

**Towards the Structural Characterisation of
AHK5, an *Arabidopsis* RedOx sensor**

and

**Structural Studies on the Bacterial Cell Wall
Modifying Enzyme GatD / MurT**

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Erik Roberto Nöldeke
aus Stuttgart

Tübingen
2021

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen

Tag der mündlichen Qualifikation:

11.02.2022

Dekan:

Prof. Dr. Thilo Stehle

1. Berichterstatter:

Prof. Dr. Thilo Stehle

2. Berichterstatter:

Prof. Dr. Remco Sprangers

ACKNOWLEDGEMENTS

This work could not have been accomplished without extensive support from many different sides and in many different shapes. It is not possible to list all of them here in a manner that reflects even remotely the support that I have received over the years. I will nonetheless attempt such a list, hoping that I do not forget anyone and hoping for forgiveness, when I inevitably do.

I wish to extend my thanks and gratitude to **Prof. Dr. Thilo Stehle** for his guidance and support throughout my time in his lab. Thilo, I have started working with you as an undergraduate assistant in 2011. Under your tutelage, first indirectly, and later directly, I have learned to become an independent scientist and to manage my own projects. I particularly thank you for the scientific freedom to operate that you have granted me while at the same time being available for any type of discussion upon short notice, despite your many other obligations. I have greatly enjoyed my time working with you both professionally and personally.

A big thanks goes to all the **members of the Stehle laboratory** for great support, discussions, and a fantastic time, both in the lab and outside of it.

Special thanks go to...

...the office gang: **Niklas Bayer, Michael Strebl, Manuel Liaci, Felix Büttner, Joana Tavares-Macedo,** and **Simon Völpel** for the great office atmosphere and inspiring scientific discussions! Additionally, I wish to thank Niklas and Michael in particular for our joint installation of an office coffee machine.

...**Volker Niemann** for his support in the early stages of my time as a doctoral candidate and for his initiation of the GatD/MurT project.

...**Lena Muckenfuß** for her work on solving the structure of unliganded GatD/MurT during her internship with Volker Niemann.

...many helping hands in the lab: **Leonie Kurz** (student assistant) for her preparative work on GatD/MurT, **Raphael Bader** (Bachelor thesis) for performing many experiments on AHK5 A3₂, **Annkathrin Scheck** (Bachelor thesis) for most of the labwork in development of the AHK5 C2 purification procedure, **Lea Hansen-Palmus** (Bachelor thesis and student assistant) for the labwork in establishing the AHK5 A2 purification procedure as well as for assistance with insect cell expression tests, and **Paul Bachmann** (Master thesis and student assistant) for his work on adapting the purification procedures for the new GatD/MurT constructs, as well as the enzymatic assays. Your help and support, both technically and scientifically has proven vital to the projects we worked on together. It has been an honour and privilege working with such motivated colleagues!

...the “robot task force”: **Felix Büttner, Irmgard Hähnlein-Schick, Georg Zocher, Christina Harprecht** for many hours spent working on the TECAN liquid handler as well as a great work in organising the acquisition, instalment, and establishment of the Gryphon. To **Elena Störk** and **Michael Strebl** many

thanks for support in the later stages of maintenance and for taking over the tutelage of the devices from Christina and myself.

...**Elena Störk** for great collaboration on the GatD/MurT manuscript, especially in helping me perform additional experiments for the revision under significant time pressure.

...**Georg Zocher** for tireless answering of questions related most of the time to either crystallography or network administration issues.

...to everybody who joined in on extracurricular activity, be it an after-work beer at Neckarmüller, an evening of concert or opera, flying, skydiving... In this context I wish to thank particularly **Niklas Bayer**, **Manuel Liaci**, and **Aleksandra Stasiak**.

...to **Simon Völpel** for caring for my projects after I left, especially GatD/MurT. It is good to know that work on these fascinating topics continues on! I wish you the best of luck!

I wish to thank **Prof. Dr. Klaus Harter** and **Thomas Drechsler** of the Center for Plant Molecular Biology at the University of Tübingen (ZMBP) for great collaboration on the AHK5 project. I particularly thank Prof. Harter for his coordinating work in the CRC1101 “Molecular Encoding of Specificity in Plant Processes”.

Many thanks also to our collaborators from the University of Bonn on the GatD/MurT project: **Prof. Dr. Tanja Schneider** and **Dr. Anna Müller** for great collaboration and discussion over the years and a fantastic support during writing of the Scientific Reports manuscript in 2018; **Prof. Dr. Dirk Menche**, and **Dr. Lukas Wingen** for their support in the search for a soluble Lipid II analogue. Many thanks to Prof. Schneider for her work in establishing the CRC-TRR261 “Antibiotic CellMAP”.

I also wish to thank **Dr. Fajun Zhang**, **Dr. Stefano DaVela**, and **Dr. Olga Matsarskaia** from the group of **Prof. Dr. Frank Schreiber** at the University of Tübingen for providing support through all stages of the SAXS experiments, including access to their in-house beamline for preliminary experiments.

Dr. Vincent Truffault (Max Planck Institute for Developmental Biology, Tübingen), thank you as well for helping with the NMR experiments on GatD/MurT.

Many thanks to **Dr. York Stierhof** of the University of Tübingen for performing negative stain TEM imaging of AHK5 A2.

Sophie Stotz (research group of **Dr. Hubert Kalbacher**, University of Tübingen), thank you for your incessant support in performing MALDI-TOF analysis on all our protein samples.

Many thanks also to **Prof. André Zapun** for discussions on GatD/MurT following the publication of our respective papers on the topic in 2018.

To **Prof. Patricia Clark** of the University of Notre Dame, many thanks for discussions on the use of the %MinMax algorithm for rare codon optimisation.

Thanks to the German Research Foundation (CRC1101, CRC-TRR261) and the Max-Planck society (International Max-Planck Research School) for funding. I also thank the Paul Scherrer Institute (Villigen, Switzerland), the German Electron Synchrotron (Hamburg, Germany), and the European Synchrotron Radiation Facility (Grenoble, France) for beam time.

I furthermore wish to thank the members of my thesis advisory committee, who accompanied my projects for the whole time: **Prof. Dr. Thilo Stehle**, **Prof. Dr. Remco Sprangers** (University of Regensburg), **Prof. Dr. Klaus Harter** (ZMBP, University of Tübingen), and **Prof. Dr. Detlef Weigel** (Max Planck Institute for Developmental Biology, Tübingen).

I thank the **Stuttgart Opera House** for student tickets, **Robert Bosch GmbH** for allowing me to partake in their chamber orchestra projects, **Valentin Flugzeugbau GmbH** for the Taifun 17E (the best touring motor glider ever built), my comrades at **Akaflieg Tübingen** for many hours on the ground and among the clouds, **Coesfelder Orchestertage**, **Brucknerorchester Coburg**, and **Bruckner Akademieorchester** and the respective participants for fantastic symphonic projects, spiced up with chamber music interludes.

Finally, to my parents **Dr. Elisabetta Nöldeke** and **Hon. Prof. Dr. Christoph Nöldeke** I wish to extend my ever-lasting gratitude. I am deeply thankful for your incessant and selfless support throughout my life. Nothing of what I have achieved would have been possible in this shape without your full support, emotionally, morally, and yes, financially. Also thank you for great travel, including to some of the remotest corners on this planet. None of this was to be taken for granted, especially not in the magnitude I was so fortunate to receive. So, thank you!

Madeleine Hollmann, Dir danke ich einfach für alles. Für Freude, Motivation, Ansporn, Ideen, Einblicke, Mut, ... Vor allem aber danke ich Dir für unsere Zukunft!

TABLE OF CONTENTS

Acknowledgements	I
Table of Contents	IV
Lists	VII
List of Tables	VII
List of Figures	VIII
List of Abbreviations	IX
Foreword	1
1 Towards the Structural Characterisation of AHK5	2
1.1 Introduction	2
1.1.1 Types of signalling receptors	2
1.1.2 Receptor histidine kinases	4
1.1.3 Histidine kinases in <i>Arabidopsis</i>	6
1.1.4 AHK5	7
1.1.5 Redox signalling	7
1.1.6 Aim of this project	8
1.1.7 Work presented in this thesis	8
1.2 Results	10
1.2.1 Computationally guided construct design	10
1.2.2 Protein expression and solubility screening	12
1.2.3 Native purification of fragment A3 ₂ (1-364)	14
1.2.4 Denaturing purification of fragment C2 (203-617)	17
1.2.5 Denaturing purification of fragment A2 (1-617)	21
1.2.6 Purified AHK5 fragments are folded and of correct identity	23
1.2.7 Purified AHK5 fragments form inconsistent oligomeric assemblies in solution	28
1.2.8 Fragments purified to date are too heterogeneous or unstable for structural studies	30
1.2.9 AHK5 fragments A2 and A3 ₂ form covalent dimers upon oxidative stress	33
1.3 Discussion	35
1.3.1 AHK5 domain architecture - a partial picture	35
1.3.2 Purified fragments allow limited inferences on AHK5 shape and function	37
1.3.3 Most AHK5 fragments are insoluble after overexpression	41
1.3.4 Structural studies are not likely to succeed with current constructs	43
1.3.5 Mode-of-action hypothesis based on redox data	44
1.3.6 Conclusion	45
1.4 Ongoing Research	46
1.4.1 Transfection of sf 9 cells with the AHK5_FL gene was successful	46
1.4.2 Initial expression tests are promising	47
1.4.3 Lysis buffer screening revealed a potential condition for native purification	48

1.4.4	Future prospects.....	48
2	Structural Studies of the Cell Wall-Modifying Enzyme GatD / MurT	50
2.1	Introduction.....	50
2.1.1	The bacterial cell wall	51
2.1.2	Antibiotics targeting the cell wall.....	53
2.1.3	Mechanisms of resistance and resistance development	54
2.1.4	Modifications of the peptidoglycan layer	55
2.1.5	The GatD / MurT bifunctional enzyme complex	56
2.1.6	Aim of this project.....	57
2.1.7	Work presented in this thesis.....	58
2.2	Results and Discussion	60
2.2.1	Publication summaries	60
2.3	Ongoing Research.....	65
2.3.1	Production and characterisation of GatD/MurT from <i>Mycobacterium tuberculosis</i> and <i>Streptococcus pyogenes</i>	65
2.3.2	Future prospects.....	69
3	Material and Methods.....	71
3.1	Instruments, chemicals, and kits.....	71
3.2	Standard protocols	76
3.2.1	Cloning.....	76
3.2.2	Culturing of <i>E. coli</i> cells.....	81
3.2.3	Protein analytics	83
3.3	Production and Purification of AHK5.....	86
3.3.1	List of buffers.....	86
3.3.2	Computational methods.....	87
3.3.3	Molecular biology and cloning	87
3.3.4	Expression of AHK5 constructs in <i>E. coli</i> cells.....	89
3.3.5	Expression verification and solubility screen	89
3.3.6	Native protein purification (fragment A3 ₂)	91
3.3.7	Denaturing protein purification from inclusion bodies (fragments A2 and C2).....	93
3.3.8	Expression of full-length AHK5 in insect cells.....	95
3.3.9	Functionality assays.....	99
3.4	Production and Purification of GatD/MurT.....	100
3.4.1	List of buffers.....	100
3.4.2	Cloning of expression constructs.....	100
3.4.3	Expression of <i>Sa</i> GatD/MurT and <i>Spy</i> GatD/MurT constructs	101
3.4.4	Expression of <i>Mtb</i> GatD/MurT constructs	101
3.4.5	Protein purification.....	101

3.4.6	Biophysical characterisation.....	104
3.4.7	Structural biology	104
3.4.8	Enzymatic activity assay	105
	Summary	106
	Zusammenfassung.....	107
4	Bibliography.....	109
5	Appendix.....	116
5.1	Superdex 200 Increase 3.2/300 Calibration	116
5.2	Original gel pictures.....	117
5.3	HHPRED search results for AHK5.....	119
5.4	Publications	125
5.5	Contributions.....	125
5.5.1	Contributions to Nöldeke et al. (2018).....	125
5.5.2	Contributions to Nöldeke and Stehle (2019).....	125

LISTS

List of Tables

Table 3-1: Instruments used in this study.	71
Table 3-2: Chromatography columns and media used in this study.	72
Table 3-3: Enzymes used in this study.	72
Table 3-4: Reagents used in this study.	72
Table 3-5: Reagents and media used for insect cell culture.	73
Table 3-6: Chemicals used in this study.	73
Table 3-7: Commercial kits used in this study.	74
Table 3-8: Solubility screens.	74
Table 3-9: Commercial crystallisation screens used in this study.	75
Table 3-10: List of standard buffers.	76
Table 3-11: Primers used for sequencing of GatD/MurT constructs.	76
Table 3-12: List of primers used for sequencing of AHK5 constructs.	77
Table 3-13: Primers used for cloning of GatD/MurT constructs.	77
Table 3-14: List of primers used for cloning of AHK5 constructs.	77
Table 3-15: Composition of PCR mix for insert amplification.	78
Table 3-16: Thermocycler protocol for insert amplification PCR.	78
Table 3-17: PCR mix for site-directed mutagenesis.	79
Table 3-18: Thermocycler protocol for site-directed mutagenesis PCR.	79
Table 3-19: Composition of restriction reactions.	79
Table 3-20: Ligation mix.	80
Table 3-21: Composition of media used for <i>E. coli</i> culture.	81
Table 3-22: Gel composition for casting of up to 13 15 % Acrylamide gels.	83
Table 3-23: Buffers used for AHK5 purification, sorted by construct.	86
Table 3-24: List of buffers used for Western Blotting.	87
Table 3-25: Constructs used for expression of AHK5 fragments in <i>E. coli</i> cells.	87
Table 3-26: Constructs used for expression of AHK5 fragments in insect cells.	88
Table 3-27: List of buffers used in sparse matrix solubility screen.	90
Table 3-28: Grid screen for lysis buffer.	90
Table 3-29: Primers for check-PCR.	96
Table 3-30: Reaction mix for check-PCR.	96
Table 3-31: Thermocycler protocol for check-PCR.	96
Table 3-32: Antibody combinations and formulations used for AHK5-FL detection. For buffer composition, see Table 3-24.	98
Table 3-33: Grid screen for AHK5_FL lysis buffer.	99
Table 3-34: Buffers used for GatD/MurT purification.	100
Table 3-35: Plasmids used for protein expression.	100
Table 5-1: Size standards used for calibration of analytical SEC column (Superdex 200 Increase 3.2/300). ..	116
Table 5-2: Top 50 HHPRED hits for full-length AHK5.	119
Table 5-3: Top 50 HHPRED hits for residues 1 – 400 of AHK5.	122

List of Figures

Figure 1-1: Representative HHPRED results	10
Figure 1-2: Overview of AHK 5 domain architecture and expression constructs.	11
Figure 1-3: Expression test for the first seven AHK5 constructs (15 h, 20 °C, 0.5 mM IPTG).	12
Figure 1-4: Identification of a suitable lysis buffer for the AHK5 A3 ₂ construct.....	13
Figure 1-5: Denaturing solubilisation tests of AHK5 fragments comprising the kinase and ATPase domains. .	13
Figure 1-6: Overview of the purification workflow for AHK5 fragment A3 ₂	14
Figure 1-7: Overview of the purification workflow for AHK5 fragment C2.	17
Figure 1-8: Near-quantitative aggregation of AHK5 C2 in previous refolding strategies.....	19
Figure 1-9: Overview of AHK5 A2 purification.	21
Figure 1-10: CD spectra of purified AHK5 constructs.	23
Figure 1-11: Protease exclusion assay performed on AHK5 fragment A3 ₂	24
Figure 1-12: Protease exclusion assay performed on AHK5 fragment C2.	25
Figure 1-13: Protease exclusion assay performed on AHK5-A2.....	27
Figure 1-14: Size and homogeneity analysis of AHK5 fragments.....	28
Figure 1-15: Example of granular precipitate in AHK5 crystallisation trials.....	30
Figure 1-16: SAXS data of AHK5 A3 ₂	31
Figure 1-17: Negative stain TEM images of AHK5 A2.	32
Figure 1-18: Oxidative dimerisation of AHK5 A3 ₂	33
Figure 1-19: Oxidative multimerisation of AHK5 A2.	34
Figure 1-20: Possible domain arrangement of the central portion of AHK5 (residues 200-620).	36
Figure 1-21: Effect of rotational ellipsoid shape on apparent stokes radius and molecular weight estimation via SEC.	40
Figure 1-22: AHK5 codon usage as calculated using the %MinMax algorithm.....	42
Figure 1-23: Hypothetical model of AHK5 oxidative activation.....	44
Figure 1-24: Signs of infection of Sf9 insect cells.....	46
Figure 1-25: Expression test of AHK5_FL adherent Hi5 cells using P2 viral stock.....	47
Figure 1-26: Excerpt of a lysis buffer screen on AHK5_FL.....	48
Figure 2-1: Overview of the steps involved in peptidoglycan biosynthesis in <i>Staphylococcus aureus</i>	52
Figure 2-2: Analytical SEC traces of purified GatD/MurT homologues.	66
Figure 2-3: CD spectra of purified GatD/MurT homologues.....	67
Figure 2-4: Glutaminase activity of <i>Sa</i> GatD/MurT and <i>Mtb</i> GatD/MurT monitored by NMR.....	68
Figure 2-5: Glutaminase activity of <i>Spy</i> GatD/MurT monitored by NMR.	69
Figure 3-1: Schematic overview of the AHK5 A3 ₂ fragment purification strategy.	91
Figure 3-2: Denaturing purification strategy for A2 and C2 fragments.....	93
Figure 3-3: Schematic overview of the GatD/MurT purification strategy.....	102
Figure 5-1: Regression analysis of calibration of Superdex 200 Increase 3.2/300 calibration.....	116
Figure 5-2: Original gel images for Figure 1-3 without labelling asterisks.	117
Figure 5-3: Original gel image for Figure 1-4a without labelling asterisks.	117
Figure 5-4: Original gel image from Figure 1-7e without cropping.	117
Figure 5-5: Original gel image from Figure 1-9f, right panel, without cropping.....	117
Figure 5-6: Original gel image for Figure 1-11a without labelling asterisks.	118
Figure 5-7: Original gel image for Figure 1-12a without labelling asterisks.	118
Figure 5-8: Original gel image for Figure 1-13a without labelling asterisks.	118
Figure 5-9: Original gel image for Figure 1-19 without labelling asterisks.	118

List of Abbreviations

Abbreviation	Full explanation
AG	Arabinogalactan
AHK	Arabidopsis Histidine kinase
AHP	Arabidopsis Histidine-containing Phosphotransfer protein
AK	Research group (German: Arbeitskreis)
AMP	Adenosine 5' monophosphate
AMPPNP	Adenosine 5'-(β,γ -imido)triphosphate
APS	Ammonium persulfate
ARNT	Aryl Hydrocarbon Receptor Nuclear Translocator
aSEC	analytical Size exclusion chromatography
ATP	Adenosine 5'-triphosphate
B. Sc.	Bachelor of Science
BLAST	Basic Local Alignment Search Tool
CD	Circular dichroism
CDC	Centers for Disease Control and Prevention
CE	Capillary electrophoresis
CHASE	Cyclases/Histidine kinases Associated Sensory Extracellular
CKI1	Cytokinin Insensitive 1
CRC	Collaborative Research Council
CV	Column Volume
DB	Data bank
DESY	Deutsches Elektronensynchrotron (German Electron Synchrotron)
DFG	Deutsche Forschungsgemeinschaft (German Research Foundation)
DHpD	Dimerisation and histidine phosphotransfer domain
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
DSF	Differential scanning fluorimetry
DTT	DL-1,4-Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EM	Electron microscopy
ER	Endoplasmic reticulum
ERS1	Ethylene response sensor 1
ESRF	European Synchrotron Radiation Facility
ETR1	Ethylene Receptor 1
FAD	Flavin adenine dinucleotide
FBS	Fetal bovine serum
FL	Full-length
FPLC	Fast protein liquid chromatography
FT	Flow-through (in context of purification steps)
FW	Forward

Abbreviation	Full explanation
GATase	Glutamine amidotransferase
GDP	Guanosine 5'-diphosphate
GPCR	G-protein coupled receptor
GST	Glutathione-S-transferase
GTP	Guanosine 5'-triphosphate
HATPase	Histidine kinase-like ATPase domain
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HGT	Horizontal gene transfer
HisKA	Histidine Kinase domain
HK	Histidine kinase
HPLC	High performance liquid chromatography
Hpt	Histidine-containing phosphotransfer protein
HRE	Hormone Response Element
IFIB	Interfaculty Institute of Biochemistry at the University of Tübingen
IJMM	International Journal of Medical Microbiology
IMAC	Immobilised metal ion affinity chromatography
IP3	Inositol 1,4,5-triphosphate
IPTG	Isopropyl- β -D-Thiogalactopyranoside
IV	Intravenous
LB	Lysogeny broth
LC-MS	Liquid chromatography coupled to mass spectrometry
LCP	Lipidic cubic phase
LOV	Light-oxygen-voltage-sensing domain
LRR-RLK	Leucine-rich repeat receptor-like kinase
M. Sc.	Master of Science
MALDI - TOF	Matrix-assisted laser desorption ionisation - time of flight
MAPK	Mitogen-activated protein kinase
mDAP	meso-Diaminopimelic acid
MES	2-(N-morpholino)ethanesulfonic acid
MOI	Multiplicity of infection
MPI	Max Planck Institute
MRE	Mean residue ellipticity
MRSA	Methicillin-resistant Staphylococcus aureus
MS	Mass spectrometry
MSP	Multistep phosphorelay
<i>Mtb</i>	prefix before a protein from <i>Mycobacterium tuberculosis</i>
MWCO	Molecular weight cut-off
NADH	Nicotinamide adenine dinucleotide, reduced form
NF- κ B	Nuclear factor kappa-light-chain-enhancer of activated B cells
NMR	Nuclear magnetic resonance
NTA	Nitrilotriacetic acid
NTP	Nucleotide 5'-triphosphate
O/N	overnight

Abbreviation	Full explanation
OD ₆₀₀	Optical density at 600 nm (equivalent to extinction at 600 nm)
PAGE	Polyacrylamide gel electrophoresis
PAMP	Pathogen-associated molecular pattern
PAS	Per-Arnt-Sim
PBP	Penicillin binding protein
PCR	Polymerase chain reaction
PDB	Protein data bank
Per	Periodic circadian protein
PES	Polyether sulfone
PI3K	Phosphoinositol-3-Kinase
PKA	Protein kinase A
PMSF	Phenylmethyl sulfonyl fluoride
PPP	Protein sample buffer for SDS-PAGE
PVDF	Polyvinylidene fluoride
QC	Quick Change PCR (site-directed mutagenesis)
Ran	Ras-related nuclear protein
RanBP	Ran binding protein
REC	Receiver domain
RMP	Redox midpoint potential
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RRT	Random reverse translation
rSAP	recombinant Shrimp alkaline phosphatase
RT	Room temperature
RV	Reverse
<i>Sa</i>	prefix before a protein from <i>Staphylococcus aureus</i>
SARS-CoV2	Severe acute respiratory syndrome Coronavirus 2
SAXS	Small-angle X-ray scattering
SDS	Sodium dodecylsulfate
SEC	Size exclusion chromatography
SH2	Src homology 2
Sim	Single-minded protein
SMART	Simple Modular Architecture Research Tool
SN	Supernatant
SOP	Standard operating procedure
<i>Sp</i>	prefix before a protein from <i>Streptococcus pneumoniae</i>
<i>Spy</i>	prefix before a protein from <i>Streptococcus pyogenes</i>
Src	a proto-oncogene tyrosine kinase associated with sarcoma
TAE	Tris-Acetic acid-EDTA
TB	Terrific broth
TBS	Tris-buffered saline
TCEP	Tris(2-carboxyethyl)phosphine

Abbreviation	Full explanation
TCS	Two component system
TEM	Transmission electron microscopy
TEMED	N,N,N',N'-Tetramethyl ethylenediamine
TEV	Tobacco etch virus
TGF-beta	Tumour growth factor beta
TLR	Toll-like receptor
Tris	Tris(hydroxymethyl)aminomethane
UDP	Uridine 5'-diphosphate
UMP	Uridine 5'-monophosphate
UV	Ultra-violet (radiation)
WHO	World Health Organisation
WTA	Wall teichoic acid
ZMBP	Center for Plant Molecular Biology at the University of Tübingen
β-ME	β-mercaptoethanol

FOREWORD

It is probably a rare occurrence that two seemingly unrelated topics are found side by side in one and the same dissertation. This happened due to one of the many beauties of structural biology and protein biochemistry. By using a dedicated portfolio of highly sophisticated analytical methods, utterly different biological questions can be addressed in the same manner. As long as a protein can be overexpressed and subsequently purified, it can then be further studied by biophysical and structural methods, regardless of whether the original donor organism is a plant, a mammal, a prokaryote, or even a virus. No complicated, time-consuming, and expensive culturing techniques like for *in vivo* studies are required.

As soon as the protein is available, studies into its physical properties, its structure, and thus its molecular mode of action can be initiated. This particular study presents recent advances on the structural and mechanistical understanding of two very different proteins: the redox-sensitive histidine kinase AHK5 from *Arabidopsis thaliana* and the cell wall-amidating enzyme complex GatD/MurT from *Staphylococcus aureus*, *Mycobacterium tuberculosis*, and *Streptococcus pyogenes*. The two projects are currently at different stages in their development and different strategies were used to answer their respective questions. In the AHK5 project biophysical methods were used to optimise expression constructs and purification strategies and traditional biochemical assays serve to assess the protein's functionality. For GatD/MurT, in contrast, crystallographic data was obtained, analysed and used to infer possible mechanisms of action.

I am very grateful to have been given the opportunity to work on these two interesting and challenging projects that have allowed me to study such fundamentally different biological pathways. I was thus confronted with adapting to different challenges and found ways to tune the methods and devices at my disposal to address the specific needs of each project. I firmly believe that the experience gained from either project helped to further the other project in ways that would have been far less likely to happen, had the projects been treated separately.

1 TOWARDS THE STRUCTURAL CHARACTERISATION OF AHK5

1.1 Introduction

One crucial feature, common to all organisms, is the ability to adapt their behaviour in response to changes in their immediate surroundings. Such changes can involve, for example, nutrient and water availability, the presence of a predator, or the availability of a potential mate. Every environmental change that is relevant to the organism is perceived, processed, and met by a very specific cellular response. The specificity of the response is achieved by a highly intricate network of signal transduction pathways. These pathways are composed, much like a computer program, of an activation command, means of information targeting, transport, and storage, and signal amplification and suppression tools. Fine tuning is then achieved by an elaborate system of positive and negative feedback loops, as well as crosstalk between different signal transduction pathways.

Activation of this sophisticated regulatory machinery requires the ability to sense and recognise chemical signals and translate them into specific physiological responses. This task is accomplished in all kingdoms of life by a wide variety of biosensors, termed receptors.

1.1.1 Types of signalling receptors

In order to reliably, specifically, and quickly achieve activation of signalling pathways, signalling receptors must possess two key features: Firstly, they must exist in at least two states, one that is dormant or inactive and one that is signalling-competent. Secondly, they must sense and react to their specific stimulus in a way that triggers transition from the dormant to the active state. This is achieved in different ways depending on the specific pathways, but many underlying mechanisms are conserved across the kingdoms of life.

1.1.1.1 Nuclear Receptors

Perhaps the most straightforward type of signalling receptor is the family of nuclear receptors. These receptors act as receptors for many freely diffusible small molecules, such as steroid hormones. In their inactive state they are found either as monomers or dimers, depending on the type and are mostly bound to inhibiting chaperones. Recognition and binding of the activating molecule induces release of the chaperone and, in the case of inactive monomers, homo- or heterodimerisation. The active dimer then influences directly the expression level of target genes by acting directly as a transcription factor. It binds to so-called Hormone Response Elements (HRE) on the DNA upstream of target genes and recruits chromatin remodelling enzymes or components of the transcription machinery [1].

1.1.1.2 Membrane-associated receptors without intrinsic enzymatic activity

In the case of signal molecules that are not able to cross a membrane on their own, receptors are required that enable transmission of the signal across the membrane. Most known receptors are,

indeed, transmembrane receptors. Depending on the scope, these receptors can be grouped in different ways, based either on the number of membrane-spanning α -helices, mechanism of activation, oligomeric state, transporter functionality, or intrinsic enzymatic activity.

Receptors without intrinsic enzymatic activity act by ligand-induced changes in conformation or oligomeric state. These changes are mechanically transferred across the membrane where they activate downstream regulatory elements. The largest family of this type of receptors is that of the G-protein-coupled receptors (GPCRs). These receptors are made up of seven transmembrane helices, connected by loops both on the inside and outside of the cytosol. Upon binding of the activating molecule (the agonist), the receptor undergoes a conformational change that leads to the activation of a heterotrimeric adaptor protein (the G-protein) that is associated with the receptor in its inactive state. Upon activation, the G-protein exchanges the GDP molecule bound to its G_α subunit for GTP and dissociates into the monomeric G_α and heterodimeric $G_{\beta\gamma}$ subunits. These, in turn activate further downstream pathways, such as the adenylate cyclase and PI3K pathways, respectively [2]–[6].

Other receptors that act without intrinsic enzymatic activity are those that create scaffolds for the assembly of large signalling complexes upon oligomer formation. The most prominent examples of these are integrins and Toll-like receptors. Integrins mediate signals initiated by contact of cells with components of the extracellular matrix. Upon activation, integrins cluster at the contact site and thus recruit their intracellular signalling complex [7]. These complexes can have direct action on downstream signalling cascades, such as the MAPK (mitogen-associated protein kinase) pathway [8], or act as modulators for other signalling receptors. Similarly, Toll-like receptors (TLRs) likely achieve their active state upon ligand-induced dimerisation. Ligands for TLRs include many pathogen-associated molecular patterns (PAMPs), as their main role is in the modulation of somatic mechanisms of the innate immune response, such as the activation of macrophages and dendritic cells. The signalling complex assembled by TLR dimerisation diversifies into multiple downstream pathways, including, again, various MAPK pathways, and the NF- κ B pathway [9].

1.1.1.3 Ligand and voltage-gated ion channels

A special case is that of ligand and voltage-gated ion channels. As opposed to other receptors described here, they do not trigger complex cascades of downstream effectors and regulators. Instead, they directly allow the influx or efflux of specific ions into or from the cell upon reception of the signal. These channels are usually specific for one particular type of ion and are heavily involved in neuronal transmission. Ligand-gated ion channels are the primary receptors at chemical synapses where they sense the concentration of neurotransmitters in the synaptic cleft (e.g. acetylcholine). Upon ligand binding, the receptor allows for cation influx and thus membrane depolarisation. This signal is then carried along the neuron's surface by voltage-gated ion channels in the form of a so-called action potential [10], [11].

A further example for a ligand-gated ion channel is the Inositol 1,4,5-triphosphate (IP3) – gated Ca^{2+} – channel, which mediates influx of calcium ions from the endoplasmic reticulum into the cytosol downstream of the α -1 adrenergic receptor, a GPCR.

1.1.1.4 Receptor kinases

Receptor kinases are protein kinases that transition from the inactive to the catalytically active form upon signal perception. The initiation of the downstream signalling pathway (often a cascade of subsequent protein kinases), is then mediated by phosphorylation of the receptor's target protein.

The most frequent types of receptor kinases are of the tyrosine-specific variety, i.e. they phosphorylate the hydroxyl group of a tyrosine residue upon activation [12], [13]. Receptor tyrosine kinases are active as dimers and signal perception leads to autophosphorylation of the cytosolic portion of the receptor. These phosphorylated variety then serves as a scaffold to recruit adaptor proteins via so-called SH2 domains (Src Homology 2) that bind specifically to phosphotyrosine-containing sequences. Prominent examples of the Receptor tyrosine kinase family are the insulin receptor [14] and the family of epidermal growth factor receptors (EGFR) [15].

Many downstream signalling kinases, such as the components of the MAPK cascade, are serine/threonine-specific kinases. This type of kinase also exists as a receptor kinase, for example the TGF- β receptor in mammals [16], [17] and the family of Leucine-rich repeat receptor-like kinases in plants (LRR-RLKs) [18]–[21]. Similarly to tyrosine-specific receptor kinases, the first step in signalling activation upon ligand perception, is autophosphorylation within a homo- or heterooligomer (dimer or higher). However, while receptor tyrosine kinases generally transmit signals via the assembly of a phosphorylation-specific signalling scaffold complex, receptor serine/threonine kinases additionally use their intrinsic kinase activity to directly phosphorylate downstream regulators, such as the Smad proteins in the context of TGF- β signalling.

1.1.2 Receptor histidine kinases

Histidine kinases take a special place among protein kinases. As the name suggests, histidine-specific protein kinases phosphorylate a histidine residue in their substrate. However, in contrast to tyrosine- or serine/threonine-specific protein kinases, the resulting modification does not merely change the behaviour of the phosphorylated protein. Instead, the phosphoryl group acts as a relay and is itself transferred along the signalling cascade, all the way from the receptor histidine kinase, directly or via one or multiple phosphotransfer proteins, and onto the final effector, termed response regulator.

1.1.2.1 Architecture and principles of signalling

Histidine kinases are dimeric proteins constructed around a functional core consisting of an N-terminal sensory input domain followed by the kinase portion [22]. In many cases, the sensory domain is located on the extracellular side of the plasma membrane and connected with the kinase domain via a transmembrane segment. The kinase portion itself consists of two separate domains. The first one, termed HisKA in the SMART protein domain annotation system [23], and also known as DHp domain [24], is a 60-70 amino acid helical domain. It consists of two α -helices of approximately equal length that are connected by a short hairpin and fold back onto one another to form an antiparallel bundle. Two such domains constitute the heart of the histidine kinase dimer by forming an antiparallel

four-helix coiled-coil. At the beginning of the HisKA domain lies a conserved histidine residue that functions as the primary acceptor substrate for the phosphorylation reaction.

The HisKA domain is C-terminally connected to a catalytic ATPase domain, termed HATPase in the SMART database. When the histidine kinase is activated, this domain uses ATP to phosphorylate the conserved histidine residue in the HisKA domain. Depending on the respective orientation of the domains, the reaction can take place either *in cis* or *in trans* within the histidine kinase dimer. Both cases have indeed been reported (reviewed in [24]).

Autophosphorylation of the histidine kinase at its conserved histidine residue creates a high energy modification in the form of a phosphoramidate. This relatively unstable modification is the central feature distinguishing histidine kinases from serine/threonine and tyrosine-specific kinases. The latter all catalyse the formation of phosphoester bonds between the γ -phosphate of ATP and the hydroxyl group in their respective substrate side chain. Thus, the ratio of phosphorylated to unphosphorylated histidine residues at equilibrium condition is very low [25], and the half-life of the modification can be expected to be by far inferior to that of phosphorylated alcoholic side chains. Histidine kinase signalling hence is characterised not by a large change in bulk phosphorylation, but rather in the modulation of phosphoryl flux [25]. The high-energy phosphoramidate in the phosphorylated histidine, in turn, is key in driving the downstream relay of the phosphoryl group to its final acceptor. After autophosphorylation, the phosphoryl group is transferred to an aspartate on a second protein, termed response regulator. This transfer can occur directly, as is the case for the evolutionarily well-established two-component system (TCS) found in bacteria. It can also take place via multiple interim steps involving shuttle proteins, giving rise to a variety of multi-step phosphorelay system (MSP) that are found in more recent organisms such as fungi and plants. The response regulator, which frequently acts as a transcription factor, finally initiates the cellular reaction to the perceived stimulus.

1.1.2.2 Two component system (TCS)

The TCS is the most common type of intracellular signalling mechanism found in eubacteria [25]. It is composed of two namesake components, the sensory histidine kinase and the response regulator. The histidine kinase contains only a specific signal-perceiving input domain and the HisKA / HATPase pair. The response regulator in turn consists of an aspartate-containing receiver domain and an effector domain [25]. The nature and identities of sensory input and effector domains are as diverse as the signals they sense. In most cases the effector domain of the response regulator is a transcription factor. Aspartate phosphorylation in the receiver domain modulates the transcription factor's affinity to its DNA target thus altering gene expression. As bacteria have no nucleus, a direct contact between the histidine kinase at the plasma membrane and the response regulator on the chromosome is possible, enabling direct phosphotransfer from the kinase to the receiver domain.

1.1.2.3 Multi-step phosphorelay (MSP) [26], [27]

In contrast to bacteria, which use TCS signalling as the most widespread means of signal transduction, only few such pathways exist in eukaryotes [25]. Fungi use histidine kinase signalling for selected pathways [26], and a number of active sensory histidine kinases are known in plants [28]–[34],

[35], [24], [36]–[40]. As in TCS, the systems are composed of a sensory histidine kinase and a (usually transcription-modulating) response regulator. However, as eukaryotes have strict topological separation of the cytosol and the genome by means of the nucleus, no direct contact between the histidine kinase and the response regulator can take place. In order to maintain the same basic signalling system of phosphoryl flux from sensor to effector, a means to shuttle the phosphate from the cytosol into the nucleus is required. This task is accomplished by two modifications to the TCS. Firstly, the use of a hybrid histidine kinase, which contains an additional C-terminal aspartate-containing receiver domain in the histidine kinase and a histidine-containing phosphotransfer protein (HPT) [41]. Hybrid histidine kinases and HPTs exist in bacteria as well but orthodox TCSs prevail (only five hybrid histidine kinases out of 30 in *E. coli* [42]). The receiver domain of hybrid histidine kinases is similar to that of the response regulators and indeed contains the same phosphorylation site in the form of an aspartic acid. Thus, instead of transferring the phosphoryl group from the histidine in the HisKA domain directly to the response regulator, the first transfer occurs to the receiver domain *in cis*. In the second step, a HPT binds to the kinase's receiver domain and the phosphoryl group is transferred to its conserved histidine. These proteins have a highly conserved structure composed around a four helix bundle with two antiparallel pairs similar to the HisKA core [24], [40] and two additional C-terminal helices. After receiving the phosphoryl group, the HPT relocates to the response regulator in the nucleus, where the third and final phosphotransfer step takes place. This three-step mechanism gave the pathway its name of multistep phosphorelay (MSP).

In addition to providing a phosphoryl shuttling mechanism from the cytosol into the nucleus, MSP offers a wider potential for cross talk between various signal transduction pathways. Indeed, some histidine kinases in *Arabidopsis* have been shown to interact with multiple HPTs [24], [43]. Additionally, many interactions with Ser/Thr kinase cascades have been reported, such as between the ethylene receptors ETR1 and ERS1 and MAPK pathways in *Arabidopsis* [44].

1.1.3 Histidine kinases in *Arabidopsis*

The *Arabidopsis thaliana* genome encodes eight functional histidine kinases [24]. These can be topologically grouped into three categories. Firstly, the ethylene sensors with three N-terminal transmembrane helices (ETR1, ERS1). Ethylene, which is a freely diffusible hydrophobic gas, is enriched within the hydrophobic environment inside the plasma and ER membranes. Thus, ETR1 and ERS1 can use a single combined sensing and membrane anchoring domain. Ethylene plays a crucial role in germination, leaf abscission, and fruit maturation, as well as organ development, and stress or pathogen response [45], [46].

The second class is that of sensory histidine kinases with one extracellular or ER-luminal sensory domain encompassed by two transmembrane helices (CKI1, AHK1-4). These comprise the cytokinin sensors AHK2-4 [28], [31], [32], [47], [48], the putative osmosensor AHK1 [35], [49], and the cytokinin-independent growth regulator CKI1 [50]. Of these, the best-characterised ones are the cytokinin receptors. These receptors are mainly located at the endoplasmic reticulum [28] and sense cytokinin levels by means of their extracellular CHASE domain (Cyclases/Histidine kinases Associated Sensory Extracellular) [51]. Cytokinin mediates cell growth and division, as well as differentiation and apical dominance. Cytokinin signalling is reviewed in [52].

AHK4 is the only *Arabidopsis* histidine kinase for which a high-resolution structure of the sensory input domain is available [31]. It consists of an N-terminal transmembrane segment, followed by a long helix that mediates homodimerisation. The polypeptide chain then loops back towards the plasma membrane and forms two subsequent PAS-like folds before the second transmembrane segment. This fold, termed a CHASE domain is common in bacterial sensory proteins [51]. Based on sequence similarity, it is likely that all *Arabidopsis* histidine kinases pertaining to this class contain this fold in their sensory input domain [53].

The third class, which has no transmembrane segment at all, contains only one member, the all-cytosolic, putatively H₂O₂-sensing AHK5 [24], [29], [30], [34].

1.1.4 AHK5

AHK5 has functionally been implicated in stomatal closure upon hydrogen peroxide stress perception [36]. Interestingly, AHK5 has been shown to be essential for stomatal closure in response to diverse stimuli that are known to cause alterations in cellular H₂O₂ levels [37], [38]. It has thus been proposed that AHK5 functions as an integrator of H₂O₂-mediated signals at multiple levels [36].

AHK5 is the only known histidine kinase in *Arabidopsis* that does not contain any transmembrane segments. It has been shown to both localise to the cytosol and to associate with the plasma membrane [36]. However, the mechanism by which AHK5 is enriched at the membrane is not known. Both N-myristoylation and interaction with scaffolding proteins have been suggested [36] but no evidence regarding either could yet be obtained.

The other feature that sets AHK5 apart from the other *Arabidopsis* HKs is its N-terminal sensing domain. The Cytokinin-sensing AHKs, AHK1, and CKI1 probably all possess CHASE or CHASE-like domains as their input domains, while the ethylene sensors have a set of three helices embedded in the membrane that serve to bind ethylene. In contrast, very little is known about the input domain of AHK5. Preliminary data obtained by Dr. Michael Heunemann and Prof. Dr. Klaus Harter (ZMBP, University of Tübingen) show that the N-terminal 350 residues are crucial for regulation. In particular, a cysteine residue at position 3 appears to be required for signal perception (Prof. Dr. Klaus Harter, personal communication). However, previous computational attempts at predicting the domain architecture of this N-terminal stretch were unsuccessful (Dr. Johannes Bauer and Prof. Dr. Thilo Stehle, IFIB, University of Tübingen, personal communication).

As AHK5 integrates H₂O₂ signals in response to many diverse primary stimuli, unravelling the mechanism of signal perception could prove key to understanding many different physiological responses and relationships between signalling pathways.

1.1.5 Redox signalling

Changes in the redox potential outside or inside the cell are dramatic developments that cells need to react to in order to survive. Hence, many mechanisms are capable to sense varying degrees of oxidative stress, from small changes to the redox potential to the presence of highly reactive oxygen species (ROS).

Redox sensing by proteins can be mediated by a number of different mechanisms, most of which involve the modification of surface-exposed cysteine residues [54]–[56]. Sensing is initiated by the reversible oxidation of the thiol function to sulfenic acid. This unstable intermediate can react with other free thiols to form disulfides, releasing one water molecule. This reaction can occur either with other cysteines or with small molecules, such as glutathione. Such modifications are termed S-thiolations. Sulfenic acids have also been shown to be S-alkylated in the presence of quinones or S-chlorinated by hypochloric acid [54], [57], [58]. These modifications can cause changes in protein activity either by directly affecting the local chemical environment, by covalently crosslinking multiple components of a signalling complex, or by providing charge patterns required for binding of additional factors, similarly to the phosphotyrosine-specific SH2 domains [59].

Furthermore, Light Oxygen and Voltage-sensing domains (LOV), a subtype of the widespread Per-Arnt-Sim (PAS) domains, have been described to act as redox sensors via oxidation / reduction of their prosthetic Flavine-Adenine-Dinucleotide (FAD) moiety and subsequent local hydrogen bond rearrangements [60]. PAS domains have also been described as signal mediators and sensory domains in many histidine kinases [61]–[63].

1.1.6 Aim of this project

In contrast to the prokaryotic two-component system, only very little is known about the molecular mechanisms of signal input of plant MSP. The only *A. thaliana* histidine kinase for which a high-resolution structure of the sensory input domain is available is the cytokinin-sensing AHK4 [31]. Given the variety of signals perceived by the eight histidine kinases, a significant variety of input domain structures is to be expected.

This project aims to obtain a high-resolution structure of the sensory input domain of AHK5, in order to unravel the its mechanism of signal perception and transduction. The main focus is to establish whether AHK5 employs canonical methods of redox sensing, such as those described in section 1.1.5. As redox signalling should be a dose-responsive reaction and not a simple on-off switch, a combination of activatory stimuli is possible and is to be investigated.

Understanding and being able to predict how an organism reacts to certain ambient stimuli requires a thorough understanding of the molecular mechanisms involved in the transduction of these stimuli into cellular reactions. The lack of clear sequence homologues raises the possibility that the sensory domain of AHK5 may contain previously undescribed protein folds, thus potentially offering the possibility to study novel mechanisms of signal perception.

1.1.7 Work presented in this thesis

The work presented here constitutes a foundation for studying the AHK5 input domain *in vitro*. Bioinformatical analyses revealed a region of low homology to PAS / LOV domains directly upstream of the Histidine Kinase domain, raising the possibility that a PAS-like sensing or transducing mechanism could be involved in AHK5 function.

Several protein fragments containing varying portions of the AHK5 input and kinase domains were successfully expressed and purified. Reproducible purification protocols were established for

three fragments yielding high amounts of pure protein. Although the protein samples produced to date exhibited varying degrees of heterogeneity that prevented structural analyses, the proteins were successfully used to assess H₂O₂ responsiveness *in vitro*. The results underline the previously postulated role of Cys3 in signal perception, as it mediates the covalent dimerisation of the N-terminal sensory domain in a H₂O₂-dependent manner. Together with sequence-based computational analyses, which suggest that two out of the four cysteines in the ATP binding domain of the kinase could be surface-exposed, the data are used to formulate a working hypothesis on how AHK5 senses the cellular redox potential.

Additionally, a successful protocol for expressing full-length AHK5 in insect cell culture is presented. Natively-purifiable full-length AHK5 can be used in the future to better characterise its functions and to assess the domain architecture *in vitro* by limited proteolysis.

Taken together, the results presented here help further the understanding of AHK5 function. They provide knowledge and tools required to push future experiments towards the ultimate goal of elucidating the mechanism of H₂O₂-responsiveness by AHK5.

1.2 Results

1.2.1 Computationally guided construct design

As there was no previous information available on the architecture of the AHK5 N-terminal sensory input region, potential domain boundaries were estimated computationally. In a first round, the full-length amino acid sequence was subjected to a secondary structure alignment-based database search using the HHPRED algorithm implemented in the MPI Bioinformatics Toolkit (toolkit.tuebingen.mpg.de, [64], Figure 1-1).

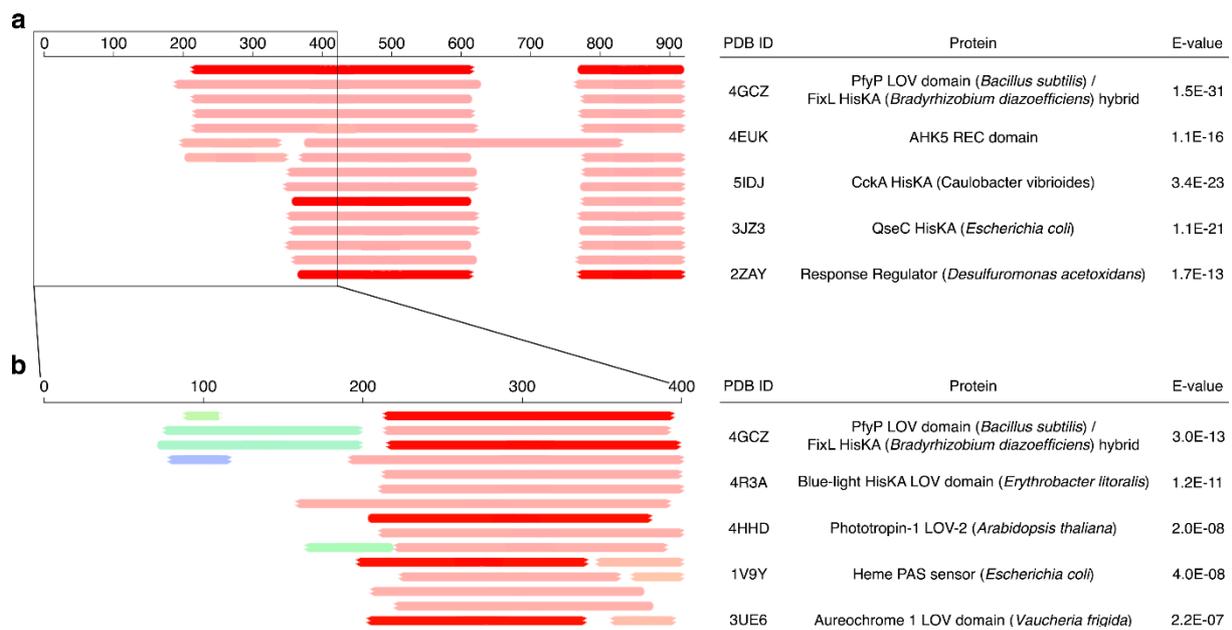


Figure 1-1: Representative HHPRED results [64]. **(a)** First 15 result lines for the full AHK5 sequence. The highlighted hits are listed with their respective E-values in the table on the right. **(b)** As in a, but restricted to the first 400 residues in AHK5. The top 50 hits of the HHPRED searches are shown in the appendix in **Table 5-2** and **Table 5-3** for panels (a) and (b), respectively. The database version used for the searches is as of 15/Oct/2019.

HHPRED database mining with the full length AHK5 sequence (Figure 1-1a) yielded multiple high confidence (E-value below 10^{-3}) secondary structure homologues covering the C-terminal receiver domain (residues 774 - 922, [40]) and the central catalytic segment (residues ~350 - 630). Interestingly, the first five hits extend N-terminally until roughly residue 200. A second iteration of HHPRED using only the first 400 residues revealed homology to multiple proteins of bacterial and plant origin. The aligned regions of these proteins contain either a Per-Arnt-Sim (PAS) or the closely related Light-oxygen-voltage-sensing (LOV) domain (Figure 1-1b). However, the E-values for these alignments are distinctly higher than for the longer histidine kinase domains.

