From chromatographic steps [Ft – flow through, Tp – target peak from AC step, Gp – target peak from GF step], Std – Precision Plus Protein Standard. The sizes of the standard proteins are provided in kDa (in red). The number in parenthesis denotes the dilution factor of the samples if relevant. Red coloring on the image marked by the Image Lab software indicates a large amount of protein. Red dashed lines in the chromatograms indicate the collected NΔ47-TPH2-R fractions. D) SDS-PAGE of purified NΔ47-TPH2-R (purified in 12 mM Na₂HPO₄/8 mM NaH₂PO₄, 100 mM (NH₄)₂SO₄, pH 7.0.) with (+) and without (%) adding β-mercaptoethanol to the samples.

2.4.1.1 Buffer formulation of NΔ47-TPH2-R

A buffer screen was carried out using nanoDSF, in search for buffer conditions that could increase the stability of NΔ47-TPH2-R. First, different salt types and salt concentrations were screened using a “base buffer” with 20 mM HEPES at pH 7.0 (Figure 11). The salt stocks were dissolved in the base buffer and pH adjusted if needed, to maintain the pH at 7.0. The results showed that increasing concentrations of Na_2SO_4 and $(\text{NH}_4)_2\text{SO}_4$ stabilizes NΔ47-TPH2-R. NaCl and KCl did not show a significant effect on the stability, and increasing amounts of NH_4Cl lowered the stability. Thus, it seems that using SO_4^{2-} as anion compared to Cl^- has a stabilizing effect on NΔ47-TPH2-R. Na_2SO_4 resulted in slightly larger melting temperatures than the equivalent amount of $(\text{NH}_4)_2\text{SO}_4$. Though working with TPH1, Haahr similarly reported a larger melting temperature for TPH1 in a Na_2SO_4 buffer compared to an $(\text{NH}_4)_2\text{SO}_4$ buffer (using DSF) ¹⁰³. However, Haahr also reported a decrease in yield (~50% less) when purifying TPH1 in a Na_2SO_4 buffer instead of an $(\text{NH}_4)_2\text{SO}_4$ buffer ¹⁰³. Using $(\text{NH}_4)_2\text{SO}_4$ in high concentrations is a common tool to perform protein precipitation. Thus, the increase in melting temperature at 900 mM $(\text{NH}_4)_2\text{SO}_4$ could be questionable. Haahr observed similar increases in melting temperature for TPH1 using up to 1800 mM $(\text{NH}_4)_2\text{SO}_4$, but also reported a decrease in yield from 16 to 7 mg/L culture when purifying in 900 mM $(\text{NH}_4)_2\text{SO}_4$ instead of 300 mM ¹⁰³. Despite the incomplete understanding of the effect of high $(\text{NH}_4)_2\text{SO}_4$ concentration on NΔ47-TPH2-R, it was chosen to continue with $(\text{NH}_4)_2\text{SO}_4$ in the buffer, as $(\text{NH}_4)_2\text{SO}_4$ has been used for studies of most TPH variants in the group ^{47,52,56,104}, making it easier to compare results. NΔ47-TPH2-R was mainly studied using NMR spectroscopy, and therefore it was chosen to use only 100 mM $(\text{NH}_4)_2\text{SO}_4$, as high salt concentrations can lower the signal-to-noise ratio in NMR spectra ¹⁰⁸. Next, the effect of pH in the range of 5.5-8.0 was investigated (Figure 11B), using 20 mM buffer component in the presence of 100 mM $(\text{NH}_4)_2\text{SO}_4$. NΔ47-TPH2-R was found to be most stable at pH 6.5 and 7.0, while a pH<6.5 decreased the melting temperature. Additionally, the results indicate that pH>7.0 slightly lowers the melting temperature. Using a sodium phosphate buffer instead of HEPES, seems to yield a small increase in melting temperature. It was chosen to maintain the pH at 7.0, to facilitate comparison to previous

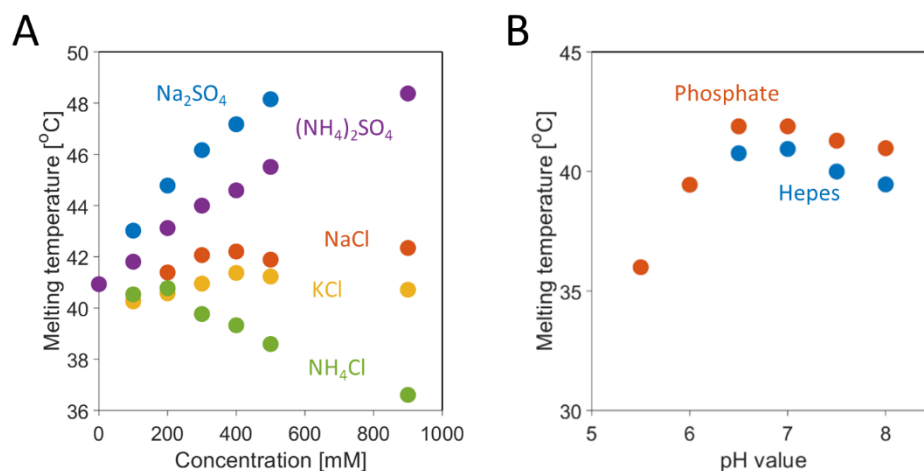


Figure 11 Melting temperatures of NΔ47-TPH2-R in different buffers determined by nanoDSF. A) In 20 mM HEPES, pH 7.0, and different salt concentrations. B) In either 20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ or 20 mM HEPES and 100 mM $(\text{NH}_4)_2\text{SO}_4$ at different pH values.

studies of TPH2 variants, and as Dextrin Sepharose columns has a pH working range >7.0. Additionally, it was chosen to use phosphate as the buffer system.

2.4.2 Purification of TPH2-R

As the first 47 residues of TPH2 are known to contain an isoform specific phosphorylation site, insights into this part of the domain is important. Therefore, it was attempted to purify the entire isolated RD, TPH2-R. However, the purification protocol did not prove to be successful for this variant. As seen in Figure 12B, MBP-TPH2-R is soluble, and binds to the affinity column as little protein elutes in the flow through. However, only two small peaks eluted after protease cleavage (P1 and P2, Figure 12A). From the SDS-PAGE analysis both P1 and P2 was shown to contain MBP-TPH2-R and MBP, but TPH2-R could not be observed. The presence of MBP in P1, P2, and especially in the elution peak (E), indicates that MBP-TPH2-R has been cleaved. Indeed, a faint band of TPH2-R is present in the elution peak (MBP-TPH2-R from the elution peak cleaved by co-eluting MBP-3CP). However, the fairly strong band of MBP-TPH2-R in the elution peak also indicates inefficient cleavage by MBP-3CP. It can be speculated that the reason for the missing TPH2-R in P1 and P2, is that TPH2-R has very low solubility without the MBP-tag in the chosen conditions, and therefore precipitates on the column. A second purification was attempted (Appendix B.5) with L-Phe included in the buffer. Although some TPH2-R was present in the target peak, it was not possible to separate the impurities using GF, as TPH2-R co-eluted with MBP, resulting in around 50% purity. The purification attempts illustrate the challenges involved in studying the full-length RD of TPH2.

2.4.3 Purification of TPH1-R

TPH1-R was purified in 20 mM Tris, 100 mM $(\text{NH}_4)_2\text{SO}_4$, 5 mM DTT, pH 8.0 (Figure 13). DTT was included in the buffer due to the tendency of TPH1-R to form dimers through disulfide bonds⁷⁵. The SDS-PAGE lane containing the pellet (P) and supernatant (S) from sample preparation, shows that MBP-TPH1-R is mainly soluble. The target peak collected after proteolytic cleavage contains almost pure TPH1-R, though a small amount of MBP-TPH1-R and MBP is also present. The impurities were removed during the GF step, where the TPH1-R protein peak (Gp) was easily identified as there was only absorption at 254 nm and not at 280.

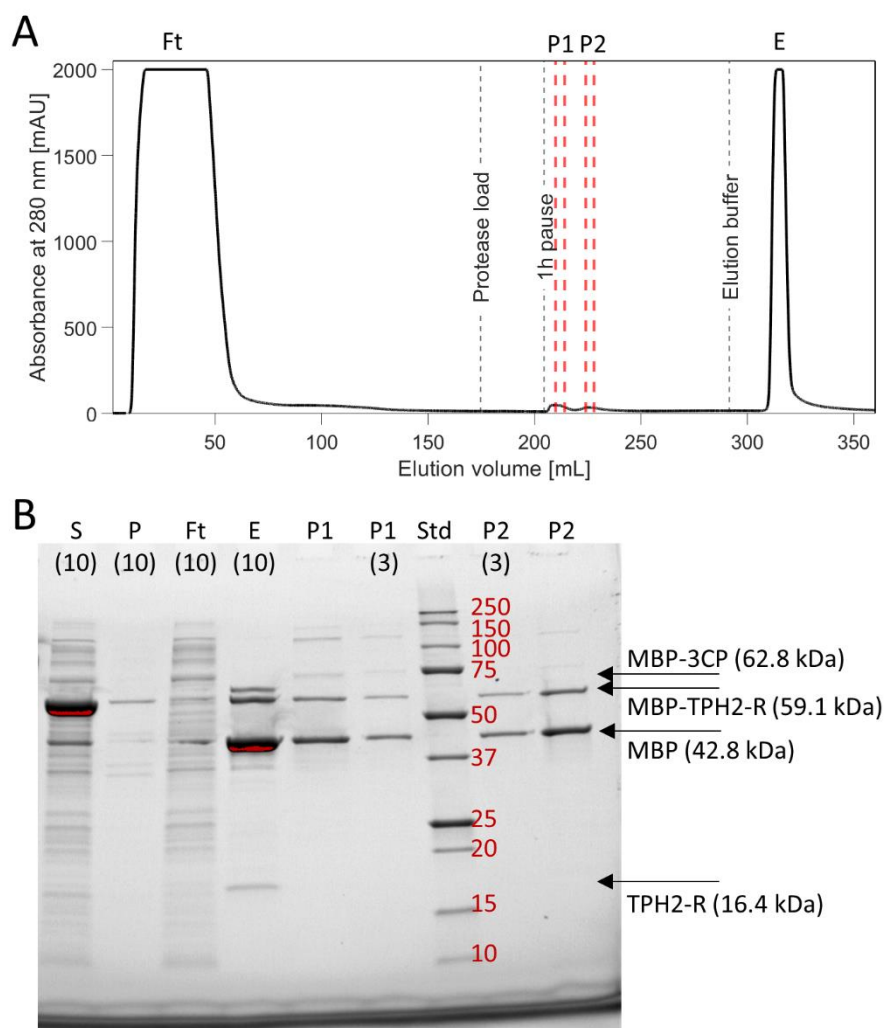


Figure 12 Purification of TPH2-R in 12 mM Na_2HPO_4 /8 mM NaH_2PO_4 , 100 mM $(\text{NH}_4)_2\text{SO}_4$, pH 7.0. A) Affinity chromatogram and B) SDS-PAGE analysis. Abbreviations: From sample preparation [S – Supernatant from centrifugation, P – pellet from centrifugation,]. From chromatographic steps [Ft – flow through, E – Elution peak], Std – Precision Plus Protein Standard. The sizes of the standard proteins are provided in kDa (in red). The number in parenthesis denotes the dilution factor of the samples if relevant. Red coloring on the image marked by the Image Lab software indicates a large amount of protein. Red dashed lines indicate the collected “P1” and “P2” fractions.

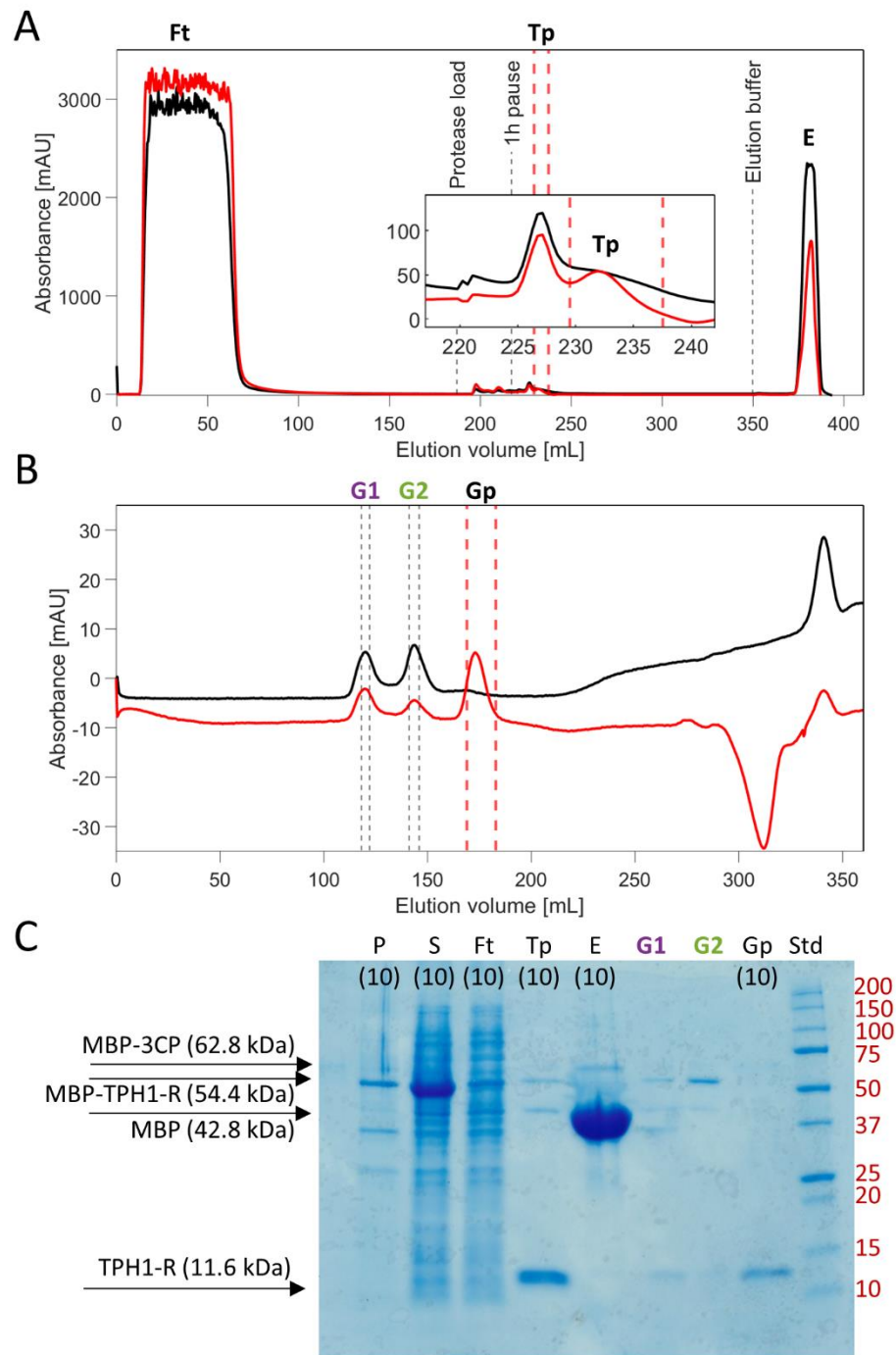


Figure 13 Purification of TPH1-R in 20 mM Tris, 100 mM $(\text{NH}_4)_2\text{SO}_4$, 5 mM DTT, pH 8.0. A) Affinity chromatogram, B) GF chromatogram, and C) SDS-PAGE analysis. Abbreviations: From sample preparation [P – pellet from centrifugation, S – Supernatant from centrifugation], from chromatographic steps [Ft – flow through, Tp – target peak from AC step, E – elution peak, Gp – target peak from GF step], Std – Precision Plus Protein Standard. The sizes of the standard proteins are provided in kDa (in red). Red dashed lines in the chromatograms indicate the collected TPH1-R fractions. In the chromatograms the black curves show the absorbance at 280 nm and the red curves at 254 nm. G1 and G2 mark two sampled fractions. The purification and SDS-PAGE analysis was performed by Master student Maja van Hees.

2.4.4 Purification yields of TPH variants

The average purification yield of the different variants are illustrated in Figure 14. Only purifications without L-Phe in the buffer were included in the calculations to enable comparison between the variants. For TPH2, the destabilizing effect of the first 47 residues is observed, as the yield greatly decreases for TPH2 compared to NΔ47-TPH2, and for TPH2-R (not possible to purify under the conditions tested) relative to NΔ47-TPH2-R. The yields of TPH2, NΔ47-TPH2, NΔ47-TPH2-RC, and TPH2-CT are similar to previous reported values ^{103,104}, though Skawinska obtained a yield around 2x larger for NΔ47-TPH2-RC ⁴⁷. For the TPH1 variants, the highest yield was obtained for TPH1-R, and the yields of TPH1-RC and TPH1-C are slightly lower than previously reported values ¹⁰⁴.

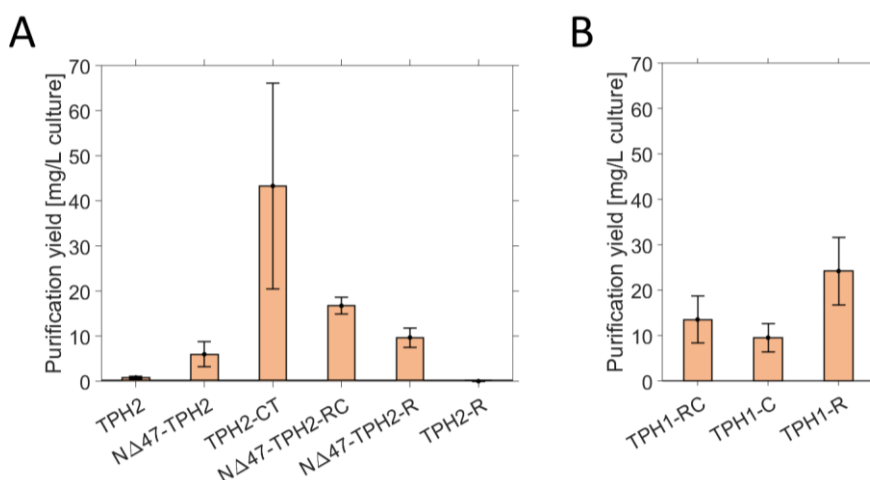


Figure 14 Average purification yield (± 1 SD) of A) TPH2 variants and B) TPH1 variants all purified without any stabilizing ligands (e.g. L-Phe) in the buffer. TPH2-R was not possible to purify under the tested conditions and thus no yield is reported.

2.5 DISCUSSION

When studying ligand binding (e.g. L-Phe or L-Trp) to the RD of TPH variants also containing the CD, it can be difficult to isolate if any observed effects of the binding are indeed a result of binding to the RD, or a result of binding to the active site in the CD where both L-Trp and L-Phe bind. The two isolated RD variants, NΔ47-TPH2-R and TPH1-R was successfully purified in quantities sufficient for protein characterization studies using an already developed procedure, which facilitated the study of ligand binding to the RDs, investigated in Chapter 4 and Chapter 5. Purification of TPH1-R has previously also been reported using a GST-tag ¹⁰² or a histidine-tag ⁷⁵.

While the purification method was successful for NΔ47-truncated TPH2 variants, efficient purifications strategies for full-length TPH2 and TPH2-R are still needed. That TPH2 could be purified in small quantities (0.6 mg/L culture), while TPH2-R could not be purified, indicates that the N-terminal tail might be stabilized to some extent by the presence of the CD. Though in low yields, TPH2-RC has also been purified using the same purification method (0.8 \pm 0.2 mg/L culture) ¹⁰⁴, supporting a possible stabilizing effect by the CD. The first 47 residues of TPH2 contains a highly hydrophobic and expectedly disordered section (2-

QPAMMMFSSKYWA-14), and it can be speculated that this region might be aggregation prone in absence of stabilizing interactions. A possible modification to the purification of TPH2-R could be to perform the proteolytic cleavage off-column, as has been done previously for other TPH variants ¹⁰³. However, if the main issue is the solubility and stability of TPH2-R, more stable conditions will need to be identified before successful purification can be achieved.

2.6 CONCLUSION

Five TPH2 and three TPH1 variants have been purified using a previously developed method. While it was possible to purify a truncated form of the isolated RD of TPH2 (NΔ47-TPH2-R), it was not possible to purify the full-length RD (TPH2-R) with the conditions used here.

3

METHODS

3.1 NUCLEAR MAGNETIC RESONANCE (NMR) SPECTROSCOPY

Solution NMR spectroscopy is a versatile technique suitable for studying both small molecules and biomolecules. In NMR spectroscopy, a sample is placed in an external magnetic field (B_0) and exposed to specific radiofrequency pulses (i.e. an RF pulse sequence) to perturb the magnetic moments of the nuclei (spins) from their equilibrium state (aligned with B_0)¹⁰⁹. The free induction decay (FID) is recorded as the final step in the pulse sequence, and Fourier transformed (FT) to yield the NMR spectrum¹⁰⁹ (Figure 15). The resonance frequency of each spin in the spectrum, or the chemical shift (δ) usually expressed in ppm, depends among other parameters on the type of nuclei and on the chemical environment surrounding the spin^{109,110}. Thus, structural changes are easily observed in NMR spectra. The RF pulse sequence determines the type of spectra that is acquired, and numerous pulse sequences are available for extracting specific information. The development of multi-dimensional NMR, protein labelling-schemes, and non-uniformly sampled data greatly expanded the application of NMR spectroscopy to study proteins^{110–113}. In multi-dimensional NMR spectroscopy, the resonances of different spins are correlated through magnetization transfer via scalar couplings (through bonds) or dipole interactions (through space)^{109,114}. Here, the use of NMR spectroscopy to study protein-ligand (PL) interactions, protein dynamics, and protein structure, being relevant for the work conducted in this project, will be briefly outlined.

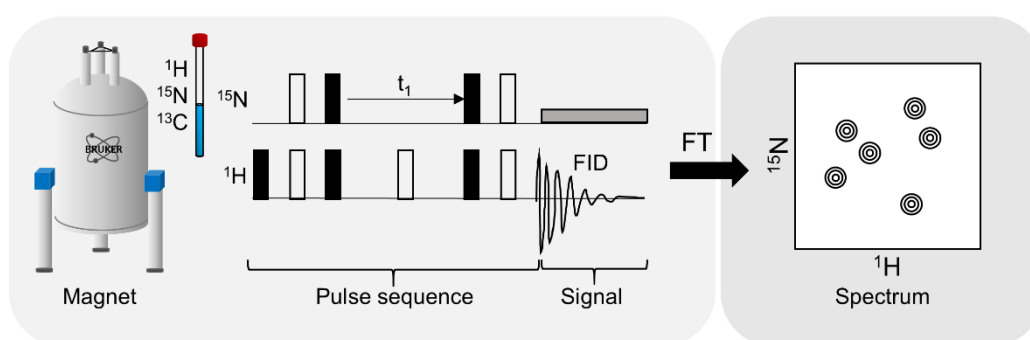


Figure 15 Schematic illustration of an NMR experiment. Proteins are typically ^{15}N or $^{15}\text{N}/^{13}\text{C}$ labelled to enable recording of multi-dimensional spectra. A spectrum specific pulse sequence is applied (solid black bars indicate 90° pulses and white bars 180° pulses), and the signal (FID) is measured at the end of the sequence. Using FT, the FID is converted into the final spectrum. Here a ^1H - ^{15}N HSQC spectrum is simplistically illustrated, where each peak represents a proton directly bonded to a ^{15}N nuclei.

3.1.1 Protein-ligand interactions

A distinctive feature of NMR spectroscopy is the ability to provide both quantitative and structural information on PL complexes¹¹⁵. As the chemical environment of a given residue

changes due to ligand binding, the peaks of affected spins will change in a spectrum. The strength of a simple two-state PL interaction is usually described by the dissociation constant (K_d), given by Eq. 1¹¹⁶.

$$K_d = \frac{[P][L]}{[PL]} = \frac{k_{off}}{k_{on}} \quad (1)$$

Where $[P]$ is the concentration of unbound protein, $[L]$ is the concentration of unbound ligand, $[PL]$ is the concentration of the PL complex, and k_{off} and k_{on} are the rate constants of dissociation and association of the complex, respectively. k_{off} , k_{on} and $[L]$ determines the exchange rate, k_{ex} (Eq. 2¹¹⁶, for a two-state binding model):

$$k_{ex} = [L]_{free}k_{on} + k_{off} \quad (2)$$

The magnitude of the exchange rate relative to the chemical shift difference (in rad/s) between the bound (PL) and the free state (P) ($\Delta\omega = |\omega_P - \omega_{PL}|$), influences the appearance of NMR spectra^{115–117}. In general, three exchange regimes or time scales are defined^{115–117}: fast exchange ($\Delta\omega \ll k_{ex}$), intermediate exchange ($\Delta\omega \approx k_{ex}$), and slow exchange ($\Delta\omega \gg k_{ex}$). In the fast exchange regime, the location of a peak in a spectrum (δ_{obs}) depends on the population (p) of P and PL¹¹⁶:

$$\delta_{obs} = \delta_P p_P + \delta_{PL} p_{PL} \quad (3)$$

Thus, in a titration the position of a peak from a given nuclei, will move gradually from the unbound state until the bound state at full saturation, as illustrated in Figure 16. In the slow exchange regime, the frequency of each state (P and PL) can be observed separately during a titration, but the intensity of the peaks depends on the population of the states. In the intermediate exchange regime, a gradual movement of the peak is observed, and at the same time, the peak becomes broadened due to the exchange process, which in some cases result in undetectable peaks^{115–117}. The three exchange regimes apply to any equilibrium between

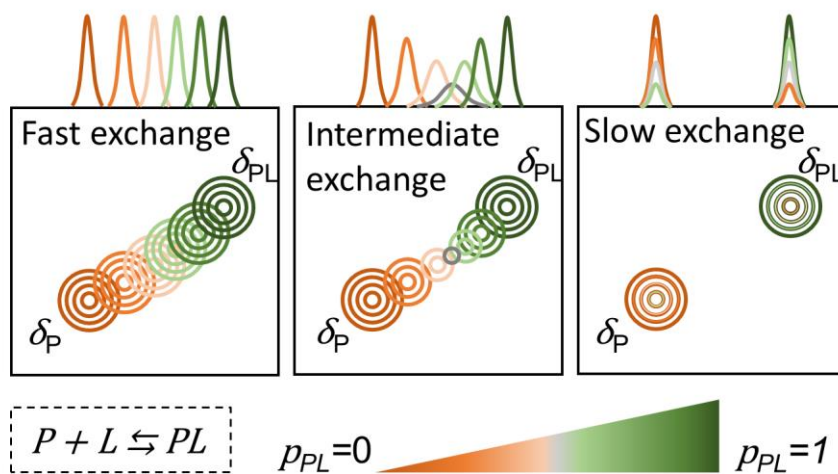


Figure 16 Schematic illustration of the chemical exchange regimes in NMR spectroscopy. The signal in one dimension is illustrated on top of the 2D spectrum. In each case, the gradual change during a titration from unbound state ($p_{PL}=0$) to bound state ($p_{PL}=1$) is illustrated. Figure made with inspiration from¹¹⁶.

different populations, e.g. conversion between a monomer and a dimer, or an open and closed conformation of a loop ¹¹⁷. The exchange regime of a given PL interaction influences how K_d can be reliably determined from NMR spectra. In the case of fast exchange, the change in chemical shift can be plotted as a function of ligand concentration, and Eq. 4 can be fitted to the data to yield K_d ¹¹⁵.

$$\Delta\delta_{obs} = \Delta\delta_{max} \frac{([P]_0 + [L]_0 + K_d) - \sqrt{([P]_0 + [L]_0 + K_d)^2 - 4[P]_0[L]_0}}{2[P]_0} \quad (4)$$

Where $\Delta\delta_{obs}$ is the observed change in chemical shift, $\Delta\delta_{max}$ is the change in chemical shift at complete saturation, $[P]_0$ is the protein concentration, and $[L]_0$ is the ligand concentration. In case of slow exchange, the intensities can be used to fit K_d instead of the chemical shift perturbations. However, it should be ensured that the peak width is constant throughout the titration for reliable fitting ¹¹⁵. In more complex cases, where the peak width is not constant or the chemical shift perturbations cannot be traced throughout the titration as often in the intermediate exchange regime, titration data can be fitted using line-shape analysis ¹¹⁸. Line-shape analysis can additionally be used to analyze more complex binding mechanisms than the simple two-state model ¹¹⁸. For proteins, a ^1H - ^{15}N HSQC is typically used to observe changes induced by possible binding partners and for protein observed titrations ^{115,116}. The spectrum is a fairly simple 2D spectrum containing one peak per backbone amide and thus per residue (except prolines), and ideally the changes of each residue in the spectrum can therefore be observed. For a suitable signal-to-noise ratio, approximately 20-50 μM protein should minimally be used ^{115,116}, and consequently NMR spectroscopy cannot be used to reliably determine the K_d of strong PL complexes (K_d smaller than $\sim 10 \mu\text{M}$) ¹¹⁶, but is still valuable for gaining structural information. To gain site-specific information, the chemical shifts need to be assigned, i.e. the nuclei giving rise to each peak need to be identified. The assignment principle will be elaborated in section 3.1.3.

3.1.2 Protein dynamics

The ability to study site-specific protein dynamics is unique to NMR spectroscopy, which can probe dynamics on time-scales ranging from ps-ns (e.g. local motion of the protein backbone or side-chain rotation), μs -ms (e.g. catalysis, ligand binding, allostery), and ms-s (e.g. unfolding events, large complex formation) ¹¹⁹. Local motions of the protein backbone on the ps-ns time-scale are usually studied by measuring the ^{15}N longitudinal relaxation rate (R_1), the ^{15}N transverse relaxation rate (R_2), and the heteronuclear nuclear Overhauser effect (het-NOE) ¹¹⁹. In short, relaxation is the process at which magnetization of a spin returns to its equilibrium (aligned with B_0) after being perturbed by RF pulses. R_1 is the recovery rate of magnetization aligned with B_0 (usually termed z-magnetization), while R_2 is the return of the transverse magnetization (x,y-magnetization) to 0 ¹⁰⁹. Both processes are driven by the unique local magnetic fields that each spin experiences due to other spins that are close in space (the dipole-dipole mechanism), and due to fluctuations in the local magnetic fields from variations in the degree of shielding by electrons (chemical shift anisotropy) ^{109,120}. In the het-NOE experiment, the effect of cross-relaxation between ^{15}N spins and the directly bonded proton is investigated ¹⁰⁹. Plotting R_1 , R_2 , and het-NOE values versus the residue number provides quick insight to the local flexibility of the backbone. The het-NOE value is close to 1 for residues with

restricted internal motion, i.e. in a folded part of the protein, while less motional restricted residues, i.e. in flexible parts of the protein, display lower het-NOE values ¹²⁰. Similarly, low values of R_2/R_1 (relative to the average value at a given site) indicates flexibility, and locally elevated R_2/R_1 values can indicate local conformational exchange (e.g. loop movement) ^{114,120}. Conformational dynamics on the μ s-ms time-scale is often studied using Carr-Purcell-Meiboom-Gill (CPMG) spin-echo experiments, and ms-s dynamics can be probed by zz-exchange experiments ¹¹⁹. The details of these experiments will not be discussed here.

3.1.3 Protein structure determination

Protein NMR structure determination ideally requires a stable sample (>1 week, since many long spectra need to be recorded) at a high concentration to increase the signal-to-noise ratio. For small proteins 200-300 μ M can be sufficient ¹¹⁶, while larger proteins require higher concentration due to their slower tumbling time resulting in increased line-broadening ^{111,114}.

The process of structure determination is schematically illustrated in Figure 17. First, backbone chemical shifts need to be sequentially assigned. This requires recording and analysis of different types of 3D triple resonance spectra, that correlates ¹H, ¹⁵N, and ¹³C nuclei ^{111,114}. For the side chain assignment, HCCH-TOCSY experiments are typically used, which links the resonances within each side-chain ^{114,121}. The assignment process becomes increasingly difficult as the size of the protein increases as (i) larger molecules increase the likelihood of spectral overlap and (ii) larger molecules tumble slower ¹¹⁴. For small proteins or peptides, up to 50 residues, 2D spectra can be sufficient for assignment, while triple resonance spectra are usually required for larger proteins up to 25-30 kDa ¹¹⁴. At higher molecular weights, deuteration is usually required as well ¹¹⁴. The assignment completeness influences the quality of the protein structure, as residues without any chemical shifts assigned are “invisible” in the NMR spectra. Thus, experimental information cannot be extracted for these residues.

After assignment, conformational restraints are collected and/or calculated. The minimum restraints include backbone dihedral angle- and distance restraints. The dihedral angle restraints (ϕ and ψ) can be derived from chemical shift data based on an empirical relationship using programs as TALOS-N (Torsion Angle Likelihood Obtained from Shift and Sequence Similarity N) ¹²². Distance restraints are generated from a NOESY spectrum ¹¹², which is essential in protein NMR structure determination. In the experiment/pulse sequence, magnetization is transferred via dipole-dipole interactions between two nuclei that are close in space (<5-6 Å) ^{111,114}. Thus, each cross-peak, or NOE, in the NOESY spectrum represents two protons that are close in space. The volume (V) of the cross-peak is related to the distance (r), $V \propto r^{-6}$, and used to determine an upper-distance limit (the distance restraint) between two protons ^{114,123}. The two protons giving rise to each cross-peak are identified based on the chemical shift assignment, increasing the importance of a high assignment completeness and accuracy. The NOE assignment and distance restraint generation can be performed automatically and simultaneously with the initial structure generation in an iterative process ¹¹⁴. In this project, the program CYANA ¹²³, was used. In seven iterative cycles, CYANA first

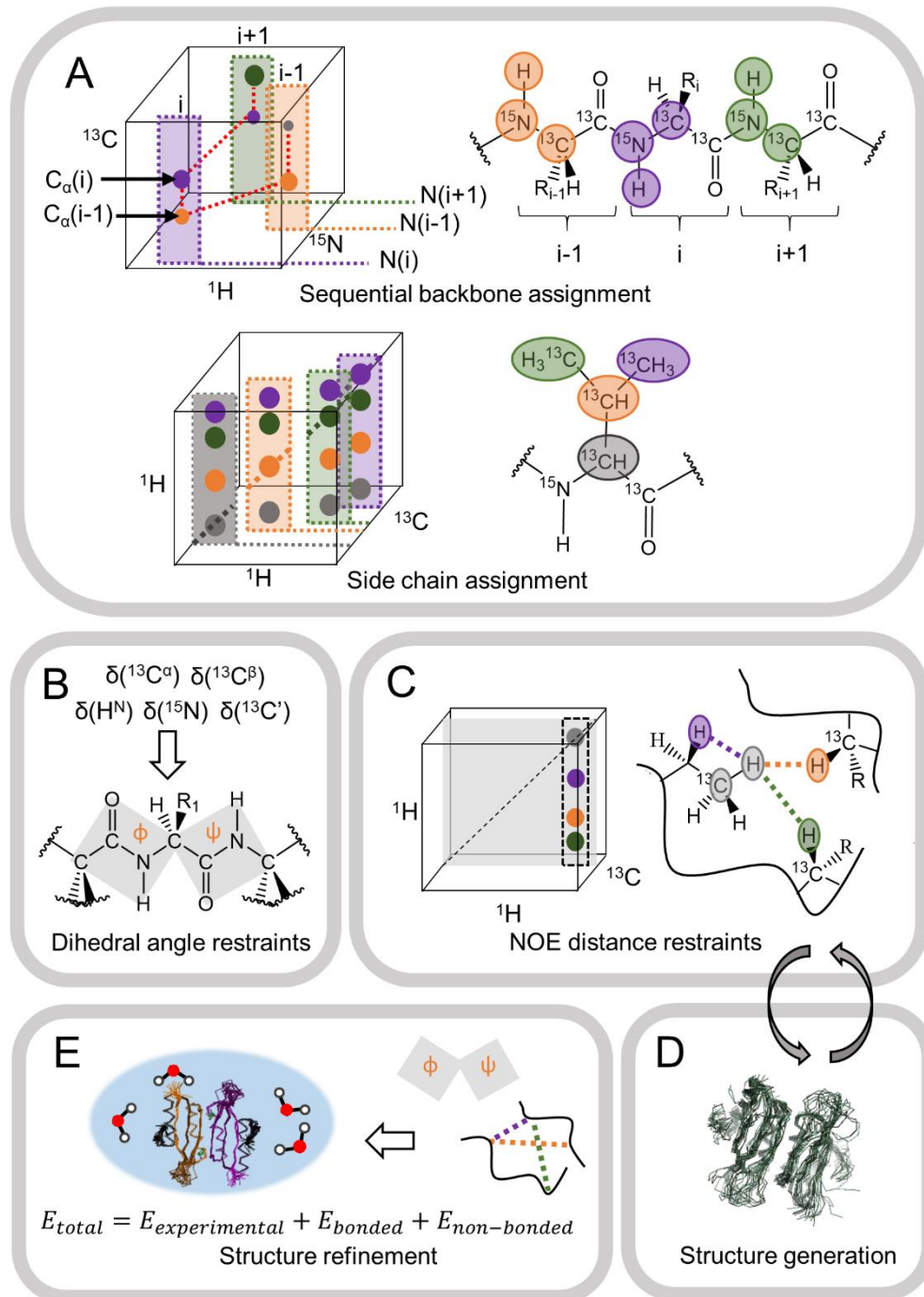


Figure 17 Schematic illustration of the NMR structure determination process. A) Assignment of chemical shifts. Sequential backbone assignment: an HNCA spectrum is illustrated. In the 3D spectrum, the NH plane of each residue (i , purple plane color) contains a peak from its own C_α atom (purple circle) and one peak from the C_α atom of the preceding residue ($i-1$, orange circle), which is used to sequentially link the resonances. An HCCH-TOCSY spectrum is typically used to assign side chain resonances. Each C-H plane of a side chain ^{13}C -nuclei contains peaks from all other H-atoms in the side-chain. B) The assigned resonances are used to generate dihedral angle restraints based on an empirical relationship. C) In the ^1H - ^{13}C -HSQC-NOESY spectrum the C-H plane of each ^{13}C nuclei contains cross peaks (NOEs) to protons within 5 Å. The volume of the cross peak is related to the distance between the two protons. The NOEs can be automatically assigned simultaneously within initial structure generation (D) in an iterative process to provide distance restraints. E) The initial structures are refined through energy minimization which includes an energy term for the experimentally derived conformational restraints.

assigns NOEs using the determined chemical shifts of the protein while allowing ambiguous NOE assignments, and next calculates a structural bundle using a simulated annealing algorithm ¹²³. The structures generated in each cycle are used as input in the following cycle to guide the NOE assignment. After the last cycle, a structure bundle is calculated using only unambiguously assigned NOEs which are gathered as the final distance restraints.

The final distance restraints, the dihedral angle restraints, and the structure bundle are used for input in the final step – structure refinement. In this step the structures are refined by energy minimization. XPLOR-NIH, which is used in this project, uses a combination of simulated annealing using molecular dynamics simulation and gradient minimization ¹²⁴. The total energy function that is minimized includes terms for experimental restraints, covalent interactions, and non-bonded interactions ^{124,125}. The XPLOR-NIH FF (EEF_x) non-bonded potential includes terms for electrostatic interactions, Van der Waals Lennard-Jones interactions, and solvation free energy, which are usually greatly simplified in the initial structure generation to save on computational costs ¹²⁵. The solvation free energy is in XPLOR-NIH calculated using an implicit water model ¹²⁵. In addition to dihedral angle and distance restraints, other experimental restraints can be included in the structure generation and/or energy refinement process. For instance, additional NMR derived restraints could include residual dipolar couplings and backbone relaxation data ¹²⁴. Furthermore, information on symmetry, and data from integrative experimental methods as SAXS/SANS and cryo-EM can also be incorporated in the refinement process ¹²⁴.

During validation of an NMR structure multiple parameters are considered, and the process typically includes data-, knowledge based-, and model versus data validation ^{126,127}. Input data can be validated by assessing assignment completeness and identifying statistically unlikely chemical shifts. Additionally, the number of (long-range) NOEs provides an estimate of the confidence in the determined protein fold. Knowledge based validation can include identification of steric clashes, dihedral angle outliers (Ramachandran outliers), ideal bond length- and angle deviations etc. The discrepancy between input conformational restraints and the final conformations of the structure (e.g. comparing distance restraints to the distances in the structure) is used to validate the consensus between the data and the calculated structures. Additionally, the RMSD of the different structures in the NMR bundle is typically reported as a measure of structure precision.

3.2 SMALL-ANGLE X-RAY SCATTERING (SAXS)

SAXS is a low resolution structural technique useful for studying proteins in solution. In SAXS, a sample is irradiated with a monochromatic X-ray beam (Figure 18). The scattering intensity of the X-rays is measured at the detector as a function of the momentum transfer (q), which is related to the scattering angle (2θ) and the wavelength of the X-rays by Eq. 5¹²⁸.

$$q = \frac{4\pi\sin\theta}{\lambda} \quad (5)$$

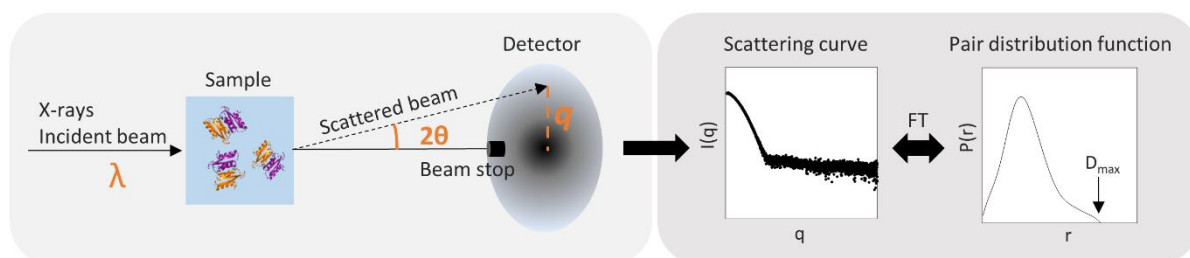


Figure 18 Schematic illustration of a SAXS experiment. A sample is irradiated with X-rays of wavelength λ , and the intensity of the scattered X-rays are measured at the detector as a function of the scattering angle 2θ . The data is radially averaged and the signal of a buffer sample is subtracted. The scattering curve is related to the pair distribution function of the protein via Fourier transformation (FT). Made with inspiration from¹²⁹.

The measured data is then radially averaged to obtain a 1D scattering profile of the intensity as a function of q ($I(q)$). SAXS is a contrast technique, and thus the signal of a buffer sample is subtracted from the protein sample to yield the scattering curve of the protein¹²⁸. The scattering curve can be Fourier transformed (using indirect Fourier transformation) to obtain the pair distribution function, $P(r)$ ^{128,129}. The $P(r)$ function describes the distribution of the interatomic distances in the studied protein, and the maximum dimension of the protein (D_{max}), is found at $P(r) = 0$ ¹²⁹. Additionally, the radius of gyration (R_g), the forward scattering ($I(0)$), the molecular weight (Mw), and the porod volume (V_p), can be determined from the SAXS data¹²⁸. R_g and $I(0)$ can be extracted by linear regression of the Guinier approximation, a Taylor expansion of the low q -region of the scattering curve (Eq. 6)¹²⁸.

$$I(q) = I(0) \exp\left(-\frac{q^2 R_g^2}{3}\right) \quad (6)$$

Once $I(0)$ is determined, the Mw can be calculated¹²⁸. $I(0)$, R_g , and the Mw can also be derived from analysis of the $P(r)$ function using software from e.g. the ATSAS package¹³⁰. Analysis of the $P(r)$ function additionally provides the parameters D_{max} and the porod volume (V_p). Comparing the molecular parameters determined from the Guinier approximation and the $P(r)$ analysis can provide an indication of the internal consistency of the data, as the Guinier analysis only evaluates the data at low q -values while the $P(r)$ analysis includes the entire q -range¹²⁹. Besides the molecular parameters, information on the flexibility of the system can be extracted from SAXS data by representing the data in a Kratky plot (q vs $q^2 I(q)$), where the shape of the curve depends on the rigidity of the protein¹²⁸. From a given atomic structure, a theoretical SAXS curve can be computed and thus compared to the experimental data¹³¹. This makes SAXS a strong technique in integrative structural biology. For example,

it can be used to model proteins and protein complexes where the crystal structure of the individual domains/proteins are already determined (rigid body modelling) ¹²⁹, to provide additional restraints in NMR structure refinement ^{124,132} or molecular dynamics simulations ¹³³, or to generate ensembles of flexible conformers ¹³⁴. To perform reliable SAXS modelling, it is important to consider if the sample is homogeneous and monodisperse, and whether interparticle interactions are present ¹²⁹. (Pre-)coupling SAXS to size exclusion chromatography can increase sample monodispersity, while measuring a protein concentration series can reveal if inter-particle interactions are occurring ¹³⁵.