In order to provide a more complete picture of the architecture of the AHK5 N-terminus, additional database searches were performed to identify secondary structure and disorder propensity, sequence conservation, and homology to structurally known domains. Results from the different search strategies were pooled and used to generate a model of the AHK5 domain architecture that served as the basis for construct design (Figure 1-2).

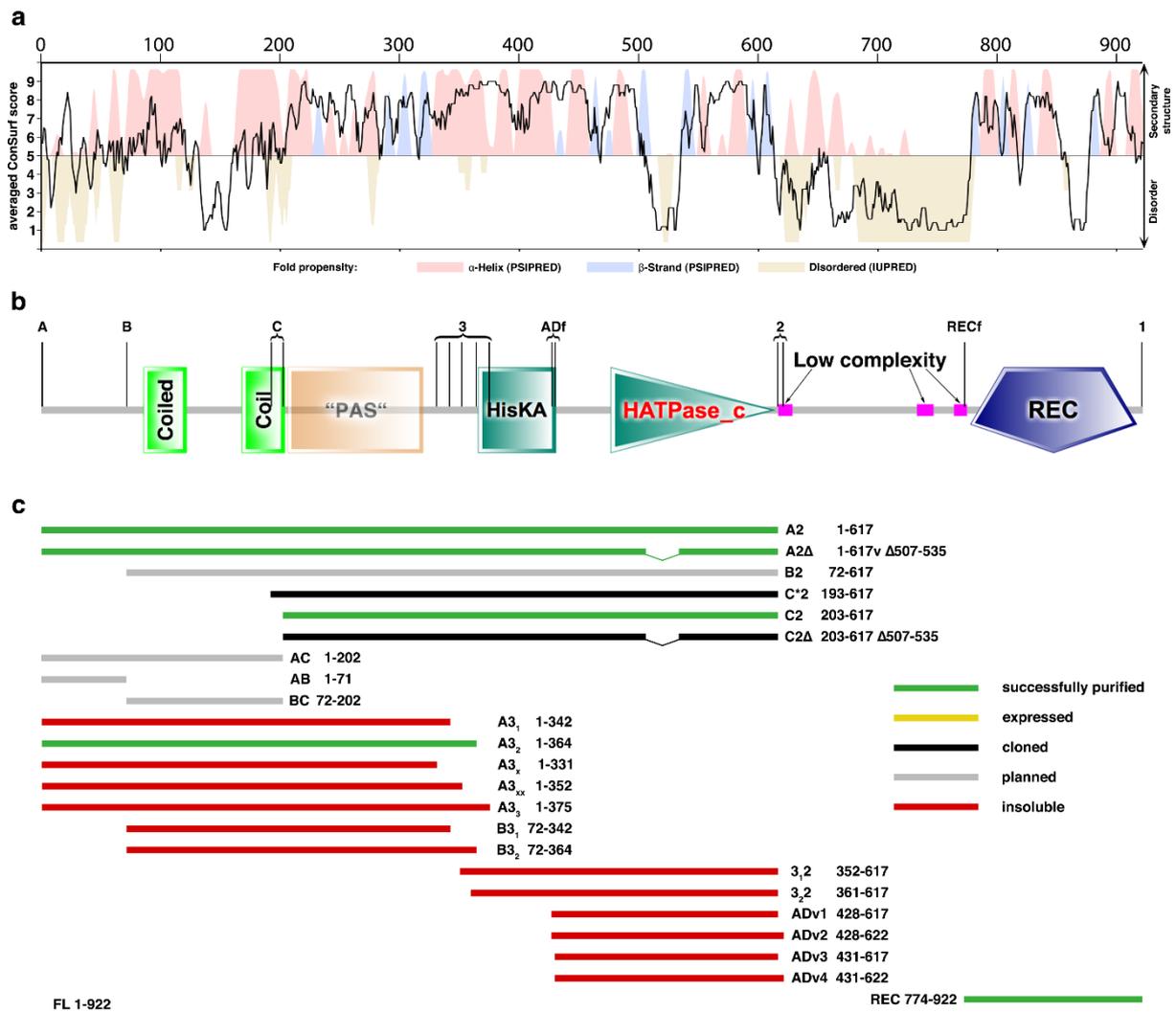


Figure 1-2: Overview of AHK5 domain architecture and expression constructs. **(a)** Overlay of the secondary structure propensity of AHK5, as calculated by PSIPRED [65], [66] and ConSurf conservation score [67]. Predicted α -helical segments are highlighted in red, β -strands in blue. Sections with increased propensity for disorder (IUPRED [68]) are shown by the brown colour. The black line plots the per-residue conservation score obtained with ConSurf [67]. All scores were smoothed using a five-residue sliding window averaging in order to better recognise domain boundaries. The residue numbering axis at the top is valid for the entire figure. **(b)** Domain architecture generated by the SMART protein domain annotation resource [23] (<https://smart.embl.de>). Segments found at high confidence are drawn and labelled as solid boxes. The stretch between residues 200 and 320 yielded multiple low-confidence hits (Figure 1-1b) and is shown as a transparent box. **(c)** Overview of the AHK5 constructs that were generated in the process of this study. The colour code represents the status of the fragments (green: successfully purified; yellow: expressed, solubility test pending; black: cloned, expression test pending; grey: to be cloned; red: expressed but insoluble).

Secondary structure and disorder analysis using Quick2D (Figure 1-2a) yielded a significant disorder propensity for the first 50 - 70 residues followed by two strongly α -helical stretches around residues 100 and 200, respectively. These stretches coincide with two separate coiled coil stretches identified by the SMART domain annotation database [23] (Figure 1-2). The segment between residues 200 and the beginning of the kinase domain around residue 350 displays a mixed α/β scoring that is compatible with the hypothesis of a PAS or LOV domain (see also Figure 1-20). Further along the sequence, the secondary structure shows the long helical section typical of the core region of histidine kinases. This region is sometimes termed dimerisation and histidine phosphotransfer domain (DHpD) [24]. Here it will be termed Histidine Kinase domain (HisKA). It consists of two antiparallel α -helices that form a four-helix bundle upon dimerisation. The reactive histidine residue is located at the beginning of the

first helix. It is followed by the ATPase domain (mixed α/β), a long disordered region, and finally the previously described receiver domain [40].

An overlay of a sequence conservation score (obtained using the ConSurf server [67]) (Figure 1-2a) reveals high sequence conservation for the HisKA / ATPase and receiver domains, as well as for the putative PAS / LOV domain. In contrast, the first 150 residues yield an average ConSurf score without clear indication as to conservation or variability. The disordered region between the ATPase and the receiver domains scores as highly variable, as does a section of roughly 20 residues within the ATPase domain.

Based on these data, domain boundaries for the input segment were selected at the very N-terminus ("A"), before the first coiled coil ("B"), as well as before the PAS/LOV domain ("C"), and at the start of the kinase domain ("3") (Figure 1-2b). Additional boundaries were chosen between the histidine kinase and ATPase domains ("ADf"), at the end of the ATPase domain ("2"), and at the very C-terminus ("1"). The N-terminus of the receiver domain ("RECF") was published previously [40]. Figure 1-2c shows all constructs that were designed, cloned, expressed, or purified during this study. All constructs were named according to their boundaries (e.g. A2 for residues 1-617). For some domain boundaries multiple versions were generated in the process of construct optimisation (e.g. A3₁ and A3₂)

1.2.2 Protein expression and solubility screening

For all *E. coli* constructs, protein expression was assessed in small scale by SDS-PAGE, and protein solubility was probed by screening for potential lysis buffers (see Material and Methods Table 3-27).

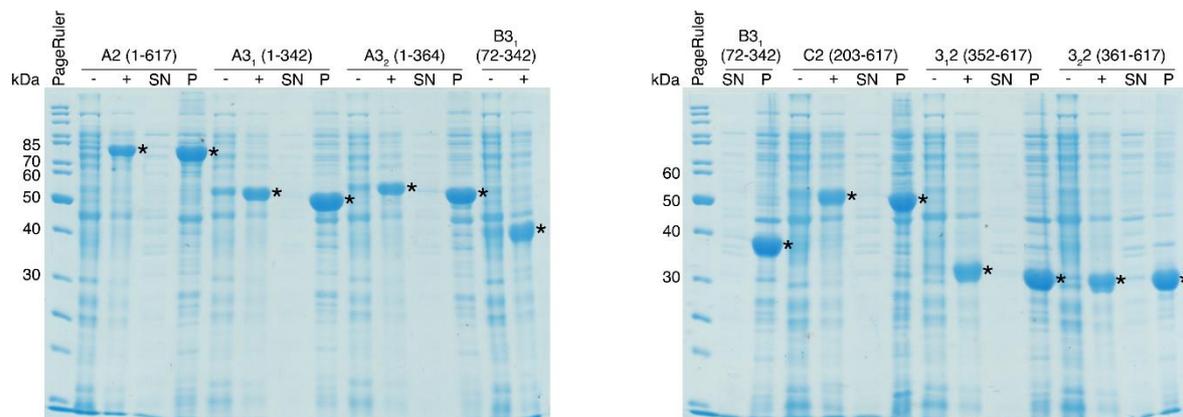


Figure 1-3: Expression test for the first seven AHK5 constructs (15 h, 20 °C, 0.5 mM IPTG). Lanes are grouped by construct. (-) Pre-induction control. (+) Whole cell pellet post induction. (SN) Supernatant and (P) pellet of preliminary lysis test in 30 mM Tris pH 7.5 @ RT, 300 mM NaCl, 20 mM imidazole. The pellet fraction was solubilised with 6 M Urea. Asterisks indicate overexpression bands of the expected sizes. The original gel pictures without labelling asterisks can be found in **Figure 5-2** in the appendix.

All tested constructs expressed very efficiently and to comparable levels (Figure 1-3). However, in all instances the overexpressed protein localised to the insoluble fraction (P). For most constructs, subsequent lysis buffer screening did not yield any conditions in which soluble protein could be natively extracted. Only one construct, A3₂, which contains the complete N-terminal segment of AHK5 preceding the kinase domain (residues 1-364), could be solubilised directly after lysis (Figure 1-4a). The

buffer identified (1-7, Material and Methods Table 3-27, [69]) to be favourable for extracting soluble protein, was the only one containing a chaotropic agent (100 mM urea).

Expression conditions were then refined by analysing the effect of induction time on the distribution of AHK5 A3₂ between the soluble and insoluble fractions (Figure 1-4b). The highest quantity of target protein in the soluble fraction was achieved after 5 h expression time. Longer expression led to sequestration of AHK5 to the pellet.

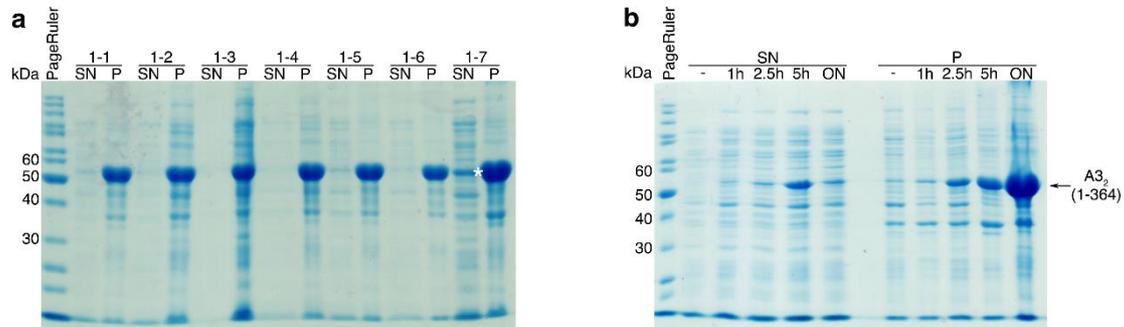


Figure 1-4: Identification of a suitable lysis buffer for the AHK5 A3₂ construct. (a) Excerpt of a lysis buffer screen conducted with O/N expression pellets and the buffer screen (Table 3-27). Lysis performed in buffer 1-7 (100 mM Tris pH 8.2 @ RT, 50 mM NaCl, 100 mM urea) yielded a band of the expected size in the soluble fraction (SN, white asterisk). The original gel picture without the labelling asterisk can be found in Figure 5-3 in the appendix. (b) Time course of expression of A3₂ and distribution into soluble (SN) and insoluble (P) fractions after lysis in buffer 1-7. The highest intensity band at the expected size in the soluble fraction is detectable after 5 h.

As protein fragments containing not only the input segment but also the kinase domain are likely required to understand the mode of AHK5 activation, two constructs were selected for purification attempts in spite of lack of soluble protein in initial lysis buffer screens. Following lysis and clarification, pellets were repeatedly washed, and denaturing solubilisation was attempted in the presence of 8 M urea. Both the C2 (residues 203-617, Figure 1-5a) and A2 (residues 1-617, Figure 1-5b) fragments were thus successfully solubilised.

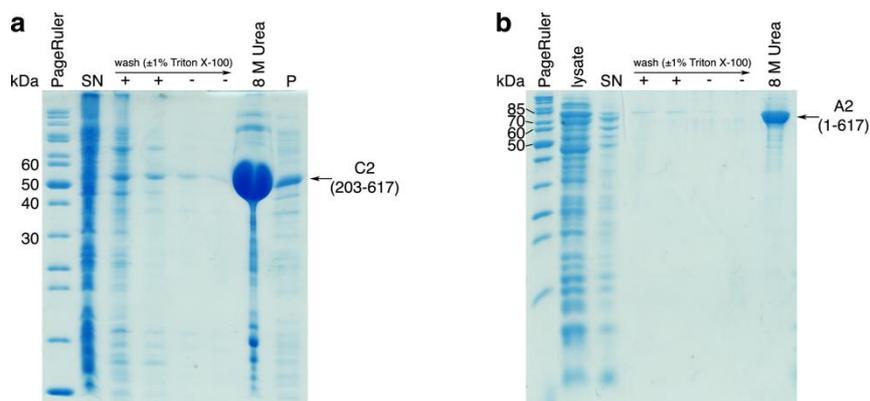


Figure 1-5: Denaturing solubilisation tests of AHK5 fragments comprising the kinase and ATPase domains. (a) C2 (residues 203-617) and (b) A2 (residues 1-617) constructs: Lysis, supernatant (SN), pellet wash fractions with (+) and without (-) 1% Triton X-100, solubilisation with 8 M urea, and remaining pellet (P) fractions were analysed, if present.

1.2.3 Native purification of fragment A3₂ (1-364)

The AHK5 fragment A3₂ was enriched by IMAC and large aggregates were removed by SEC. The affinity tag was then proteolytically removed and a second SEC run was used for final polishing.

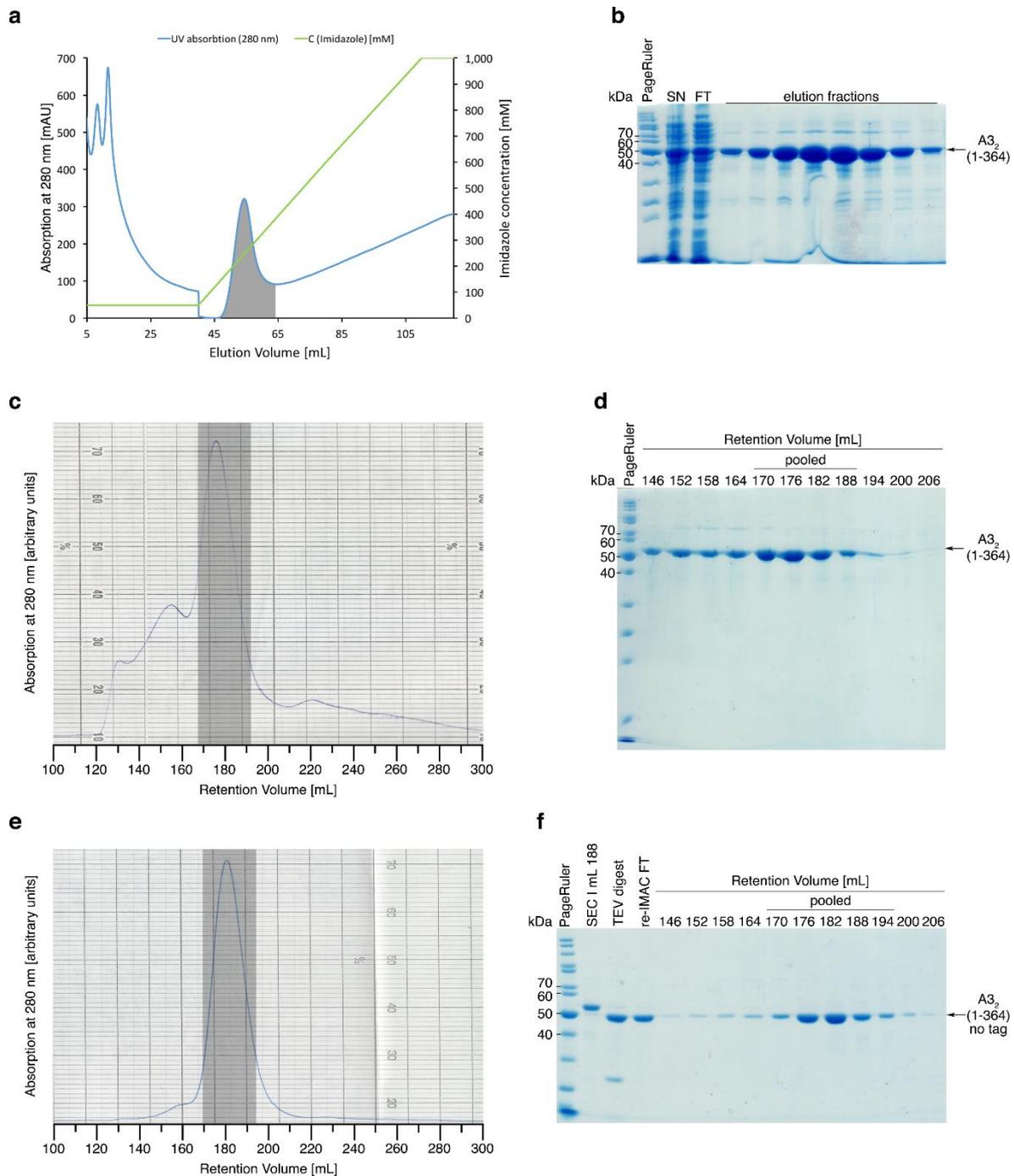


Figure 1-6: Overview of the purification workflow for AHK5 fragment A3₂. **(a)** Enrichment of AHK5 from crude lysate by IMAC. Chromatogram of column wash and elution steps. The imidazole concentration is shown in green; the blue line denotes UV absorption at 280 nm. Target protein-containing fractions pooled for further purification are highlighted in grey. **(b)** SDS-PAGE of lysis supernatant (SN), IMAC flowthrough (FT), and elution fractions. **(c)** Chromatogram of first gel filtration (SEC I), using a HiLoad 26/60 column containing Superdex 200 (GE Healthcare, Uppsala, Sweden). The UV absorption trace was recorded at 280 nm. The pooled peak fractions are marked in grey. **(d)** SDS-PAGE analysis of relevant SEC fractions. **(e)** UV absorption trace of polishing SEC (SEC II) elution recorded at 280 nm with pooled fractions marked in grey. **(f)** SDS-PAGE analysis of affinity tag removal by proteolytic cleavage and subsequent relevant SEC II fractions. Lanes contain pre-cleavage reference (SEC I mL 186), digestion with TEV protease (+TEV), protease removal by flowing the digest over an IMAC column (re-IMAC FT), and the relevant SEC II fractions.

1.2.3.1 Protein enrichment by IMAC

The overexpressed protein could successfully be enriched by IMAC, as is visible by the very prominent bands at 50 kDa in the SDS-PAGE analysis of the elution fractions (Figure 1-6a, right). Only a small portion of the available protein successfully bound to the Ni-NTA matrix, as suggested by the comparable intensities of the cognate bands in lysis supernatant and IMAC flowthrough (lanes L and FT, respectively). This could be caused either by saturation of the column material or by the binding kinetics of AHK5. Indeed, a single pass of the lysate over the column was not sufficient to bind any significant amount of protein. Continuous circular loading O/N was required. Elution of AHK5 A3₂ was successfully carried out with a gradient of imidazole, ranging from 0 to 1 M imidazole. Target protein elution started at concentrations around 150 mM (Figure 1-6a, left). However, in many cases, AHK5 A3₂ could be detected in the eluate over a wide range of imidazole concentrations, including 1 M. This suggests indeed slow association and dissociation kinetics to and from the Ni-NTA matrix as a contributing factor to the observed behaviour of this AHK5 fragment.

1.2.3.2 Aggregate and high-order oligomer removal by SEC

Following removal of imidazole and urea by overnight dialysis against buffer SEC I (see Material and Methods Table 3-23), the protein was subjected to a first round of gel filtration in order to remove aggregation nuclei. When using unconcentrated sample immediately after dialysis, SEC chromatograms as shown in Figure 1-6c could reproducibly be obtained. A trace of the UV absorption at 280 nm revealed three discernible elution events: A first one after 130 mL, corresponding to the column's void volume and thus containing particles with molecular weights of over 600 kDa or aggregate, a second one with a broad distribution and a peak at approximately 155 mL, and a third and more narrow one at an elution volume of 177 mL. As is apparent from SDS-PAGE analysis, the second and third event both contain predominantly a protein of the expected size (Figure 1-6d). Based on the comparison with elution volumes observed for other proteins available in the laboratory, neither of the observed discrete elution events for AHK5 A3₂ is consistent with the expected dimeric assembly (expected particle size of roughly 100 kDa). Such a particle should elute approximately after 200 mL. Both elution events thus likely arose from higher molecular weight protein species, hence suggesting that overexpressed AHK5 A3₂ – specifically or unspecifically – forms larger oligomeric assemblies.

1.2.3.3 Affinity tag removal and polishing SEC

To complete the purification of AHK5 A3₂ only the protein species eluting at 177 mL from the first SEC was further processed. First the His₆ affinity tag was proteolytically removed and the remaining protein concentrated and subjected to a second round of gel filtration in order to maximise sample homogeneity (Figure 1-6e). Tag removal by TEV protease was highly efficient, as can be seen by the comparison of the first two lanes on the SDS-PAGE in Figure 1-6f. Owing to the His₆ tag of the TEV protease, it could in turn be removed from the solution together with undigested AHK5 A3₂ by flowing the sample over a second Ni-NTA-agarose-column, as only the digested target protein is detectable in the flowthrough (Figure 1-6f, lane re-IMAC FT). Concentration and subsequent polishing SEC using the same strategy as in the first gel filtration step reproducibly yielded a single peak at 177 mL. The small

shoulder preceding it could either be attributed to a possible oligomerisation equilibrium, or to carryover from SEC I, as the lack of baseline separation of the two species makes precise pooling of the fractions very difficult.

Overall, AHK5 A3₂ could successfully and reproducibly be purified to average yields of 4.3 mg of pure protein per litre of TB culture processed. Interestingly, the C3A mutant, which was purified according to the same protocol, showed a nearly six-fold increase in yield. Both versions of the protein can be stored at -80 °C at concentration of 2-4 mg/mL.

1.2.4 Denaturing purification of fragment C2 (203-617)

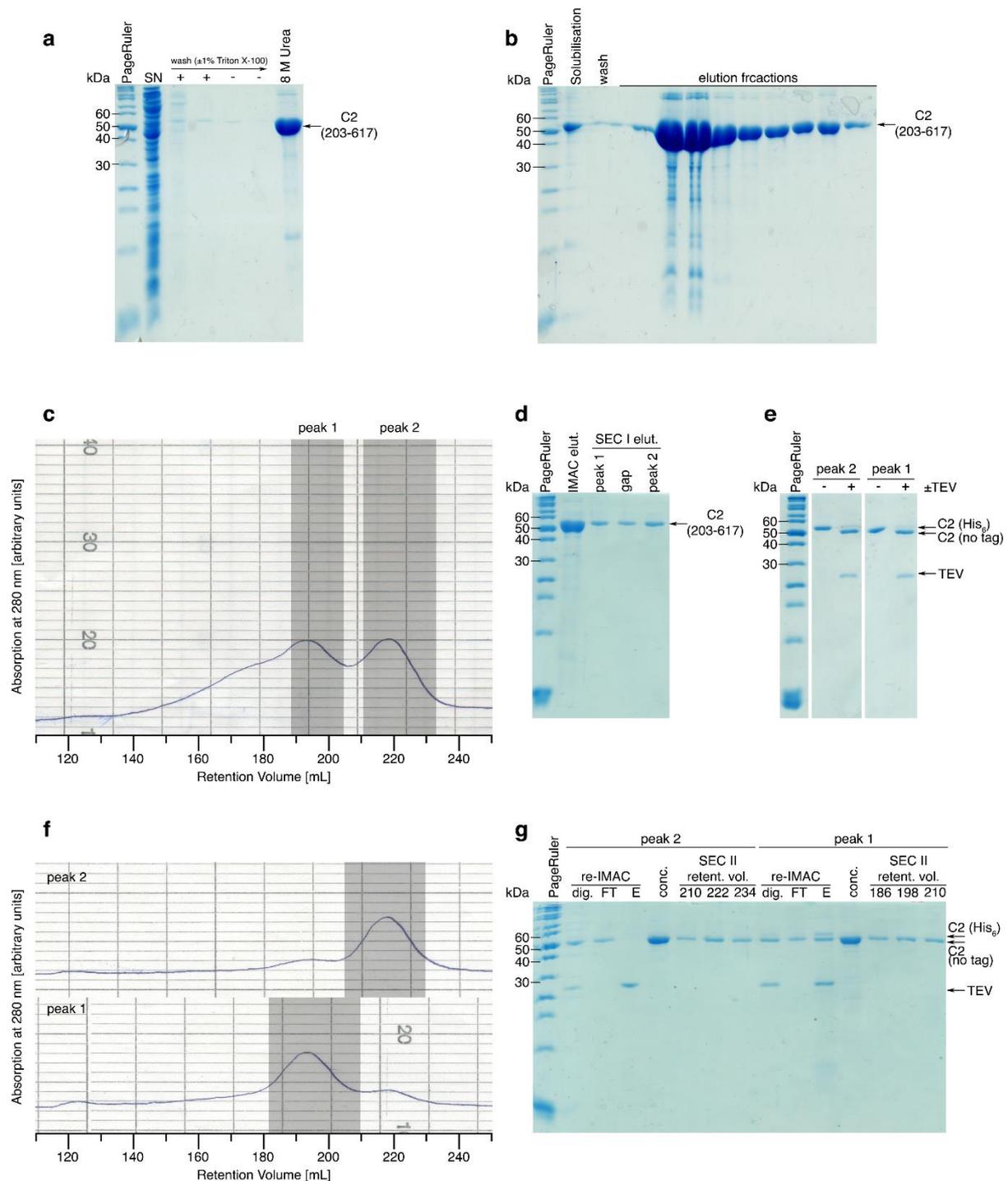


Figure 1-7: Overview of the purification workflow for AHK5 fragment C2. **(a)** SDS-PAGE of Urea-solubilisation of the insoluble fraction following lysis. Lane SN contains the lysis supernatant, followed by four wash steps with (+) and without (-) Triton X-100 and the final inclusion body fraction solubilised with 8 M urea. **(b)** Target protein enrichment by denaturing IMAC visualised by SDS-PAGE. Solubilised inclusion bodies are followed by the IMAC wash and the elution fractions. **(c)** Trace of UV absorption at 280 nm of first SEC run after refolding. Grey areas represent separately pooled peak fractions. **(d)** SDS-PAGE of (c) with pooled fractions of peak 1 and peak 2, as well as one fraction separating the peaks (gap). IMAC elution prior to refolding is included for reference. **(e)** SDS-PAGE of TEV protease-mediated affinity tag cleavage. The uncropped gel image is shown in **Figure 5-4** in the appendix. **(f)** UV absorption traces (280 nm) of SEC II (peak 2 and peak 1 separately). Grey areas denote pooled fractions. **(g)** SDS-PAGE analysing reverse IMAC (re-IMAC) after TEV digest showing target protein in flowthrough (FT) lane, followed by relevant peak fractions from SEC II. SEC I peak 2 and peak 1 were treated equally and separately.

AHK5 C2 was solubilised from inclusion bodies using a buffer containing 8 M urea. The denatured protein was then subjected to target protein enrichment by IMAC, followed by refolding via stepwise dialysis. Protein homogeneity was then improved by two iterations of SEC and removal of the His₆ purification tag (Figure 1-7).

1.2.4.1 Solubilisation with 8 M Urea and denaturing IMAC

Whole-cell SDS-PAGE analysis of the protein expression cultures 20 h post-induction show a distinct band of just over 50 kDa ($M_{w(C2)} = 49.6$ kDa) signalling high-yield overexpression of the target protein (Figure 1-7a). In contrast, after lysis in buffer C2 main (see Material and Methods Table 3-23) and centrifugation, no band of that size is detectable (Figure 1-7a, lane SN). Indeed, the target protein nearly quantitatively localised to the pellet, and no amount of washing with or without Triton X-100 could solubilise it. Only when the pellet was homogenised in buffer containing 8 M urea, the corresponding band could be detected on SDS-PAGE (Figure 1-7a, lane 8 M urea). The high intensity of the band as well as the near complete lack of other bands suggest a very high yield of solubilised AHK5 C2, as well as an extremely high initial purity. The protein was further enriched by IMAC while under denaturing conditions. Most of the target protein successfully bound to the Ni-NTA matrix and could again be eluted with solubilisation buffer containing 250 mM imidazole (Figure 1-7b). The resulting eluate contained a high concentration of denatured AHK5 C2, as suggested by SDS-PAGE.

1.2.4.2 Refolding and first SEC

Refolding was carried out by dialysing the diluted (1 : 4) protein solution against buffers containing decreasing concentrations of urea (see Material and Methods Table 3-23). This procedure was immediately followed by a first iteration of gel filtration, in order to remove aggregation and aggregation nuclei. The SEC chromatogram (Figure 1-7c) shows two discrete elution events at retention volumes of 193 and 218 mL, respectively. These two events are preceded by a continuous signal, starting at the column's void volume of 120 mL and growing towards the 193 mL peak. SDS-PAGE analysis of fractions covering the entire elution range confirmed the presence of only the target protein band in all of the elution events (Figure 1-8c). This suggests the presence of at least two different discrete oligomeric species of AHK5 C2, in addition to a continuum of unspecific multimers, up to large aggregates that elute in the column's void volume.

Fractions pertaining to either of the two discrete elution events were pooled separately for further processing and termed peak 1 (193 mL), and peak 2 (218 mL). Size estimation of the two species, as carried out by analytical SEC (Figure 1-14b), suggests the smaller species to be monomeric AHK5 C2, while the larger one to be a dimer.

A preliminary attempt to refold AHK5 C2 in the presence of 5 mM Riboflavin was performed by Annkathrin Scheck during her bachelor's studies in order to assess whether a Flavin-binding PAS domain is present. No selective localisation of the characteristic yellow colour to the proteinaceous fractions was observed. However, this experiment was performed using the preliminary rapid dilution protocol instead of the later, more efficient dialysis strategy.

1.2.4.3 Aggregation depends on refolding strategy and incubation time

In an early version of the protocol, rapid dilution was used for refolding instead of dialysis. This led to very large volumes that required a lengthy concentration procedure via a second IMAC. The resulting protein displayed a radically different behaviour on preparative SEC, with most of the protein localising to the column's void volume suggesting nearly quantitative aggregation (Figure 1-8a+c). These experiments, however, confirmed the high degree of purity of the target protein, as SDS samples from across the entire elution range contained a unique single band, compatible with the molecular weight of AHK5 C2 (Figure 1-8c). Interestingly, a similar effect was observed when incubating dialysis-refolded protein at 4 °C O/N prior to SEC (Figure 1-8b), suggesting the timely removal of aggregation nuclei to be a critical step.

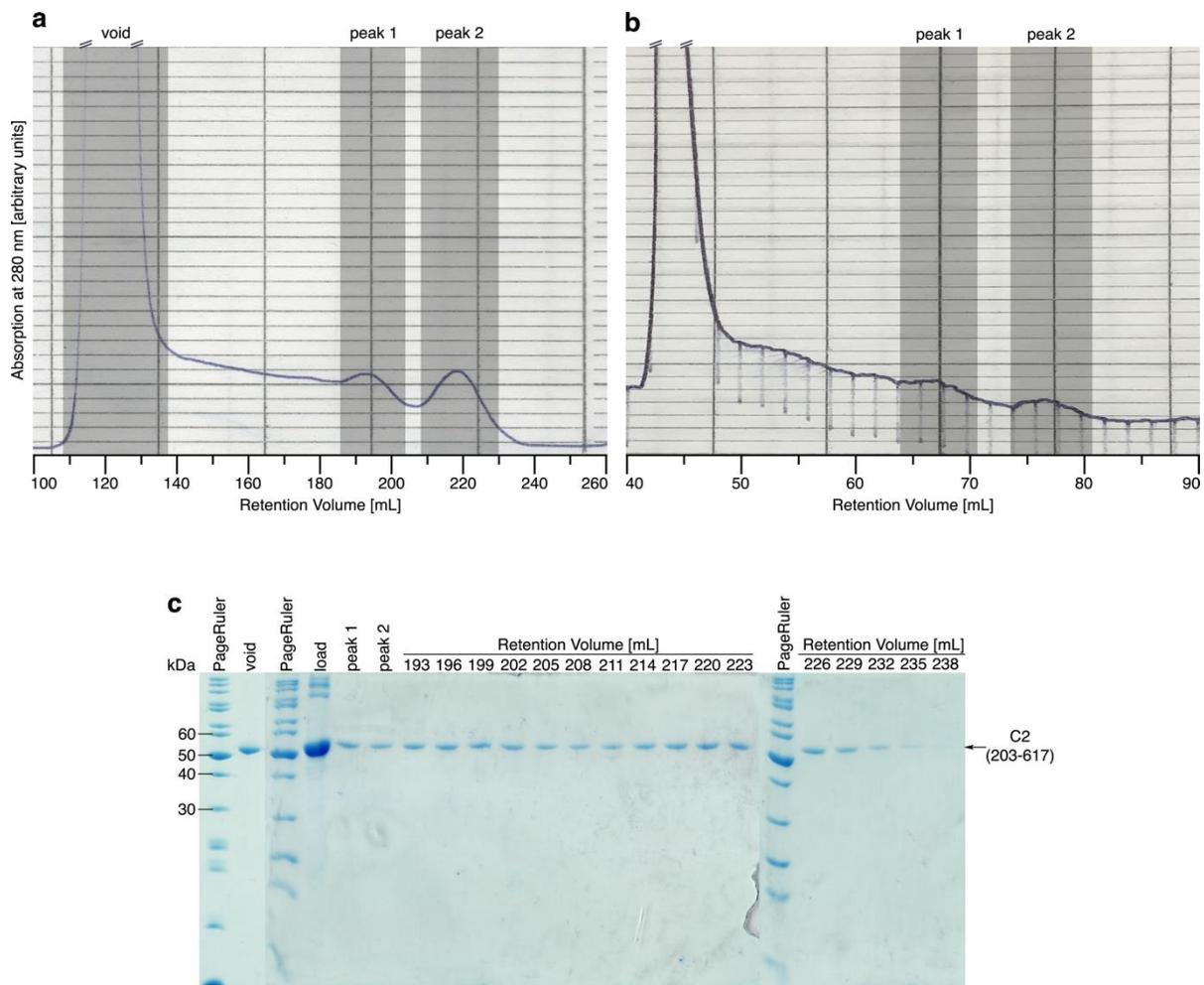


Figure 1-8: Near-quantitative aggregation of AHK5 C2 in previous refolding strategies. **(a)** SEC I trace (UV absorption at 280 nm) of protein refolded by 1:100 rapid dilution and subsequent second IMAC for concentration. The vast majority of the signal localises to the column's void volume. Grey areas signal samples analysed by SDS-PAGE. **(b)** SEC I elution profile obtained from a sample refolded by dialysis but stored O/N at 4 °C prior to SEC yielding a similar elution profile to (a). This experiment was carried out on a smaller 16/60 HiLoad SD200 column. **(c)** SDS-PAGE analysis of (a) showing only one species of protein across the entire chromatogram. The IMAC I elution (load) was included for reference. Fractions obtained from (b) could not be detected on SDS-PAGE, due to low concentration.

1.2.4.4 Affinity tag removal and polishing SEC

His₆ tag removal by digestion with TEV protease and subsequent reverse IMAC yielded results comparable to those obtained for A3₂. The digestion was nearly quantitative. No band of the size of

the original protein could be detected on SDS-PAGE for the peak 2 sample and only a very small fraction of protein remained undigested in the peak 1 sample (Figure 1-7e). In both cases, the digested protein could successfully be separated from its undigested precursor, as well as from the protease, by means of reverse IMAC. Both the undigested protein and the TEV protease bound to the Ni-NTA column and are visible in the subsequent imidazole-containing elution step, while the digested protein was directly collected in the flowthrough fraction (Figure 1-7g).

Polishing SEC performed with the concentrated C2 peak 2 and peak 1 samples obtained after TEV digestion and reverse IMAC yielded appreciably low aggregation at 120 mL retention volume and much less high order oligomers than in the first SEC. Both elution profiles show a distinct shoulder pertaining to the respective other species (Figure 1-7f). This could either be due to carryover from SEC I, as the peaks are not baseline-separated and thus hard to delimit, or be indicative of a dynamic equilibrium in solution.

Successful and reproducible, purification of AHK5 C2 yielded approximately 2 mg of monomeric C2 ("peak 2") and 1 mg of dimeric C2 ("peak 1") per gram of cell pellet wet weight. The samples could be stored at -80 °C at concentrations between 1 and 2.5 mg/mL.

1.2.5 Denaturing purification of fragment A2 (1-617)

The AHK5 fragment A2 was purified following the same general strategy used for C2. The protein was solubilised in 8 M urea, then enriched by IMAC under denaturing conditions. Following re-folding by dialysis, AHK5 A2 was subjected to preparative SEC, His₆ tag cleavage with TEV protease and a second round of SEC.

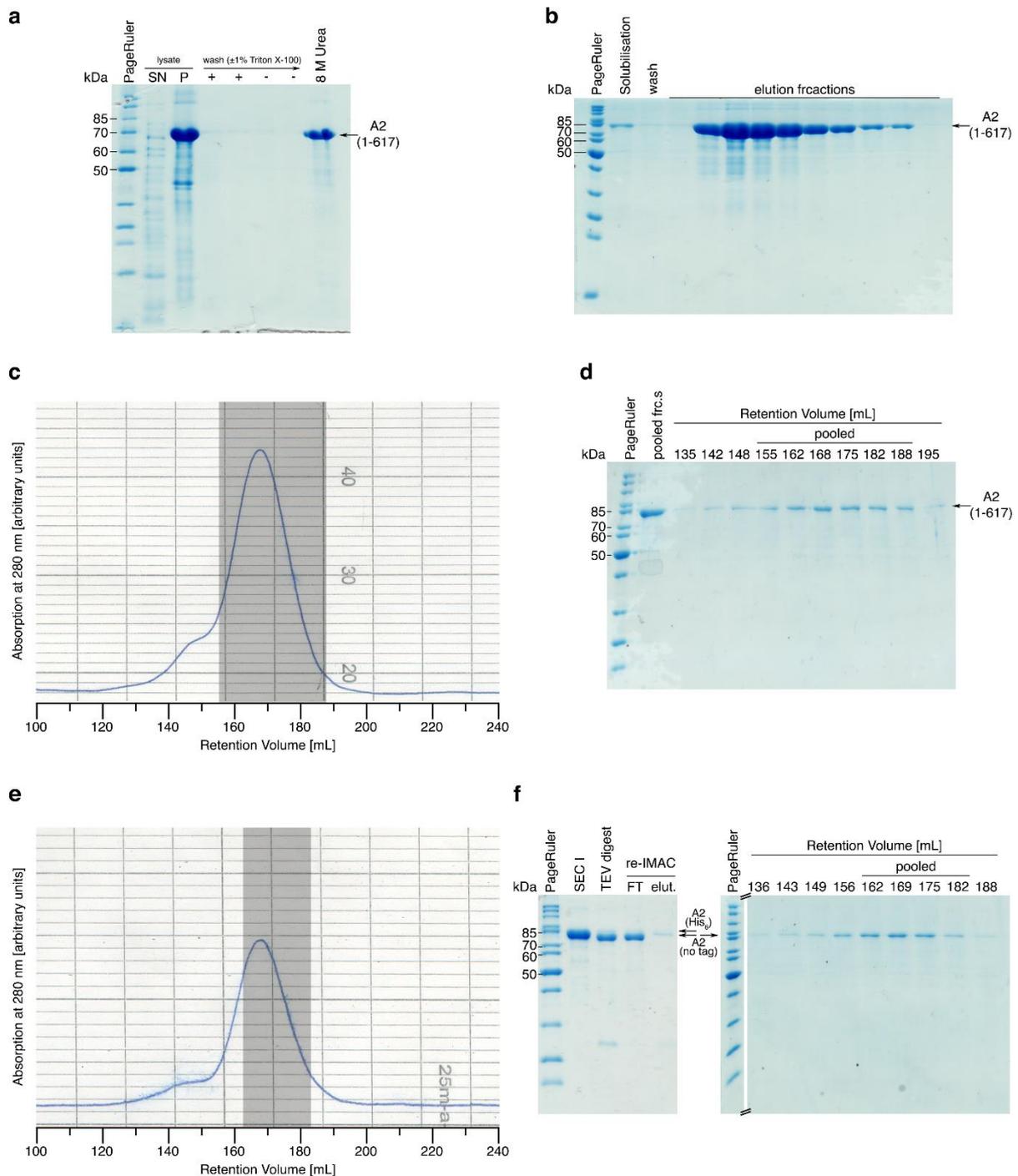


Figure 1-9: Overview of AHK5 A2 purification. **(a)** SDS-PAGE visualising washing and subsequent solubilisation of lysis pellet (lane P). **(b)** IMAC under denaturing conditions. SDS-PAGE shows solubilised protein prior to application to Ni-NTA agarose, the IMAC wash step, and imidazole elution containing high amounts of target protein. **(c)** SEC I chromatogram (UV absorption at 280 nm). Grey area denotes pooled fractions. **(d)** SDS-PAGE analysis of (c). **(e)** A₂₈₀ chromatogram of polishing SEC. Grey areas denote pooled fractions. **(f)** SDS-PAGE of TEV and reverse IMAC (left) and relevant SEC II fractions (right). The uncropped gel image is shown in **Figure 5-5** in the appendix.

1.2.5.1 Solubilisation and denaturing IMAC

Similarly to fragment C2 (section 1.2.4), inclusion bodies derived from 20 h of AHK5 A2 over-expression in *E. coli* (TB) could readily be solubilised in A2 main buffer (see Material and Methods Table 3-23) containing 8 M urea, as suggested by the presence of a prominent band on SDS-PAGE at a molecular weight of approximately 70 kDa (Figure 1-9a, $M_{w(A2)}=73.27$ kDa). Subsequent IMAC, conducted under denaturing conditions, yielded high quantities of over 95% pure target protein (Figure 1-9b).

1.2.5.2 Refolding and SEC

AHK5 A2 was successfully refolded using the step dialysis strategy originally developed for the C2 fragment (see section 1.2.4.2). Subsequent preparative SEC revealed far less tendency to aggregate compared to C2. The largest portion of the sample eluted with peak intensity at 168 mL and a small shoulder preceding it at 148 mL. Virtually no UV absorption could be detected at the column's void volume of 110 mL. SDS-PAGE confirmed the peak fractions to contain the same protein species as the relevant IMAC elution (Figure 1-9c and d).

Size estimation by analytical SEC (Figure 1-14c) revealed a theoretical molecular weight of 307 kDa for the main peak and 610 kDa for the shoulder. These values are compatible with an elongated dimer of AHK5 A2 and a dimer of dimers, respectively.

1.2.5.3 Affinity tag removal and polishing SEC

Removal of the purification tag by digestion with TEV protease and subsequent reverse IMAC was successful and near quantitative (Figure 1-9f, left panel). Final sample polishing by a second round of SEC yielded the elution profile displayed in Figure 1-9e. The main species elutes again at 168 mL, indicating that purification tag removal, although successful, has no detectable impact on the protein's hydrodynamic radius. The main peak in Figure 1-9e appears to exhibit a slight tailing effect compared to Figure 1-9c, but this effect could be subjectively overjudged due to the different overall intensities of the chromatograms. Integration and subsequent numeric evaluation would be required and is not available for the chromatograms in question. Despite conservative fractioning during the first round of SEC (Figure 1-9c, grey area) aimed at exclusion of the 148 mL shoulder, the overall shape of the elution peak remained virtually identical, suggesting the possible formation of dimers of dimers to be driven by a dynamic equilibrium in solution. SDS-PAGE analysis of the relevant peak fraction (Figure 1-9f, right panel) is again consistent with the size of AHK5 A2 and confirms the protein to be resilient against proteolytic decay over the course of the purification.

Final purification yields of 2-3 mg of pure protein per gram of pellet wet weight were reproducibly achieved and stored at -80 °C at concentrations of 0.5-0.6 mg/mL.

1.2.6 Purified AHK5 fragments are folded and of correct identity

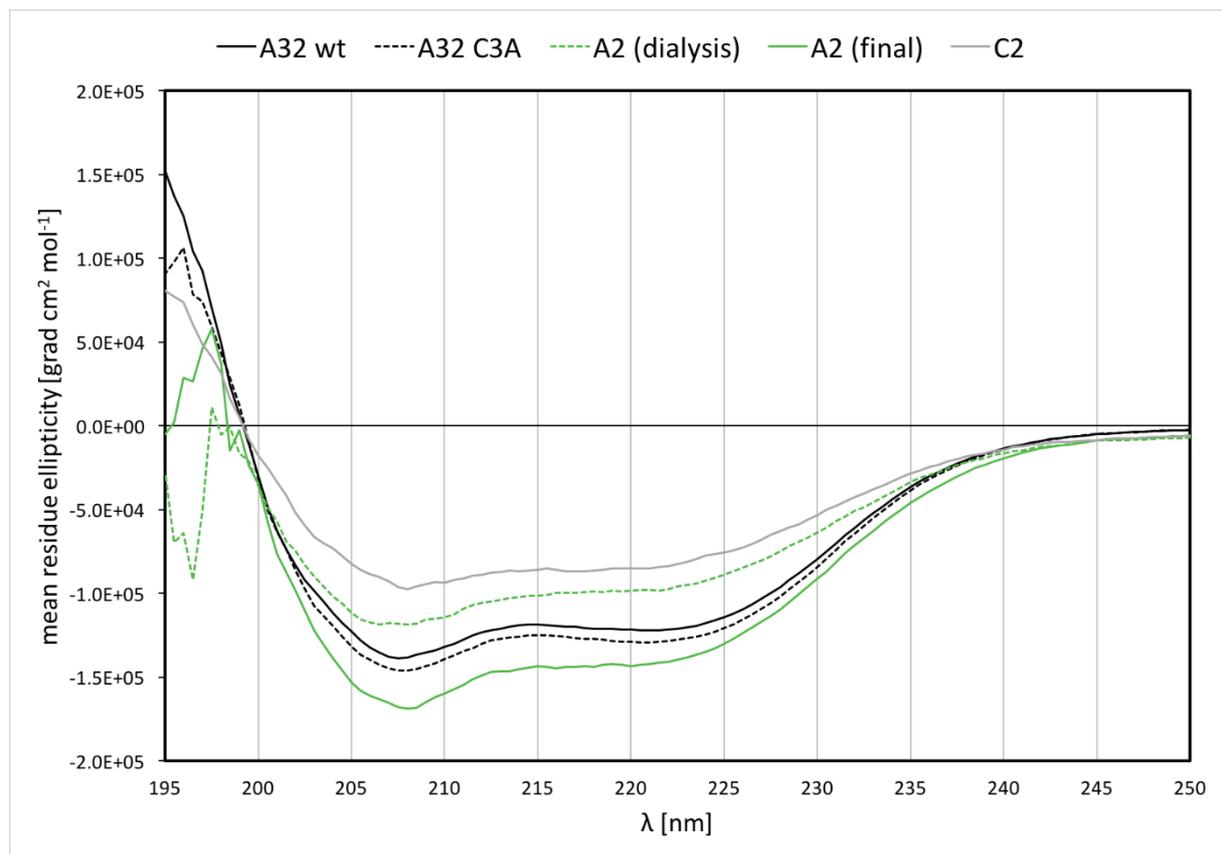


Figure 1-10: CD spectra of purified AHK5 constructs. Spectra were normalised for protein size, concentration, and cuvette thickness. The spectra are colour-coded as denoted on top.

Circular dichroism (CD) spectroscopy was employed to assess protein fold after purification. All three successfully purified AHK5 fragments yielded spectra typical for mixed α/β folds (Figure 1-10), based on the characteristic minima at 207 nm and 222 nm for α -helices and at 217 nm for β -strands. Although helices generally dominate CD spectra, the higher negative value of the 207 nm minimum relative to the 222 nm minimum is strongly indicative of both an increased β -sheet content and, possibly, a significant random coil fraction. This phenomenon applies mainly to the two fragments purified using denaturing agents, namely C2 (grey) and A2 (green, solid and dashed). By comparison, A3₂ has a slightly less pronounced intensity difference for the two minima (black solid line, wt, black dashed line, C3A mutant). A2 and A3₂ display comparable per-residue signal (mean residue ellipticity) suggesting similar overall folding compactness. In contrast, C2 shows a weaker signal.

In order to ascertain the compactness of the proteins' fold, a protease exclusion assay was performed using the ProtiAce kit (Hampton Research, Aliso Viejo, CA, USA). SDS-PAGE-separated protein fragments were excised from the gel and sent to Sophie Stotz, Kalbacher Lab, University of Tübingen for MALDI-TOF analysis. Undigested samples were included for protein identity verification.

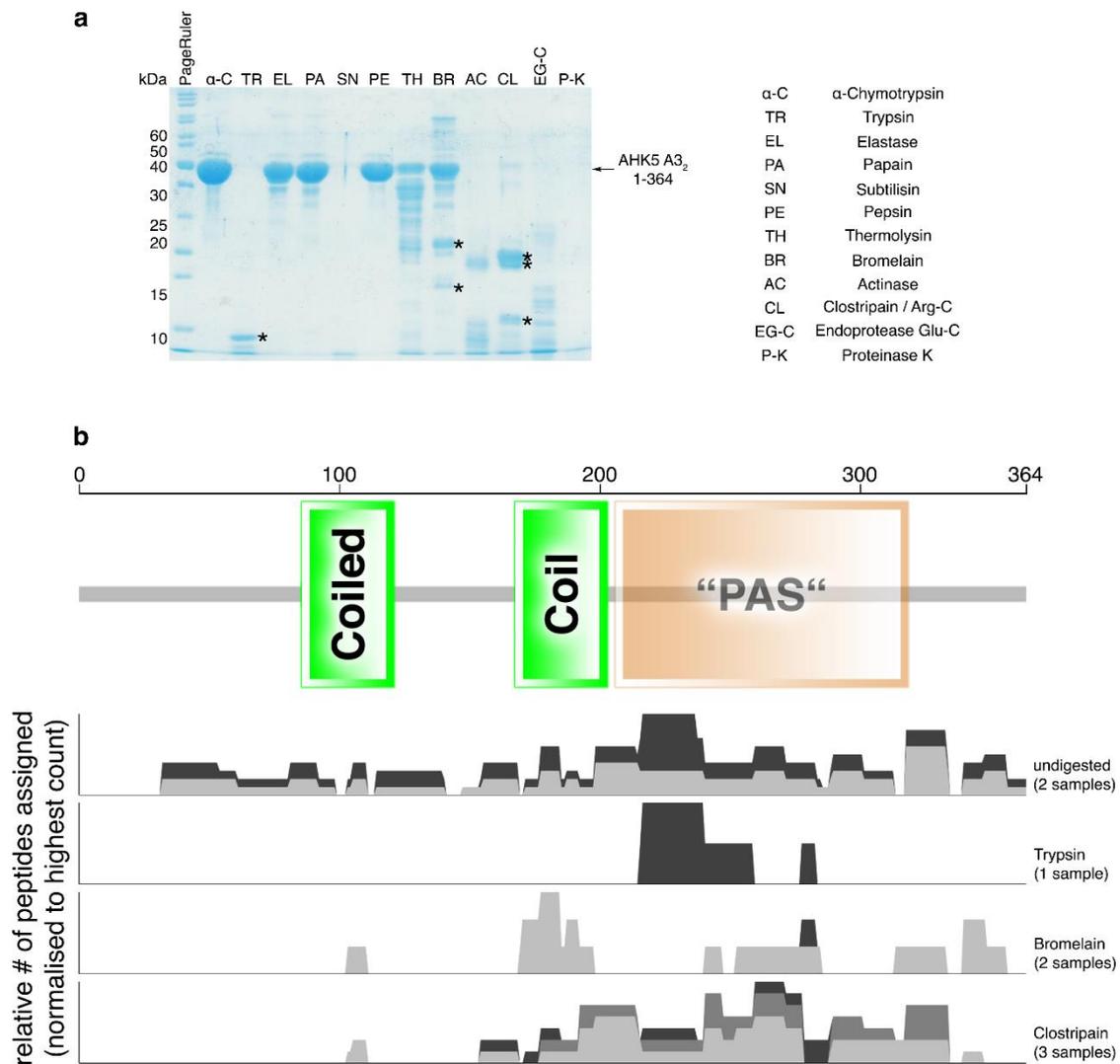


Figure 1-11: Protease exclusion assay performed on AHK5 fragment A3₂. **(a)** SDS-PAGE of single-protease digests. Lanes are labelled according to the respective protease. Abbreviations are expanded on the right. Bands subjected to MALDI-TOF are marked with an asterisk. The original gel picture without labelling asterisks can be found in **Figure 5-6** in the appendix. **(b)** Peptide assignments mapped onto the putative A3₂ domain architecture (see **Figure 1-2**). Each sequence position was assigned a value equal to the number of peptides covering that position as detected by MALDI-TOF. The thus generated sequence coverage histogram was then normalised to the respective highest count for better comparability. Greyscale differences delimit sub-histograms derived from individual gel bands, if more than one was analysed for one digest.

AHK5 A3₂ appears to be resistant against proteolytic cleavage to a certain extent (Figure 1-11a). Out of the twelve proteases in the kit only three led to complete or near complete degradation, namely Proteinase K, Subtilisin, and Trypsin. A number of defined degradation products were detected for Thermolysin, Actinase, Clostripain, and Endoprotease Glu-C. The remaining four, including the high-efficiency-low-specificity digestive enzyme α-Chymotrypsin, showed no or nearly no degradative activity. Of the discrete fragments obtained, six were successfully subjected to mass spectrometry analysis by MALDI-TOF.

Due to imperfect separation of the fragments by SDS-PAGE, precise sequence boundary assignment is not advisable. Hence, a normalised cumulative representation of peptides automatically assigned to spectral peaks (BioTools Version 2.2 software (Bruker Daltonik GmbH, Bremen, Germany)) was chosen, resulting in sequence coverage histograms (Figure 1-11b).

Processing of the undigested sample revealed nearly complete sequence coverage, confirming protein identity (Figure 1-11b, top histogram). Analysis of the digested samples show the highest number of peptide reads clustering towards the region between residues 170 and 320 (Figure 1-11b). This segment is thought to contain a coiled coil stretch as well as a PAS or LOV domain (or a similar domain, Figure 1-2b). Cumulatively, MALDI-TOF results suggest indeed the presence of one or more folded domains in that region.

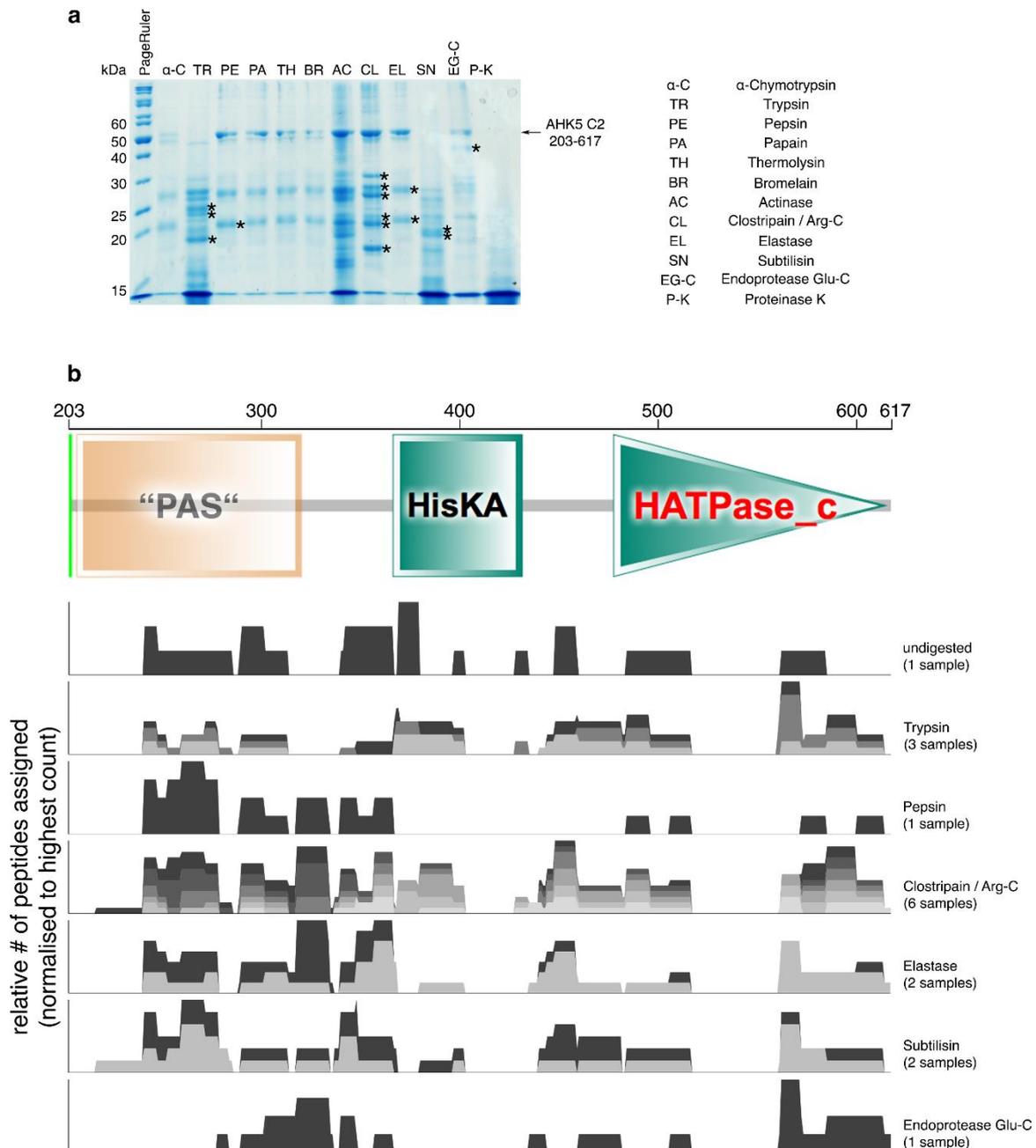


Figure 1-12: Protease exclusion assay performed on AHK5 fragment C2. **(a)** SDS-PAGE of single-protease digests. Lanes are labelled according to the respective protease. Abbreviations are expanded on the right. Bands subjected to MALDI-TOF are marked with an asterisk. The original gel picture without labelling asterisks can be found in **Figure 5-7** in the appendix. **(b)** Peptide assignments mapped onto the putative C2 domain architecture (see **Figure 1-2**). Each sequence position was assigned a value equal to the number of peptides covering that position as detected by MALDI-TOF. The thus generated sequence coverage histogram was then normalised to the respective highest count for better comparability. Greyscale differences delimit sub-histograms derived from individual gel bands, if more than one was analysed for one digest.

In contrast to A3₂, the C2 fragment displays diminished stability against proteolytic digestion, as some extent of degradation is visible in all digestion samples (Figure 1-12a). This, however, led to a higher number of degradation products suitable for MALDI-TOF fingerprinting. A total of 15 fragments were selected for analysis (Figure 1-12a, black arrows).

As before, protein identity was confirmed by the spectral fingerprint of an undigested sample (Figure 1-12b, top histogram), although at lower sequence coverage of just under 60%. Sequence coverage histograms obtained from proteolytic digests indicate the presence of folded domains across the entire sequence of AHK5 C2, with a particular emphasis on the putative PAS / LOV domain (approx. residues 207-320). The section connecting this domain to the actual histidine kinase (HisKA) domain is also strongly represented in the histograms. This is of particular interest, as the segment had previously not been classified as pertaining to any domain. Further regions of high sequence coverage are located to the linker region between the HisKA domain and the subsequent ATPase, as well as within the ATPase. The helix-turn-helix fold located centrally in the HisKA is largely missing, as well as a large portion within the ATPase.

Although SDS-PAGE analysis of the proteolytic digests suggests diminished stability of C2 in comparison to A3₂, mass spectrometry revealed protected (and thus probably folded) sequences to exist across the entire fragment length.

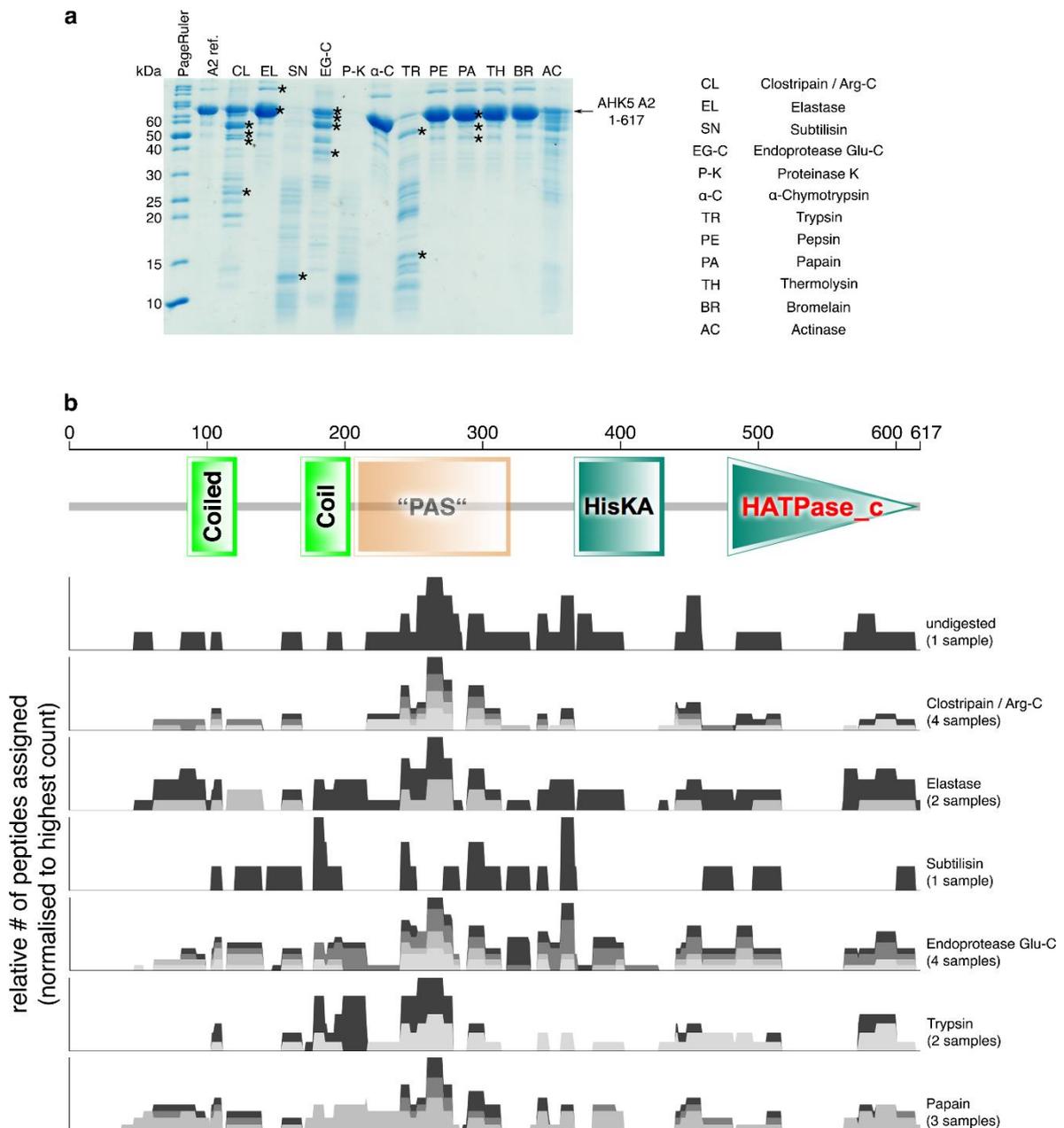


Figure 1-13: Protease exclusion assay performed on AHK5-A2. **(a)** SDS-PAGE of single-protease digests. Lanes are labelled according to the respective protease. Abbreviations are expanded on the right. Bands subjected to MALDI-TOF are marked with an asterisk. The original gel picture without labelling asterisks can be found in **Figure 5-8** in the appendix. **(b)** Peptide assignments mapped onto the putative A2 domain architecture (see **Figure 1-2**). Each sequence position was assigned a value equal to the number of peptides covering that position as detected by MALDI-TOF. The thus generated sequence coverage histogram was then normalised to the respective highest count for better comparability. Greyscale differences delimit sub-histograms derived from individual gel bands, if more than one was analysed for one digest.

Of the three AHK5 fragments purified to date, A2 is the most resistant to proteolytic degradation. In nine out of 12 digests, the full length fragment was the major species detected on SDS-PAGE following incubation, suggesting a remarkable stability (Figure 1-13a). A total of 15 bands were excised from the PAGE and sent for MALDI-TOF. Due to the high stability of the protein, most of the fragments fingerprinted by mass spectrometry were very close or at the full size of A2. The thus resulting extensive sequence coverage (Figure 1-13b) is not surprising, but strongly confirms protein identity. Two particular histograms are noteworthy in terms of fragment separation by SDS-PAGE and the resulting accuracy in domain boundary assignment achievable by this method. Both the single read from the

Subtilisin digest and the lower fragment from the Trypsin digest (black sub-histogram) reveal sequence coverage far beyond the reach of a 13-18 kDa fragment. It is thus likely that the excised gel portions contained more than one fragment of that size. The faint band at the very top of the gel may consist of residual disulfide-linked A2 dimers and be thus indicative of insufficient amounts of reducing agents in the sample buffer (Figure 1-13a, EL lane analysed).

All in all CD spectroscopy and mass spectrometry fingerprinting confirm that the correct protein fragments were purified and that they exhibit discrete folds likely covering the entire sequence length.

1.2.7 Purified AHK5 fragments form inconsistent oligomeric assemblies in solution

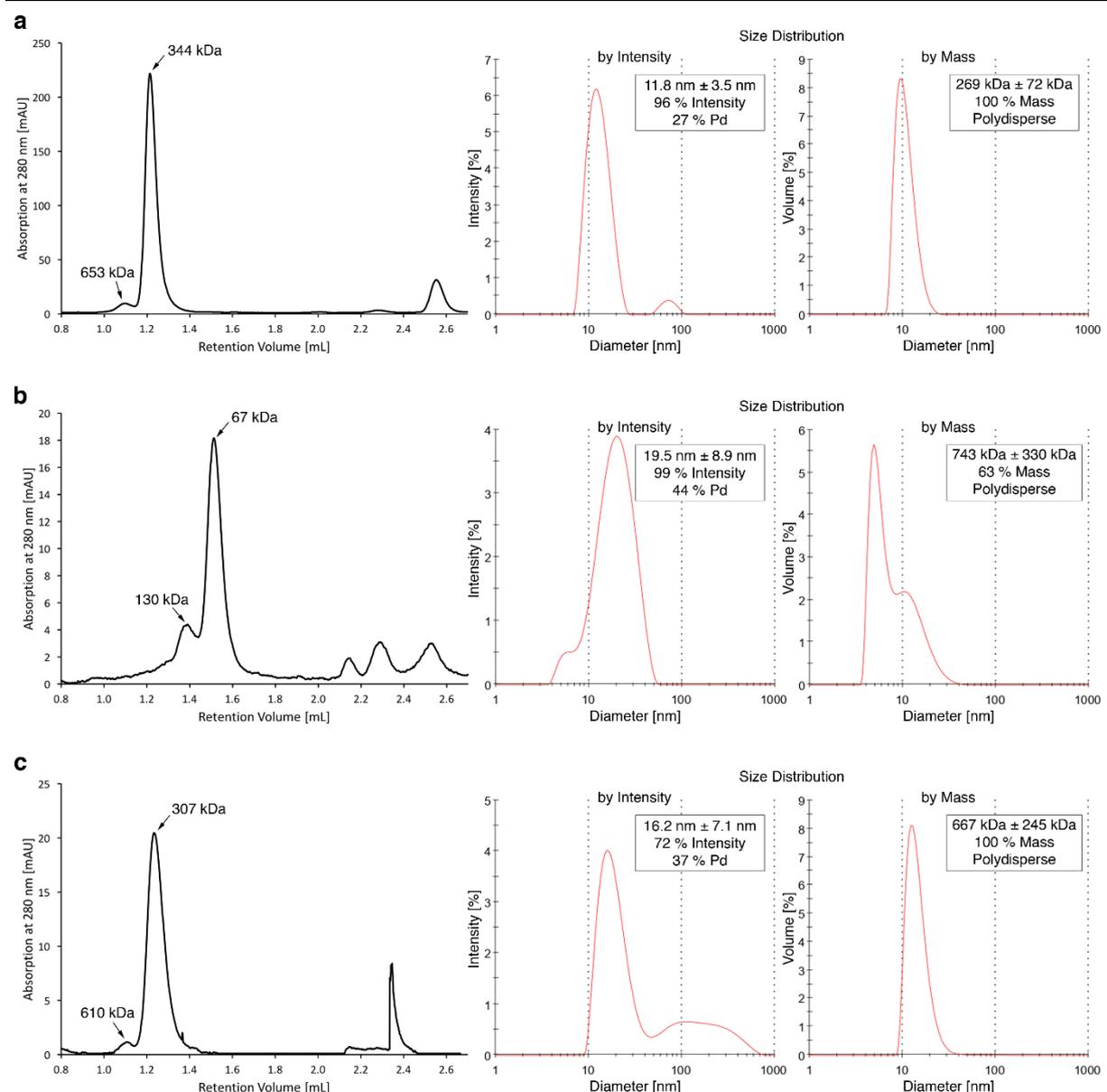


Figure 1-14: Size and homogeneity analysis of AHK5 fragments. **(a)** A3₂, **(b)** C2, and **(c)** A2: Analytical SEC traces (left) are plotted by UV absorption at 280 nm. Mass assignment to retention volumes was performed according to in-house calibration (see Appendix, **Table 5-1** and **Figure 5-1**). DLS distributions by scattering intensity (centre) and calculated mass fraction (right) are shown as automatically computed by the Zetasizer software, using cumulant analyses of 3 experiments consisting of 20 single measurements each. Inset boxes contain the calculated details for the major peak species.

In order to assess sample homogeneity, oligomeric assembly, and ultimately suitability for crystallisation, purified AHK5 fragments were subjected to analytical SEC (aSEC) and particle size estimation by dynamic light scattering (DLS). In all cases, results were inconclusive and pointed towards a strong inherent heterogeneity of AHK5 fragments *in vitro* (Figure 1-14). All three constructs migrate on aSEC as a single discrete peak, preceded by a shoulder corresponding to approximately double the molecular weight, based on the column's calibration (see Appendix, Table 5-1 and Figure 5-1). Estimates as to the oligomeric assemblies of the fragments are at best ambiguous. The main peak for A3₂ (Figure 1-14a, left) elutes after a volume of 1.213 mL, corresponding to a calculated molecular weight of 344 kDa. Given the monomeric mass of 42.6 kDa, this would signify an octameric spheroid or ellipsoid or a less compact (i.e. elongated) assembly of an unknown number of (but less than eight) monomers. The estimated particle weight is corroborated by DLS (Figure 1-14a, right). The calculated weight of 269 kDa allows for a hexameric or lower assembly. However, sample polydispersity and the resulting large error in size estimation may be indicative of a non-compact assembly. This could be caused by intrinsic flexibility in the protein. As the size calculation is based on the Stokes radius, which assumes a spherical particle, the values obtained here could also be indicative of AHK5's deviation from a solid spherical shape.

Size estimation is even more ambiguous for the C2 fragment. While the calculated elution volumes on aSEC (67 kDa and 130 kDa, respectively, Figure 1-14b, left) are compatible with an equilibrium between monomer and dimer ($M_{w(C2)} = 46.6$ kDa), size distributions calculated from DLS experiments (Figure 1-14b, centre and right) are strongly indicative of a very heterogeneous sample, leading to a highly unreliable weight estimation of $743 \text{ kDa} \pm 330 \text{ kDa}$.

The major species for A2 has an apparent molecular weight of 307 kDa, according to aSEC (Figure 1-14c, left), corresponding to a compact tetramer or lower ($M_{w(A2)} = 70.3$ kDa). Again, DLS paints a very heterogeneous picture with high polydispersity and very broad distributions. The resulting molecular weight estimate of $667 \text{ kDa} \pm 245 \text{ kDa}$ are thus not reliable.

Regardless of the high discrepancies between the expected dimeric assemblies and the highly divergent molecular weight estimates calculated by different methods, the small peaks preceding the main elution events on aSEC are always compatible with an assembly of twice the main species (Figure 1-14, left panels). This strongly suggests the presence of discrete oligomeric assemblies rather than of a continuum of mono-, di-, trimers (etc.). Under this light, the highly polydisperse nature of the DLS distributions can be attributed to protein flexibility or partial foldedness, rather than ambiguous oligomers.

1.2.8 Fragments purified to date are too heterogeneous or unstable for structural studies

1.2.8.1 AHK5 A3₂ forms unusual precipitate in crystallisation trials

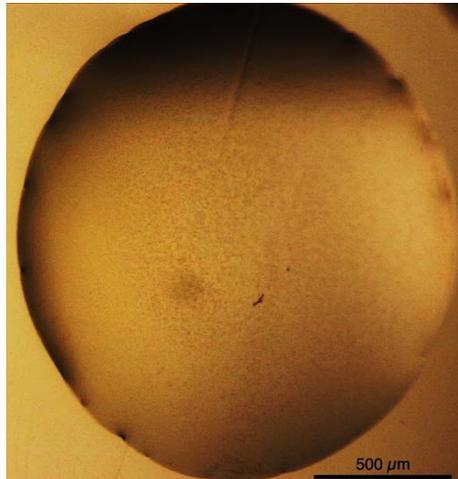


Figure 1-15: Example of granular precipitate in AHK5 crystallisation trials. Similar precipitates were found in the vast majority of AHK5 crystallisation drops, independently of protein concentration.

In spite of the aforementioned propensity of AHK5 fragments to form heterogeneous solutions, hopeful crystallisation trials were performed using commercial screens (see Material and Methods Table 3-9). Unfortunately, light granular precipitate appeared in virtually all conditions after about one day. This precipitation event was shown to be independent of protein concentration (0.5 mg/mL to 5 mg/mL) and protein to precipitant volume ratios. After appearing in previously clear drops, no further development could be detected in any condition. Figure 1-15 shows an exemplary photograph of this phenomenon.

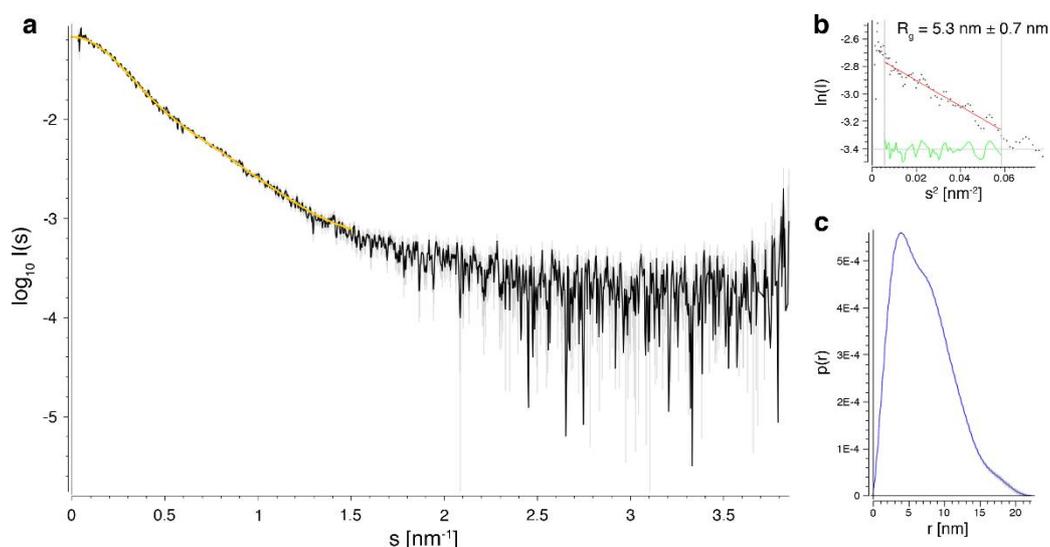
1.2.8.2 SAXS analysis of AHK5 A3₂ yields weak data

Figure 1-16: SAXS data of AHK5 A3₂. **(a)** SAXS profile average extrapolated to zero concentration. **(b)** Guinier region and fit, determination of radius of gyration (R_g) **(c)** Distance distribution function obtained from scattering profile using GNOM [70]. The back-calculated partial fit to the scattering profile is traced in (a) as solid orange line). SAXS data were kindly collected by Dr. Olga Matsarskaia, University of Tübingen.

Small-angle X-ray scattering (SAXS) profiles of AHK5 A3₂ collected at beamline ID02, ESRF (Grenoble, France) yielded an extremely weak and noisy signal (Figure 1-16a). The sharp upwards turn of the data at very low scattering angles in the Guinier representation (Figure 1-16b) could be a sign of aggregation. Nonetheless, a value of $5.3 \text{ nm} \pm 0.7 \text{ nm}$ for the radius of gyration (R_g) could be extracted by linear fitting of the Guinier region. This value is compatible with the hydrodynamic diameters of just over 10 nm obtained by DLS (section 1.2.7). The distance distribution or $p(r)$ function obtained by automated processing with GNOM [70] (Figure 1-16c) shows an at least bimodal distribution of interatomic distances suggesting a multi-domain arrangement. While this finding is in agreement with the predicted domain architecture (Figure 1-1, Figure 1-2), the weak nature of the data, the sharp upturn in the Guinier plot, as well as the extension of the $p(r)$ function to over 20 nm are suggestive of sample stability issues, and potentially of aggregation.

1.2.8.3 AHK5 A2 does not adopt discrete conformations in negative stain TEM

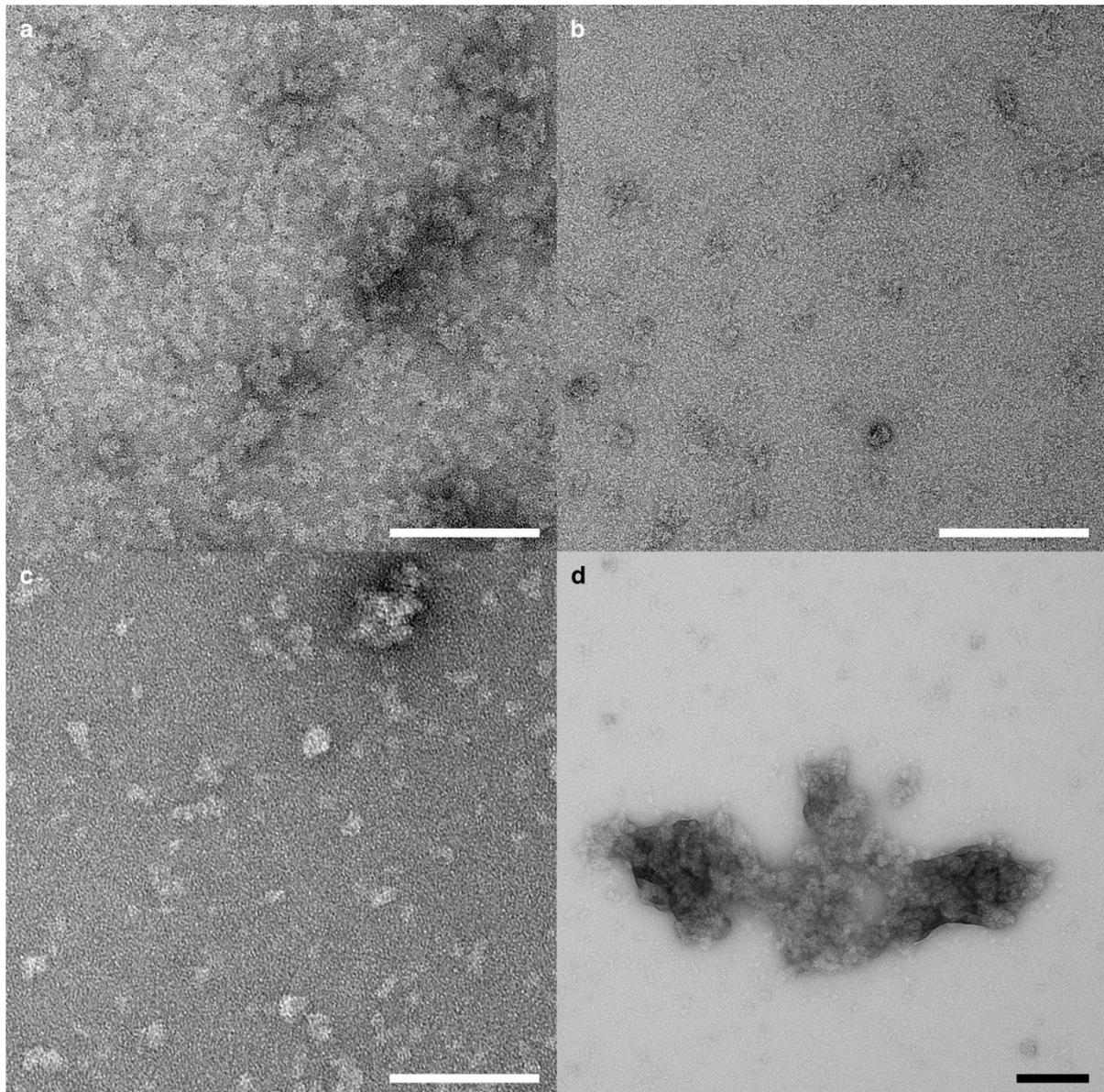


Figure 1-17: Negative stain TEM images of AHK5 A2. Concentrations are (a) 7 $\mu\text{g}.\text{mL}$ and (b-d) 70 ng/mL , courtesy of Dr. York Stierhof, ZMBP, University of Tübingen. Scale bars are 100 nm.

Negative stain transmission electron microscopy (TEM) images were kindly collected by Dr. York Stierhof (ZMBP, University of Tübingen) (Figure 1-17). AHK5 A2 forms particles of approximately 10-20 nm diameter, but no preferred shape or conformations could be discerned. In addition to the vast majority of particles being separated from one another, larger conglomerates or aggregates were found (Figure 1-17c, top, Figure 1-17d). These results highlight the protein's heterogeneity and tendency to aggregate. As A2 is purified in denaturing conditions and refolded *in vitro*, these findings could also indicate folding problems.

1.2.9 AHK5 fragments A2 and A3₂ form covalent dimers upon oxidative stress

1.2.9.1 AHK A3₂ oxidative dimerisation depends on the presence of a single cysteine residue

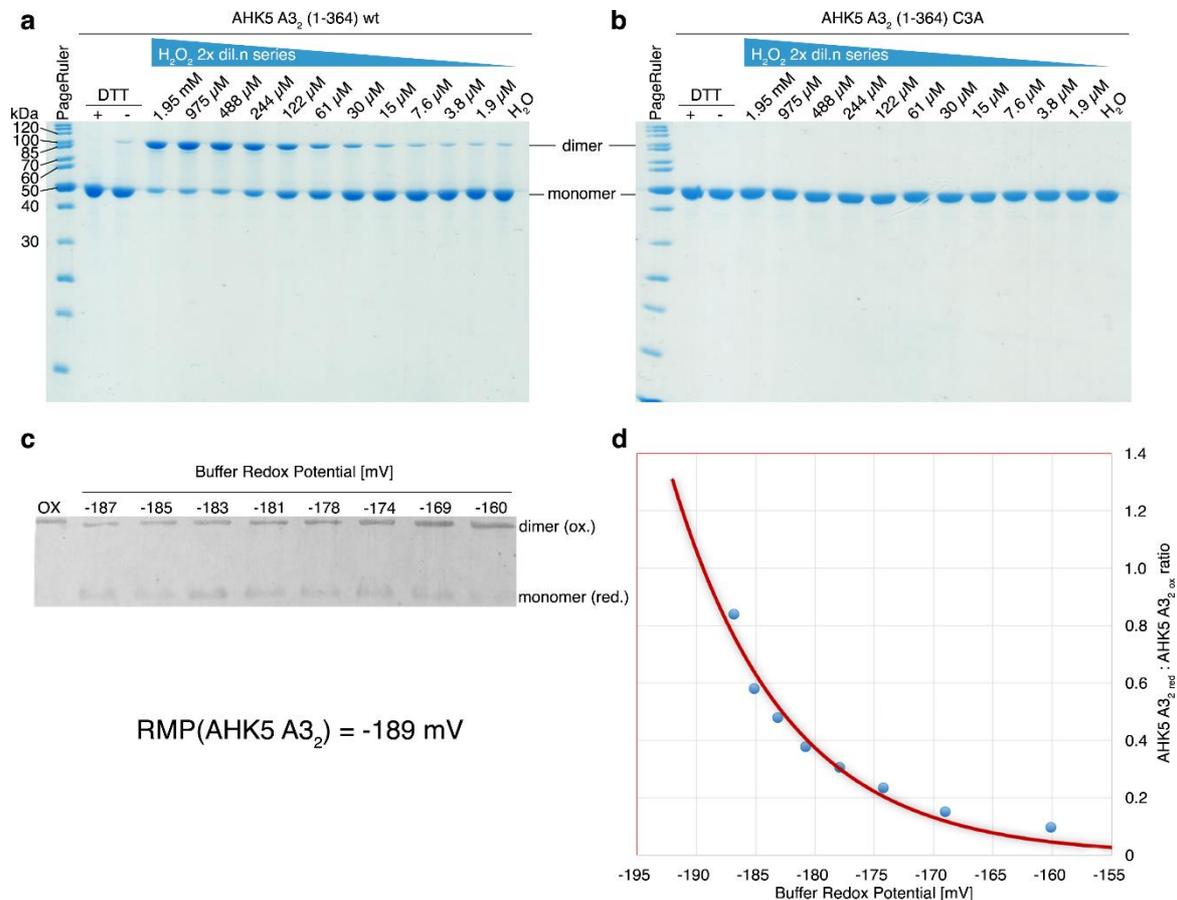


Figure 1-18: Oxidative dimerisation of AHK5 A3₂. **(a)** Non-reducing SDS-PAGE of A3₂ samples pre-incubated in varying concentrations of hydrogen peroxide showing an oxidation-dependent shift from the monomeric to a covalently linked dimeric form. **(b)** Control experiment for (a) performed with the A3₂ Cys3Ala (C3A) mutant. **(c)** Western Blot of A3₂ dimerisation induced by varying ratios of oxidised to reduced glutathione. A custom-made polyclonal rabbit- α -AHK5 antibody was used for primary staining, detection was performed via an IRDye680LT goat- α -rabbit fluorescent antibody (Chromotek, Martinsried, Germany). Courtesy of Thomas Drechsler, ZMBP, University of Tübingen. **(d)** Ratios of reduced to oxidised AHK5 A3₂ extracted from intensity-integrated western blots (c). A least squares fit (red line) yields a redox midpoint potential (RMP) of -189 mV (left). Courtesy of Thomas Drechsler, ZMBP, University of Tübingen.

As AHK5 is involved in the regulation of redox homeostasis [36], the reaction of the input domain (fragment A3₂) to oxidative stress by hydrogen peroxide titration was assayed. AHK5 A3₂ forms covalent dimers in a reaction that is dependent on the hydrogen peroxide concentration. This is evident from the appearance of a band migrating at approx. 100 kDa on a non-reducing SDS-PAGE (Figure 1-18a). This matches double the size of the monomer, which migrates just under 50 kDa. The more hydrogen peroxide is added, the more the band intensities shift towards the dimer. This effect is completely abolished when the only cysteine in A3₂ is replaced by an alanine (C3A mutant, Figure 1-18b), strongly suggesting this residue to be functionally involved in AHK5 redox sensing. These findings agree with this mutant showing the same phenotype as the complete knockout *in planta* (Prof. Dr. Klaus Harter, ZMBP, University of Tübingen, personal communication).

The effect was quantified using purified A3₂ fragment by Thomas Drechsler (ZMBP, University of Tübingen) by using defined ratios of reduced and oxidised glutathione instead of hydrogen peroxide

to generate controllable buffer redox potentials. The resulting AHK5 dimer : monomer ratios could be quantified by Western Blot (Figure 1-18c) and fitted using a least-squares fit of the form

$$R_{red:ox} = c_1 \cdot e^{-c_2 \cdot E_{redox}} \quad (1-1)$$

with $R_{red:ox}$ as the ratio between oxidised (dimeric) and reduced (monomeric) AHK5, E_{redox} as the sample's electrochemical or redox potential, and the fitting parameters c_1 and c_2 . Solving Equation (1-1) for equal quantities of oxidised and monomeric AHK5 ($R_{red:ox} = 1$), yields a Redox Midpoint Potential (RMP) of -189 mV. For comparison, the steady state redox potential in *Arabidopsis thaliana* leaf cells has been previously suggested to be -320 mV [71].

1.2.9.2 AHK5 A2 is sensitive to oxidative stress and forms various covalent multimers

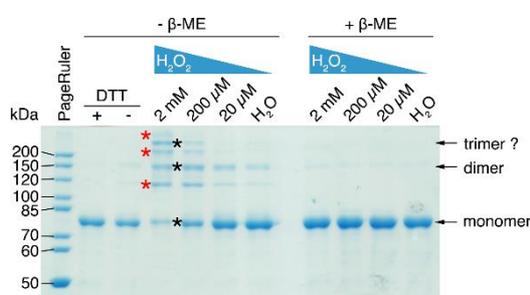


Figure 1-19: Oxidative multimerisation of AHK5 A2. As in **Figure 1-18**, purified protein was incubated with varying concentrations of hydrogen peroxide prior to non-reducing (left, no β -mercaptoethanol in the sample buffer) or reducing (right, with β -mercaptoethanol) SDS-PAGE. Black arrows and black asterisks denote bands corresponding to covalent monomers, dimers, or trimers, respectively. Red asterisks indicate protein bands that do not migrate at integer multiple values of the monomer band. . The original gel picture without labelling asterisks can be found in **Figure 5-9** in the appendix

The longer AHK5 fragment A2, encompassing the histidine kinase and ATPase domains in addition to the input domain, also shows depletion of its monomeric form on non-reducing SDS-PAGE upon treatment with hydrogen peroxide (Figure 1-19). In contrast to A3₂, this does not correlate with the appearance of one discrete band for a covalent dimer, but rather with a multitude of higher-order oligomers, in addition to several bands that do not match the size of any discrete oligomer (Figure 1-19, red arrows). As treatment of half of the sample with reducing agents prior to SDS-PAGE leads to all high molecular weight bands disappearing and full intensity recovery of the monomer band (Figure 1-19, + β -ME lanes), any hypotheses involving protein degradation can be disregarded. As AHK5 contains four cysteines in the ATPase domain, it is more likely that these additional bands correspond to partially unfolded versions of the protein stabilised by intramolecular disulfides. The presence of these additional cysteines, in contrast to the single one for A3₂, may also account for the presence of trimeric and possibly higher oligomeric covalent assemblies *in vitro*.

Initial *in vitro* transphosphorylation assays performed by Thomas Drechsler (ZMBP, University of Tübingen) using AHP1 as a phosphoryl acceptor [40] showed that A2 possesses the desired histidine kinase activity and modulates this activity depending on the buffer's redox potential (Thomas Drechsler, personal communication).

1.3 Discussion

Any endeavour to understand a sensor protein's function must, by default, include a detailed inspection of its sensory input domain. In the case of AHK5, this sensory region was mapped to the first 300 - 360 residue prior to the beginning of this study (Prof. Dr. Klaus Harter, personal communication). N-terminally truncated constructs starting with the Kinase domain and including the C-terminal receiver domain displayed a constitutively active behaviour both *in vivo* and *in vitro*, thus assigning the N-terminal portion an inhibitory role in the resting or non-activated state. The attempts to elucidate form, function, and structure of this region, performed in the study presented here, paint a preliminary and speculative, but nonetheless highly intriguing, picture of mechanisms potentially involved in AHK5 signal perception.

1.3.1 AHK5 domain architecture - a partial picture

AHK5 conservation analysis using ConSurf [67] as well as secondary structure and disorder predictions using PSIPRED and IUPRED [65], [66], [68] yield results that agree well with the presence of canonical Histidine kinase (HisKA), HATPase, and receiver domains. The predictions obtained for these regions contain the expected secondary structure fingerprints for the respective domains (see Results Figure 1-2a, Figure 1-20c,d). The regions in question also coincide with stretches of high conservation scores and are clearly delimited against regions of predicted disorder, such as the long low-complexity linker between HATPase and receiver domains (residues 620-770). The prominent dip in conservation in the middle of the receiver domain covers one flexible loop (residues 507-535) that is unresolved in the crystal structure [40]. Interestingly, the secondary structure fingerprint for the HATPase domain (Figure 1-20c,e) is interrupted between the fourth and fifth β strands (β 4 and β 5) by a 20 - 25 residue-long disordered loop of lowest conservation (Results Figure 1-2a) and unknown purpose.

Using the previous metrics, it is more difficult to divide the entire N-terminal sensory region into domains and variable regions. The conservation score is not as clearly pronounced towards conservation on the one hand or variability on the other. It appears however likely that the first 50 to 70 residues are flexible, based on the IUPRED disorder score. This first segment is followed by a 30 residue stretch with helical propensity, matching the region assigned by SMART to the first of two coiled-coil regions. The next segment (residues 120 - 163) scores neither for secondary structure nor for disorder. This apparent contradiction highlights the limitations of predictive approaches to determine domain boundaries of unknown protein structures. The stretch exhibits a borderline low complexity characteristic with twelve negatively and nine positively charged residues for a total of 21 charges out of 44 residues. Owing to its low conservation score, this segment has never been functionally investigated. Such a high concentration of charges is however intriguing, and its function should be probed. The remainder of the sensory input region scores as a second short coiled-coil followed by a stretch of higher conservation (residues 200 - 340). This section exhibits a secondary structure fingerprint that nearly matches that of PAS / LOV domains (see Results Figure 1-2a, Figure 1-20d,f). Upon comparison of the secondary structure fingerprint derived from the LOV domain of YF1 (PDB ID 4GCZ, Figure 1-20, [72]) with the PSIPRED secondary structure prediction for AHK5 (Results Figure 1-2a), the same pattern

is recognisable, with the exception of strand $\beta 1$, which is scored as an extension of helix $\alpha 1$. Whether this observation can be attributed to more stable secondary structure fingerprinting in the second half of this putative domain than it is in the first half or it is instead indicative of deviations from the canonical PAS fold cannot be stated.

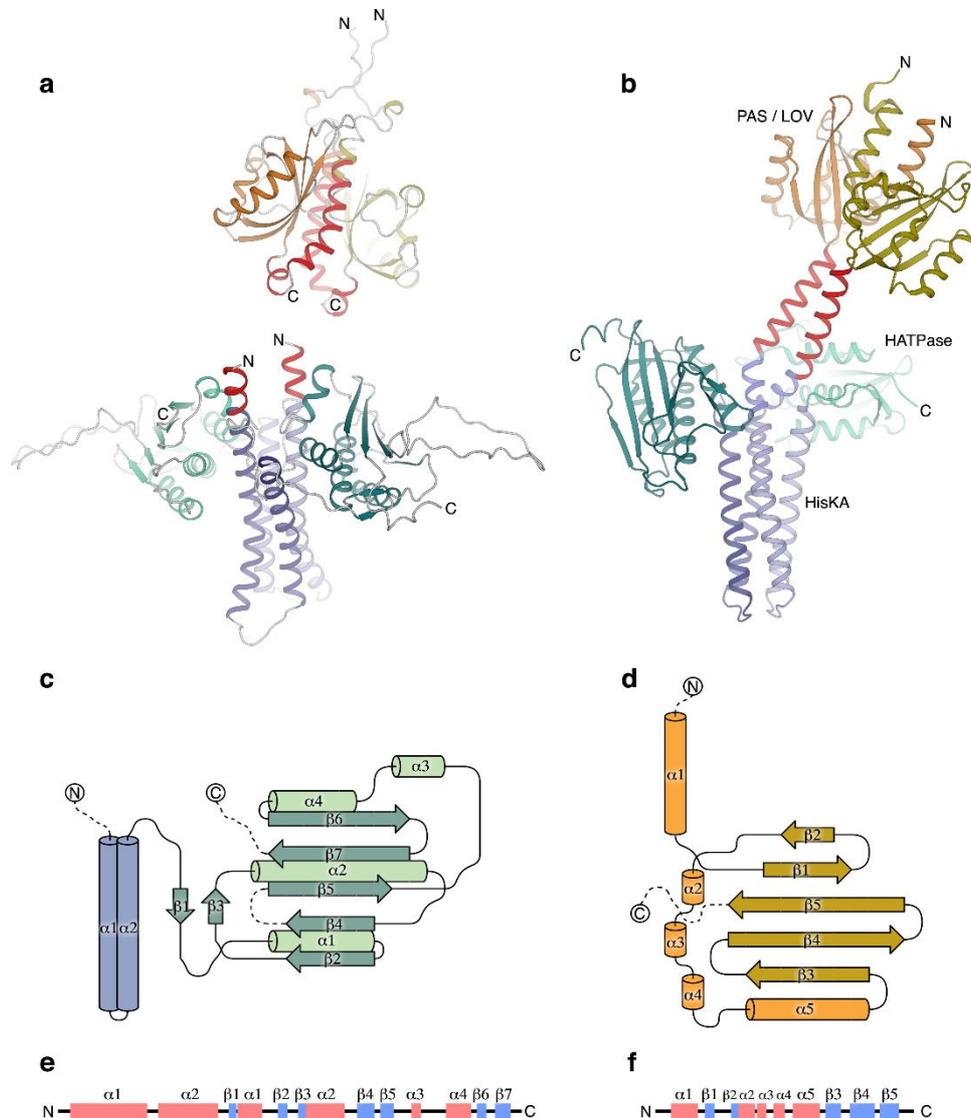


Figure 1-20: Possible domain arrangement of the central portion of AHK5 (residues 200-620). **(a)** Homology models of AHK5 PAS (top) and histidine kinase (bottom) domains calculated using MODELLER [73], [74]. The putative PAS domain is coloured in yellow and orange, the HisKA portion of the catalytic domain in slate, and the HATPase domain in turquoise. The residues pertaining to a putative helix connecting the PAS and HisKA domains are coloured red. The relative orientation of the domains is modelled on **(b)**. **(b)** Crystal structure of a PfyP/FixL hybrid blue light sensor (PDB ID 4GCZ, [72]). Colouring as in **(a)**. **(c)** Schematic domain architecture of the histidine kinase domain (HisKA / HATPase domains). Colouring as in **(a)**. **(d)** Schematic domain architecture of PAS / LOV domains. **(e)**, **(f)** Secondary structure fingerprints of the histidine kinase and PAS / LOV domains, respectively.

PAS domains have been reported in many bacterial sensory histidine kinases. In these enzymes, they fulfil various functions ranging from primary stimulus sensing [60] to signal relay to the catalytic domain by means of structural rearrangements [61], [63]. In *Arabidopsis thaliana*, PAS-like domains are also found in the context of histidine kinases. The extracellular sensory domain of AHK4 consists of two tandem PAS-like folds that constitute a CHASE (Cyclases/Histidine kinases Associated Sensory Extracellular) domain. This high frequency of PAS domains and PAS-like folds in the context of

sensory histidine kinases suggest that, while the secondary structure fingerprinting for such a domain between residues 200 and 340 in AHK5 is not perfect, the presence of such a domain is not unlikely.