3.3 SIZE EXCLUSION CHROMATOGRAPHY COUPLED TO MULTI-ANGLE LIGHT SCATTERING (SEC-MALS)

While conventional SEC with UV detection only results in an estimate of the protein molecular weight using protein standard calibration curves, coupling SEC to a MALS detector and either a refractive index (RI) detector or a UV detector allows the absolute molecular weight to be determined ¹³⁶. Thus, uncertainties in the estimated molecular weight from SEC that can arise due to unexpected interactions with the column material or differences in the shape of the protein compared to the standard proteins are avoided. Additionally, RI and MALS detection enables the usage of buffer components that would otherwise interfere with conventional UV detection, as L-Trp used in this project.

In MALS, the static light scattering of the sample is measured at different deflection angles (θ). The intensity of the scattered light ($I(\theta)_{scattered}$) is proportional to the Mw and the concentration (c) ¹³⁷:

$$I(\theta)_{scattered} \propto M_w \cdot c \quad (7)$$

As SAXS, MALS is a contrast based technique and the buffer signal at each deflection angle ($I(\theta)_{scattered\ solvent}$) is therefore subtracted. The scattering signal from the solute/protein is expressed by the excess Rayleigh ratio ($R(\theta)$) described via the following equation ¹³⁷:

$$R(\theta) = \frac{I(\theta)_{scattered} - I(\theta)_{scattered\ solvent}}{I_0} \frac{r^2}{V} \quad (8)$$

Where (I_0) is the incident light intensity, r is the distance from the sample to the detector, and V is the volume of the sample being illuminated. The Mw can be related to the Rayleigh ratio using the following equation ¹³⁷:

$$\frac{K \cdot c}{R(\theta)} = \frac{1}{M_w P(\theta)} + 2cA_2 \quad (9)$$

Where $P(\theta)$ is a form factor, K is an optical constant, and A_2 is the second virial coefficient ¹³⁷. For dilute solutions, as often assumed in the case of SEC-MALS, the second virial coefficient term in Eq. 9 can be neglected ¹³⁸. Molecular weights can be calculated from SEC-MALS data using a Debye plot ¹³⁷, which in this project was done using the software ASTRA (Wyatt). To

obtain a reliable molecular weight, the preceding separation of sample proteins using SEC (or other techniques e.g., ion exchange chromatography or asymmetric flow field-flow fractionation) is important to ensure monodisperse samples as one otherwise obtains the weight-averaged molar mass of the different proteins. Additionally, accurate determination of the protein concentrations in the eluting fractions is required, which is usually achieved using an RI detector coupled in parallel or in series to the MALS detector ¹³⁷.

3.4 SINGLE PARTICLE CRYO-ELECTRON MICROSCOPY (CRYO-EM)

Cryo-EM was performed in collaboration with Ph.D. Zhenwei Zhang (Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany), and therefore only a brief conceptual description of the method will be provided here.

Cryo-EM is a technique that can be used to determine atomic resolution structures of (large) biomolecules ^{139,140}. The method is well suited to study e.g. membrane proteins, large complexes, or dynamic proteins that are very difficult to crystallize and thus unsuitable for X-ray crystallography and too large to study by NMR spectroscopy ^{141,142}. Although high-resolution structures (~ 3 Å) of proteins down to 52 kDa have been determined, cryo-EM is not yet a routine method for high-resolution structure determination of proteins <200 kDa ¹⁴³. In cryo-EM, a protein sample is applied in a thin film onto a grid that is then flash frozen with liquid ethane, thereby trapping the protein particles in their hydrated state in a thin layer of vitreous ice ^{144,145}. Currently, the most standard method to create the thin film is blotting, illustrated in Figure 19A, though other methods are available ¹⁴⁶: First, the protein sample is pipetted onto a grid, and excess liquid is wicked/blotted away with filter paper to create a thin film of sample, after which the grid is plunged into liquid ethane. The process can be done both manually or semi-automated. Finding the correct settings of blot time and blot force are important for ensuring the optimal ice-thickness, which may vary from sample to sample ¹⁴⁶. The optimal ice-thickness is a balance between decreasing background noise (thinner ice)

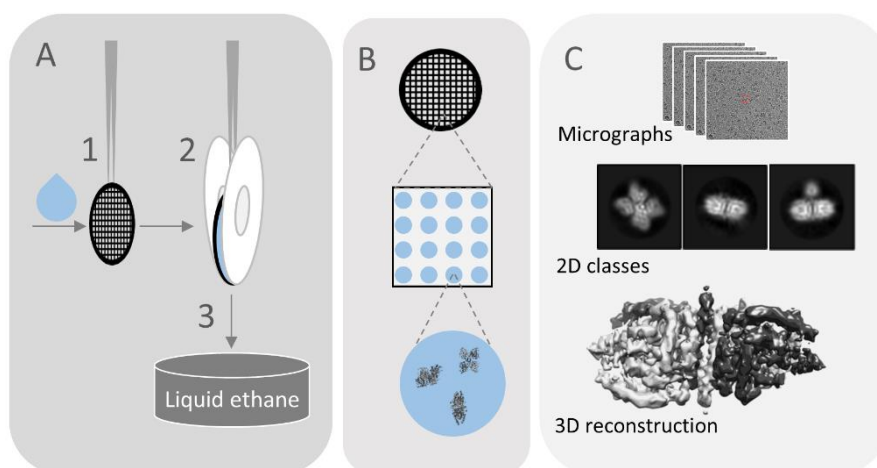


Figure 19 A) Simplistic illustration of “blotting” grid preparation in cryo-EM. 1: A sample drop is pipetted onto the grid. 2: Excess liquid is blotted away with filter paper. 3: The grid is plunged into liquid ethane, trapping the protein particles in hydrated states in the vitreous ice. B) Schematic illustration of a grid. Each square in the grid contains holes in which the protein particles are frozen in different (ideally random) orientations. C) Images of relevant holes in the grid (micrographs) are collected in the electron microscope. The 2D particle projections are computationally selected and aligned into 2D averages, and used to generate a 3D reconstruction of the protein.

while avoiding e.g. protein denaturation ^{146,147}. The grid is imaged in an electron microscope, and each micrograph (image) contains 2D projections of the protein (i.e. particles) in many different, ideally random, orientations ¹⁴⁵ (Figure 19C). Images of particles in the same orientation are computationally classified and aligned into 2D class averages to increase the signal-to-noise ratio ¹⁴⁵. The images are combined to generate a 3D reconstruction: A coulomb potential energy map corresponding to the electron density map in X-ray crystallography ¹⁴⁵. As discussed by Weissenberger et al., factors such as e.g. protein aggregation, denaturation, dynamics, and preferential orientation of the protein complex are currently large limitations in routine high-resolution structure determination of proteins ¹⁴⁶.

4 TRYPTOPHAN HYDROXYLASE 2 (MANUSCRIPT I)

This chapter concerns a structural study of *human* tryptophan hydroxylase 2. The study investigates an N-terminally truncated variant of the isolated regulatory domain using NMR spectroscopy, and an N-terminally truncated variant of the complete tetrameric enzyme using cryo-EM, and focuses on ligand binding to the regulatory domain. The results are presented in the form of a manuscript (under preparation).

NMR experiments were carried out at the Department of Biology, University of Copenhagen in collaboration with Professor Birthe B. Kragelund and NMR manager Andreas Prestel. Cryo-EM experiments were conducted at the Department of Structural Dynamics at the Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany in collaboration with Professor Holger Stark and Ph.D. Zhenwei Zhang. The cryo-EM experiments were conducted and analyzed by Zhenwei Zhang.

Supporting information is provided at the end of the chapter

In relation to the study, additional experiments have been conducted that are not included in the manuscript. A selected set of these experiments are provided in the appendix, with the hope that the results can be useful for future students. This includes ^1H - ^{15}N HSQC spectra of NΔ47-TPH2-R in the presence of possible ligands (Appendix D) and kinetic experiments of TPH2-CT, NΔ47-TPH2, and TPH2 in the presence of L-Phe (Appendix E).

Structural characterization of L-Phe binding to a potential allosteric site in human tryptophan hydroxylase 2

Ida M. Vedel¹, Andreas Prestel², Zhenwei Zhang³, Natalia T. Skawinska¹, Holger Stark³, Pernille Harris⁴, Birthe B. Kragelund^{2}, Günther H. J. Peters^{1*}*

¹*Department of Chemistry, Technical University of Denmark, Kemitorvet, 2800 Kgs. Lyngby, Denmark*

²*Department of Biology, University of Copenhagen, Ole Maaløes vej 5, 2200 Copenhagen N, Denmark*

³*Department of Structural Dynamics, Max Planck Institute for Biophysical Chemistry, Am Faßberg 11, 37077 Göttingen, Germany*

⁴*Department of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark*

Corresponding authors

**Birthe B. Kragelund, bbk@bio.ku.dk*

**Günther H. J. Peters, ghp@kemi.dtu.dk*

Abstract

Tryptophan hydroxylase 2 (TPH2) catalyzes the rate-limiting step in the biosynthesis of serotonin in the brain. Consequently, regulation of TPH2 is relevant for serotonin related diseases, yet, TPH2's regulatory mechanism is poorly understood. Here, we use NMR spectroscopy to determine the structure of a 47 N-terminally truncated variant of the regulatory domain (RD) dimer of *human* TPH2 in complex with L-Phe, and show that L-Phe is the preferred RD ligand compared to the natural substrate, L-Trp. Using cryo-EM we obtain a low-resolution structure of a similarly truncated variant of the complete tetrameric enzyme with dimerized RDs. The cryo-EM 2D class averages additionally indicate that the RDs are dynamic in the tetramer and likely exist in a monomer-dimer equilibrium. Our results provide structural information on the RD both as an isolated domain and in the TPH2 tetramer, which will facilitate future elucidation of TPH2's regulatory mechanism.

Introduction

Tryptophan hydroxylase (TPH, EC 1.14.16.4) is a non-heme, iron-dependent enzyme that catalyzes the first and rate-limiting step in the biosynthesis of the neurotransmitter and peripheral hormone, serotonin. TPH hydroxylates L-Trp into 5-hydroxy-L-Trp (5-HTP) using Fe(II) as co-factor and molecular oxygen and tetrahydrobiopterin (BH₄) as co-substrates. 5-HTP is subsequently decarboxylated by aromatic amino acid decarboxylase to form serotonin¹. There are two isoforms of TPH, independently regulating each their respective serotonin system²: TPH1, mainly expressed in the gut and in the pineal gland, and TPH2, mainly expressed in the brain³. Irregular levels of serotonin in the peripheral system (TPH1 associated) are related to e.g. irritable bowel syndrome and ulcerative colitis diseases⁴, while irregular levels of serotonin in the brain (TPH2 associated) have been connected to e.g. schizophrenia⁵ and major depression⁶, highlighting TPH as a relevant drug target.

TPH constitutes together with phenylalanine hydroxylase (PAH, EC 1.14.16.1) and tyrosine hydroxylase (TH, EC 1.14.16.2) the family of aromatic amino acid hydroxylases (AAAH)⁷. As TPH, PAH and TH catalyze the hydroxylation of their respective amino acid, L-Phe and L-Tyr. The AAAHs form homo-tetramers with each monomer consisting of an N-terminal regulatory domain (RD), a catalytic domain (CD), and a small tetramerization domain (TD, Figure 1A)⁸⁻¹¹. Of these, the RDs are the most divergent with relatively low sequence identity among the enzymes, while the CDs are very homologous¹². Between TPH1 and TPH2, the RDs are 48% identical (Figure 1A), and TPH2 contains an additional 46 N-terminal residues (N-terminal tail) in its RD associated with instability^{13,14}. The N-terminal tail contains an isoform specific phosphorylation site (Ser19) known to activate the enzyme upon phosphorylation and subsequent binding to 14-3-3 protein¹⁵. The divergence between the RDs of TPH1 and TPH2 is promising in regards to selective targeting of one isoform over the other. However, to date there is no detailed structure available of either the RD of TPH2 or TPH1 and their exact regulatory mechanisms are yet to be determined. Therefore, insight into the RDs and their function is crucially needed, to not only enable selective drug targeting, but also to further elucidate how serotonin levels in the body are regulated.

Current knowledge on the RD of TPH2 is extrapolated from PAH as its structure-function relationship is relatively well understood: PAH is allosterically activated by its substrate L-Phe, which has been confirmed by multiple kinetic studies¹⁶⁻¹⁹. Structurally, the activation can be described by an equilibrium between a resting state (RS) and an active state (AS) with L-Phe bound²⁰, as shown in Figure 1B. In the RS, the RDs are docked between two CDs, and residues from the N-terminal tails hinders substrate access to the active site^{9,20,21}. In the AS, the RDs of two diagonally opposed monomers form a dimer with 2 L-Phe molecules bound at the dimer interface²², resulting in the N-terminal tails moving away from the catalytic sites activating the enzyme^{20,23}. In solution, the isolated RD of PAH exists in an L-Phe affected monomer-dimer equilibrium^{22,24}. A similar equilibrium was observed by Tidemand et al., for an N-terminally truncated TPH2 variant containing the RD and CD (NΔ47-TPH2-RC)¹⁴, and differential scanning fluorimetry (DSF) showed that both L-Phe and L-Trp have a stabilizing effect on the variant indicating the presence of an RD binding site¹⁴. This is supported by a hydrogen/deuterium exchange study of NΔ47-TPH2-RC which has pointed to the dimer

interface as a location of binding similar to PAH (Skawinska et al., in preparation ²⁵). However, kinetic studies have not reported substrate induced activation of TPH2 ^{26,27}, indicating that TPH2's regulatory mechanism is different to PAH's.

In this study, we structurally characterize the regulatory domain of *human* TPH2 and L-Phe binding to the domain. Using solution nuclear magnetic resonance (NMR) spectroscopy, we determine the structure of a 47 N-terminally truncated variant of the isolated RD (NΔ47-TPH2-R) in complex with L-Phe, and we observe that L-Phe binds stronger to NΔ47-TPH2-R than L-Trp. Furthermore, using cryo-electron microscopy (cryo-EM), we show that the RDs mainly form dimers in tetrameric TPH2 in a structure similar to the AS of PAH and additionally observe that the RDs are dynamic likely existing in a monomer-dimer equilibrium. Combined, these results point to L-Phe as the preferred ligand of the isolated RD, and additionally provide a structural scaffold that will be important for future studies of TPH2 regulation.

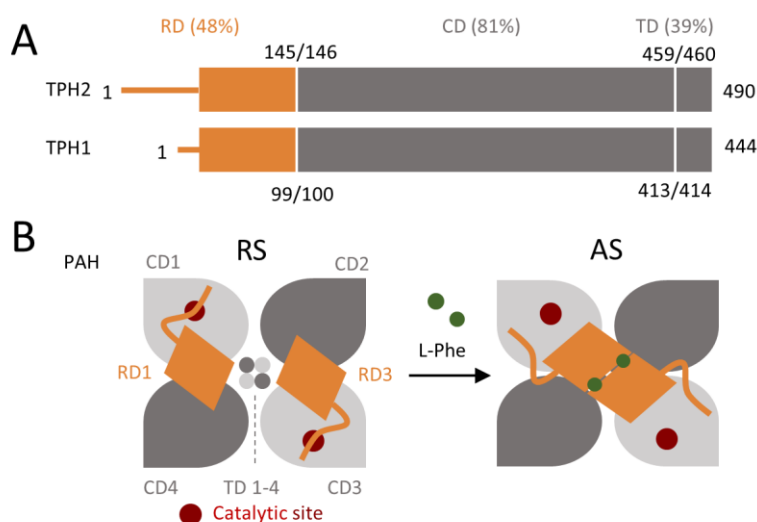


Figure 1 A) Domain alignment of human TPH2 and human TPH1 indicating the relative sizes of the regulatory domains (RDs), the catalytic domains (CDs), and the tetramerization domains (TDs). The sequence identities are indicated above and were determined using Clustal Omega ⁶⁴, and numbers (in black) indicate which residue positions define the domain borders. Despite a low sequence identity, the TDs are expected to be structurally similar consisting of a single α -helix, which in the tetramer forms a 4- α -helix bundle with the other TDs ⁷ B) Illustration of the activation of PAH by binding of L-Phe, made with inspiration from ²⁰. L-Phe shifts the conformation from the resting state (RS) to the active state (AS)²⁰. The RS and AS have been confirmed by multiple crystal structures ^{9,20,21} and small angle X-ray scattering (SAXS) data^{20,23}.

Results

The isolated regulatory domain of TPH2 is a dimer and L-Phe increases its thermal stability

To investigate if NΔ47-TPH2-R exists in a similar L-Phe-affected monomer-dimer equilibrium as NΔ47-TPH2-RC, analytical size exclusion chromatography was run at various concentrations of NΔ47-TPH2-R in the presence and absence of L-Phe (10 mM). L-Phe was used as the ligand since the natural substrate L-Trp interferes with the absorption at 280 nm. In the concentration range of 3-100 μ M, NΔ47-TPH2-R eluted as a dimer for all protein loading concentrations and independent of the addition of L-Phe (Figure 2A-B). This is similar to the behavior of the isolated RD of TPH1, which also has been shown to exist as a dimer at protein concentrations above 2.7 μ M²⁸. Although L-Phe did not affect the oligomeric state of NΔ47-TPH2-R, nanoDSF measurements showed that L-Phe had a stabilizing effect on the dimeric complex with an increase in melting temperature (T_m) of up to 18 °C, consistent with the presence of a binding site as previously proposed¹⁴ (Figure 2C). For the isolated RD of PAH (PAH-R) a similar increase in melting temperature of ~20 °C has been observed in the presence of 1 mM L-Phe²².

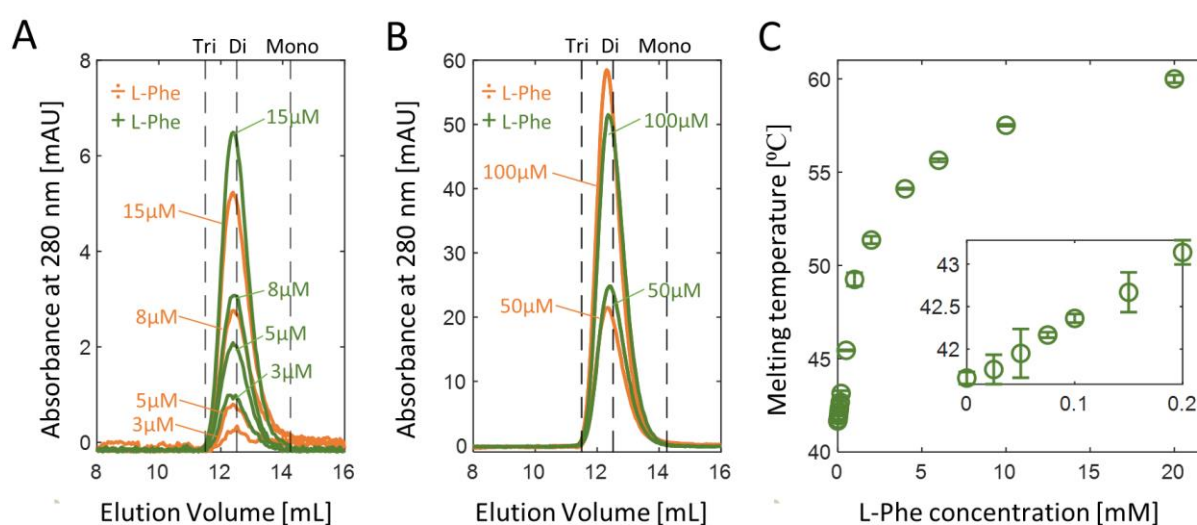


Figure 2 SEC runs with A) low protein concentrations (3-15 μ M) and B) high protein concentrations (50-100 μ M) with 0 mM L-Phe (orange) and with 10 mM L-Phe (green). The dashed lines indicate the expected elution volume of a monomer (Mono), dimer (Di), and trimer (Tri). C) The effect of L-Phe on T_m of NΔ47-TPH2-R determined by nanoDSF. The mean of triplicate measurements (green circles) is shown with error bars representing the standard deviation. The insert shows a zoom-in of the low L-Phe concentration range.

Solution structure of NΔ47-TPH2-R in complex with L-Phe reveals a binding site similar to the allosteric site of PAH

While free NΔ47-TPH2-R aggregated over time even at 4 °C, the presence of L-Phe in the buffer increased T_m and caused the domain to be stable at room temperature, which facilitated a detailed structural characterization. Here, we used solution NMR spectroscopy to obtain the three-dimensional structure of the isolated RD of TPH2 in complex with L-Phe.

A ^1H - ^{15}N HSQC spectrum of NΔ47-TPH2-R with L-Phe (Extended data figure 1A) showed 84 distinct backbone peaks out of the expected 96, confirming that the isolated RD forms symmetric dimers like the RDs of both PAH and TH^{22,29}. Additionally, an aromatic ^1H - ^{13}C HSQC spectrum of ^{13}C -labeled L-Phe and unlabeled NΔ47-TPH2-R, showed three peaks from the free L-Phe and three peaks from bound L-Phe, confirming that there is only one type of binding site (Extended data figure 1B). The quality of the L-Phe:NΔ47-TPH2-R NMR spectra enabled assignment of the majority of resonances of the complex and allowed for structure calculations. Inter- and intramolecular nuclear Overhauser effects (NOEs) were obtained from semi-automated NOE assignment providing 2171 distance restraints for the structure calculations (Extended data table 1). During the NOE assignment, P2 symmetry was imposed on both the NΔ47-TPH2-R dimer and the two L-Phe ligands. A total of 100 structures were calculated and refined with P2 symmetry imposed on the folded region, and the 20 lowest energy structures selected to represent the L-Phe:NΔ47-TPH2-R structure (Extended data table 1). The structure showed that the RD of TPH2 contains a folded core with $\beta\alpha\beta\beta\alpha$ topology ($\beta 1$: 65-71, $\alpha 1$: 76-85, $\beta 2$: 90-98, $\beta 3$: 105-112, $\alpha 2$: 116-129) and disordered N- and C-terminal tails (Figure 3A-B). The two monomers are arranged side-by-side forming an extended six stranded β -sheet. Although two residues (Thr134 and Leu135) show β -strand propensity based on their chemical shifts (Extended data figure 1C), a fourth β -strand is not fully formed (Figure 3B-C). Backbone ^{15}N relaxation rates of L-Phe:NΔ47-TPH2-R support a folded core and flexible N- and C-terminal tails (Extended data figure 2). Additionally, modelling the N- and C-terminal residues (46-60 and 136-145) using SAXS data resulted in a good fit ($\chi^2=0.39$, extended data figure 3 and extended data table 2), increasing the confidence in the NMR structure. The folded core confirms that the RD of TPH2 contains an ACT domain (aspartate kinase, chorismate mutase, and TyrA domain), a domain found in several proteins involved in amino acid metabolism including PAH and TH³⁰. The domain is characterized by $\beta\alpha\beta\beta\alpha\beta$ topology, and the ACT domain often has an allosteric function in its parent enzyme by ligand binding to the domain³⁰. In some ACT domain enzymes ligand binding is observed to cause large conformational changes as the binding of L-Phe to the RD of PAH, while more subtle allosteric changes occur in other ACT domain enzymes^{30,31}.

The two L-Phe ligands are bound at the dimer interface in hydrophobic pockets. Pocket 1 is formed by the side chains of Leu70, Leu77, Ser96, Val106, and Ile108 from monomer 1 and Phe84, Met91, and Ile94 from monomer 2 (Figure 3D). Additionally, the backbone atoms of residues 71-76 (monomer 1) and residues 90-93 (monomer 2) are in close proximity to the ligand, and the side chains of Asn72 and Asn90 are in positions to form hydrogen bonds to the amine group of L-Phe. Due to symmetry, pocket 2 is constituted by identical residues but from opposite monomers. The distance restraints for orienting L-Phe in the binding pockets

were determined from a $^{13}\text{C}/^{15}\text{N}$ -filtered ^{13}C -NOESY-HSQC spectrum, which exclusively yields NOESY cross peaks from ^{13}C -labeled protein to unlabeled L-Phe. Five residues (Leu77, Met91, Ile94, Val106, Ile108) showed distinct cross peaks to the ligand (Extended data figure 4), establishing these residues as important parts of the L-Phe binding pockets. The binding

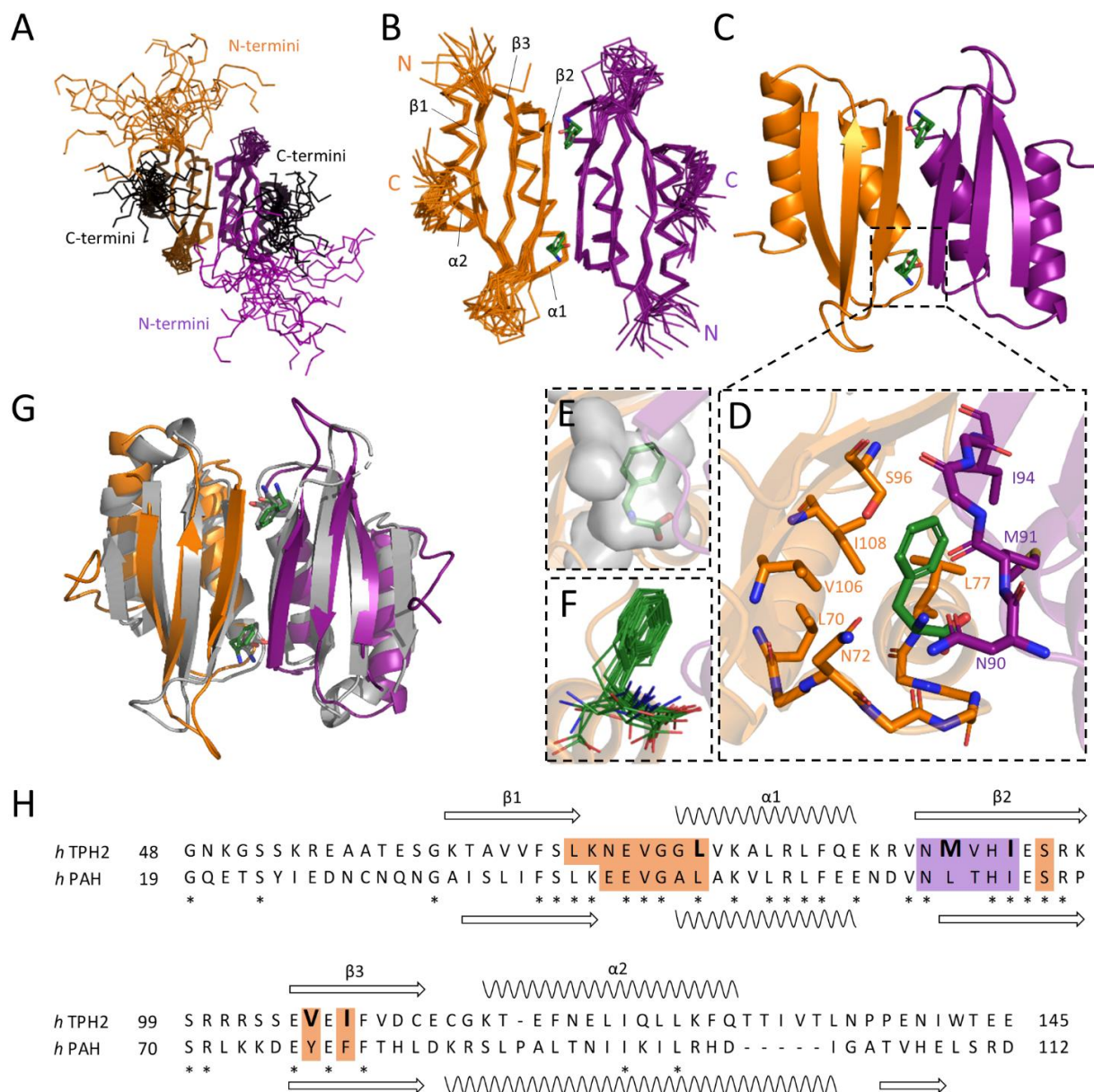


Figure 3 A) Superimposition of the 20 lowest energy conformers of the calculated NMR ensemble showing all residues (46-145) of monomer 1 (orange) and monomer 2 (purple). B) Superimposition of the 20 lowest energy conformers of the calculated NMR ensemble showing only the folded core of monomer 1 (orange) and monomer 2 (purple). L-Phe ligands are displayed as green sticks. C) Representative conformer of the calculated ensemble displayed as cartoon with monomer 1 (orange), monomer 2 (purple), and L-Phe ligands (green sticks). D) zoom-in on the residues in the binding pocket of the conformer in C. E) The void available in the binding pocket displayed as a grey blob (Illustrated using PyMOL 2.3.3 (Schrödinger), standard settings). F) Overlay of the position of L-Phe in the 20 lowest energy conformers. G) Superimposition of the conformer in C (orange/purple) with the crystal structure of human PAH-R (grey, PDB: 5FII). L-Phe bound in TPH2 (green sticks), L-Phe bound in PAH (grey sticks) H) sequence alignment between residues 48-145 of human TPH2 and residues 19-112 of human PAH. The secondary structure elements of N447-TPH2-R are marked above the alignment. Residues involved in the binding of L-Phe in the dimer interface are marked with orange (monomer 1) and purple (monomer 2). Residues that showed cross peaks in the filtered ^{13}C -NOESY-HSQC are highlighted in bold. * marks identical residues.

pockets are small ($164 \pm 41 \text{ \AA}^3$ ¹) with only little excess space available to the ligands, resulting in a tight fit (Figure 3E). The ligands are positioned very similarly in all 20 models, increasing the confidence in their position (Figure 3F).

A comparison of the structure of NΔ47-TPH2-R to the crystal structure of PAH-R (PDB: 5FII) showed a relatively high structural identity, including the position of the L-Phe ligands (Figure 3G). A sequence alignment of the two regulatory domains (Figure 3H) showed that they are just 33% identical. However, the identical residues are clustered around the dimer interface where the binding pockets are located, a feature which has likely conserved the ability of TPH2 to bind L-Phe. Indeed, mapping the binding sites from each structure onto the sequence alignment showed that the binding pockets of L-Phe are located at almost the exact same place in the two enzymes, though the residues involved in binding are not completely conserved.

L-Phe binds stronger to the regulatory domain of TPH2 than L-Trp

To investigate the binding of both L-Phe and L-Trp to the RD of TPH2, ¹H-¹⁵N HSQC spectra of NΔ47-TPH2-R were recorded in the presence and absence of either L-Phe or L-Trp. The spectra show that the addition of 1 mM L-Phe causes large global changes to the spectrum, while addition of 20 mM L-Trp only causes very small chemical shift perturbations (CSPs) (Figure 4A-B). The gradual movement of the peaks in the L-Trp spectra from 0 mM-3 mM-20 mM shows that L-Trp binds in fast exchange and indicates a K_d in the mM range. In contrast, an NMR titration with L-Phe showed that L-Phe binds with slow exchange kinetics and that the complex is saturated at a concentration of 1 mM L-Phe (Extended data figure 5A-B). This shows that L-Phe binds stronger to the RD dimer than L-Trp, despite L-Trp being the natural substrate of the enzyme. Due to the slow exchange kinetics, the assigned peaks from the L-Phe:NΔ47-TPH2-R complex could not be traced through the titration, and the spectrum of free NΔ47-TPH2-R was therefore assigned to enable calculation of CSPs and K_d determination by 2D NMR line-shape analysis. However, the aggregation propensity of free NΔ47-TPH2-R limited the feasible experimental conditions both in relation to temperature and protein concentration and thus resulted in low quality 3D spectra. Nonetheless, it was possible to assign 75 out of 93 peaks and CSPs were calculated and mapped onto the NMR structure (Figure 4C-F). For both L-Phe and L-Trp binding, almost all of the identified residues have some CSP, but with vast differences in amplitudes, where L-Phe binding causes around ten times larger CSPs than binding of L-Trp. It should be noted that the highest L-Trp concentration used was 20 mM, and it is possible that the L-Trp:NΔ47-TPH2-R complex is still not saturated. In both cases, the largest CSPs occur at the dimer interface where the binding pockets are located. For L-Phe binding there are especially large CSPs around loop 1 and loop 2 (L1 and L2), while L-Trp causes the largest CSPs around L2. It is possible that L-Trp is too large to fit properly into the binding pockets or that particular chemical complementarity is missing to anchor L-Trp in the pockets, giving rise to weaker binding and different CSP profile. Despite the large differences in the ¹H-¹⁵N HSQC spectra of free and L-Phe bound NΔ47-

¹ The volume is calculated based on the 5 lowest energy structures using SiteMap (Schrödinger)

TPH2-R, R_1 and R_2 backbone ^{15}N relaxation rates and secondary structures based on secondary chemical shifts of the free and bound form are all similar, indicating that the overall size and shape of NΔ47-TPH2-R remains the same (Extended data figure 6).

The titration of NΔ47-TPH2-R with L-Phe, revealed peaks from two different states: The initial peaks from the free NΔ47-TPH2-R dimer and the final peaks from the NΔ47-TPH2-R dimer with two L-Phe molecules bound (Extended data figure 5B-C). Despite that a third state, the NΔ47-TPH2-R dimer with only one L-Phe bound was expected to exist, its population was too low to result in any observable peaks. Therefore, the titration data was fitted to a two state

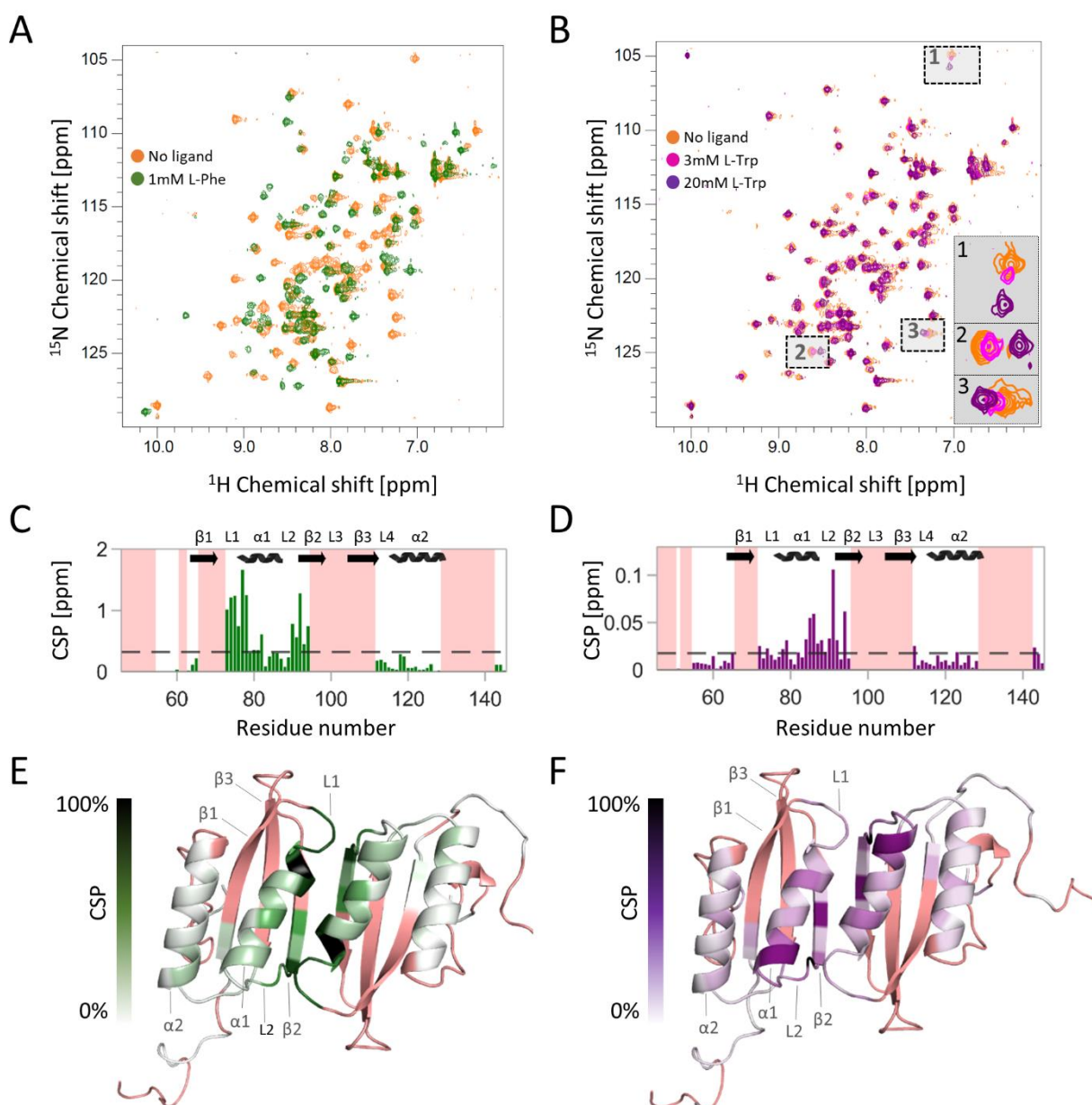


Figure 4 A) ^1H - ^{15}N HSQC of NΔ47-TPH2-R in the presence and absence of L-Phe (1 mM) B) ^1H - ^{15}N HSQC of NΔ47-TPH2-R in the presence and absence of L-Trp (3 mM and 20 mM). Inserts show zoom-in of areas 1, 2, and 3. C) Chemical shift perturbations caused by L-Phe binding. D) Chemical shift perturbations caused by 20 mM L-Trp binding. E) L-Phe perturbations mapped on the NΔ47-TPH2-R structure F) L-Trp perturbations mapped on the NΔ47-TPH2-R structure.

binding model using the 2D line-shape analysis program TITAN³² (Extended data figure 5C). The resulting K_d was $24.8 \pm 0.6 \mu\text{M}$ and k_{off} $0.03 \pm 0.05 \text{ s}^{-1}$ (see Table 1). As the exchange rate is very slow, the k_{off} is at the limit of what can be reliably determined by line-shape analysis, and the k_{off} is therefore associated with a relatively large error. However, k_{off} is reliably small and likely to be less than 1 s^{-1} . To increase the confidence in the determined K_d , another titration was performed, by varying the concentration of unlabeled protein at a fixed concentration of ^{13}C -labeled L-Phe. At low protein concentrations, the recorded aromatic ^1H - ^{13}C HSQC spectra showed only three peaks arising from the free L-Phe. Only at the two highest protein concentrations, three peaks from bound L-Phe started to appear. Therefore, despite expected differences in R_2 between bound and free L-Phe, a global fit was performed to the decreasing intensities of the three free L-Phe peaks (Extended data figure 7). The K_d was determined to $99 \pm 8 \mu\text{M}$, which is larger than the one determined by line-shape analysis ($24.8 \pm 0.6 \mu\text{M}$), but still within the same order of magnitude, supporting a K_d in the μM range.

Table 1 Calculated dissociation constants (K_d) of the L-Phe:NΔ47-TPH2-R complex

	K_d (μM)	k_{off} (s^{-1})
Protein observed titration (<i>line-shape analysis</i>)	24.8 ± 0.6	0.03 ± 0.05
Ligand observed titration (<i>Intensity fit</i>)	99 ± 8	NA

Cryo-EM images show dynamic regulatory domains in tetrameric TPH2

To address the structural behavior of the RD of TPH2 in a more native context, we next used cryo-EM to investigate an N-terminally truncated TPH2 variant containing the RD, the CD, and the TD (NΔ47-TPH2). Cryo-EM grids with this variant were set up both without L-Phe and, as large amounts of free L-Phe lower the image contrast, with $50 \mu\text{M}$ L-Phe, thus not reaching full saturation of the complex.

3D reconstruction of the NΔ47-TPH2 particles from the grids with L-Phe resulted in a tetrameric structure, with the four CDs and TDs creating an x-shaped planar core, and RDs from two diagonally opposed monomers forming dimers above and below the plane (Figure 5A). The dimeric conformation of the RDs is in agreement with a previous SAXS model of NΔ47-TPH2 in the presence of L-Phe (Skawinska et al., in preparation²⁵). While reconstruction of just the CDs and TDs reached 3.9 Å (Figure 5B and Extended data table 3), the complete structure was limited at 8.9 Å (ranging from $6\text{--}9 \text{ Å}$ of the CDs to $15\text{--}20 \text{ Å}$ of the RDs). However, our NMR structure of the RD dimer and the crystal structure of the CD and TD (PDB 4V06) could be well fitted into the density. The structure reveals a gap of $\sim 10 \text{ Å}$ between the RDs and CDs (Figure 5A, zoom in) likely causing a considerable degree of flexibility between the domains that could have contributed to the relatively low resolution.

Assessing the 2D class averages of side view images (Figure 5C), showed three different particle types: The most abundant, P2, with RD density on both sides of the CDs, P1, with density on one side, and P0, with no visible RD density. The missing RD density in P1 and P0 indicates that the RDs are dynamic, and likely move randomly relative to the CDs when they are not dimerized. The dynamic behavior of the RDs is consistent with the monomer-dimer equilibrium of NΔ47-TPH2-RC observed by Tidemand et al ¹⁴. Comparing the 2D class averages from the grids without L-Phe (Figure 5D) to the 2D classes with 50 μM L-Phe, did not reveal any significant differences in the overall structure, (Figure 5C-D). The top view images confirm that NΔ47-TPH2 is a tetrameric complex both with and without L-Phe, and the side view images revealed the same three particle types (P2, P1, and P0) in similar abundances. Thus, either L-Phe does not affect the RDs in tetrameric TPH2, or 50 μM L-Phe is not enough to fully induce RD dimerization. In support of the latter, L-Phe induces dimerization of NΔ47-TPH2-RC ¹⁴, and stabilizes both NΔ47-TPH2-RC ¹⁴ and NΔ47-TPH2-R observed in this study. It is therefore likely that a higher L-Phe concentration would have an effect on RD dimerization. We additionally set up grids with full-length TPH2, in an attempt to get structural information on the N-terminal tail (residues 1-46). However, the 2D class averages were of low quality, likely due to the aggregation propensity of full-length TPH2, and

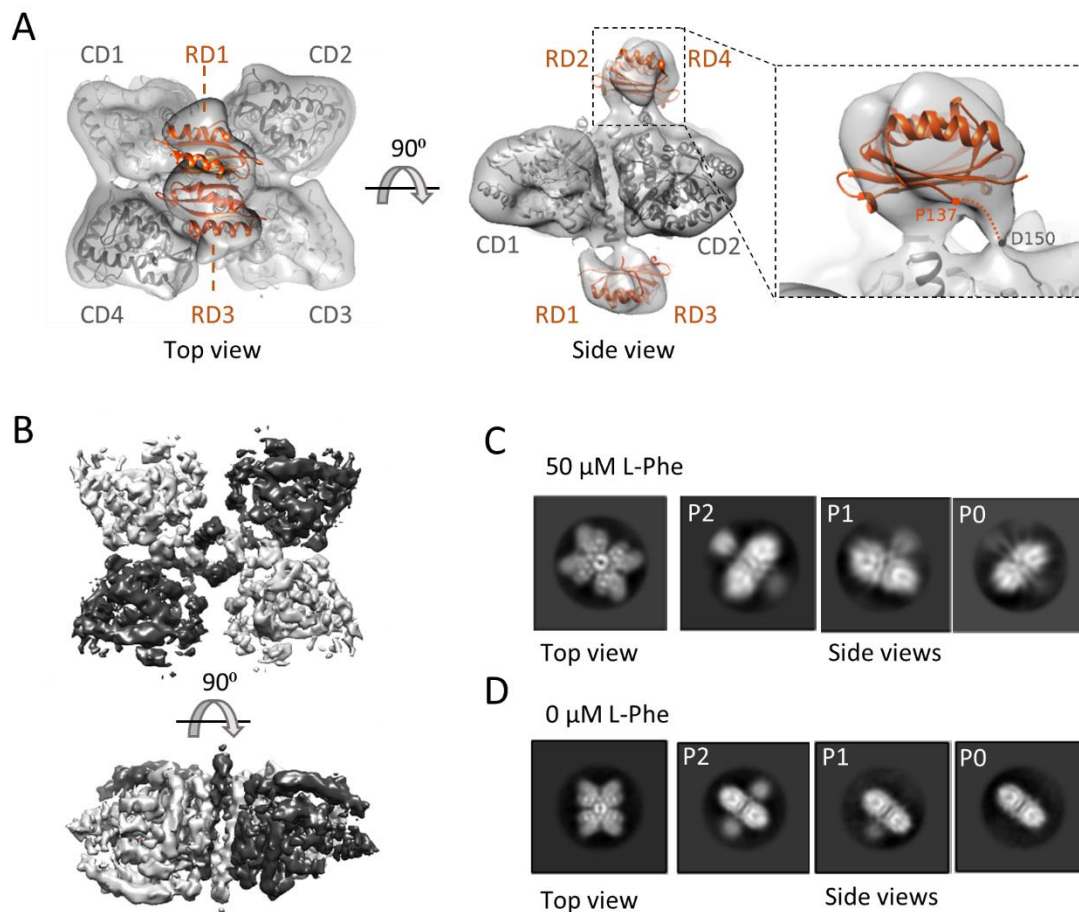


Figure 5 A) 3D reconstruction of NΔ47-TPH2 (Particle type P2) with 50 μM L-Phe. 8.9 Å EM map resolution. B) 3D reconstruction of the CDs and TDs of NΔ47-TPH2. 3.9 Å EM map resolution. C) Top view 2D class and side view 2D classes of the three different particle types of NΔ47-TPH2 with 50 μM L-Phe. D) Top view 2D class and side view 2D classes of the three different particle types of NΔ47-TPH2 without L-Phe.

thus it would not be possible to reach a resolution were the density of the tail could be clearly identified. The 2D classes still showed similar particle types (Extended data figure 8) indicating that the N-terminal tail does not strongly affect the conformation of the RDs in tetrameric TPH2.