The homology model built on the sequence of AHK5 using HHPRED and MODELLER [64], [73]–[75] terminates in an extended α -helix that folds back onto the domain (Figure 1-20a, red helix). Complementarily, the N-terminus of the kinase domain scores as an extended α -helix that could not be reliably modelled. However, the helices marked in red in Figure 1-20a contain the same primary sequence. It is thus conceivable that they may be directly connected and form an extension to the HisKA coiled coil helix bundle, similarly to that of the hybrid blue light sensor from *Bradyrhizobium japonicum* and *Bacillus subtilis* shown in Figure 1-20b. The folding of the connecting helix onto the PAS domain in the AHK5 homology model may be due to penalties for loose packing in the MODELLER algorithm. In the case of such a direct connection between the PAS or PAS-like domain and the catalytic domain, the PAS section may play a signal mediating or amplifying role. Various activation mechanisms for histidine kinases have been postulated over the years [76]–[78] and many rely on mechanical rearrangements of the central four helix bundle as well as the adjacent HATPase domains. These movements can be symmetric or asymmetric helix rotations as well as helix piston movements. It is well possible that the PAS domain's function in AHK5 is to initiate or enhance such subtle structural rearrangements upon signal sensing by the N-terminal portion of the protein. There are to date no data suggesting a direct involvement of the PAS section in AHK5 signal perception.

Based on the available data, a partial hypothetical model of the AHK5 sensory input and kinase domains can be derived. The first identifiable folded element is a first instance of a coiled coil between residues 70 and 100. This segment is only poorly connected to the second detected coiled coil, which starts approximately at residue 160. The region in between displays a high net charge and low sequence complexity. Starting with the second coiled coil region, the protein likely arranges around a central helix bundle until the canonical histidine kinase domain, with a PAS-like fold between residues 200 and 340 acting as a mediator. The histidine kinase is followed by a 170 amino acid long disordered linker that allows for receiver domain (residues 774-922) mobility. The first 70 residues and the low complexity linker between the two coiled coil segments, though likely of functional importance, could not be structurally classified to date.

1.3.2 Purified fragments allow limited inferences on AHK5 shape and function

As of today, the endeavour to obtain soluble, stable, and homogeneous AHK5 fragments for structural studies was met with only partial success. Three constructs were successfully purified, in addition to the previously described C-terminal receiver domain [40]. The newly obtained fragments, aimed at elucidating the mechanism by which AHK5 senses changes in the cellular redox potential, span the first two thirds of the primary sequence.

The shortest construct, termed A3₂, comprises the entire N-terminal portion of the protein preceding the catalytic histidine kinase domain. This section has internally been termed the input domain. In the past, experiments carried out by Michael Heunemann in Prof. Dr. Klaus Harter's group (ZMBP, University of Tübingen) revealed constitutive kinase activity of constructs lacking this portion of the protein (Prof. Harter, personal communication). Fragment A3₂ was chosen for its potential to

shed light on the sensory mechanism, as well as in order to identify potential domain boundaries by limited proteolysis.

C2, the second fragment purified, contains the second part of the input domain and the entire catalytic kinase domain (HisKA and HATPase). The start of this fragment, residue 203, was chosen based on an increase in confidence of the fold prediction algorithms. It is situated immediately C-terminally of the second predicted coiled coil stretch (results Figure 1-2a)) and coincides with a sharp rise in sliding-window-corrected conservation score. It is also situated immediately after a significant spike in the IUPRED disorder prediction. Thus, it contains the entire potential PAS domain. This construct was selected for its potential to contain the portion of the input domain that was deemed most likely to be folded, as well as the connection to the catalytic domain and the catalytic domain itself. Thus, it may be suitable to study the mechanism of signal transduction from the sensory input domain to the activation of histidine kinase activity.

The final construct, A2, spans the entirety of both previous constructs. It was selected in the hope that domain truncations, which may have been present in the previously used fragments, could be avoided. This hypothesis assumes that the C-terminus of the HATPase is accurately predicted.

1.3.2.1 AHK5 fragments are folded but oligomeric assemblies are difficult to judge

Of the three fragments, A3₂ was the only one that could be natively purified. The other two required denaturing lysis and subsequent refolding. All three protein fragments were reproducibly purifiable both in yield and quality. All three displayed strong mean-residue ellipticity signals in CD experiments, suggesting the proteins to be folded. The magnitude of the signal is similar to that observed for *SaGatD/MurT* [79], which was successfully crystallised. Nevertheless, all fragments displayed unexpected behaviour on preparative and analytical SEC. Although the observed elution profiles strongly suggest the presence of one discrete major oligomeric assembly for each fragment (see Results Figure 1-14), the identity of these assemblies is, at best, ambiguous. The two fragments containing the full N-terminus, A2 and A3₂, exhibit very early main elution events on analytical SEC. Based on the in-house calibration of the column used (see Appendix, Table 5-1 and Figure 5-1), the observed retention volumes suggest very high oligomeric assemblies of six (A2) or eight (A3₂) monomers.

However, sample mobility on SEC is dictated by the particle's hydrodynamic volume and not directly by its molecular weight. The calibration experiment performed on the equipment used in this study was carried out using globular (i.e. near-spherical) proteins. Thus, the molecular weight estimates available assume a spherical assembly of a given density. Any deviation from a spherical shape (i.e. a rod-shaped or disk-shaped protein) or from the assumed average density (i.e. incomplete folding of the protein) leads inevitably to an increase in the effective hydrodynamic volume and thus to an overestimation of the molecular weight. In order to assess the plausibility of the results obtained for the AHK5 samples, it is necessary to quantify the expected deviations.

Assessing the compactness of the fold is, unfortunately, not conclusively possible, based only on the available data. As the normalised CD signals reach a magnitude that is comparable with that of fully folded *SaGatD/MurT* [79], complete unfoldedness can be ruled out. However, as CD spectroscopy assesses only secondary structure, even a molten globule-like state would probably yield a high intensity signal, while displaying a significantly lower compactness and thus lower density than a fully folded

protein. It must thus be assumed that fold compactness and flexibility have an impact of unknown magnitude on the apparent hydrodynamic volume, thus likely leading to an overestimation of the molecular weight of AHK5 fragments.

The impact of shape variations, in contrast, has been extensively studied [80]–[82]. Indeed, a protein's shape deviating from a spherical body can be easily described as a prolate (cigar-shaped) or oblate (pancake-shaped) rotational ellipsoid with two distinct semi axes a and b (Figure 1-21c). According to Francis Perrin (obtained from [82]), the correction factors for the frictional coefficients f of such objects (Figure 1-21a) can be described by Equation (1-1) for prolate objects ($b < a$) and Equation (1-3) for oblate objects ($b > a$).

$$\frac{f_{pro}}{f_0} = \frac{\sqrt{1 - \left(\frac{b}{a}\right)^2}}{\left(\frac{b}{a}\right)^{\frac{2}{3}} \cdot \ln\left(\frac{1 + \sqrt{1 - \left(\frac{b}{a}\right)^2}}{\frac{b}{a}}\right)} \quad (1-2)$$

$$\frac{f_{obl}}{f_0} = \frac{\sqrt{\left(\frac{b}{a}\right)^2 - 1}}{\left(\frac{b}{a}\right)^{\frac{2}{3}} \cdot \tan^{-1}\sqrt{\left(\frac{b}{a}\right)^2 - 1}} \quad (1-3)$$

Using the description of the frictional coefficient by Stokes (obtained from [80], [81])

$$f = 6\pi\eta r \quad (1-4)$$

with η denoting the viscosity of the medium, it follows that:

$$\frac{f_{pro\ or\ obl}}{f_0} = \frac{f_{app}}{f_0} = \frac{r_{app}}{r_0} \quad (1-5)$$

In other words, the Stokes radius r_{app} of a prolate or an oblate object is the Stokes radius of a spherical object of the same mass multiplied by the Perrin frictional correction factor.

The apparent molecular weight of a protein obtained from a calibrated SEC experiment is thus that of a hypothetical spherical object with a Stokes radius r_{app} . The relationship between the spherical radius and the molecular weight is given by the volume of the (hypothetical) sphere multiplied by its density:

$$M_{w,0} = \rho_{protein} \cdot \frac{4}{3}\pi \cdot r_0^3 \quad (1-6)$$

It follows from Equations (1-5) and (1-6) that the apparent molecular weight of any prolate or oblate ellipsoid protein, as determined by SEC, is given by:

$$M_{w,app} = \rho_{protein} \cdot \frac{4}{3} \pi \cdot \left(r_0 \cdot \frac{f_{app}}{f_0} \right)^3 \quad (1-7)$$

or, plainly, assuming proteins of equal density:

$$M_{w,app} = M_{w,0} \cdot \left(\frac{f_{app}}{f_0} \right)^3 \quad (1-8)$$

where $M_{w,0}$ is the actual molecular weight of the protein.

Using Equation (1-8), an apparent molecular weight correction factor can be plotted as a function of the rotational ellipsoid axial ratio (long over short semi axis, Figure 1-21b).

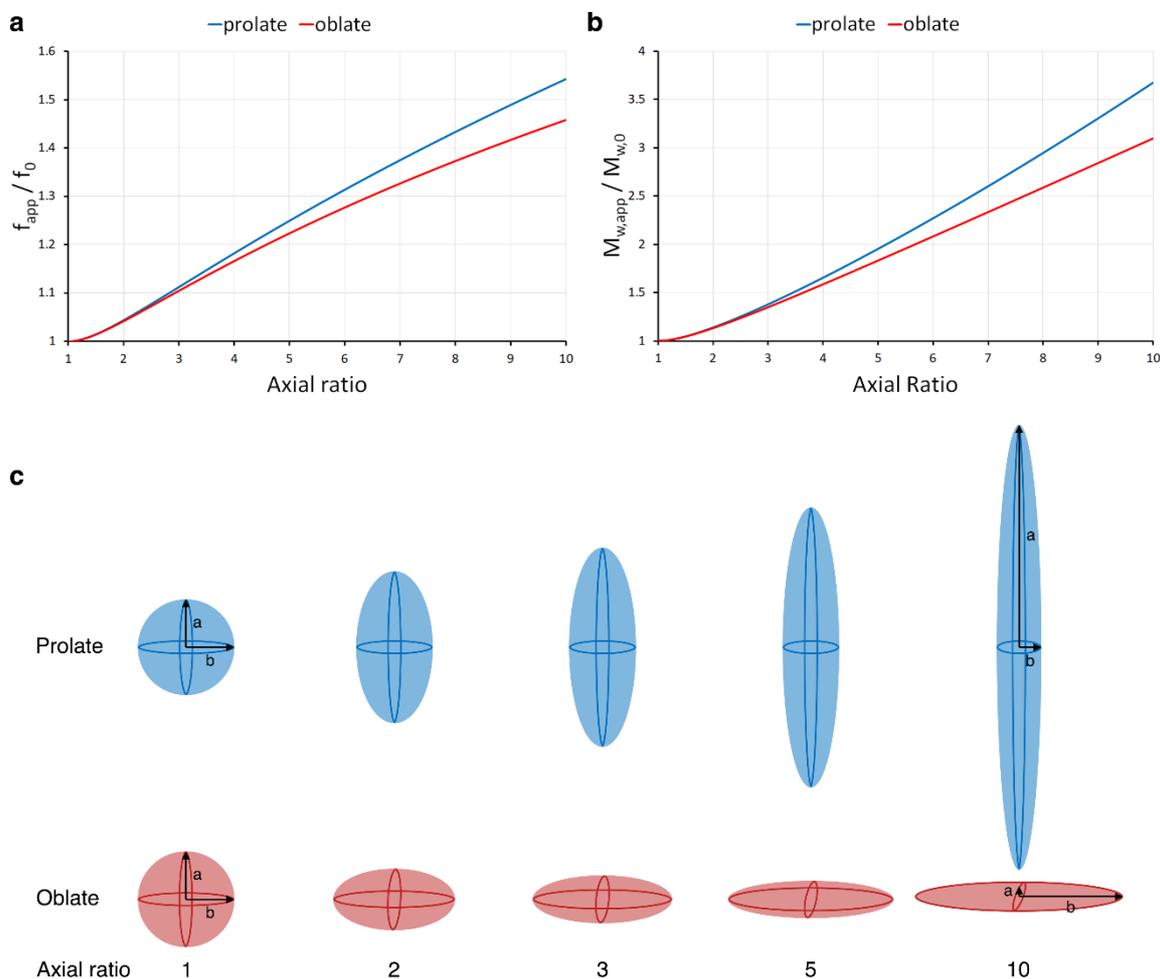


Figure 1-21: Effect of rotational ellipsoid shape on apparent stokes radius and molecular weight estimation via SEC. **(a)** Dependency of frictional correction factor f_{app}/f_0 on rotational ellipsoid shape for prolate (blue, Equation (1-1)) and oblate (red, Equation (1-3)) objects. **(b)** SEC – molecular weight overestimation factors derived from Equation (1-8) for prolate (blue) and oblate (red) objects. **(c)** Prolate (blue, top) and oblate (red, bottom) rotational ellipsoids of equal volume shown for selected axial ratios.

In the case of fragment A3₂ an $M_{w,app}$ of 344 kDa, given the monomeric M_w of 42.6 kDa, a compact rotational ellipsoid containing only two monomers (as is to be expected for a histidine kinase) would require an axial ratio far higher than 10. Such an object is not likely to be mechanically stable. Unless A3₂ forms stable and specific assemblies of four or more monomers, a completely compact fold

is unlikely. If however the N-terminal 100-200 residues were of high flexibility, this would significantly lower the protein's density (Equation (1-7)), hence leading to a much higher hydrodynamic volume and thus $M_{w,app}$.

The C-terminal extension of A3₂, A2, displays an $M_{w,app}$ value of 307 kDa, with a monomeric M_w of 70.3 kDa. Under the assumption of a dimeric assembly and compact fold, these values lead to a prolate object with an axial ratio of 5.7 or an oblate object with an axial ratio of 6.4. This is intriguing, since A3₂ alone shows a higher $M_{w,app}$, despite a much smaller $M_{w,0}$. It could be indicative of either completely different folding (i.e. problems with the refolding procedure) or an attractive effect of the histidine kinase portion on the flexible parts of the N-terminal region, thus forming a more compact assembly. Indeed, this hypothesis is strengthened by the redox-modulable kinase activity of A2 observed by Thomas Drechsler, which suggests a functional protein. It is also compatible with the homology model shown in Figure 1-20a, with the remaining 200 N-terminal residues in part extending the rod-shaped structure and in part folded back onto its surface in a manner yet to be determined.

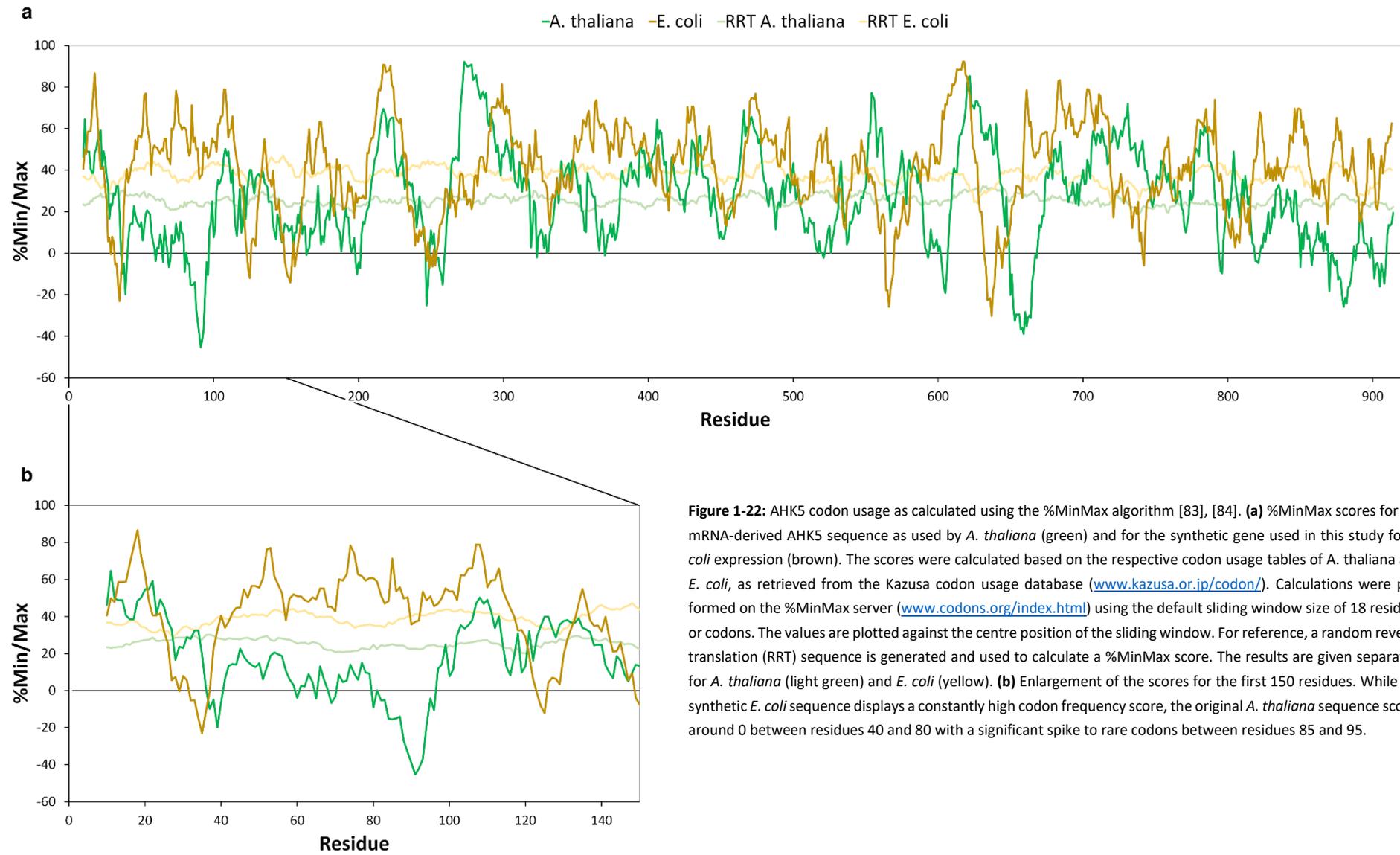
Interestingly, purified C2, which consists only of the histidine kinase portion of AHK5, is mostly monomeric. The $M_{w,app}$ of 67 kDa ($M_{w,C2}=46.6$ kDa) cannot accommodate any oligomers. Unless the fold is not correct, this suggests the N-terminus to play a major role in AHK5 oligomer assembly.

Taken together, these findings paint a highly flexible picture of the AHK5 N-terminal fragments purified to date. The disordered character is further underlined by the high level of sample polydispersity observed in DLS experiments (see Results Figure 1-14).

1.3.3 Most AHK5 fragments are insoluble after overexpression

Although all AHK5 constructs tested to date overexpress at remarkably high levels (see Results Figure 1-3), nearly all the protein mass is typically found in the insoluble fraction after *E. coli* lysis, with the A3₂ fragment being the only exception. There are multiple plausible reasons why the AHK5 fragments tend to aggregate or precipitate when expressed in *E. coli*. The most obvious explanation is that AHK5, a plant protein, is being expressed in a prokaryotic host. This host may lack key chaperones or other stability-inducing factors required for correct AHK5 folding. Along similar lines, the larger AHK5 fragments used in this study may be too large for *E. coli* to handle. The smaller ones, aimed at expressing single domains, could have been incorrectly designed, leading to incomplete domains that cannot properly fold. Additionally, yet unidentified post-translational modifications, such as O-glycosylation or phosphorylation, may play a role in AHK5 structure.

Perhaps one of the most interesting, though least considered, reasons for failure to ectopically express soluble protein is codon usage. In the days of affordable de-novo gene synthesis, it has become standard procedure to use synthetic codon-optimised genes for target protein expression, rather than genomic fragments or cDNA clones. This option offers the possibility to switch from codons used by the donor organism to codons more suitable for the expression organism, thus apparently eliminating translation problems due to rare codons. This strategy usually works by backtranslating the amino acid sequence and assigns the most often used codon to each amino acid. The resulting gene then contains only frequently used codons.



However, rare codons may be required for proper protein folding. It has been repeatedly described that some proteins can fold and unfold reversibly (i.e. by addition and removal of chaotropic agents), while others only fold co-translationally and attempts at unfolding and refolding invariably end in aggregation or precipitation (summarised in [85]). When looking at protein folding as a co-translational event, the addition of each single amino acid leads to a change in the shape the conformational energy landscape, commonly referred to as the folding funnel. Thus, a conformation that may be easily reached for a nascent fragment of a protein, may be kinetically or even thermodynamically hindered in the context of the full-length polypeptide, thus resulting in aggregation rather than in a folded protein [85]. It has been proposed that selective use of rare codons at selected positions can serve to slow down translation at critical steps, in order to allow for a slow-folding domain to reach its target conformation prior to synthesising the rest of the protein [86], [87]. Thus, a gene consisting solely of frequent codons increases the net translation yield, but can drastically decrease the portion of correctly folded protein [87].

Given the very high amounts of AHK5 fragments detected after overexpression, such a mechanism may indeed contribute to AHK5 insolubility. Using the %MinMax tool [83], [84], *A. thaliana* codon usage in AHK5 can be computed from the mRNA sequence (retrieved from the European Nucleotide Archive (www.ebi.ac.uk/ena/) accession number DQ167579.1) and compared with the *E. coli*-optimised synthetic gene used in this study (Figure 1-22a). Not surprisingly, the synthetic sequence is situated nearly completely in the %Max half (i.e. positive scores) of the chart, signifying an above average use of frequent codons. While the original *A. thaliana* sequence also lies mostly in the positive range, one striking difference within the first 100 codons may be of relevance. Between residues 40 and 80, its %MinMax score lies around 0 and displays a sharp peak down to -45 between residues 80 and 100 (Figure 1-22b). This could be indicative of a folding pause for an intermediate nascent polypeptide that cannot take place using the high codon frequency of the synthetic construct. It is thus conceivable that both fragments incorporating the AHK5 N-terminus investigated in this study (A3₂ and A2) may be hampered in their folding due to translation taking place too quickly.

It may be worth to design future studies while keeping in mind the potential roles of synonymous rare codons. A codon harmonisation tool has been programmed by the group of Prof. Patricia Clark at the University of Notre Dame and is available as a Python script (Prof. Patricia Clark, personal communication).

1.3.4 Structural studies are not likely to succeed with current constructs

Based on the data gathered on the AHK5 constructs currently in use, it is unlikely that any of these constructs will be amenable to conduct high resolution structural studies. While CD spectroscopy yielded signals amenable with a compactly folded protein and the protease exclusion assays revealed protected stretches across the entire length of the constructs, further biophysical characterisations raise multiple concerns. Firstly, the ambiguous oligomeric assemblies discussed above (section 1.3.2.1) make the quality of the purified protein difficult to judge. Even if one oligomer were reproducibly purified, it is not possible to assess whether the observed assembly is physiologically relevant. This is especially true for the fragments purified by refolding. Secondly, all fragments displayed increased sample heterogeneity. Even after thorough filtration, DLS measurements never yielded stable results.

All samples measured displayed high sample polydispersity. This indicates, at best, high protein flexibility and, at worst, major issues with the stability of AHK5 oligomeric states. Indeed, SEC elution profiles, recorded at various stages throughout the respective purification procedures, revealed an invariable tendency of all fragments to form higher order oligomers over time. Even though elution fractions were always pooled so as to select for a single species, up-concentration and incubation times between the purification steps led to additional larger species. Finally, attempts at conducting structural studies performed to date all yielded disappointing results. SAXS experiments carried out with A3₂ led to only very weak signal intensities. Initial Guinier analysis revealed the hallmarks of a polydisperse and potentially aggregation-prone sample. Preliminary negative stain TEM images of A2 recorded by Dr. York Stierhof (ZMBP, University of Tübingen) failed to identify any predominant protein conformations but at the same time revealed a high level of heterogeneity, such as aggregation clusters. Finally, all attempts to crystallise either A3₂ or A2 led to sample precipitation, independently of protein concentration.

Taken together, these findings suggest major problems in the stability and oligomeric assemblies of purified AHK5 fragments. Such protein samples, while potentially catalytically active, lack the most crucial quality for structural studies, regardless of the method used: sample homogeneity.

1.3.5 Mode-of-action hypothesis based on redox data

Although the purified AHK5 fragments likely will not allow immediate structural studies, biochemical analyses reveal clues as to the mode of AHK5 activation.

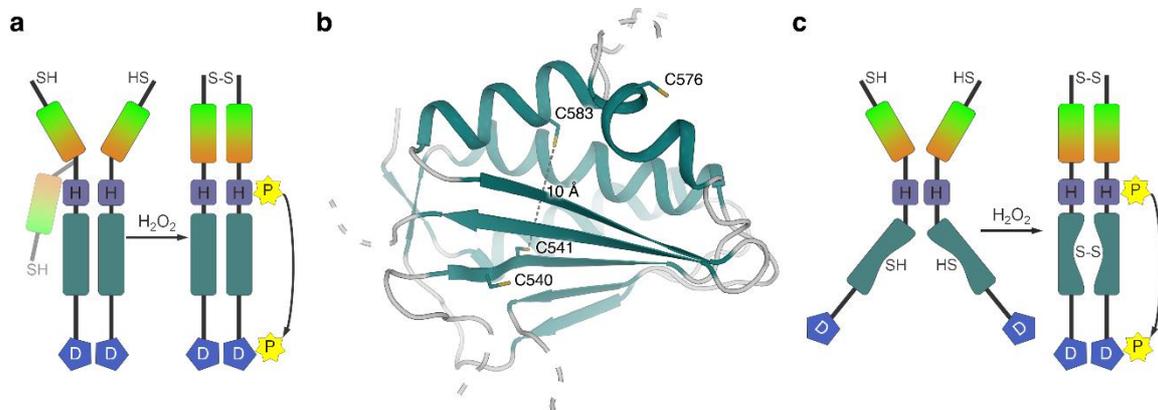


Figure 1-23: Hypothetical model of AHK5 oxidative activation. **(a)** In the reduced, inactive state, the N-terminal sensing domain (green brown) may be flexible and either adopt multiple conformations or even fold back (transparent portion) onto the kinase domain (violet, turquoise) in order to inhibit its action. Upon oxidation (e.g. by H₂O₂ exposure) the input domain dimerises via disulfide bond formation at Cys3, locking the input domain in an activation-competent conformation. **(b)** Homology model of AHK5 HATPase domain, calculated by MODELLER [73], [74]. The four cysteines are shown as sticks. The distance of 10 Å between the only pair of cysteines facing one another within the domain precludes intramolecular disulfide formation. **(c)** Extended model from (a). Additionally to Cys3 intermolecular disulfide formation, a disulfide between the two HATPase domains in the dimer could be envisioned as redox-dependent modulators of AHK5 activity.

Qualitative and quantitative *in vitro* dimerisation assays performed using the AHK5 A3₂ fragment revealed a strong dependency between the redox potential of the buffer solution and the formation of covalent AHK5 dimers. This effect is solely dependent on the presence of a single cysteine at position 3 (Cys3). Additionally, the midpoint equilibrium of this oxidative dimerisation is situated

close to and on the oxidative side of a previously reported value for steady-state cytosolic redox potential in *Arabidopsis* leaf cells [71]. These observations prompt a potential model for AHK5 activation (Figure 1-23a). In the resting state (i. e. under reducing conditions) the AHK5 N-terminus is flexible and may perhaps even serve an inhibitory function by folding back onto the kinase domain. In the event of a rise in the cellular redox potential, for example by generation of reactive oxygen species such as H₂O₂, Cys3 forms an intermolecular disulfide, thus locking AHK5 in a different, catalysis-competent conformation.

A repeat of this experiment with the longer A2 construct, which also contains the kinase domain, revealed the formation of higher order covalent oligomers. An inspection of the amino acid sequence revealed the presence of four cysteines within the HATPase domain. (Figure 1-23b). While the distance between the closest of the two cysteines in this homology model, C541 and C583, is too long to form an intramolecular disulfide bond (10 Å), two cysteines (C540 and C576) are situated on the surface of the domain. It is thus conceivable that these residues may play a similar role to Cys3 in bringing AHK5 into a catalysis-competent conformation as a response to increasing oxidative stress (Figure 1-23c). Even without covalent dimerisation, oxidation of surface-exposed cysteine residues such as C540 or C576 to sulfenic acid may play a role in ATPase activation, either directly or as an intermediate for further modifications, such as S-glutathionylation. Indeed, many proteins become sulfenylated upon hydrogen peroxide treatment in *Arabidopsis*. Either hypothesis needs to be thoroughly investigated by *in vitro* and *in vivo* mutagenesis and mass spectrometry analyses.

1.3.6 Conclusion

Taken together, the results presented in this study constitute a foundation from which to study the mode of activation of the redox-sensitive histidine kinase AHK5 *in vitro*. A plausible initial model for AHK5 activation upon increase of the redox potential could be derived. It includes one or more events of intermolecular oxidative dimerisation by exposed cysteine residues and, potentially, a yet undescribed N-terminal domain of the PAS/LOV variety as a signal mediator.

High resolution structural studies of AHK5 are unlikely to succeed with the protein fragments presented here due to heterogeneity and probable flexibility. Structural analysis of AHK5 will most likely require better knowledge of domain boundaries in order to improve construct design. Potentially, protein expression can be improved by harmonising host codon usage to mimic the original codon usage distribution, thus facilitating co-translational folding. Ultimately, in order to obtain a comprehensive picture of AHK5 structure-function relations, a multi-pronged approach synergising X-ray crystallography, cryoEM, as well as *in vitro* and *in vivo* biochemical techniques will be essential.

1.4 Ongoing Research

In order to better understand AHK5 domain architecture and domain boundaries, access to high amounts of pure full-length protein would be immensely beneficial. To this end, efforts are being pushed forward to express full length AHK5 (AHK5_FL) from Sf9 and Hi5 insect cells.

1.4.1 Transfection of sf 9 cells with the AHK5_FL gene was successful

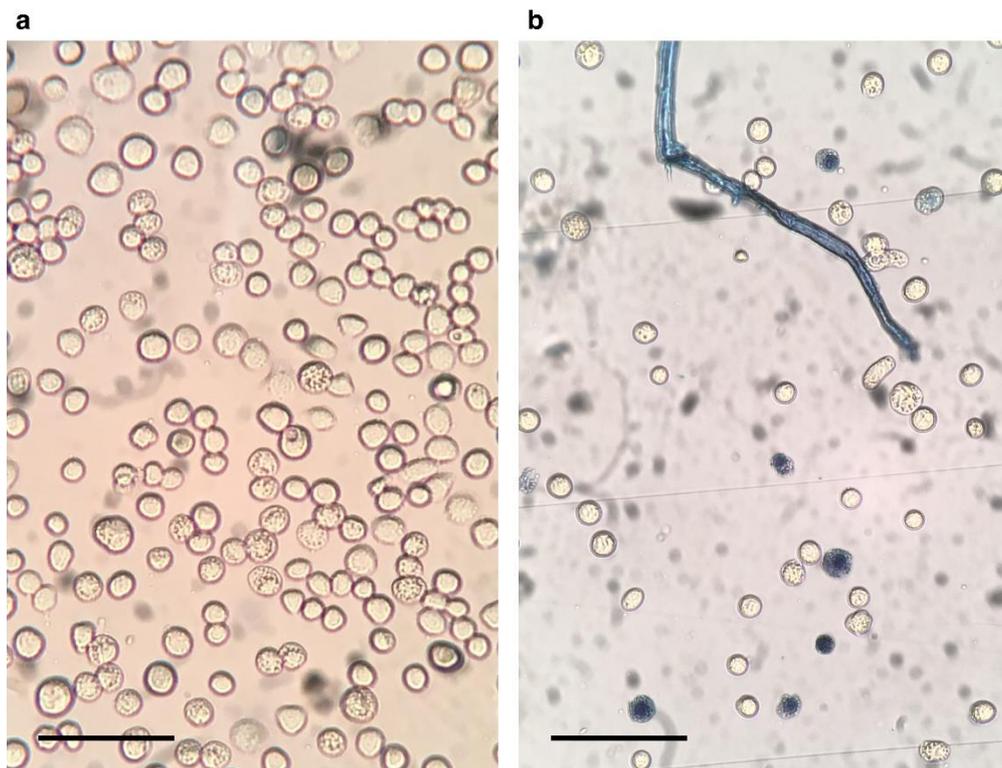


Figure 1-24: Signs of infection of Sf9 insect cells. Scale bars are 100 μm . **(a)** Adherent Sf9 cells 5 days after transfection with AHK5-c-containing bacmid. Cells are enlarged and display a characteristic granular structure. **(b)** Suspension Sf9 cells 5 days after infection with recombinant virus derived from (a). Again, enlarged cells with granular structure are visible. Trypan blue staining highlights dead cells.

After transfection with recombinant baculoviral genome (bacmid) containing the gene for AHK5, adherent Sf9 cells showed the distinct features of viral infection (Figure 1-24a). Many cells were dead and detached (shadows outside of microscope focal plane) and the remaining ones were large and round and most of them displayed a characteristic granular structure. The supernatant from this culture was harvested and stored at 4 $^{\circ}\text{C}$ as the P1 stock. Further enrichment of the transgenic baculovirus was achieved by infecting Sf9 cells in suspension with the P1 stock. These cells again showed signs of infection after five days (Figure 1-24b) with the dead cells clearly visible in the trypan blue stain.

These images confirm that Sf9 cells transfected with the recombinant bacmid could generate infectious virions.

1.4.2 Initial expression tests are promising

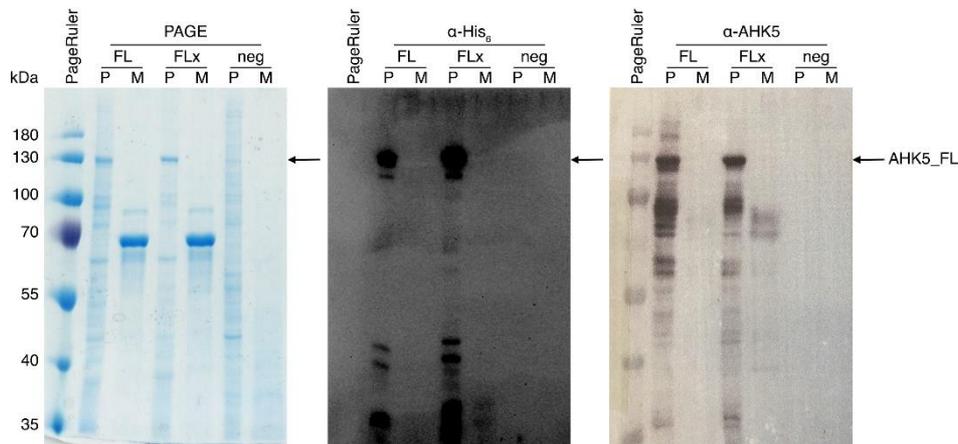


Figure 1-25: Expression test of AHK5_FL adherent Hi5 cells using P2 viral stock. Post-infection samples of cells (P) and cell culture medium (M) were analysed by SDS-PAGE (left) and western blot using anti-His₆ tag (centre) and anti-AHK5 (right) primary antibodies. Two constructs were tested, without (FL) and with (FLx) a Factor X_a cleavage site between AHK5 and the His₆ tag. The arrows mark the height of the AHK5 bands confirmed by MALDI-TOF. Uninfected samples were included as negative controls.

Five-day expression tests using both available AHK5_FL constructs (with and without a Factor X_a cleavage site before the affinity tag) were performed using P2 viral stock and adherent Hi5 cells. The cells as well as the cell culture media were subsequently subjected to SDS-PAGE and Western blot analysis (Figure 1-25). Both constructs led to the presence of a prominent band in the whole cell pellet lane on SDS-PAGE at around 130 kDa. This band was stained with both commercial anti His₆ antibody (Thermo Fisher Scientific, Waltham, MA, USA) and custom-made polyclonal anti AHK5 input domain antibodies provided by Thomas Drechsler (ZMBP, University of Tübingen). As the expected molecular weight of AHK5_FL is of approximately 100 kDa, this band appears to be too large, but mass spectrometry analysis of excised bands performed by Sophie Stotz (Kalbacher group, University of Tübingen) confirmed the identity of the protein to be AHK5. Indeed, full-length AHK5 from whole plant cells has been reported to migrate at larger apparent molecular weights (Prof. Harter, ZMBP, University of Tübingen, personal communication), although the reason for this behaviour is not known to date.

AHK5 can thus be successfully be expressed from insect cells using recombinant baculovirus vectors.

1.4.3 Lysis buffer screening revealed a potential condition for native purification

After obtaining the confirmation of successful expression of AHK5_FL from insect cells, a lysis buffer screen and subsequent initial IMAC purification was performed by Lea Hansen-Palmus (Student assistant, B. Sc.).

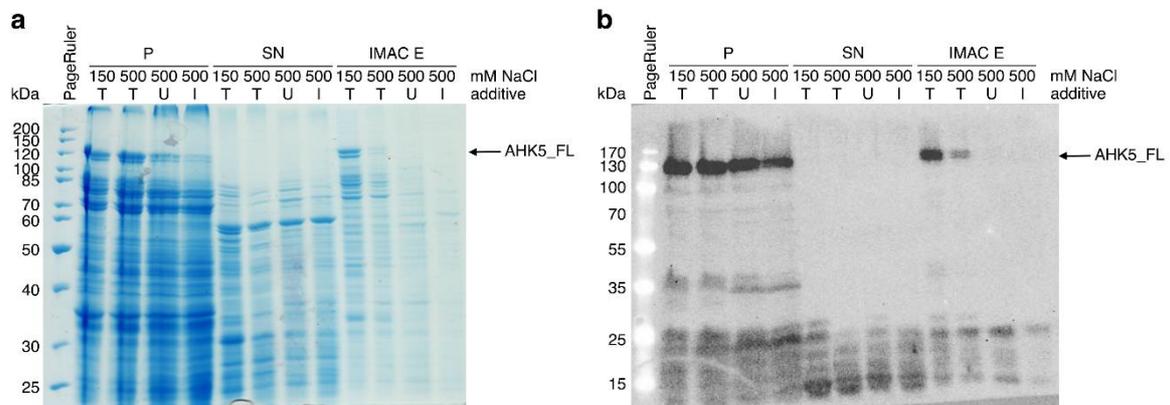


Figure 1-26: Excerpt of a lysis buffer screen on AHK5_FL. Screening was performed by Lea Hansen-Palmus (Student assistant, B. Sc.) on Hi5 cells infected with AHK5_FL-containing baculoviral stocks. **(a)** SDS-PAGE and **(b)** Western Blot stained with a mouse α His6 tag antibody and detected with a goat α-mouse secondary antibody coupled to horseradish peroxidase. Lanes contain lysis pellet (P), supernatant (SN), and the elution from a His-spin trap (IMAC E) (GE Healthcare). Lysis buffer conditions were 50 mM MES, pH 6.0, with NaCl concentrations as indicated by the lane labels. Additives used were 0.1 % Triton X-100 (T), 100 mM urea (U), and 20 mM imidazole (I). A black arrow indicates a double band, which likely consists of multiple isoforms of AHK5_FL.

This initial lysis buffer screen, which covered a pH range from 6 to 8, NaCl concentrations from 150 mM to 1.5 M and tested the effects of imidazole, Triton X-100, and non-denaturing amounts of urea on AHK5 native extraction from cell pellets and subsequent enrichment by IMAC revealed a very promising buffer condition at 50 mM MES, pH 6.0, 150 mM NaCl, 0.1 % Triton X-100 (Figure 1-26). Preliminary upscaling experiments performed by Simon Völpel in the group of Prof. Dr. Thilo Stehle have shown promising results. These data will be used to design a full purification strategy for AHK5-FL. Additionally, phosphorylation studies are underway in the hands of Thomas Drechsler (Harter group, ZMBP, University of Tübingen) with the protein obtained here. Why AHK5-FL expressed in Hi5 cells shows a double band on SDS-PAGE and Western Blot has yet to be addressed. Potentially, it could be indicative of differential phosphorylation or even glycosylation.

1.4.4 Future prospects

The results presented here constitute a solid basis on which to further the work towards elucidating the mechanism of AHK5 signal perception. Refolded A2 fragment as well as the FL protein from insect cell expression will be used in *in vitro* phosphorylation assays in the group of Prof. Dr. Klaus Harter at the ZMBP, University of Tübingen.

Furthermore, the hypothetical activation hotspots, namely Cys3, the putative PAS/LOV domain, and the surface-exposed cysteines in the HATPase domain should be tested for their relevance in AHK5 activation. Cysteine modifications in particular should be studied in a qualitative and quantitative approach using mass spectrometry coupled to either liquid chromatography (LC-MS) for native

conditions or capillary electrophoresis (CE-MS) for denaturing studies. Such efforts will help to fine-tune the hypotheses presented here and direct future structural studies.

The ultimate goal remains to achieve high resolution structural data on the AHK5 sensory input domain. Further refinement of the construct design as well as protein handling conditions are required to reach this milestone. Limited proteolysis and mass spectrometry analyses on AHK5_FL will be used to precisely pinpoint the domain boundaries and guide the design of more stable fragments that are suitable for crystallisation. It is important to note that LC-MS or CE-MS should be preferred over the traditional manual SDS-PAGE with subsequent gel excision and MALDI-TOF analysis.

Taken together, the work presented here lays the foundations to further the understanding of a central regulator of cellular responses in *Arabidopsis thaliana*. Understanding the modulation of this process will help to mechanistically understand many cellular processes that, to date, are known only at the phenomenological level.

2 STRUCTURAL STUDIES OF THE CELL WALL-MODIFYING ENZYME GATD / MURT

2.1 Introduction

Since the discovery of penicillin in 1929 [88], many different types of antibiotic have been developed. The so-called antibiotic revolution during the 1960s and 1970s saw the advent of new molecules capable of targeting the vast majority of bacterial pathogens at very high rates. As a consequence, together with other life-prolonging development of the era after World War 2, life expectancy increased dramatically from under 50 years at the beginning of the 20th century to over 70 years in North America and Europe by the early 1980s. Over the same period of time, infectious diseases were replaced by non-transmissible diseases, such as cardiovascular diseases and cancer, as the most common cause of death in developed countries. Unsurprisingly, people living in the developed countries, especially the younger generations, have grown accustomed to regard infectious diseases caused by bacterial pathogens as little more than an occasional nuisance. Even *Yersinia pestis*, which wreaked devastation on multiple occasions during the middle ages, killing as much as 30 – 60 % of the European population at times, can safely be killed by antibiotics.

However, as we should be reminded by the current COVID-19 pandemic caused by the coronavirus SARS-CoV2, threats from infectious diseases are far from banned. While the emergence of novel pathogens, such as SARS-CoV2, poses a constant unknown threat we have to live with and react to when times call for it, a very different threat is posed by established and well-understood pathogens. Since the early 1980s, no fundamentally new antibiotics have been discovered. All new molecules approved for clinical use since then are modified versions of existing drugs. As is reviewed in [89], in many instances it took only a few years from the introduction of a given antibiotic to the market until the emergence of a bacterial strain that had developed a resistance mechanism against that molecule and was thus capable of surviving antibiotic treatment. Widespread use of antibiotics in clinical settings and misuse as growth boosters in livestock farming have accelerated the development of resistant strains. With the number of antibiotic resistances constantly rising, so-called multi-resistant or even pan-resistant microbes have become a serious threat. Many such strains are found in increasing numbers in hospital environments where they cause an ever-increasing medical challenge and financial burden [90]. According to projections performed on behalf of the government of the United Kingdom in 2016 [91], this development could lead to bacterial infections by resistant strains surpassing cancer in terms of yearly mortality by the middle of the century. In this scenario, antimicrobial resistance would be responsible for 10 million deaths per year worldwide and could cause cumulative costs to the global economy of up to US\$ 100 trillion.

It is hence imperative to push forward research into novel antimicrobial agents in every possible way.

2.1.1 The bacterial cell wall

In contrast to protozoa, bacteria exhibit very discrete cellular shapes and assemblies. For example, the enterobacterium *Escherichia coli* forms rod-shaped cells, whereas cocci such as *Staphylococcus aureus* have spherical cells, and spirilla, such as *Campylobacter sp.* assume a corkscrew shape. The ability to assume and maintain these relatively rigid shapes stems from a sophisticated cell wall assembly.

As in all cells, the bacterial cell is delimited by a plasma membrane composed of a phospholipid bilayer. Outside of the membrane, the bacterium is then encapsulated in a carbohydrate peptide copolymer termed peptidoglycan. This mesh is differentially functionalised depending on the bacterial species. In the case of Gram-positive bacteria, the layer is relatively thick and highly functionalised, for example with wall teichoic acids [92] in *Staphylococci* or mycolic acids in *Mycobacteria* [93] (which react negatively to Gram staining due to the mycolic acids, but are architecturally similar to other Gram-positive bacteria). In contrast, Gram-negative bacteria have a comparatively simple and thin peptidoglycan layer that is then covered by a second membrane.

It is the layer of peptidoglycan that functions as the primary structural component of the bacterial cell wall. It confers the cell rigidity as well as protection from mechanical and chemical damage. It also allows bacteria to establish high turgor pressures, ranging from 0.3 – 3.0 bar for the Gram-negative *Escherichia coli* up to as much as 20 bar for *Staphylococcus aureus*, a Gram-positive bacterium [94]–[96]. It is constituted by a highly crosslinked network of polysaccharide chains and short peptides. The polysaccharide is surprisingly homogeneous across all bacteria [97]. It invariably consists of alternating β -1-4 linked N-acetyl-glucosamine (GlcNAc) and N-acetyl-muramic acid (MurNAc) moieties. While the identity of the sugars is constant, the length of the polysaccharide chains varies significantly from species to species [97].

Plasticity is conferred to the polymer by crosslinking the polysaccharide chains with short peptides. These peptides differ significantly in their identity, depending on the bacterial species. In general, they are composed of a stem peptide of 3 – 5 mixed L- and D-amino acids that is connected to the D-lactate moiety of the N-acetylmuramic acid. Typically, the fourth amino acid is a diamino acid (L-lysine or meso-diaminopimelic acid (mDAP) in most cases), enabling the crosslinking between adjacent stem peptides. The identity of the crosslink also varies and can consist of either a direct connection or be mediated via a bridging peptide that contains small amino acids, such as glycines or serines. The different types of peptidoglycan have been systematically reviewed in the excellent work of Schleifer and Kandler in 1972 [97].

2.1.1.1 Peptidoglycan biosynthesis

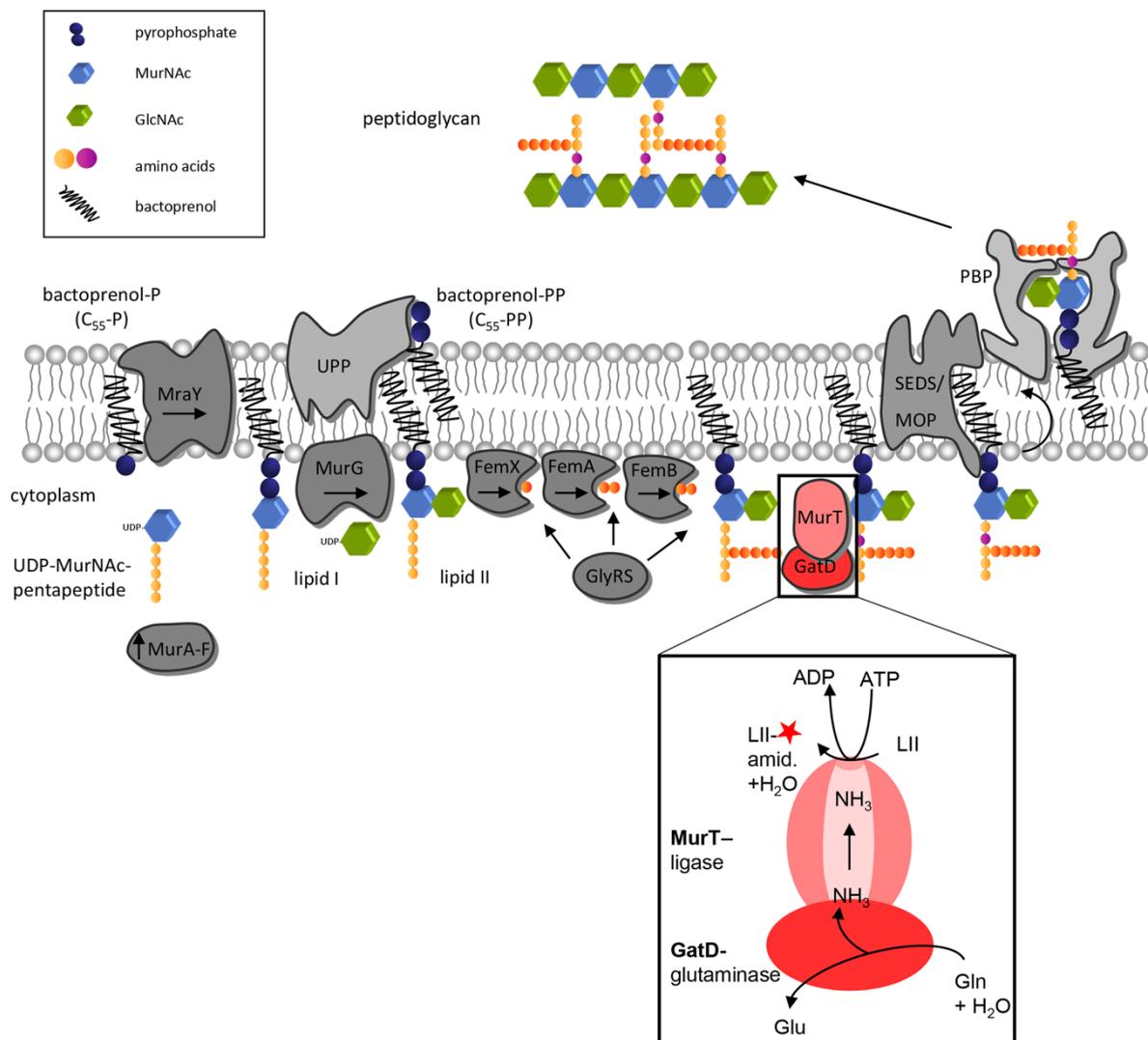


Figure 2-1: Overview of the steps involved in peptidoglycan biosynthesis in *Staphylococcus aureus*. MurA-F synthesise Park's nucleotide (UDP-MurNAC-pentapeptide), which is then anchored to the membrane via an undecaprenyl pyrophosphate tail by MraY, to form Lipid I. MurG adds GlcNAc to form Lipid II, and the FemXAB enzymes add the pentaglycine bridge to the lysine side chain in the stem peptide. GatD. MurT then acts on Lipid II to amidate the D-isoglutamate in the stem peptide (box). Hereby, GatD deaminates glutamine to form ammonia and glutamate. Ammonia is then channelled to the active site of MurT, where it is used to amidate the peptide substrate in an ATP-consuming reaction. Afterwards, the modified Lipid II is flipped to the periplasm and used for peptidoglycan polymerisation by penicillin-binding proteins. Taken from [79].