Discussion

In this work, we observe that NΔ47-TPH2-R exists as a dimer independent of the presence of L-Phe, whereas Tidemand et al. found that when the CD is included (NΔ47-TPH2-RC), the enzyme is in an L-Phe affected monomer-dimer equilibrium¹⁴. Combined, this indicates that the presence of the CD weakens the RD dimer interface, highlighting the need for supplementary studies of full-length TPH2 to fully elucidate L-Phe binding to the RD. Nonetheless, the NMR structure of the L-Phe:NΔ47-TPH2-R complex presented in this study is the first structure of the isolated RD of TPH2 and it reveals the exact location of the RD binding sites, valuable for further characterization of the regulatory mechanism of TPH2. The binding sites are very similar to the allosteric sites of PAH, possibly resulting in L-Phe binding stronger to the NΔ47-TPH2-R dimer than the natural substrate L-Trp. While the K_d of the L-Phe:NΔ47-TPH2-R complex is in the μM -range (25-100 μM), our NMR spectra indicate a K_d of the L-Trp:NΔ47-TPH2-R complex in the mM-range. The K_d of L-Phe:PAH-R has been determined to 8 μM ²⁴ and 15 μM ²⁰, and thus L-Phe binds slightly weaker to NΔ47-TPH2-R than to the isolated RD of PAH. The exact physiological levels of L-Phe in the brain of healthy individuals is, to the best of our knowledge, not known. However, Waisbren et al. found that elevated brain L-Phe levels of nine adults with early treated phenylketonuria (PKU, elevated L-Phe levels caused by a deficiency of PAH) ranged from 76-185 μM ³³, whereas elevated brain L-Phe levels from 140-780 μM have been reported in other PKU patients³⁴. Thus, the K_d of L-Phe:NΔ47-TPH2-R reported in this study is within a physiological relevant range. Yet, we acknowledge that the K_d for both L-Phe and L-Trp is likely different for full-length TPH2, where the monomer-dimer equilibrium of the RD is shifted, as observed in the cryo-EM images. Still, our results show that L-Phe is preferred to L-Trp as a ligand for the isolated RD of TPH2.

Additionally, future elucidation of the regulatory mechanism of TPH2 will be facilitated by the cryo-EM data reported here. Our data indicate that the RDs are dynamic in tetrameric TPH2, likely existing in a monomer-dimer equilibrium as illustrated in Figure 6, where RDs either form motionally restricted dimers or move randomly relative to the CDs as monomers in “undocked states”. The term undocked states was used by Khan et al. and Arturo et al. to describe the intermediate conformations between the RS and AS in PAH, where RD monomers are neither interacting with the CDs or another RD^{17,35}. Importantly, we do not observe any particles resembling the RS of PAH, but find a large part of the particles with dimerized RDs even without L-Phe in a conformation similar to the AS of PAH. Thus indicating that the conformational landscape of TPH2 is different to that of PAH. For PAH and TH, the N-terminal tail is creating an auto-inhibitory state by occluding the active site. While the N-tail of PAH is displaced by dimerization of the RDs induced by L-Phe binding²⁰, the N-tail of TH can reach

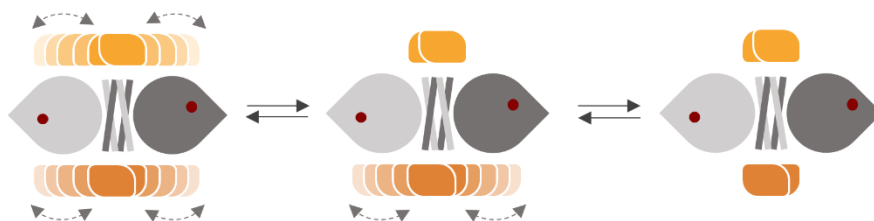


Figure 6 Schematic illustration of the proposed conformational landscape of TPH2, investigated using cryo-EM. The RDs (orange) can exist either as dynamic monomers randomly moving relative to the CDs (grey), or form more motionally restricted dimers.

and cover the active site with the RDs dimerized, and the tail is instead displaced upon phosphorylation of Ser40³⁶. The N-tail of TPH2 is, similar to TH, longer than that of PAH. It is thus possible that residues in the end of the tail can cover the active site in a state with dimerized RDs, as also discussed by Zhang et al.²⁸. Although with low confidence, the Alpha-fold 2 predicts residues 1-21 of TPH2 to cover the entrance to the active site (Extended data figure 9)³⁷, suggesting at least that this region has a propensity to interact with the CD. This is consistent with activation by phosphorylation of Ser19 and 14-3-3 binding observed by Winge et al.¹⁵, which could cause displacement of the N-tail. However, the binding of L-Phe to the RD dimer of TPH2 raises the question if TPH2 is also allosterically affected by L-Phe as PAH. Although we did not observe changes in the 2D classes of NΔ47-TPH2 in the presence of 50 μM L-Phe, the binding of L-Phe stabilized the RD (Figure 2C), and L-Phe affected the monomer-dimer equilibrium of TPH2-RC observed by Tidemand et al.¹⁴, which jointly indicates that increasing concentrations of L-Phe would induce RD dimerization in tetrameric TPH2. L-Phe is already known to influence serotonin production to some extent, as reduced serotonin levels have been reported in PKU patients and is associated with some of the neurological symptoms of PKU^{38,39}. The mechanism of serotonin depletion in relation to PKU is not completely understood, but is hypothesized to be caused by either a lowered transport of L-Trp into the brain or increased competitive inhibition of L-Phe in the active site of TPH2 at elevated L-Phe levels^{38,39}, both resulting in reduced L-Trp hydroxylation. Thus, L-Phe binding to the RD could be an activation mechanism of TPH2, to keep a steady serotonin production regardless of minor increases in the L-Phe level. Though using TPH2 with an N-terminal MBP-tag expected to influence the N-tail, Ogawa et al. observed an inhibitory effect of L-Phe on TPH2 activity⁴⁰, indicating that L-Phe alone does not activate TPH2. Yet L-Phe RD binding could have a stabilizing effect, as also observed in this and in an earlier study¹⁴, potentially influencing e.g. phosphorylation of the RD and subsequent 14-3-3 binding. However, with the current knowledge available on TPH2, a function of L-Phe binding to the RD remains speculative. Future experiments mapping the effect of L-Phe on TPH2 kinetics and phosphorylation will therefore be important to elucidate if there is a functional effect of ligand binding to the RD. Nonetheless, the current study has shown L-Phe to be the preferred ligand of the RD of TPH2 compared to L-Trp, and shown that L-Phe affects the thermal stability of the RD dimer. Importantly, the study has provided detailed structural information on the RD of TPH2, which will be valuable in the elucidation of TPH2- and consequently serotonin regulation.

Materials and Methods

Protein Expression

An LB agar plate with 30 μ M/mL kanamycin sulfate (30 Kan) was streaked with a frozen stock of an *E. Coli* BL21 (DE3) strain containing the relevant MBP-tagged TPH2 construct (either MBP-N Δ 47-TPH2-R, MBP-N Δ 47-TPH2, or MBP-TPH2). The plate was incubated at 37 °C overnight. A single colony from the plate was used to inoculate 50 mL LB 30 Kan media in a 250 mL shaking flask and was incubated at 37 °C and 140 rpm shaking for around 4 h until the OD₆₀₀ was 0.6-1.0. The cells were pelleted by centrifugation (1800 g, 4 °C, 10 min) and resuspended in 50 mL fresh LB 30 Kan media. 650 mL LB 30 Kan media in a 2 L shaking flask was then inoculated with 6.5 mL of the resuspended cells and incubated (30 °C, 140 rpm) for around 3h until the OD₆₀₀ was 0.3. At this point, the temperature was lowered to 20 °C, the shaking set to 240 rpm, and the incubation continued for around 1.5 h until the OD₆₀₀ was 0.6. Expression was induced by addition of isopropyl- β -D-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM. If the catalytic domain was part of the construct (MBP-N Δ 47-TPH2 or MBP-TPH2), 650 μ L of a sterile filtered freshly prepared (NH₄)₂Fe(SO₄)₂ solution was added to a final Fe(II) concentration of 0.2 mM. The culture was incubated for 16 h over-night (20 °C, 240 rpm) and cells harvested by centrifugation (4 °C, 3000 g, 15 min). The cells were resuspended in 20 mL specific binding buffer (Table 2), transferred to a 50 mL falcon tube, pelleted by centrifugation (4 °C, 3000 g, 15 min), and stored at -80 °C until purification. For expression of ¹⁵N- and ¹⁵N/¹³C-labeled MBP-N Δ 47-TPH2-R for NMR experiments, the procedure was essentially the same but with the following modifications: two pre-cultures (50 mL LB 30 Kan media) were prepared and at OD₆₀₀ 0.6-0.1, the cells were pelleted and resuspended in 20 mL 30 Kan M9 minimal media (22 mM KH₂PO₄, 53 mM Na₂HPO₄, 86 mM NaCl, 1 mM MgSO₄, 1:1000 V/V M2 trace element solution, 22 mM glucose, 19 mM NH₄Cl). If needed the glucose was ¹³C labeled and the NH₄Cl ¹⁵N-labeled. The resuspended cells from both pre-cultures were used to inoculate 650 mL 30 Kan M9 minimal media in a 2L shaking flask and incubated as described above.

Protein purification

To purify the N Δ 47-TPH2-R, frozen cells from a 650 mL culture containing MBP-N Δ 47-TPH2-R were thawed and re-suspended in 40 mL specific binding buffer (Table 2). The cells were lysed while kept in ice-water using a Labsonic P sonicator (Sartorius), with 80% amplitude and 1 cycle for 3x30 s with 1 min cooling in-between. The sample was centrifuged (2x20 min, 4 °C, 16000 g), and the supernatant recovered and filtered through a 0.45 μ m syringe filter (Whatman™). The supernatant was loaded onto a Dextrin Sepharose affinity column (GE Healthcare) pre-equilibrated with 5 CV of binding buffer, using an ÄKTA start chromatography system (GE Healthcare). The flow rate was 5 mL/min, and the progress of the purification was followed by UV absorbance at 280 nm. After loading the sample, un-bound proteins were removed with 5-6 CV of binding buffer. Hereafter, 30 mL of 2.2 μ M MBP-tagged 3C protease (MBP-3CP, expressed and purified in-house, and diluted from a stock into binding buffer) was

loaded on the column, and the system was paused for 1 h to ensure sufficient cleavage, leaving two additional residues from the linker on the N-terminal of the protein (Gly-Pro). After cleavage, NΔ47-TPH2-R was eluted from the column using binding buffer and ~8 mL protein collected. Elution buffer (binding buffer + 1 mM Maltose) was run through the column to elute MBP-3CP and cleaved MBP. The collected 8 mL protein was filtered through a 0.45 μm syringe filter (PALL), and loaded onto a HiLoad 26/600 Superdex 75 prep grade column (GE Healthcare), pre-equilibrated with 2 CV of binding buffer and with a flow rate of 2.6 mL/min. NΔ47-TPH2-R was collected from the peak eluting at the expected elution volume (confirmed by SDS-PAGE). The protein was concentrated to the desired concentration using a Vivaspinn Turbo centrifugal concentrator (Sartorius) with a 3 kDa membrane by centrifugation at 4000 g and 4 °C. The protein was aliquoted, flash frozen in liquid N₂, and stored at -80 °C until further use.

To purify NΔ47-TPH2 and TPH2 for use in cryo-EM experiments, the following modifications were done: The chromatographic system was an ÄKTA Pure (GE Healthcare) and a very narrow fraction of the top of the affinity chromatography peak after protease cleavage was collected. The concentration was measured on a NanoDrop One Microvolume spectrophotometer (Thermo Scientific™) and 1.5x molar equivalent of (NH₄)₂Fe(SO₄)₂ was added from a freshly prepared stock solution and allowed to incubate for around 15 min before filtering the sample through a 0.22 μm centrifuge tube filter (Corning). The protein was immediately used to set up cryo-EM grids.

Table 2 Binding buffers used for purification of TPH2 variants

TPH2 variant	Binding buffer
NΔ47-TPH2-R*	Std: 12 mM Na ₂ HPO ₄ /8 mM NaH ₂ PO ₄ , 100 mM (NH ₄) ₂ SO ₄ , pH 7.0 (adjusted with NH ₄ OH)
	L-Phe: 12 mM Na ₂ HPO ₄ /8 mM NaH ₂ PO ₄ , 100 mM (NH ₄) ₂ SO ₄ , 10 mM L-Phe pH 7.0 (adjusted with NH ₄ OH)
NΔ47-TPH2 TPH2	20 mM Hepes, 300 mM (NH ₄) ₂ SO ₄ , pH 7.0 (adjusted with NH ₄ OH)

All buffers were filtered using a 1000 mL filter unit with a 0.2 μm membrane (Fisher Scientific)

** NΔ47-TPH2-R was purified in standard buffer (std) except for protein purified for NMR structure determination and SAXS experiments where the L-Phe buffer was used.*

Sample preparation

If not stated otherwise, frozen samples of NΔ47-TPH2-R were prepared for experiments in the following way: A frozen sample of the relevant variant was thawed on ice and filtered through a 0.22 μm centrifuge tube filter (Corning). The protein concentration was determined as a mean of triplicate measurements using an ND-1000 NanoDrop Spectrophotometer (Saveen

Werner). The extinction coefficient of NΔ47-TPH2-R at 280 nm were estimated based on the protein sequence using ProtParam from the ExPASy server ⁴¹.

Analytical size exclusion chromatography (SEC)

The oligomeric state of NΔ47-TPH2-R was investigated by analytical size exclusion chromatography using a Superdex 75 10/300 GL column (GE Healthcare) and an ÅKTA Explorer chromatographic system (GE Healthcare). The expected elution volumes of a monomer, dimer, and trimer were estimated based on a calibration curve constructed from the column manual (Superdex 75 10/300 GL, GE Healthcare). The column had been pre-equilibrated with 2 CV of running buffer (Std buffer or L-Phe buffer, Table 2). Samples of 500 μL with the desired protein concentrations were prepared by diluting NΔ47-TPH2-R with either Std or L-Phe buffer. The samples were loaded using a 500 μL sample loop and eluted with 1.5 CV of running buffer at a flow rate of 0.8 mL/min.

Nano differential scanning fluorimetry (nanoDSF)

The influence of L-Phe on the thermal stability of NΔ47-TPH2-R was measured using nanoDSF. 15 μL samples with a final concentration of 100 μM NΔ47-TPH2-R were prepared by mixing the relevant volumes of NΔ47-TPH2-R, Std buffer (Table 2), and an L-Phe stock solution (1 mM L-Phe for low concentrations/46 mM L-Phe for high concentrations in Std buffer). The measurements were performed with a Prometheus NT.48 (NanoTemper Technologies) using standard grade capillaries (NanoTemper Technologies) and run with a constant temperature ramp of 0.5 °C/min from 20 °C to 95 °C. All measurements were run in triplicates, and data analyzed with PR.Thermcontrol software (NanoTemper Technologies). The inflection point of the 350 nm/330 nm ratio was used to determine the T_m of the individual measurements. The final T_m was calculated as a mean of the three measurements with standard deviation using Matlab (Mathworks Inc).

Nuclear Magnetic Resonance (NMR) spectroscopy

NMR spectra were acquired on a Bruker Avance III 600 MHz or 750 MHz spectrometer with TCI cryo-probes. Spectra for structure determination, titrations, and backbone relaxation parameters were referenced internally to DSS at 0 ppm in the ¹H dimension and by relative gyromagnetic ratios in the ¹³C and ¹⁵N dimensions. All data sets were processed and phase corrected manually with either Topspin 3.6 for 2D spectra or qMDD ⁴² and NMRPipe ⁴³ for non-uniformly sampled 3D spectra and pseudo 3D spectra. The buffer conditions of NMR samples were 12 mM Na₂HPO₄/8 mM NaH₂PO₄, 100 mM (NH₄)₂SO₄, pH 7.0 (adjusted with NH₄OH), 5% (v/v) D₂O (+125 μM DSS for referencing) unless specified otherwise. Secondary structure propensities were calculated based on chemical shifts using Talos-N ⁴⁴, and secondary ¹³C^α chemical shifts were calculated by use of random coil shifts estimated using the online tool from Kjaergaard et al. ⁴⁵. Chemical shift perturbations were calculated using the following equation:

$$\Delta\delta_{NH}(ppm) = \sqrt{(\Delta\delta^1H)^2 + (0.154 \cdot \Delta\delta^{15}N)^2}$$

NMR assignment

Resonances of free and L-Phe bound NΔ47-TPH2-R were assigned manually using CcpNmr analysis 2.4.2 ⁴⁶. Backbone and sidechain resonances of L-Phe bound NΔ47-TPH2-R were assigned using following triple resonance spectra: BEST TROSY (BT)-HNCA, BT-HNCACB, BT-HN(CO)CACB, BT-HNCO, BT-HN(CA)CO, HcCH-TOCSY, hCCH-TOCSY, ¹⁵N-NOESY-HSQC, and ¹³C-NOESY-HSQC (all measured at 30 °C) and BT-HNCA (measured at 40 °C). Additionally, three versions of amino acid specific ¹H-¹³C HSQC spectra were used in the assignment process of the methyl resonances. The spectra were measured on one sample (A1) containing 1.1 mM NΔ47-TPH2-R and 10 mM L-Phe. Chemical shifts for the L-Phe ligands were obtained from a ¹H-¹³C HSQC spectrum and an aromatic ¹H-¹³C HSQC spectrum on a sample (A2) containing 450 μM unlabeled NΔ47-TPH2-R and 2.6 mM ¹³C/¹⁵N-labeled L-Phe. Backbone resonances of free NΔ47-TPH2-R were assigned using following triple resonance spectra collected on multiple samples: BT-HNCA (25 °C), BT-HNCO (25 °C), BT-HN(CO)CA (25 °C), BT-HNCACB (25 °C), HcCH-TOCSY (20 °C), hCCH-TOCSY (20 °C), BT-HN(CA)CO (20 °C), BT-HNCACB (20 °C), ¹⁵N-NOESY-HSQC (20 °C), ¹³C-NOESY-HSQC (20 °C). Additionally, an HNCA spectrum (20 °C) recorded using a gradient optimized CO decoupling pulse (GOODCOP ⁴⁷), and three HNCA spectra (20 °C) with different variations of a beta/alpha decoupling pulse (BADCOP1, BADCOP2, BADCOP3 ⁴⁷) were used in the assignment process. Details on the samples used are provided in Table 3.

Table 3 NMR samples of free NΔ47-TPH2-R

Sample	NΔ47-TPH2-R concentration*	Spectra
B1	1000 μM	BT-HNCA (25 °C)
B2	200 μM	BT-HNCO (25 °C), BT-HNCOCA (25 °C)
B3	275 μM	BT-HNCACB (25 °C), HcCH-TOCSY (20 °C)
B4	250 μM	hCCH-TOCSY (20 °C), BT-HN(CA)CO (20 °C)
B5	240 μM	¹³ C-NOESY-HSQC (20 °C), ¹⁵ N-NOESY-HSQC (20 °C)
B6	800 μM	BT-HNCACB (20 °C), BADCOP1 (20 °C), BADCOP2 (20 °C), BADCOP3 (20 °C), GOODCOP (20 °C)

*The stated concentration is the initial concentration of the sample. Due to aggregation of the protein, the concentration gradually decreased over time. ¹H-¹⁵N HSQC spectra were recorded before and after each of the 3D spectra to ensure that the aggregation had not caused any structural changes in the remaining soluble protein.

NMR Structure calculations and refinements

Intra- and intermolecular (protein – protein) restraints for the structure calculations of NΔ47-TPH2-R in complex with L-Phe were obtained from a ¹⁵N-NOESY-HSQC spectrum and a ¹³C-NOESY-HSQC spectrum with 120 ms mixing time recorded on sample A1. Intermolecular restraints (protein – ligand) were obtained from a ¹³C/¹⁵N-filtered ¹³C-NOESY-HSQC and a ¹³C/¹⁵N-filtered ¹⁵N-NOESY-HSQC with 120 ms mixing time also recorded on sample A1. In all cases, NOE peaks were picked manually. Backbone dihedral restraints were calculated from the chemical shifts using TALOS-N ⁴⁴. Using CYANA 3.98.5 ⁴⁸, the NOESY spectra were semi-automatically assigned and initial structures were calculated to generate distance restraints. This process was done iteratively, and assignments and structures were evaluated manually and recalculated with modified inputs in each iteration. To unambiguously assign the peaks of the ¹³C/¹⁵N-filtered ¹³C-NOESY-HSQC spectrum and the ¹⁵N-filtered ¹³C/¹⁵N-NOESY-HSQC spectrum to the L-Phe ligands, the chemical shifts in the indirect ¹H dimension of the filtered NOESY spectra and the ¹H chemical shifts of the L-Phe ligands were shifted with +12 ppm. P2 symmetry were imposed on both the NΔ47-TPH2-R dimer and the two L-Phe ligands. XPLOR-NIH v2.44 ^{49,50} was subsequently used to calculate and refine 100 structures with the distance restraints generated from the initial structure calculations. The structures were refined using an implicit water potential (EEFx ⁵¹), and an energy term was added to restrain the folded regions (residues 63-135) in P2 symmetry. The 20 lowest energy structures were selected to represent the dimer structure of NΔ47-TPH2-R in complex with L-Phe.

NMR titrations and *K_d* determination

For the titration of ¹⁵N-labeled NΔ47-TPH2-R with L-Phe, the concentration of NΔ47-TPH2-R was kept constant at 200 μM and ¹H-¹⁵N HSQC spectra of the following NΔ47-TPH2-R:L-Phe molar ratios were recorded: 1:0, 1:0.25, 1:0.5, 1:0.75, 1:1, 1:1.25, 1:1.5, 1:1.8, 1:2.3, 1:3.5, 1:5. For lineshape analysis, the spectra were processed using an exponential window function with 4 Hz and 8 Hz line broadening in the ¹⁵N and ¹H dimensions, respectively. The data was fitted to a 2-state binding model using the 2D lineshape analysis software TITAN ³². 21 peak pairs were used for the analysis and the errors were determined by a bootstrap analysis included in the TITAN software using 100 replicas.

For the titration of ¹³C-labeled L-Phe with NΔ47-TPH2-R the concentration of L-Phe was kept constant at 200 μM and aromatic ¹H-¹³C HSQC spectra of the following L-Phe:NΔ47-TPH2-R molar ratios were recorded: 1:0, 1:0.075, 1:0.15, 1:0.225, 1:0.3, 1:0.375, 1:0.45, 1:0.5, 1:0.6, 1:0.75, 1:1, 1:2.65. The spectra were processed using a QSINE window function with 1 Hz and 0.3 Hz line broadening in the ¹H and ¹³C dimensions, respectively. Though differences in *R*₂ between bound and free L-Phe is expected ⁵², the intensities of the three decreasing peaks were extracted using CcpNmr analysis 2.4.2 ⁴⁶ and fitted to the following equation using Origin2019 (OriginLabs) with a shared *K_d* parameter for all three peaks.

$$I_{obs} = I_{free} - I_{max,bound} \cdot \frac{([TPH2] + [Phe]_0 + K_d) - \sqrt{([Phe]_0 + [TPH2] + K_d)^2 - 4 \cdot [Phe]_0 \cdot [TPH2]}}{2 \cdot [Phe]_0}$$

Where *I_{obs}* is the intensity of a free L-Phe peak at a certain concentration of NΔ47-TPH2-R, *I_{free}* is the intensity of a free L-Phe peak at 0 μM NΔ47-TPH2-R, *I_{max,bound}* is the difference

between I_{free} and I_{obs} at complete saturation, $[TPH2]$ is the concentration of NΔ47-TPH2-R, $[Phe]_0$ is the initial concentration of free L-Phe, and K_d is the dissociation constant.

Backbone ^{15}N relaxation parameters

Longitudinal (R_1) relaxation rates, transverse (R_2) relaxation rates, and steady state heteronuclear 1H - ^{15}N NOE relaxations were determined for L-Phe bound NΔ47-TPH2-R at 30 °C. Data sets were recorded on a sample containing 750 μM $^{13}C/^{15}N$ labeled NΔ47-TPH2-R and 10 mM L-Phe on both a 600 MHz and a 750 MHz spectrometer. A single data set was recorded for both the determination of R_1 and R_2 relaxation rates, while data sets for 1H - ^{15}N NOEs were recorded in triplicates on the 600 MHz spectrometer and duplicates on the 750 MHz spectrometer. To determine R_1 relaxation rates, the experimental delay times were: 0, 100, 200, 400, 600, 900, and 1300 ms and for R_2 relaxation rates the delay times were: 0, 8.48, 16.96, 33.92, 50.88, 76.32, 101.76, and 135.68 ms. The intensity of the spectra as a function of delay times were fitted to an exponential function using CcpNmr Analysis v2.5.1⁴⁶. 1H - ^{15}N NOEs were calculated from the ratio of spectra with and without 1H pre-saturation and each data set was analyzed individually with CcpNmr Analysis 2.5.1⁴⁶. The final 1H - ^{15}N NOEs and errors were calculated as a mean of the individually determined values with a standard deviation using Matlab (Mathworks Inc.).

R_1 and R_2 relaxation rates were determined for both free NΔ47-TPH2-R (150 μM ^{15}N labeled NΔ47-TPH2-R) and L-Phe bound NΔ47-TPH2-R (150 μM ^{15}N labeled NΔ47-TPH2-R and 10 mM L-Phe) on a 750 MHz spectrometer at 20 °C. All data sets were recorded in duplicates. To determine R_1 relaxation rates, the experimental delay times were: 0, 100, 200, 400, 600, 900, and 1300 ms and for R_2 relaxation rates the delay times were: 0, 8.48, 16.96, 33.92, 50.88, 76.32, 101.76, and 135.68 ms. The peak intensity with respect to delay time was fitted to an exponential function using CcpNmr Analysis 2.5.1⁴⁶.

Small angle X-ray scattering (SAXS) data collection

Different concentrations of NΔ47-TPH2-R (1.3, 2.5, 5.9, 8.2, and 11.4 mg/mL) were prepared by diluting the stock solution of NΔ47-TPH2-R with L-Phe buffer collected from the filtrate from the concentration-step during the protein purification. Measurements (in batch mode) were performed at the German Electron Synchrotron DESY at the P12 EMBL BioSAXS beamline (Extended data table 2). Data processing and analysis were performed using the ATSAS software package version 2.8.2⁵³, and final plots were made using Matlab (Mathworks Inc.).

SAXS modelling

10000 different conformations of the N-terminal (residues 46-60) and C-terminal (residues 136-145) of our NMR structure of L-Phe:NΔ47-TPH2-R were generated using RANCH⁵⁴ with the following parameters: The chain type was set to random coil, the symmetry type to mixture, and the percentage of symmetric conformers to 50%. Subsequently, GAJOE⁵⁴ was run on the RANCH structure pool using default settings and fitting to the 11.4 mg/mL SAXS curve. To estimate the volume fraction of each conformer, the output structures from GAJOE were submitted to OLIGOMER⁵⁵ with a constant added as additional component and keeping the remaining parameters as the default settings. In case the error on the volume fraction of a

given conformer was larger than the estimated volume fraction, this conformer was taken out of the ensemble from GAJOE, and OLIGOMER was re-run with the reduced set of conformers.

EM sample preparation and EM data acquisition

The L-Phe sample for cryo-EM (21 μ M N Δ 47-TPH2, 50 μ M L-Phe, 20 mM Hepes, 300 mM (NH₄)₂SO₄, pH 7.0) was prepared by adding L-Phe from a stock solution to freshly purified N Δ 47-TPH2. To prepare the grids, 4 μ L of sample were applied onto glow-discharged R2/2 QUANTIFOIL grids and excess sample was removed using an FEI Vitrobot loaded with pre-wet filter paper with a blotting time of 6.5 s, a blotting force of 9, and at 100% humidity and 4 °C. The grid was subsequently flash-frozen by plunging into liquid ethane. Cryo-EM grids were screened, and EM images from the grids with optimal ice thickness and particle distribution were acquired at 300 kV on an FEI Titan Krios electron microscope (Thermo Fisher Scientific, USA), equipped with a Falcon III direct electron detector. Images were recorded via Thermo Fischer EPU 2.1 in integration mode at 132,000x magnification, corresponding to a calibrated pixel size of 1.06 Å at the specimen level, using an exposure time of 1.02 s with 40 movie frames and a total dose of 78 e⁻ per Å². The cryo-EM grids of N Δ 47-TPH2 without L-Phe was prepared essentially in the same way, except that the sample contained (18 μ M N Δ 47-TPH2, 20 mM Hepes, 300 mM (NH₄)₂SO₄, pH 7.0). The EM images were acquired at 300kV on an FEI Titan Krios electron microscope (Thermo Fisher Scientific, USA), equipped with a Cs corrector and a Falcon III direct electron detector. Images were recorded via Thermo Fischer EPU 2.1 in integration mode at 153,750x magnification, corresponding to a calibrated pixel size of 1.16 Å at the specimen level, using an exposure time of 1.02 second with 40 movie frames and a total dose of 60 e⁻ per Å². The cryo-EM grids of full-length TPH2 without L-Phe (16 μ M TPH2, 20 mM Hepes, 300 mM (NH₄)₂SO₄, pH 7.0) was prepared in the same way and recorded with the same settings as N Δ 47-TPH2 without L-Phe.

Image processing

Frames were dose-weighted, aligned, and summed by MotionCor 2.0⁵⁶. The defocus values were determined using Gctf⁵⁷. Summed micrographs were manually evaluated in the COW-MicrographQualityChecker (<http://www.cow-em.de>). Micrographs with isotropic Thong rings and a defocus range of 1.5 μ m to 5 μ m that showed clear particles were selected and processed further.

For N Δ 47-TPH2 with L-Phe, a total of 1639 summed micrographs were retained for processing, from which about 1.6 million particles were picked using Gautomatch (<https://www2.mrc-lmb.cam.ac.uk/research/locally-developed-software/zhang-software/>). The picked particles were extracted with a box size of 160x160 pixels and binned to 80 x 80 pixels (pixel size of 2.12 Å) in RELION 3.1

(http://www2.mrc-lmb.cam.ac.uk/relion/index.php/Main_Page). Several rounds of reference-free 2D classification were performed, and classes showing empty boxes, ice/ethane contaminations, or uninterpretable features were removed, yielding 697,589 “good particles”. A subset of 100,000 particles were used to generate an initial 3D map via the *ab initio* reconstruction in cryoSPARC⁵⁸, without symmetry imposed (i.e. C1 symmetry). The *ab initio* model from cryoSPARC was low-pass filtered to 30 Å resolution and used as the starting

reference for subsequent 3D classification in RELION 3.0. The 697,589 “good particles” from 2D classification were 3D classified into eight classes without any mask applied nor symmetry imposed. The 3D classes showing clear density for both CDs and RDs were selected, yielding 310,165 particles. These particles were further classified into five classes without symmetry or mask, and the one class containing 58,616 particles that shows clear RD densities on both sides of the CDs was selected and further refined by Refine3D in RELION 3.1 into a map of the entire NΔ47-TPH2 with an average resolution of 8.9 Å (ranging from ~7 Å of the CDs to ~15 Å of the RDs), without symmetry imposed. To reconstruct a higher resolution map for the CDs, the 697,589 “good particles” from 2D classification were 3D classified into 8 classes with a mask applied around the CDs and TDs (hereafter referred to as the CD tetramer) without symmetry imposed. The best classes showing clear CD density were selected and combined, yielding 241,504 particles, which were further classified into eight classes with a mask applied around the CD tetramer but without symmetry imposed. The best classes containing 197,917 particles were selected, re-extracted with a box size of 360×360 pixels (pixel size of 1.06 Å) and refined with a mask around the CD tetramer and D2 symmetry imposed, resulting in a 4.1 Å map of the CD tetramer. After polishing, the resolution of the CD tetramer map was further improved to 3.9 Å (D2 symmetry imposed). For NΔ47-TPH2 without L-Phe, a total of 1149 summed micrographs were retained, from which 313,195 particles were extracted with a box size of 210×210 pixels and binned to 70 × 70 pixels (pixel size of 3.48 Å) in RELION 3.1 (http://www2.mrc-lmb.cam.ac.uk/relion/index.php/Main_Page). After two rounds of 2D classification, the best 2D class averages showing clear features of TPH2 were selected, containing 131,117 “good particles”. For Full-length TPH2, a total of 628 micrographs were recorded, from which 147,940 particles extracted and subjected to 2D classification. However, the 2D class averages did not show reasonable features of typical TPH2, likely due to the aggregation of particles.

Model building and refinement

For the model of the CD tetramer, the crystal structure of the *human* TPH2 CD and TD (PDB: 4V06) was used as the template for model building. The crystal structure was rigid body docked into the 3.9 Å EM map of the NΔ47-TPH2 CD tetramer using Chimera v.1.13.1⁵⁹. All the side chains were truncated to poly-alanine and the carbon backbone was manually adjusted into the EM density using Coot v. 0.8.9.2⁶⁰ with secondary structure restraints. The model was then refined in PHENIX⁶¹ via real-space refinement with secondary structure restraints. For the model of the entire NΔ47-TPH2, the refined model of the CD tetramer and the NMR solution structure of the RD dimer obtained in this study were rigid body docked into the EM map of the entire NΔ47-TPH2 and combined without further adjustment or refinement.

Author contributions

GHJP and IMV conceived the study. GHJP, PH, and BBK supervised the study. BBK and AP designed the NMR experiments and provided input and feedback on NMR data analysis. AP supervised and assisted with the NMR data collection and processing. IMV performed all experiments and data analysis with exception of the cryo-EM data collection and processing. ZZ and IMV prepared cryo-EM samples, and ZZ performed cryo-EM data collection and processing. HS provided feedback on the cryo-EM analysis. NTS performed preliminary cryo-EM experiments. IMV wrote the manuscript apart from the cryo-EM method section, which was written by ZZ. GHJP, BBK, AP, ZZ, and PH provided feed-back and revisions for the manuscript.

Competing interests

The authors declare no competing interests.

Acknowledgements

The synchrotron SAXS data was collected at beamline P12 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). We would like to thank Melissa Graevert for the assistance in using the beamline. The authors thank Danscatt for funding the SAXS trip. The NMR infrastructure was supported by Novo Nordisk Foundation infrastructure grant [#NNF18OC0032996, cOpenNMR]. The work was in part funded by the Novo Nordisk Foundation Challenge grant REPIN [#NNF18OC0033926; to B.B.K] and NTS' Ph.D. study was financially supported by the Independent Research Fund Denmark [DFF-6108-00247; to G.H.P.]. The authors would additionally like to thank David F. Nielsen and Signe Sjørup for technical assistance, and Yulian Gavrilov and Alina Kulakova for useful discussions.

References

1. Udenfriend, S., Clark, C. T. & Titus, E. 5-Hydroxytryptophan Decarboxylase: A New Route of Metabolism of Tryptophan. *J. Am. Chem. Soc.* **75**, 501–502 (1953).
2. Walther, D. J. & Bader, M. A unique central tryptophan hydroxylase isoform. *Biochem. Pharmacol.* **66**, 1673–1680 (2003).
3. Matthes, S., Mosienko, V., Bashammakh, S., Alenina, N. & Bader, M. Tryptophan Hydroxylase as Novel Target for the Treatment of Depressive Disorders. *Pharmacology* **85**, 95–109 (2010).
4. Coates, M. D. *et al.* Molecular defects in mucosal serotonin content and decreased serotonin reuptake transporter in ulcerative colitis and irritable bowel syndrome. *Gastroenterology* **126**, 1657–1664 (2004).
5. Quednow, B. B., Geyer, M. A. & Halberstadt, A. L. Serotonin and schizophrenia. in *Handbook of Behavioral Neuroscience* **31**, 711–743 (2020).
6. Haghighi, F. *et al.* Genetic architecture of the human tryptophan hydroxylase 2 Gene: existence of neural isoforms and relevance for major depression. *Mol. Psychiatry* **13**, 813–820 (2008).
7. Fitzpatrick, P. F. Tetrahydropterin-Dependent Amino Acid Hydroxylases. *Annu. Rev. Biochem.* **68**, 355–381 (1999).
8. Carkaci-Salli, N. *et al.* Functional Domains of Human Tryptophan Hydroxylase 2 (hTPH2). *J. Biol. Chem.* **281**, 28105–28112 (2006).
9. Arturo, E. C. *et al.* First structure of full-length mammalian phenylalanine hydroxylase reveals the architecture of an autoinhibited tetramer. *Proc. Natl. Acad. Sci.* **113**, 2394–2399 (2016).
10. Yohrling, G. J., Jiang, G. C.-T., Mockus, S. M. & Vrana, K. E. Intersubunit binding domains within tyrosine hydroxylase and tryptophan hydroxylase. *J. Neurosci. Res.* **61**, 313–320 (2000).
11. Vrana, K. E., Walker, S. J., Rucker, P. & Liu, X. A Carboxyl Terminal Leucine Zipper Is Required for Tyrosine Hydroxylase Tetramer Formation. *J. Neurochem.* **63**, 2014–2020 (2002).
12. Fitzpatrick, P. F. The Aromatic Amino Acid Hydroxylases. in *Advances in Enzymology and Related Areas of Molecular Biology* **74**, 235–294 (2000).
13. Murphy, K. L., Zhang, X., Gainetdinov, R. R., Beaulieu, J.-M. & Caron, M. G. A Regulatory Domain in the N Terminus of Tryptophan Hydroxylase 2 Controls Enzyme Expression. *J. Biol. Chem.* **283**, 13216–13224 (2008).
14. Tidemand, K. D. *et al.* Stabilization of tryptophan hydroxylase 2 by L-phenylalanine-induced dimerization. *FEBS Open Bio* 1–13 (2016). doi:10.1002/2211-5463.12100
15. Winge, I. *et al.* Activation and stabilization of human tryptophan hydroxylase 2 by phosphorylation and 14-3-3 binding. *Biochem. J.* **410**, 195–204 (2008).
16. Shiman, R. Relationship between the substrate activation site and catalytic site of phenylalanine hydroxylase. *J. Biol. Chem.* **255**, 10029–10032 (1980).
17. Arturo, E. C. *et al.* Manipulation of a cation- π sandwich reveals conformational flexibility in phenylalanine hydroxylase. *Biochimie* **183**, 63–77 (2021).
18. Tourian, A. Activation of phenylalanine hydroxylase by phenylalanine. *Biochim.*

- Biophys. Acta - Enzymol.* **242**, 345–354 (1971).
19. Knappskog, P. M., Flatmark, T., Aarden, J. M., Haavik, J. & Martinez, A. Structure/Function Relationships in Human Phenylalanine Hydroxylase. Effect of Terminal Deletions on the Oligomerization, Activation and Cooperativity of Substrate Binding to the Enzyme. *Eur. J. Biochem.* **242**, 813–821 (1996).
 20. Meisburger, S. P. *et al.* Domain movements upon activation of phenylalanine hydroxylase characterized by crystallography and chromatography-coupled small-angle X-ray scattering. *J. Am. Chem. Soc.* **138**, 6506–6516 (2016).
 21. Flydal, M. I. *et al.* Structure of full-length human phenylalanine hydroxylase in complex with tetrahydrobiopterin. *Proc. Natl. Acad. Sci.* **116**, 11229–11234 (2019).
 22. Patel, D., Kopec, J., Fitzpatrick, F., McCorvie, T. J. & Yue, W. W. Structural basis for ligand-dependent dimerization of phenylalanine hydroxylase regulatory domain. *Sci. Rep.* **6**, 23748 (2016).
 23. Arturo, E. C., Gupta, K., Hansen, M. R., Borne, E. & Jaffe, E. K. Biophysical characterization of full-length human phenylalanine hydroxylase provides a deeper understanding of its quaternary structure equilibrium. *J. Biol. Chem.* **294**, 10131–10145 (2019).
 24. Zhang, S., Roberts, K. M. & Fitzpatrick, P. F. Phenylalanine binding is linked to dimerization of the regulatory domain of phenylalanine hydroxylase. *Biochemistry* **53**, 6625–6627 (2014).
 25. Skawinska, N. Allosteric regulation of human tryptophan hydroxylase isoform 2 (hTPH2). (PhD thesis, Department of Chemistry, Technical University of Denmark, 2020).
 26. McKinney, J., Knappskog, P. M. & Haavik, J. Different properties of the central and peripheral forms of human tryptophan hydroxylase. *J. Neurochem.* **92**, 311–320 (2005).
 27. Tenner, K., Walther, D. & Bader, M. Influence of human tryptophan hydroxylase 2 N- and C-terminus on enzymatic activity and oligomerization. *J. Neurochem.* **102**, 1887–1894 (2007).
 28. Zhang, S., Hinck, C. S. & Fitzpatrick, P. F. The regulatory domain of human tryptophan hydroxylase 1 forms a stable dimer. *Biochem. Biophys. Res. Commun.* **476**, 457–461 (2016).
 29. Zhang, S., Huang, T., Ilangoan, U., Hinck, A. P. & Fitzpatrick, P. F. The Solution structure of the Regulatory domain of Tyrosine Hydroxylase. *J. Mol. Biol.* 1483–1497 (2014). doi:10.1016/j.pestbp.2011.02.012. Investigations
 30. Lang, E. J. M., Cross, P. J., Mittelstädt, G., Jameson, G. B. & Parker, E. J. Allosteric ACTion: the varied ACT domains regulating enzymes of amino-acid metabolism. *Curr. Opin. Struct. Biol.* **29**, 102–111 (2014).
 31. Cross, P. J., Dobson, R. C. J., Patchett, M. L. & Parker, E. J. Tyrosine Latching of a Regulatory Gate Affords Allosteric Control of Aromatic Amino Acid Biosynthesis. *J. Biol. Chem.* **286**, 10216–10224 (2011).
 32. Waudby, C. A., Ramos, A., Cabrita, L. D. & Christodoulou, J. Two-Dimensional NMR Lineshape Analysis. *Sci. Rep.* **6**, 24826 (2016).
 33. Waisbren, S. E. *et al.* Improved Measurement of Brain Phenylalanine and Tyrosine Related to Neuropsychological Functioning in Phenylketonuria. in *JIMD Reports* **4**, 77–86 (2016).

34. Möller, H. E., Ullrich, K. & Weglage, J. In vivo proton magnetic resonance spectroscopy in phenylketonuria. *Eur. J. Pediatr.* **159**, S121–S125 (2000).
35. Khan, C. A., Meisburger, S. P., Ando, N. & Fitzpatrick, P. F. The phenylketonuria-associated substitution R68S converts phenylalanine hydroxylase to a constitutively active enzyme but reduces its stability. *J. Biol. Chem.* **294**, 4359–4367 (2019).
36. Bueno-Carrasco, M. T. *et al.* Structural mechanism for tyrosine hydroxylase inhibition by dopamine and reactivation by Ser40 phosphorylation. *Nat. Commun.* **13**, 74 (2022).
37. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
38. de Groot, M. J., Hoeksma, M., Blau, N., Reijngoud, D. J. & van Spronsen, F. J. Pathogenesis of cognitive dysfunction in phenylketonuria: Review of hypotheses. *Mol. Genet. Metab.* **99**, S86–S89 (2010).
39. Winn, S. R. *et al.* Blood phenylalanine reduction corrects CNS dopamine and serotonin deficiencies and partially improves behavioral performance in adult phenylketonuric mice. *Mol. Genet. Metab.* **123**, 6–20 (2018).
40. Ogawa, S. & Ichinose, H. Effect of metals and phenylalanine on the activity of human tryptophan hydroxylase-2: Comparison with that on tyrosine hydroxylase activity. *Neurosci. Lett.* **401**, 261–265 (2006).
41. Gasteiger, E. *et al.* Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook* 571–607 (Humana Press, 2005). doi:10.1385/1-59259-890-0:571
42. Kazimierczuk, K. & Orekhov, V. Y. Accelerated NMR Spectroscopy by Using Compressed Sensing. *Angew. Chemie Int. Ed.* **50**, 5556–5559 (2011).
43. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
44. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
45. Kjaergaard, M., Brander, S. & Poulsen, F. M. Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *J. Biomol. NMR* **49**, 139–149 (2011).
46. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins Struct. Funct. Genet.* **59**, 687–696 (2005).
47. Coote, P. W. *et al.* Optimal control theory enables homonuclear decoupling without Bloch–Siegert shifts in NMR spectroscopy. *Nat. Commun.* **9**, 3014 (2018).
48. Güntert, P. & Buchner, L. Combined automated NOE assignment and structure calculation with CYANA. *J. Biomol. NMR* **62**, 453–471 (2015).
49. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73 (2003).
50. Schwieters, C. D., Kuszewski, J. J. & Marius Clore, G. Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **48**, 47–62 (2006).
51. Tian, Y., Schwieters, C. D., Opella, S. J. & Marassi, F. M. A practical implicit solvent potential for NMR structure calculation. *J. Magn. Reson.* **243**, 54–64 (2014).

52. Teilum, K., Kunze, M. B. A., Erlandsson, S. & Kragelund, B. B. (S)Pinning down protein interactions by NMR. *Protein Sci.* **26**, 436–451 (2017).
53. Franke, D. *et al.* ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212–1225 (2017).
54. Tria, G., Mertens, H. D. T., Kachala, M. & Svergun, D. I. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207–217 (2015).
55. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282 (2003).
56. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
57. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
58. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
59. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
60. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
61. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 352–367 (2012).
62. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 12–21 (2010).
63. Trewhella, J. *et al.* 2017 publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution: an update. *Acta Crystallogr. Sect. D Struct. Biol.* **73**, 710–728 (2017).
64. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).

Extended data

Structural characterization of L-Phe binding to a potential allosteric site in human tryptophan hydroxylase 2

Extended data figure 1: ^1H - ^{15}N HSQC spectrum of NΔ47-TPH2-R with 10 mM L-Phe, ^1H - ^{13}C HSQC spectrum of ^{13}C -labelled L-Phe with NΔ47-TPH2-R, and secondary structure prediction based on chemical shift values.

Extended data table 1: NMR structural constraints and refinement statistics

Extended data figure 2: R_1 and R_2 backbone relaxation rates and heteronuclear ^1H - ^{15}N NOE relaxations of NΔ47-TPH2-R with L-Phe.

Extended data figure 3: SAXS modelling fit, modelled SAXS conformers, and SAXS scattering curves.

Extended data table 2: Details on SAXS data collection, data analysis, and modelling.

Extended data figure 4: Panels of a $^{13}\text{C}/^{15}\text{N}$ -filtered ^{13}C -NOESY-HSQC of NΔ47-TPH2-R with L-Phe

Extended data figure 5: Representative spectra from titration of NΔ47-TPH2-R with L-Phe, and TITAN line-shape analysis.

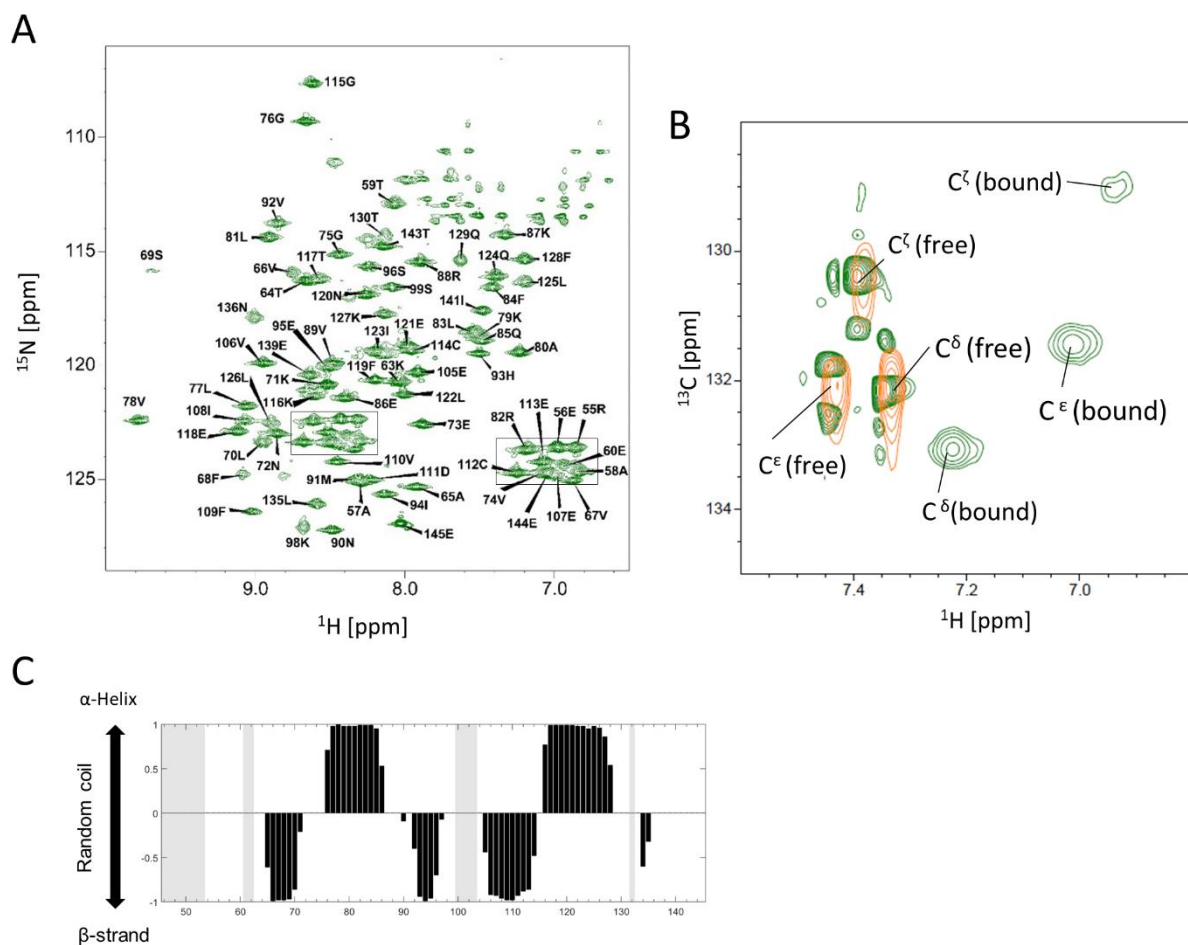
Extended data figure 6: R_1 and R_2 backbone relaxation rates of NΔ47-TPH2-R with and without L-Phe. Secondary structure propensity and secondary chemical shifts of NΔ47-TPH2-R with and without L-Phe.

Extended data figure 7: Titration data from the titration of ^{13}C , ^{15}N -labelled L-Phe with un-labelled NΔ47-TPH2-R.

Extended data table 3: Cryo-EM structural statistics

Extended data figure 8: Cryo-EM 2D class averages of full-length TPH2

Extended data figure 9: AlphaFold prediction of full-length TPH2



Extended data figure 1 A) ^1H - ^{15}N HSQC spectrum of NΔ47-TPH2-R in the presence of 10 mM L-Phe. B) ^1H - ^{13}C HSQC spectrum of ^{13}C -labelled L-Phe with NΔ47-TPH2-R (green) and free ^{13}C -labelled L-Phe (orange). C) Secondary structure prediction based on chemical shift values. Calculated using TALOS-N. Grey areas indicate residues where there is no chemical shift information that allows for calculation of secondary structure propensity.

NMR distance and dihedral constraints

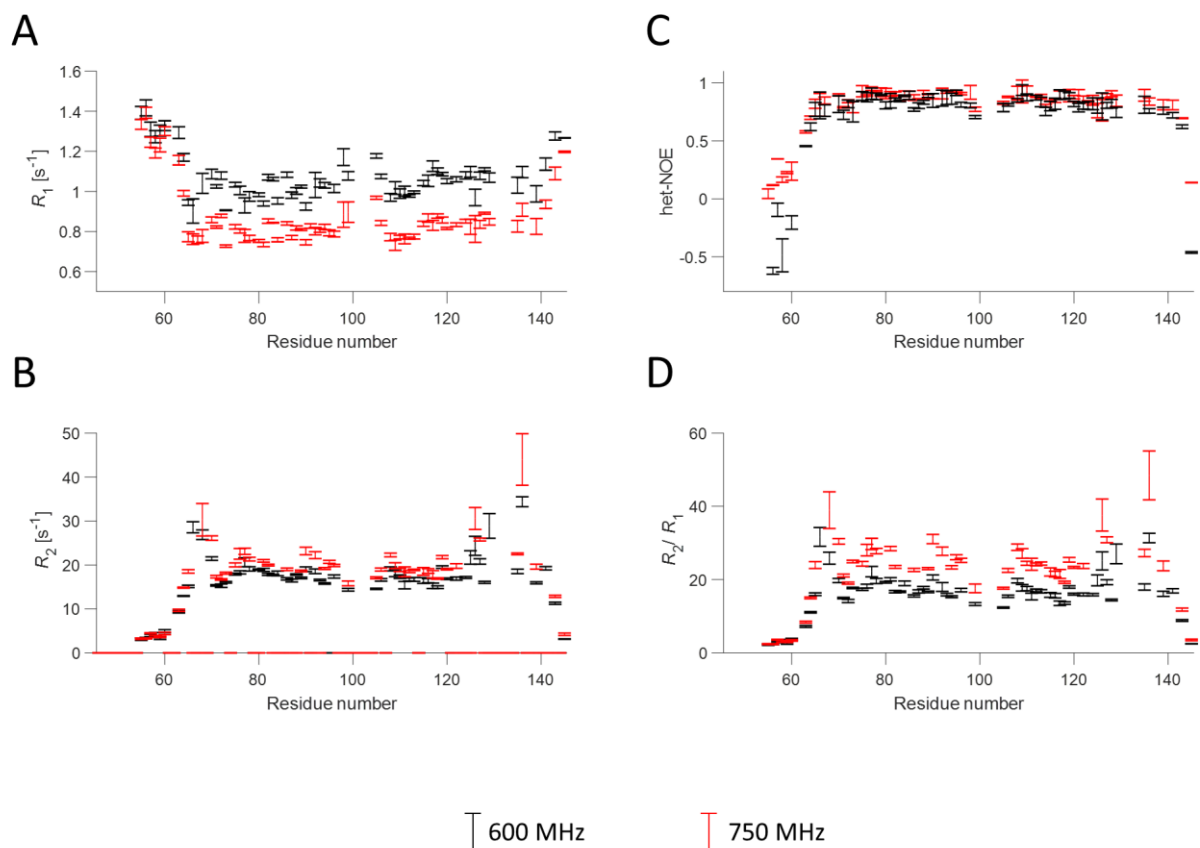
Distance constraints (dimer)	
Total NOE	2171
Intra-residue	554
Inter-residue	1617
Sequential ($ i - j = 1$)	541
Medium-range ($1 < i - j < 5$)	348
Long-range ($ i - j > 5$)	728
Intermolecular (protein-protein)	118
Intermolecular (protein-ligand)	68
Total dihedral angle restraints (monomer)	
ϕ	54
ψ	54

Structure statistics

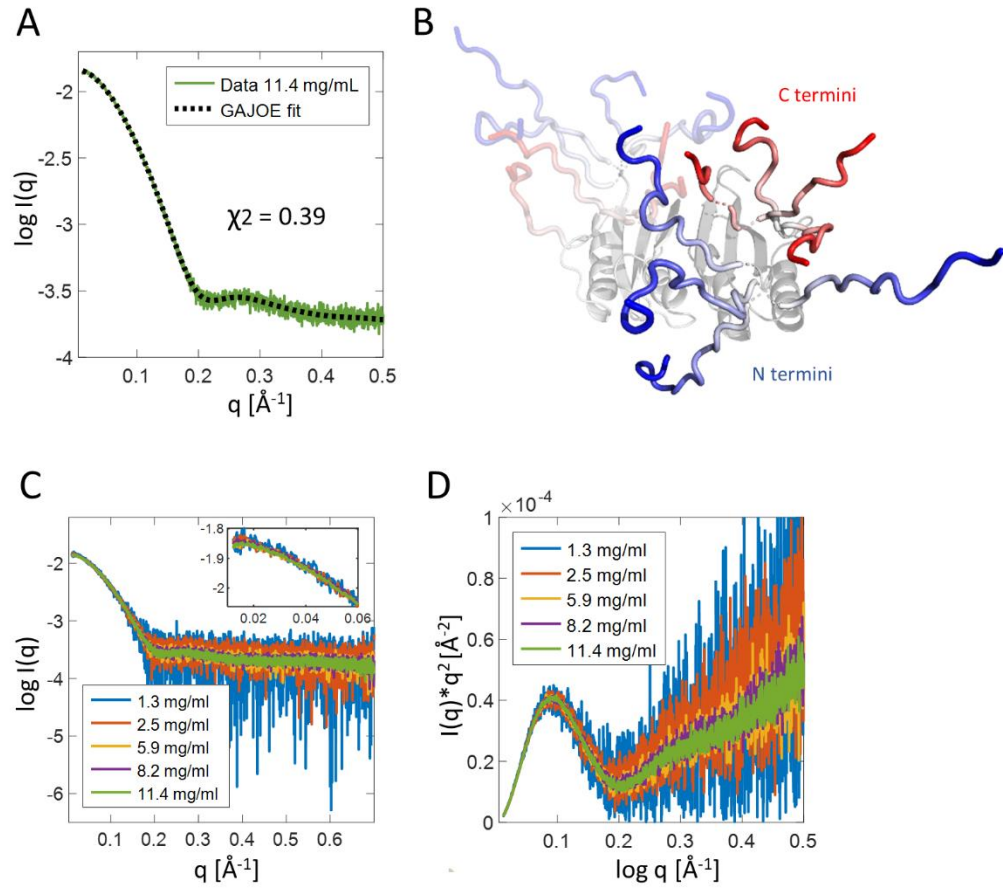
Violations (mean and s.d.)	
Distance constraints (Å)	0.043±0.003
Dihedral angle constraints (°)	0.53±0.07
Max. dihedral angle violation (°)	6.17
Max. distance constraint violation (Å)	0.635
Deviations from idealized geometry	
Bond lengths (Å)	0
Bond angles (°)	2.29±0.01
Impropers (°)	2.436±0.006
Average pairwise r.m.s. deviation** (Å)	
Protein (All heavy)	1.3±0.2
Protein (Backbone)	0.5±0.1
L-Phe ligands (All heavy)	1.2±0.6
Ramachandran statistics***	
Favoured region	99%
Allowed region	1%
Outliers	0%

**Pairwise r.m.s. deviation was calculated among 20 refined structures using residues 64-97 and 105-130 for the protein (protein core defined by wwPDB server upon structure validation) and including both L-Phe ligands.

***Ramachandran statistics were obtained from the wwPDB structure validation report ⁶², which included residues 64-97 and 105-130 in the analysis.



Extended data figure 2 Backbone ^{15}N relaxation parameters of N Δ 47-TPH2-R with 10 mM L-Phe at 30 $^{\circ}C$. A) Longitudinal (R_1) relaxation rates, B) Transverse (R_2) relaxation rates, C) heteronuclear 1H - ^{15}N NOE relaxations, and D) R_2/R_1



Extended data figure 3 A) GAJOE⁵⁴ fit using a RANCH⁵⁴ structure pool of 10000 conformers of residues 46-60 and 136-145, to an 11.4 mg/mL SAXS curve of NΔ47-TPH2-R with L-Phe, resulting in a χ^2 of 0.39. B) Overlay of the four different N- and C-terminal RANCH conformers determined by GAJOE (resulting in the fit in A). C) SAXS curves of NΔ47-TPH2-R in different protein concentrations in the presence of 10 mM L-Phe, showing that NΔ47-TPH2-R is in the same conformation in the range 1.3-11.4 mg/mL. The insert shows a zoom-in of the low q -region showing no attractive or repulsive interactions in the concentration range. D) Kratky representation of the SAXS curves in A indicating a folded but flexible protein.

Extended Data Table 2 SAXS data collection and processing details. The table was prepared according to the guidelines from⁶³. C) Details on the samples and the data collection parameters, and the software used for data analysis. D) Structural parameters of NΔ47-TPH2-R calculated based on the SAXS data. F) Details on the SAXS modelling.

A) Sample details

Organism	NΔ47-TPH2-R
Expression host	<i>Homo sapiens</i>
Source	E. Coli BL21
Uniprot sequence ID	Expressed and purified in-house
Extinction coefficient (A_{280} , $M^{-1}cm^{-1}$)	Q8IWU9, residues 48-145 + additional N-terminal Gly-Pro from MBP purification tag
Molecular mass M from chemical composition (kDa)	5500
Sample concentration range (mg/mL)	11.266
Solvent composition	1.3-11.4
	12 mM Na_2HPO_4 /8 mM NaH_2PO_4 , 100 mM $(NH_4)_2SO_4$, 10 mM L-Phe pH 7.0
	(adjusted with NH_4OH)

B) SAXS data collection parameters

Instrument	P12 BioSAXS beamline (PETRAIII)
Date	Dec-2019
Detector	Pilatus6m
Wavelength (Å)	1.23984
Beam size (mm^2)	At the detector: 0.2x0.12
Detector distance (m)	3.0
q -measurement range (nm^{-1})	0.03-7.29
Absolute scaling method	Comparison with scattering from pure H_2O
Normalization	To transmitted intensity by beam-stop counter
Monitoring for radiation damage	Frame-by-frame comparison
Exposure time (s)	0.045
Number of frames	20
Sample configuration	Quartz glass capillary
Sample temperature ($^{\circ}C$)	20

C) Software employed for SAXS data reduction, analysis, and interpretation

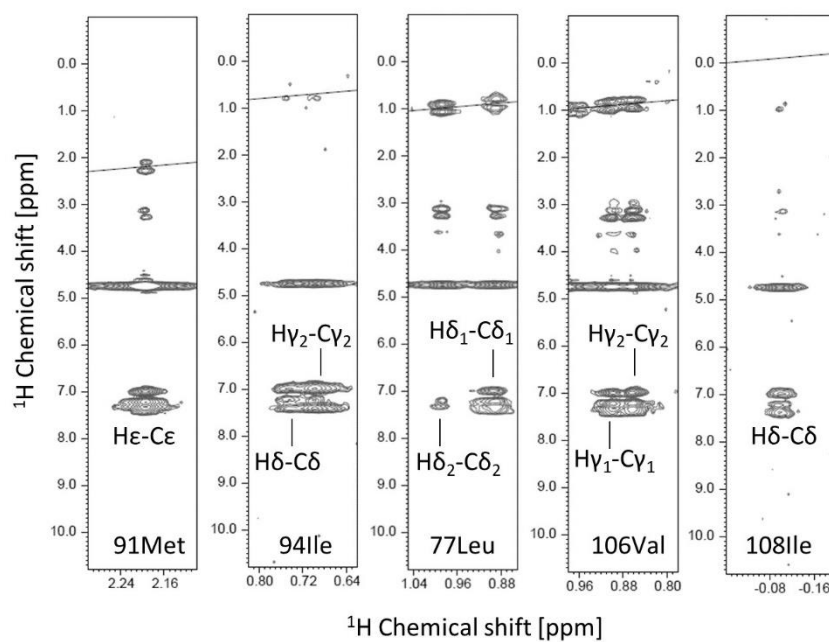
SAS data reduction	Primusqt from Atsas ⁵³
Calculation of ϵ from sequence	ProtParam from ExPaSy ⁴¹
Atomic structure modelling	EOM from Atsas ⁵⁴
Molecular graphics	PyMOL 2.3.3 (Schrödinger)

D) Structural Parameters

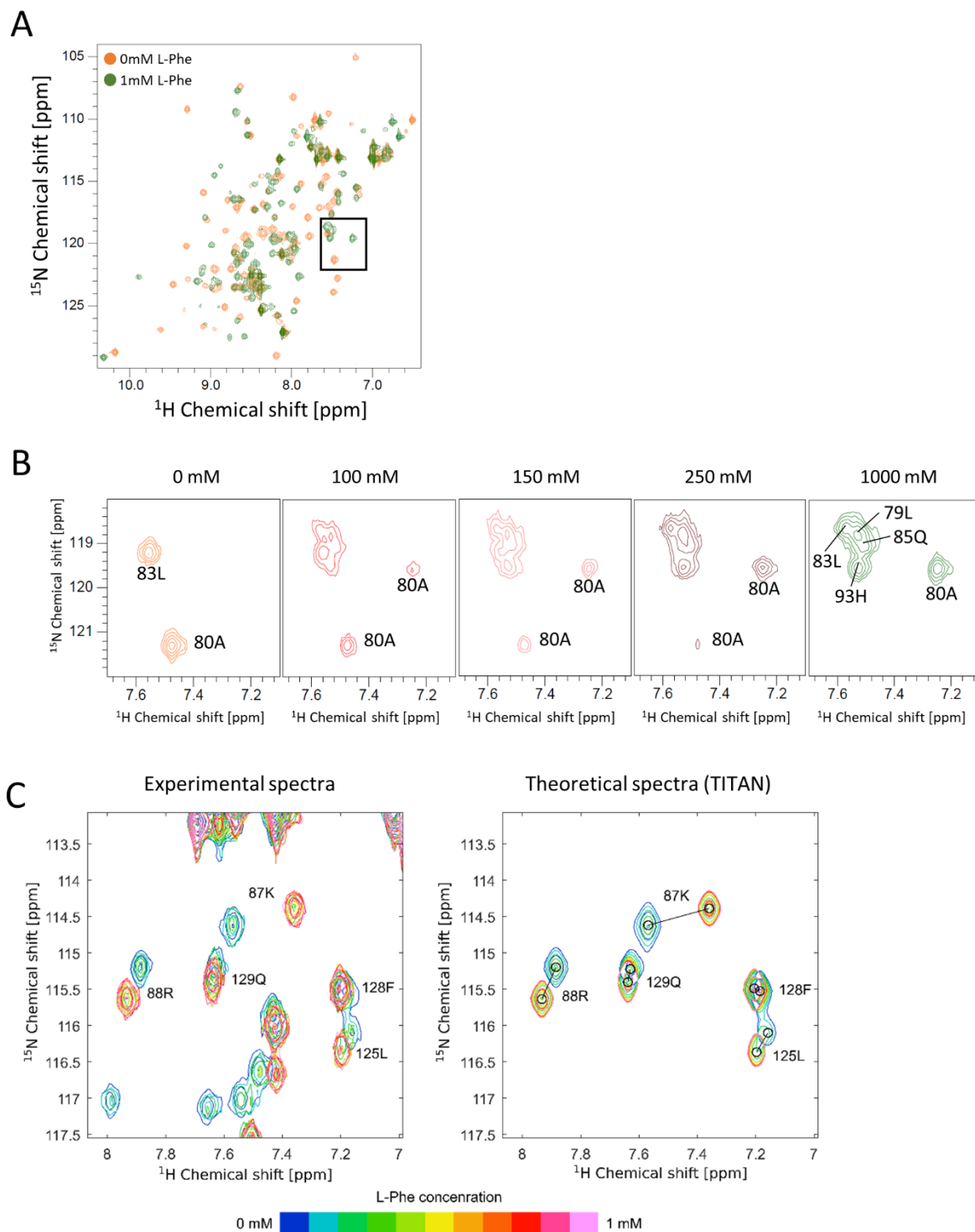
Sample concentration (mg/mL)	1.3	2.5	5.6	8.2	11.4
Guinier Analysis					
$I(0)/c$	0.015	0.015	0.015	0.015	0.015
R_g (Å)	20.9±0.5	20.9±0.3	20.7±0.3	20.8±0.1	20.6±0.1
q -range ($Å^{-1}$)	0.01470-0.06234	0.01221-0.06234	0.01526-0.06234	0.01277-0.06261	0.01498-0.06261
Fidelity (%)	80	96	98	99	99
M from $I(0)$ (kDa)	21	21	21	21	21
$P(r)$ analysis					
$I(0)/c$	0.0151±0.0001	0.01484±0.00005	0.01470±0.00004	0.01484±0.00003	0.01462±0.0003
R_g (Å)	21.2±0.2	21.1±0.1	20.85±0.07	20.99±0.05	20.77±0.05
d_{max} (Å)	67.5	70.0	68.9	69.5	69.7
q -range ($Å^{-1}$)	0.01443-0.39160	0.01194-0.39326	0.01498-0.39326	0.01249-0.39326	0.01470-0.43756
Quality estimate (%)	81	79	75	79	76
M from $I(0)$ (kDa)	21	21	20	21	20
Porod Volume ($Å^3$)	48380	50300	49440	51830	50780

F) Atomistic modelling

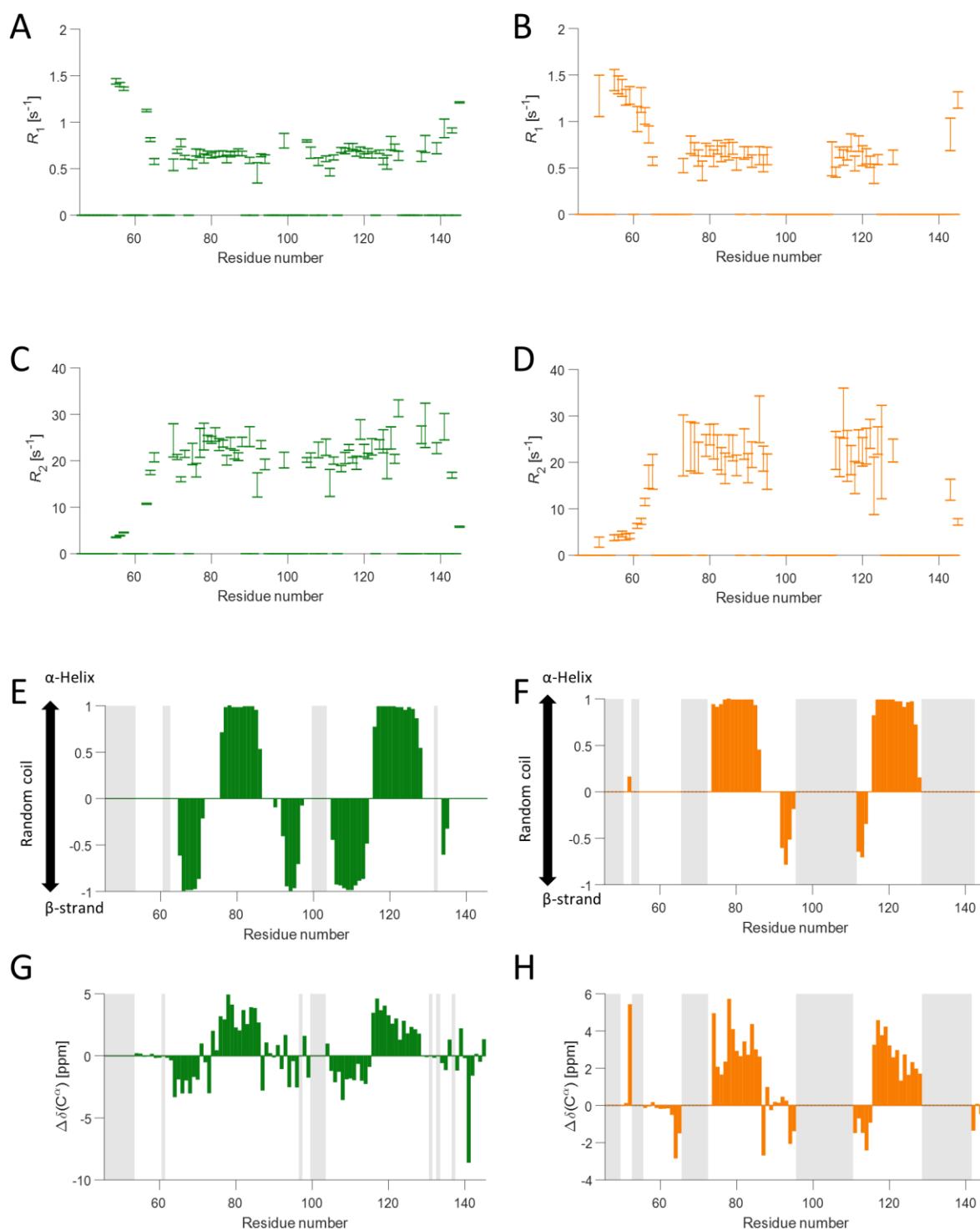
Method	11.4 mg/mL NΔ47-TPH2-R
q -range for fitting ($Å^{-1}$)	RANCH + GAJOE (Ensemble)
Symmetry	0.01-0.5
χ^2	P2
Constant adjustment to intensities	0.39
	0.001



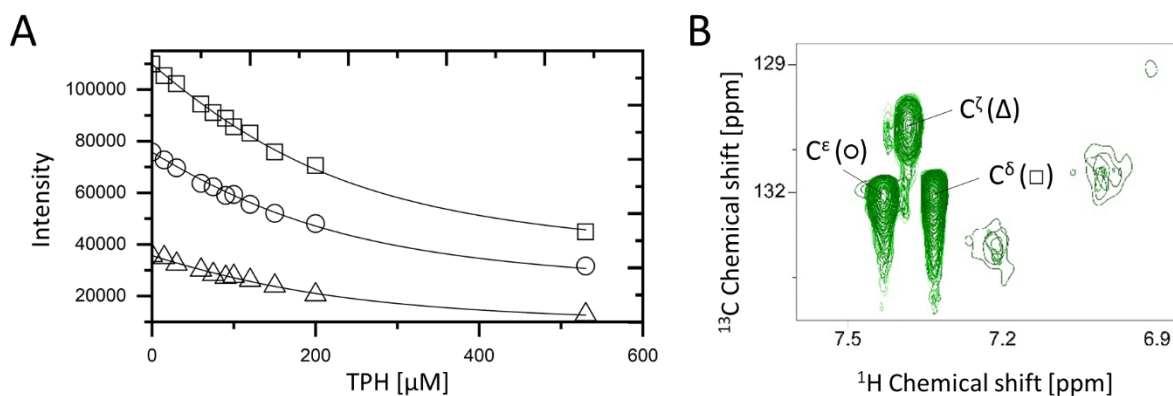
Extended data figure 4 Panels of a $^{13}\text{C}/^{15}\text{N}$ -filtered ^{13}C -NOESY-HSQC of N447-TPH2-R with cross-peaks to the L-Phe ligands.



Extended data figure 5 A) ^1H - ^{15}N HSQC spectra of N Δ 47-TPH2-R with 0 mM L-Phe (orange) and 1 mM L-Phe (green). B) zoom in of the area marked by a black square in A, showing selected spectra with different L-Phe concentrations illustrating the slow exchange kinetics of the system. C) Side-by-side comparison from the titration analysis with TITAN showing a selected region of experimental (left) and theoretical (right) ^1H - ^{15}N HSQC spectra of 200 μM N Δ 47-TPH2-R with 0-1 mM L-Phe.



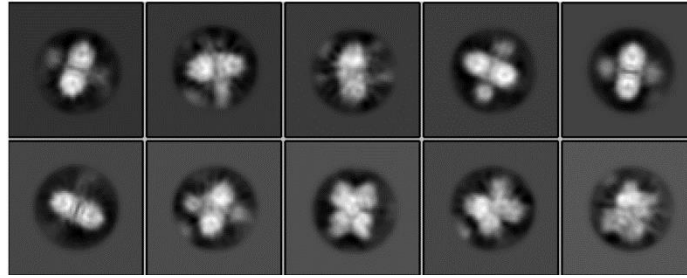
Extended data figure 6 A-D Backbone ^{15}N relaxation parameters of free and L-Phe bound NΔ47-TPH2-R at 20 °C recorded on a 750 MHz spectrometer. Longitudinal (R_1) relaxation rates of A) L-Phe bound NΔ47-TPH2-R and B) Free NΔ47-TPH2-R. Transverse (R_2) relaxation rates of C) L-Phe bound NΔ47-TPH2-R and D) Free NΔ47-TPH2-R. Secondary structure propensity calculated by TALOS-N of E) L-Phe bound NΔ47-TPH2-R and F) free NΔ47-TPH2-R. Secondary chemical shifts of G) L-Phe bound NΔ47-TPH2-R and H) free NΔ47-TPH2-R



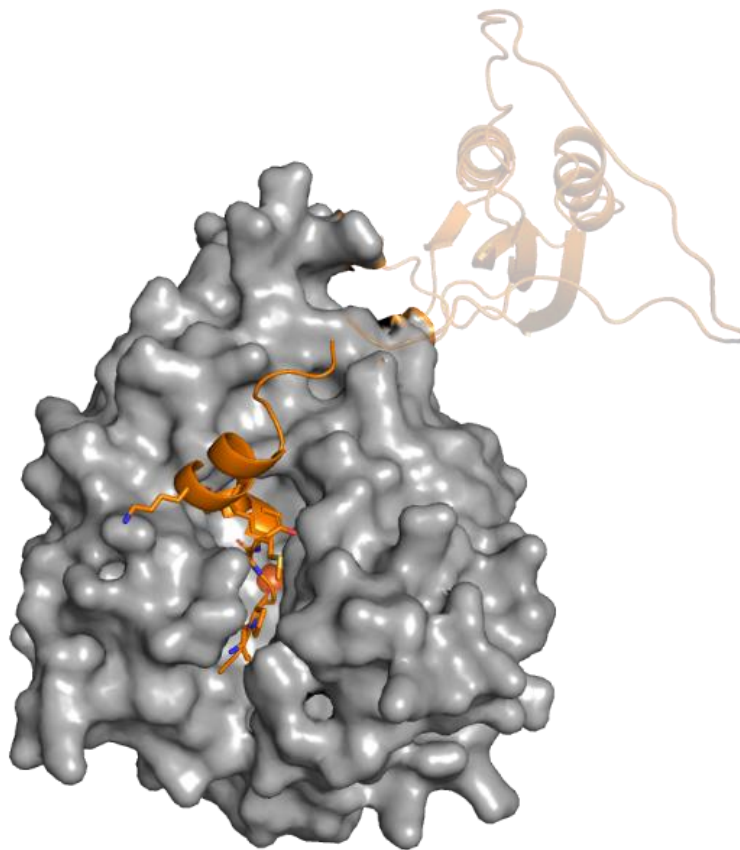
Extended data figure 7 A) Titration data from the titration of ^{13}C , ^{15}N -labelled L-Phe with un-labeled NΔ47-TPH2-R. The intensity decays of the 3 peaks (shown in B) are plotted as squares, circles, and triangles, respectively. The fits are shown as solid lines. B) ^1H - ^{13}C aromatic HSQC spectra of ^{13}C -labeled L-Phe and different concentrations of un-labeled NΔ47-TPH2-R.

Extended Data Table 3 Cryo-EM structural statistics

	Catalytic core of NΔ47-TPH2 with 50 μM L-Phe (EMD-xxxxx) (PDB xxxx)	Complete map of NΔ47-TPH2 with 50 μM L-Phe (EMD-xxxxx) (PDB xxxx)
Data collection and processing		
Magnification	120,700	120,700
Voltage (kV)	300	300
Electron exposure (e ⁻ /Å ²)	72	72
Defocus range (μm)	1-4	1-4
Pixel size (Å)	1.16	1.16
Symmetry imposed	C1	C1
Initial particle images (no.)	1.6 million	1.6 million
Final particle images (no.)	197,917	58,616
Map resolution (Å)	3.9	8.9
FSC threshold	0.143	0.143
Map resolution range (Å)		
Catalytic domain	3.9	6-9
Regulatory domain	---	15-20
Refinement		
Initial model used (PDB code)	4V06	4V06
Model resolution (Å)	4.8	---
FSC threshold	0.5	---
Model resolution range (Å)	---	---
Map sharpening B factor (Å ²)	-250	-300
Model composition		
Non-hydrogen atoms	6873	
Protein residues	1362	
Ligands	0	
B factors (Å ²)		
Protein	---	---
Ligand	---	---
R.m.s. deviations		
Bond lengths (Å)	0.005	---
Bond angles (°)	1.17	---
Validation		
MolProbity score	0.78	---
Clashscore	0.49	---
Poor rotamers (%)	0	---
Ramachandran plot		
Favored (%)	97.42	---
Allowed (%)	2.58	---
Disallowed (%)	0	---



Extended data figure 8 2D class averages from grids with full-length TPH2.



Extended data figure 9 Structure prediction of full-length TPH2 from alphaFold³⁷. The catalytic domain is shown as a grey surface, residues 1-21 as orange cartoon, and the rest of the regulatory domain as faded orange cartoon.

5 TRYPTOPHAN HYDROXYLASE 1 (MANUSCRIPT II)

This chapter presents a study on *human* tryptophan hydroxylase 1. The study investigates the influence of L-Phe and L-Trp on the oligomeric states of three different C-terminally truncated TPH1 variants (The catalytic domain, the combined regulatory and catalytic domain, and the regulatory domain), and additionally investigates ligand binding to the regulatory domain. The results are presented in the form of an early stage manuscript that is currently under preparation. The manuscript is intended to follow after publication of the manuscript presented in Chapter 4, and references to the chapter are therefore included in the manuscript.

NMR experiments were carried out both at Copenhagen University and at the Technical University of Denmark in collaboration with Professor Birthe B. Kragelund, NMR manager Andreas Prestel, and Associate Professor Charlotte H. Gotfredsen. Maja van Hees is acknowledged for preliminary experiments and for protein purifications conducted during her MSc project.

Supporting information is provided at the end of the chapter

L-Phe binds to the regulatory domain of human tryptophan hydroxylase 1

Ida M. Vedel¹, Andreas Presteß, Maja van Hees¹, Charlotte H. Gotfredsen¹, Pernille Harris³, Birthe B. Kragelund², Günther H. J. Peters^{1*}

¹Department of Chemistry, Technical University of Denmark, Kemitorvet, 2800 Kgs. Lyngby, Denmark

²Department of Biology, University of Copenhagen, Ole Maaløes vej 5, 2200 Copenhagen N, Denmark

³Department of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark

Corresponding authors

*Günther H. J. Peters: ghp@kemi.dtu.dk

Abstract

The enzyme tryptophan hydroxylase 1 (TPH1) catalyzes the rate-limiting step in the biosynthesis of peripheral serotonin, an important hormone and neurotransmitter associated with a variety of diseases. Here, three different variants of C-terminally truncated *human* TPH1 were investigated: The isolated catalytic domain (TPH1-C), the isolated regulatory domain (TPH1-R), and both domains combined (TPH1-RC). The oligomeric states of the variants were analyzed using size exclusion chromatography coupled to multi-angle light scattering in the presence of two suspected regulatory domain (RD) ligands, L-Phe and L-Trp. TPH1-R and TPH1-RC were found to be mainly dimeric at protein concentrations >25 µM, while the elution profiles at lower protein concentrations indicated the presence of a monomer-dimer equilibrium. L-Trp was not found to affect the oligomeric state of either TPH1-R or TPH1-RC, while L-Phe affected the elution time of both variants indicating a stabilization of the dimer. TPH1-C was found to be monomeric in agreement with previous studies and independent of L-Phe and L-Trp, which is consistent with dimer formation occurring between RDs similar to the second TPH isoform (TPH2). Nuclear magnetic resonance (NMR) spectroscopy of ¹⁵N-TPH1-R showed large chemical shift perturbations upon addition of L-Phe, confirming that L-Phe can bind to the RD dimer of TPH1.

Introduction

The first and rate-limiting step in the two-step biosynthesis of the peripheral hormone and neurotransmitter serotonin is catalyzed by the enzyme tryptophan hydroxylase (TPH, EC 1.14.16.4), which converts L-Trp into 5-hydroxy-L-Trp. There are two known isoforms of TPH, TPH1 and TPH2, which are expressed mainly in enterochromaffin cells in the gut and in the Raphe nuclei in the brain, respectively ¹. As serotonin cannot pass the blood-brain barrier (BBB), the two enzymes regulate each their independent serotonin system. Increased levels of serotonin in the peripheral system is related to several diseases e.g. carcinoid syndrome and inflammatory bowel disease ²⁻⁴, while decreased levels of serotonin in the brain is associated with e.g. depression and schizophrenia ⁵. The duality of the serotonergic system enables selective therapeutic targeting of one system over the other, either by isoform specific targeting of TPH, or by ensuring that the potential drug cannot pass the BBB ². The latter is however only an option when targeting TPH1, and isoform specific regulation of TPH1 and TPH2 are thus important to characterize for drug development purposes, but also to gain an increased understanding of serotonin regulation. TPH1 and TPH2 are very homologous. Both enzymes are homo-tetramers and each monomer contains three domains: An N-terminal regulatory domain (RD), a catalytic domain (CD), and a tetramerization domain (TD) ⁶⁻⁸. While the CD of TPH1 and TPH2 are both sequentially and structurally homologous ⁹, the biggest differences between the two enzymes are found in the regulatory domain. The RD of TPH2 consists of an expectedly disordered N-terminal tail, and an ordered core (Unpublished data, Chapter 4). The structure of the RD of TPH1 has not been determined, but sequential analysis indicates that the overall architecture of the RD is similar to TPH2, although the N-terminal tail of TPH1 is 46 residues shorter (Figure 1). The regulatory mechanisms of both TPH1 and TPH2 have not yet been fully mapped. Therefore, the enzymes are often compared to the two other more thoroughly studied members of the aromatic amino acid hydroxylase (AAAH) family: Phenylalanine hydroxylase (PAH, EC 1.14.16.1) and tyrosine hydroxylase (TH, 1.14.16.2), which both have the same overall domain architecture and oligomeric state as TPH ^{10,11}. PAH is allosterically activated by formation of symmetric RD dimers induced by binding of two L-Phe molecules at the dimer interface ¹²⁻¹⁴. Importantly, the RD dimer of TPH2 contains two identical RD binding sites for L-Phe similar to PAH, but it is not known if there is a function of L-Phe binding (Unpublished data, Chapter 4). An RD binding site for TPH2 was suggested by Tidemand et al., as a TPH2 variant consisting of the RD and CD (TPH2-RC) existed in an L-Phe affected monomer-dimer equilibrium analogous to the RD of PAH, and as L-Phe and L-Trp thermally stabilized the variant ¹⁵. The isolated RD of TPH1 has differently been found to exist as a dimer ¹⁶, which is not thermally stabilized by L-Trp ¹³. This indicates that the RD of TPH1 does not contain a binding site and consequently suggests that the RD binding sites of TPH2 are isoform specific. However, assessing the sequences of TPH1 and TPH2 (Figure 1) shows that the residues forming the L-Phe binding pockets in the RD of TPH2 are mostly conserved in TPH1. Thus, it is possible that TPH1 contains similar L-Phe binding sites in its RD. In this study, we use size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) to study the oligomeric states of three different C-terminally truncated *human* TPH1 variants: TPH1-R (isolated RD), TPH1-C (isolated CD), and TPH1-RC (RD and CD combined). We observe that TPH1-C is a monomer independent of L-Phe, while the SEC-

MALS profiles of TPH1-RC and TPH1-R indicate a monomer-dimer equilibrium affected by L-Phe. Additionally, using nuclear magnetic resonance (NMR) spectroscopy, we confirm that L-Phe binds to TPH1-R, and observe that the binding site is specific towards L-Phe compared to L-Trp. The findings indicate that the TPH1-R dimer contains similar RD binding sites as TPH2-R.