As depicted in Figure 2-1, peptidoglycan biosynthesis occurs in multiple steps. Firstly, the building blocks of the polymer are generated in the cytosol. Secondly, the building blocks are transferred to the periplasmic space by means of an isoprenoid membrane anchor. Thirdly, specialised transglycosidases and transpeptidases catalyse the polymerisation while recycling the membrane anchor. The biosynthetic pathway will be exemplified here by that of *Staphylococcus aureus*. Although single enzymes are named differently in other organisms, the general architecture of the pathway is applicable to most bacteria.

The assembly of the building blocks begins in the soluble phase in the cytosol with UDP-activated N-acetylglucosamine (GlcNAc). A pyruvate moiety is then added to the 3' hydroxyl group of GlcNAc by the enzyme MurA to form 3'-enolpyruvyl-GlcNAc. The NADH-dependent oxidoreductase MurB

then reduces the enolpyruvyl moiety to lactate, forming 3'-lactyl-GlcNAc or N-acetylmuramic acid (MurNAc). A set of Mur ligases, termed MurC-F then catalyse the subsequent addition of a total of five amino acids to the carboxyl group of MurNAc thus building the stem peptide. The resulting molecule, UDP-MurNAc-pentapeptide, often called Park's nucleotide, is the final cytosolic soluble peptidoglycan precursor. Park's nucleotide is then tethered to the plasma membrane by means of an undecaprenyl phosphate anchor. This molecule, which is also known as bactoprenol, is attached to the soluble precursor via a pyrophosphate linker under the release of UMP by MraY. This first membrane-bound intermediate product is termed Lipid I. Addition of a GlcNAc moiety from the cytosolic pool of UDP-GlcNAc by MurG yields Lipid II [98]. Further, the enzymes FemX, A, and B add a stretch of five glycines to the lysine- ϵ -amine in the Lipid II peptide stem. This Lipid II pentapeptide is the final peptidoglycan building block. It is subsequently flipped across the membrane to the periplasm [99], [100]. There, transglycosylases and transpeptidases termed Penicillin-Binding Proteins (PBP) add the finished building blocks to the peptidoglycan mesh [101].

2.1.2 Antibiotics targeting the cell wall

The very first antibiotic ever to be discovered and to be used in a clinical setting, penicillin, functions by inhibiting peptidoglycan polymerisation. Over the years, a wealth of different antimicrobially active compounds have been discovered and adapted for medical use. Although many of them act on very different portions of the bacterial metabolism, the cell wall remains the target of the vast majority of antibiotics. As of 2010, 50 % of prescription antibiotic doses were molecules targeting the cell wall, in the case of intravenous (IV) application, the portion reaches as high as 75 % [102].

Modern antibiotics that target the cell wall can be roughly grouped into three categories: those interfering with the cytosolic preparation of Lipid II, those directly interacting with Lipid II, thus preventing it from contacting any components of the downstream assembly machinery, and those that inhibit the final polymerisation reactions.

The first group comprises amino acid analogues and sugar analogues. The sugar analogues, such as tunicamycin, inhibit the assembly of Lipid I, i.e. the reaction of MraY. Unfortunately, they also inhibit eukaryotic N-glycosylation and are still under investigation with regard to their potential clinical applications [103]. The amino acid analogues D-cycloserine and fosfomycin are the only two pharmacologically approved antibiotics to inhibit the cytosolic steps of peptidoglycan biosynthesis [102]. D-cycloserine acts as a structural analogue of D-alanine, thus inhibiting the assembly of the D-Ala-D-Ala dipeptide. It is used as a second line antibiotic for the treatment of resistant *Mycobacterium tuberculosis* infections [104]. Fosfomycin is the most widely used molecule from this first group. It is the only one interfering with the early steps of Park's nucleotide assembly. It acts as a phosphoenol pyruvate mimic and alkylates the catalytic cysteine in MurA, thus inhibiting the enzyme [104], [105]. Although it has been termed "the only antibiotic currently in clinical use that targets a Mur enzyme" [104], it should be pointed out that MurA is not a Mur ligase. This term refers to the ATP-dependent amidotransferases MurC-F [106], and MurT [79], [107]–[109]. There are currently no candidate molecules acting against any of these enzymes.

The second group consists of, generally speaking, modified peptides. Among other, more experimental molecules [102], the most prominent representatives of this group are the glycopeptides

and, to a lesser extent, the lantibiotics. Lantibiotics, or lanthionine-containing antibiotic peptides associate with the membrane in a yet undescribed fashion. They then proceed to bind to Lipid II in various ways [110]. Nisin for instance binds to the pyrophosphate of Lipid II [111]. Although these molecules have been known for a considerable amount of time, none have yet been approved for medical use. The most prospective candidate, nisin, is however being successfully used as a food conservation agent [111]. The most famous glycopeptide is, undoubtedly, vancomycin. Like lantibiotics, glycopeptides act by binding to Lipid II, thus interrupting the peptidoglycan biosynthetic pathway. Specifically, vancomycin binds to the terminal D-Ala-D-Ala moiety of the stem peptide [112], [113]. Vancomycin is a last-resort antibiotic that is used to treat infections with otherwise resistant bacteria, such as MRSA [114].

The third group could be also termed the archetype group for antibiotics. It is, of course, the β -lactam group. β -lactams come in various shapes and functionalisations. They comprise the penicillins, cephalosporins, monobactams, and carbapenems [102]. The core functional unit is the namesake β -lactam ring, which exhibits structural similarity to the terminal D-Ala-D-Ala in the stem peptide. It can covalently bind to the active site of the dedicated transpeptidase enzymes (aptly termed the penicillin binding proteins, PBPs [101]) thus irreversibly inactivating them and impeding peptidoglycan synthesis [102], [115]. β -lactams are, until today, the most widely used class of antibiotics [115].

2.1.3 Mechanisms of resistance and resistance development

As soon as a population of bacteria is exposed to an antimicrobial agent or antibiotic, they are subjected to an evolutionary selective pressure to develop means and abilities to withstand the effect of the drug. In other words, should any bacterium within that population develop or acquire any means of tolerance or even resistance against the drug, it will increase its chance of survival. Fortunately for the bacteria, but very much unfortunately for us humans wishing to employ antibiotics, bacteria have quite short generation times and the advantage of large numbers of individuals. These features enormously increase their statistical odds at successfully attaining resistance.

Resistance to antibiotics can be achieved in two basic ways [116]: (a) A mutation of an existing gene or (b) the uptake of exogenous DNA coding for a resistance determinant. Thus achieved genetic alterations can lead to five general mechanisms of resistance: (i) The target of the antibiotic can decrease its affinity to the antimicrobial molecule, (ii) the antibiotic uptake can be reduced or (iii) its efflux increased, or (iv) the overall metabolic pathway can be shifted so as to tolerate the presence of the antimicrobial agent. Finally, (v) the antibiotic molecule itself can be targeted by a modifying or degrading enzyme. While the last one can only be achieved by the uptake of a pre-evolved gene, the first four mechanisms can be achieved either by a transgene or endogenously, with relatively few point mutations. Although most mutationally acquired resistances will invariably decrease the bacterial fitness, for example by generally reducing the uptake of small molecules, thus also reducing the uptake of nutrients [116], such events are relatively frequent. Thus, the more often one particular antibiotic is employed, the higher the statistical probability of a successful mutation. Additionally, pre-existing mutations can further evolve, thus fine-tuning the resistance mechanisms. Examples for such mutationally acquired antibiotic resistances are rifamycin resistance by a single point mutation in the bacterial RNA

polymerase [117], fluoroquinolone resistance by mutating a set of DNA gyrases [118], or the tolerance of oxazolidinones by modification of various ribosomal proteins [119].

Most antibiotics are derived from molecules generated by certain fungi or even bacteria in order to prevent the growth of harmful or concurring microorganisms. Hence, sophisticated resistance mechanisms have already existed for a long time in bacteria that share an ecological niche with an antibiotic-producing organism. If the source organism is a bacterium itself, such as actinobacteria from the genus *Streptomyces* for the antibiotic streptomycin [104], a resistance mechanism must be in place in that organism in order to tolerate its own product. Resistance mechanisms originating from such settings can be extremely specific, very efficient, and highly refined. Genetic information can be shared between bacteria via a number of horizontal gene transfer (HGT) mechanisms. These include transformation of naked DNA derived from decayed cells, transduction via an intermediate shuttle, such as phages, or directly via conjugation. For a comprehensive review on bacterial HGT, refer to [120]. HGT has helped pathogenic bacteria to acquire sophisticated resistance machineries such as β -lactamases (a group of enzymes that degrades β -lactam antibiotics) [121], the MRSA PBP2a transpeptidase variant with reduced affinity for methicillin [122], or the D-Ala-D-Lac-dipeptide-synthesizing enzymes responsible for vancomycin resistance [123]. Once the hurdle of transitioning from the source organism to a clinically relevant pathogen has been overcome, transposon-mediated gene integration and shift of genes between the chromosome and plasmids quickly spreads genetic resistance determinants within a bacterial community [89]–[91], [104], [116], [124], [125].

General overuse of antibiotics both as prescription drugs in clinical settings and in agriculture accelerate these mechanisms of resistance generation. The small number of antibiotics currently under development as well as the total lack of molecules targeting yet unchallenged bacterial metabolic pathways project a grim future for the treatability of bacterial infections.

2.1.4 Modifications of the peptidoglycan layer

Many resistances to antimicrobial agents reside in the structure of the bacterial cell wall itself. In fact, the finally assembled peptidoglycan layer harbours a multitude of chemical modifications to its basic structure. These range from the addition or removal of small functional groups, such as acetyl or amide groups, to the anchoring of large extracellular extensions like teichoic acids in many Gram-positive bacteria [126], [127] or the extensive arabinogalactan matrix found in mycobacteria [93].

Large modifications have very diverse functions. Indeed, it is perhaps somewhat misleading to term mycobacterial arabinogalactan (AG) a peptidoglycan modification. More correctly, peptidoglycan serves as an anchor to the AG matrix and provide a connection to the plasma membrane. It is then the combination of the two matrices, together with the terminal mycolic acids that confer mycobacteria their unique polymer sacculus. In a way, this “modification” confers the bacterium its characteristic plasticity and chemical impenetrability [93].

Wall teichoic acids (WTA) serve a similar purpose. However, they achieve this purpose by a much more subtle mechanism than a brute-force encapsulation. Owing to their polyanionic nature, WTAs attract cations, thus affecting the overall envelope architecture, porosity, and plasticity. Additionally, they serve as scaffolds for many different molecules, such as endogenous autolysins used in cell remodelling. For a comprehensive review on WTA function, please refer to [92].

Finally, the purpose of small modifications is to render the peptidoglycan layer chemically more stable [128]. Bacteria are often exposed to exogenous enzymatic attacks, like an attacking attempting to pierce the cell wall with a dedicated set of endolysins for infection or a host immune system deploying lysozymes to kill the bacteria [129], [130]. Such specialised enzymes oftentimes recognise and bind to a very defined pattern of charges, hydrogen bonds, and hydrophobic patches. Any alterations to this pattern will result in a failure of the enzyme to bind, and subsequently digest its substrate. Thus, modifications in the peptidoglycan structure that reduce its affinity to exogenous peptidoglycan-degrading enzymes are beneficial to the bacterium. The most common small modifications are sugar N-deacetylation or C₆-O-acetylation [131], [132], and the amidation of free carboxyl groups.

2.1.4.1 Peptidoglycan amidation

It has long been known that especially many Gram-positive bacteria display a high degree of amidation in their peptidoglycan [133]. Amidation occurs most frequently at the α -carboxylate of D-isoglutamate at the second position in the stem peptide or, when present, the ϵ -carboxylate of mDAP at position three [97]. Other peptidoglycan amidations also exist, but they are rather rare by comparison, such as at a D-aspartate present in the interpeptide bridge of certain *Lactococci* [134]. Collectively, the function of these modifications has been reported to be an enhancement of cell wall stability against attacks from exogenous enzymes. Particularly, they increase resistance against lysozyme, a key player of the innate immune system [128], [134], [135].

In spite of the apparent medical relevance of this phenomenon, the enzymatic machineries responsible for peptidoglycan amidation have long remained elusive. Only very few amidating enzymes have been identified in selected bacterial species [134]–[136]. At this point, possibly the best-studied peptidoglycan amidating enzyme is the bifunctional, heterodimeric glutaminase / ligase GatD/MurT, which was first identified in *Staphylococcus aureus* [107], [108].

2.1.5 The GatD / MurT bifunctional enzyme complex

GatD/MurT was identified just under a decade ago by two research groups independently working on the cell wall biosynthesis of the pathogenic bacterium *Staphylococcus aureus* [107], [108]. The first player in the heterodimeric enzyme complex is the Type I glutamine amidotransferase (GA-Tase) GatD. It deamidates glutamine from the cytosolic pool, thus generating glutamate and ammonia. The nascent ammonia is then used by the Mur ligase-like [106] MurT to amidate the α -carboxyl of D-isoglutamate at position two in the peptidoglycan stem peptide, in an ATP-dependent reaction. Hereby, GatD/MurT is active on the membrane-bound peptidoglycan precursors Lipid I and Lipid II. Although *in vitro* analyses revealed that the enzyme is able to also perform the amidation reaction on Park's nucleotide [107], no such activity could be observed *in vivo*. Shortly thereafter, the presence of a homologous enzyme complex was reported in *Streptococcus pneumoniae* and shown to be required for an *in vitro* reconstitution of peptidoglycan assembly [137].

Crucially, conditional knockout and knockdown studies performed in both species revealed the enzyme to be essential for bacterial survival. Attenuation phenotypes showed severe growth de-

fects and increased susceptibility to lysozyme and β -lactam antibiotics [107], [108], [128]. These observations have led to suggest a difference in affinity of the relevant PBPs for the amidated versus the unamidated form of Lipid II [108]. Complete knockout of GatD/MurT was lethal.

Interestingly, genomic sequence analyses revealed that a wide range of Gram-positive bacteria including Staphylococci, Streptococci, Clostridia, and the more distantly related Mycobacteria, likely possess homologous proteins [107]. In addition to the already characterised GatD/MurT instances in *S. aureus* and *S. pneumoniae*, knockout screening has confirmed the existence and essentiality of the enzyme complex in *Mycobacterium tuberculosis* [138]. The potentially widespread presence of this enzyme in many pathogenic bacteria render it an interesting target for the development of novel antimicrobials. Being able to selectively inhibit GatD/MurT could be a game changer in the fight against antimicrobial resistance. The fact that beneficial commensal bacteria, such as *E. coli*, do not amidate their peptidoglycan matrix and hence do not possess a homologous enzyme, offer the added bonus of targeting specificity. In contrast, broad action antibiotics kill all bacteria, not only the pathogenic ones. This can have adverse effects on various bodily function. Additionally, as more bacteria are affected, it also promotes resistance development.

2.1.6 Aim of this project

The aim of this project, which is currently being funded by the German Research Foundation (Deutsche Forschungsgesellschaft, DFG) under the grant CRC-TRR261 “Cellular Mechanisms of Antibiotic Action and Production”, a cooperation between the universities of Tübingen and Bonn, is to elucidate the molecular mechanism behind the amidation reaction catalysed by GatD/MurT. This aim is to be attained with a combinational approach of structural biology, *in vitro*, *in vivo*, and *in silico* investigations.

Ultimately, the goal is to gain information about the structural aspects as well as the dynamics of GatD/MurT-catalysed peptidoglycan amidation. This information, founded on high-resolution crystallographic data, could be used in a structure-guided drug design strategy to search for potential inhibitors. For this, co-crystal structures of the active enzyme in complex with its ammonia donor glutamine, an ATP analogue and a soluble acceptor substrate, i. e. a soluble Lipid II analogue are required. In parallel, solution scattering techniques (DLS, SAXS) are to be used to investigate the dynamics of potential ligand-induced conformational rearrangements. These structural data shall be complemented with thorough biochemical characterisation of the enzyme, with the main focus on kinetic and thermodynamic parameters of the catalysed reaction. Finally, phylogeny and sequence variability analyses within different strains of *S. aureus* as well as across many different putative and confirmed GatD/MurT homologues are to be used to identify key residues and modules within the protein complex.

It has been shown that deficiency in peptidoglycan amidation leads to increased susceptibility of methicillin resistant *S. aureus* (MRSA) strains to β -lactam antibiotics and lysozyme [128]. Hence, detailed knowledge of the molecular machinery responsible for this modification could result in the identification of new potential drug targets not only against resistant *Staphylococcus* strains, but also against a wide host of human pathogenic bacteria like *Streptococcus pneumoniae* or *Mycobacterium tuberculosis*. Such new leads in antibiotic development are desperately needed in order to combat the ever-rising problem of multi-drug-resistant bacteria, especially in hospital environments [124].

2.1.7 Work presented in this thesis

Previous work conducted in our laboratory by Volker Niemann, a former PhD student, and two undergraduate students yielded an initially refined unliganded crystal structure of the GatD/MurT complex. I have fully refined this structure and complemented it by a complex structure obtained in the presence of an ATP analogue as well as solution studies on the overall conformation by means of Small-Angle X-ray Scattering (SAXS). The data obtained from these studies, together with functional data generated by Dr. Anna Müller in Prof. Dr. Tanja Schneider's group at the University of Bonn help to paint a compelling, if yet incomplete picture of the likely mechanism of action of this enzyme complex. GatD/MurT adopts an open, crescent-shaped conformation in solution that is likely to transition to a closed state upon substrate binding. It harbours strong structural similarities to class I glutamine amidotransferases and Mur ligases. I have published these findings in *Nature Scientific Reports* in 2018 [79].

The nearly contemporaneous publication of a homologous crystal structure from *Streptococcus pneumoniae* by the group of Prof. André Zapun at the University of Grenoble les Alpes [109] and subsequent discussions with him led to a thorough review of the available literature on the topic. Additionally, sequence databases were analysed as to potential homologues of GatD/MurT in other Gram-positive bacteria. I have used the results of these studies to write a minireview for a special edition of the *International Journal of Medical Microbiology (IJMM)* [139]. Herein I cover the similarities and differences between the GatD/MurT structures of *Staphylococcus aureus* and *Streptococcus pneumoniae* as well as discussing possible links between structural observations and previously published functional data [79], [107]–[109], [128], [137], [140], [141].

These two publications constitute the core of this chapter and are summarised in the Results and Discussion section. The full texts can be found in the appendix.

2.1.7.1 Recent developments

Based on sequence data identified as potential GatD/MurT homologues, expression constructs for homologues from *Streptococcus pyogenes* and *Mycobacterium tuberculosis* were designed. The existing purification protocol was successfully adapted for the two new proteins by Paul Bachmann during his master's thesis under my supervision. Initial *in vitro* activity assays confirmed part of the proteins' function thus paving the road for expanding the field to other organisms. Interspecies similarities and differences between homologous proteins can provide powerful insights into functional relationships.

In the long run, it is crucial to exhaustively describe the mode of substrate peptide engagement and amidation in order to provide a basis for future drug design. This requires a synergistic approach of *in vitro* and *in vivo* experiments as well as structural studies. Functional studies investigating the role of the MurT N-terminus, which I identified as potentially determining factor for specificity during the research on the IJMM review, are on the way in the hands of Dr. Anna Müller in Bonn. As soon as N-terminally truncated versions of MurT are available for expression, they will be used to search for alternative crystal forms.

With multi-drug-resistant bacteria strains on the rise, the need for novel therapeutic strategies has increased in urgency. Hence, detailed structural information on yet undescribed components of

the cell wall biosynthesis machinery of highly pathogenic bacteria, such as *S. aureus* or *M. tuberculosis*, may pave the way towards a new edge in the arms race of antibiotics against resistances.

2.2 Results and Discussion

2.2.1 Publication summaries

This chapter provides summarised versions of the published results obtained for the work conducted on GatD/MurT. The full text reprints are provided in the appendix, please refer to that section for details.

2.2.1.1 Structural Basis of Cell Wall Peptidoglycan Amidation by the GatD/MurT Complex of *Staphylococcus aureus*

E. R. Nöldeke, L. M. Muckenfuss, V. Niemann, A. Müller, E. Störk, G. Zocher, T. Schneider, and T. Stehle, “Structural basis of cell wall peptidoglycan amidation by the GatD/MurT complex of *Staphylococcus aureus*,” *Sci. Rep.*, vol. 8, no. May, p. 12953, **2018**. doi: [10.1038/s41598-018-31098-x](https://doi.org/10.1038/s41598-018-31098-x)

GatD/MurT has been identified as the primary determinant of peptidoglycan isoglutamate amidation [107], [108], [137]. As the absence of this modification severely decreases bacterial viability, the enzyme complex is a potential target for the development of novel antimicrobial agents. A thorough understanding of the molecular mechanism of action of GatD/MurT is required for the rational development of such agents in order to avoid resource-consuming broad screening approaches.

Key to the understanding of any enzymatic mechanism, beyond the characterisation of substrates and products, is the precise architecture of the enzyme’s catalytic core. To this end, a crystal structure of *S. aureus* GatD/MurT was solved to 2.04 Å. The protein complex is composed of three domains: the class I glutamine amidotransferase-type fold of GatD and the Mur ligase middle and C-terminal domains of MurT. GatD forms a stable contact with the MurT C-terminal domain, giving the complex an overall crescent shape. All three domains are mixed α/β folds with a central β -sheet sandwiched between α -helices interconnecting the β -strands.

The MurT middle domain shows high sequence and structure similarity to the middle domain of the canonical Mur ligases C-F and contains a conserved ATP binding pocket. Binding to ATP was confirmed by solving a second GatD/MurT crystal structure to 2.14 Å, obtained from a crystal soaked with the ATP analogue AMPPNP. Modelling of the ligand into the resulting electron density revealed a nucleotide conformation similar to that reported for ATP analogues bound to *S. aureus* MurE and *P. aeruginosa* MurF [142], [143]. Site-directed mutagenesis of key residues involved in AMPPNP binding followed by *in vitro* amidation studies confirmed the involvement of the residues MurT-T60, MurT-E108, and MurT-N265 in GatD/MurT catalysis. Curiously, on the opposite side of the middle domain, facing away from the putative substrate binding cleft, lies a Cys₄-type Zinc finger motif of unknown function.

The MurT C-terminal domain exhibits the same fold observed for that domain in Mur ligases C-F. Residues that were previously suggested to be relevant for catalysis [107] are indeed situated on the surface of the domain and face towards the ATP binding site in the neighbouring middle domain. The C-terminal domain also provides the contact surface for the formation of the stable heterodimeric complex with GatD.

Possessing similarity to class I glutamine amidotransferases (GATases), GatD has a catalytic core arranged around a central cysteine (GatD-C94). In most GATases, this residue is the first of a catalytic triad that also includes a histidine and a glutamic acid to form a proton relay chain and ultimately increase the nucleophilicity of the cysteine. However, GatD lacks the third residue. Instead, the histidine (GatD-H189) is shifted out of position with respect to homologous structures and towards the interface with MurT. MurT then contributes an aspartic acid (MurT-D349) to complete the triad. Mutagenic studies carried out on the triad residues confirmed the importance of this residue for catalysis. A MurT-D349N mutant exhibited severely decreased catalytic rates with little to no *in vitro* turnover when compared to the wild type. Replacement of GatD-C94 with a serine or glycine completely abolished activity. The same observation was previously reported for GatD-H189 [141].

Additionally, all three residues are fully conserved throughout a range of confirmed and putative GatD/MurT homologues covering most of the phylum of the Firmicutes as well as many Actinobacteria. Conformational analysis carried out by means of the CONSURF webserver [67] furthermore classified the entirety of the inner surface of the crescent-shaped complex as more conserved than average. This surface stretches all the way from the catalytic site of GatD to the ATP-binding pocket of MurT. In contrast, the rest of the surface exhibits an above average sequence variability.

The overall crescent-like shape of GatD/MurT results in a large distance between the catalytic sites of GatD and MurT (approx. 40 Å). It is unlikely that nascent ammonia can bridge this distance from GatD to MurT just by diffusion. It appears likely that a domain rearrangement to bring the active sites into close spatial proximity is required for catalysis. As the conformation observed *in crystallo* is stabilised by crystal contacts, the solution conformation was studied by SAXS. Rigid-body refinement of the crystal structure against SAXS profiles resulted in a slight expansion of the domains and a movement of GatD and the MurT middle domain towards one another, decreasing the shortest distance between the two domains from 32 Å to 22 Å. These findings confirm that the GatD/MurT conformation in solution, while being indeed open and elongated, exhibits a certain degree of flexibility concerning the reciprocal orientation of the respective domains.

Prompted by these findings, a putative closed conformation of GatD/MurT was simulated based on the crystal structure of *S. aureus* MurE in complex with its substrate and an ATP analogue. The MurT middle and C-terminal domains were superimposed with their respective counterparts in MurE, while maintaining the geometry between GatD and the MurT C-terminal domain. The resulting hypothetical model exhibits a drastically reduced distance between the active sites of GatD and MurT, from over 40 Å to just over 20 Å. Most of the inner surface of the crescent is now buried in an extended interdomain contact between GatD and the MurT middle domain. The remaining inner surface forms a secluded cavity with the catalytic site of GatD at one end and the ATP binding pocket at the other. Superimposition of MurE's mucopeptide substrate places it nicely into this cavity, between the two active sites. While this model is speculative, it nevertheless provides a plausible picture of how GatD/MurT mediated peptidoglycan amidation might take place. This could involve a concerted reaction starting with the phosphorylation of the target carboxyl group in the mucopeptide from one site, and subsequent substitution with ammonia that is generated on the opposite site of the substrate by GatD, to form the amide product.

The findings reported in Nöldeke et al. provide a solid foundation from which to further the characterisation of GatD/MurT, both as a novel enzyme and as a potential drug target.

2.2.1.2 Unraveling the mechanism of peptidoglycan amidation by the bifunctional enzyme complex GatD/MurT: a comparative structural approach

E. R. Nöldeke and T. Stehle, "Unraveling the mechanism of peptidoglycan amidation by the bifunctional enzyme complex GatD/MurT: A comparative structural approach.," *Int. J. Med. Microbiol.*, vol. 309, no. 6, p. 151334, Sep. 2019. doi: [10.1016/j.ijmm.2019.151334](https://doi.org/10.1016/j.ijmm.2019.151334)

This minireview considers published results on GatD/MurT enzymes identified to date, mainly from *Staphylococcus aureus* and *Streptococcus pneumoniae*, and performs a structure-based comparison. It identifies key similarities and differences as well as important questions to be addressed by future research.

Many efforts have been made into the characterisation of the peptidoglycan-amidating enzyme complex GatD/MurT [79], [107]–[109], [137], [141]. As this enzyme's activity has been linked to changes in bacterial viability and resistance to β -lactam antibiotics, it constitutes a potential target for the development of novel antimicrobial therapeutics. The recent and nearly simultaneous publication of two GatD/MurT crystal structures from *Staphylococcus aureus* [79] and *Streptococcus pneumoniae* [109] have provided immensely valuable initial insights into the architecture of this enzyme complex. The structures exhibit a large number of similarities as well as a few highly intriguing differences.

In both bacteria, GatD/MurT forms a heterodimeric assembly, with GatD exhibiting a class I glutamine-amidotransferase (GATase)-like fold and MurT folding into two domains homologous to Mur ligase middle and C-terminal domains. While in *S. aureus* the assembly is arranged in an open crescent with no direct contact between the MurT middle domain and GatD, the *S. pneumoniae* enzyme is more compact, with the three domains in a triangular arrangement and a direct contact between the MurT middle domain and GatD. Although the overall conformation differs, the single domains superimpose very well, with C- α RMSD values between 1.9 Å and 3.21 Å. Both studies identified a cysteine-containing helix-turn-helix motif in GatD, a neighbouring histidine residue, as well as a canonical ATP-binding site in the MurT middle domain as the key catalytic regions. Based on published functions of structurally homologous proteins, as well as *in vivo* and *in vitro* data on GatD/MurT, the cysteine – histidine pair in GatD, together with a third residue, likely mediates glutamine deamidation and thus the generation of ammonia. Nascent ammonia is then probably used by MurT to amidate the Lipid II substrate in an ATP-dependent reaction.

As structural data published on other members of the Mur ligase family suggest a significant degree of flexibility in the orientation of the middle and C-terminal domains [106], it is likely that the two GatD/MurT crystal structures represent two distinct states of the enzyme. The open conformation observed in the *S. aureus* enzyme (*Sa*GatD/MurT) could depict the resting state of the protein in the absence of substrates. Substrate engagement would then lead to domain rearrangement, generating a closed state similar to that observed for the *S. pneumoniae* enzyme (*Sp*GatD/MurT). Although the distance between the two active sites of GatD and MurT is reduced from around 40 Å in the open

conformation to about 20 Å in the closed state, it remains too large to be bridged by nascent ammonia without considerable diffusive loss. Thus, additional movements are likely involved in the full catalytic cycle, such as the formation of a yet unidentified ammonia channel or the relocation of the substrate from one active site to the other.

One striking difference between *SaGatD/MurT* and *SpGatD/MurT* is the third residue in the GatD active site. In addition to the aforementioned cysteine and histidine, class I GATases contain a glutamate residue. Together, these residues form a catalytic triad, with the cysteine being deprotonated to enhance its nucleophilicity, the histidine acting as a proton relay, and the glutamate acting as the proton sink. Neither *SaGatD* nor *SpGatD* contain this residue. Instead, two aspartic acids in the C-terminal domain of MurT (MurT-D349, *S. aureus*) and in a neighbouring helix of GatD (GatD-D32, *S. pneumoniae*) have been suggested to fulfil the role of the proton sink. Crucially, both residues are conserved and could be brought into proximity of the central histidine by minor rearrangements in either crystal structure. However, the intermolecular triad postulated in the *SaGatD/MurT* is the only one forming a linear arrangement amenable to concerted proton transfer. *In vitro* and *in vivo* mutagenesis data on *SaMurT*-D349 and its homologue in *S. pneumoniae* confirm its importance for catalysis. Due to a different orientation of the helix containing the intramolecular triad candidate in *S. aureus*, this residue was not investigated by mutagenesis. Clearly, further research is required to resolve this matter, including obtaining structures of GatD/MurT complexes from other organisms.

MurT contains a structural motif on the side of its middle domain that is not found in any other known Mur ligase. In both *S. aureus* and *S. pneumoniae* an insertion of around 50 residues was identified. The motif harbours the structural hallmarks of RanBP type Zn ribbons with four cysteines coordinating a Zinc ion. While such an ion was modelled only in *SaMurT*, its presence was inferred from acetylation experiments and structural similarity in *SpMurT*. RanBP Zinc ribbons have been shown to mediate protein-RNA and protein-protein interactions [144], [145]. It is thus conceivable that this motif may play a scaffolding role in guiding GatD/MurT to its membrane-bound substrate by interacting with other components of the peptidoglycan biosynthesis pathway. Interestingly, this motif is the least well conserved portion of the enzyme complex. While it is always present in sequences from bacteria belonging to the Bacilli class, in most Actinobacteria, the insertion contains no cysteines, thus precluding it from forming the same Zinc finger. Actinobacteria of the *Streptomyces* genus have the four cysteines, but the insertion is much shorter than in Bacilli. These significant differences across taxa may stem from distinct mechanistical requirements due to the very different cell wall compositions. Elucidation of the structures of these alternate versions of the MurT middle domain insertion may help to shed light on its role in peptidoglycan amidation.

The largest difference between MurT and the other canonical Mur ligases is the N-terminal domain. MurE-F have an N-terminal domain of 90-100 residues arranged around a Rossmann fold [106]. This domain serves to bind the uridine moiety of the substrate (UDP-MurNAc-bound peptides) and to correctly position it in the active site. In contrast, MurT has an N-terminal segment of 30-40 residues of high conformational variability. No interpretable electron density was reported for the region in either *S. aureus* or *S. pneumoniae*. *SaGatD/MurT* was shown to amidate soluble UDP-MurNAc-pentapeptide (Park's nucleotide) *in vitro* as well as Lipid I and Lipid II with similar efficiency [107] but no soluble amidated peptidoglycan precursors were observed *in vivo* [108]. It is thus likely that the nature of the molecule to which MurNAc is anchored (UDP in Park's nucleotide or bactoprenol in Lipid

I / II) is relevant for selectivity but plays no distinct role in the catalytic mechanism. It is possible that the N-terminus of MurT plays a role in orienting the active site towards the substrate like in the other Mur ligases. However, the substrate being membrane-anchored, MurT may do so by interacting with the membrane or with membrane-embedded components of the peptidoglycan biosynthetic machinery rather than with the substrate itself. Indeed, it was shown that *S. aureus* membrane fractions and not cytosolic extracts exhibit Lipid II amidating activity *in vitro* [107]. Perhaps the MurT N-terminus, which is predicted to form an α -helix and contains a high density of positive charges, interacts with the phospholipids of the cytosolic leaflet of the plasma membrane, thus establishing the sterical constraints required for GatD/MurT positioning. Future investigations into GatD/MurT mutants lacking the MurT N-terminus as well as the investigation of activity patterns of homologous enzymes will help to characterise the function of this domain in depth.

The comparison and review of available structural, biochemical, and microbiological data on GatD/MurT performed by Nöldeke and Stehle has identified features in the enzyme complex that may have gone unnoticed by analysing data on only one instance of the protein. This information will help to guide future research into *Sa*GatD/MurT and *Sp*GatD/MurT, as well as into homologous enzymes from other organisms. Hopefully, this will help in the design of a specific inhibitor molecule that can be used as a basis for novel antibiotics to step into the breach opened by bacterial resistances to existing drugs.

2.3 Ongoing Research

Building on the results published in Nature Scientific Reports [79] and the subsequent mini-review for IJMM [139], various follow-up studies are under way. These include a search for different crystal forms, co-crystallisation with a synthetic Lipid I or Lipid II analogue, and investigation of the role of the MurT N-terminus. Additionally, efforts are being made to characterise putative homologous enzymes from other bacteria.

2.3.1 Production and characterisation of GatD/MurT from *Mycobacterium tuberculosis* and *Streptococcus pyogenes*

Following the identification of potential homologues in many different Gram-positive bacteria [79], [107], [137], [138], it became apparent that a wealth of information could be derived from comparing *S. aureus* GatD/MurT to its cognate enzymes from other organisms. In order to further our understanding of the structure and mechanism of action of GatD/MurT, two previously unstudied putative homologues were selected, namely the genes Spy0877 and Spy0876 from *Streptococcus pyogenes* strain MGAS10270 and Rv3712 and Rv3713 from *Mycobacterium tuberculosis* strain H37Rv. All sequences were retrieved from the EnsemblBacteria database (<https://bacteria.ensembl.org/> [146]). The rationale behind this was, firstly, to ascertain whether these genes produce soluble proteins, and secondly, whether the protein products are comparable to the already characterised enzymes from *S. aureus* and *S. pneumoniae*. Thirdly, and most importantly, the comparison of homologous enzymes from very different bacteria, such as *M. tuberculosis*, and quite similar ones, such as *S. pyogenes*, may help to better understand the specificities and the mechanism of catalysis.

Coding sequences were codon-optimised for *E. coli* expression and synthesised by a commercial supplier (BioCat GmbH, Heidelberg, Germany, see section 3.4.2).

As his master thesis project, Paul Bachmann successfully expressed both complexes and adapted the existing purification strategy. Both complexes express well in *E. coli*, but variants with N-terminal truncations in MurT exhibited much better solubility than the full-length constructs. The *S. pyogenes* protein (SpGatD/MurT) could only be purified as a truncated protein. When the full-length construct was used, the protein nearly quantitatively localised to the insoluble fraction after cell lysis. However, both full-length and truncated versions of *M. tuberculosis* GatD/MurT (*Mtb*GatD/MurT) and the truncated *Sp*GatD/MurT showed similar behaviour in analytical SEC as the *S. aureus* complex (*Sa*GatD/MurT, Figure 2-2).

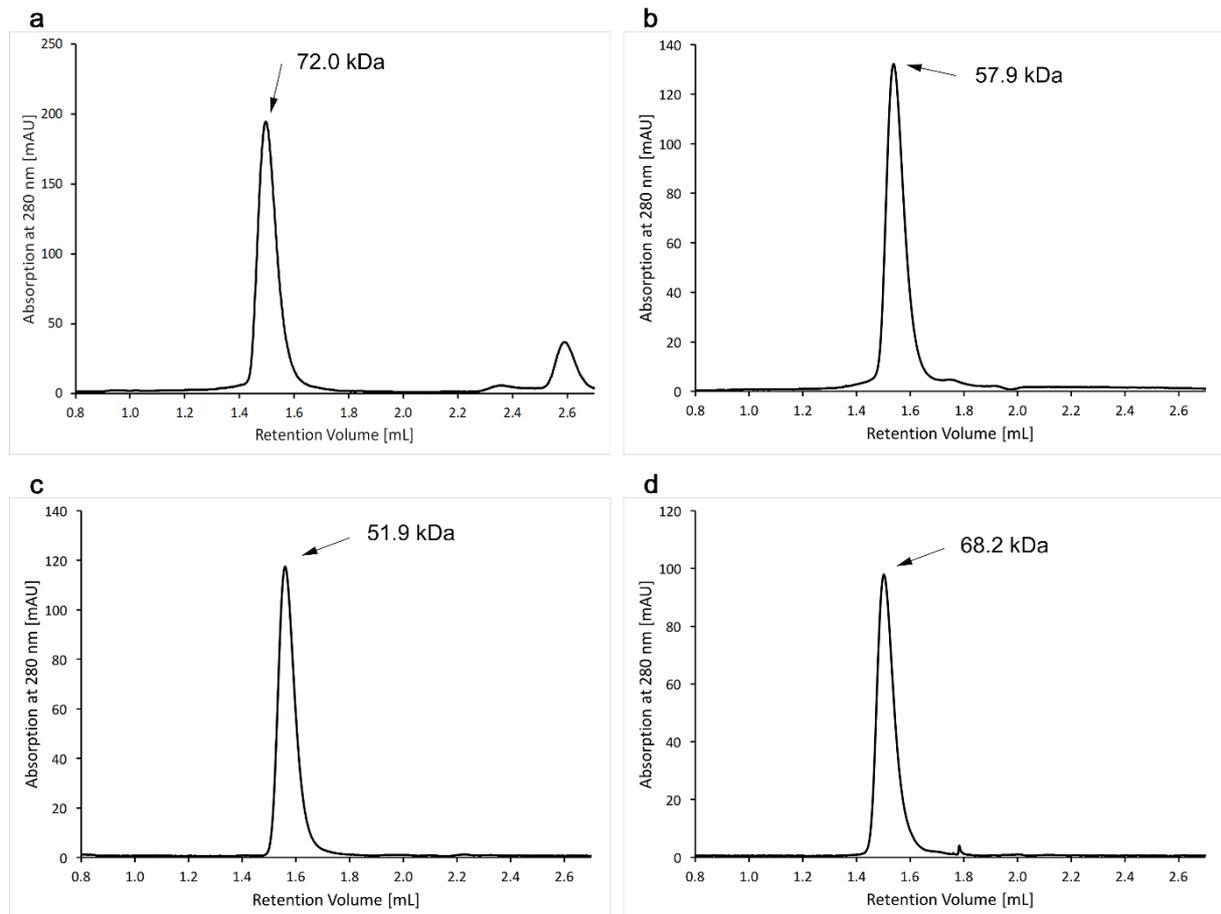


Figure 2-2: Analytical SEC traces of purified GatD/MurT homologues. Enzymes from *S. aureus* (a), *M. tuberculosis* (b), *M. tuberculosis* with N-terminal truncation in MurT (c), and *S. pyogenes* with N-terminal truncation in MurT (d). The SEC profiles of the new constructs are compatible with that of the *S. aureus* complex. All have apparent molecular weights of about 10 kDa less than the respective expected values (78.8 kDa for *S. aureus*, 68.4 kDa for *M. tuberculosis* full-length, 64.9 kDa for *M. tuberculosis* with N-terminal truncation, and 75.8 kDa for *S. pyogenes* with N-terminal truncation).

Circular dichroism (CD) spectra of full-length *Mtb*GatD/MurT and truncated *Spy*GatD/MurT compared with the published spectrum for *Sa*GatD/MurT [79] show similar magnitudes of mean residue weight-corrected signals, suggesting a similar compactness of fold. However, the spectra suggest a distinct secondary structure composition, with a higher percentage of α -helices (Figure 2-3). It remains to be investigated whether these differences originate in misfolding due to the process of recombinant protein production or whether they are rooted in the protein's function. Potential differences in homologous enzymes from different organisms would provide crucial information towards understanding the mechanism of GatD/MurT action.

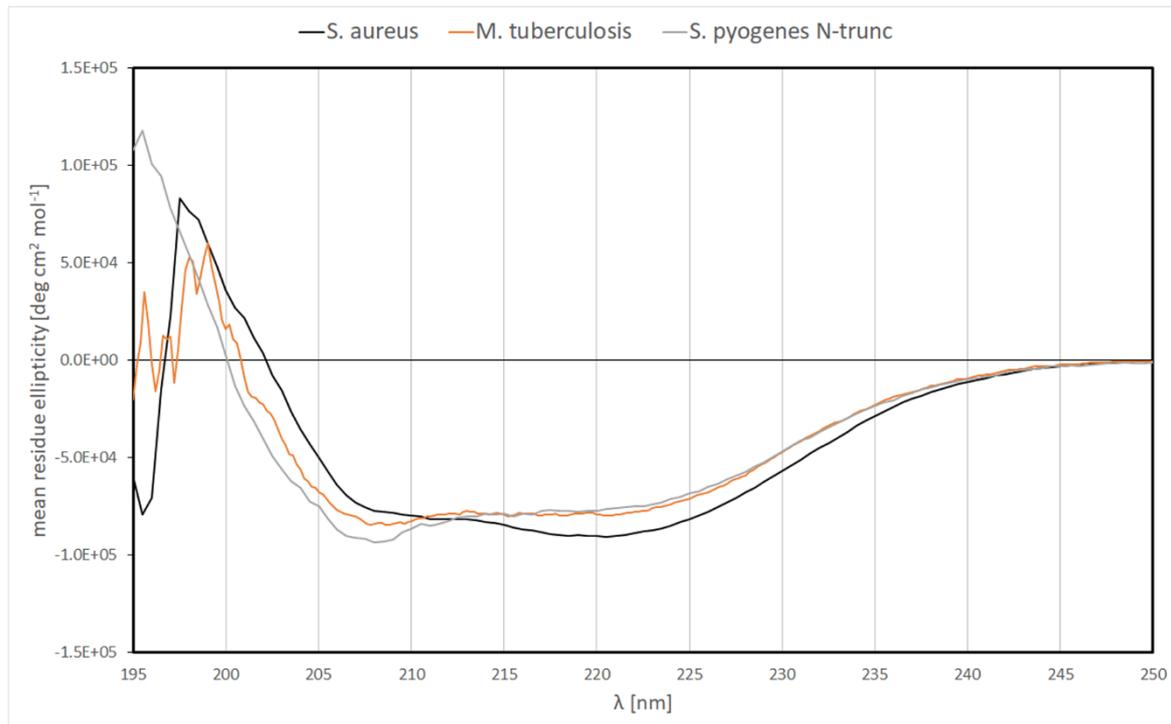


Figure 2-3: CD spectra of purified GatD/MurT homologues. Complexes from *S. aureus* (black), *M. tuberculosis* (orange), and *S. pyogenes* with N-terminal truncation (grey) were compared. The magnitude of mean residue ellipticity (MRE) for the new constructs is comparable to that of the original *S. aureus* one. However, the spectra show distinct differences, suggesting differences in secondary structure composition.

2.3.1.1 *MtbGatD/MurT* and *SpyGatD/MurT* exhibit glutaminase activity *in vitro*.

Paul Bachmann also studied Glutaminase activity based on a protocol published by Francisco Leisico et al. [141]. The experiments were performed with the help of Dr. Vincent Truffault at the Max Planck Institute for Developmental Biology in Tübingen and clearly show glutaminase activity for *MtbGatD/MurT* as well as *SpyGatD/MurT* (Figure 2-4). Although the data are preliminary and more stringent experiments are required to fully describe the catalytic activities, the results strongly suggest that these two GatD/MurT homologues harbour the same catalytic activity as the *S. aureus* enzyme. The same decrease of glutamine γ -CH₃ signals and contemporaneous increase of glutamate γ -CH₃ signals was observed in the reaction catalysed by *MtbGatD/MurT* as in that catalysed by *SaGatD/MurT*. However, the speed of the reaction is much slower. In order to fully characterise the enzyme's kinetics, more thorough experiments are required.

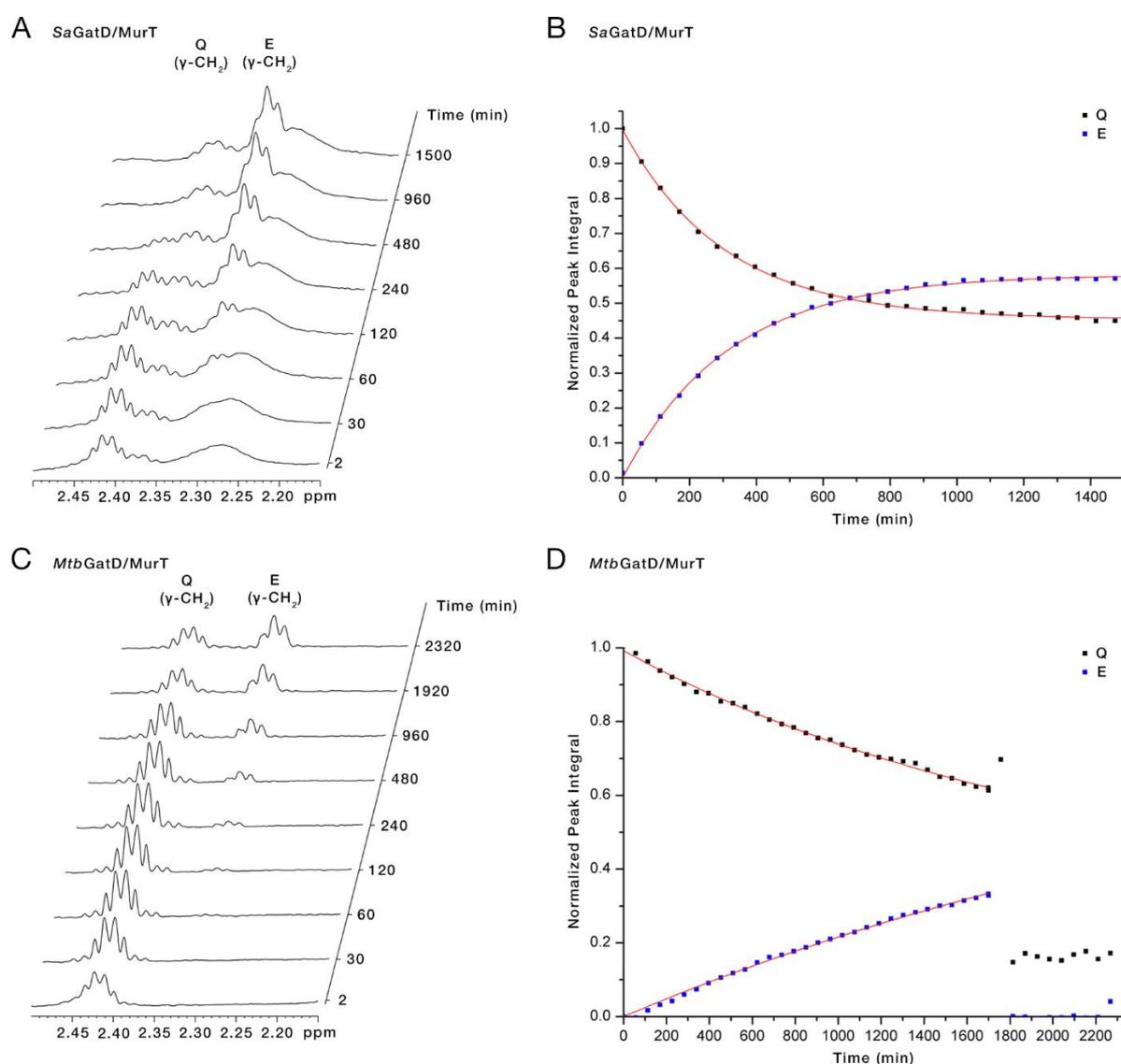


Figure 2-4: Glutaminase activity of *SaGatD/MurT* and *MtbGatD/MurT* monitored by NMR. **(A)** Sequential ¹H NMR spectra of the region containing the signal for glutamine and glutamate γ -CH₂ (2.42 and 2.31 ppm, respectively). Glutamine deamination was catalysed by *SaGatD/MurT*. **(B)** Normalised peak integrals of glutamine and glutamate γ -CH₂ signals over time. The decrease of glutamine signal coincides with the increase of the glutamate signal. **(C)** As in (A), for the reaction catalysed by *MtbGatD/MurT*. **(D)** As in (B), for *MtbGatD/MurT*. After 1800 min the signal lock was lost, and no more meaningful data could be recorded.