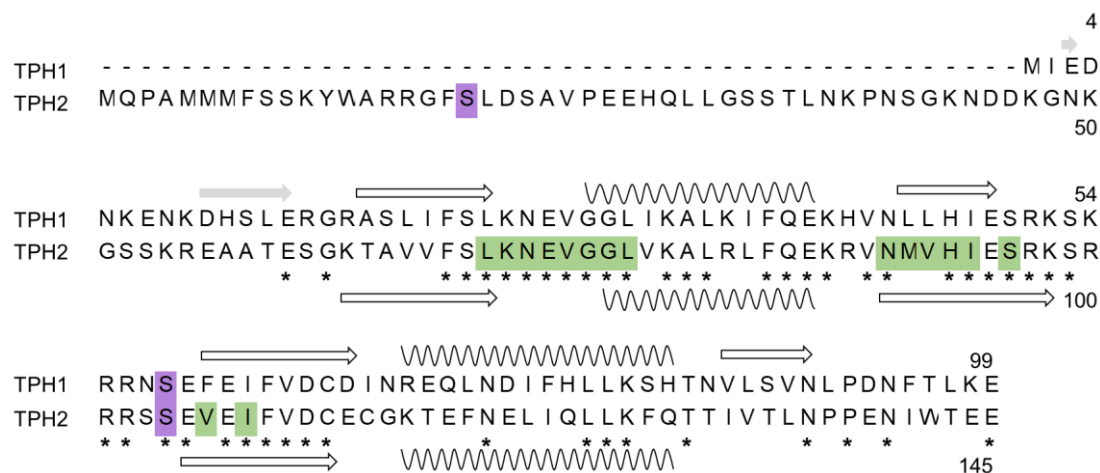


Figure 1 Sequence alignment between the RDs of human TPH1 and human TPH2 using Clustal Omega²⁹. The secondary structure based on the structure of TPH2 is indicated at the bottom, while the secondary structure of TPH1 estimated by Jpred4³⁰ is indicated at the top. The first two predicted β -strands are colored grey as their prediction had a very low score. * Marks identical residues in the alignment. Purple coloring marks a phosphorylation site^{22,23}. Green coloring on the TPH2 sequence indicates residues forming the L-Phe binding pockets (Unpublished data, Chapter 4). There is no structural information available on the first 46 residues of TPH2.

Results & discussion

SEC-MALS reveal L-Phe dependent elution profiles of TPH1-RC and TPH1-R

SEC-MALS was used to investigate the oligomeric states of TPH1-RC, TPH1-R, and TPH1-C at various protein concentrations and either without any ligand, with 10 mM L-Phe, or with 10 mM L-Trp. All the elution curves are provided in supplementary figure 1. In the absence of ligand, the SEC elution volume of TPH1-RC was found to be protein concentration dependent (Figure 2A-B). Only a single peak with tailing was observed at each concentration, which gradually moved to larger elution volumes at lower protein concentrations. Calculating the molecular weight based on the light scattering (LS) signals at the top of each SEC peak showed that the species were primarily dimeric at concentrations >25 μ M (Figure 2C, TPH1-RC MW_{dimer}: 95.6 kDa). Additionally, calculating the molecular weight across the entire SEC peak showed lower molecular weight values at higher elution volumes (Supplementary figure 2A), which in combination with the later elution times at lower protein concentrations (Figure 2B) indicated the presence of monomeric species. Thus, TPH1-RC was found to be in a monomer-dimer equilibrium. This is in agreement with a study from Mockus et al., who found that C-terminal truncations of rabbit TPH1 led to dimeric and monomeric species¹⁷. The gradual concentration dependent movement of the peak observed here, could be attributed to a rapid exchange between monomers and dimers which prevented their separation on the column as also observed for insulin and SecA^{18,19}. It should be noted that the elution volume and calculated molecular weight at higher TPH1-RC concentrations did not completely reach a plateau ([TPH1-RC]>50 μ M), and it is therefore possible that TPH1-RC can exist in higher order oligomer species as well. The monomer-dimer equilibrium of TPH1-RC is similar to an equivalent variant of TPH2 (N47-TPH2-RC, the RD and CD of TPH2 without the first 47 N-terminal residues)¹⁵. Different to TPH1-RC, N47-TPH2-RC could be separated into monomers and dimers on a gel filtration column¹⁵, which could be due to slower exchange between the monomers and dimers and consequently a stronger dimer interface of N47-

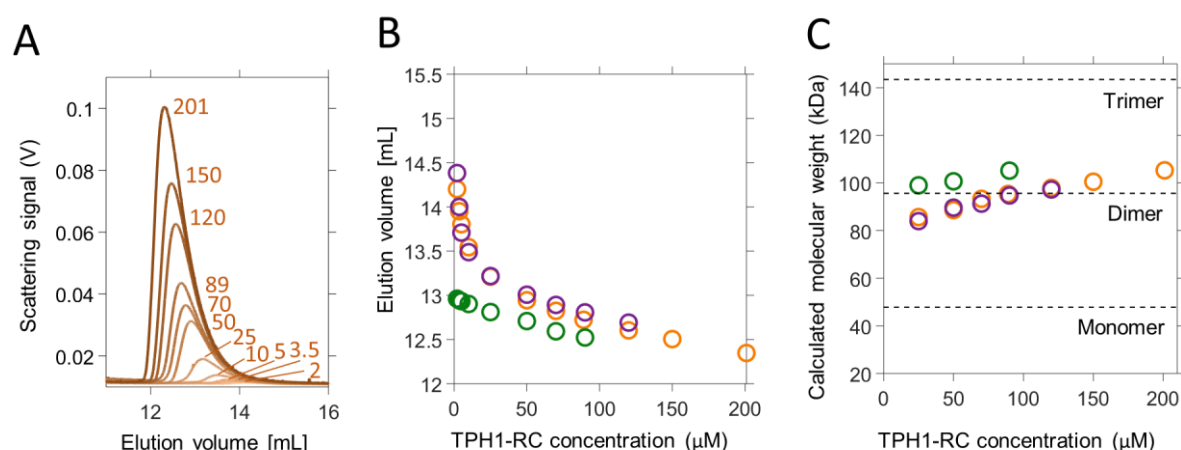


Figure 2 SEC-MALS analysis of TPH1-RC A) Elution profiles of TPH1-RC (LS signal) without any ligand. The numbers next to the curves indicate the protein loading concentration in μ M. B) Elution volumes (at peak maximum) from SEC-MALS runs of TPH1-RC and C) the calculated molecular weight at peak maximum based on the LS and dRI signals either without ligand (orange), with 10 mM L-Phe (green), or 10 mM L-Trp (purple). The theoretically calculated molecular weights of a TPH1-RC monomer, dimer, and trimer are indicated by dashed lines. The calculated molecular weight is shown only for the higher protein concentrations, as the signal-to-noise ratio of the LS- and RI signal is too small at lower concentrations to yield reliable results.

TPH2-RC. Adding L-Trp to the buffer did not affect the oligomeric state of TPH1-RC (Figure 2B-C), while adding L-Phe resulted in a decreased elution volume and a corresponding larger molecular weight, easiest observed at lower TPH1-RC concentrations (Figure 2B). Additionally, tailing of the eluting peaks was less pronounced when L-Phe was present (Supplementary figure 2B), and both observations point to L-Phe stabilizing the dimer. With L-Phe in the buffer, an additional small peak appeared in front of the main peak (Supplementary figure 1E), with a calculated molecular weight of around 150 kDa, corresponding to a trimer of TPH1-RC or an “impure” dimer of TPH1-RC and residual MBP-TPH1-RC from the purification (Supplementary figure 3).

Removing the CD, led TPH1-R in the absence of ligand to elute in a more symmetric peak than TPH1-RC (Figure 3A). Similar to TPH1-RC the molecular weight was calculated to be a dimer at concentrations above 100 μM . (Figure 3C, TPH1-R MW_{dimer} : 23.2 kDa). The signal-to-noise ratio of the LS and dRI signal at lower protein concentrations was too low to calculate a reliable molecular weight in this range. The elution volume was stable for TPH1-R without ligand and with L-Trp at higher tested concentrations, but increased slightly at lower concentrations (10-25 μM), which could be attributed to the presence of a monomeric species. Zhang et al. previously observed TPH1-R to be a stable dimer from 2.7-30 μM using analytical ultracentrifugation, but did not see any indication of a monomeric species¹⁶. However, they worked with a variant where the only cysteine was mutated to a serine to avoid the formation of di-sulfide bridges and used different buffer conditions, which both could influence the behavior of the domain. In this work, we instead added 5 mM DTT to the buffers, to keep the only cysteine reduced. In the presence of 10 mM L-Phe, TPH1-R eluted at a similar volume for all tested concentrations (10-300 μM), consistent with L-Phe stabilizing the dimer. The elution volumes with L-Phe were more stable for TPH1-R than those observed for TPH1-RC,

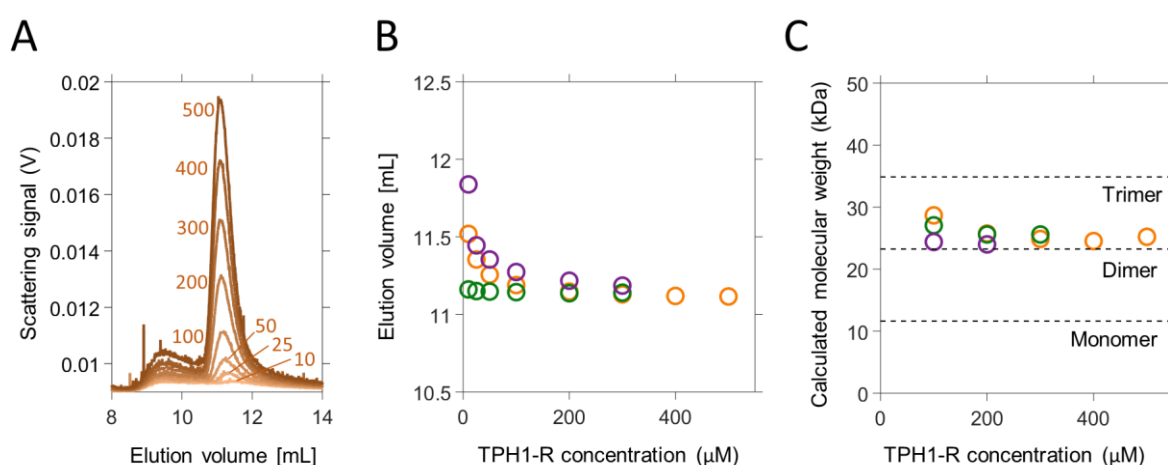


Figure 3 SEC-MALS analysis of TPH1-R A) Elution profiles of TPH1-R (LS signal) without any ligand. The numbers next to the curves indicate the protein loading concentration in μM . B) Elution volumes (at peak maximum) from SEC-MALS runs of TPH1-R and C) the calculated molecular weight at peak maximum based on the LS and RI signals either without ligand (orange), with 10 mM L-Phe (green), or 10 mM L-Trp (purple). The theoretically calculated molecular weights of a TPH1-RC monomer, dimer, and trimer are indicated by dashed lines. The calculated molecular weight is only shown for the higher protein concentrations, as the signal to noise ratio of the LS- and RI signal is too small at lower concentrations to yield reliable results. The baseline of the SEC-MALS run of 300 μM TPH1-R with L-Trp was not stable, and therefore no MW was calculated for this concentration.

indicating that the TPH1-R dimer interface is stabilized to a larger extent by L-Phe and further suggesting the CD to impose restraints to dimer formation.

Alone, TPH1-C eluted as a monomer for all loading concentrations and independent on addition of L-Trp or L-Phe (Figure 4, TPH1-C MW_{monomer}: 36.2 kDa) in agreement with a previous study ²⁰, and consistent with the RD being responsible for dimerization, as also observed for TPH2 ¹⁵.

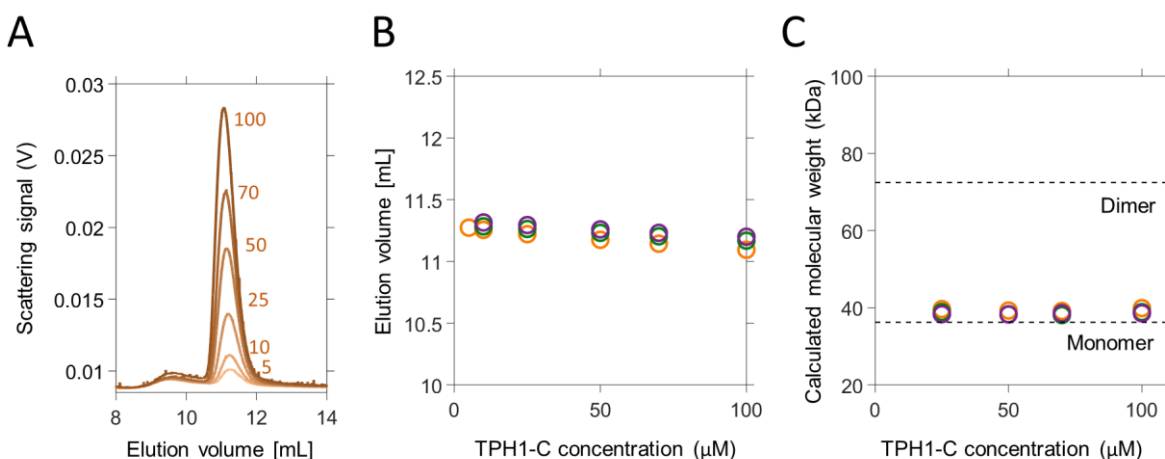


Figure 4 SEC-MALS analysis of TPH1-C A) Elution profiles of TPH1-C (LS signal) without any ligand. The numbers next to the curves indicate the protein loading concentration in μM. B) Elution volumes (at peak maximum) from SEC-MALS runs of TPH1-C and C) the calculated molecular weight at peak maximum based on the LS and RI signals either without ligand (orange), with 10 mM L-Phe (green), or 10 mM L-Trp (purple). The theoretically calculated molecular weights of a TPH1-C monomer and dimer are indicated by dashed lines. The calculated molecular weight is only shown for the higher protein concentrations, as the signal-to-noise ratio of the LS- and dRI signal is too small at lower concentrations to yield reliable results.

NMR analyses confirm L-Phe binding to TPH1-R

Ligand binding to TPH1-R was further investigated using solution NMR spectroscopy. ¹H-¹⁵N HSQC spectra were recorded of ¹⁵N-TPH1-R in the absence of ligand, and with either L-Phe, L-Trp, or L-Tyr. The ¹H-¹⁵N HSQC spectrum of ¹⁵N-TPH1-R alone (Figure 5A), showed 102 peaks out of the expected 97 peaks for a symmetric dimer, and the peaks were all very broad and not well defined. This could be due to an exchange process in the intermediate-to-slow time regime, and linked to TPH1-R monomer-dimer exchange in-line with the gradual elution behavior observed in SEC. Though the spectra were not assigned, a temperature series (5-35 °C) showed two peaks both likely to belong to the C-terminal residue, again indicating exchange between two different conformations (Supplementary figure 4). Addition of 20 mM L-Phe caused large global changes to the spectrum (Figure 5B) affecting both the chemical shifts, the intensity of the peaks, and the number of peaks. The ¹H-¹⁵N HSQC spectrum of ¹⁵N-TPH1-R with L-Phe yielded 61 well dispersed peaks, out of which most were much more intense than the peaks in the spectrum without any ligand. The chemical shift perturbations (CSPs) confirmed the binding of L-Phe to TPH1-R, and the fewer more intense peaks could indicate a stabilization of the dimer leading to less exchange. Another possibility could be a stabilization of the protein fold, however, as TPH1-R elutes at the same volume in SEC with

and without L-Phe at higher concentrations, the overall size of the dimer stays the same. For the RD of *rat* PAH, the ^1H - ^{15}N HSQC spectrum has been observed to change from 120 mainly broad and weak peaks without L-Phe when the system is in a monomer-dimer equilibrium, to fewer more intense peaks upon L-Phe binding where only a dimer is observed ²¹. Addition of 20 mM L-Trp to TPH1-R yielded only small changes to the spectrum (Figure 5C), and the overall appearance remained the same. Thus, it is likely that L-Trp binds very weakly, or not at all, to the TPH1-R dimer, in agreement with L-Trp not increasing the melting temperature of TPH1-R as observed by Patel et al. ¹³. Addition of 1 mM L-Tyr due to the limited solubility of the amino acid, did not cause any CSPs in the spectrum (Figure 5D). While intensity differences were observed, it should be noted that the TPH1-R concentration was 150 μM with L-Tyr, compared to 300 μM in the other spectra, hampering direct comparison. An NMR titration of ^{15}N -TPH1-R with L-Phe (up to 15000 μM L-Phe) showed that the complex is in slow-to-intermediate exchange on the NMR time-scale, complicating a reliable determination of K_d . Upon addition of L-Phe, the peaks from the unbound form gradually disappeared until a certain concentration was reached, where peaks from the bound form started to appear. The spectra mostly resembled the free form of TPH1-R at 400 μM L-Phe (Figure 6A), while peaks from both the free state and the bound state were observed at 800 μM (Figure 6B). At 1500 μM the spectrum mostly resembled the bound form (Figure 6C). To gain more information on the behavior of the complex, and to get a rough idea of the K_d range, an additional NMR titration was performed using ^{13}C -L-Phe. Here, the concentration of ^{13}C L-Phe was kept constant (150 μM) and TPH1-R was added in increasing concentrations. The recorded aromatic ^1H - ^{13}C HSQC spectra of ^{13}C -L-Phe showed three peaks from free L-Phe, while weak peaks from the bound state were visible only at the highest TPH1-R concentration (695 μM , Supplementary figure 5). Although the line-broadening observed in our spectra as a result of the intermediate exchange makes it impossible to determine a reliable K_d from a fit to the intensities, we nonetheless made a global fit to the three decreasing L-Phe peaks and obtained a rough K_d estimate of 590-850 μM (Supplementary figure 5). Combined with the changes in the protein ^1H - ^{15}N HSQC spectrum of ^{15}N -TPH1-R from more “free-like” to more “bound-like” from 400-800-1500 μM added L-Phe, it is probable that the K_d is in the high μM to low mM range, and not smaller. Thus, the binding of L-Phe to TPH1-R is likely weaker than L-Phe binding to N Δ 47-TPH2-R which has a K_d in the lower μM range (~25-100 μM , Unpublished data, Chapter 4).

Although the data presented here provide evidence for L-Phe binding to the RD of TPH1, the exact location of the binding site(s) has not been confirmed. However, the SEC-MALS data show that the TPH1 variants containing the RDs (TPH1-R and TPH1-RC) form dimers both affected by the presence of L-Phe, and considering the sequential analysis between TPH1 and TPH2 presented in the introduction, it is likely that L-Phe binding sites are located at a similar RD dimer interface as in TPH2 and PAH. Additionally, the preference for L-Phe compared to L-Trp, as also observed for N Δ 47-TPH2-R (Unpublished data, Chapter 4), supports similar binding sites. Consequently, the L-Phe binding sites in the RD dimer of TPH2 cannot be assumed to be isoform specific. Still, we do observe different properties of TPH1-R and TPH1-RC compared to the equivalent variants of TPH2. The relatively larger K_d range for the L-Phe TPH1-R complex estimated here compared to that of N Δ 47-TPH2-R indicates different properties of the RD binding pocket(s) of TPH1-R and N Δ 47-TPH2-R.

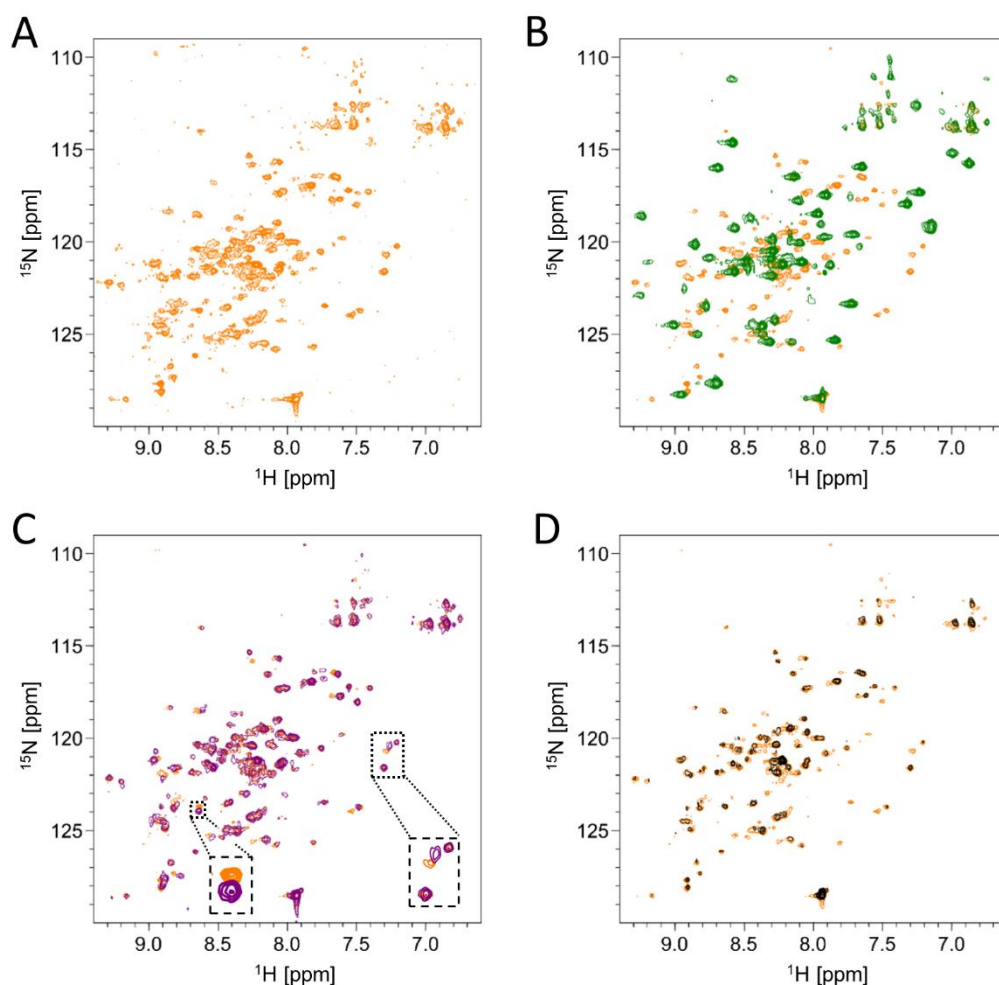


Figure 5 NMR analyses of TPH1-R. ^1H - ^{15}N HSQC spectra of TPH1-R with A) 300 μM TPH1-R and no ligand (Orange), B) 300 μM TPH1-R + 20 mM L-Phe (Green), C) 300 μM TPH1-R + 20 mM L-Trp (Purple), and D) 150 μM TPH1-R + 1 mM L-Tyr (Black). Only 1 mM was added due to the limited solubility of L-Tyr.

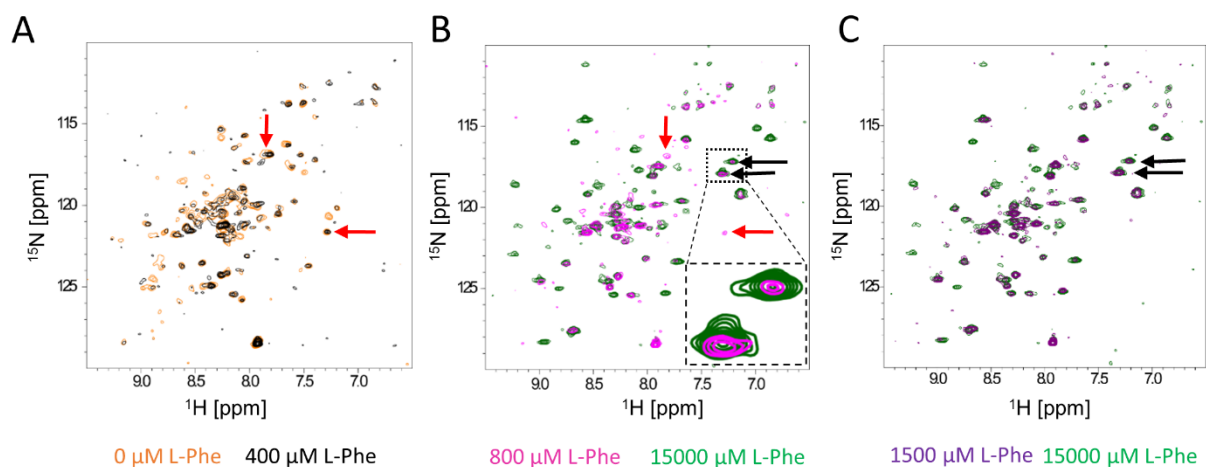


Figure 6 NMR titration of TPH1-R with L-Phe. ^1H - ^{15}N HSQC spectra from a titration of ^{15}N -labeled TPH1-R (400 μM) with L-Phe. A) with 0 μM L-Phe (orange) and 400 μM L-Phe (black). B) with 800 μM L-Phe (Magenta) and with 15000 μM L-Phe (green). C) with 1500 μM L-Phe (purple) and 15000 μM L-Phe (green). Red arrows indicate two peaks representative for the unbound state, while black arrows indicate two peaks representative of the bound state.

In line with this observation, the TPH1-RC variants were not separated into monomers and dimers on the gel filtration column, as was the case for the equivalent variant of TPH2 (NΔ47-TPH2-R)¹⁵, indicating a more rapid equilibrium and a weaker dimer interface, which could contribute to a weaker binding of L-Phe. These subtle differences will be important to further characterize in order to determine the isoform specific properties of the RDs of TPH1 and TPH2, which can additionally aid in the elucidation of isoform specific regulatory mechanisms of the enzymes. Importantly, it is not yet known if the RD binding sites are physiologically relevant in the regulation of either TPH1 or TPH2. The RD cores of TPH1 and TPH2 contain a conserved phosphorylation site, while TPH2 has an isoform specific phosphorylation site in its 46 residue longer N-tail important for 14-3-3 binding^{22,23}. In combination with the similar, yet not identical properties of TPH1-R and NΔ47-TPH2-R and TPH1-RC and NΔ47-TPH2-RC, this indicates that the differences in the N-tail are a main factor in isoform specific regulation of TPH.

Conclusion

The current study show that the RD of TPH1 can bind L-Phe, and additionally show that L-Phe affects dimerization of C-terminally truncated variants of TPH1 including the RD. This is similar to the behavior of the RD of TPH2 reported in previous studies, though differences are observed relating to the strength of L-Phe binding and the dimer interface. These subtle differences will be important to unravel for a more complete understanding of the RDs of TPH.

Methods

Protein expression and purification

TPH1-R (residues 2-99), TPH1-RC (residues 2-413), and TPH1-C (residues 100-413) were expressed with an N-terminal maltose binding protein (MBP)-tag essentially as described by Tidemand et al.¹⁵, using LB media with 30 µg/mL kanamycin sulfate (30 Kan). Upon induction with isopropyl-β-D-thiogalactopyranoside, freshly prepared (NH₄)₂Fe(SO₄)₂ was added to a final concentration of 0.2 mM for TPH1-RC and TPH1-C. When expressing ¹⁵N-labelled TPH1-R, pre-cultures (50 mL) were grown in LB 30 Kan media, pelleted and resuspended in M9 minimal media (22 mM KH₂PO₄, 53 mM Na₂HPO₄, 86 mM NaCl, 1 mM MgSO₄, 1:1000 V/V M2 trace element solution, 22 mM glucose, 11 mM ¹⁵N-(NH₄)₂SO₄). M9 30 Kan media (650 mL) was then inoculated with the resuspended cells from two precultures. Purifications, including cleaving of the MBP-tag, of all three variants were performed as described by Tidemand et al.¹⁵ with the following modifications: The buffer conditions were for TPH1-R 20 mM Tris/HCl, 100 mM (NH₄)₂SO₄, 5 mM DTT, pH 8.0, for TPH1-RC 20 mM Tris/HCl, 300 mM (NH₄)₂SO₄, 5 mM DTT, pH 8.0, and for TPH1-C 20 mM Tris/HCl, 100 mM (NH₄)₂SO₄, pH 8.0. Cleaving off the MBP-tag leaves an additional N-terminal “GP” residues on TPH1-R and TPH1-C and “GPGS” on TPH1-RC. After purification, the protein was concentrated using Vivaspın Turbo centrifugal concentrators (Sartorius) with either 3 kDa or 10 kDa cut-off membranes. The protein was stored at -80 °C after flash freezing in liquid N₂. The purities of

the variants were assessed by SDS-PAGE. The purity of TPH1-RC samples were estimated to 90-95% purity, TPH1-C to 98% purity, while no impurities in the TPH1-R samples were visible on the gel (Supplementary figure 3).

Sample preparation

Frozen protein samples (TPH1-R, TPH1-RC, or TPH1-C) were thawed slowly on ice, and filtered through a 0.22 μm centrifuge tube filter (Corning). The protein concentration was measured using an ND-1000 NanoDrop Spectrophotometer (Saveen Werner) and an extinction coefficient estimated by ProtParam²⁴ based on the protein sequence (for TPH1-RC and TPH1-C) or using an Abbemat 550 refractometer (Anton Paar) and $dn/dc=0.185$ ^{25,26} (for TPH1-R, since the sequence contains no tryptophans or tyrosines). Concentration measurements were performed in triplicates and the average value used as the final protein concentration.

Size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS)

SEC-MALS was used to investigate the oligomeric state of TPH1-R, TPH1-RC, and TPH1-C. The column was coupled to a DAWN Heleos-II MALS detector (Wyatt), an 1260 Infinity II Multiple Wave Length Detector (Agilent Technologies), and an Optilab T-rex refractive index detector (Wyatt). The flow was controlled using an Agilent 1260 Infinity II LC system (Agilent). The column (Superdex 75 10/300 GL for TPH1-R and TPH1-C or Superdex 200 increase 10/300 for TPH1-RC, GE Healthcare) had been pre-equilibrated with at least three column volumes (CV) of running buffer. The running buffer, which had been filtered through a 1000 mL filter unit with a 0.2 μm membrane (Fisher Scientific), was identical to the binding buffer during purification, but with addition of 10 mM L-Phe or 10 mM L-Trp when relevant. Frozen protein samples were prepared as described but filtered through a 0.1 μm centrifuge tube filter (Millipore). Samples of the relevant concentrations were prepared by diluting the concentrated protein sample with running buffer and kept at 10 °C until loading. 50 μL samples were loaded using a 1260 Infinity II autosampler (Agilent Technologies) and eluted at room temperature with 1 CV running buffer at a flow rate of 0.4 mL/min. The data were analyzed using Astra 7.1.4 (Wyatt), using the dRI measurements and $dn/dc=0.185$ to determine the protein concentration. Final plots were made in Matlab (Mathworks Inc.).

Nuclear magnetic resonance spectroscopy

^1H - ^{15}N HSQC spectra of TPH1-R were measured at 35 °C (unless otherwise specified) on a Bruker AVANCE 800 MHz spectrometer with a 5mm TCI cryoprobe. ^1H - ^{15}N HSQCs for the protein observed titration of TPH1-R with L-Phe were measured at 35 °C on a Bruker Avance Neo 800 MHz spectrometer equipped with a 5 mm CPTXO cryoprobe. Aromatic ^1H - ^{13}C HSQC spectra for the L-Phe observed titration with TPH1-R were recorded on a Bruker Avance III HD 600 MHz spectrometer equipped with a 5 mm QCI cryoprobe. Topspin 3.6 was used for processing, phase corrections, and referencing of spectra. For the titration spectra, referencing was done internally to DSS at 0 ppm in the ^1H dimension and by relative gyromagnetic ratios in the ^{13}C and ^{15}N dimensions. The buffer condition for the NMR samples was 20 mM Tris/HCl, 100 mM $(\text{NH}_4)_2\text{SO}_4$, 5 mM DTT, pH 8.0, 5% D_2O (+125 μM DSS for titration samples). A titration of ^{15}N -TPH1-R with L-Phe at 35 °C was performed by maintaining

a constant protein concentration (400 μM) and recording ^1H - ^{15}N HSQC spectra with the following L-Phe concentrations: 0, 100, 200, 400, 800, 1100, 1500, 2000, 3000, 5000, 15000 μM . For the titration of ^{13}C -labeled L-Phe with TPH1-R aromatic ^1H - ^{13}C HSQCs were recorded on samples with 150 μM L-Phe and the following TPH1-R concentrations: 37.5, 75, 112.5, 150, 225, 300, 695 μM . The intensity of the decaying peaks were extracted with CcpNmr Analysis 2.5.1²⁷ and fitted using Origin2019 (OriginLabs) to the following equation with a shared K_d parameter²⁸:

$$I_{obs} = I_{free} - I_{max,bound} \cdot \frac{([TPH1] + [Phe]_0 + K_d) - \sqrt{([Phe]_0 + [TPH1] + K_d)^2 - 4 \cdot [Phe]_0 \cdot [TPH1]}}{2 \cdot [Phe]_0}$$

Where I_{obs} is the intensity of a free L-Phe peak at a certain concentration of TPH1-R, I_{free} is the intensity of a free L-Phe peak without TPH1-R, $I_{max,bound}$ is the difference between I_{free} and I_{obs} at complete saturation, $[TPH1]$ is the concentration of TPH1-R, $[Phe]_0$ is the initial concentration of free L-Phe, and K_d is the dissociation constant.

Acknowledgements

NMR spectra were recorded at cOpenNMR, an infrastructure supported by the Novo Nordisk Foundation (#NNF18OC0032996). B.B.K was supported by the Novo Nordisk Foundation Challenge Program – REPIN (#NNF18OC0033926). The NMR Center • DTU and the Villum Foundation are acknowledged for access to the 800 MHz spectrometer. The authors would additionally like to thank David F. Nielsen for technical assistance and for help with protein expression and purification, and Mariusz Kubus for assistance with the SEC-MALS equipment.

Author contributions

GHJP and IMV conceived the study. IMV performed experiments and data analysis, and wrote the initial draft of the manuscript. MH contributed to protein purification and performed preliminary SEC experiments. BBK, AP, and CHG provided input and feedback on NMR experiments and analysis. GHJP, BBK, CHG, AP, and PH provided input on data interpretation and feedback and revisions for the manuscript.

References

1. Walther, D. J. & Bader, M. A unique central tryptophan hydroxylase isoform. *Biochem. Pharmacol.* **66**, 1673–1680 (2003).
2. Bader, M. Inhibition of serotonin synthesis: A novel therapeutic paradigm. *Pharmacol. Ther.* **205**, 107423 (2020).
3. Coates, M. D. *et al.* Molecular defects in mucosal serotonin content and decreased serotonin reuptake transporter in ulcerative colitis and irritable bowel syndrome. *Gastroenterology* **126**, 1657–1664 (2004).
4. Tsoli, M. & Kaltsas, G. Carcinoid Syndrome. in *Encyclopedia of Endocrine Diseases* 13–20 (Elsevier, 2019). doi:10.1016/B978-0-12-801238-3.65346-1
5. Torrente, M. P., Gelenberg, A. J. & Vrana, K. E. Boosting serotonin in the brain: is it time to revamp the treatment of depression? *J. Psychopharmacol.* **26**, 629–635 (2012).
6. Carkaci-Salli, N. *et al.* Functional Domains of Human Tryptophan Hydroxylase 2 (hTPH2). *J. Biol. Chem.* **281**, 28105–28112 (2006).
7. Yohrling, G. J., Jiang, G. C.-T., Mockus, S. M. & Vrana, K. E. Intersubunit binding domains within tyrosine hydroxylase and tryptophan hydroxylase. *J. Neurosci. Res.* **61**, 313–320 (2000).
8. D'Sa, C. M., Arthur, R. E. & Kuhn, D. M. Expression and Deletion Mutagenesis of Tryptophan Hydroxylase Fusion Proteins: Delineation of the Enzyme Catalytic Core. *J. Neurochem.* **67**, 917–926 (2002).
9. Tidemand, K. D., Peters, G. H., Harris, P., Stensgaard, E. & Christensen, H. E. M. Isoform-Specific Substrate Inhibition Mechanism of Human Tryptophan Hydroxylase. *Biochemistry* **56**, 6155–6164 (2017).
10. Arturo, E. C. *et al.* First structure of full-length mammalian phenylalanine hydroxylase reveals the architecture of an autoinhibited tetramer. *Proc. Natl. Acad. Sci.* **113**, 2394–2399 (2016).
11. Bueno-Carrasco, M. T. *et al.* Structural mechanism for tyrosine hydroxylase inhibition by dopamine and reactivation by Ser40 phosphorylation. *Nat. Commun.* **13**, 74 (2022).
12. Meisburger, S. P. *et al.* Domain movements upon activation of phenylalanine hydroxylase characterized by crystallography and chromatography-coupled small-angle X-ray scattering. *J. Am. Chem. Soc.* **138**, 6506–6516 (2016).
13. Patel, D., Kopec, J., Fitzpatrick, F., McCorvie, T. J. & Yue, W. W. Structural basis for ligand-dependent dimerization of phenylalanine hydroxylase regulatory domain. *Sci. Rep.* **6**, 23748 (2016).
14. Tourian, A. Activation of phenylalanine hydroxylase by phenylalanine. *Biochim. Biophys. Acta - Enzymol.* **242**, 345–354 (1971).
15. Tidemand, K. D. *et al.* Stabilization of tryptophan hydroxylase 2 by L-phenylalanine-induced dimerization. *FEBS Open Bio* 1–13 (2016). doi:10.1002/2211-5463.12100
16. Zhang, S., Hinck, C. S. & Fitzpatrick, P. F. The regulatory domain of human tryptophan hydroxylase 1 forms a stable dimer. *Biochem. Biophys. Res. Commun.* **476**, 457–461 (2016).
17. Mockus, S. M., Kumer, S. C. & Vrana, K. E. Carboxyl terminal deletion analysis of tryptophan hydroxylase. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1342**,

- 132–140 (1997).
18. Wahlund, P., Roessner, D., Jocks, T. Identification of Insulin Oligomeric States Using SEC-MALS. *Appl. note (AN1605)*, *Wyatt Technol.* (2013).
 19. Woodbury, R. L. Complex behavior in solution of homodimeric SecA. *Protein Sci.* **11**, 875–882 (2002).
 20. Moran, G. R., Daubner, S. C. & Fitzpatrick, P. F. Expression and Characterization of the Catalytic Core of Tryptophan Hydroxylase. *J. Biol. Chem.* **273**, 12259–12266 (1998).
 21. Li, J., Ilangovan, U., Daubner, S. C., Hinck, A. P. & Fitzpatrick, P. F. Direct evidence for a phenylalanine site in the regulatory domain of phenylalanine hydroxylase. *Arch. Biochem. Biophys.* **505**, 250–255 (2011).
 22. Winge, I. *et al.* Activation and stabilization of human tryptophan hydroxylase 2 by phosphorylation and 14-3-3 binding. *Biochem. J.* **410**, 195–204 (2008).
 23. Kuhn, D. M., Arthur, R. & States, J. C. Phosphorylation and activation of brain tryptophan hydroxylase: identification of serine-58 as a substrate site for protein kinase A. *J. Neurochem.* **68**, 2220–3 (1997).
 24. Gasteiger, E. *et al.* Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook* 571–607 (Humana Press, 2005). doi:10.1385/1-59259-890-0:571
 25. Barer, R., Joseph, S. Refractometry of living cells.1. Basic Principles. *Q. J. Microsc. Sci.* **95**, 399–423 (1954).
 26. EMBL Hamburg. Refractometer measurements. Available at: <https://www.embl-hamburg.de/biosaxs/manuals/equipment/refractometer.html>.
 27. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins Struct. Funct. Genet.* **59**, 687–696 (2005).
 28. Teilum, K., Kunze, M. B. A., Erlendsson, S. & Kragelund, B. B. (S)Pinning down protein interactions by NMR. *Protein Sci.* **26**, 436–451 (2017).
 29. Sievers, F. *et al.* Fast, scalable generation of high- quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
 30. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.* **43**, W389–W394 (2015).