Taken from: Paul Bachmann: "Towards a Functional and Structural Comparison of the Enzyme Complex GatD/MurT among Gram-positive Bacteria" – Master thesis, Tübingen, March 2019.

The experiment performed with *SpyGatD/MurT* cannot be correctly quantified, as a drift in the signal lock decreased the spectra quality significantly. However, the increase of the glutamate γ -CH₃ signal over time can clearly be discerned, confirming the catalytical activity. This marks a major breakthrough in the GatD/MurT field, as it appears more and more likely that indeed many Gram-positive bacteria employ this enzyme, thus boosting its potential as an antimicrobial target.

As for *MtbGatD/MurT*, a thorough quantification of enzyme kinetics is required.

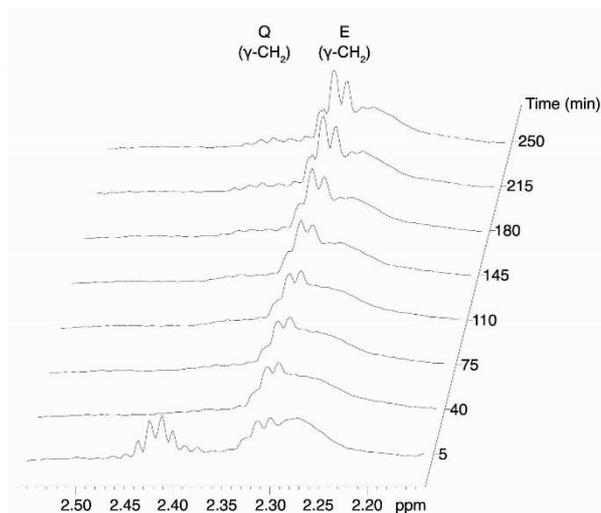


Figure 2-5: Glutaminase activity of *SpyGatD/MurT* monitored by NMR. Sequential ¹H NMR spectra of the region containing the signal for glutamine and glutamate γ -CH₂ (2.42 and 2.31 ppm, respectively) are shown. Although the quality of the spectra is lost after few minutes, the increase of the glutamate signal at 2.3 ppm over time is clearly discernible. The immediate drop to zero of the glutamine signal at 2.42 ppm is likely exaggerated by the loss of the locking signal.

2.3.2 Future prospects

The ultimate goal of the project is to obtain a crystal structure of the complete, catalytically-competent complex of GatD/MurT with all three substrates: glutamine, ATP, and Lipid II or analogues thereof. To date, soaking trials using the well-characterised crystal form of *SaGatD/MurT* [79] and a minimal substrate consisting of a commercially available *S. aureus* pentapeptide stem (D-Ala-D-Glu- γ -L-Lys-D-Ala-D-Ala) failed to yield any meaningful ligand electron density. As discussed in [79] and [139], it is likely that the established crystal form reflects a non-catalysis-competent conformation of the enzyme. It is likely that a major conformational change is required for catalysis. Thus, co-crystallisation with a pre-saturated enzyme complex is required. Crystallisation screens performed with the commercial pentapeptide as well as with purified Park's nucleotide from *S. aureus* failed to identify any novel crystal forms. It is likely that the sugar moiety of Lipid II is required for GatD/MurT activity and that the UDP function in Park's nucleotide has a detrimental effect (Dr. Anna Müller, University of Bonn, personal communication). Until recently, no soluble Lipid I or Lipid II analogues were available in purity sufficient for crystallographic studies and quantities suitable for a screening approach. With the development of a modular synthetic pathway for a version of Lipid I with shortened prenyl tail by Lukas Wingen from the lab of Prof. Dr. Dirk Menche (University of Bonn), such molecules have now become available. Initial feasibility studies are underway at the hands of Simon Völpe, a new PhD student who is continuing this project. Currently only a Lipid I version of the molecule is available, but experiments to add the GlcNAc moiety for a Lipid II version are ongoing.

With these molecules available, the search for a new crystal form reflecting the active catalytic complex can be resumed. Obtaining the full picture of the catalytic mechanism will enable to launch a structure-guided design studies to identify potential inhibitor molecules. Future studies will also include variants of GatD/MurT that lack the N-terminal section of MurT [139] in order to understand the role this yet unstudied portion of the protein.

An effective and, more crucially, selective inhibitor for GatD/MurT would be a very powerful drug candidate for the narrow spectrum treatment of infections by Gram-positive bacteria, including MRSA and *Mycobacterium tuberculosis*, both of which represent severe challenges to public health. In contrast to COVID19, which likely constitutes a pandemic wave that may last up to several years, we cannot expect the threats by pathogenic bacteria to subside at all. The more antibiotics become ineffective, the greater the risks will become. Only by exploring all opportunities and all potential new drug targets will we be able to devise an effective strategy to combat this incumbent threat.

3 MATERIAL AND METHODS

3.1 Instruments, chemicals, and kits

Table 3-1: Instruments used in this study.

Instrument	Manufacturer
ÄKTA BASIC FPLC	GE Healthcare (Uppsala, Sweden)
ÄKTA Ettan LC FPLC	GE Healthcare (Uppsala, Sweden)
ÄKTAprime plus FPLC	GE Healthcare (Uppsala, Sweden)
Analytical balance Genius	Sartorius (Göttingen, Germany)
Autoclave Systec VX-95	Systec (Wettenberg, Germany)
Centrifuge 5415D	Eppendorf (Hamburg, Germany)
Centrifuge Heraeus, Multifuge 1 L-R	Thermo Fisher Scientific (Waltham, MA, USA)
Centrifuge Sorvall RC6+ equipped with SS-34 and SLC-4000 rotors	Thermo Fisher Scientific (Waltham, MA, USA)
Crystal Gryphon LCP liquid handling robot	Art Robbins Instruments (Sunnyvale, CA, USA)
Digital microscope	Leica (Wetzlar, Germany)
Digital sonifier 450	Branson Ultrasonics (Danbury, CT, USA)
Emulsiflex C3 high pressure homogenizer	Avestin (Ottawa, ON, Canada)
Freedom EVO® liquid handling robot	Tecan Group Ltd. (Männedorf, Switzerland)
HisTrap FF Crude affinity column (5 mL)	GE Healthcare (Uppsala, Sweden)
Ice machine AF-80	Scotsman (Milano, Italy)
iCycler™, Thermal Cycler	BioRad (Hercules, CA, USA)
Intelli-Mixer	NeoLab. (Heidelberg, Germany)
J-720 Spectrometer	Jasco (Easton, MD, USA)
MAR345 dtb image plate detector	marXperts GmbH (Norderstedt, Germany)
MicroMax-007HF X-ray generator	Rigaku (Tokyo, Japan)
MilliQ Biocel A10	Millipore (Schwabach, Germany)
NanoDrop ND-100 spectral photometer	Thermo Fisher Scientific (Waltham, MA, USA)
Peristaltic Pump	EconoPump BioRad (Hercules, CA, USA)
pH electrode PB-11	Sartorius (Göttingen, Germany)
Pilatus3 R 300K HPC detector	Dectris (Baden-Daettwil, Switzerland)
Pipets Pipetman L	Gilson (Middleton, WI, USA)
PowerPac HC (Electrophoresis)	BioRad (Hercules, CA, USA)
QuantStudio 5 Real-Time PCR System	Thermo Fisher Scientific (Waltham, MA, USA)
Scanner LiDE 600F	Canon (Tokyo, Japan)
Smart Spec™ Plus, Spectrophotometer	BioRad (Hercules, CA, USA)
Thermomixer comfort	Eppendorf (Hamburg, Germany)
Vortex Genie 2™	Bender & Hobein (Zurich, Switzerland)
Xeuss 2.0 X-ray diffractometer	Xenocs SAS (Sassenage, France)
Zetasizer Nano ZS	Malvern Instruments (Malvern, UK)

Table 3-2: Chromatography columns and media used in this study.

Column	Source
HiLoad 16/60 packed with Superdex 200 column material	GE Healthcare (Uppsala, Sweden)
HiLoad 26/60 packed with Superdex 200 column material	GE Healthcare (Uppsala, Sweden)
Ni-NTA agarose	Qiagen (Hilden, Germany)
PD-10 desalting column	GE Healthcare (Uppsala, Sweden)
HiPrep 26/10 desalting column	GE Healthcare (Uppsala, Sweden)
His SpinTrap™ centrifugal IMAC columns	GE Healthcare (Uppsala, Sweden)
Superdex 200 Increase 3.2/300	GE Healthcare (Uppsala, Sweden)

Table 3-3: Enzymes used in this study.

Enzyme	Source
Benzonase® Nuclease	Sigma-Aldrich (St. Louis, MO, USA)
Bsal-HF v2	New England Biolabs (Ipswich, MA, USA)
DpnI	New England Biolabs (Ipswich, MA, USA)
ExactRun proofreading polymerase	Genaxxon bioscience (Ulm, Germany)
Lysozyme from chicken egg white	Sigma-Aldrich (St. Louis, MO, USA)
NcoI-HF	New England Biolabs (Ipswich, MA, USA)
Q5® High-Fidelity DNA polymerase	New England Biolabs (Ipswich, MA, USA)
rSAP	New England Biolabs (Ipswich, MA, USA)
T4 DNA ligase (1 U/μL)	Thermo Fisher Scientific (Waltham, MA, USA)
XhoI-HF	New England Biolabs (Ipswich, MA, USA)

Table 3-4: Reagents used in this study.

Reagent	Source
CutSmart® buffer (10X)	New England Biolabs (Ipswich, MA, USA)
dNTP mix – Sodium salt (10 mM)	Genaxxon bioscience (Ulm, Germany)
ExactRun buffer (5X)	Genaxxon bioscience (Ulm, Germany)
PageRuler Prestained Protein Ladder	Thermo Fisher Scientific (Waltham, MA, USA)
PageRuler Unstained Protein Ladder	Thermo Fisher Scientific (Waltham, MA, USA)
O'GeneRuler DNA Ladder Mix	Thermo Fisher Scientific (Waltham, MA, USA)
Orange DNA loading dye (6X)	Thermo Fisher Scientific (Waltham, MA, USA)
T4 DNA ligase buffer (10X)	Thermo Fisher Scientific (Waltham, MA, USA)

Table 3-5: Reagents and media used for insect cell culture.

Reagent	Source
Sf-900 II SFM insect cell medium	(Gibco, Thermo Fisher Scientific, Waltham, MA, USA)
Express Five™ SFM insect cell medium	(Gibco, Thermo Fisher Scientific, Waltham, MA, USA)
Grace's Insect Medium	(Gibco, Thermo Fisher Scientific, Waltham, MA, USA)
Penicillin-Streptomycin	(Gibco, Thermo Fisher Scientific, Waltham, MA, USA)
Cellfectin® II	(Gibco, Thermo Fisher Scientific, Waltham, MA, USA)
Fetal Bovine Serum	(Gibco, Thermo Fisher Scientific, Waltham, MA, USA)

Table 3-6: Chemicals used in this study.

Chemical	Source
Adenylyl-imidodiphosphate (AMPPNP)	Roche Diagnostics GmbH (Mannheim, Germany)
Bromphenol blue, Na-salt	Serva Electrophoresis GmbH (Heidelberg, Germany)
Ethanol ≥ 99.5 %, Ph.Eur., extra pure	Roth (Karlsruhe, Germany)
Ethylenediamine tetraacetic acid (EDTA) disodium salt dihydrate	Roth (Karlsruhe, Germany)
L-Glutathione oxidized disodium salt	Sigma-Aldrich (St. Louis, MO, USA)
L-Glutathione reduced	Sigma-Aldrich (St. Louis, MO, USA)
Glycerol, ROTIPURAN® (≥ 98 %)	Roth (Karlsruhe, Germany)
Hydrogen peroxide 30 %, ROTIPURAN®	Roth (Karlsruhe, Germany)
LB Broth (Lennox)	Roth (Karlsruhe, Germany)
β-Mercaptoethanol	Sigma-Aldrich (St. Louis, MO, USA)
Phenylmethyl Sulfonylfluoride (PMSF)	Sigma-Aldrich (St. Louis, MO, USA)
Potassium phosphate monobasic	Sigma-Aldrich (St. Louis, MO, USA)
Potassium phosphate dibasic	Sigma-Aldrich (St. Louis, MO, USA)
Sodium acetate	Sigma-Aldrich (St. Louis, MO, USA)
Triton X-100	Sigma-Aldrich (St. Louis, MO, USA)
Urea	Roth (Karlsruhe, Germany)
Acrylamide-Bisacrylamide solution Rotiphorese® Gel 30	Roth (Karlsruhe, Germany)
Adenosine 5'-triphosphate (ATP) magnesium salt	Sigma-Aldrich (St. Louis, MO, USA)
Agar-Agar, bacteriological	Roth (Karlsruhe, Germany)
Agarose NEEO ultra-quality	Roth (Karlsruhe, Germany)
Ala-D-γ-Glu-Lys-D-Ala-D-Ala (≥ 97 % (HPLC))	Sigma-Aldrich (St. Louis, MO, USA)
Ammonium chloride	Sigma-Aldrich (St. Louis, MO, USA)
Ammonium persulfate (APS)	Roth (Karlsruhe, Germany)
Ammonium sulfate	Roth (Karlsruhe, Germany)
Ampicillin sodium salt	Roth (Karlsruhe, Germany)
Antifoam 204	Sigma-Aldrich (St. Louis, MO, USA)
cComplete™, EDTA-free Protease Inhibitor Cocktail	Roche Diagnostics GmbH (Mannheim, Germany)
D-(+)-Glucose monohydrate	Sigma-Aldrich (St. Louis, MO, USA)
D ₂ O (99.8 % ² H atom)	CortecNet (Voisins-Le-Bretonneux, France)

Chemical	Source
DL-Dithiothreitol (DTT)	Sigma-Aldrich (St. Louis, MO, USA)
Glycerol, > 99 %	Sigma-Aldrich (St. Louis, MO, USA)
HEPES ≥ 99.5 % p.a.	Roth (Karlsruhe, Germany)
Imidazole ≥ 99 %	Sigma-Aldrich (St. Louis, MO, USA)
InstantBlue stain	Expedeon (Cambridgeshire, UK)
Isopropanol puriss.	Sigma-Aldrich (St. Louis, MO, USA)
Isopropyl β-D-thiogalactopyranoside (IPTG)	Peqlab (Erlangen, Germany)
Kanamycin sulfate	Sigma-Aldrich (St. Louis, MO, USA)
LB Broth (Miller)	Sigma-Aldrich (St. Louis, MO, USA)
Magnesium chloride hexahydrate ≥ 99 %	Sigma-Aldrich (St. Louis, MO, USA)
Magnesium sulfate heptahydrate	Roth (Karlsruhe, Germany)
N-Z-Amine® A	Sigma-Aldrich (St. Louis, MO, USA)
Potassium phosphate monobasic	Sigma-Aldrich (St. Louis, MO, USA)
Rotiphorese. 10x Running Buffer	Roth (Karlsruhe, Germany)
SDS ≥ 99 %	Roth (Karlsruhe, Germany)
Sodium chloride ≥ 99.8 %	Roth (Karlsruhe, Germany)
Sodium hypochlorite, 13 % (w/v)	AppliChem (Darmstadt, Germany)
Sodium phosphate dibasic	Sigma-Aldrich (St. Louis, MO, USA)
Sodium sulfate decahydrate	Roth (Karlsruhe, Germany)
TEMED, puriss.	Serva (Heidelberg, Germany)
Tris(2-carboxethyl)phosphine (TCEP) hydrochloride	Sigma-Aldrich (St. Louis, MO, USA)
Trizma base, ≥ 99 %	Sigma-Aldrich (St. Louis, MO, USA)
Tryptone/Peptone ex casein	Roth (Karlsruhe, Germany)
Yeast extract	Roth (Karlsruhe, Germany)
α-Lactose monohydrate	Sigma-Aldrich (St. Louis, MO, USA)

Table 3-7: Commercial kits used in this study.

Kit	Source
Wizard® Plus SV Minipreps DNA Purification System	Promega (Madison, WI, USA)
Protein Thermal Shift™ Dye Kit	Thermo Fisher Scientific (Waltham, MA, USA)
Wizard® SV Gel and PCR Clean-Up System	Promega (Madison, WI, USA)
Proti-Ace™ Kit 1+2	Hampton Research (Aliso Viejo, CA, USA)

Table 3-8: Solubility screens.

Screen	Source
Solubility & Stability Screen 1	Hampton Research (Aliso Viejo, CA, USA)
Solubility & Stability Screen 2	Hampton Research (Aliso Viejo, CA, USA)

Table 3-9: Commercial crystallisation screens used in this study.

Screen	Manufacturer
Wizard 1 - 4	Jena Biosystems (Jena, Germany)
JCSG+	Molecular Dimensions Limited (Maumee, OH, USA)
Morpheus	Molecular Dimensions Limited (Maumee, OH, USA)
Peglon	Hampton Research (Aliso Viejo, CA, USA)
Crystal Screen 1 & 2	Hampton Research (Aliso Viejo, CA, USA)
Additive Screen HT	Hampton Research (Aliso Viejo, CA, USA)

3.2 Standard protocols

Table 3-10: List of standard buffers.

Name	Composition	
50 x TAE	2 M	Tris pH 8.0
	6 % (w/v)	Acetic acid
	5 mM	EDTA
4x Protein sample buffer	30 % (w/v)	Glycerol
	0.3 M	Tris pH 6.8
	1.5 %	SDS
	12 mM	EDTA
	0.03 % (w/v)	Bromophenol blue
	± 6 % (v/v)	β-mercaptoethanol

3.2.1 Cloning

The primers used for cloning and later sequencing of expression constructs were manually designed and synthesized upon commission by either biomers.net (Ulm, Germany) or Microsynth AG (Balgach, Switzerland). Primers intended for sequencing are listed in Table 3-11 (GatD/MurT) and Table 3-12 (AHK5), primers for cloning in Table 3-13 (GatD/MurT) and Table 3-14 (AHK5). Primer ID numbers refer to the primer database on the AK Stehle server. Melting temperatures were calculated with Equation (1-1) [147].

$$T_m = \left(81.5 + 0.41 \cdot (\%GC) - \frac{675}{N} - (\%mismatch) \right) ^\circ\text{C} \quad (3-1)$$

Table 3-11: Primers used for sequencing of GatD/MurT constructs. Primers without database number were stored at SeqLab (Göttingen, Germany) for sequencing.

ID	name	Sequence	Length / % GC / T _m (°C)			target for sequencing
50	murT_fwd_4	5'-atgccatgaatttgaacaaagtacg-3'	26	38	71	<i>Sa</i> MurT nt 560 f
	GatD_FW	5'-gtcgaacagaaagtaatcg-3'	19	42	49	<i>Sa</i> GatD f
	mytb_MT_FW1	5'-gaatctgagccgtgatcag-3'	19	53	53	<i>Mtb</i> MurT nt 270 f
	mytb_MT_RV1	5'-ataacatccggcagcac-3'	17	53	53	<i>Mtb</i> GatD nt 30 r
N/A	pET_up	5'-tgcgtccggcgtaga-3'	15	67	57	pET vector start
	spy_MT_FW1	5'-tcagatggatcgctatgg-3'	18	50	51	<i>Spy</i> MurT 490 f
	spy_MT_RV1	5'-tacagatgcgcaatggtc-3'	18	50	53	<i>Spy</i> GatD nt 60 r
	T7 term	5'-gctagttattgctcagcgg-3'	19	53	54	pET vector end

Table 3-12: List of primers used for sequencing of AHK5 constructs. Primers without database number were stored at Seqlab (Göttingen, Germany) for sequencing.

ID	name	Sequence	Length / % GC / Tm (°C)			target sequence
16	AHK5_160_f	5'-ctacgaagagtttgatacagttgc-3'	24	42	70	<i>E. coli</i> bp 470
17	AHK5_350_f	5'-gtgaagataatgcggtacgcaaag-3'	24	46	72	<i>E. coli</i> bp 1050
27	AHK5sf9_500f	5'-attggaagcagaagactctctc-3'	22	45	69	sf9 bp 500
28	AHK5sf9_1101f	5'-ggccaaactgagagaggacaatg-3'	23	52	74	sf9 bp 1101
29	AHK5sf9_1515f	5'-attacaagtaatttccgagccgtc-3'	24	42	70	sf9 bp 1515
30	AHK5sf_2000f	5'-tgccgcacaagtgatgctaac-3'	22	55	73	sf9 bp 2000
	pET_up	5'-tgcgtccggcgtaga-3'	15	67	64	pET vector start
	T7 term	5'-gctagtattgctcagcgg-3'	19	53	68	pET vector end

Table 3-13: Primers used for cloning of GatD/MurT constructs.

ID	name	Sequence	nt / % GC / Tm (°C)			purpose
85	42_Cterm_*_FW	5'-gcctgagtgcccgttagcatcatcaccattagctcgag-3'	41	54	85	QC tag*
86	42_Cterm_*_RV	5'-ctcgagctaattggtgatgatgctaacgggcactcaggc-3'	41	54	85	QC tag*
89	42_N-trunc_FW	5'-gattggtggtctggtggccatggccctggatcgtagtattctgc-3'	44	57	87	QC NcoI
90	42_N-trunc_RV	5'-cccagctggcgcagaatactacgatccagggccatggccacc-3'	42	64	89	QC NcoI
91	44_Bsal_FW1	5'-ctttaagaaggagggtctccatgaaaatgaaaccctgctgggc-3'	44	48	83	QC Bsal
92	44_Bsal_RV1	5'-gccagcagggttttcatttcatgggagacctcttctaaag-3'	44	48	83	QC Bsal
93	44_Cterm_*_FW	5'-cttaacaaagtgtagcaccatcaccatcactagctcagtgctgg-3'	45	47	83	QC tag*
94	44_Cterm_*_RV	5'-gtgatggtgatggtgctacactttgtaaagctcgtttgc-3'	41	46	82	QC tag*
95	44_N-trunc_FW	5'-ctatccgggcccctggccatggcgtgcatgataaag-3'	35	66	86	QC NcoI
96	44_N-trunc_RV	5'-gaatatcttatcgcacgcatggccaggcggcccgg-3'	37	62	86	QC NcoI

Table 3-14: List of primers used for cloning of AHK5 constructs.

ID	name	Sequence	nt / % GC / Tm (°C)			purpose
4	AHK5_coli_AF	5'-ccggtctcacatggtagtgagatggaactg-3'	32	50	81	M1 fwd
5	AHK5_coli_BF	5'-ccggtctcacatggtagcagccagtaacc-3'	31	58	84	G72 fwd
6	AHK5_coli_CF	5'-ccggtctcacatgcaatccgctccggtac-3'	29	62	84	Q203 fwd
8	AHK5_coli_3F1	5'-ccggtctcacatgctggaattgaacaaaccatc-3'	34	47	81	S352 fwd
9	AHK5_coli_3F2	5'-ccggtctcacatgaccgaagaaccatgcg-3'	30	57	82	T361 fwd
10	AHK5_coli_3R1	5'-ccgcgccgcttattcacgtaactggcc-3'	29	62	84	E342 rev
11	AHK5_coli_3R2	5'-ccgcgccgcttaggtttctcggtaatg-3'	29	59	82	T364 rev
12	AHK5_coli_2R	5'-ccgcgccgcttaactacggcccactttg-3'	29	66	85	S617 rev
45	AHK5pETM_C3Af	5'-cagggcgccatgtagctgagatggaac-3'	29	59	79	QC C3A
46	AHK5pETM_C3Ar	5'-gttccatctcagctaccatggcgccctg-3'	29	59	79	QC C3A
69	AHK5 V332*fwd	5'-gtcactgatcagtgagtaaacgcgagaaaatggcc-3'	36	50	80	QC V332*
70	AHK5 V332*rev	5'-cgcgttccactcactgatcagtgacttccatccc-3'	35	54	82	QC V332*
71	AHK5 E353*fwd	5'-gccatggaatcgtaattgaacaaaaccatccacattacc-3'	39	41	78	QC E353*
72	AHK5 E353*rev	5'-gatggtttgttaactacgattccatggctttgctgctacc-3'	40	43	80	QC E353*

ID	name	Sequence	nt / % GC / Tm (°C)			purpose
77	AHK5 H376*fwd	5'-ggcgactatgtcttaggaaatccgttcccctctgagcg-3'	38	55	84	QC H376*
78	AHK5 H376*rev	5'-ggggaacggatttctaagacatagtcgccagcatttgcttc-3'	42	48	83	QC H376*
79	AHK5 193_fwd	5'-ccggtctcccatgctgtcagaatcgctcaaaacg-3'	34	56	85	L193 fwd
81	AHK5 506_rev_loopleft	5'-cgcactggccgatgagtgatgataccgacgttgcc-3'	36	61	88	Δ507- 535
82	AHK5 536_fwd_loopright	5'-gcatcggccagtcgagcgtctggatctgctgtgac-3'	36	64	89	Δ507- 535

3.2.1.1 PCR protocols

Polymerase chain reaction (PCR) was used to amplify selected stretches of DNA in vitro both in order to generate fragments to be used in restriction-ligation protocols for the generation of protein expression constructs or to perform site-directed mutagenesis on existing constructs.

For insert amplification, components were mixed according to Table 3-15. PCR was performed in an iCycler thermocycler (BioRad, Hercules, CA, USA) using gradually increasing annealing temperatures (Table 3-16). The protocol was originally developed by Prof. Dr. Remco Sprangers (University of Regensburg).

Table 3-15: Composition of PCR mix for insert amplification.

template	200 ng
primer fwd (100 μM)	0.25 μL
primer rev (100 μM)	0.25 μL
dNTP mix (10 mM)	1 μL
Q5 DNA polymerase	0.5 μL
5x Q5 reaction buffer	10 μL
ddH ₂ O	ad 50 μL

Table 3-16: Thermocycler protocol for insert amplification PCR.

Block	(repeats)	Step	Temperature (°C)	Time (s)
0	(1)	Initial melting	98	300
1	(3)	Melting	98	10
		Annealing	40	30
		Extension	72	80
2	(6)	Melting	98	10
		Annealing	45	30
		Extension	72	80
3	(9)	Melting	98	10
		Annealing	50	30
		Extension	72	80
4	(12)	Melting	98	10
		Annealing	55	30
		Extension	72	80

5	(15)	Melting	98	10
		Annealing	60	30
		Extension	72	80
6	(1)	Final extension	72	600
		Storage	8	∞

In contrast, site-directed mutagenesis was performed using a lower template DNA concentration (Table 3-17). The thermocycler program (Table 3-18) employed longer elongation steps in order to efficiently amplify the entire plasmid. Two separate but constant annealing temperatures were used in parallel. After PCR the condition with the highest yield was selected by agarose gel electrophoresis (section 3.2.1.3).

Table 3-17: PCR mix for site-directed mutagenesis.

template	50 ng
primer fwd (10 μ M)	0.9 μ L
primer rev (10 μ M)	0.9 μ L
dNTP mix (10 mM)	1 μ L
Q5 DNA polymerase	0.5 μ L
5x Q5 reaction buffer	10 μ L
ddH ₂ O	ad 50 μ L

Table 3-18: Thermocycler protocol for site-directed mutagenesis PCR.

Block	(repeats)	Step	Temperature ($^{\circ}$C)	Time (s)
0	(1)	Initial melting	98	300
1	(20)	Melting	98	30
		Annealing	55 or 65	30
		Extension	72	240
6	(1)	Final extension	72	600
		Storage	8	∞

3.2.1.2 Generation of expression plasmids by restriction and ligation

Following insert amplification by PCR the reaction product was extracted using the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA). The resulting DNA was quantified photometrically (section 3.2.1.4) and used for restriction digestion. Parallel reactions were set up for the insert and the target vector (purification see section 3.2.2.2) with compatible restriction enzymes. The reactions were set up according to Table 3-19 and incubated for 90 min at 37 $^{\circ}$ C.

Table 3-19: Composition of restriction reactions.

	insert	vector
DNA	4 μ g	10 μ g
10x CutSmart Buffer	5 μ L	10 μ L
restriction endonucleases	1 μ L each	2 μ L each
ddH₂O	ad 50 μ L	ad 100 μ L

After digestion, restriction endonucleases were heat-inactivated at 65 °C for 20 min. In order to prevent target vector religation, 2 µL of recombinant shrimp alkaline phosphatase (rSAP, New England Biolabs, Ipswich, MA, USA) were added to the vector and again incubated at 37 °C for 60 min. Both the insert and the vector backbone were purified using the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA).

Finally, ligation reactions were set up using 75 ng of vector and varying molar excesses of insert, as listed in Table 3-20. Ligation was allowed to occur at 4 °C for 10-18 h (O/N). The next day, the reaction was used to transform *E. coli* cloning cells (DH5- α , section 3.2.2.1).

Table 3-20: Ligation mix.

target vector	75 ng
insert	3x, 5x, or 10x molar excess
10x T4 reaction buffer	2 µL
T4 DNA ligase	0.25 µL
ddH ₂ O	ad 20 µL

3.2.1.3 Agarose gel electrophoresis

Agarose gel electrophoresis was employed to assess efficacy and quality of PCR amplification as well as for preparative separation of inserts and vector backbone in subcloning experiments.

Briefly, 0.8 % (w/v) or 1 % (w/v) agarose were dissolved by heating in 1x TAE buffer (Table 3-10) for preparative and analytical purposes, respectively. Two µL of GelRed (Genaxxon bioscience, Ulm, Germany) were added to one 50 mL agarose gel. Gels were cast into Mini-Sub Cells GT (BioRad, Hercules, CA, USA) and appropriate combs were inserted to generate sample pockets. After polymerisation, the gels were transferred into the cognate running chamber filled with 1x TAE buffer, and the samples, premixed 1:6 in 6x loading dye (Thermo Fisher Scientific, Waltham, MA, USA) were applied. A DNA ladder of appropriate size (O'GeneRuler 100+500 bp or GeneRuler 1kb plus, Thermo Fisher Scientific, Waltham, MA, USA) was included in a separate pocket. Electrophoresis was performed for 30 - 90 min at 100 V. DNA detection was carried out with a GelDoc system (BioRad, Hercules, CA, USA), imaging was performed with a smartphone camera.

If required, target gel bands were excised using a sterile scalpel and DNA was purified by means of the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA).

3.2.1.4 Determination of DNA concentration

DNA concentrations were determined photometrically using a NanoDrop ND1000 (Thermo Fisher Scientific, Waltham, MA, USA). Typically, sample volumes of 2 µL were applied to the sample pedestal and full absorption spectra were recorded between 220 and 350 nm. Peak absorption intensity at 260 nm was converted into mass concentrations using the specific extinction coefficient for DNA of $\epsilon_{260} = 0.02 \mu\text{L ng}^{-1} \text{cm}^{-1}$. Protein and RNA contaminations were estimated using the absorption ratios of A_{260}/A_{230} and A_{260}/A_{230} , respectively [148].

3.2.2 Culturing of *E. coli* cellsTable 3-21: Composition of media used for *E. coli* culture.

Medium	Components			
Lysogeny Broth (LB)			10 g/L	tryptone
			5 g/L	NaCl
			5 g/L	yeast extract
			autoclaved	
LB agar			10 g/L	tryptone
			5 g/L	NaCl
			5 g/L	yeast extract
			20 g/L	agarose
			autoclaved	
Terrific Broth (TB)	YTG base	(900 mL)	12 g	tryptone
			24 g	yeast extract
			4 mL	glycerol
			autoclaved	
	10x phosphate buffer	(100 mL)	0.17 M	KH ₂ PO ₄
		0.72 M	K ₂ HPO ₄	
		autoclaved		
Simplified Studier Auto-induction medium	ZY base	(938 mL)	5 g	yeast extract
			10 g	N-Z amino acids
			autoclaved	
	25x M	(40 mL)	625 mM	Na ₂ HPO ₄
			625 mM	KH ₂ PO ₄
			125 mM	NH ₄ Cl
pH 6.75 @ RT, autoclaved				
1 M MgSO ₄	(2 mL)			
50x 5052	(20 mL)	250 g/L	glycerol	
		25 g/L	glucose	
		100 g/L	α-lactose	
		filter sterilised		

3.2.2.1 Transformation of *E. coli* cells

Purified plasmids, QC reactions, and ligations were transformed into chemically competent *E. coli* cells. DH5- α cells (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) were used for plasmid amplification and cloning while protein expression was carried out in BL21 (DE3)-RIL cells (Stratagene, La Jolla, CA, USA). 50 μ L aliquots were transformed according to the following protocol:

- thaw cells on ice
- add 100-200 ng of plasmid DNA or 2 μ L purified QuikChange PCR or ligation mix
- incubate 15 min on ice
- heat-shock at 42 °C for 60-90 sec
- cool on ice
- add 700 μ L LB w/o antibiotics
- shake at 37°C for 1 hr
- spread on LB-Agar plates with appropriate antibiotics for clone picking or
- add to 25 mL LB culture with appropriate antibiotics for plasmid amplification and protein expression

3.2.2.2 Plasmid amplification and isolation

For plasmid amplification, newly transformed cultures or clones picked from LB-Agar plates were grown O/N at 37° C while shaking at 180 rpm in 25 mL LB with appropriate antibiotics. Cells were then harvested by centrifugation at 3 345 x g in a Heraeus Multifuge 1L-R table top centrifuge (Thermo Fisher Scientific, Waltham, MA, USA) and DNA extraction was performed using the Wizard® *Plus* SV Minipreps DNA Purification System (Promega, Madison, WI, USA).

3.2.2.3 Generation of glycerol stocks

Glycerol stocks for medium-term storage of freshly transformed cloning cells (DH5- α or DH10Bac™) were generated by adding glycerol to a final concentration of 15 % (v/v) to 1 mL O/N culture destined for plasmid isolation. Thus cryo-protected cells were flash-frozen in liquid nitrogen and stored at -80 °C.

3.2.3 Protein analytics

3.2.3.1 Determination of protein concentration

Protein concentrations were determined photometrically using a NanoDrop ND1000 (Thermo Fisher Scientific, Waltham, MA, USA). Depending on sample availability and surface tension, sample volumes ranging from 1.5 to 2 μL were applied to the sample pedestal and full absorption spectra were recorded between 220 and 350 nm. Spectra were inspected for possible contaminations by nucleic acids, as visible by an absorption peak at 260 nm, in addition to the protein-specific ones at 215 and 280 nm. Pure protein signals were converted into molar or specific concentrations using the absorption at 280 nm and extinction coefficients as well as molecular weights calculated by the ExPasy ProtParam tool (<https://web.expasy.org/protparam/>, [149]).

3.2.3.2 Denaturing Polyacrylamide gel electrophoresis (SDS-PAGE)

Protein purity and size were assessed by SDS-PAGE. Gels were prepared according to Table 3-22. Briefly, mixes for stacking and resolving gels were prepared separately with the omission of APS and TEMED which were added immediately prior to casting in order to initiate polymerisation. The resolving gel was cast into a multicast chamber (Mini-PROTEAN 3, BioRad, Hercules, CA, USA) containing 10-13 gel sandwiches (BioRad, Hercules, CA, USA) and overlaid with isopropanol. After allowing the gel to polymerise for 20-30 min, isopropanol was removed and the stacking gel was cast. Sample pockets were generated by inserting appropriate combs into the freshly cast stacking gel. After polymerisation, gels were stored at 4 °C wrapped in a moist paper towel to prevent exsiccation.

Table 3-22: Gel composition for casting of up to 13 15 % Acrylamide gels.

	Resolving gel (15 %)	Stacking gel (4 %)
30 % Acrylamide : Bisacrylamide 37:1	30.2 mL	2.6 mL
1.5 M Tris pH 8.8	15.1 mL	-
1.5 M Tris pH 6.8	-	4.8 mL
ddH₂O	13.8 mL	11.5 mL
10 % (w/v) SDS	600 μL	190 μL
10 % (w/v) APS	600 μL	190 μL
TEMED	60 μL	19 μL

Samples were prepared by mixing 15 μL of protein solution with 5 μL of 4x protein sample buffer (4xPPP). Samples were boiled at 95 °C for at least 3 min (Note: Boiling was omitted in GatD/MurT analytics, as GatD selectively precipitates in the process). Electrophoresis was performed in 1x Roti-phoresis™ SDS-PAGE buffer (Roth, Karlsruhe, Germany) at 250 V 47 min using a Powerpac HC power source (BioRad, Hercules, CA, USA). The gel was then boiled in deionised water to remove SDS. Protein gel stain (InstantBlue™ (Expedeon, Heidelberg, Germany)) was added and the gel boiled again and incubated for at least 5 min. The staining solution was discarded and background staining was removed by incubation in deionised water.

3.2.3.3 Analytical SEC

Native protein size and sample homogeneity as well as their stability over time were assessed by analytical SEC. Depending on target protein size, either a Superose 6 3.2/300 or a Superdex 200 Increase 3.2/200 analytical gel filtration column (GE Healthcare, Uppsala, Sweden) fitted to an ETTAN FPLC system (GE Healthcare, Uppsala, Sweden) was equilibrated with the appropriate protein buffer. Protein sample volumes of 30 μL containing at least 50 μg of protein were clarified of potential aggregates by centrifugation for 5 min at 16 100 x g and 4 $^{\circ}\text{C}$ in an Eppendorf table-top centrifuge and loaded onto the FPLC system. SEC was performed at 4 $^{\circ}\text{C}$ and a flow rate of 40 $\mu\text{L}/\text{min}$ or 50 $\mu\text{L}/\text{min}$ for the Superose 6 and Superdex 200 Increase resins, respectively. UV absorption at 280 nm, 260 nm, and 215 nm as well as sample electrical conductivity were monitored.

3.2.3.4 Dynamic light scattering (DLS)

Size and homogeneity estimates by analytical SEC were further consolidated by means of dynamic light scattering using a Zetasizer Nano ZS (Malvern Instruments, Malvern, UK). Samples were filtered (Costar Spin-X Centrifuge Tube Filter, 0.22 μm Cellulose Acetate membrane) prior to measurement. Measurements were performed in a 3 mm quartz cuvette in triplicates. Automated processing was carried out in the Zetasizer Software v.7.11 (Malvern Instruments, Malvern, UK).

3.2.3.5 Circular dichroism (CD) spectroscopy

CD spectroscopy experiments were performed on a JASCO J-720 (Jasco, Easton, MD, USA) spectrometer under nitrogen flow. Protein samples were diluted tenfold with deionised water in order to avoid detector saturation due to high salt concentrations. Typical protein concentrations range from 0.1 to 0.3 mg/mL. The resulting spectra were corrected for concentration, protein molecular weight, and cuvette thickness.

3.2.3.6 Protein identity determination by mass spectrometry

In order to ascertain protein or fragment identities after purification or limited proteolysis (section 3.2.3.7), MALDI-TOF mass spectrometry was employed. Protein samples were separated by SDS-PAGE and the relevant bands excised from the gel and sent to Sophie Stotz (Kalbacher laboratory, University of Tübingen) for MALDI-TOF analysis.

3.2.3.7 Protease exclusion assay

In order to investigate protein domain boundaries as well as compactness of fold protease exclusion assays (limited proteolysis) were performed. To this end, the target protein was incubated at RT for one hour or at 4 $^{\circ}\text{C}$ O/N with a selection of proteases obtained from the Proti-Ace™ Kit 1+2 (Hampton Research, Aliso Viejo, CA, USA). Protease and target protein amounts were chosen as recommended in the kit's handbook.

3.2.3.8 Melting temperature measurement by Differential Scanning Fluorimetry (DSF)

DSF was performed on a QuantStudio 5 real-time PCR cycler (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA). The protein of interest was supplied at 0.1 - 1 mg/mL and Protein Thermal Shift Dye™ (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) was added as a DSF fluorophore. The samples were equilibrated at 4 °C and gradually heated to 95 °C over 30 min while monitoring dye fluorescence. When applicable, protein melting temperatures were derived from the melting curve inflection point using the Protein Thermal Shift™ Software v1.3 (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA).

In order to screen for stabilising buffer conditions, DSF was performed in a 96 well format in 1:4 dilutions of the Solubility & Stability Screen™ 1+2 (Hampton Research, Aliso Viejo, CA, USA). Melting temperatures were compared and the most stable conditions (i.e. highest melting temperatures) selected for further experiments.

3.3 Production and Purification of AHK5

3.3.1 List of buffers

Table 3-23: Buffers used for AHK5 purification, sorted by construct.

Construct	Name	Composition		
A3₂	IMAC A	100 mM	Tris pH 8.2 @ RT	
		50 mM	NaCl	
		100 mM	Urea	
		1 mM	DTT	
	IMAC B	as IMAC A		
	SEC I	1 M	Imidazole	
		100 mM	Tris pH 8.2 @ RT	
		150 mM	NaCl	
SEC II	1 mM	DTT		
	25 mM	Tris pH 8.2 @ RT		
	150 mM	NaCl		
	1 mM	DTT		
C2	C2 main	50 mM	Tris pH 8.2 @ RT	
		150 mM	NaCl	
		1 mM	DTT	
	Triton buffer	as C2 main	1 % (v/v) TRITON X-100	
	Solubilisation buffer	as C2 main	8 M Urea	
	IMAC I B	as solubilisation buffer	250 mM Imidazole	
	IMAC II B	as C2 main	500 mM Imidazole	
	SEC II	as C2 main	500 mM NaCl	
	A2	A2 main	50 mM	Tris pH 7.9 @ RT
			1.5 M	NaCl
1 mM			DTT	
Triton buffer		as A2 main	1 % (v/v) TRITON X-100	
Solubilisation buffer		as A2 main	8 M Urea	
IMAC I B	as solubilisation buffer	250 mM Imidazole		

Table 3-24: List of buffers used for Western Blotting.

Name	Composition	
Transfer Buffer	25 mM	Tris base
	1.44 % (w/v)	Glycerol
	20 % (v/v)	Methanol
10x TBS	200 mM	Tris pH 7.5 @ RT
	1.5 M	NaCl
1x TBS-T		1x TBS
	+ 0.5 % (v/v)	Tween 20
1x TBS-T-MP5		1x TBS-T
	+ 5 % (w/v)	milk powder

3.3.2 Computational methods

In order to identify possible domain boundaries within the AHK5 input domain, its amino acid sequence was analysed with different web-based bioinformatics tools. In addition to traditional BLAST searches (<https://blast.ncbi.nlm.nih.gov/>), secondary structure prediction-based methods were employed. The core work was performed using the MPI Tübingen Bioinformatics Toolkit (<https://toolkit.tuebingen.mpg.de>, [64]).

Briefly, sequence searches and alignments were performed based on secondary structure prediction using the HHpred algorithm. These alignments were then inspected manually for possible domain boundaries as well as employed as templates for tertiary structure prediction using MODELLER [74]. Additionally, pure secondary structure predictions were performed using Quick2D [64].

Finally, a BLAST-based conservation search was performed using Consurf [67] in order to identify sequences of increased conservation to further guide construct design.

3.3.3 Molecular biology and cloning

Table 3-25: Constructs used for expression of AHK5 fragments in *E. coli* cells.

ID	fragment	residues / mutation	vector	tag	comments
11	FL	1-922	pEX-K4	-	codon-optimised source vector
19	A2	1-617	pETM-11	N-His ₆ -TEV	used for all experiments on A2 fragment (sections 1.2.5, 1.2.6, 1.2.7, 1.2.8, and 1.2.9)
20	A3 ₁	1-342	pETM-11	N-His ₆ -TEV	expression tests only
21	A3 ₂	1-364	pETM-11	N-His ₆ -TEV	used for all experiments on A3 ₂ fragment (sections 1.2.3, 1.2.6, 1.2.7, 1.2.8, and 1.2.9)
22	B3 ₁	72-342	pETM-11	N-His ₆ -TEV	expression tests only
23	C2	203-617	pETM-11	N-His ₆ -TEV	all experiments on C2 fragment (sections 1.2.4, 1.2.6, 1.2.7, and 1.2.9)

ID	fragment	residues / mutation	vector	tag	comments
24	3 ₁ 2	352-617	pETM-11	N-His ₆ -TEV	expression tests only
25	3 ₂ 2	361-617	pETM-11	N-His ₆ -TEV	expression tests only
28	A3 ₂	1-364 / C3A	pETM-11	N-His ₆ -TEV	used for oxidation studies (section 1.2.9)
31	B3 ₂	72-364	pETM-11	N-His ₆ -TEV	expression tests only
32	A3 _x	1-331	pETM-11	N-His ₆ -TEV	expression tests only
33	A3 _{xx}	1-352	pETM-11	N-His ₆ -TEV	expression tests only
34	A3 ₃	1-375	pETM-11	N-His ₆ -TEV	expression tests only
54	REC	774-922	pETM-11	N-His ₆ -TEV	planned use in <i>in vitro</i> phosphorelay experiments
60	A2Δ	1-617 Δ507-535	pETM-11	N-His ₆ -TEV	version of A2 without insert in HATPase domain – purification successful
61	A2	1-617	pETM-30	N-His ₆ -GST-TEV	A2 with GST tag – not used
62	C*2	193-617	pETM-11	N-His ₆ -TEV	version of C2 extending further N-terminally – not used
63	C*2	193-617	pETM-30	N-His ₆ -GST-TEV	version of C2 extending further N-terminally – not used
64	C2Δ	203-617 Δ507-535	pETM-30	N-His ₆ -GST-TEV	version of C2 without insert in HATPase domain – not used
65	ADv1	428-617	pETM-11	N-His ₆ -TEV	HATPase V1 – for investigation of role of Cysteines – insoluble
66	ADv2	428-622	pETM-11	N-His ₆ -TEV	HATPase V2 – for investigation of role of Cysteines – insoluble
67	ADv3	431-617	pETM-11	N-His ₆ -TEV	HATPase V3 – for investigation of role of Cysteines – insoluble
68	ADv4	431-622	pETM-11	N-His ₆ -TEV	HATPase V4 – for investigation of role of Cysteines – insoluble

Table 3-26: Constructs used for expression of AHK5 fragments in insect cells.

ID	fragment	residues / mutation	vector	tag	comments
12	FL	1-922	pEX-K4	-	codon-optimised source vector
29	FL	1-922	pFastBac1	C-His ₆	used for expression tests (section 1.4)
30	FL	1-922	pFastBac1	C-FactorXa-His ₆	used for expression tests (section 1.4)

3.3.4 Expression of AHK5 constructs in *E. coli* cells

Plasmids containing the genes for varying AHK5 fragments (Table 3-25) were transformed into BL21 (DE3) RIL *E. coli* cells (Stratagene, La Jolla, CA, USA) as described in Section 3.2.2.1. 20 mL LB suspension cultures containing the appropriate selection antibiotics were grown O/N at 37 °C. The next day, 1.5 L expression cultures (TB-kan) were inoculated with the O/N-culture each and grown at 37 °C until an OD₆₀₀ value of 0.4-0.5 was reached. The temperature was subsequently decreased to 20 °C and 0.5 mM IPTG were added for induction.

For native purification (fragment A3₂), protein expression was terminated after 5 h, for denaturing purifications from inclusion bodies (fragments C2, A2), expression was carried out O/N. The cells were subsequently harvested by centrifugation using a Sorvall RC6+ (Thermo Fisher Scientific, Waltham, MA, USA) centrifuge equipped with an SLC-4000 rotor (10 000 x g, 15 min) and stored at -20 °C until purification.

For preliminary expression and solubility tests, small (100 mL) cultures were used and harvesting was carried out as described above, but in 1.5 mL aliquots and using a table top Eppendorf centrifuge.

In all cases, OD₆₀₀ was determined prior to harvesting.

3.3.5 Expression verification and solubility screen

Pellets derived from small scale expression (1.5 mL aliquots) were resuspended in (50 · OD₆₀₀) µL 1x protein sample buffer and boiled at 95 °C for 10 min. Five µL of the resulting suspension were then analysed by SDS-PAGE. Pellets obtained before and after induction by IPTG addition were compared side by side.

In order to assess protein solubility upon confirmation of expression, equal pellets obtained from the same 100 mL culture were resuspended in 1 mL of buffer and lysed by sonication (Branson Digital Sonifier 450, microtip mode, 20 % amplitude, 0.5 s pulse alternating with 3.5 s pause, total pulse length 30 s) and cell debris was removed by centrifugation at 16 100 x g in an Eppendorf table top centrifuge. The two different sets of buffers used in this screening method are listed in Table 3-27 and Table 3-28, respectively. After lysis and clarification, the supernatant fraction of each sample was analysed by SDS-PAGE.

Table 3-27: List of buffers used in sparse matrix solubility screen. Selected from [69]

Buffer ID	Composition	
1-1	100 mM	Tris pH 7.6 @ RT
	10 % (v/v)	Glycerol
1-2	100 mM	Tris pH 7.6 @ RT
	50 mM	LiCl
1-3	100 mM	KH ₂ PO ₄ pH 4.3
	2.5 M	ZnCl ₂
1-4	100 mM	HEPES pH 7.0
	100 mM	KCl
1-5	100 mM	Tris pH 8.2 @ RT
	50 mM	NaCl
	10 % (v/v)	Isopropanol
1-6	100 mM	KH ₂ PO ₄ pH 6.0
	50 mM	(NH ₄) ₂ SO ₄
	1 % (v/v)	Triton X-100
1-7	100 mM	Tris pH 8.2 @ RT
	50 mM	NaCl
	100 mM	Urea
1-8	100 mM	HEPES pH 7.0
	100 mM	Na-Glutamate
	5 mM	DTT
1-9	100 mM	Na-Acetate pH 5.5
	100 mM	KCl
	0.1 % (w/v)	β-D-Octylglucoside
1-10	100 mM	Na-Acetate pH 5.5
	1 M	MgSO ₄
1-11	100 mM	HEPES pH 7.0
	1 M	MgSO ₄
1-12	100 mM	Triethanolamine pH 8.5
	50 mM	LiCl
	50 mM	EDTA
1-13	30 mM	Tris pH 7.5 @ RT
	300 mM	NaCl
1-14	30 mM	Tris pH 7.5 @ RT
	300 mM	NaCl
	1 % (w/v)	N-Lauroylsarcosinate

Table 3-28: Grid screen for lysis buffer. pH is screened against salt concentrations, numbers in table represent buffer IDs

Buffer	NaCl	150 mM	0.5 M	1.5 M
50 mM MES pH 6.5		2-1	2-2	2-3
50mM HEPES pH 7.0		2-4	2-5	2-6
50mM HEPES pH 7.5		2-7	2-8	2-9
50mM HEPES pH 8.0		2-10	2-11	2-12
50 mM Tris pH 8.5		2-13	2-14	2-15

3.3.6 Native protein purification (fragment A3₂)

The A3₂ construct, encompassing the entire input domain, was purified in a three-step process comprising initial Ni²⁺-NTA-affinity chromatography and two subsequent rounds of size exclusion chromatography. The purification strategy is depicted schematically in Figure 3-1.