Supplementary figures and tables

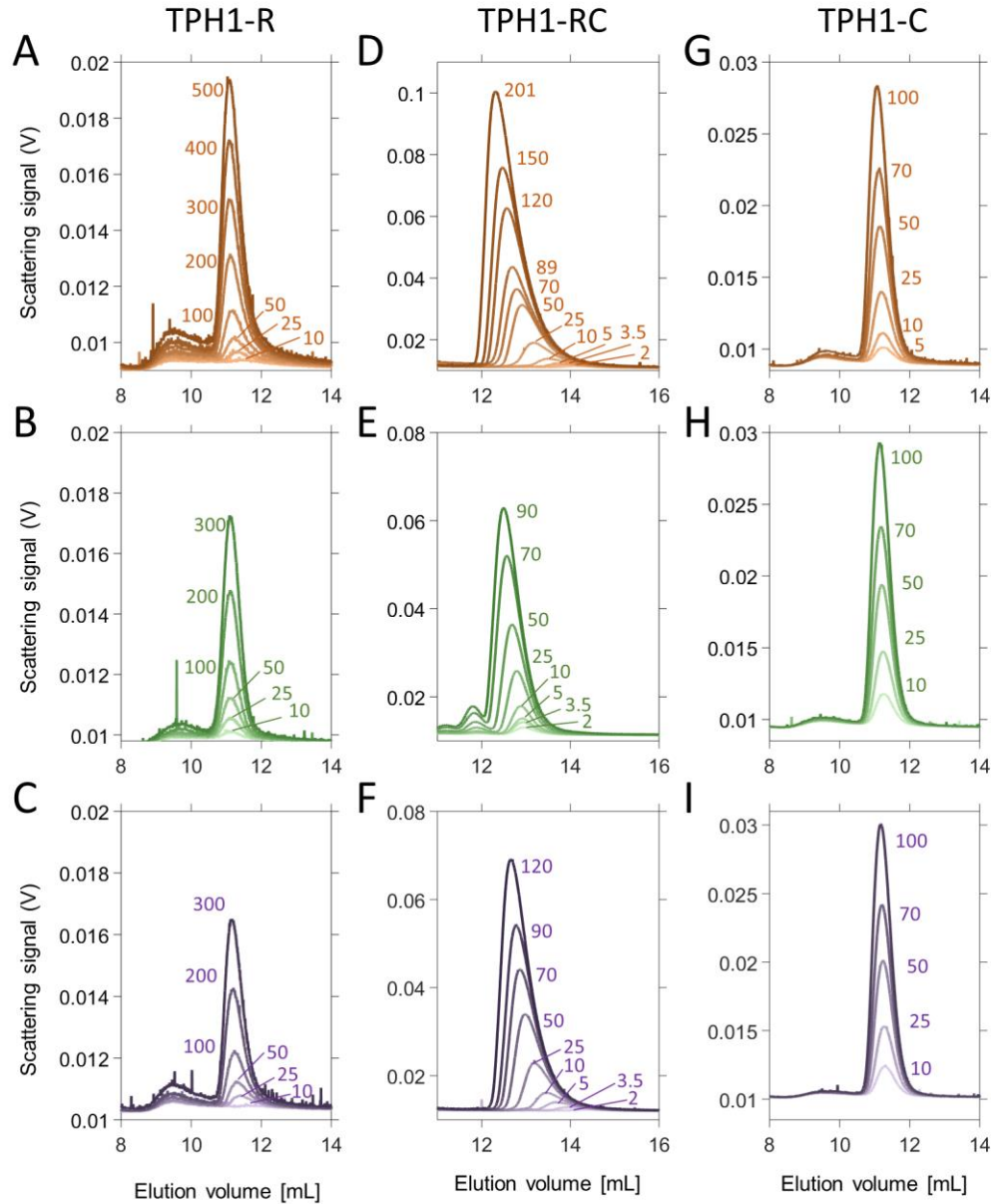
Supplementary figure 1: SEC-MALS elution profiles of TPH1-RC, TPH1-R, and TPH1-RC

Supplementary figure 2: SEC-MALS elution profiles of TPH1-RC with molecular weight traces

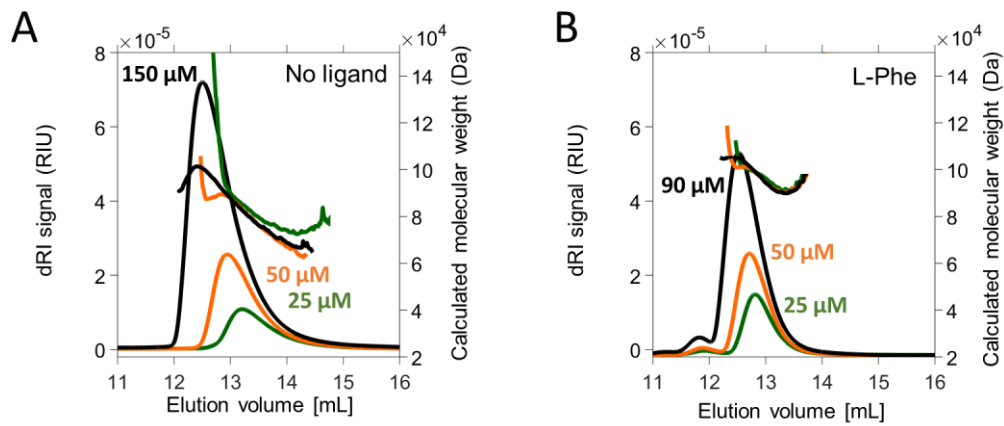
Supplementary figure 3: SDS-PAGE analysis of protein used for SEC-MALS

Supplementary figure 4: ^1H - ^{15}N HSQC spectra of C-terminal peak at different temperatures

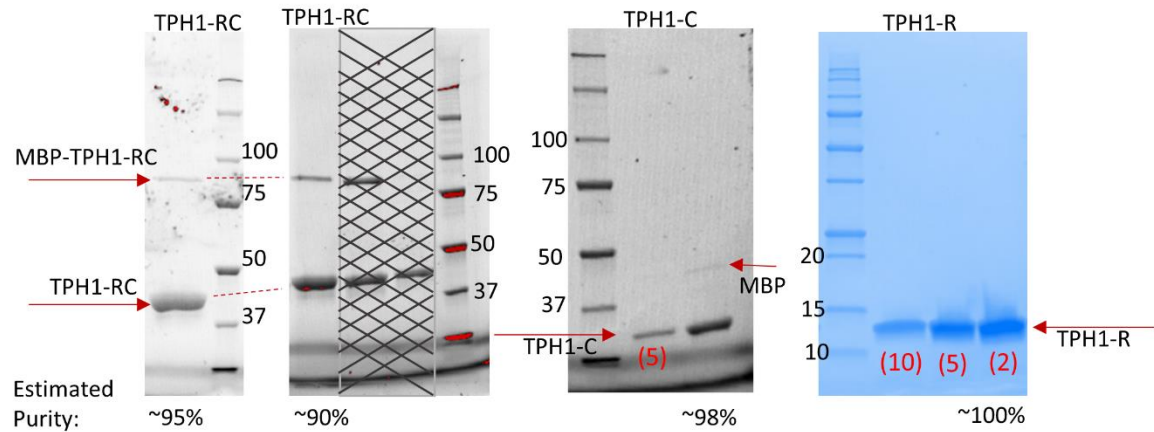
Supplementary figure 5: Titration of ^{13}C - L-Phe with TPH1-R



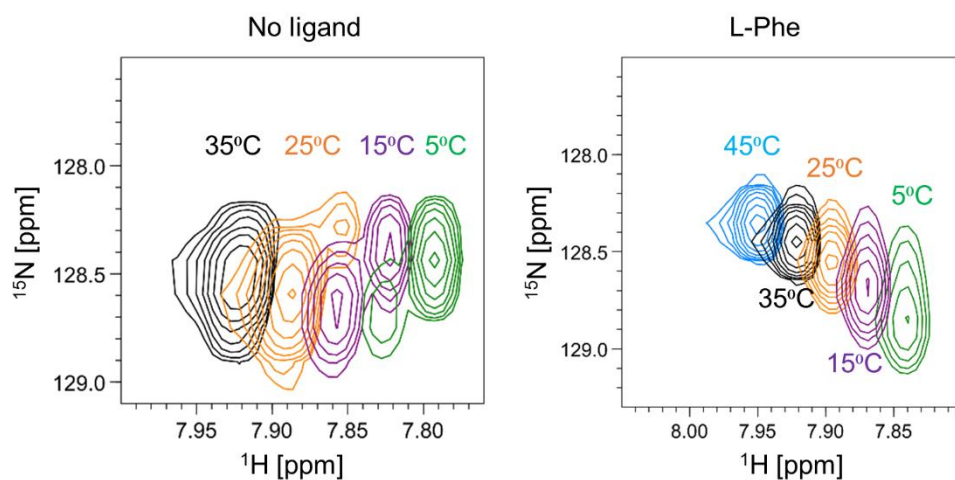
Supplementary figure 1 SEC-MALS elution profiles. A-C SEC-MALS elution profiles of TPH1-R with A) no ligand B) L-Phe C) L-Trp. D-F SEC-MALS elution profiles of TPH1-RC with D) no ligand E) L-Phe F) L-Trp. G-I SEC-MALS elution profiles of TPH1-C with G) no ligand H) L-Phe I) L-Trp. The numbers next to the curves indicate the protein concentration (μM) of the loaded samples



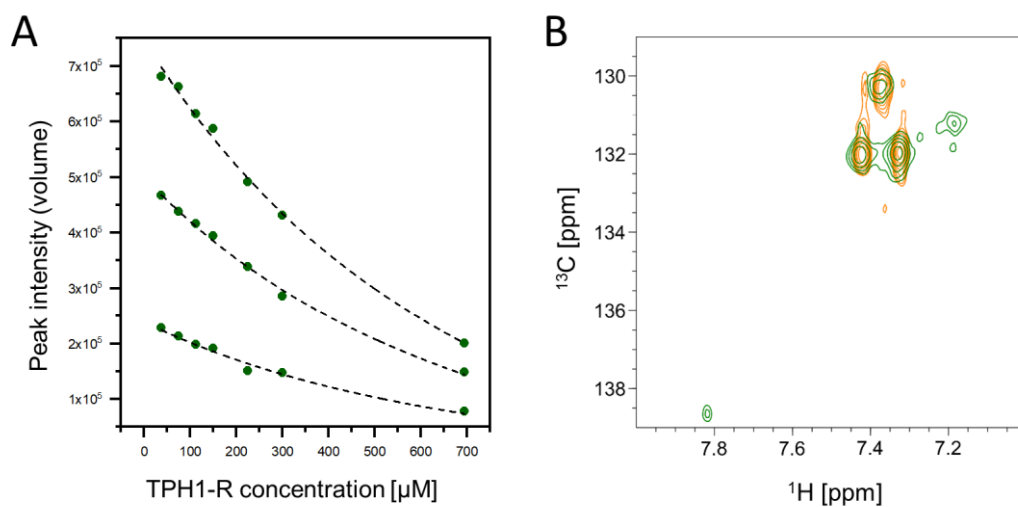
Supplementary figure 2 SEC-MALS elution profiles of TPH1-RC with A) no ligand and B) 10 mM L-Phe. The calculated molecular weight is plotted on top of the curves and are colored according to the equivalent elution curve.



Supplementary figure 3 SDS-PAGES of the pooled fractions from purifications of TPH1-RC, TPH1-C, and TPH1-R. The numbers (in black) indicate the molecular weight of the marker proteins. The numbers in parenthesis (in red) indicates how many times the sample was diluted if relevant. The grey mesh on the second TPH1-RC gels marks other samples from protein purification not relevant for these experiments. TPH1-RC was estimated to 90-95% purity, TPH1-C 98% purity, and no impurities could be detected on the gel for TPH1-R.



Supplementary figure 4 Zoom-in on the C-terminal peak in the ^1H - ^{15}N HSQC spectra of TPH1-R at different temperatures with and without L-Phe.



Supplementary figure 5 Titration of ^{13}C -L-Phe with TPH1-R. A) fit (black dashed line) to the decreasing intensities (Circles) of the three free L-Phe peaks. B) Aromatic ^1H - ^{13}C HSQC spectrum of $150\ \mu\text{M}$ ^{13}C -L-Phe with $0\ \mu\text{M}$ unlabeled TPH1-R (orange) and $695\ \mu\text{M}$ unlabeled TPH1-R (green).

6

CONCLUDING REMARKS

This project concerned the enzymes tryptophan hydroxylase 1 and tryptophan hydroxylase 2 that catalyze the rate-limiting step in the production of serotonin in the peripheral tissues and in the brain, respectively. Despite the importance of TPH1 and TPH2 in relation to serotonin related diseases, characterization of the regulatory domains and regulatory mechanisms of the enzymes have remained sparse. In this project, structural biology techniques and biophysical characterization methods have been used to gain an increased understanding of the structure and behavior of the RDs of TPH1 and TPH2.

In the presence of L-Phe, a 47 N-terminally truncated variant of the isolated RD of TPH2 (N Δ 47-TPH2-R) was well-behaved (less aggregation prone in protein concentrations up to at least 1.1 mM / 12.4 mg/mL), enabling a detailed structural study using NMR spectroscopy. Previous studies had pointed towards the RD dimer interface as the location of L-Phe binding^{47,52}, and the NMR structure of the N Δ 47-TPH2-R dimer determined in this study, confirmed that L-Phe was bound in two identical pockets in the RD dimer interface. Furthermore, the NMR spectroscopy experiments showed that the natural substrate of TPH2, L-Trp, was a poor ligand of the RD compared to L-Phe. Cryo-EM was used to obtain structural information on the RDs in tetrameric TPH2 (N Δ 47-TPH2). The results showed that the RDs form dimers also in the absence of L-Phe, and indicated that the RDs exist in a monomer-dimer equilibrium in the tetramer.

Investigating C-terminally truncated TPH1 variants using SEC-MALS showed that variants containing the RD (TPH1-R and TPH1-RC) mainly form dimers and indicated that the variants exist in a monomer-dimer equilibrium at lower protein concentrations. Furthermore, the presence of L-Phe was found to affect the elution volumes of TPH1-R and TPH1-RC, while L-Trp had no effect. Binding of L-Phe to the TPH1-R dimer was additionally confirmed using NMR spectroscopy, revealing that both TPH isoforms contain L-Phe binding site(s) in their RD, though the exact location of the binding site in TPH1 was not determined.

7

FUTURE PERSPECTIVES

In this study, the conformational landscape of the NΔ47-TPH2 tetramer has been further elucidated, and detailed structural information of the NΔ47-TPH2-R dimer showed the location of the L-Phe binding pockets. However, it is yet to be determined if ligand binding to the RD dimer has any functional effect on TPH2, which will be important to elucidate in further studies. This includes confirming whether L-Phe induces RD dimerization in NΔ47-TPH2 and full-length TPH2, and investigating if L-Phe binding to the RD affects e.g. the stability, kinetics, and phosphorylation of the enzyme. Additionally, due to the instability of the RD without L-Phe, detailed structural information of the un-bound form of the RD is still missing. A detailed structure of the RD without L-Phe would provide insights to the local structural effects of L-Phe binding, which could be important in mediating a potential allosteric response.

A detailed structural characterization of the RD of TPH1, both in context of L-Phe binding to the isolated domain and the RD position in the full-length enzyme, will be important to further increase the understanding of the isoform specific properties of TPH1 and TPH2. Here, the methods that have proven successful for the studies of TPH2 could be employed. Thus, the location of the RD in full-length TPH1 could be investigated using cryo-EM, or alternatively SAXS could be used to obtain a structural envelope of the TPH1 tetramer. NMR spectroscopy could be used to study the isolated RD in presence of L-Phe in order to investigate the exact location of RD binding site(s).

The most clear difference between TPH1 and TPH2 is the 46-residue longer N-tail of TPH2, of which no structural information is available. The longer N-tail in TPH2 could allow it to reach and cover the active site in the full-length enzyme with dimerized RDs, while the same would not be possible in TPH1, as also discussed by Zhang et al.⁷⁵. Thus, structural information on the first 46-residues in the N-tail of TPH2 is needed to elucidate isoform specific regulation of TPH2. However, this remains hampered by the decreased stability of variants containing the full N-tail, as also observed in Chapter 2 and Chapter 4. An alternative approach could be to express the N-tail of TPH2 alone, and investigate its possible interactions with the other domains. As the N-tail contains an isoform specific phosphorylation site, the effect of phosphorylation on the interactions of the N-tail could additionally provide pieces to the puzzle of TPH2 regulation. Structural effects of phosphorylation at the non-isoform specific site Ser104 (in TPH2), could be studied using the NΔ47-TPH2-R variant and NMR spectroscopy, to probe which residues that are affected by phosphorylation.

BIBLIOGRAPHY

1. Popova, N. K. & Kulikov, A. V. Targeting tryptophan hydroxylase 2 in affective disorder. *Expert Opin. Ther. Targets* **14**, 1259–1271 (2010).
2. Zhang, X. *et al.* Loss-of-Function Mutation in Tryptophan Hydroxylase-2 Identified in Unipolar Major Depression. *Neuron* **45**, 11–16 (2005).
3. Coates, M. D. *et al.* Molecular defects in mucosal serotonin content and decreased serotonin reuptake transporter in ulcerative colitis and irritable bowel syndrome. *Gastroenterology* **126**, 1657–1664 (2004).
4. Tsoli, M. & Kaltsas, G. Carcinoid Syndrome. in *Encyclopedia of Endocrine Diseases* 13–20 (Elsevier, 2019). doi:10.1016/B978-0-12-801238-3.65346-1
5. Fitzpatrick, P. F. Mechanism of Aromatic Amino Acid Hydroxylation. *Biochemistry* **42**, 14083–14091 (2003).
6. Walther, D. J. & Bader, M. A unique central tryptophan hydroxylase isoform. *Biochem. Pharmacol.* **66**, 1673–1680 (2003).
7. Bader, M. Inhibition of serotonin synthesis: A novel therapeutic paradigm. *Pharmacol. Ther.* **205**, 107423 (2020).
8. Veenstra-VanderWeele, J., Anderson, G. M. & Cook, E. H. Pharmacogenetics and the serotonin system: initial studies and future directions. *Eur. J. Pharmacol.* **410**, 165–181 (2000).
9. Walther, D. J. *et al.* Synthesis of Serotonin by a Second Tryptophan Hydroxylase Isoform. *Science (80-.)*. **299**, 76–76 (2003).
10. Udenfriend, S., Clark, C. T. & Titus, E. 5-Hydroxytryptophan Decarboxylase: A New Route of Metabolism of Tryptophan. *J. Am. Chem. Soc.* **75**, 501–502 (1953).
11. Fischer, M., Thöny, B. & Leimkühler, S. The Biosynthesis of Folate and Pterins and Their Enzymology. in *Comprehensive Natural Products II* 599–648 (Elsevier, 2010). doi:10.1016/B978-008045382-8.00150-7
12. Kaufman, S. The enzymatic conversion of phenylalanine to tyrosine. *J. Biol. Chem.* **226**, 511–524 (1957).
13. Molinoff, P. B. & Axelrod, J. Biochemistry of Catecholamines. *Annu. Rev. Biochem.* **40**, 465–500 (1971).
14. Weiner, N. Regulation of Norepinephrine Biosynthesis. *Annu. Rev. Pharmacol.* **10**, 273–290 (1970).
15. Kaneda, N. *et al.* Isolation of a novel cDNA clone for human tyrosine hydroxylase: Alternative RNA splicing produces four kinds of mRNA from a single gene. *Biochem. Biophys. Res. Commun.* **146**, 971–975 (1987).
16. Cansev, M. & Wurtman, R. J. Aromatic Amino Acids in the Brain. in *Handbook of Neurochemistry and Molecular Neurobiology* 59–97 (Springer US, 2007). doi:10.1007/978-0-387-30373-4_4

17. Cleary, M. *et al.* Fluctuations in phenylalanine concentrations in phenylketonuria: A review of possible relationships with outcomes. *Mol. Genet. Metab.* **110**, 418–423 (2013).
18. de Groot, M. J., Hoeksma, M., Blau, N., Reijngoud, D. J. & van Spronsen, F. J. Pathogenesis of cognitive dysfunction in phenylketonuria: Review of hypotheses. *Mol. Genet. Metab.* **99**, S86–S89 (2010).
19. Fernando, J. C. R., Knott, P. J., Curzon, G. The relevance of both plasma free tryptophan and insulin to rat brain tryptophan concentration. *J. Neurochem.* **27**, 343–345 (1976).
20. Berger, M., Gray, J. A. & Roth, B. L. The Expanded Biology of Serotonin. *Annu. Rev. Med.* **60**, 355–366 (2009).
21. Lucki, I. The spectrum of behaviors influenced by serotonin. *Biol. Psychiatry* **44**, 151–162 (1998).
22. Waider, J., Araragi, N., Gutknecht, L. & Lesch, K.-P. Tryptophan hydroxylase-2 (TPH2) in disorders of cognitive control and emotion regulation: A perspective. *Psychoneuroendocrinology* **36**, 393–405 (2011).
23. Immadisetty, K., Geffert, L. M., Surratt, C. K. & Madura, J. D. New design strategies for antidepressant drugs. *Expert Opin. Drug Discov.* **8**, 1399–1414 (2013).
24. Masab, M. & Saif, M. W. Telotristat ethyl: proof of principle and the first oral agent in the management of well-differentiated metastatic neuroendocrine tumor and carcinoid syndrome diarrhea. *Cancer Chemother. Pharmacol.* **80**, 1055–1062 (2017).
25. Bae, E. J. *et al.* Peripheral Selective Oxadiazolylphenyl Alanine Derivatives as Tryptophan Hydroxylase 1 Inhibitors for Obesity and Fatty Liver Disease. *J. Med. Chem.* **64**, 1037–1053 (2021).
26. Petrassi, M. *et al.* Identification of a Novel Allosteric Inhibitory Site on Tryptophan Hydroxylase 1 Enabling Unprecedented Selectivity Over all Related Hydroxylases. *Front. Pharmacol.* **8**, (2017).
27. Scriver, C. R. The PAH gene, phenylketonuria, and a paradigm shift. *Hum. Mutat.* **28**, 831–845 (2007).
28. Blau, N., Hennermann, J. B., Langenbeck, U. & Lichter-Konecki, U. Diagnosis, classification, and genetics of phenylketonuria and tetrahydrobiopterin (BH4) deficiencies. *Mol. Genet. Metab.* **104**, S2–S9 (2011).
29. van Spronsen, F. J. Phenylketonuria: a 21st century perspective. *Nat. Rev. Endocrinol.* **6**, 509–514 (2010).
30. Daubner, S. C., Le, T. & Wang, S. Tyrosine hydroxylase and regulation of dopamine synthesis. *Arch. Biochem. Biophys.* **508**, 1–12 (2011).
31. Tripp, G. & Wickens, J. R. Neurobiology of ADHD. *Neuropharmacology* **57**, 579–589 (2009).
32. Tabrez, S. *et al.* A Synopsis on the Role of Tyrosine Hydroxylase in Parkinson's Disease. *CNS Neurol. Disord. - Drug Targets* **11**, 395–409 (2012).
33. Carkaci-Salli, N. *et al.* Functional Domains of Human Tryptophan Hydroxylase 2 (hTPH2). *J. Biol. Chem.* **281**, 28105–28112 (2006).
34. Arturo, E. C. *et al.* First structure of full-length mammalian phenylalanine hydroxylase reveals the architecture of an autoinhibited tetramer. *Proc. Natl. Acad. Sci.* **113**, 2394–

- 2399 (2016).
35. Yohrling, G. J., Jiang, G. C.-T., Mockus, S. M. & Vrana, K. E. Intersubunit binding domains within tyrosine hydroxylase and tryptophan hydroxylase. *J. Neurosci. Res.* **61**, 313–320 (2000).
 36. Bueno-Carrasco, M. T. *et al.* Structural mechanism for tyrosine hydroxylase inhibition by dopamine and reactivation by Ser40 phosphorylation. *Nat. Commun.* **13**, 74 (2022).
 37. Andersen, O. A., Flatmark, T. & Hough, E. High resolution crystal structures of the catalytic domain of human phenylalanine hydroxylase in its catalytically active Fe(II) form and binary complex with tetrahydrobiopterin. *J. Mol. Biol.* **314**, 279–291 (2001).
 38. Arturo, E. C., Gupta, K., Hansen, M. R., Borne, E. & Jaffe, E. K. Biophysical characterization of full-length human phenylalanine hydroxylase provides a deeper understanding of its quaternary structure equilibrium. *J. Biol. Chem.* **294**, 10131–10145 (2019).
 39. Patel, D., Kopec, J., Fitzpatrick, F., McCorvie, T. J. & Yue, W. W. Structural basis for ligand-dependent dimerization of phenylalanine hydroxylase regulatory domain. *Sci. Rep.* **6**, 23748 (2016).
 40. Kobe, B. *et al.* Structural basis of autoregulation of phenylalanine hydroxylase. *Nat. Struct. Biol.* **6**, 442–8 (1999).
 41. Meisburger, S. P. *et al.* Domain movements upon activation of phenylalanine hydroxylase characterized by crystallography and chromatography-coupled small-angle X-ray scattering. *J. Am. Chem. Soc.* **138**, 6506–6516 (2016).
 42. Erlandsen, H., Flatmark, T., Stevens, R. C. & Hough, E. Crystallographic Analysis of the Human Phenylalanine Hydroxylase Catalytic Domain with Bound Catechol Inhibitors at 2.0 Å Resolution. *Biochemistry* **37**, 15638–15646 (1998).
 43. Martinez, A. *et al.* Expression of recombinant human phenylalanine hydroxylase as fusion protein in *Escherichia coli* circumvents proteolytic degradation by host cell proteases. Isolation and characterization of the wild-type enzyme. *Biochem. J.* **306**, 589–597 (1995).
 44. Szigetvari, P. D. *et al.* The quaternary structure of human tyrosine hydroxylase: effects of dystonia-associated missense variants on oligomeric state and enzyme activity. *J. Neurochem.* **148**, 291–306 (2019).
 45. Kopec, J., Oberholzer, A., Fitzpatrick, F., Newman, J., Tallant, C., Kiyani, W., Shrestha, L., Burgess-brown, N., von Delft, F., Arrowsmith, C., Edwards, A., Bountra, C., Yue, W. W. Crystal Structure of Human Tryptophan Hydroxylase 2 (Tph2), Catalytic Domain, pdb entry 4V06. *Protein Databank* (2014).
 46. Bezem, M. T. *et al.* Stable preparations of tyrosine hydroxylase provide the solution structure of the full-length enzyme. *Sci. Rep.* **6**, 30390 (2016).
 47. Skawinska, N. T. Allosteric regulation of human tryptophan hydroxylase isoform 2 (hTPH2). (Technical University of Denmark, 2020).
 48. Flydal, M. I. *et al.* Structure of full-length human phenylalanine hydroxylase in complex with tetrahydrobiopterin. *Proc. Natl. Acad. Sci.* **116**, 11229–11234 (2019).
 49. Wang, L., Erlandsen, H., Haavik, J., Knappskog, P. M. & Stevens, R. C. Three-dimensional structure of human tryptophan hydroxylase and its implications for the biosynthesis of the neurotransmitters serotonin and melatonin. *Biochemistry* **41**, 12569–12574 (2002).

50. Muiz, J. R. C., Cooper, C. D. O, Yue, W. W., Krysztofinska, E., von Delft, F., Knapp, S., Gileadi, O., Arrowsmith, C. H., Edwards, A. M., Weigelt, J., Bountra, C., Kavanagh, K. L., Oppermann, U. Crystal Structure of Human Tyrosine Hydroxylase Catalytic Domain, PDB entry 2XSN. *Protein Databank* (2010).
51. Windahl, M. S., Boesen, J., Karlsen, P. E. & Christensen, H. E. M. Expression, Purification and Enzymatic Characterization of the Catalytic Domains of Human Tryptophan Hydroxylase Isoforms. *Protein J.* **28**, 400–406 (2009).
52. Tidemand, K. D. *et al.* Stabilization of tryptophan hydroxylase 2 by L-phenylalanine-induced dimerization. *FEBS Open Bio* 1–13 (2016). doi:10.1002/2211-5463.12100
53. Moran, G. R., Daubner, S. C. & Fitzpatrick, P. F. Expression and Characterization of the Catalytic Core of Tryptophan Hydroxylase. *J. Biol. Chem.* **273**, 12259–12266 (1998).
54. Windahl, M. S., Petersen, C. R., Christensen, H. E. M. & Harris, P. Crystal Structure of Tryptophan Hydroxylase with Bound Amino Acid Substrate. *Biochemistry* **47**, 12087–12094 (2008).
55. Roberts, K. M. & Fitzpatrick, P. F. Mechanisms of tryptophan and tyrosine hydroxylase. *IUBMB Life* **65**, 350–357 (2013).
56. Tidemand, K. D., Peters, G. H., Harris, P., Stensgaard, E. & Christensen, H. E. M. Isoform-Specific Substrate Inhibition Mechanism of Human Tryptophan Hydroxylase. *Biochemistry* **56**, 6155–6164 (2017).
57. McKinney, J., Knappskog, P. M. & Haavik, J. Different properties of the central and peripheral forms of human tryptophan hydroxylase. *J. Neurochem.* **92**, 311–320 (2005).
58. Cichon, S. *et al.* a functional Pro206Ser substitution and variation in the 5'0'-region are associated with bipolar affective disorder. **17**, (2008).
59. Winge, I., McKinney, J. A., Knappskog, P. M. & Haavik, J. Characterization of wild-type and mutant forms of human tryptophan hydroxylase 2. *J. Neurochem.* **100**, 1648–1657 (2007).
60. McKinney, J. *et al.* Conformation of the Substrate and Pterin Cofactor Bound to Human Tryptophan Hydroxylase. Important Role of Phe313 in Substrate Specificity. *Biochemistry* **40**, 15591–15601 (2001).
61. Mordhorst, A. *et al.* Phenylalanine hydroxylase contributes to serotonin synthesis in mice. *FASEB J.* **35**, (2021).
62. McKinney, J. *et al.* Expression and purification of human tryptophan hydroxylase from *Escherichia coli* and *Pichia pastoris*. *Protein Expr. Purif.* **33**, 185–194 (2004).
63. Murphy, K. L., Zhang, X., Gainetdinov, R. R., Beaulieu, J.-M. & Caron, M. G. A Regulatory Domain in the N Terminus of Tryptophan Hydroxylase 2 Controls Enzyme Expression. *J. Biol. Chem.* **283**, 13216–13224 (2008).
64. Zhang, S., Huang, T., Ilangovan, U., Hinck, A. P. & Fitzpatrick, P. F. The Solution structure of the Regulatory domain of Tyrosine Hydroxylase. *J. Mol. Biol.* 1483–1497 (2014). doi:10.1016/j.pestbp.2011.02.012. Investigations
65. Aravind, L. & Koonin, E. V. Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. *J. Mol. Biol.* **287**, 1023–1040 (1999).
66. Lang, E. J., Cross, P. J., Mittelstädt, G., Jameson, G. B. & Parker, E. J. Allosteric ACTion: the varied ACT domains regulating enzymes of amino-acid metabolism. *Curr.*

- Opin. Struct. Biol.* **29**, 102–111 (2014).
67. Grant, G. A., Schuller, D. J. & Banaszak, L. J. A model for the regulation of D-3-phosphoglycerate dehydrogenase, a Vmax-type allosteric enzyme: Regulation of phosphoglycerate dehydrogenase. *Protein Sci.* **5**, 34–41 (1996).
 68. Cross, P. J., Dobson, R. C. J., Patchett, M. L. & Parker, E. J. Tyrosine Latching of a Regulatory Gate Affords Allosteric Control of Aromatic Amino Acid Biosynthesis. *J. Biol. Chem.* **286**, 10216–10224 (2011).
 69. Cho, Y., Sharma, V. & Sacchettini, J. C. Crystal Structure of ATP Phosphoribosyltransferase from *Mycobacterium tuberculosis*. *J. Biol. Chem.* **278**, 8333–8339 (2003).
 70. Schreiter, E. R. *et al.* Crystal structure of the nickel-responsive transcription factor NikR. *Nat. Struct. Mol. Biol.* **10**, 794–799 (2003).
 71. Grant, G. A. The ACT Domain: A Small Molecule Binding Domain and Its Role as a Common Regulatory Element. *J. Biol. Chem.* **281**, 33825–33829 (2006).
 72. Liberles, J. S., Thorolfsson, M. & Martinez, A. Allosteric mechanisms in ACT domain containing enzymes involved in amino acid metabolism. *Amino Acids* **28**, 1–12 (2005).
 73. Zhang, S., Roberts, K. M. & Fitzpatrick, P. F. Phenylalanine binding is linked to dimerization of the regulatory domain of phenylalanine hydroxylase. *Biochemistry* **53**, 6625–6627 (2014).
 74. Neira, J. L., Hornos, F., Bacarizo, J., Cámara-Artigás, A. & Gómez, J. The Monomeric Species of the Regulatory Domain of Tyrosine Hydroxylase Has a Low Conformational Stability. *Biochemistry* **55**, 3418–3431 (2016).
 75. Zhang, S., Hinck, C. S. & Fitzpatrick, P. F. The regulatory domain of human tryptophan hydroxylase 1 forms a stable dimer. *Biochem. Biophys. Res. Commun.* **476**, 457–461 (2016).
 76. Tourian, A. Activation of phenylalanine hydroxylase by phenylalanine. *Biochim. Biophys. Acta - Enzymol.* **242**, 345–354 (1971).
 77. Thórolfsson, M. *et al.* L-Phenylalanine Binding and Domain Organization in Human Phenylalanine Hydroxylase: A Differential Scanning Calorimetry Study. *Biochemistry* **41**, 7573–7585 (2002).
 78. Andersen, O. A., Stokka, A. J., Flatmark, T. & Hough, E. 2.0Å Resolution Crystal Structures of the Ternary Complexes of Human Phenylalanine Hydroxylase Catalytic Domain with Tetrahydrobiopterin and 3-(2-Thienyl)-L-alanine or L-Norleucine: Substrate Specificity and Molecular Motions Related to Substrate Binding. *J. Mol. Biol.* **333**, 747–757 (2003).
 79. Carluccio, C., Fraternali, F., Salvatore, F., Fornili, A. & Zagari, A. Towards the identification of the allosteric Phe-binding site in phenylalanine hydroxylase. *J. Biomol. Struct. Dyn.* **34**, 497–507 (2016).
 80. Arturo, E. C. *et al.* Manipulation of a cation- π sandwich reveals conformational flexibility in phenylalanine hydroxylase. *Biochimie* **183**, 63–77 (2021).
 81. Ge, Y. *et al.* Simulations of the regulatory ACT domain of human phenylalanine hydroxylase (PAH) unveil its mechanism of phenylalanine binding. *J. Biol. Chem.* **293**, 19532–19543 (2018).
 82. Jaffe, E. K., Stith, L., Lawrence, S. H., Andrade, M. & Dunbrack, R. L. A new model for allosteric regulation of phenylalanine hydroxylase: Implications for disease and

- therapeutics. *Arch. Biochem. Biophys.* **530**, 73–82 (2013).
83. Jennings, I. G., Teh, T. & Kobe, B. Essential role of the N-terminal autoregulatory sequence in the regulation of phenylalanine hydroxylase. *FEBS Lett.* **488**, 196–200 (2001).
 84. Khan, C. A. & Fitzpatrick, P. F. Phosphorylation of Phenylalanine Hydroxylase Increases the Rate Constant for Formation of the Activated Conformation of the Enzyme. *Biochemistry* **57**, 6274–6277 (2018).
 85. Fitzpatrick, P. F. Allosteric regulation of phenylalanine hydroxylase. *Arch. Biochem. Biophys.* **519**, 194–201 (2012).
 86. Andersson, K. K., Cox, D. D., Que, L., Flatmark, T. & Haavik, J. Resonance Raman studies on the blue-green-colored bovine adrenal tyrosine 3-monooxygenase (tyrosine hydroxylase). Evidence that the feedback inhibitors adrenaline and noradrenaline are coordinated to iron. *J. Biol. Chem.* **263**, 18621–18626 (1988).
 87. Almas, B., Bourdelles, B., Flatmark, T., Mallet, J. & Haavik, J. Regulation of recombinant human tyrosine hydroxylase isozymes by catecholamine binding and phosphorylation. Structure/activity studies and mechanistic implications. *Eur. J. Biochem.* **209**, 249–255 (1992).
 88. Sura, G. R., Daubner, S. C. & Fitzpatrick, P. F. Effects of phosphorylation by protein kinase A on binding of catecholamines to the human tyrosine hydroxylase isoforms. *J. Neurochem.* **90**, 970–978 (2004).
 89. Ramsey, A. J. & Fitzpatrick, P. F. Effects of Phosphorylation of Serine 40 of Tyrosine Hydroxylase on Binding of Catecholamines: Evidence for a Novel Regulatory Mechanism. *Biochemistry* **37**, 8980–8986 (1998).
 90. Haycock, J. W., Ahn, N. G., Cobb, M. H. & Krebs, E. G. ERK1 and ERK2, two microtubule-associated protein 2 kinases, mediate the phosphorylation of tyrosine hydroxylase at serine-31 in situ. *Proc. Natl. Acad. Sci.* **89**, 2365–2369 (1992).
 91. Bevilacqua, L. R. M., Graham, M. E., Dunkley, P. R., von Nagy-Felsobuki, E. I. & Dickson, P. W. Phosphorylation of Ser19 Alters the Conformation of Tyrosine Hydroxylase to Increase the Rate of Phosphorylation of Ser40. *J. Biol. Chem.* **276**, 40411–40416 (2001).
 92. Kleppe, R. *et al.* Phosphorylation Dependence and Stoichiometry of the Complex Formed by Tyrosine Hydroxylase and 14-3-3 γ . *Mol. Cell. Proteomics* **13**, 2017–2030 (2014).
 93. Aitken, A. 14-3-3 proteins: A historic overview. *Semin. Cancer Biol.* **16**, 162–172 (2006).
 94. Steinacker, P., Aitken, A. & Otto, M. 14-3-3 proteins in neurodegeneration. *Semin. Cell Dev. Biol.* **22**, 696–704 (2011).
 95. Skjevik, Å. A. *et al.* The N-Terminal Sequence of Tyrosine Hydroxylase Is a Conformationally Versatile Motif That Binds 14-3-3 Proteins and Membranes. *J. Mol. Biol.* **426**, 150–168 (2014).
 96. Winge, I. *et al.* Activation and stabilization of human tryptophan hydroxylase 2 by phosphorylation and 14-3-3 binding. *Biochem. J.* **410**, 195–204 (2008).
 97. Kuhn, D. M., Arthur, R. & States, J. C. Phosphorylation and activation of brain tryptophan hydroxylase: Identification of serine-58 as a substrate site for protein kinase A. *J. Neurochem.* **68**, 2220–2223 (1997).
 98. Jiang, G. C.-T., Yohrling, G. ., Schmitt, I. V. J. . & Vrana, K. . Identification of substrate

- orienting and phosphorylation sites within tryptophan hydroxylase using homology-based molecular modeling. *J. Mol. Biol.* **302**, 1005–1017 (2000).
99. Kuhn, D. M., Sakowski, S. A., Geddes, T. J., Wilkerson, C. & Haycock, J. W. Phosphorylation and activation of tryptophan hydroxylase 2: identification of serine-19 as the substrate site for calcium, calmodulin-dependent protein kinase II. *J. Neurochem.* **103**, 1567–1573 (2007).
 100. Tenner, K., Walther, D. & Bader, M. Influence of human tryptophan hydroxylase 2 N- and C-terminus on enzymatic activity and oligomerization. *J. Neurochem.* **102**, 1887–1894 (2007).
 101. Nielsen, M. S. *et al.* A simple two step procedure for purification of the catalytic domain of chicken tryptophan hydroxylase 1 in a form suitable for crystallization. *Protein Expr. Purif.* **57**, 116–126 (2008).
 102. Boesen, J. Cloning, expression, purification and characterization of tryptophan hydroxylase variants. (Technical University of Denmark, 2010).
 103. Haahr, L. T. Purification and Characterization of Tryptophan Hydroxylase. *Technical University of Denmark* (Technical University of Denmark, 2012).
 104. Tidemand, K. D. Characterization of the human tryptophan hydroxylase isoforms. (Technical University of Denmark, 2017).
 105. Gasteiger, E. *et al.* Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook* 571–607 (Humana Press, 2005). doi:10.1385/1-59259-890-0:571
 106. Barer, R., Joseph, S. Refractometry of living cells.1. Basic Principles. *Q. J. Microsc. Sci.* **95**, 399–423 (1954).
 107. EMBL Hamburg. Refractometer measurements. Available at: <https://www.embl-hamburg.de/biosaxs/manuals/equipment/refractometer.html>.
 108. King, G. F. & Mobli, M. Derivation of Peptide and Protein Structure using NMR Spectroscopy. in *Comprehensive Natural Products II* 279–325 (Elsevier, 2010). doi:10.1016/B978-008045382-8.00653-5
 109. Keeler, J. *Understanding NMR spectroscopy*. (Wiley, 2005).
 110. Wider, G. Structure Determination of Biological Macromolecules in Solution Using Nuclear Magnetic Resonance Spectroscopy. *Biotechniques* **29**, 1278–1294 (2000).
 111. Kanelis, V., Forman-Kay, J. D. & Kay, L. E. Multidimensional NMR Methods for Protein Structure Determination. *IUBMB Life (International Union Biochem. Mol. Biol. Life)* **52**, 291–302 (2001).
 112. Zuiderweg, E. R. P. & Fesik, S. W. Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. *Biochemistry* **28**, 2387–2391 (1989).
 113. Bostock, M. & Nietlispach, D. Compressed sensing: Reconstruction of non-uniformly sampled multidimensional NMR data. *Concepts Magn. Reson. Part A* **46A**, e21438 (2017).
 114. Ferella, L., Rosato, A. & Turano, P. Determination of Protein Structure and Dynamics. in *NMR of Biomolecules* 51–94 (Wiley-VCH Verlag GmbH & Co. KGaA, 2012). doi:10.1002/9783527644506.ch4
 115. Teilum, K., Kunze, M. B. A., Erlendsson, S. & Kragelund, B. B. (S)Pinning down protein interactions by NMR. *Protein Sci.* **26**, 436–451 (2017).

116. Orts, J. & Gossert, A. D. Structure determination of protein-ligand complexes by NMR in solution. *Methods* **138–139**, 3–25 (2018).
117. Furukawa, A., Konuma, T., Yanaka, S. & Sugase, K. Quantitative analysis of protein–ligand interactions by NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **96**, 47–57 (2016).
118. Waudby, C. A., Ramos, A., Cabrita, L. D. & Christodoulou, J. Two-Dimensional NMR Lineshape Analysis. *Sci. Rep.* **6**, 24826 (2016).
119. Chao, F.-A. & Byrd, R. A. Protein dynamics revealed by NMR relaxation methods. *Emerg. Top. Life Sci.* **2**, 93–105 (2018).
120. Jarymowycz, V. A. & Stone, M. J. Fast Time Scale Dynamics of Protein Backbones: NMR Relaxation Methods, Applications, and Functional Consequences. *Chem. Rev.* **106**, 1624–1671 (2006).
121. Bax, A., Clore, G. M. & Gronenborn, A. M. 1H-1H correlation via isotropic mixing of ¹³C magnetization, a new three-dimensional approach for assigning 1H and ¹³C spectra of ¹³C-enriched proteins. *J. Magn. Reson.* **88**, 425–431 (1990).
122. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
123. Güntert, P. & Buchner, L. Combined automated NOE assignment and structure calculation with CYANA. *J. Biomol. NMR* **62**, 453–471 (2015).
124. Schwieters, C. D., Bermejo, G. A. & Clore, G. M. Xplor-NIH for molecular structure determination from NMR and other data sources. *Protein Sci.* **27**, 26–40 (2018).
125. Tian, Y., Schwieters, C. D., Opella, S. J. & Marassi, F. M. High quality NMR structures: a new force field with implicit water and membrane solvation for Xplor-NIH. *J. Biomol. NMR* **67**, 35–49 (2017).
126. Montelione, G. T. *et al.* Recommendations of the wwPDB NMR Validation Task Force. *Structure* **21**, 1563–1570 (2013).
127. Vranken, W. F. NMR structure validation in relation to dynamics and structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **82**, 27–38 (2014).
128. Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40**, 191–285 (2007).
129. Jacques, D. A. & Trehwella, J. Small-angle scattering for structural biology-Expanding the frontier while avoiding the pitfalls. *Protein Sci.* **19**, 642–657 (2010).
130. Franke, D. *et al.* ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212–1225 (2017).
131. Koch, M. H. J., Vachette, P. & Svergun, D. I. Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q. Rev. Biophys.* **36**, 147–227 (2003).
132. Trehwella, J. Small-angle scattering and 3D structure interpretation. *Curr. Opin. Struct. Biol.* **40**, 1–7 (2016).
133. Hermann, M. R. & Hub, J. S. Interpreting SAXS/WAXS Data with Explicit-Solvent Simulations: A Practical Guide. *Methods Mol. Biol.* **2168**, 199–215 (2020).
134. Tria, G., Mertens, H. D. T., Kachala, M. & Svergun, D. I. Advanced ensemble modelling

- of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207–217 (2015).
135. Pérez, J. & Nishino, Y. Advances in X-ray scattering: from solution SAXS to achievements with coherent beams. *Curr. Opin. Struct. Biol.* **22**, 670–678 (2012).
 136. Some, D., Amartely, H., Tsadok, A. & Lebendiker, M. Characterization of Proteins by Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS). *J. Vis. Exp.* (2019). doi:10.3791/59615
 137. Wyatt, P. J. Light scattering and the absolute characterization of macromolecules. *Anal. Chim. Acta* **272**, 1–40 (1993).
 138. Minton, A. P. Recent applications of light scattering measurement in the biological and biopharmaceutical sciences. *Anal. Biochem.* **501**, 4–22 (2016).
 139. Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157–161 (2020).
 140. Nakane, T. *et al.* Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152–156 (2020).
 141. Cheng, Y. Membrane protein structural biology in the era of single particle cryo-EM. *Curr. Opin. Struct. Biol.* **52**, 58–63 (2018).
 142. Liu, Y., Huynh, D. T. & Yeates, T. O. A 3.8 Å resolution cryo-EM structure of a small protein bound to an imaging scaffold. *Nat. Commun.* **10**, 1864 (2019).
 143. Fan, X. *et al.* Single particle cryo-EM reconstruction of 52 kDa streptavidin at 3.2 Angstrom resolution. *Nat. Commun.* **10**, 2386 (2019).
 144. Dubochet, J., Chang, J. J., Freeman, R., Lepault, J. & McDowell, A. W. Frozen aqueous suspensions. *Ultramicroscopy* **10**, 55–61 (1982).
 145. Cheng, Y. Single-Particle Cryo-EM at Crystallographic Resolution. *Cell* **161**, 450–457 (2015).
 146. Weissenberger, G., Henderikx, R. J. M. & Peters, P. J. Understanding the invisible hands of sample preparation for cryo-EM. *Nat. Methods* **18**, 463–471 (2021).
 147. Glaeser, R. M. *et al.* Factors that Influence the Formation and Stability of Thin, Cryo-EM Specimens. *Biophys. J.* **110**, 749–755 (2016).
 148. Moran, G. R. & Fitzpatrick, P. F. A continuous fluorescence assay for tryptophan hydroxylase. *Anal. Biochem.* **266**, 148–152 (1999).
 149. Carkaci-Salli, N. *et al.* Functional Domains of Human Tryptophan Hydroxylase 2 (hTPH2). *J. Biol. Chem.* **281**, 28105–28112 (2006).
 150. Ogawa, S. & Ichinose, H. Effect of metals and phenylalanine on the activity of human tryptophan hydroxylase-2: Comparison with that on tyrosine hydroxylase activity. *Neurosci. Lett.* **401**, 261–265 (2006).
 151. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
 152. Flydal, M. I. & Martinez, A. Phenylalanine hydroxylase: Function, structure, and regulation. *IUBMB Life* **65**, 341–349 (2013).

APPENDIX



SEQUENCE ALIGNMENTS

A1 ALIGNMENT OF THE *HUMAN* AAHs

The alignment continues on the following page

TH isoform 2 (P07101-3)	MPTPDATTPQAKGFRRVSELDAKQAEAIM-----	30
TH isoform 4 (P07101-4)	MPTPDATTPQAKGFRRVSELDAKQAEAIMVRGQ-----	34
TH isoform 3 (P07101-1)	MPTPDATTPQAKGFRRVSELDAKQAEAIMVRGQGAPGSLTGSPWPGTAAAPAASYTPTP	60
TH isoform 1 (P07101-2)	MPTPDATTPQAKGFRRVSELDAKQAEAIM---GAPGSLTGSPWPGTAAAPAASYTPTP	56
PAH (P00439-1)	-----	0
TPH isoform 1 (P17752-1)	-----	0
TPH isoform 2 (Q8IWU9-1)	-----MQPAMM-----MFSSKYWA-----	14
TH isoform 2 (P07101-3)	-SPRFIGRRQSLIEDARKEREAAVAAAAAAVPSEPGDP--LEAVAF-----EEKEGKAV	81
TH isoform 4 (P07101-4)	-SPRFIGRRQSLIEDARKEREAAVAAAAAAVPSEPGDP--LEAVAF-----EEKEGKAV	85
TH isoform 3 (P07101-1)	RSPRFIGRRQSLIEDARKEREAAVAAAAAAVPSEPGDP--LEAVAF-----EEKEGKAV	112
TH isoform 1 (P07101-2)	RSPRFIGRRQSLIEDARKEREAAVAAAAAAVPSEPGDP--LEAVAF-----EEKEGKAV	108
PAH (P00439-1)	-----MST---AVLENPGLGRKLSDFGQETS YIEDNCNQNGA	34
TPH isoform 1 (P17752-1)	-----MIEDNKENKDHSLERGR	17
TPH isoform 2 (Q8IWU9-1)	-----RRGFSLDSAVPEEHQLLGS---STLNKPN SGK-NDDKGNKGSSKREAAATESGK	63
TH isoform 2 (P07101-3)	LNLLFSPRATKPSALSRAVKVFETFEAKIHHLETRPAQRPRAGGPHLEYFVRLEVRRG-D	140
TH isoform 4 (P07101-4)	LNLLFSPRATKPSALSRAVKVFETFEAKIHHLETRPAQRPRAGGPHLEYFVRLEVRRG-D	144
TH isoform 3 (P07101-1)	LNLLFSPRATKPSALSRAVKVFETFEAKIHHLETRPAQRPRAGGPHLEYFVRLEVRRG-D	171
TH isoform 1 (P07101-2)	LNLLFSPRATKPSALSRAVKVFETFEAKIHHLETRPAQRPRAGGPHLEYFVRLEVRRG-D	167
PAH (P00439-1)	ISLIFSLK-EEVGALAKVLRLEFENDVNLTHIESRPSRLK---KDEYEFFTHLDKRS LPA	90
TPH isoform 1 (P17752-1)	ASLIFSLK-NEVGGLIKALKIFQEKHVNLLHIESRKSRR---NSEFEIFVDCDINRE-Q	72
TPH isoform 2 (Q8IWU9-1)	TAVVFSLK-NEVGGLVKALRLFQEKRVNMVHIESRKSRR---SSEVEIFVDCCEGKT-E	118
TH isoform 2 (P07101-3)	LAALLSGV-----RQVSEDVRSPAGPKVPWFPRKVSELDKCHHLVTKFDPDLDL	190
TH isoform 4 (P07101-4)	LAALLSGV-----RQVSEDVRSPAGPKVPWFPRKVSELDKCHHLVTKFDPDLDL	194
TH isoform 3 (P07101-1)	LAALLSGV-----RQVSEDVRSPAGPKVPWFPRKVSELDKCHHLVTKFDPDLDL	221
TH isoform 1 (P07101-2)	LAALLSGV-----RQVSEDVRSPAGPKVPWFPRKVSELDKCHHLVTKFDPDLDL	217
PAH (P00439-1)	LTNIIKILRHDIG-----ATVHELSDRDKKDTVPWFPRTIQELDRFANQILSYGAELDAD	145
TPH isoform 1 (P17752-1)	LNDIFHLLKSHNTNVLSPNFTLTKEDGMETVPWFPRKISLDLHCANRVL MYGSEL DAD	132
TPH isoform 2 (Q8IWU9-1)	FNELIQLLK FQTTIVTLNPPENIWTEEELEDVPWFPRKISELDKCSHRVLMYGSEL DAD	178
TH isoform 2 (P07101-3)	HPGFSDQVYRQRRLIAEIAFQYRHGDPIPRVEYTAEEIATWKEVYTTLKGLYATHACGE	250
TH isoform 4 (P07101-4)	HPGFSDQVYRQRRLIAEIAFQYRHGDPIPRVEYTAEEIATWKEVYTTLKGLYATHACGE	254
TH isoform 3 (P07101-1)	HPGFSDQVYRQRRLIAEIAFQYRHGDPIPRVEYTAEEIATWKEVYTTLKGLYATHACGE	281
TH isoform 1 (P07101-2)	HPGFSDQVYRQRRLIAEIAFQYRHGDPIPRVEYTAEEIATWKEVYTTLKGLYATHACGE	277
PAH (P00439-1)	HPGFKDPVYRARRKQFADIAYNYRHGQIPRVEYMEEEKKTWGT VFKTLKSLYKTHACYE	205
TPH isoform 1 (P17752-1)	HPGFKDNVYRKRKYFADLAMNYKHGQIPKVEFTEEEKTWGTVFQELNKLKYPHACRE	192
TPH isoform 2 (Q8IWU9-1)	HPGFKDNVYRQRKYFVDVAMGYKYGQIPRVEYTEEEKTWGVVFRELSKLYPTHACRE	238
TH isoform 2 (P07101-3)	HLEAFALLERFSGYREDNIPQLEDVSRFLKERTGFQLRPVAGLLSARDFLASLAFRVFQC	310
TH isoform 4 (P07101-4)	HLEAFALLERFSGYREDNIPQLEDVSRFLKERTGFQLRPVAGLLSARDFLASLAFRVFQC	314
TH isoform 3 (P07101-1)	HLEAFALLERFSGYREDNIPQLEDVSRFLKERTGFQLRPVAGLLSARDFLASLAFRVFQC	341
TH isoform 1 (P07101-2)	HLEAFALLERFSGYREDNIPQLEDVSRFLKERTGFQLRPVAGLLSARDFLASLAFRVFQC	337
PAH (P00439-1)	YNHIFP LLEKYCGFHEDNIPQLEDVSQFLQTCTGFRLRPVAGLLSSRDFLGGLAFRVFHC	265
TPH isoform 1 (P17752-1)	YLKNLPLLSKYCYGREDNIPQLEDVSNFLKERTGFSIRPVAGYLSPRDFLSGLAFRVFHC	252
TPH isoform 2 (Q8IWU9-1)	YLKNFPLLT KYCYGREDNVPQLEDVSMFLKERSGFTVRPVAGYLSPRDFLAGLAYRVFHC	298
TH isoform 2 (P07101-3)	TQYIRHASSPMHSPEPDCCHELLGHVPMLADRTFAQFSQDIGLASLGASDEEIEKLSTLY	370
TH isoform 4 (P07101-4)	TQYIRHASSPMHSPEPDCCHELLGHVPMLADRTFAQFSQDIGLASLGASDEEIEKLSTLY	374
TH isoform 3 (P07101-1)	TQYIRHASSPMHSPEPDCCHELLGHVPMLADRTFAQFSQDIGLASLGASDEEIEKLSTLY	401
TH isoform 1 (P07101-2)	TQYIRHASSPMHSPEPDCCHELLGHVPMLADRTFAQFSQDIGLASLGASDEEIEKLSTLY	397
PAH (P00439-1)	TQYIRHGSKPMYTPEPDICHELLGHVPLFSDRSFAQFSQDIGLASLGAPDEYIEKLAT IY	325
TPH isoform 1 (P17752-1)	TQYVRHSSDPFYTPEPDTCHELLGHVPLLAEPSFAQFSQDIGLASLGASEEAVQKLATCY	312
TPH isoform 2 (Q8IWU9-1)	TQYIRHGSDPLYTPEPDTCHELLGHVPLLADPKFAQFSQDIGLASLGASDEDVQKLATCY	358

TH isoform 2 (P07101-3)	WFTVEFGLCKQNGEVKAYGAGLLSSYGELLHCLSEEPEIRAFDPEAAAVQPYQDQTYQSV	430
TH isoform 4 (P07101-4)	WFTVEFGLCKQNGEVKAYGAGLLSSYGELLHCLSEEPEIRAFDPEAAAVQPYQDQTYQSV	434
TH isoform 3 (P07101-1)	WFTVEFGLCKQNGEVKAYGAGLLSSYGELLHCLSEEPEIRAFDPEAAAVQPYQDQTYQSV	461
TH isoform 1 (P07101-2)	WFTVEFGLCKQNGEVKAYGAGLLSSYGELLHCLSEEPEIRAFDPEAAAVQPYQDQTYQSV	457
PAH (P00439-1)	WFTVEFGLCKQGDSEIKAYGAGLLSSFGELQYCLSEKPKLLPLELEKTAIQNYTVTEFQPL	385
TPH isoform 1 (P17752-1)	FFTVEFGLCKQDQGQLRVFGAGLLSSI SELKHALSGHAKVKPFDPK I TCKQECL I TTFQDV	372
TPH isoform 2 (Q8IWU9-1)	FFTIEFGLCKQEGQLRAYGAGLLSSI GELKHALSDKACVKA FDKPTTCLQECL I TTFQEA	418
	:*:***** ..:::***** .** :.*. . : : : . *	
TH isoform 2 (P07101-3)	YFVSESFSDAKDKLRSYASRIQRPFSVKFDPYTLAIDVLDSPQAVRRSLEGVQDELDTLA	490
TH isoform 4 (P07101-4)	YFVSESFSDAKDKLRSYASRIQRPFSVKFDPYTLAIDVLDSPQAVRRSLEGVQDELDTLA	494
TH isoform 3 (P07101-1)	YFVSESFSDAKDKLRSYASRIQRPFSVKFDPYTLAIDVLDSPQAVRRSLEGVQDELDTLA	521
TH isoform 1 (P07101-2)	YFVSESFSDAKDKLRSYASRIQRPFSVKFDPYTLAIDVLDSPQAVRRSLEGVQDELDTLA	517
PAH (P00439-1)	YYVAESFNDAKEKVRNFAATIPRPFSVRYDPYTQRIEVL DNTQQLK I LADSINSEIGILC	445
TPH isoform 1 (P17752-1)	YFVSESFEDAKEKMREFTKT I KRPFGVKYNPYTRS I Q I LKDTKS I T SAMNELQHDLDVVS	432
TPH isoform 2 (Q8IWU9-1)	YFVSESFEEAKEKMRDFAKSI TRPFSVYFNPYTQSI E I LKDTRS I ENVVQDLRSDLNTVC	478
	*:****.**:*:*.:: * ***. * :*** *:*. . : : : . : . : .	
TH isoform 2 (P07101-3)	HALSAIG - - - -	497
TH isoform 4 (P07101-4)	HALSAIG - - - -	501
TH isoform 3 (P07101-1)	HALSAIG - - - -	528
TH isoform 1 (P07101-2)	HALSAIG - - - -	524
PAH (P00439-1)	SALQKIK - - - -	452
TPH isoform 1 (P17752-1)	DALAKVSRKPSI	444
TPH isoform 2 (Q8IWU9-1)	DALNKMNQYLG I	490
	* * :	

Figure 20 alignment of the human AAHs performed with Clustal Omega ¹⁵¹. The Uniprot sequence entry is noted in parenthesis.

A2 ALIGNMENT OF THE CATALYTIC DOMAINS OF *HUMAN* TPH1, *HUMAN* TPH2 AND *CHICKEN* TPH1

<i>human</i> TPH2	- EELEDVPWFPRK I SELDKCSHRVLMYGSELDADHPGFKDNVYRQRRKYFVDVAMGYKYG	204
<i>human</i> TPH1	- DGMETVPWFPPK I SLDLHCANRVLMYGSELDADHPGFKDNVYRKRKYFADLAMNYKHG	158
<i>chicken</i> TPH1	DGDMENIPWYPKK I SLDLHCANRVLMYGSDLDADHPGFKDNVYRKRKYFADLAMNYKHG	159
	:* :*	
<i>human</i> TPH2	QP I PRVEYTEEETKTWGVVFRELSKLYPTHACREYLKNFPLLT KYCGYREDNVPQLEDVS	264
<i>human</i> TPH1	DP I PKVEFTEEE IKTWGTVFQELNKLYP THACREYLKNLPLLSKYCGYREDNIPQLEDVS	218
<i>chicken</i> TPH1	DP I PE I EFTEEE IKTWGTVYRELNKLYP THACREYLKNLPLLT KYCGYREDNIPQLEDVS	219
	:***. :*:***** ***. :*:*. ***** :*** :***** :*****	
<i>human</i> TPH2	MFLKERSGFTVRPVAGYLSPRDFLAGLAYRVFHCTQYIRHGSDPLYTPEPDTCH ELLGHV	324
<i>human</i> TPH1	NFLKERTGFSIRPVAGYLSPRDFLSGLAFRVFHCTQYVRHSSDPFYTPEPDTCH ELLGHV	278
<i>chicken</i> TPH1	RFLKERTGFTIRPVAGYLSPRDFLAGLAFRVFHCTQYVRHSSDPLYTPEPDTCH ELLGHV	279
	*****:*. :*****:*. :*****:*. :*****:*. :*****:*. :*****:*	
<i>human</i> TPH2	PLLADPKFAQFSQE IGLASLGASEADVQKLATCYFFFTIEFGLCKQEGQLRVFGAGLLSSI	384
<i>human</i> TPH1	PLLAEPSFAQFSQE IGLASLGASEADVQKLATCYFFFTVEFGLCKQDQQLRVFGAGLLSSI	338
<i>chicken</i> TPH1	PLLAEPSFAQFSQE IGLASLGASEADVQKLATCYFFFTVEFGLCKQEGQLRVFGAGLLSSI	339
	*****:*. *****:*. *****:*. *****:*. *****:*. *****:	
<i>human</i> TPH2	GELKHALSDKACVKA FDKPTTCLQECL I TTFQEAYFVSES FEEAKEKMRDFAKSI TRPFS	444
<i>human</i> TPH1	SELKHALSGHAKVKPFDPK I TCKQECL I TTFQDVYFVSES FEDAKEKMREFTKT I KRPFG	398
<i>chicken</i> TPH1	SELKHSLSGSAKVKPFDPKVTCKQECL I TTFQEVYFVSES FEEAKEKMREFAKT I KRPFG	399
	. *****:*. * ***. * *****:*****:*****:*****:*. :*. :*. :*. :*	
<i>human</i> TPH2	VYFNPTYQS I E I LKD	459
<i>human</i> TPH1	VKYNPYTRS I Q I LKD	413
<i>chicken</i> TPH1	VKYNPYTQSVQ I LKD	414
	* :*****:*:*****	

Figure 21 Sequential alignment of the catalytic domains of *human* TPH1 (residues 100-413), *human* TPH2 (residues 146-459), and *chicken* TPH1 (residues 100-414) determined by use of Clustal Omega ¹⁵¹. The CD of *chicken* TPH1 was determined to be 91 % identical to the CD of *human* TPH1 and 83 % identical to the CD of *human* TPH2. Residues involved in binding of L-Trp and BH₂ identified from structures of *chicken* TPH1 ⁴⁹ and *human* TPH1 ⁵⁴ are highlighted in orange.

B PURIFICATION OF TPH VARIANTS

B1 PURIFICATION OF TPH2

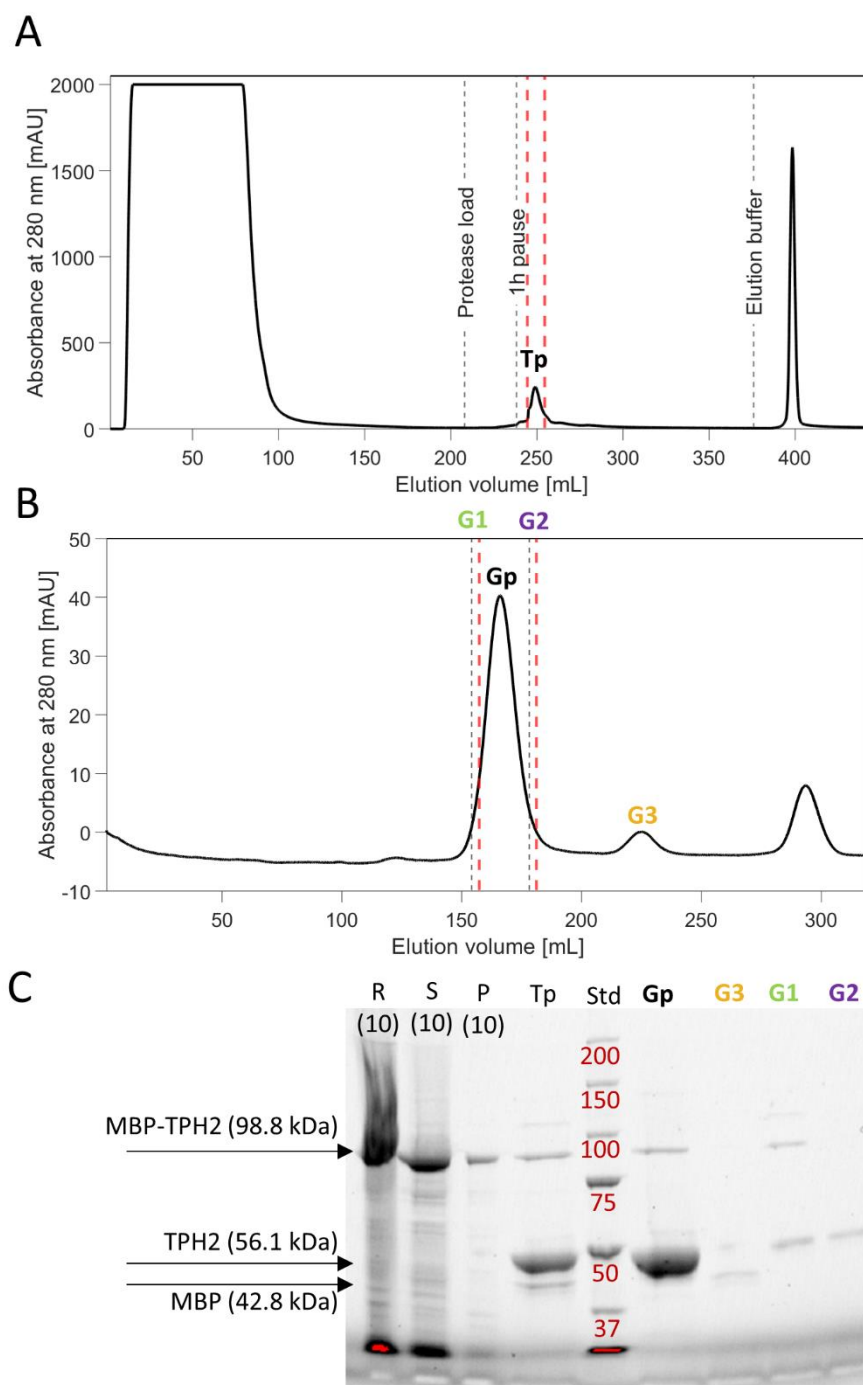


Figure 22 Purification of TPH2 in 20 mM HEPES, 300 mM $(\text{NH}_4)_2\text{SO}_4$, pH 7.0. A) Affinity chromatogram, B) GF chromatogram, and C) SDS-PAGE. Abbreviations: From sample preparation [R – lysate, S – Supernatant from centrifugation, P – pellet from centrifugation]. From chromatographic steps [Tp – target peak from AC step, Gp – target peak from GF step] Std – Precision Plus Protein Standard. The sizes of the standard proteins are provided in kDa (in red). The number in parenthesis denotes the dilution factor of the sample if relevant. Red coloring on the image marked by the Image Lab software indicates a large amount of protein. Red dashed lines in the chromatograms mark the collected TPH2 fractions. G1, G2, and G3 marks three sampled fractions.

B2 PURIFICATION OF NΔ47-TPH2

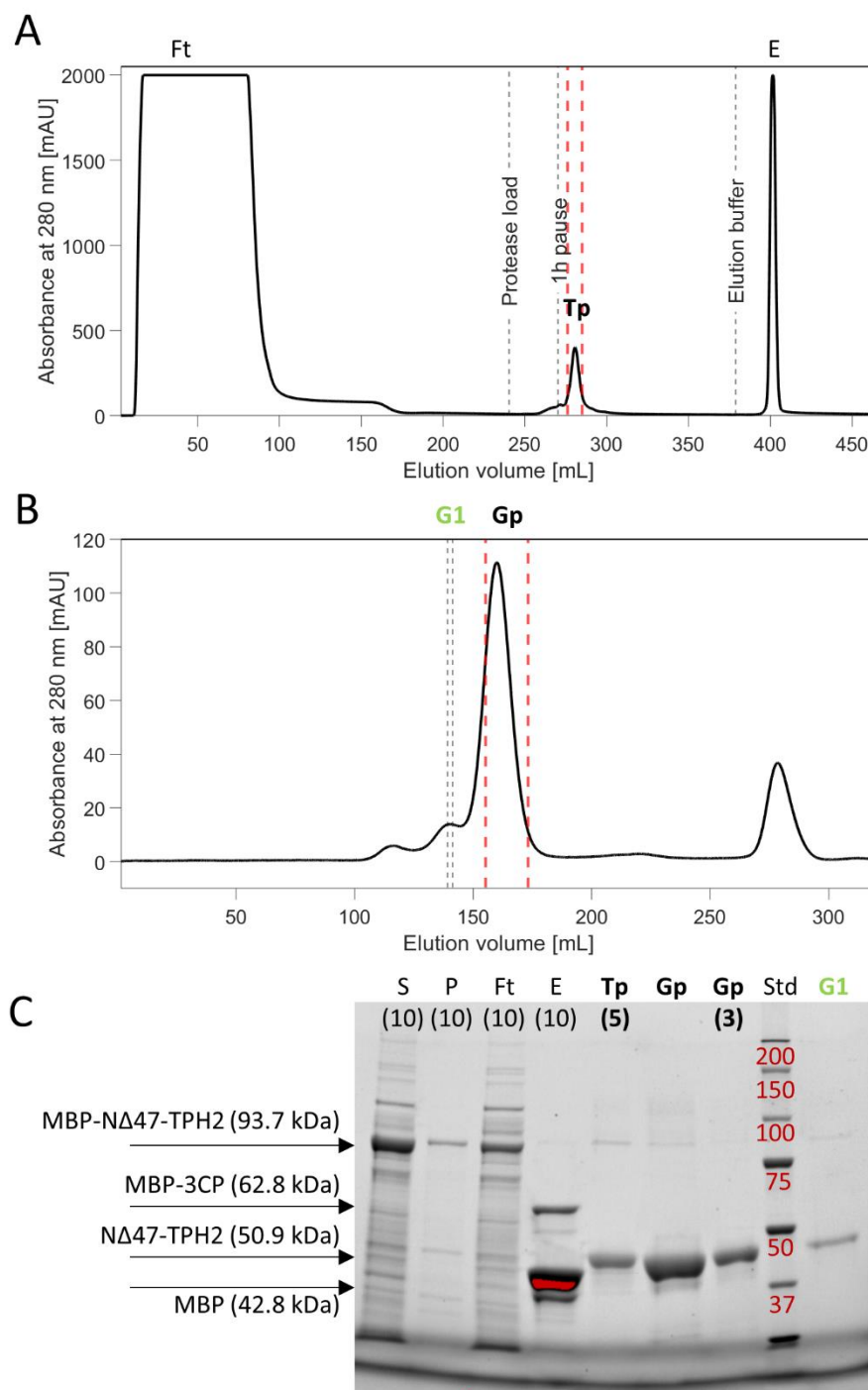


Figure 23 Purification of NΔ47-TPH2 in 20 mM HEPES, 300 mM (NH₄)₂SO₄, pH 7.0. A) Affinity chromatogram, B) GF chromatogram, and C) SDS-PAGE analysis. Abbreviations: From sample preparation [S – Supernatant from centrifugation, P – pellet from centrifugation]. From chromatographic steps [Ft – Flow through, Tp – target peak from AC step, E – Elution peak, Gp – target peak from GF step] Std – Precision Plus Protein Standard. The sizes of the standard proteins are provided in kDa (in red). The number in parenthesis denotes the dilution factor of the samples if relevant. Red coloring on the image marked by the Image Lab software indicates a large amount of protein. Red dashed lines in the chromatograms mark the collected NΔ47-TPH2 fractions. G1 marks a sampled fraction.

B3 PURIFICATION OF NΔ47-TPH2-RC

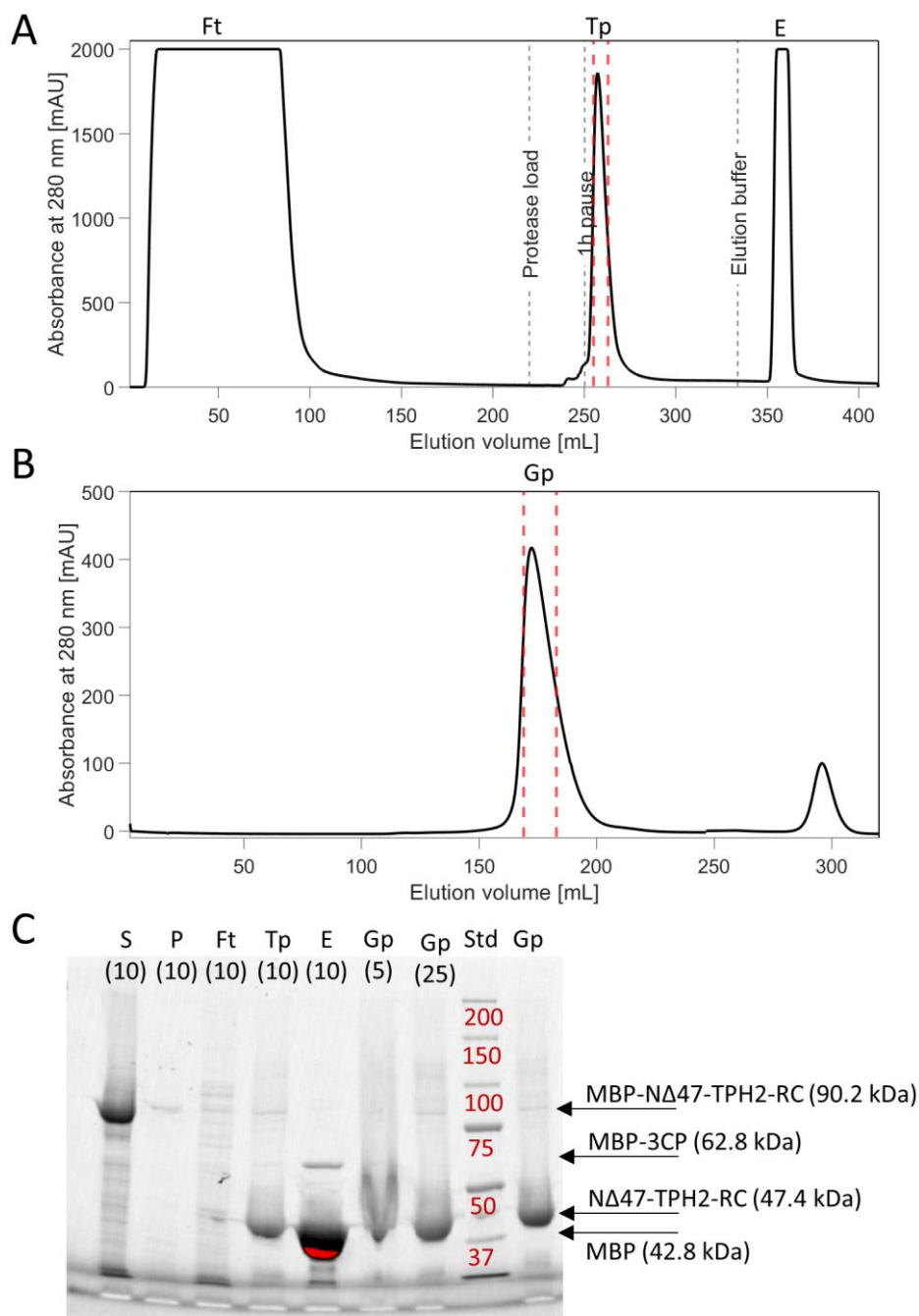


Figure 24 Purification of NΔ47-TPH2-RC in 20 mM HEPES, 300 mM (NH₄)₂SO₄, pH 7.0. A) Affinity chromatogram, B) GF chromatogram, and C) SDS-PAGE analysis. Abbreviations: From sample preparation [S – Supernatant from centrifugation, P – pellet from centrifugation]. From chromatographic steps [Ft – Flow through, Tp – target peak from AC step, E – Elution peak, Gp – target peak from GF step] Std – Precision Plus Protein Standard. The sizes of the standard proteins are provided in kDa (in red). The number in parenthesis denotes the dilution factor of the samples if relevant. Red coloring on the image marked by the Image Lab software indicates a large amount of protein. Red dashed lines in the chromatograms mark the collected NΔ47-TPH2-RC fractions.

B4 PURIFICATION OF TPH2-CT

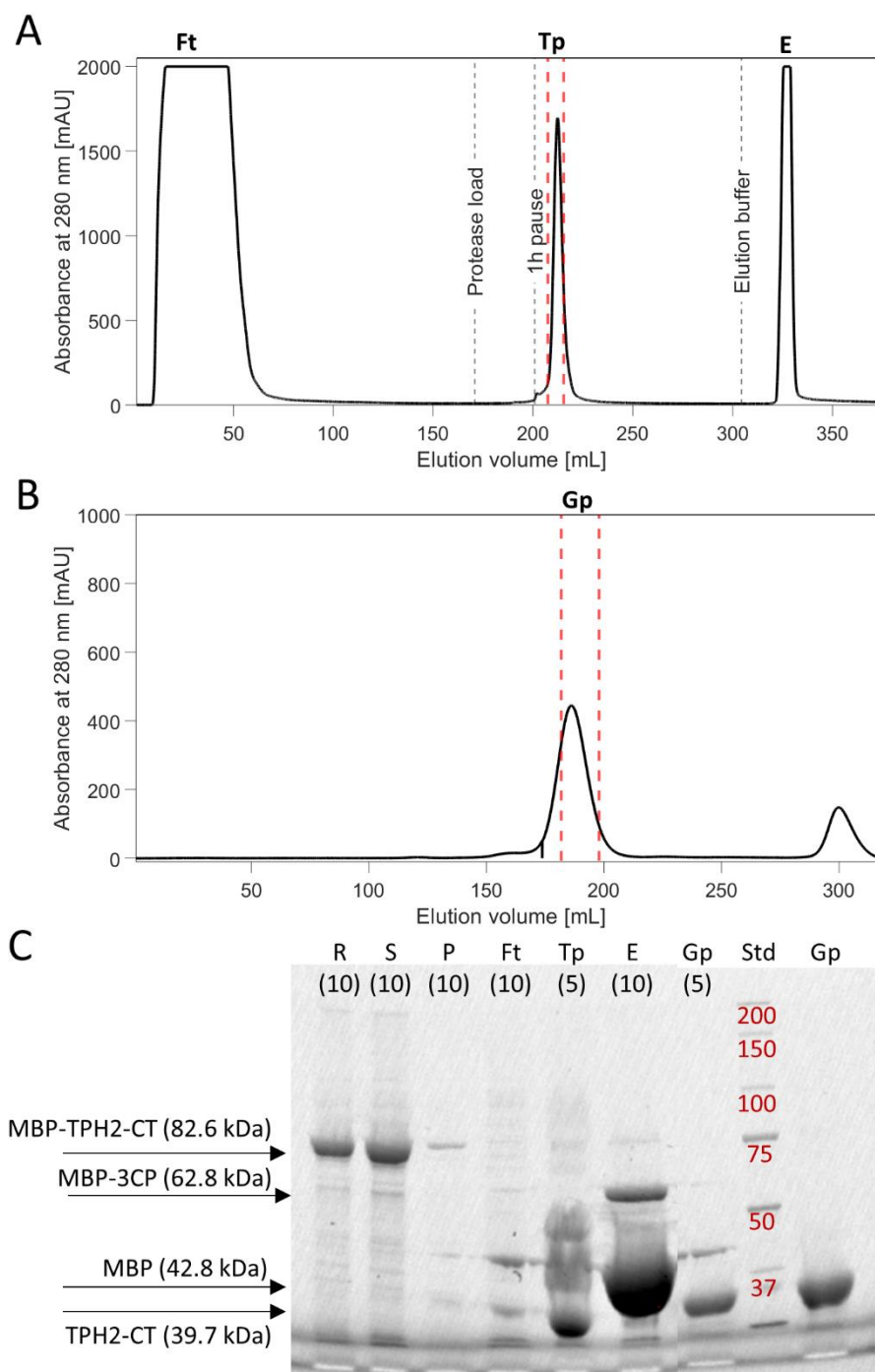


Figure 25 Purification of TPH2-CT in 20 mM HEPES, 100 mM $(\text{NH}_4)_2\text{SO}_4$, pH 7.0. A) Affinity chromatogram, B) GF chromatogram, and C) SDS-PAGE analysis. Abbreviations: From sample preparation [R – lysate, S – Supernatant from centrifugation, P – pellet from centrifugation]. From chromatographic steps [Ft – Flow through, Tp – target peak from AC step, E – Elution peak, Gp – target peak from GF step], Std – Precision Plus Protein Standard. The sizes of the standard proteins are provided in kDa (in red). The number in parenthesis denotes the dilution factor of the samples if relevant. Red dashed lines in the chromatograms mark the collected TPH2-CT fractions.

B5 PURIFICATION OF TPH2-R

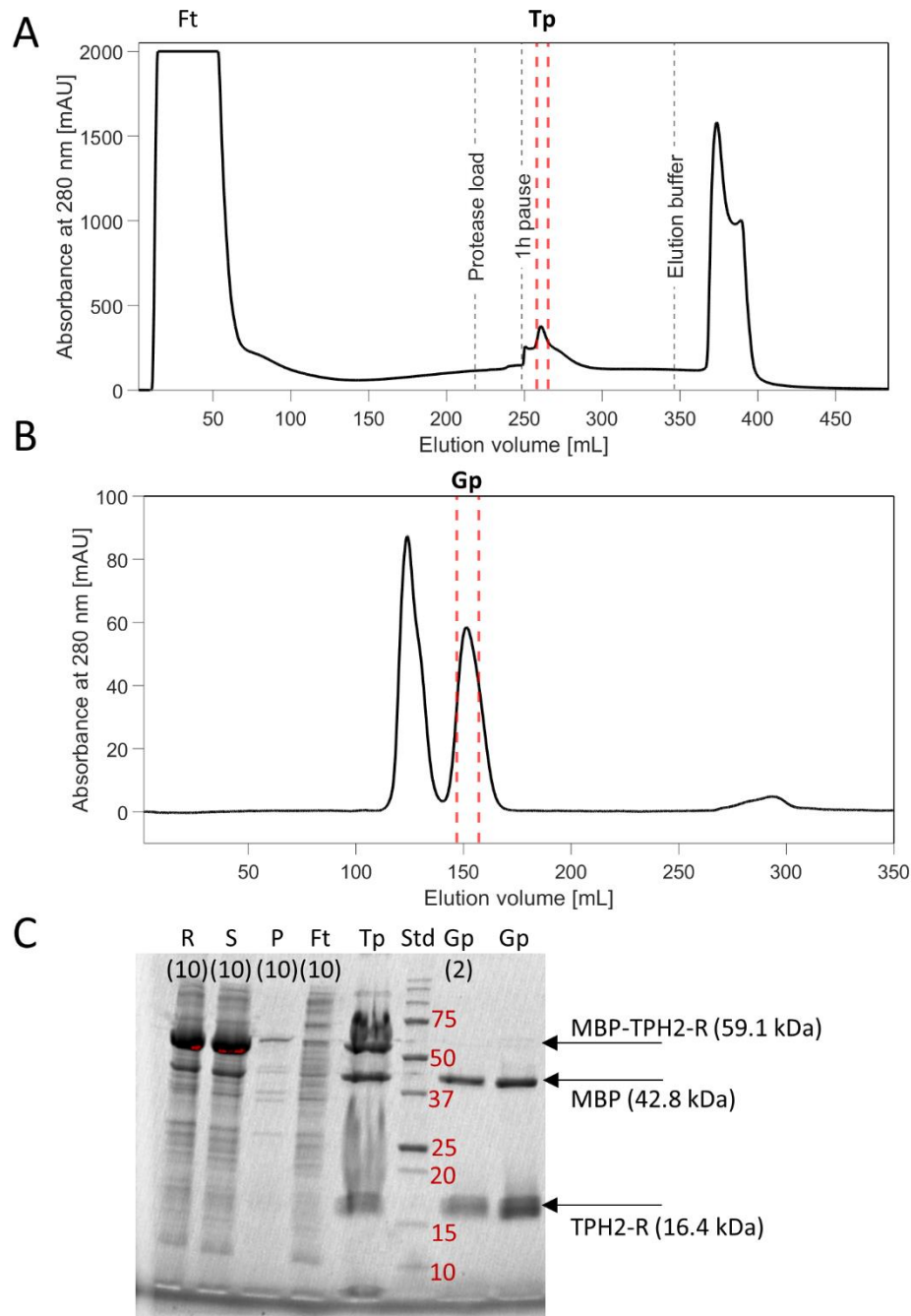


Figure 26 Purification of TPH2-R in 20 mM HEPES, 300 mM $(\text{NH}_4)_2\text{SO}_4$, 3 mM L-Phe, pH 7.0. A) Affinity chromatogram, B) GF chromatogram, and C) SDS-PAGE analysis. Abbreviations: From sample preparation [R – lysate, S – Supernatant from centrifugation, P – pellet from centrifugation]. From chromatographic steps [Ft – Flow through, Tp – target peak from AC step, Gp – peak from GF step], Std – Precision Plus Protein Standard. The sizes of the standard proteins are provided in kDa. The number in parenthesis denotes the dilution factor of the samples if relevant. Red coloring on the image marked by the Image Lab software indicates a large amount of protein. Red dashed lines in the chromatograms mark the collected TPH2-R fractions.

B6 PURIFICATION OF NΔ47-TPH2-R

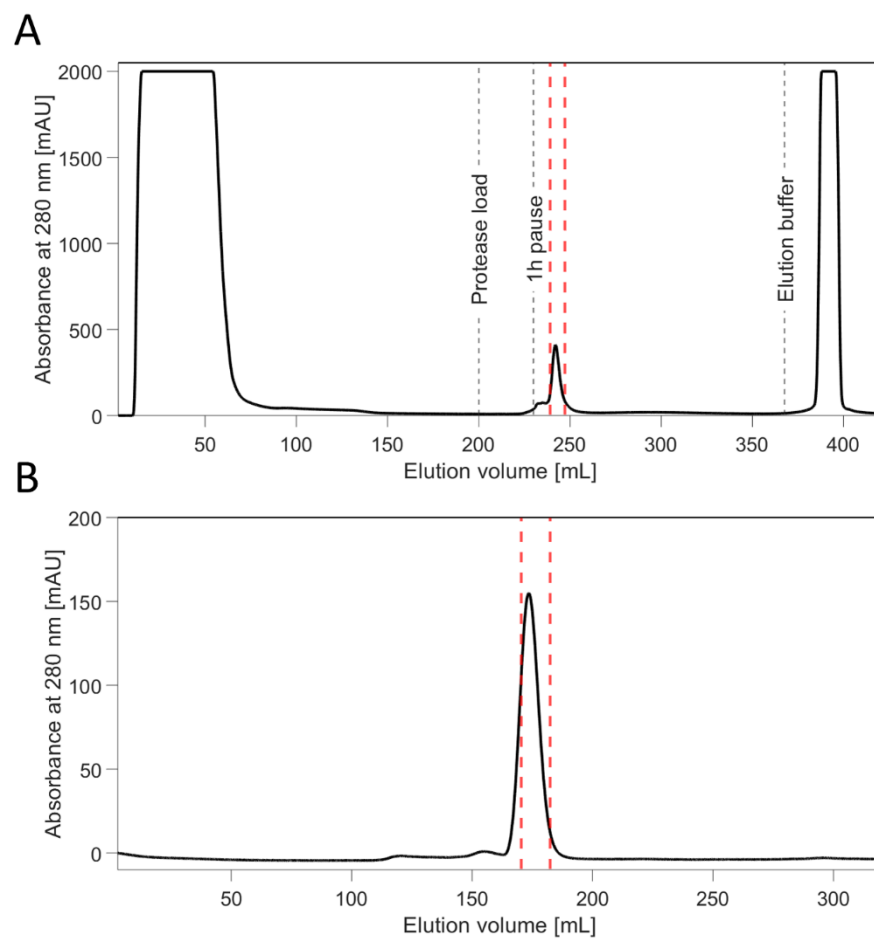


Figure 27 purification of NΔ47-TPH2-R in 12 mM Na₂HPO₄/8 mM NaH₂PO₄, 100 mM (NH₄)₂SO₄, pH 7.0. A) Affinity chromatogram and B) GF chromatogram. Red dashed lines mark the collected NΔ47-TPH2-R fractions.