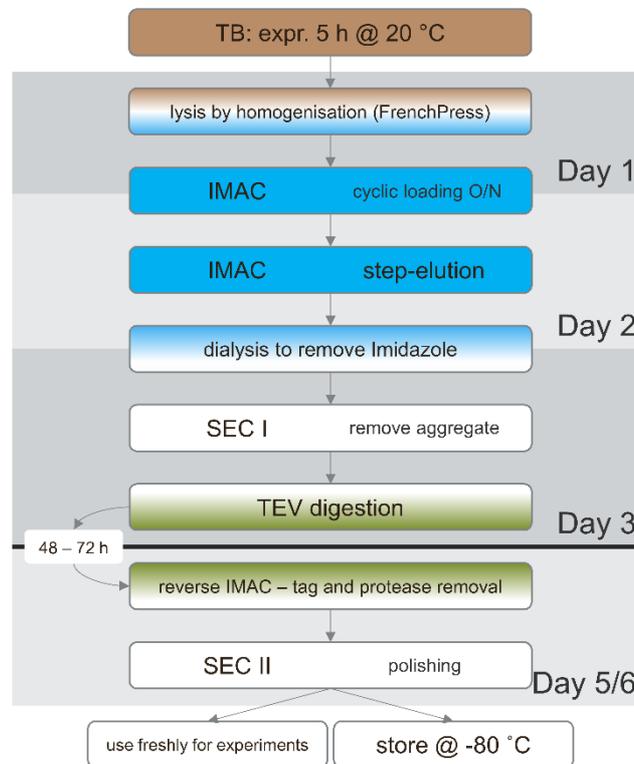


Figure 3-1: Schematic overview of the AHK5 A3₂ fragment purification strategy. Imidazole-containing steps are coloured in blue, all steps performed in standard SEC buffer (Table 3-23) are white. The presence of TEV protease is denoted in green. Grey boxes in the background highlight the recommended timeline.

3.3.6.1 Lysis by high-pressure homogenisation

E. coli pellets containing AHK5 A3₂ were lysed by high pressure homogenisation using an Emulsi-flex C3 (Avestin, Canada). Pellets were resuspended in 10 mL of Lysis buffer (Table 3-23) per gram of cell wet weight (~4 g for wt, 2 g for C3A mutant). After addition of 1:10 000 volume units of Benzonase® (Sigma-Aldrich, St. Louis, MO, USA) and 0.1 mM PMSF, the slurry was incubated at 4 °C for 15 min while stirring. Cells were then lysed by repeatedly flowing them through the ice-cooled homogeniser. The device had previously been equilibrated with Lysis buffer and the air pressure applied to the homogenising valve was constantly adjusted to maintain a hydrostatic pressure on the cells of 800-1 000 bar.

The lysate was subsequently clarified by centrifugation (Sorvall RC6+, 4 °C, 60 min, 34 000x g).

3.3.6.2 Immobilised Metal Ion Affinity Chromatography (IMAC)

AHK5 was enriched from the lysate by IMAC. To this end, a 5 mL Ni-NTA-Agarose cartridge (Hitrap Crude FF, GE Healthcare, Uppsala, Sweden) was equilibrated with IMAC A buffer (Table 3-23) and the clarified lysate was flowed over the cartridge cyclically O/N at a flow rate of 0.75 mL/min. On

the next day, the cartridge was fitted onto an ÄKTA Prime FPLC system (GE Healthcare, Uppsala, Sweden) that had previously been washed with the appropriate buffers. Unspecifically bound proteins were removed by washing the Ni²⁺-NTA matrix with 10 CV of IMAC A followed by 50 mL of IMAC A with 50 mM imidazole. Elution was carried out by first increasing the imidazole concentration to 150 mM for 50 mL and, finally, applying a gradient from 150 mM to 1 M imidazole. Throughout washing and elution, the UV absorption at 280 nm was monitored and relevant peak fractions were analysed by SDS-PAGE. Fractions containing only protein of correct size were pooled based on SDS-PAGE results.

In order to remove imidazole from the protein, the pooled fractions were dialysed O/N (Standard RC Tubing, 6 – 8 kDa, 1.7 mL/cm, 23 mm, SpectrumLabs) against 2 L of SEC I buffer (Table 3-23).

3.3.6.3 SEC I

After dialysis, the protein was concentrated to a maximum of 2.6 mg/mL using centrifugal concentrators (Sartorius VIVASPIN 20, 50 kDa MWCO, 1 300 x g). Precipitated protein and particulate aggregates were then removed by centrifugation at 16 100 x g in an Eppendorf table-top centrifuge. The remaining sample was then applied to a HiLoad 26/60 column packed with Superdex 200 gel filtration media (GE Healthcare, Uppsala, Sweden) and equilibrated with buffer SEC I. After emptying the loading loop with a flow rate of 1 mL/min, elution was performed with an isocratic flow of 3 mL/min. Throughout the run, UV absorption at 280 nm was monitored. Relevant peak fractions were pooled based on the UV trace and SDS-PAGE analysis.

3.3.6.4 Affinity tag removal

After the first gel filtration step, the N-terminal His₆-tag was removed by proteolytic cleavage using TEV-protease. To this end, 1 mg protease was added per 5 mg AHK5. After supplementing the mixture with 1 mM fresh DTT, cleavage was allowed to occur at 4 °C O/N while gently mixing the sample in a tilt shaker.

The next day, successfully digested protein was separated from undigested sample as well as the protease and the cleaved tags by means of a second IMAC step. The mixture was flowed over a 5 mL IMAC cartridge (separate from that used in step 3.3.6.2) and the flow-through containing only cleaved protein was collected. The cartridge was then regenerated by washing with 1 M imidazole. All steps were monitored by SDS-PAGE.

3.3.6.5 Polishing SEC

In order to achieve the highest possible sample homogeneity, a second size exclusion chromatography step was carried out as described in 3.3.6.3. Finally, the protein was concentrated to ~ 3 mg/mL and flash-frozen in liquid nitrogen in 100 - 200 µL aliquots for long-term storage at -80 °C.

3.3.7 Denaturing protein purification from inclusion bodies (fragments A2 and C2)

Constructs containing the input domain as well as the kinase domain (A2, and C2) were purified from inclusion bodies by means of urea solubilisation and subsequent refolding. An overview of the purification strategy is shown in Figure 3-2.

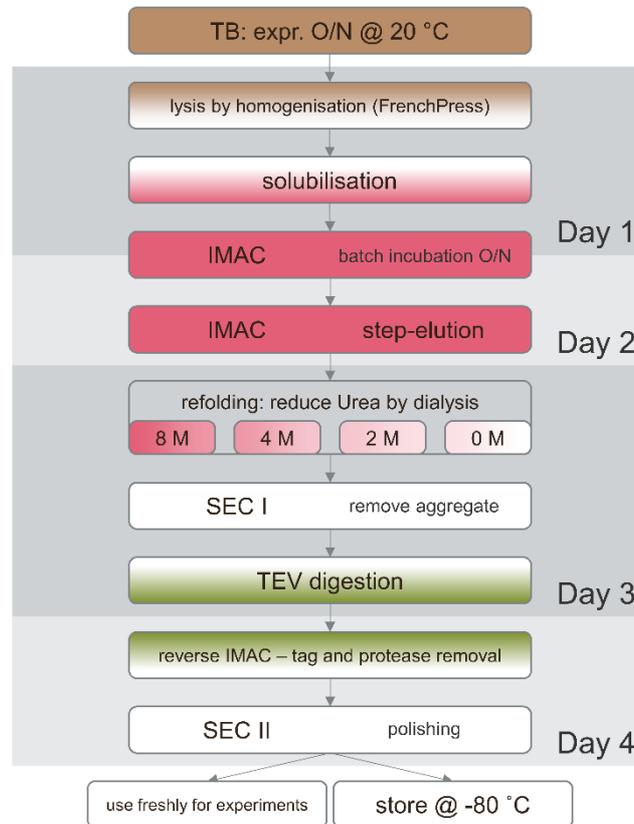


Figure 3-2: Denaturing purification strategy for A2 and C2 fragments. The pink colour denotes the presence of urea, all steps performed in main buffer (Table 3-23) are white, TEV protease steps are green.

3.3.7.1 Lysis and solubilisation

E. coli cell pellets containing inclusion bodies of AHK5 constructs were lysed by high-pressure homogenisation using an Avestin Emulsiflex C3. Pellets were resuspended in 10 mL of the appropriate main buffer (A2 or C2, Table 3-23) per gram of cell wet weight. Following addition of 1 : 10 000 Benzonase® (Sigma-Aldrich, St. Louis, MO, USA), 0.1 mM PMSF and one tablet of cComplete® EDTA-free Protease Inhibitor Cocktail (Roche), the cell slurry was incubated on ice for 30 min while stirring. The cells were subsequently lysed by repeated passages through the high-pressure homogeniser while applying air pressures of 1.8 - 2.2 bar (resulting in hydrostatic pressures of 800 - 1000 bar).

In order to isolate the inclusion bodies, the lysate was then centrifuged at 34 000 x g (Sorvall RC6+, SS-34 rotor, 4 °C, 60 min) and the supernatant was removed. After removal of the membrane fraction located at the top of the pellet, the pellet was transferred to a Potter-Elvehjem tissue grinder (Wheaton, Millville, NJ, USA). In order to achieve the highest possible inclusion body purity, the pellet was then washed twice by resuspension in 35 mL of the appropriate Triton wash buffer and twice in main buffer. After each resuspension the pellet was sedimented by centrifugation at 34 000 x g for 30 min and the supernatant was discarded. Finally, the washed inclusion bodies were solubilised in 35 mL solubilisation buffer, and insoluble contaminants were again removed by centrifugation.

3.3.7.2 Immobilised Metal Ion Affinity Chromatography (IMAC)

The urea-solubilised protein was added to 5-10 mL Ni²⁺-NTA agarose beads (Qiagen, Hilden, Germany) that had previously been washed and resuspended in solubilisation buffer. The mixture was then incubated O/N at 4 °C while rotating.

The next day, the Ni²⁺-NTA beads were pelleted by centrifugation at 1 880 x g at 4 °C for 15 min in a Heraeus Multifuge 1L-R table top centrifuge, resuspended in solubilisation buffer and transferred to an EconoPac gravity flow column (BioRad, Hercules, CA, USA). Unbound protein and contaminants were removed by washing with 30 - 40 mL solubilisation buffer. The bound protein was then eluted using IMAC I B buffer in 12 x 2 mL fractions. Fractions containing a high amount of ≥ 80 % pure target protein, as estimated by SDS-PAGE, were pooled for further processing.

Denatured protein in IMAC I B buffer was then stored at 4 °C for up to 2 weeks until further processing.

3.3.7.3 Refolding

In order to achieve correct folding of the target protein, urea concentration in the buffer was reduced stepwise by dialysis. For best results, 2.5 mL IMAC I eluate were diluted 1 : 4 with solubilisation buffer and placed in to a pre-wetted dialysis bag (Standard RC Tubing, 6 – 8 kDa, 1.7 mL/cm, 23 mm, SpectrumLabs). In a first step, dialysis was then carried out O/N against solubilisation buffer (8M urea) in order to remove imidazole. The next day, refolding was performed by 2 h dialysis in main buffer containing 4 M urea followed by 1 h in 2 M urea and finally 1 h in main buffer without urea.

In order to minimise aggregation, AHK5 fragments were subjected to preparative SEC immediately after refolding without further concentration.

3.3.7.4 SEC I

Aggregation nuclei were removed immediately after the final refolding dialysis. The content of the dialysis bag was clarified by centrifugation at 16 100 x g in a cooled Eppendorf table top centrifuge and applied to a HiLoad 26/60 column packed with Superdex 200 gel filtration media (GE Healthcare, Uppsala, Sweden) that had previously been equilibrated with main buffer (Table 3-23). Loading and elution parameters were chosen as described in 3.3.6.3.

3.3.7.5 Affinity tag removal and polishing SEC

TEV-protease digestion, reverse IMAC and polishing SEC (SEC II) were performed as described for the A3₂ fragment in sections 3.3.6.4 and 3.3.6.5 with the following fragment-specific variations: 0.5 mg TEV protease were employed to digest 20 mL pooled SEC I fractions and all concentration steps were carried out using 10 K MWCO centrifugal concentrators (VIVASPIN 20, PES membrane, Sartorius, Göttingen, Germany) at a centrifugal force of 1 300 x g. The maximum protein concentration never to

exceed throughout the purification in order to limit aggregation was estimated to be 1 mg/mL, based on analytical SEC (Section 3.2.3.3).

While the main buffer was used for both SEC steps for the A2 fragment, AHK5 C2 required a dedicated SEC II buffer for the polishing step (Table 3-23).

3.3.8 Expression of full-length AHK5 in insect cells

In order to obtain empirical data on the overall architecture and domain boundaries of AHK5, stable, folded, and homogeneous full-length protein is essential. Due to the large size of the protein of 922 residues, insect cells were chosen as the expression system.

3.3.8.1 Generation of recombinant bacmids

The recombinant baculoviral genome (bacmid) containing the gene coding for AHK5 was generated using the Bac-to-Bac system (Thermo Fisher Scientific, Waltham, MA, USA) [150], [151]. Codon-optimised full-length AHK5 DNA was obtained from Eurofins Genomics GmbH (Ebersberg, Germany, DB ID #12, Table 3-26) and sub-cloned into a pFastBac-1 entry vector by General Biosystems (Durham, NC, USA).

For recombination of the transgene from the entry vector into the infectious bacmid, entry vectors containing AHK5-FL were transformed into DH10Bac™ *E. coli* cells. These cells contain the acceptor bacmid and express the relevant transposases. DH10Bac™ require a dedicated transformation protocol:

- thaw cells on ice
- add 1 ng of plasmid DNA
- incubate 30 min on ice
- heat-shock at 42 °C for 45 sec
- cool on ice for 2 min
- add 900 µL LB w/o antibiotics
- shake at 37°C for 4 hr
- spread on LB-Agar plates containing 50µg/mL kanamycin, 7 µg/mL gentamycin10 µg/mL tetracycline, 100 µg/mL X-gal, and 40 µg/mL IPTG (LB-Agar-KGTXI) for white-blue screening [152]

After incubation at 37 °C for 48 h, white colonies resulting from lacZ gene disruption by successful recombination were picked and re-streaked on LB-Agar-KGTXI for amplification and blue-white confirmation. Doubly confirmed clones were amplified in suspension culture (LB-KGT), and bacmid DNA was isolated using the PureLink™ HiPure Plasmid Miniprep Kit (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's manual.

3.3.8.2 Clone confirmation by PCR

Successful insertion of the transgene was additionally confirmed by PCR. Two primers flanking the insertion site for the transgene (Table 3-29) were used to amplify the region of the bacmid potentially containing the insert. PCR composition and thermocycler settings are shown in Table 3-30 and Table 3-31, respectively. Following the reaction, 5 µl of the mix were subjected to analytical agarose gel electrophoresis (Section 3.2.1.3). Primer design and PCR parameters were chosen according to the recommendations in the Bac-to-Bac manual [153].

Table 3-29: Primers for check-PCR.

M13 fwd	5'-gtttccagtcacgac-3'
M13 rev	5'-caggaacagctatgac-3'

Table 3-30: Reaction mix for check-PCR.

recombinant bacmid	100 ng
each primer (10 µM)	1.25 µL
dNTP mix (10 mM)	1 µL
Q5 DNA polymerase	0.5 µL
5x Q5 reaction buffer	10 µL
ddH ₂ O	ad 50 µL

Table 3-31: Thermocycler protocol for check-PCR.

Block	(repeats)	Step	Temperature (°C)	Time (s)
0	(1)	Initial melting	93	180
		Melting	94	45
1	(30)	Annealing	55	45
		Extension	72	300
		Final extension	72	420
6	(1)	Storage	8	∞

3.3.8.3 Culturing of sf9 and Hi5 insect cell suspension cultures

Insect cells used for baculovirus generation (Sf9) and protein expression (Hi5) were kept in serum-free suspension cultures in the appropriate growth media (Sf-900 II SFM (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) and Express Five™ SFM (Gibco, Thermo Fisher Scientific, Waltham, MA, USA), respectively) supplemented with Penicillin-Streptomycin (Gibco, Thermo Fisher Scientific, Waltham, MA, USA). Cells were split 1 :4 - 1: 10 every 3 - 5 days at cell counts of $3.0 \cdot 10^6$ - $3.5 \cdot 10^6$ per mL and grown at 28 °C while shaking. Cells were kept for up to 30 passages or until they displayed severe growth defects.

3.3.8.4 Transfection of Sf9 cells for baculovirus production

Transfection of Sf9 cells was performed according to the Bac-to-bac® user guide (Thermo Fisher Scientific, Waltham, MA, USA), with minor modifications.

Briefly, Sf9 cells had been seeded at a density of $1 \cdot 10^6 \text{ mL}^{-1}$ and grown O/N at 28 °C. The next day, after verification that log-phase growth was reached ($1.5\text{-}2.5 \cdot 10^6 \text{ mL}^{-1}$), $8 \cdot 10^5$ cells were seeded in one well of a 6-well tissue culture plate containing 2 mL Grace's Insect Medium (Gibco, Thermo Fisher Scientific, Waltham, MA, USA). Cells were allowed to attach for 15 min at RT.

In the meantime the transfection mix was prepared. To this end, 8 μL Cellfectin® II (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) were added to 100 μL Grace's Insect Medium and mixed vigorously. In a separate vial, 100 μL Grace's Insect Medium were supplemented with 1-2 μg bacmid DNA and mixed gently. Both vials were then combined and incubated at RT for 15 min. The entire transfection mix was then added dropwise to the previously prepared cells. The transfection reaction was then incubated at 27 °C. After five hours the transfection mix was removed and replaced with Sf-900 II SFM growth medium. The cells were stored at 27 °C for 72 - 96 h and regularly inspected for signs of successful infection.

3.3.8.5 Harvesting and amplification of the baculoviral P1 stock

As soon as clear signs of infection were visible (swelling and high granularity of cells, detached cells), the medium, now containing recombinant baculovirus, was harvested and supplemented with 5 % fetal bovine serum (FBS) (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) for storage at 4 °C. This will be referred to as the P1 stock.

In order to obtain virus quantities suitable for expression, at least one round of amplification in Sf9 suspension culture was performed. To this end, P1 virus stock was used to infect a 24 mL suspension culture in log-phase growth (seeded accordingly one day prior). The volume of inoculum required was calculated using Equation (3-2).

$$V_{\text{Inoculum}} = \frac{\text{MOI} \left(\frac{\text{pfu}}{\text{cell}} \right) \cdot N_{\text{cells}}}{\text{Virus titer} \left(\frac{\text{pfu}}{\text{mL}} \right)} \text{ mL} \quad (3-2)$$

A MOI or multiplicity of infection 0.075 plaque forming units (pfu) per cell was chosen and the virus titer estimated to be $5 \cdot 10^6$ pfu/mL. Cells were kept in suspension culture at 28 °C for 72 - 96 h post infection and monitored for signs of infection. After visual inspection revealed clear signs of baculoviral infection, the cells were discarded by centrifugation (10 min, 500 x g) and 5 % FBS (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) were added to the supernatant for storage at 4 °C. This is the P2 stock with an estimated titer of $5 \cdot 10^7$ pfu/mL.

3.3.8.6 Expression test with baculoviral P2 stock

In order to assess whether AHK5 is successfully expressed by Sf9 or High Five™ cells upon infection with recombinant baculovirus stocks derived from the transfection of AHK5-containing bacmids, a small-scale expression test in 24-well plates was performed using the P2 stock as inoculum.

To this end, $6 \cdot 10^5$ log-phase Sf9 or High Five™ cells were seeded per well of a 24-well plate and allowed to attach for 15 min. Attached cells were rinsed once with fresh growth media and then overlaid with 300 μL media. P2 stock was added for MOI of 0.5, 1, 2, 5, or 10 with one negative control

included and the cells were incubated at 27 °C. The experiment was performed in duplicates for expression times of 72 and 120 h, respectively.

After the pre-established expression time, cells were visually inspected for signs of baculoviral infection and harvested. To this end, the growth and infection medium was removed, and the cells were resuspended in 300 µL fresh growth medium by gently pipetting up and down. The cells were then pelleted by centrifugation at 500 x g for 10 min. Both supernatants and the pellets were stored at -20 °C upon analysis.

3.3.8.7 Assessment of expression success by Western Blot and MALDI-TOF

Success of AHK5 expression was ultimately probed by Western Blot. To this end, pellet samples obtained as described in 3.3.8.6 were resuspended in 50 µL of 1x protein sample buffer and boiled at 95 °C for 5 min. The supernatant samples were mixed with appropriate volumes of 4x protein sample buffer and boiled in the same way. Five µL of each sample were then separated by SDS-PAGE in triplicates. One gel was used for InstantBlue® staining, the other two for Western Blotting.

Western Blotting was performed in a semi-dry setup using PVDF membranes (Roth, Karlsruhe, Germany) as a target. In preparation, the membrane was soaked for 10 min in pure methanol followed by a 5 min wash step in ddH₂O and a 10 min incubation in transfer buffer (Table 3-24). The SDS-gel as well as the Whatman papers were soaked in transfer buffer for 15 and 5 min, respectively. The blot sandwich was then assembled with the SDS-PAGE on the cathode side and the membrane on the anode side, flanked by four layers of Whatman papers on either side. Transfer was performed by applying 20 V for 1 h.

After the transfer the membrane was placed into a 50 mL Falcon tube and blocked with 20 mL 1x TBS-T-MP5 (Table 3-24) for 1 h @ RT while rotating. 10 mL primary antibody were added after removal of the blocking solution and decoration was allowed to proceed O/N @ 4 °C. On the next day, the blot was washed three times with 1x TBS-T for 15 min followed by a 1 h incubation with the secondary antibody @ RT. The membrane was then washed four times with 1x TBS-T for 10 min.

Detection was performed by chemiluminescence based on the horseradish peroxidase-coupled secondary antibodies using the Amersham™ ECL™ Western Blot Detection Reagents. The membrane was placed on the imaging table of a ChemiDoc® Imaging System (BioRad, Hercules, CA, USA). and the two reagents were mixed in equal volumes, as specified in the manufacturer's manual and applied to the entire surface of the membrane. Imaging was performed using the device's built-in auto-exposure function. Merged images of InstantBlue-stained SDS-PAGE and Western Blot were generated using Adobe® PhotoShop® CS3.

The antibody combinations used for detection and their formulations are listed in Table 3-32.

Table 3-32: Antibody combinations and formulations used for AHK5-FL detection. For buffer composition, see **Table 3-24**

primary antibody	secondary antibody
Mouse-α-His, 1 : 5 000 in TBS-T-MP5	Goat-α-Mouse, 1 : 5 000 in TBS-T-MP5
Rabbit-α-AHK5-ID, polyclonal, 1 : 1 000 in TBS-T	Goat-α-Rabbit, 1 : 10 000 in TBS-T-MP5

Additionally, well-visible bands in the InstantBlue-stained SDS-PAGE were excised and sent to Sophie Stotz (Kalbacher lab, University of Tübingen) for mass spectrometry analysis using MALDI-TOF.

3.3.8.8 Lysis buffer screen

In order to assess whether AHK5_FL expressed in insect cells can be natively extracted, lysis buffer screening was conducted. To this end, suspension cultures of Hi5 cells were infected with P2 baculoviral stock at a MOI of 10 and incubated for five days at 27 °C. The cells were then harvested in 0.5 mL aliquots. The cells were resuspended in 1 mL lysis buffer and lysed by sonication. Samples were clarified and supernatants were applied to His SpinTrap™ centrifugal IMAC columns (Ge, Healthcare, Uppsala, Sweden). After washing with lysis buffer, bound proteins were eluted with 1:1 mixtures of lysis buffer and 1 M imidazole, pH 8.0. Lysis supernatant, pellet, and IMAC elution were subjected to SDS-PAGE and Western Blot analysis. The buffers used are shown in matrix representation in Table 3-33.

Table 3-33: Grid screen for AHK5_FL lysis buffer. pH is screened against salt concentrations in the presence of 0.1 % Triton X-100 (T). The 0.5 M NaCl buffers were additionally prepared with 100 mM Urea (U) and 20 mM Imidazole (I).

Buffer	NaCl	150 mM	0.5 M	1.5 M
50 mM MES pH 6.0		T	T / U / I	T
50 mM MES pH 6.5		T	T / U / I	T
50 mM HEPES pH 7.0		T	T / U / I	T
50 mM HEPES pH 7.5		T	T / U / I	T
50 mM Tris pH 8.0		T	T / U / I	T

3.3.9 Functionality assays

3.3.9.1 Qualitative assessment of oxidative dimerisation

In order to probe the hypothesised participation of the cysteine at position 3 (Cys3) in AHK5 signalling, the propensity of AHK5 input domain-containing proteins to form covalent dimers upon induction of oxidative stress was assayed by non-reducing denaturing SDS-PAGE. To this end, AHK5 A3₂ wild-type (wt) and a Cys3-deficient mutant (C3A) as well as the A2 wt were transferred to A3₂-SEC II and A2 main buffers, respectively, lacking any reduction agents. The proteins were subsequently challenged with hydrogen peroxide dilution series ranging from 1 pM to 2 mM for 15 min at RT. The reactions were then mixed with appropriate amounts of 4x protein sample buffer without β-mercaptoethanol, heated to 95 °C for 2 min, and analysed by SDS-PAGE.

3.4 Production and Purification of GatD/MurT

3.4.1 List of buffers

Table 3-34: Buffers used for GatD/MurT purification.

Name	Composition	
Lysis	50 mM	Tris pH 8.5 @ 4°C (7.9 @ RT)
	300 mM	NaCl
	10 mM	Imidazole
	1 mM	DTT (add freshly)
Wash	as Lysis buffer	
	50 mM	Imidazole
Elution	as Lysis buffer	
	250 mM	Imidazole
SEC	50 mM	Tris pH 8.5 @ 4°C (7.9 @ RT)
	500 mM	NaCl
	50 mM	MgCl ₂
	5 mM	KCl
	5 mM	DTT (add freshly, use 5 mM TCEP for storage)

3.4.2 Cloning of expression constructs

GatD/MurT constructs used in this study are shown in Table 3-35. All *S. aureus* constructs were generated in the group of Prof. Dr. Tanja Schneider (University of Bonn) as described in [107] and [79]. Constructs 41-44 were bought from BioCat GmbH (Heidelberg, Germany), the remaining constructs were cloned by Paul Bachmann as a part of his M. Sc. project.

Table 3-35: Plasmids used for protein expression.

DB #	protein	mutation	vector	tag	comments
26	<i>Sa</i> GatD/MurT	wt	pET21b	GatD-C-His ₆	original construct used for crystallographic studies [79]
35	<i>Sa</i> GatD/MurT	GatD C94S	pET21b	GatD-C-His ₆	used in [79]
36	<i>Sa</i> GatD/MurT	MurT D349N	pET21b	GatD-C-His ₆	used in [79]
37	<i>Sa</i> GatD/MurT	MurT T60A	pET21b	GatD-C-His ₆	used in [79]
38	<i>Sa</i> GatD/MurT	MurT E108A	pET21b	GatD-C-His ₆	used in [79]
39	<i>Sa</i> GatD/MurT	MurT N267Y	pET21b	GatD-C-His ₆	used in [79]
40	<i>Sa</i> GatD/MurT	MurT T60A E108A	pET21b	GatD-C-His ₆	used in [79]
41	<i>Mtb</i> MurT	wt	pETDuet-1	-	intermediate cloning product – not used
42	<i>Mtb</i> GatD/MurT	wt	pETDuet-1	GatD-C-His ₆	codon-optimised expression vector, used in section 2.3
43	<i>Spy</i> MurT	wt	pETDuet-1	-	intermediate cloning product – not used

44	<i>SpyGatD/MurT</i>	wt	pETDuet-1	GatD-C-His ₆	codon-optimised expression vector, very low yield
46	<i>MtbGatD/MurT</i>	MurT Δ 1-36	pETDuet-1	GatD-C-His ₆	used in section 2.3, increased yield
50	<i>MtbGatD/MurT</i>	wt	pETM-11	MurT-N-His ₆ -TEV	not used
58	<i>SpyGatD/MurT</i>	wt	pETM-11	MurT-N-His ₆ -TEV	not used
59	<i>SpyGatD/MurT</i>	MurT Δ 1-34	pETM-11	MurT-N-His ₆ -TEV	high yield, used in section 2.3

3.4.3 Expression of *SaGatD/MurT* and *SpyGatD/MurT* constructs

Plasmids containing the genes for *SaGatD/MurT* or *SpyGatD/MurT* (Table 3-35) were transformed into BL21 (DE3) RIL *E. coli* cells (Stratagene, La Jolla, CA, USA) as described in Section 3.2.2.1. 10 mL LB suspension cultures containing the appropriate selection antibiotics were grown for 5-8 h at 37 °C and then used to inoculate 350 mL LB-amp/kan. This large pre-culture was grown O/N at 37 °C. The next day, 1.5 L expression cultures (LB-amp/kan) were inoculated with 30-40 mL of the O/N-culture each and grown at 37 °C until an OD₆₀₀ value of 0.4-0.5 was reached. The temperature was subsequently decreased to 30 °C and 0.25 mM IPTG were added for induction. Protein expression was carried out for 5 h and the cells were subsequently harvested by centrifugation using a Sorvall RC6+ (Thermo Fisher Scientific, Waltham, MA, USA) centrifuge equipped with an SLC-4000 rotor (10 000 x g, 15 min). The cells were stored at -20 °C until purification.

3.4.4 Expression of *MtbGatD/MurT* constructs

O/N LB pre-cultures were prepared as described above. 20 mL of the pre-culture were used to inoculate 1 L Simplified Studier autoinduction medium [154] (Table 3-21). Autoinduction cultures were grown at 20 °C, typically over the weekend (72 h). The cells were harvested as described above and stored at -20 °C until purification.

3.4.5 Protein purification

All GatD/MurT constructs were purified according to the general protocol published in [79], with minor variations. The protocol aims to initially enrich the target protein by Immobilised Metal ion Affinity Chromatography (IMAC) followed by immediate buffer exchange to remove imidazole. The resulting protein was then subjected to two subsequent rounds of preparative size exclusion chromatography (SEC) to reach a high degree of homogeneity. Proteolytically cleavable affinity tags (if present) were removed between the two SEC steps (Figure 3-3).

Adaptation of the original protocol to *M. tuberculosis* and *S. pyogenes* constructs was mainly performed in collaboration with Paul Bachmann as a part of his M. Sc. project.

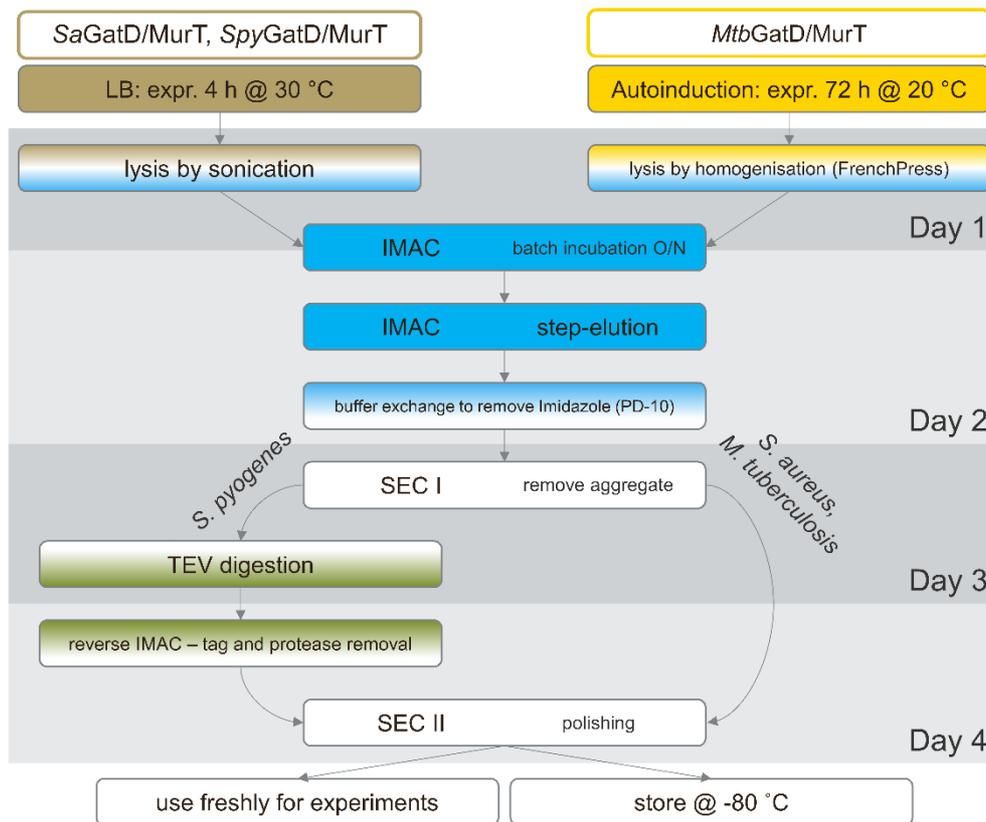


Figure 3-3: Schematic overview of the GatD/MurT purification strategy. Imidazole-containing steps are coloured in blue, all steps performed in standard SEC buffer (Table 3-34) are white. The presence of TEV protease is denoted in green. Grey boxes in the background highlight the timeline used.

3.4.5.1 Lysis by high-pressure homogenisation

E. coli pellets derived from autoinduction expression (i.e. all *Mtb* constructs) were lysed by high pressure homogenisation using an Emulsiflex C3 (Avestin, Canada). Pellets were resuspended in 20 mL of Lysis buffer (Table 3-34) per gram of cell wet weight. Hen egg lysozyme (200 µg/mL) and one tablet of cComplete[®] EDTA-free Protease Inhibitor Cocktail (Roche Diagnostics GmbH, Mannheim, Germany) were added, and the slurry was incubated at 4 °C for 30 min while stirring. After addition of 1:10 000 volume units of Benzonase[®] (Sigma-Aldrich, St. Louis, MO, USA), cells were lysed by repeatedly flowing them through the ice-cooled homogeniser. The device had previously been equilibrated with Lysis buffer, and the air pressure applied to the homogenising valve was constantly adjusted to maintain a hydrostatic pressure on the cells of 800-1 000 bar.

The lysate was subsequently clarified by centrifugation (Sorvall RC6+, 4 °C, 15 min, 12 000x g).

3.4.5.2 Lysis by sonication

Cells harvested from LB expression (*Sa* and *Spy* constructs) were lysed by sonication. The cells were resuspended in 10 mL Lysis buffer per gram of cell wet weight and additives and incubation were performed as described in 3.3.6.1. Cell disruption was performed in a Branson Digital Sonifier 450 using 20 % amplitude and 1 s pulses alternating with 2 s pauses for a total pulse duration of 10 min. The

sample was kept on ice throughout sonication. The lysate was clarified by centrifugation as described above.

3.4.5.3 Immobilised Metal Ion Affinity Chromatography (IMAC)

Following lysis, all GatD/MurT constructs were subjected to IMAC as a first purification step using Ni-NTA agarose beads (Qiagen, Hilden, Germany). 2 mL of 50 % bead slurry, pre-equilibrated with Lysis buffer, were added to 15 mL of clarified lysate and incubated for 17 h at 4 °C while stirring. The mixture was subsequently transferred to EconoPac gravity flow columns (BioRad, Hercules, CA, USA), unspecifically bound proteins were removed by washing with 10 CV of Wash buffer containing 50 mM imidazole, and the remaining bound protein was eluted with Elution buffer containing 250 mM imidazole. For *SpyGatD/MurT* washing was carried out with 10 CV of Lysis buffer. Elution was carried out with 5 CV of Wash buffer followed by 5 CV of Elution buffer. Appropriate fraction sizes were collected and pooled after protein size estimation by SDS-PAGE.

3.4.5.4 Buffer exchange

A semi-automated buffer exchange strategy was employed to remove imidazole from the protein. To this end, the protein was applied to a HiPrep 26/10 Desalting column that had previously been equilibrated with buffer SEC (5 mM DTT). The column was operated via an ÄKTAprime FPLC system (GE Healthcare, Uppsala, Sweden) at a flow rate of 10 mL/min. 15 mL IMAC elution were applied and fractions were collected from 14 to 36 mL. Separation of protein and imidazole was monitored via changes in the absorption at 280 nm and the electrical conductivity.

3.4.5.5 Size Exclusion Chromatography (SEC)

The protein sample was concentrated to 2.5 mg/mL using centrifugal concentrators (Amicon Ultra-15, 50 kDa MWCO, 1 880 x g) and applied to a HiLoad 26/60 column packed with SD200 (GE Healthcare, Uppsala, Sweden) that had previously been equilibrated with buffer SEC (5 mM DTT). Elution was carried out at 3 mL/min while monitoring UV absorption at 280 nm. Relevant peak fractions were checked for protein purity by SDS-PAGE and pooled accordingly.

To further enhance sample quality, a second iteration of SEC was performed for polishing as described above but using 5 mM TCEP instead of DTT as a reducing agent for long-term storage.

3.4.5.6 Affinity tag cleavage

In the case of *SpyGatD/MurT*_{Δ1-34}, which was expressed from a pETM-11 vector, the N-terminal His₆-tag for Ni-affinity purification was removed by site-specific proteolysis between the first and second SEC steps using the Tobacco Etch Virus (TEV) protease. 0.5 mg of protease were added per 5 mg of target protein and incubated O/N at 4 °C. The digested protein was separated from undigested one as well as the residual tag and the TEV protease (also His-tagged) by means of a 1 mL pre-packed Ni-NTA cartridge (HiTrap Crude FF, GE Healthcare, Uppsala, Sweden).

3.4.6 Biophysical characterisation

Sample homogeneity and stability were assessed by analytical SEC, DLS, CD and DSA, as described in Sections 3.2.3.3, 3.2.3.4, 3.2.3.5, and 3.2.3.8, respectively.

3.4.7 Structural biology

3.4.7.1 Crystallisation screening

GatD/MurT samples were concentrated to 2-5 mg/mL. Suitability of protein concentration for initial crystallisation trials was estimated using an ammonium sulfate precipitation assay. Briefly, 1 μ L of protein was mixed with 1 μ L of a $(\text{NH}_4)_2\text{SO}_4$ solution of either 1, 1.5, 2, 2.5, or 3 M concentration, and the propensity to precipitation was assessed. Protein concentration was deemed suitable for crystallisation if the top two ammonium sulfate conditions lead to immediate precipitation while the bottom two remained clear for at least one minute.

The protein was then filtered using Costar Spin-X Centrifuge Tube Filters with 0.22 μ m Cellulose Acetate membrane (Sigma-Aldrich, St. Louis, MO, USA). Sitting-drop vapour diffusion screens were set up with a Crystal Gryphon LCP (Art Robbins Instruments, Sunnyvale, CA, USA) using protein drop sizes ranging from 100 to 300 nL and protein to precipitant ratios of 2:1, 1:1, or 1:2.

Initial screening was carried out using commercially available crystallisation screens (Table 3-9). After setup, crystallisation trays were stored either at 4 °C or at 20 °C and manually inspected after one day, three days, one week, two weeks, one, and two months using a Leica Digital Microscope (Leica, Wetzlar, Germany). Upon identification of a promising condition, fine screens were designed to refine crystallisation. 96-well-based refinement screens were prepared from stock solutions using a Tecan Freedom EVO liquid handler (Männedorf, Switzerland). Crystallisation trays were again pipetted with the Crystal Gryphon LCP.

In order to screen a wide variety of additives, such as various salts or alcohols, a previously determined crystallisation condition was supplemented 1:10 with Additive Screen (Hampton Research, Aliso Viejo, CA, USA) in a 96 well layout and used for sitting drop crystallisation.

Crystal growth was often enhanced by seeding. To this end, crystals grown from vapour diffusion experiments were suspended in 50 μ L well solution and crushed by vortexing using SeedBeads[®] (Hampton Research, Aliso Viejo, CA, USA). The thus obtained seed stock was either used directly for streak-seeding using a cat whisker (Provided by Dr. Ancilla Neu, collected from her cat "Pauli") or diluted and added 1:10 to newly set crystallisation drops.

3.4.7.2 Crystallisation

Crystals for diffraction experiments were mostly obtained from hanging drop vapour diffusion trays. Herein, 1 μ L of protein solution was mixed with 1 μ L of well solution and suspended over a reservoir containing approximately 500 μ L well solution. If required, the drops were seeded as described above. The composition of the crystallisation conditions for *Sa*GatD/MurT as well as crystal derivatisation strategies are published in [79].

3.4.7.3 Diffraction data collection, data processing, model building and refinement

All crystallographic procedures were carried out as described in [79].

3.4.7.4 Small-angle X-ray scattering (SAXS)

Preliminary SAXS experiments were performed in the laboratory of Prof. Frank Schreiber with the kind support of Dr. Olga Matsarskaia, Dr. Stefano da Vela, and Dr. Fajun Zhang on a Xeuss 2.0 X-ray diffractometer (Xenocs, Grenoble, France) equipped with a Pilatus 300k detector (Dectris, Baden-Daettwil, Switzerland) and a microfocus X-ray source. Protein concentration ranging from 0.5 to 11 mg/mL were supplied in 1 mm quartz capillaries in a temperature-controlled sample holder (Linkam FFSX350, Linkam, Waterfield, UK) pre-cooled to 12 °C. Measurements were performed at full beam transmission over multiple hours subdivided in single frames of 30 min exposure each in order to assess radiation damage. Frames were integrated with the Foxtrot data reduction software (Synchrotron Soleil, Saint-Aubin, France) and the scattering profiles were analysed using components of the ATSAS suite [155].

SAXS scattering profiles for GatD/MurT were recorded at beamline P12 at the German Electron Synchrotron (DESY, Hamburg, Germany) as described in [79].

Additionally, profiles for AHK5 A₃ were kindly recorded by Dr. Olga Matsarskaia at beamline ID02 at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France. Data obtained from different sample concentrations were used to extrapolate a theoretical scattering profile at zero concentration to remove possible concentration effects using the PRIMUS interface of the ATSAS program suite [155], [156]. Radius of gyration (R_g) determination by linear fitting of the Guinier region and extraction of a distance distribution function were also performed in PRIMUS.

3.4.8 Enzymatic activity assay

An NMR-based glutaminase assay was used to assess the capability of GatD/MurT proteins to hydrolyse glutamine to glutamate in vitro. The protocol was adapted from [141].

SEC buffer (Table 3-34) without any reducing agent was lyophilised and reconstituted in 99.8 % D₂O to yield NMR buffer. GatD/MurT was transferred into NMR buffer by means of a PD-10 desalting column (GE healthcare, Uppsala, Sweden) using the manufacturer's gravity flow protocol.

GatD/MurT was supplied at a concentration of 30 µM in NMR buffer and the glutaminase reaction was started by adding 750 µM of glutamine (25-fold excess). ¹H NMR spectra were recorded by Dr. Vincent Truffault (Max Planck Institute for Developmental Biology, Tübingen, Germany) on a Bruker AVIII-600 spectrometer with a room temperature sample holder at 300 K with a spectral resolution of 0.66 Hz. Scans were repeated with every 68 s for a total of 2400 scans in a pseudo-2D experiment. The spectra were processed in the Bruker TopSpin® software (Bruker, Billerica, MA, USA). Protein-only spectra were recorded and subtracted to yield the clean substrate and product spectra. The pseudo-2D experiments were plotted over the entire time course of the experiment and compared with glutamine-only as well as glutamate-only reference spectra. Relevant peak changes were tracked using the kinetic analysis function in the Bruker Dynamics Center software.

SUMMARY

This thesis presents recent advances on the structural and mechanistical characterisation of two ATP-binding proteins or protein complexes; AHK5 from *Arabidopsis thaliana*, and GatD/MurT from *Staphylococcus aureus*, *Mycobacterium tuberculosis*, and *Streptococcus pyogenes*.

AHK5 is a sensory histidine kinase that is involved in multiple physiological pathways, such as the modulation of stomatal closure, or the response to PAMPs. Its activity is likely triggered by increases in hydrogen peroxide concentrations. In this project, a tentative model for hydrogen peroxide sensing by the N-terminal portion of AHK5 was postulated. Based on *in-vitro* dimerisation studies, a cysteine residue at position 3 was shown to covalently link two AHK5 monomers under oxidative conditions, with a redox midpoint potential compatible with previously published physiological values. The study additionally lays a comprehensive foundation for future work to elucidate the structure and thus the details of AHK5 activation. Three different fragments which, together, span the sensory N-terminus and the histidine kinase and ATPase domains were successfully purified. Bioinformatic methods as well as a protease exclusion assay indicated the probable presence of a yet undescribed PAS-like domain, immediately preceding the histidine kinase domain. This domain could play a key role in transducing the input signal from the N-terminus to the catalytically active part. However, biophysical analyses highlight potential heterogeneity with respect to the purified proteins' oligomeric state and stability problems. These issues will be the subject of follow-up studies before structural work can begin.

Understanding the mechanism of action of AHK5 will not only provide additional understanding of how histidine kinases work; due to its key role in numerous physiological processes, it could prove an anchor point from which to deepen our understanding of how *Arabidopsis* reacts to exogenous stresses as well as plant physiology in general.

GatD/MurT is a cell wall amidating enzyme complex that is essential in many Gram-positive bacteria, including highly pathogenic ones like *Staphylococcus aureus*, *Streptococcus pneumoniae*, and even the actinobacterium *Mycobacterium tuberculosis*. In this project, the first fully refined crystal structure of the *S. aureus* enzyme was obtained and a complex crystal structure containing an ATP analogue mimicking one of the three substrates was solved. Together with solution scattering data obtained by SAXS, these data indicate an intriguing crescent-shaped open conformation of the heterodimer. This conformation likely represents the enzyme's resting state and probably transitions to a more compact conformation upon substrate binding, in order to achieve catalysis. The enzyme complex exhibits a canonical ATP binding pocket in MurT. GatD achieves intermediate ammonia generation by deamidation of cytosolic glutamine. This reaction is catalysed by an unusual intermolecular catalytic triad composed of a cysteine and a histidine in GatD and an aspartate in MurT. Additionally, MurT contains a Cys₄-type Zinc finger of unknown function. Expression of putative homologous complexes from *M. tuberculosis* and *Streptococcus pyogenes* was successful, and preliminary glutaminase activity monitored by ¹H-NMR confirmed both enzymes to be active in this respect.

These insights raise the stakes in understanding the molecular mechanism of action of GatD/MurT, as it could offer the possibility to develop novel antimicrobial drugs that specifically target it. Such drugs could be active not only against *S. aureus* but rather a wide range of pathogenic bacteria, thus providing a new edge in combating antibiotic resistances.

ZUSAMMENFASSUNG

Diese Dissertation präsentiert neue Erkenntnisse in Bezug auf die strukturelle und mechanistische Beschreibung zweier ATP-bindender Proteine oder Komplexe; AHK5 aus *Arabidopsis thaliana* und GatD/MurT aus *Staphylococcus aureus*, *Mycobacterium tuberculosis* und *Streptococcus pyogenes*.

AHK5 ist eine sensorische Histidinkinase, die an zahlreichen physiologischen Signalwegen beteiligt ist. Beispielsweise moduliert sie die Schließung der Stomata oder reguliert die Antwort auf PAMPs. Vermutlich wird AHK5 durch einen Anstieg der Wasserstoffperoxidkonzentration aktiviert. Im Rahmen dieses Projektes wurde ein mögliches mechanistisches Modell postuliert, wie Erkennung von H₂O₂ durch AHK5 ablaufen könnte. Basierend auf *in-vitro*-Dimerisierungsexperimenten konnte gezeigt werden, dass ein Cystein auf Position 3 unter oxidativen Bedingungen zwei AHK5-Monomere kovalent verbinden kann. Das Redox-Midpoint-Potential liegt herbei in einem Bereich, der kompatibel ist mit veröffentlichten physiologischen Werten. Diese Arbeit liefert zudem eine umfangreiche Grundlage für künftige Strukturaufklärungsarbeit an AHK5 und damit für das mechanistische Verständnis der H₂O₂-Steuerung der Kinaseaktivität. Drei verschiedene Fragmente, die zusammen den gesamten sensorischen N-Terminus sowie die Kinase- und ATPase-Domänen enthalten, wurden erfolgreich aufgereinigt. Sowohl bioinformatische Methoden als auch ein Proteasenexklusionsassay deuten darauf hin, dass sich unmittelbar N-terminal zur Kinasedomäne eine vermutlich PAS-ähnliche Domäne befindet. Diese Domäne könnte eine zentrale Rolle im Weiterreichen des Aktivierungssignals vom N-Terminus hin zu Kinasedomäne spielen. Allerdings zeigten biophysikalische Analysen Heterogenität in der oligomeren Zusammensetzung der gereinigten Proteine sowie Stabilitätsprobleme. Diese Schwierigkeiten werden der Inhalt weiterführender Studien sein, bevor die Arbeit an der Strukturaufklärung beginnen kann.

Den molekularen Mechanismus von AHK5 zu entschlüsseln wird nicht nur das allgemeine Verständnis von Histidinkinasen erweitern; aufgrund der zentralen Rolle von AHK5 könnte es auch einen Ankerpunkt darstellen von dem aus das Wissen darüber wie *Arabidopsis* auf exogenen Stress reagiert, sowie unser Verständnis von Pflanzenphysiologie im Allgemeinen, zu erweitern.