B7 PURIFICATION OF TPH1-C

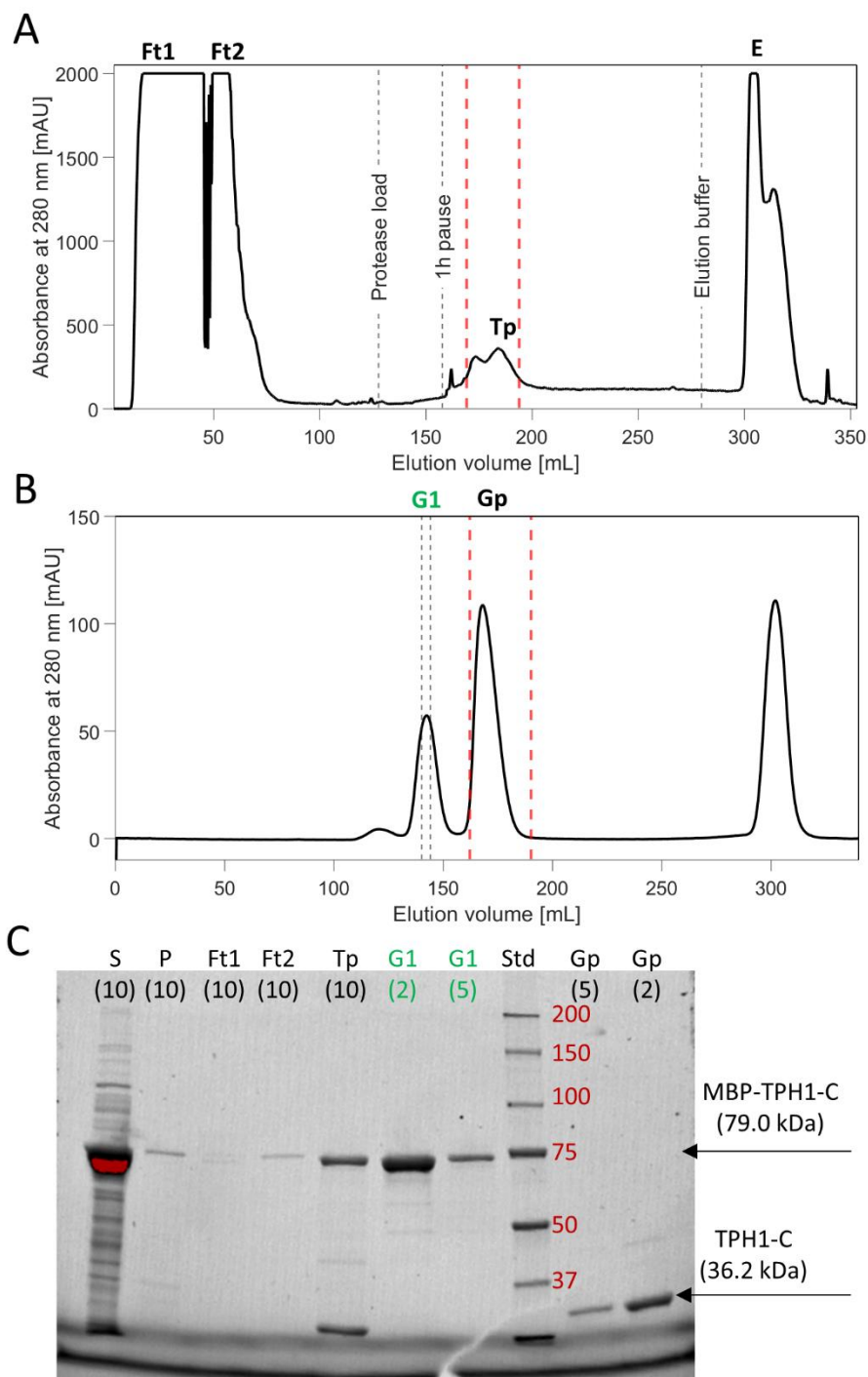


Figure 28 Purification of TPH1-C in 20 mM Tris, 100 mM $(\text{NH}_4)_2\text{SO}_4$, pH 8.0. A) Affinity chromatogram, B) GF chromatogram, and C) SDS-PAGE analysis. Abbreviations: From sample preparation [S – Supernatant from centrifugation, P – pellet from centrifugation]. From chromatographic steps [Ft – Flow through, Tp – target peak from AC step, Gp – target peak from GF step], Std – Precision Plus Protein Standard. The sizes of the standard proteins are provided in kDa (in red). The number in parenthesis denotes the dilution factor of the samples if relevant. Red coloring on the image marked by the Image Lab software indicates a large amount of protein. The affinity chromatogram in A) is a bit noisy, as air was introduced on the column during sample load. Red dashed lines in the chromatograms mark the collected TPH1-C fractions. G1 marks a sampled fraction. The purification was performed by David F. Nielsen.

B8 PURIFICATION OF TPH1-RC

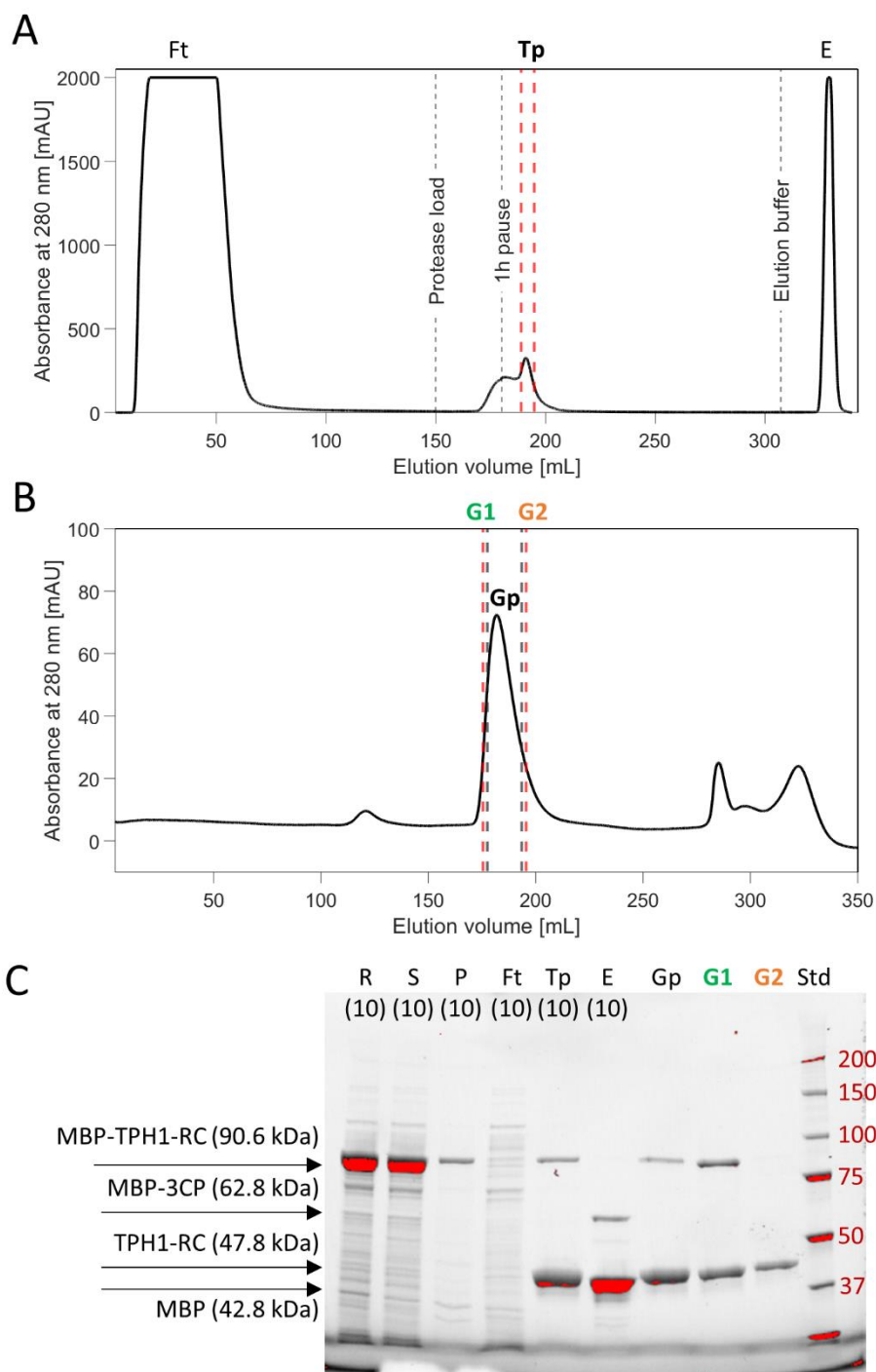


Figure 29 Purification of TPH1-RC in 20 mM Tris, 300 mM $(\text{NH}_4)_2\text{SO}_4$, 5 mM DTT, pH 8.0. A) Affinity chromatogram, B) GF chromatogram, and C) SDS-PAGE analysis. Abbreviations: From sample preparation [R – lysate, S – Supernatant from centrifugation, P – pellet from centrifugation]. From chromatographic steps [Ft – Flow through, Tp – target peak from AC step, E – elution peak, Gp – target peak from GF step]. Std – Precision Plus Protein Standard. The sizes of the standard proteins are provided in kDa (in red). The number in parenthesis denotes the dilution factor of the sample if relevant. Red coloring on the image marked by the Image Lab software indicates a large amount of protein. Red dashed lines in the chromatograms mark the collected TPH1-RC fractions. G1 and G2 mark two sampled fractions. Excluding the G1 fraction, thus a more asymmetric fractioning, would lead to a higher purity.

C SEQUENCES OF PROTEIN VARIANTS

All TPH sequences include the additional N-terminal “GP” residues from the affinity tag (“GPGS” for TPH1-RC).

TPH2

GPQPAMMMFSSKYWARRGFSLDSAVPEEHQLLSSTLNKPNSGKNDDKGNKGSSSKREAA
TESGKTAVVFSKNEVGGLVKALRLFQEKRVMNVHIESRKSRRRSSEVEIFVDCECGKTEF
NELIQLLKFFQTTIVTLNPPENIWTEEELEDPWFPRKISELDKCSHRVLMYGSELDADHPGF
KDNVYRQRRKYFVDVAMGYKYGQPIPRVEYTEEETKTWGVVFRELSKLYPTHACREYLKN
FPLLTKYCGYREDNVPQLEDVSMFLKERSGFTVRPVAGYLSPRDFLAGLAYRVFHCTQYIR
HGSDPLYTPEPDTCHELLGHVPLLADPKFAQFSQEIGLASLGASDEDVQKLATCYFFTIEFG
LCKQEGQLRAYGAGLLSSIGELKHALSDKACVKAFDPKTTCLQECLITTFQEAYFVSESFEE
AKEKMRDFAKSITRPFSVYFNPYTQSIEILKDTRSIENVVQDLRSDLNTVCDALNKMNQYLGI

NΔ47-TPH2

GPGNKGSSSKREAATESGKTAVVFSKNEVGGLVKALRLFQEKRVMNVHIESRKSRRRSSE
VEIFVDCECGKTEFNELIQLLKFFQTTIVTLNPPENIWTEEELEDPWFPRKISELDKCSHRVL
MYGSELDADHPGFKDNVYRQRRKYFVDVAMGYKYGQPIPRVEYTEEETKTWGVVFRELS
KLYPTHACREYLKNFPLLTKYCGYREDNVPQLEDVSMFLKERSGFTVRPVAGYLSPRDFLA
GLAYRVFHCTQYIRHGSDPLYTPEPDTCHELLGHVPLLADPKFAQFSQEIGLASLGASDEDV
QKLATCYFFTIEFGLCKQEGQLRAYGAGLLSSIGELKHALSDKACVKAFDPKTTCLQECLITT
FQEAYFVSESFEEAKEKMRDFAKSITRPFSVYFNPYTQSIEILKDTRSIENVVQDLRSDLNTV
CDALNKMNQYLGI

NΔ47-TPH2-RC

GPGNKGSSSKREAATESGKTAVVFSKNEVGGLVKALRLFQEKRVMNVHIESRKSRRRSSE
VEIFVDCECGKTEFNELIQLLKFFQTTIVTLNPPENIWTEEELEDPWFPRKISELDKCSHRVL
MYGSELDADHPGFKDNVYRQRRKYFVDVAMGYKYGQPIPRVEYTEEETKTWGVVFRELS
KLYPTHACREYLKNFPLLTKYCGYREDNVPQLEDVSMFLKERSGFTVRPVAGYLSPRDFLA
GLAYRVFHCTQYIRHGSDPLYTPEPDTCHELLGHVPLLADPKFAQFSQEIGLASLGASDEDV
QKLATCYFFTIEFGLCKQEGQLRAYGAGLLSSIGELKHALSDKACVKAFDPKTTCLQECLITT
FQEAYFVSESFEEAKEKMRDFAKSITRPFSVYFNPYTQSIEILKD

NΔ47-TPH2-R

GPGNKGSSSKREAATESGKTAVVFSKNEVGGLVKALRLFQEKRVMNVHIESRKSRRRSSE
VEIFVDCECGKTEFNELIQLLKFFQTTIVTLNPPENIWTEE

TPH2-R

GPQPAMMMFSSKYWARRGFSLDSAVPEEHQLLGSSTLNKPNSGKNDDKGNKGSSKREAA
TESGKTAVVFSLKNEVGGLVKALRLFQEKRVMVHIESRKSRRRSSEVEIFVDCECGKTEF
NELIQLLKFTTIVTLNPPENIWTEE

TPH2-CT

GPEELEDVPWFPRKISELDKCSHRVLMYGSELDADHPGFKDNVYRQRRKYFVDVAMGYKY
GQPIPRVEYTEEETKTWGVVFRELSKLYPTHACREYLKNFPLLTKYCGYREDNVPQLEDVS
MFLKERSGFTVRPVAGYLSPRDFLAGLAYRVFHCTQYIRHGSDPLYTPEPDTCHELLGHVP
LLADPKFAQFSQEIGLASLGASEDEVQKLATCYFFTIEFGLCKQEGQLRAYGAGLLSSIGELK
HALSDKACVKAFDPKTTCLQECLITTFQEAYFVSESFEEAKEKMRDFAKSITRPFVYFNPY
TQSIEILKDTRSIENVVQDLRSDLNTVCDALNKMNQYLG

TPH1-C

GPDMETVPWFPPKISDLHDHCANRVLMYGSELDADHPGFKDNVYRKRRKYFADLAMNYK
HGDPIPKVEFTEEEIKTWGTVFQELNKLYPTHACREYLKNLPLLSKYCGYREDNIPQLEDVS
NFLKERTGFSIRPVAGYLSPRDFLSGLAFRVFHCTQYVRHSSDPFYTPEPDTCHELLGHVPL
LAEPSFAQFSQEIGLASLGASEEAVQKLATCYFFTVEFGLCKQDGQLRVFGAGLLSSISELK
HALSGHAKVKPFDPKITCKQECLITTFQDVYFVSESFEDAKEKMREFTKTIKRPFGVKYNPY
TRSIQILKD

TPH1-RC

GPGSIEDNKENKDHSLERGRASLIFSLKNEVGGLIKALKIFQEKHVNLLHIESRKSRRNSEF
EIFVDCDINREQLNDIFHLLKSHTNVLSVNLPDNFTLKEDGMETVPWFPPKISDLHDHCANRVL
MYGSELDADHPGFKDNVYRKRRKYFADLAMNYKHGDPIPKVEFTEEEIKTWGTVFQELNKL
YPTHACREYLKNLPLLSKYCGYREDNIPQLEDVSNFLKERTGFSIRPVAGYLSPRDFLSGLA
FRVFHCTQYVRHSSDPFYTPEPDTCHELLGHVPLLAEPSFAQFSQEIGLASLGASEEAVQK
LATCYFFTVEFGLCKQDGQLRVFGAGLLSSISELKHALSGHAKVKPFDPKITCKQECLITTFQ
DVYFVSESFEDAKEKMREFTKTIKRPFGVKYNPYTRSIQILKD

TPH1-R

GPIEDNKENKDHSLERGRASLIFSLKNEVGGLIKALKIFQEKHVNLLHIESRKSRRNSEFEIF
VDCDINREQLNDIFHLLKSHTNVLSVNLPDNFTLKE

MBP-3CP

MKIEEGKLVWINGDKGYNGLAIEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIF
WAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLP
NPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGV
NAGAKAGLTFLVDLIKHKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVT
VLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSY
EEELAKDPRIAATMENAQKGEIMPNIQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNS
SSNNNNNNNNNNLGGIPGPEHEFLNALIRRNCHIITTDKGEFNLLGIYSNCAVVPHTAEPGD
VVDIDGRLVRVLKQQVLTDMNDVDTEVTVLWLDQNEKFRDIRRFIPEHQQDWHNIHLATNV
TKFPMNLNVEVGHTVPYGEINLSGNATCRLYKYDYPTQPGQCGAVLANTGNIIGIHVGGNGR
VGYAAALLRKYFAEEQ

D ^1H - ^{15}N HSQC SPECTRA OF NΔ47-TPH2-R

This appendix contains ^1H - ^{15}N HSQC spectra of NΔ47-TPH2-R in presence of different possible ligands: L-4-fluorophenylalanine, L-5-fluorotryptophan, L-5-hydroxytryptophan (5-HTP), and L-Tyr. All spectra were recorded at 20 °C, processed with topspin 3.6.2 and referenced to DSS. The protein concentration varies from sample to sample, hampering comparison of the intensities.

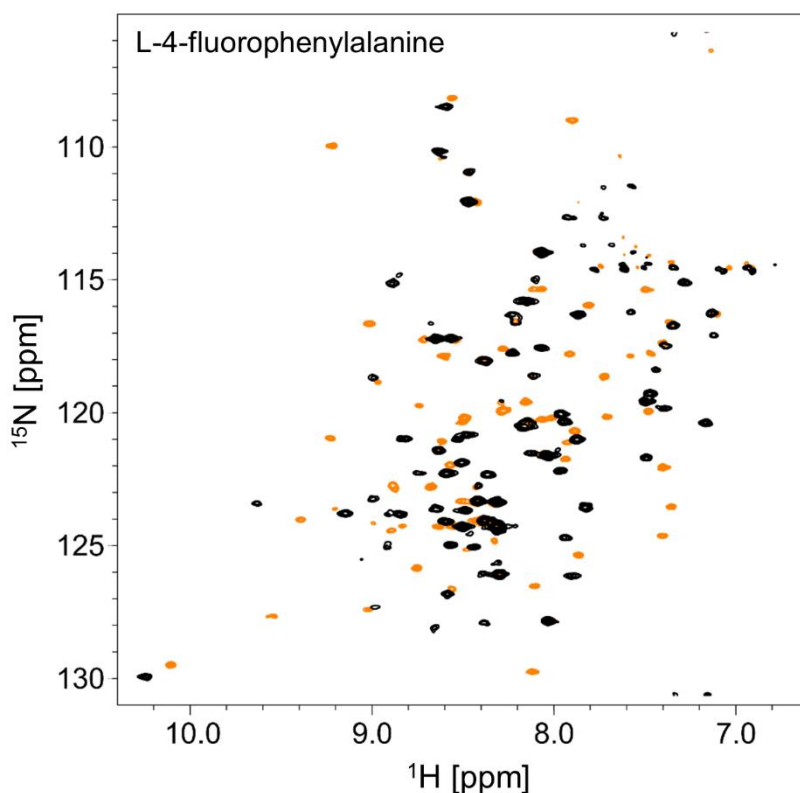


Figure 30 ^1H - ^{15}N HSQC spectra of NΔ47-TPH2-R in 12 mM Na_2HPO_4 /8 mM NaH_2PO_4 , 100 mM $(\text{NH}_4)_2\text{SO}_4$, pH 7.0, 125 μM DSS, 5% D_2O . Orange: Without ligand. Black: 5.6 mM L-4-fluorophenylalanine

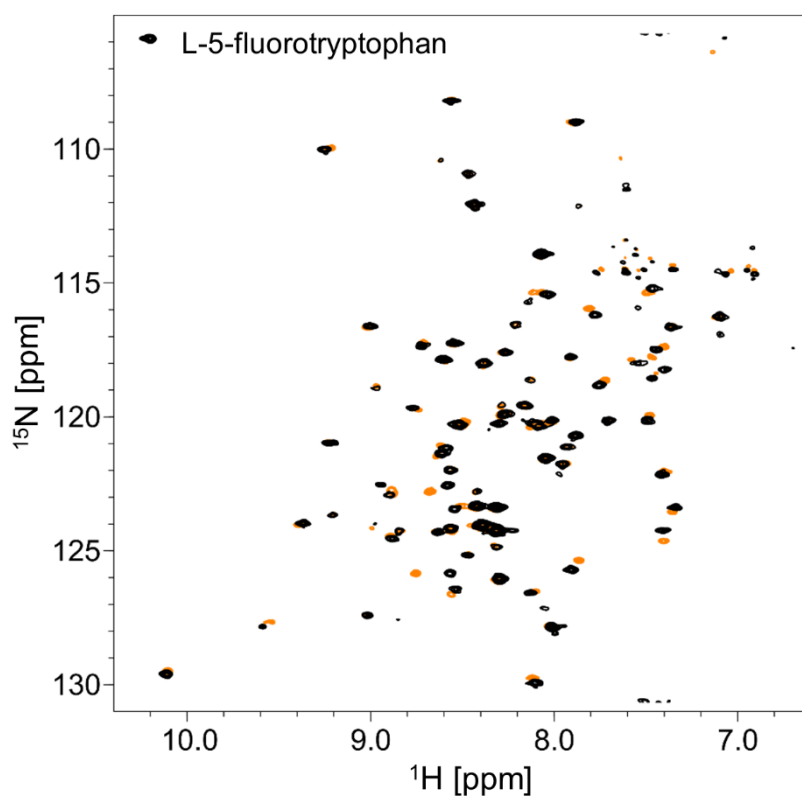


Figure 31 ^1H - ^{15}N HSQC spectra of NΔ47-TPH2-R in 12 mM Na_2HPO_4 /8 mM NaH_2PO_4 , 100 mM $(\text{NH}_4)_2\text{SO}_4$, pH 7.0, 125 μM DSS, 5% D_2O . Orange: Without ligand. Black: 10 mM L-5-fluorotryptophan.

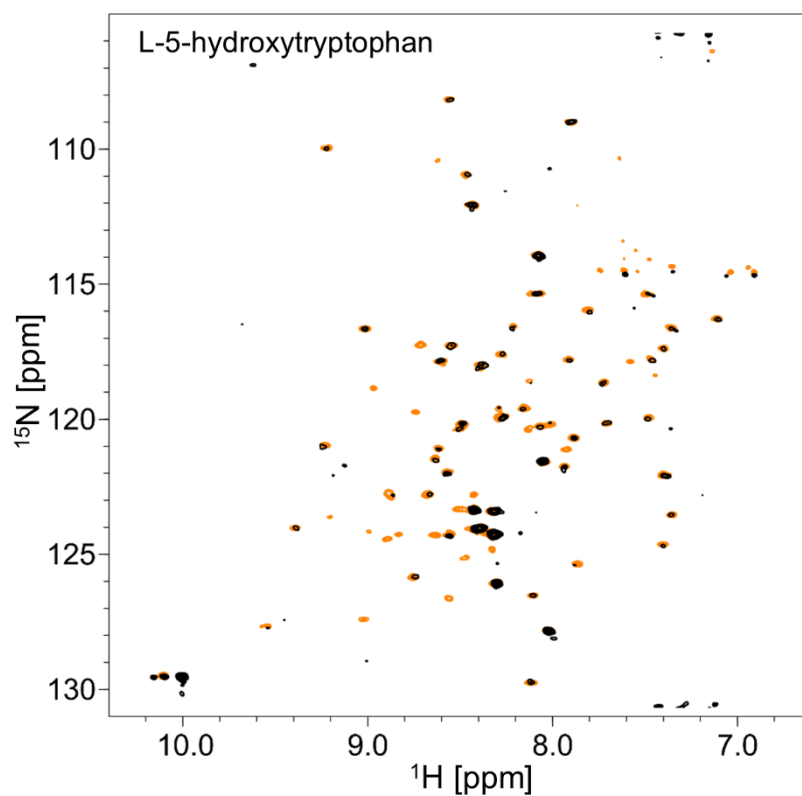


Figure 32 ^1H - ^{15}N HSQC spectra of NΔ47-TPH2-R in 12 mM Na_2HPO_4 /8 mM NaH_2PO_4 , 100 mM $(\text{NH}_4)_2\text{SO}_4$, pH 7.0, 125 μM DSS, 5% D_2O . Orange: Without ligand. Black: 12 mM L-5-hydroxytryptophan.

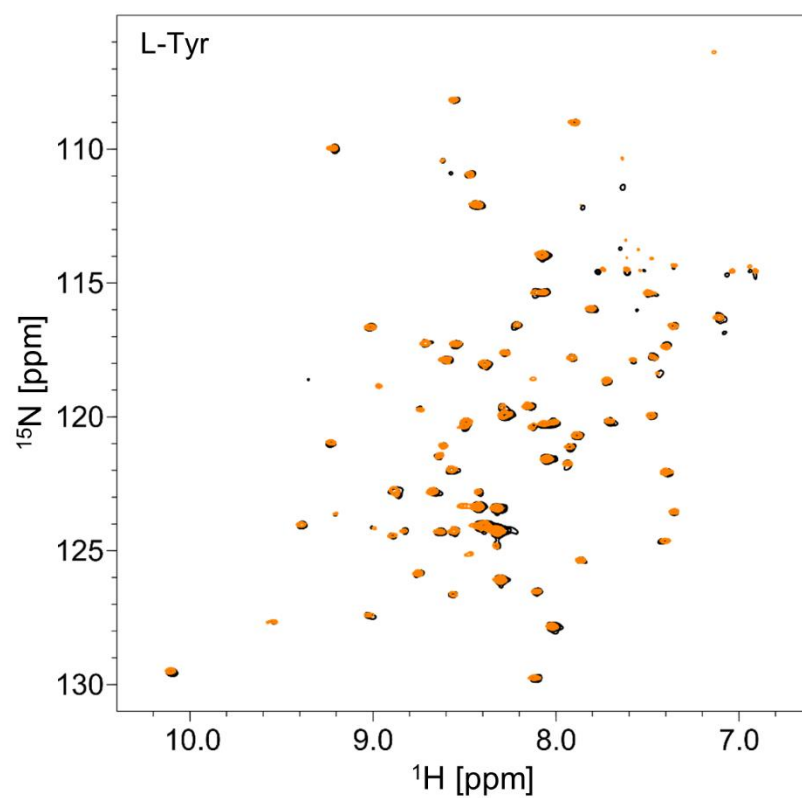


Figure 33 ^1H - ^{15}N HSQC spectra of NΔ47-TPH2-R in 12 mM Na_2HPO_4 /8 mM NaH_2PO_4 , 100 mM $(\text{NH}_4)_2\text{SO}_4$, pH 7.0, 125 μM DSS, 5% D_2O . Orange: Without ligand. Black: 1.5 mM L-Tyr.

E A KINETIC STUDY OF *HUMAN* TPH2 VARIANTS

This appendix contains a short presentation of the results from kinetic experiments. A brief introduction is provided, along with a description of the methodology. Following this the results are presented.

Introduction

In Chapter 4, it was shown that L-Phe binds stronger than L-Trp to a truncated version of the isolated RD of TPH2. However, it remains unknown if L-Phe binding to the RD has a physiological function for TPH2. To investigate if L-Phe affects the kinetics of TPH2, a kinetic study of different TPH2 variants was conducted. Three different tetrameric TPH2 variants were investigated: TPH2-CT, NΔ47-TPH2, and TPH2. The purpose of the three variants was to compare the effect of including either the truncated RD or the full RD in the TPH2 tetramer. First, the substrate specificity of TPH2-CT and TPH2 with respect to L-Trp and L-Phe was investigated. Next, kinetic parameters with L-Trp was determined for TPH2-CT, NΔ47-TPH2, and TPH2 and finally, the rate of L-Trp hydroxylation in presence of increasing L-Phe concentrations was investigated for all three variants.

Methods

Expression and purification

TPH2-CT, NΔ47-TPH2, and TPH2 were expressed and purified as described in Chapter 2.

Sample preparation

Batches of TPH2-CT, NΔ47-TPH2, and TPH2 were thawed slowly on ice and filtered through a 0.22 μm spin filter (Corning). The concentration was determined by triplicate measurements of the UV absorbance at 280 nm using an ND-1000 NanoDrop Spectrophotometer (Saveen Werner) and theoretical extinction coefficients calculated using the ExPASy server¹⁰⁵ and the protein sequence (Provided in appendix C).

Kinetic experiments using nuclear magnetic resonance (NMR) spectroscopy

1D proton NMR spectra were used to monitor the formation of 5-HTP and L-Tyr from hydroxylation of L-Trp and L-Phe respectively. The spectra were recorded with 64 scans at 15 °C on a Bruker Avance Neo 800 MHz spectrometer equipped with a 5 mm CPTXO cryoprobe. Processing was performed in Topspin 3.6.2.

The final 500 μL samples contained the following: Assay buffer (50 mM Hepes, 200 mM (NH₄)₂SO₄, pH 7.0), 0.025 mg/mL Catalase, 25 μM (NH₄)₂Fe(II)(SO₄)₂, 7 mM dithiothreitol

(DTT), 0.1 μM TPH2, 300 μM BH_4 , [0 or 300 μM] L-Trp, [0 or 300 μM] L-Phe, 5% V/V D_2O . The experiments were performed in the following way: Assay buffer, TPH2, DTT, L-Phe/L-Trp, Catalase, $(\text{NH}_4)_2\text{Fe}(\text{II})(\text{SO}_4)_2$, water for volume adjustment and D_2O was mixed thoroughly in an Eppendorf tube and transferred to a 5 mm NMR tube. An initial 1D proton spectrum was measured before BH_4 was added to the NMR tube. After BH_4 addition the sample was mixed thoroughly before re-inserting the tube in the spectrometer. 6 min after addition of BH_4 , to allow time to achieve the correct sample temperature and to perform shimming, 30 consecutive 1D proton spectra were recorded to monitor the product formation.

Initial rate kinetic experiments using a continuous fluorescent assay

The initial rates of L-Trp hydroxylation were measured by monitoring the formation of 5-HTP using a continuously fluorescent assay described by Moran et al.¹⁴⁸. The measurements were conducted in a semi-micro cell quartz cuvette (Hellma) using a Cary Eclipse fluorescence spectrophotometer (Varian) at 15 $^\circ\text{C}$ maintained by a Heto HMT 200 thermostat (Holm&Halby) and a Heto CBN 8-30 cooling bath (Holm&Halby). The excitation and emission wavelengths were 300 nm and 330 nm, respectively. The final 1000 μL assay mixture contained the following: Assay buffer (50 mM Hepes, 200 mM $(\text{NH}_4)_2\text{SO}_4$, pH 7.0), 0.025 mg/mL Catalase, 25 μM $(\text{NH}_4)_2\text{Fe}(\text{II})(\text{SO}_4)_2$, 7 mM dithiothreitol (DTT), 0.8 μM TPH2, 300 μM BH_4 , [5-200 μM] L-Trp, and [0-750 μM] L-Phe. The assay was conducted in the following way: Assay buffer, DTT, $(\text{NH}_4)_2\text{Fe}(\text{II})(\text{SO}_4)_2$, and MilliQ water for volume adjustment were added to the cuvette and allowed to incubate for 5 min to achieve the correct temperature. Then, L-Trp, L-Phe, Catalase, and TPH2 were added and the incubation continued for 2 min before the reaction was initiated by addition of BH_4 . All measurements were performed at least in triplicates. The initial rate was determined from linear regression of the first 0.06 min after addition of BH_4 . A standard curve was used to convert the rate to a unit of μM 5-HTP/min. The standard curve was constructed by measuring the fluorescence intensity of different 5-HTP concentrations [0.5, 2, 5, 10, 20, 50, 100 μM] in assay buffer with 300 μM BH_4 . The initial rates were fitted simultaneously to the standard Michaelis-Menten equation and the Michaelis-Menten equation with competitive inhibition using Origin (OriginLab) to determine K_m , V_{max} , and K_i .

Results

Investigating substrate specificity of TPH2-CT and TPH2

As described in Chapter 1 TPH2 is not fully specific towards hydroxylation of L-Trp but can also hydroxylate L-Phe⁵⁷. To investigate the substrate specificity, kinetic experiments were conducted using NMR spectroscopy, allowing us to follow the product formation of both 5-HTP and L-Tyr simultaneously. The experiments were run with two different TPH2 variants, TPH2 and TPH2-CT, and 3 different combinations of substrates were used: 300 μ M L-Trp, 300 μ M L-Phe, and 300 μ M L-Trp + 300 μ M L-Phe. TPH2 and TPH2-CT were both able to hydroxylate L-Phe, as also observed by McKinney et al.⁵⁷ (Figure 34 and Figure 35). However, when both L-Trp and L-Phe were present 5-HTP was the main product while only small amounts of L-Tyr was formed (Figure 34 and Figure 35). Thus, when equal amounts of L-Phe and L-Trp are present, it can be assumed that L-Trp will be the main substrate.

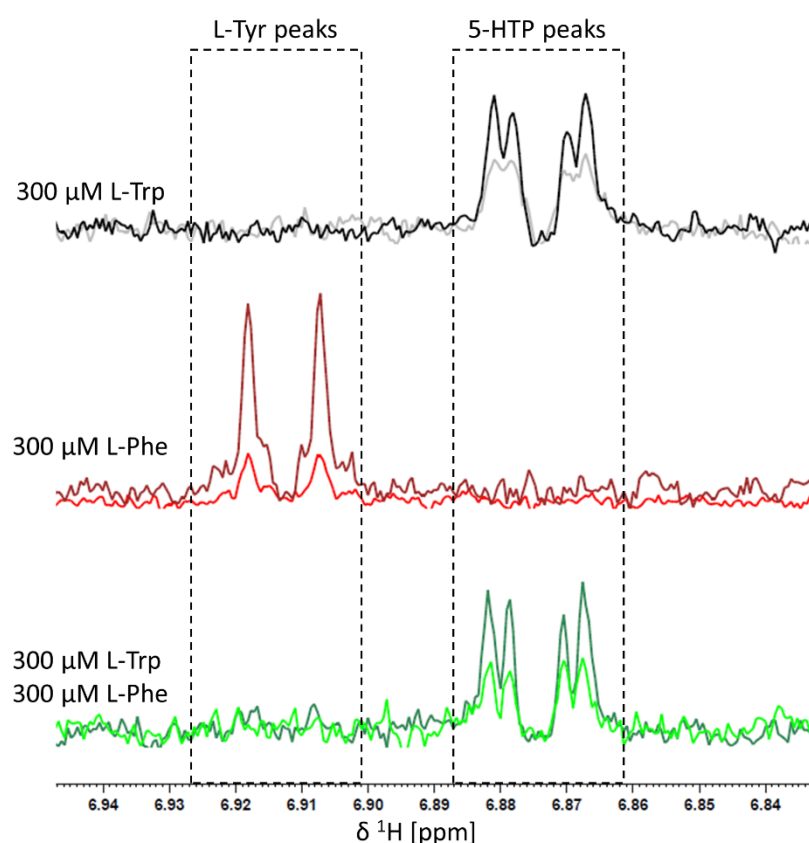


Figure 34 Kinetic experiments with TPH2 followed by NMR spectroscopy. 1D proton spectra following the conversion of 300 μ M L-Trp: (grey curve: 6 min after BH_4 addition, black curve: 87 min after BH_4 addition). 300 μ M L-Phe: (red curve: 6 min after BH_4 addition, dark red curve: 87 min after BH_4 addition). 300 μ M L-Trp + 300 μ M L-Phe: (green curve: 6 min after BH_4 addition, dark green curve: 87 min after BH_4 addition). The dashed boxes indicate the peaks belonging to L-Tyr (product of L-Phe hydroxylation) and 5-HTP (product of L-Trp hydroxylation), respectively.

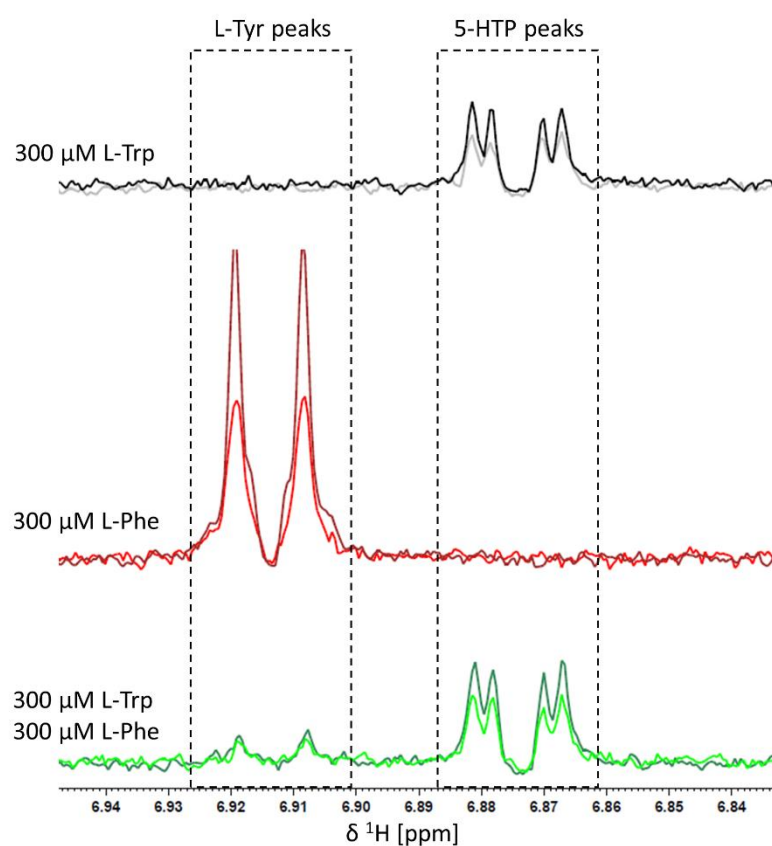


Figure 35 Kinetic experiments with TPH2-CT followed by NMR spectroscopy. 1D proton spectra following the conversion of 300 μM L-Trp: (grey curve: 6 min after BH_4 addition, black curve: 87 min after BH_4 addition). 300 μM L-Phe: (red curve: 6 min after BH_4 addition, dark red curve: 87 min after BH_4 addition). 300 μM L-Trp + 300 μM L-Phe: (green curve: 6 min after BH_4 addition, dark green curve: 87 min after BH_4 addition). The dashed boxes indicate the peaks belonging to L-Tyr (product of L-Phe hydroxylation) and 5-HTP (product of L-Trp hydroxylation), respectively.

Investigating the effect of L-Phe on the rate of L-Trp conversion

We next investigated if L-Phe binding to the RD affects the kinetics of TPH2. First, steady-state Michaelis-Menten kinetic constants were determined using the natural substrate, L-Trp, for TPH2-CT, NΔ47-TPH2, and TPH2. While the K_m values of the three variants are very similar (Table 4, 27 ± 2 μM for TPH2-CT, 25 ± 4 μM for NΔ47-TPH2, and 26 ± 3 μM for TPH2), there is a significant difference in the V_{\max} values (200 ± 7 μM 5-HTP/min for TPH2-CT, 142 ± 6 μM 5-HTP/min for NΔ47-TPH2, and 84 ± 2 μM 5-HTP/min for TPH2). This shows that inclusion of the RD has an inhibitory effect, which is in agreement with previous data on different TPH2 variants also showing smaller V_{\max} values when either the full RD or part of the RD is included^{100,149}. Next, to investigate the effect of L-Phe, the L-Trp concentration was kept constant at 100 μM , adding increasing concentrations of L-Phe (100-750 μM). We expect that L-Phe can bind in two different places: In the active site as a competitive inhibitor/alternative substrate (for TPH2-CT, NΔ47-TPH2, and TPH2) and in the RD dimer with un-known function (for NΔ47-TPH2 and TPH2). For all three variants, it is observed that L-Phe has an inhibitory effect with competitive K_i values between 200-400 μM (Table 4), showing that L-Phe binding to the RDs does not activate TPH2 as observed in the regulatory mechanism of PAH. This is in-line with a study from Ogawa et al., who also observed an inhibitory effect of L-Phe on L-Trp conversion for MBP-tagged TPH2¹⁵⁰.

Table 4 Kinetic constants of TPH2-CT, NΔ47-TPH2, and TPH2. K_m and V_{max} are the Michaelis-Menten constants of L-Trp and K_i is the competitive inhibition constant of L-Phe.

Variant	K_m [μ M]	V_{max} [μ M 5-HTP/min]	K_i [μ M]
TPH2-CT	27 \pm 2	200 \pm 7	300 \pm 50
NΔ47-TPH2	25 \pm 4	142 \pm 6	230 \pm 40
TPH2	26 \pm 3	84 \pm 2	390 \pm 60

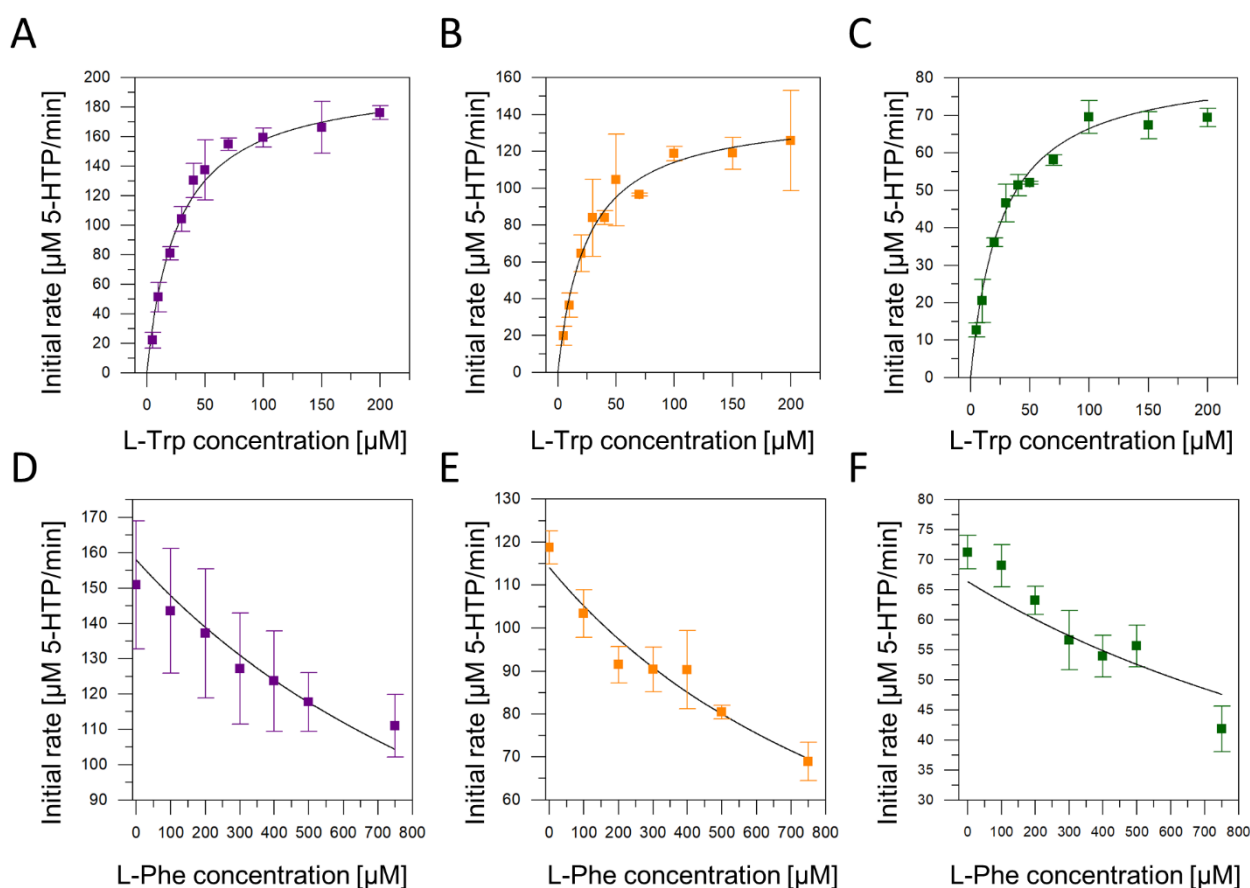


Figure 36 Kinetic data of the three TPH2 variants. A-C: initial rates of 5-HTP formation at different L-Trp concentrations using A) TPH2-CT, B) NΔ47-TPH2, and C) TPH2. D-F: Initial rates of 5-HTP formation with 100 μ M L-Trp and different L-Phe concentrations using D) TPH2-CT, E) NΔ47-TPH2, and F) TPH2. The average initial rate at each condition is shown as a square with error bars, and the fit is shown as a black line.