GatD/MurT ist ein Enzymkomplex, der für Amidierungen in der Zellwand verantwortlich ist. Dieses Enzym ist in vielen Gram-positiven Bakterien essenziell, beispielsweise in hochpathogenen Spezies wie *Staphylococcus aureus*, *Streptococcus pneumoniae* und sogar im Actinobakterium *Mycobacterium tuberculosis*. Im Rahmen dieses Projektes wurde die erste vollständig verfeinerte Kristallstruktur des *S. aureus*-Enzyms erhalten und eine Komplexstruktur mit AMPPNP anstelle von ATP als eines der drei physiologischen Substrate wurde gelöst. Zusammen mit Daten über das Molekül in Lösung, die durch Röntgenkleinwinkelstreuexperimente erhalten wurden, zeigen die kristallographischen Daten auf, dass *S. aureus* GatD/MurT eine sehr interessante offene sichelförmige Konformation besitzt. Diese Konformation stellt vermutlich den freien Grundzustand des Enzyms dar und wird sich wahrscheinlich nach der Bindung aller Substrate zu einer kompakteren geschlossenen Form wandeln, in welcher die katalysierte Reaktion stattfinden kann. Der Komplex enthält eine kanonische ATP-Bindetasche in MurT. GatD generiert vorübergehend Ammoniak durch die Deamidierung von freiem cytosolischem Glutamin. Diese Reaktion wird durch eine ungewöhnliche intermolekulare katalytische Triade bewerkstelligt, die aus einem Cystein und einem Histidin aus GatD, sowie einem Aspartat aus MurT besteht. Zudem besitzt MurT einen Cys₄-Zinkfinger unbekannter Funktion. Die Expression möglicher homologer Komplexe aus *M. tuberculosis* und *Streptococcus pyogenes* war erfolgreich. Vorversuche

zur Überprüfung der Glutaminaseaktivität mittels $^1\text{H-NMR}$ bestätigten, dass beide Enzyme diesbezüglich aktiv sind.

Diese Erkenntnisse erhöhen die Wichtigkeit, den molekularen Katalysemechanismus von GatD/MurT zu verstehen, da hierauf basierend neue antimikrobielle Wirkstoffe entwickelt werden könnten um selektiv GatD/MurT zu inhibieren. Solche Wirkstoffe könnten dann nicht nur gegen *S. aureus*, sondern gegen eine Vielzahl an pathogenen Bakterien wirksam sein und somit eine neue Waffe im Kampf gegen antibiotikaresistente Erreger darstellen.

4 BIBLIOGRAPHY

- [1] R. Sever and C. K. Glass, "Signaling by nuclear receptors.," *Cold Spring Harb. Perspect. Biol.*, vol. 5, no. 3, p. a016709, Mar. 2013.
- [2] M. J. Lohse and D. Calebiro, "Cell biology: Receptor signals come in waves.," *Nature*, vol. 495, no. 7442, pp. 457–8, Mar. 2013.
- [3] E. Rozengurt, J. Sinnott-Smith, and K. Kisfalvi, "Crosstalk between insulin/insulin-like growth factor-1 receptors and G protein-coupled receptor signaling systems: a novel target for the antidiabetic drug metformin in pancreatic cancer.," *Clin. Cancer Res.*, vol. 16, no. 9, pp. 2505–11, May 2010.
- [4] B. Vanhaesebroeck, J. Guillermet-Guibert, M. Graupera, and B. Bilanges, "The emerging mechanisms of isoform-specific PI3K signalling.," *Nat. Rev. Mol. Cell Biol.*, vol. 11, no. 5, pp. 329–41, May 2010.
- [5] N. C. Law, M. F. White, and M. E. Hunzicker-Dunn, "G protein-coupled receptors (GPCRs) That Signal via Protein Kinase A (PKA) Cross-talk at Insulin Receptor Substrate 1 (IRS1) to Activate the phosphatidylinositol 3-kinase (PI3K)/AKT Pathway.," *J. Biol. Chem.*, vol. 291, no. 53, pp. 27160–27169, 2016.
- [6] D. M. Rosenbaum, S. G. F. Rasmussen, and B. K. Kobilka, "The structure and function of G-protein-coupled receptors.," *Nature*, vol. 459, no. 7245, pp. 356–63, May 2009.
- [7] B. Luo, C. V. Carman, and T. A. Springer, "Structural basis of integrin regulation and signaling.," *Annu. Rev. Immunol.*, vol. 25, pp. 619–47, 2007.
- [8] D. K. Morrison, "MAP kinase pathways.," *Cold Spring Harb. Perspect. Biol.*, vol. 4, no. 11, pp. 892–908, Nov. 2012.
- [9] T. Kawasaki and T. Kawai, "Toll-like receptor signaling pathways.," *Front. Immunol.*, vol. 5, no. SEP, p. 461, 2014.
- [10] M. B. Phillips, A. Nigam, and J. W. Johnson, "Interplay between Gating and Block of Ligand-Gated Ion Channels.," *Brain Sci.*, vol. 10, no. 12, pp. 1–22, Dec. 2020.
- [11] F. H. Yu, V. Yarov-Yarovoy, G. A. Gutman, and W. A. Catterall, "Overview of molecular relationships in the voltage-gated ion channel superfamily.," *Pharmacol. Rev.*, vol. 57, no. 4, pp. 387–95, Dec. 2005.
- [12] J. Schlessinger, "Cell signaling by receptor tyrosine kinases.," *Cell*, vol. 103, no. 2, pp. 211–25, Oct. 2000.
- [13] Z. Du and C. M. Lovly, "Mechanisms of receptor tyrosine kinase activation in cancer.," *Mol. Cancer*, vol. 17, no. 1, p. 58, 2018.
- [14] J. Boucher, A. Kleinriders, and C. R. Kahn, "Insulin receptor signaling in normal and insulin-resistant states.," *Cold Spring Harb. Perspect. Biol.*, vol. 6, no. 1, p. a009191, Jan. 2014.
- [15] A. Tomas, C. E. Futter, and E. R. Eden, "EGF receptor trafficking: consequences for signaling and cancer.," *Trends Cell Biol.*, vol. 24, no. 1, pp. 26–34, Jan. 2014.
- [16] R. Derynck and E. H. Budi, "Specificity, versatility, and control of TGF- β family signaling.," *Sci. Signal.*, vol. 12, no. 570, 2019.
- [17] C. H. Heldin and A. Moustakas, "Signaling receptors for TGF- β family members.," *Cold Spring Harb. Perspect. Biol.*, vol. 8, no. 8, pp. 1–34, 2016.
- [18] A. J. Afzal, A. J. Wood, and D. A. Lightfoot, "Plant receptor-like serine threonine kinases: roles in signaling and plant defense.," *Mol. Plant. Microbe. Interact.*, vol. 21, no. 5, pp. 507–17, May 2008.
- [19] B. Kemmerling, T. Halter, S. Mazzotta, S. Mosher, and T. Nürnberger, "A genome-wide survey for Arabidopsis leucine-rich repeat receptor kinases implicated in plant immunity.," *Front. Plant Sci.*, vol. 2, no. NOV, p. 88, 2011.
- [20] B. S. Blaum *et al.*, "Structure of the pseudokinase domain of BIR2, a regulator of BAK1-mediated immune signaling in Arabidopsis.," *J. Struct. Biol.*, vol. 186, no. 1, pp. 112–21, Apr. 2014.
- [21] G.-Z. Han, "Origin and evolution of the plant immune system.," *New Phytol.*, vol. 222, no. 1, pp.

- 70–83, 2019.
- [22] P. V. Attwood, “Histidine kinases from bacteria to humans,” *Biochem. Soc. Trans.*, vol. 41, no. 4, pp. 1023–1028, 2013.
- [23] I. Letunic and P. Bork, “20 years of the SMART protein domain annotation resource,” *Nucleic Acids Res.*, vol. 46, no. D1, pp. D493–D496, 2018.
- [24] B. Pekárová, A. Szmítowska, R. Dopitová, O. Degtjarik, L. Žídek, and J. Hejátko, “Structural Aspects of Multistep Phosphorelay-Mediated Signaling in Plants,” *Molecular Plant*, vol. 9, no. 1, pp. 71–85, 2016.
- [25] A. M. Stock, V. L. Robinson, and P. N. Goudreau, “Two-component signal transduction,” *Annu. Rev. Biochem.*, vol. 69, no. 1, pp. 183–215, 2000.
- [26] G. E. Schaller, S.-H. Shiu, and J. P. Armitage, “Two-Component Systems and Their Co-Option for Eukaryotic Signal Transduction,” *Curr. Biol.*, vol. 21, no. 9, pp. R320–R330, 2011.
- [27] R. Gao and A. M. Stock, “Biological Insights from Structures of Two-Component Proteins,” *Annu. Rev. Microbiol.*, vol. 63, no. 1, pp. 133–154, 2009.
- [28] K. Caesar *et al.*, “Evidence for the localization of the Arabidopsis cytokinin receptors AHK3 and AHK4 in the endoplasmic reticulum,” *J. Exp. Bot.*, vol. 62, no. 15, pp. 5571–5580, 2011.
- [29] J. Lohrmann and K. Harter, “Plant two-component signaling systems and the role of response regulators,” *Plant Physiol.*, vol. 128, no. 2, pp. 363–369, 2002.
- [30] T. Urao, K. Yamaguchi-Shinozaki, and K. Shinozaki, “Plant histidine kinases: an emerging picture of two-component signal transduction in hormone and environmental responses,” *Sci. STKE*, vol. 2001, no. December 2007, p. re18, 2001.
- [31] M. Hothorn, T. Dabi, and J. Chory, “Structural Basis for Cytokinin Recognition by Arabidopsis thaliana Histidine Kinase 4,” *Nat. Chem. Biol.*, vol. 7, no. 11, pp. 766–8, 2011.
- [32] R. Dautel, X. N. Wu, M. Heunemann, W. X. Schulze, and K. Harter, “The sensor histidine kinases AHK2 and AHK3 proceed into multiple serine/threonine/tyrosine phosphorylation pathways in Arabidopsis thaliana,” *Mol. Plant*, 2015.
- [33] C. Grefen and K. Harter, “Plant two-component systems: Principles, functions, complexity and cross talk,” *Planta*, vol. 219, no. 5, pp. 733–742, 2004.
- [34] I. Hwang, H. H. Chen, and J. Sheen, “Two-component signal transduction pathways in Arabidopsis,” *Plant Physiol.*, vol. 129, no. 2, pp. 500–515, 2002.
- [35] T. Urao *et al.*, “A Transmembrane Hybrid-Type Histidine Kinase in Arabidopsis Functions as an Osmosensor,” *Plant Cell Online*, vol. 11, no. 9, pp. 1743–1754, 1999.
- [36] R. Desikan *et al.*, “The histidine kinase AHK5 integrates endogenous and environmental signals in Arabidopsis guard cells,” *PLoS One*, vol. 3, no. 6, p. e2491, 2008.
- [37] E. Marchadier and A. M. Hetherington, “Involvement of two-component signalling systems in the regulation of stomatal aperture by light in Arabidopsis thaliana,” *New Phytol.*, vol. 203, no. 2, pp. 462–468, 2014.
- [38] J. Pham, J. Liu, M. H. Bennett, J. W. Mansfield, and R. Desikan, “Arabidopsis histidine kinase 5 regulates salt sensitivity and resistance against bacterial and fungal infection,” *New Phytol.*, vol. 194, no. 1, pp. 168–80, 2012.
- [39] V. Mira-Rodado, M. Veerabagu, J. Witthöft, J. Teply, K. Harter, and R. Desikan, “Identification of two-component system elements downstream of AHK5 in the stomatal closure response of Arabidopsis thaliana,” *Plant Signal. Behav.*, vol. 7, no. 11, pp. 1467–76, 2012.
- [40] J. Bauer, K. Reiss, M. Veerabagu, M. Heunemann, K. Harter, and T. Stehle, “Structure–Function Analysis of Arabidopsis thaliana Histidine Kinase AHK5 Bound to Its Cognate Phosphotransfer Protein AHP1,” *Mol. Plant*, vol. 6, no. 3, pp. 959–970, 2013.
- [41] C. E. Hutchison and J. J. Kieber, “Signaling via Histidine-Containing Phosphotransfer Proteins in Arabidopsis,” *Plant Signal. Behav.*, vol. 2, no. 4, pp. 287–9, Jul. 2007.
- [42] T. Mizuno, “Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of Escherichia coli,” *DNA Res.*, vol. 4, no. 2, pp. 161–8, Apr. 1997.
- [43] C. E. Hutchison *et al.*, “The Arabidopsis histidine phosphotransfer proteins are redundant positive regulators of cytokinin signaling,” *Plant Cell*, vol. 18, no. 11, pp. 3073–87, Nov. 2006.
- [44] R. Nongpiur, P. Soni, R. Karan, S. Singla-Pareek, and A. Pareek, “Histidine kinases in plants: Cross

- talk between hormone and stress responses," *Plant Signal. Behav.*, vol. 7, no. 10, pp. 1230–1237, 2012.
- [45] A. B. Bleeker and H. Kende, "Ethylene: a gaseous signal molecule in plants.," *Annu. Rev. Cell Dev. Biol.*, vol. 16, no. 1, pp. 1–18, 2000.
- [46] C. Merchante, J. M. Alonso, and A. N. Stepanova, "Ethylene signaling: simple ligand, complex regulation.," *Curr. Opin. Plant Biol.*, vol. 16, no. 5, pp. 554–60, Oct. 2013.
- [47] C. Ueguchi, S. Sato, T. Kato, and S. Tabata, "The AHK4 Gene Involved in the Cytokinin-Signaling Pathway as a Direct Receptor Molecule in *Arabidopsis thaliana*," *Plant Cell Physiol.*, vol. 42, no. 7, pp. 751–755, Jul. 2001.
- [48] M. Higuchi *et al.*, "In planta functions of the *Arabidopsis* cytokinin receptor family," *Proc. Natl. Acad. Sci.*, vol. 101, no. 23, pp. 8821–8826, Jun. 2004.
- [49] L.-S. P. Tran *et al.*, "Functional analysis of AHK1/ATHK1 and cytokinin receptor histidine kinases in response to abscisic acid, drought, and salt stress in *Arabidopsis*," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 51, pp. 20623–20628, 2007.
- [50] J. Hejatko *et al.*, "The histidine kinases CYTOKININ-INDEPENDENT1 and ARABIDOPSIS HISTIDINE KINASE2 and 3 regulate vascular tissue development in *Arabidopsis* shoots.," *Plant Cell*, vol. 21, no. 7, pp. 2008–21, Jul. 2009.
- [51] V. Anantharaman and L. Aravind, "The CHASE domain: a predicted ligand-binding module in plant cytokinin receptors and other eukaryotic and bacterial receptors.," *Trends Biochem. Sci.*, vol. 26, no. 10, pp. 579–82, Oct. 2001.
- [52] J. J. Kieber and G. E. Schaller, "Cytokinins.," *Arab. B.*, vol. 12, p. e0168, Jan. 2014.
- [53] Z. Zhang and W. a. Hendrickson, "Structural Characterization of the Predominant Family of Histidine Kinase Sensor Domains," *J. Mol. Biol.*, vol. 400, no. 3, pp. 335–353, 2010.
- [54] M. Hillion and H. Antelmann, "Thiol-based redox switches in prokaryotes.," *Biol. Chem.*, vol. 396, no. 5, pp. 415–44, May 2015.
- [55] S. Akter *et al.*, "Cysteines under ROS attack in plants: a proteomics view.," *J. Exp. Bot.*, vol. 66, no. 10, pp. 2935–44, May 2015.
- [56] C. Waszczak *et al.*, "Sulfenome mining in *Arabidopsis thaliana*," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 31, pp. 11545–50, Aug. 2014.
- [57] H. S. Marinho, C. Real, L. Cyrne, H. Soares, and F. Antunes, "Hydrogen peroxide sensing, signaling and regulation of transcription factors.," *Redox Biol.*, vol. 2, no. 1, pp. 535–62, 2014.
- [58] C. Waszczak, M. Carmody, and J. Kangasjarvi, "Reactive Oxygen Species in Plant Signaling.," *Annu. Rev. Plant Biol.*, vol. 69, no. 15, pp. 209–236, 2018.
- [59] R. B. Russell, J. Breed, and G. J. Barton, "Conservation analysis and structure prediction of the SH2 family of phosphotyrosine binding domains.," *FEBS Lett.*, vol. 304, no. 1, pp. 15–20, Jun. 1992.
- [60] J. Key, M. Hefti, E. B. Purcell, and K. Moffat, "Structure of the Redox Sensor Domain of *Azotobacter vinelandii* NifL at Atomic Resolution: Signaling, Dimerization, and Mechanism † , ‡," *Biochemistry*, vol. 46, no. 12, pp. 3614–3623, Mar. 2007.
- [61] A. Moglich, R. a. Ayers, and K. Moffat, "Structure and Signaling Mechanism of Per-ARNT-Sim Domains," *Structure*, vol. 17, no. 10, pp. 1282–1294, 2009.
- [62] A. Moglich, R. A. Ayers, and K. Moffat, "Design and signaling mechanism of light-regulated histidine kinases.," *J. Mol. Biol.*, vol. 385, no. 5, pp. 1433–44, Feb. 2009.
- [63] R. A. Ayers and K. Moffat, "Changes in Quaternary Structure in the Signaling Mechanisms of PAS Domains † , ‡," *Biochemistry*, vol. 47, no. 46, pp. 12078–12086, Nov. 2008.
- [64] L. Zimmermann *et al.*, "A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core," *J. Mol. Biol.*, no. Table 1, pp. 1–7, 2017.
- [65] D. T. Jones, "Protein secondary structure prediction based on position-specific scoring matrices.," *J. Mol. Biol.*, vol. 292, no. 2, pp. 195–202, Sep. 1999.
- [66] D. W. A. Buchan and D. T. Jones, "The PSIPRED Protein Analysis Workbench: 20 years on.," *Nucleic Acids Res.*, vol. 47, no. W1, pp. W402–W407, Jul. 2019.
- [67] H. Ashkenazy *et al.*, "ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules.," *Nucleic Acids Res.*, vol. 44, no. W1, pp. W344–

- 50, 2016.
- [68] B. Mészáros, G. Erdos, and Z. Dosztányi, "IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding.," *Nucleic Acids Res.*, vol. 46, no. W1, pp. W329–W337, 2018.
- [69] G. Lindwall, M. Chau, S. R. Gardner, and L. A. Kohlstaedt, "A sparse matrix approach to the solubilization of overexpressed proteins.," *Protein Eng.*, vol. 13, no. 1, pp. 67–71, Jan. 2000.
- [70] D. I. Svergun, "Determination of the Regularization Parameter in Indirect- Transform Methods Using Perceptual Criteria," *J. Appl. Crystallogr.*, vol. 25, pp. 495–503, 1992.
- [71] M. Schwarzländer *et al.*, "Confocal imaging of glutathione redox potential in living plant cells," *J. Microsc.*, vol. 231, no. 2, pp. 299–316, Aug. 2008.
- [72] R. P. Diensthuber, M. Bommer, T. Gleichmann, and A. Möglich, "Full-Length Structure of a Sensor Histidine Kinase Pinpoints Coaxial Coiled Coils as Signal Transducers and Modulators," *Structure*, vol. 21, no. 7, pp. 1127–1136, 2013.
- [73] J. Söding, A. Biegert, and A. N. Lupas, "The HHpred interactive server for protein homology detection and structure prediction," *Nucleic Acids Res.*, vol. 33, no. SUPPL. 2, pp. 244–248, 2005.
- [74] A. Sali, L. Potterton, F. Yuan, H. Van Vlijmen, and M. Karplus, "Evaluation of comparative protein modeling by MODELLER," *Proteins Struct. Funct. Genet.*, vol. 23, no. 3, pp. 318–326, 1995.
- [75] V. Alva, S.-Z. Nam, J. Söding, and A. N. Lupas, "The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis," *Nucleic Acids Res.*, vol. 44, no. April 2016, p. gkw348, 2016.
- [76] F. Jacob-Dubuisson, A. Mechaly, J.-M. Betton, and R. Antoine, "Structural insights into the signalling mechanisms of two-component systems," *Nat. Rev. Microbiol.*, vol. 16, no. 10, pp. 585–593, 2018.
- [77] P. Casino, V. Rubio, and A. Marina, "Structural insight into partner specificity and phosphoryl transfer in two-component signal transduction.," *Cell*, vol. 139, no. 2, pp. 325–36, Oct. 2009.
- [78] P. Casino, V. Rubio, and A. Marina, "The mechanism of signal transduction by two-component systems," *Curr. Opin. Struct. Biol.*, vol. 20, no. 6, pp. 763–771, 2010.
- [79] E. R. Nöldeke *et al.*, "Structural basis of cell wall peptidoglycan amidation by the GatD/MurT complex of *Staphylococcus aureus*," *Sci. Rep.*, vol. 8, no. May, p. 12953, 2018.
- [80] B. M. Schulze, D. L. Watkins, J. Zhang, I. Ghiviriga, and R. K. Castellano, "Estimating the shape and size of supramolecular assemblies by variable temperature diffusion ordered spectroscopy.," *Org. Biomol. Chem.*, vol. 12, no. 40, pp. 7932–6, Oct. 2014.
- [81] H. P. Erickson, "Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy.," *Biol. Proced. Online*, vol. 11, no. 1, pp. 32–51, May 2009.
- [82] E. R. S. Kunji, M. Harding, P. J. G. Butler, and P. Akamine, "Determination of the molecular mass and dimensions of membrane proteins by size exclusion chromatography.," *Methods*, vol. 46, no. 2, pp. 62–72, Oct. 2008.
- [83] T. F. Clarke and P. L. Clark, "Rare codons cluster.," *PLoS One*, vol. 3, no. 10, p. e3412, 2008.
- [84] A. Rodriguez, G. Wright, S. Emrich, and P. L. Clark, "%MinMax: A versatile tool for calculating and comparing synonymous codon usage and its impact on protein folding.," *Protein Sci.*, vol. 27, no. 1, pp. 356–362, 2018.
- [85] P. L. Clark, "Protein folding in the cell: reshaping the folding funnel.," *Trends Biochem. Sci.*, vol. 29, no. 10, pp. 527–34, Oct. 2004.
- [86] J. L. Chaney and P. L. Clark, "Roles for Synonymous Codon Usage in Protein Biogenesis.," *Annu. Rev. Biophys.*, vol. 44, no. 1, pp. 143–66, 2015.
- [87] G. N. Jacobson and P. L. Clark, "Quality over quantity: optimizing co-translational protein folding with non-'optimal' synonymous codons.," *Curr. Opin. Struct. Biol.*, vol. 38, pp. 102–10, 2016.
- [88] A. Fleming, "On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*," *Br. J. Exp. Pathol.*, vol. 10, pp. 226–236, 1929.
- [89] C. L. Ventola, "The antibiotic resistance crisis: part 1: causes and threats.," *P T*, vol. 40, no. 4, pp. 277–83, Apr. 2015.
- [90] CDC, "Antibiotic resistance threats in the United States," *Centers for Disease Control and*

- Prevention website*, 2013. [Online]. Available: <http://www.cdc.gov/drugresistance/threat-report-2013>. [Accessed: 19-Dec-2018].
- [91] J. O'Neill, "Tackling Drug-Resistant Infections Globally : Final Report and Recommendations," *The Review on Antimicrobial Resistance*, 2016. [Online]. Available: [https://amr-review.org/sites/default/files/160525_Final paper_with cover.pdf](https://amr-review.org/sites/default/files/160525_Final%20paper_with%20cover.pdf). [Accessed: 26-Jul-2020].
- [92] J. G. Swoboda, J. Campbell, T. C. Meredith, and S. Walker, "Wall teichoic acid function, biosynthesis, and inhibition.," *Chembiochem*, vol. 11, no. 1, pp. 35–45, Jan. 2010.
- [93] M. Jankute, J. A. G. Cox, J. Harrison, and G. S. Besra, "Assembly of the Mycobacterial Cell Wall.," *Annu. Rev. Microbiol.*, vol. 69, no. 1, pp. 405–23, 2015.
- [94] J. V Höltje, "Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*.,," *Microbiol. Mol. Biol. Rev.*, vol. 62, no. 1, pp. 181–203, 1998.
- [95] A. M. Whatmore and R. H. Reed, "Determination of turgor pressure in *Bacillus subtilis*: a possible role for K⁺ in turgor regulation.,," *J. Gen. Microbiol.*, vol. 136, no. 12, pp. 2521–6, Dec. 1990.
- [96] D. S. Cayley, H. J. Guttman, and M. T. Record, "Biophysical characterization of changes in amounts and activity of *Escherichia coli* cell and compartment water and turgor pressure in response to osmotic stress.,," *Biophys. J.*, vol. 78, no. 4, pp. 1748–64, Apr. 2000.
- [97] K. H. Schleifer and O. Kandler, "Peptidoglycan types of bacterial cell walls and their taxonomic implications.,," *Bacteriol. Rev.*, vol. 36, no. 4, pp. 407–77, Dec. 1972.
- [98] D. Scheffers and M. Pinho, "Bacterial cell wall synthesis: new insights from localization studies," *Microbiol. Mol. Biol. Rev.*, vol. 69, no. 4, pp. 585–607, 2005.
- [99] A. Bouhss, A. E. Trunkfield, T. D. H. Bugg, and D. Mengin-Lecreulx, "The biosynthesis of peptidoglycan lipid-linked intermediates," *FEMS Microbiol. Rev.*, vol. 32, no. 2, pp. 208–233, 2008.
- [100] H. Barreteau, A. Kovač, A. Boniface, M. Sova, S. Gobec, and D. Blanot, "Cytoplasmic steps of peptidoglycan biosynthesis," *FEMS Microbiol. Rev.*, vol. 32, no. 2, pp. 168–207, 2008.
- [101] E. Sauvage, F. Kerff, M. Terrak, J. A. Ayala, and P. Charlier, "The penicillin-binding proteins: Structure and role in peptidoglycan biosynthesis," *FEMS Microbiol. Rev.*, vol. 32, no. 2, pp. 234–258, 2008.
- [102] T. Schneider and H. G. Sahl, "An oldie but a goodie - cell wall biosynthesis as antibiotic target pathway," *Int. J. Med. Microbiol.*, vol. 300, no. 2–3, pp. 161–169, 2010.
- [103] K. Yamamoto and S. Ichikawa, "Tunicamycin: chemical synthesis and biosynthesis.,," *J. Antibiot. (Tokyo).*, vol. 72, no. 12, pp. 924–933, 2019.
- [104] I. Nikolaidis, S. Favini-Stabile, and A. Dessen, "Resistance to antibiotics targeted to the bacterial cell wall," *Protein Sci.*, vol. 23, no. 3, pp. 243–259, 2014.
- [105] M. E. Falagas, E. K. Vouloumanou, G. Samonis, and K. Z. Vardakas, "Fosfomycin.,," *Clin. Microbiol. Rev.*, vol. 29, no. 2, pp. 321–47, Apr. 2016.
- [106] C. A. Smith, "Structure, Function and Dynamics in the mur Family of Bacterial Cell Wall Ligases," *J. Mol. Biol.*, vol. 362, no. 4, pp. 640–655, 2006.
- [107] D. Münch, T. Roemer, S. H. Lee, M. Engeser, H. G. Sahl, and T. Schneider, "Identification and in vitro Analysis of the GatD/MurT Enzyme-Complex Catalyzing Lipid II Amidation in *Staphylococcus aureus*," *PLoS Pathog.*, vol. 8, no. 1, 2012.
- [108] T. A. Figueiredo *et al.*, "Identification of genetic determinants and enzymes involved with the amidation of glutamic acid residues in the peptidoglycan of *Staphylococcus aureus*," *PLoS Pathog.*, vol. 8, no. 1, 2012.
- [109] C. Morlot *et al.*, "Structure of the essential peptidoglycan amidotransferase MurT/GatD complex from *Streptococcus pneumoniae*," *Nat. Commun.*, vol. 9, no. 1, pp. 1–12, 2018.
- [110] L. A. Draper, P. D. Cotter, C. Hill, and R. P. Ross, "Lantibiotic resistance.,," *Microbiol. Mol. Biol. Rev.*, vol. 79, no. 2, pp. 171–91, Jun. 2015.
- [111] J. M. Shin, J. W. Gwak, P. Kamarajan, J. C. Fenno, A. H. Rickard, and Y. L. Kapila, "Biomedical applications of nisin.,," *J. Appl. Microbiol.*, vol. 120, no. 6, pp. 1449–65, Jun. 2016.
- [112] D. Zeng *et al.*, "Approved Glycopeptide Antibacterial Drugs: Mechanism of Action and Resistance.,," *Cold Spring Harb. Perspect. Med.*, vol. 6, no. 12, Dec. 2016.

- [113] R. C. Moellering, "Vancomycin: a 50-year reassessment.," *Clin. Infect. Dis.*, vol. 42 Suppl 1, no. Suppl 1, pp. S3-4, Jan. 2006.
- [114] R. Álvarez, L. E. López Cortés, J. Molina, J. M. Cisneros, and J. Pachón, "Optimizing the Clinical Use of Vancomycin.," *Antimicrob. Agents Chemother.*, vol. 60, no. 5, pp. 2601–9, 2016.
- [115] K. Bush and P. A. Bradford, "β-Lactams and β-Lactamase Inhibitors: An Overview.," *Cold Spring Harb. Perspect. Med.*, vol. 6, no. 8, 2016.
- [116] J. M. Munita and C. A. Arias, "Mechanisms of Antibiotic Resistance.," *Microbiol. Spectr.*, vol. 4, no. 2, pp. 1–24, 2016.
- [117] H. G. Floss and T.-W. Yu, "Rifamycin-mode of action, resistance, and biosynthesis.," *Chem. Rev.*, vol. 105, no. 2, pp. 621–32, Feb. 2005.
- [118] D. C. Hooper, "Fluoroquinolone resistance among Gram-positive cocci.," *Lancet. Infect. Dis.*, vol. 2, no. 9, pp. 530–8, Sep. 2002.
- [119] R. E. Mendes, L. M. Deshpande, and R. N. Jones, "Linezolid update: stable in vitro activity following more than a decade of clinical use and summary of associated resistance mechanisms.," *Drug Resist. Updat.*, vol. 17, no. 1–2, pp. 1–12, Apr. 2014.
- [120] C. M. Thomas and K. M. Nielsen, "Mechanisms of, and barriers to, horizontal gene transfer between bacteria.," *Nat. Rev. Microbiol.*, vol. 3, no. 9, pp. 711–21, Sep. 2005.
- [121] M. S. Wilke, A. L. Lovering, and N. C. J. Strynadka, "Beta-lactam antibiotic resistance: a current structural perspective.," *Curr. Opin. Microbiol.*, vol. 8, no. 5, pp. 525–33, Oct. 2005.
- [122] K. V. Mahasanen *et al.*, "Conformational Dynamics in Penicillin-Binding Protein 2a of Methicillin-Resistant *Staphylococcus aureus* , Allosteric Communication Network and Enablement of Catalysis," *J. Am. Chem. Soc.*, vol. 139, no. 5, pp. 2102–2110, 2017.
- [123] C. Gorrie, C. Higgs, G. Carter, T. P. Stinear, and B. Howden, "Genomics of vancomycin-resistant *Enterococcus faecium*.," *Microb. genomics*, vol. 5, no. 7, 2019.
- [124] WHO, "Antimicrobial resistance.," *Bull. World Health Organ.*, vol. 61, no. 3, pp. 383–94, 2014.
- [125] C. L. Ventola, "The antibiotic resistance crisis: part 2: management strategies and new agents.," *P T*, vol. 40, no. 5, pp. 344–52, May 2015.
- [126] W. Vollmer, O. Massidda, and A. Tomasz, "The Cell Wall of *Streptococcus pneumoniae*.," *Microbiol. Spectr.*, vol. 7, no. 3, pp. 230–240, May 2019.
- [127] D. Gerlach *et al.*, "Methicillin-resistant *Staphylococcus aureus* alters cell wall glycosylation to evade immunity.," *Nature*, vol. 563, no. 7733, pp. 705–709, Nov. 2018.
- [128] T. a Figueiredo, A. M. Ludovice, and R. G. Sobral, "Contribution of Peptidoglycan Amidation to Beta-Lactam and Lysozyme Resistance in Different Genetic Lineages of *Staphylococcus aureus*.," *Microb. Drug Resist.*, vol. 20, no. 3, pp. 1–12, 2014.
- [129] S. A. Ragland and A. K. Criss, "From bacterial killing to immune modulation: Recent insights into the functions of lysozyme.," *PLoS Pathog.*, vol. 13, no. 9, p. e1006512, Sep. 2017.
- [130] T. Mirski, M. Lidia, A. Nakonieczna, and R. Gryko, "Bacteriophages, phage endolysins and antimicrobial peptides - the possibilities for their common use to combat infections and in the design of new drugs.," *Ann. Agric. Environ. Med.*, vol. 26, no. 2, pp. 203–209, Jun. 2019.
- [131] P. J. Moynihan, D. Sychantha, and A. J. Clarke, "Chemical biology of peptidoglycan acetylation and deacetylation," *Bioorg. Chem.*, vol. 54, pp. 44–50, 2014.
- [132] M. Rajagopal and S. Walker, "Envelope Structures of Gram-Positive Bacteria.," *Curr. Top. Microbiol. Immunol.*, vol. 404, no. 3, pp. 1–44, 2017.
- [133] G. Siewert and J. L. Strominger, "Biosynthesis of the peptidoglycan of bacterial cell walls. XI. Formation of the isoglutamine amide group in the cell walls of *Staphylococcus aureus*.," *J. Biol. Chem.*, vol. 243, no. 4, pp. 783–90, Feb. 1968.
- [134] P. Veiga, M. Erkelenz, E. Bernard, P. Courtin, S. Kulakauskas, and M.-P. Chapot-Chartier, "Identification of the asparagine synthase responsible for D-Asp amidation in the *Lactococcus lactis* peptidoglycan interpeptide crossbridge.," *J. Bacteriol.*, vol. 191, no. 11, pp. 3752–7, Jun. 2009.
- [135] M. Levefaudes *et al.*, "Diaminopimelic Acid Amidation in Corynebacteriales: NEW INSIGHTS INTO THE ROLE OF LtsA IN PEPTIDOGLYCAN MODIFICATION.," *J. Biol. Chem.*, vol. 290, no. 21, pp. 13079–94, May 2015.

- [136] E. Bernard, T. Rolain, P. Courtin, P. Hols, and M.-P. Chapot-Chartier, "Identification of the amidotransferase AsnB1 as being responsible for meso-diaminopimelic acid amidation in *Lactobacillus plantarum* peptidoglycan.," *J. Bacteriol.*, vol. 193, no. 22, pp. 6323–30, Nov. 2011.
- [137] A. Zapun *et al.*, "In vitro reconstitution of peptidoglycan assembly from the gram-positive pathogen *Streptococcus pneumoniae*," *ACS Chem. Biol.*, vol. 8, no. 12, pp. 2688–2696, 2013.
- [138] M. A. DeJesus *et al.*, "Comprehensive Essentiality Analysis of the *Mycobacterium tuberculosis* Genome via Saturating Transposon Mutagenesis.," *MBio*, vol. 8, no. 1, pp. 1–17, 2017.
- [139] E. R. Nöldeke and T. Stehle, "Unraveling the mechanism of peptidoglycan amidation by the bifunctional enzyme complex GatD/MurT: A comparative structural approach.," *Int. J. Med. Microbiol.*, vol. 309, no. 6, p. 151334, Sep. 2019.
- [140] D. Münch *et al.*, "Structural variations of the cell wall precursor lipid II and their influence on binding and activity of the lipoglycopeptide antibiotic oritavancin," *Antimicrob. Agents Chemother.*, vol. 59, no. 2, pp. 772–781, 2015.
- [141] F. Leisico *et al.*, "First insights of peptidoglycan amidation in Gram-positive bacteria-The high-resolution crystal structure of *Staphylococcus aureus* glutamine amidotransferase GatD," *Sci. Rep.*, vol. 8, no. 1, pp. 1–13, 2018.
- [142] K. M. Ruane *et al.*, "Specificity determinants for lysine incorporation in *Staphylococcus aureus* peptidoglycan as revealed by the structure of a MurE enzyme ternary complex," *J. Biol. Chem.*, vol. 288, no. 46, pp. 33439–33448, 2013.
- [143] S. S. Cha, Y. J. An, C. S. Jeong, J. H. Yu, and K. M. Chung, "ATP-binding mode including a carbamoylated lysine and two Mg²⁺ ions, and substrate-binding mode in *Acinetobacter baumannii* MurF," *Biochem. Biophys. Res. Commun.*, vol. 450, no. 2, pp. 1045–1050, 2014.
- [144] R. Gamsjaeger, C. K. Liew, F. E. Loughlin, M. Crossley, and J. P. Mackay, "Sticky fingers: zinc-fingers as protein-recognition motifs," *Trends Biochem. Sci.*, vol. 32, no. 2, pp. 63–70, 2007.
- [145] G. Malgieri, M. Palmieri, L. Russo, R. Fattorusso, P. V. Pedone, and C. Isernia, "The prokaryotic zinc-finger: Structure, function and comparison with the eukaryotic counterpart," *FEBS J.*, vol. 282, no. 23, pp. 4480–4496, 2015.
- [146] K. L. Howe *et al.*, "Ensembl Genomes 2020-enabling non-vertebrate genomic research.," *Nucleic Acids Res.*, vol. 48, no. D1, pp. D689–D695, 2020.
- [147] Agilent Technologies, "QuikChange II Site-Directed Mutagenesis Kit User Manual." .
- [148] W. W. Wilfinger, K. Mackey, and P. Chomczynski, "Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity.," *Biotechniques*, vol. 22, no. 3, pp. 474–6, 478–81, Mar. 1997.
- [149] E. Gasteiger *et al.*, "Protein Analysis Tools on the ExpASY Server," in *The Proteomics Protocols Handbook Protein Identification and Analysis Tools on the ExpASY Server*, 2005, pp. 571–607.
- [150] V. C. Ciccarone, D. A. Polayes, and V. A. Luckow, "Generation of Recombinant Baculovirus DNA in *E. coli* Using a Baculovirus Shuttle Vector.," *Methods Mol. Med.*, vol. 13, pp. 213–35, 1998.
- [151] D. Anderson *et al.*, "Rapid Generation of Recombinant Baculovirus and Expression of Foreign Genes using the Bac-To-Bac™ Baculovirus Expression System," *Focus (Madison)*, vol. 17, no. January 1995, pp. 53–58, 1996.
- [152] U. Rüter, "Construction and properties of a new cloning vehicle, allowing direct screening for recombinant plasmids.," *Mol. Gen. Genet.*, vol. 178, no. 2, pp. 475–7, 1980.
- [153] Invitrogen, "Bac-to-Bac® Baculovirus Expression System." .
- [154] F. W. Studier, "Protein production by auto-induction in high density shaking cultures.," *Protein Expr. Purif.*, vol. 41, no. 1, pp. 207–34, May 2005.
- [155] M. V. Petoukhov *et al.*, "New developments in the ATSAS program package for small-angle scattering data analysis," *J. Appl. Crystallogr.*, vol. 45, no. 2, pp. 342–350, 2012.
- [156] P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch, and D. I. Svergun, "PRIMUS: a Windows PC-based system for small-angle scattering data analysis," *J. Appl. Crystallogr.*, vol. 36, no. 5, pp. 1277–1282, Oct. 2003.

5 APPENDIX

5.1 Superdex 200 Increase 3.2/300 Calibration

$$K_{av} = \frac{V_e - V_0}{V_c - V_0} \quad (5-1)$$

Table 5-1: Size standards used for calibration of analytical SEC column (Superdex 200 Increase 3.2/300). The elution coefficient K_{av} is given by equation (5-1), where V_e is the elution volume of the analyte, V_c the total column volume, and V_0 the column's void volume. Calibration was performed by Aleandra Thor.

	V_e	M_w	K_{av}	$\log(M_w)$
Ribonuclease A	1.8 mL	13.7 kDa	0.59	1.14
Carbonic anhydrase	1.69 mL	29 kDa	0.52	1.46
Ovalbumin	1.57 mL	44 kDa	0.44	1.64
Conalbumin	1.49 mL	75 kDa	0.38	1.88
Aldolase	1.36 mL	158 kDa	0.30	2.20
Ferritin	1.15 mL	440 kDa	0.15	2.64
Thyroglobulin	1.01 mL	669 kDa	0.06	2.83
Blue Dextran 2000	0.92 mL = V_0			

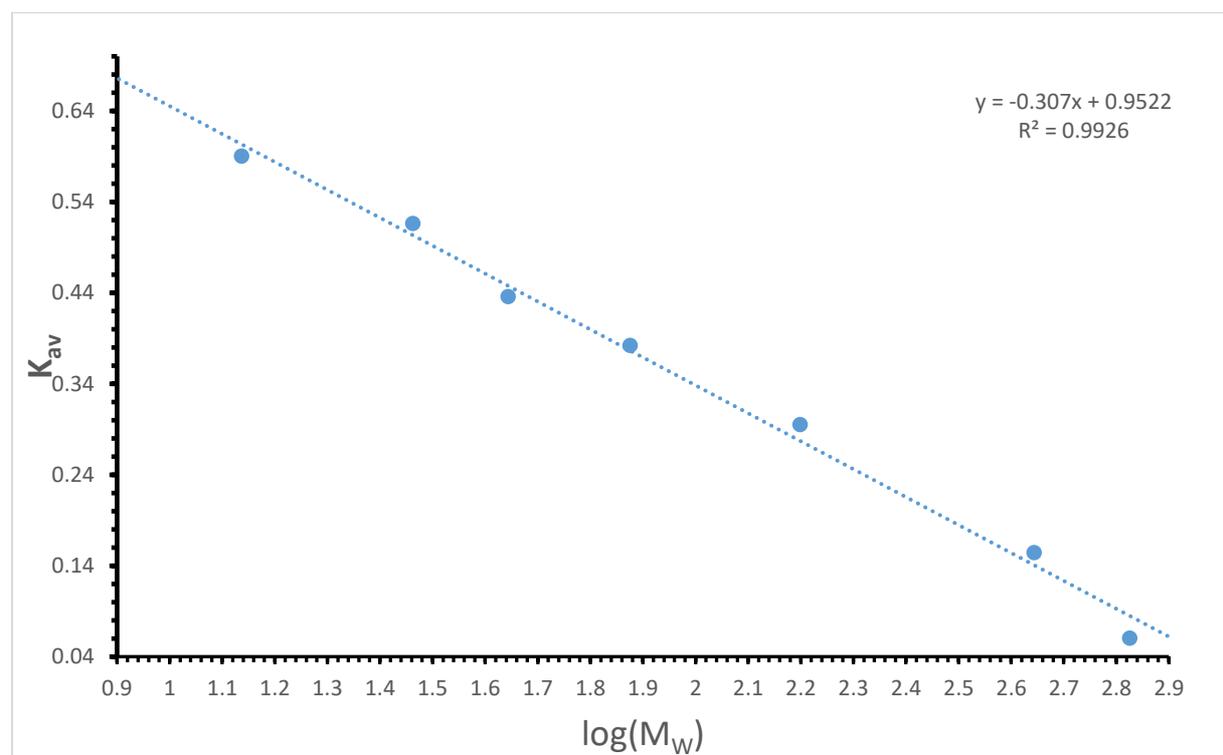


Figure 5-1: Regression analysis of calibration of Superdex 200 Increase 3.2/300. The elution coefficient K_{av} is plotted against the logarithm of the molecular weight of the analyte. Linear regression yields the dependency shown in the top right corner. Within the calibrated area it can be used to predict elution coefficients (and thus elution volumes using equation (5-1)) of analytes of known molecular weight, or to calculate the molecular weight of a spherical object, based on a measured elution volume. Calibration was performed by Alexandra Thor.

5.2 Original gel pictures.

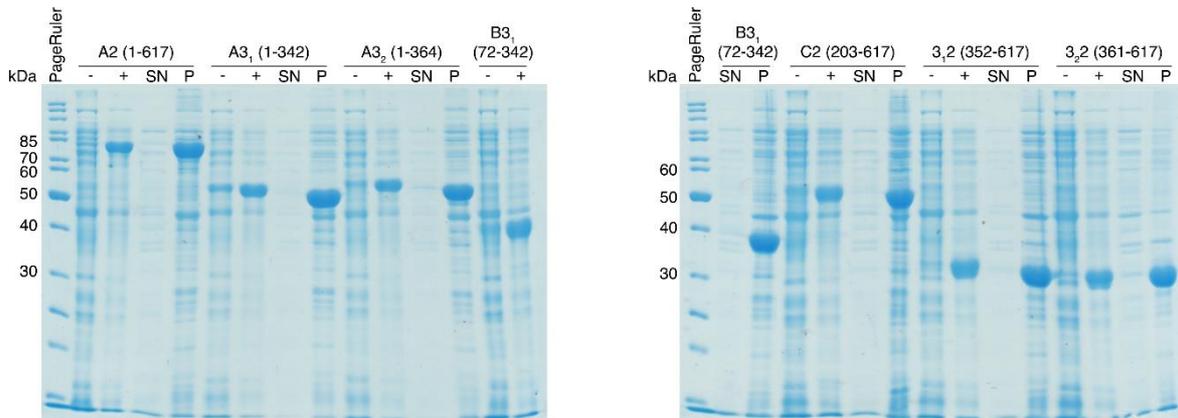


Figure 5-2: Original gel images for Figure 1-3 without labelling asterisks.

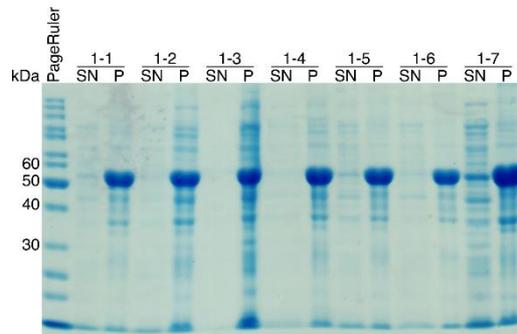


Figure 5-3: Original gel image for Figure 1-4a without labelling asterisks.

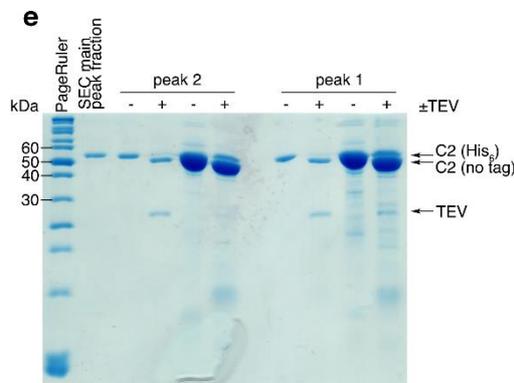


Figure 5-4: Original gel image from Figure 1-7e without cropping.

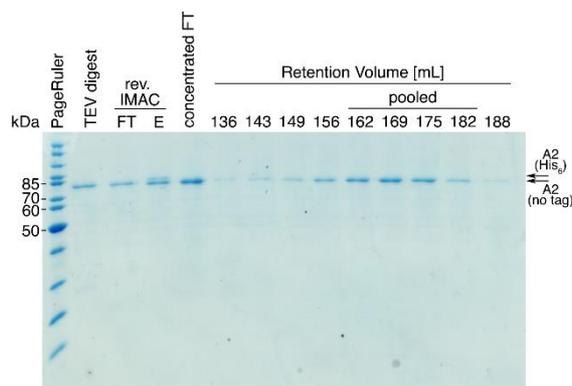


Figure 5-5: Original gel image from Figure 1-9f, right panel, without cropping.

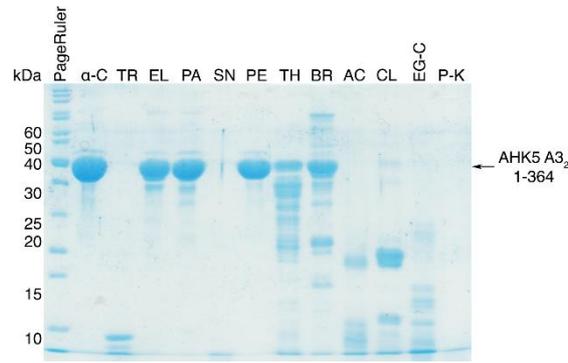


Figure 5-6: Original gel image for Figure 1-11a without labelling asterisks.

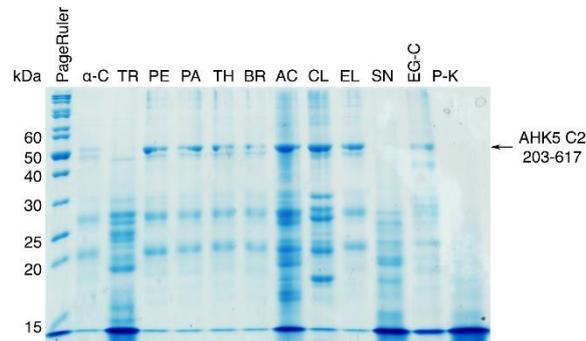


Figure 5-7: Original gel image for Figure 1-12a without labelling asterisks.

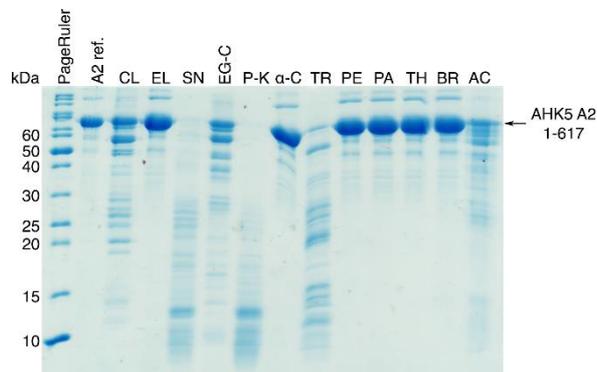


Figure 5-8: Original gel image for Figure 1-13a without labelling asterisks.

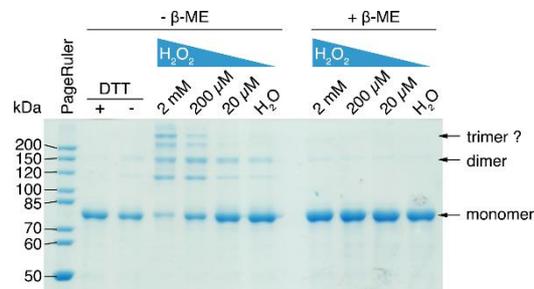


Figure 5-9: Original gel image for Figure 1-19 without labelling asterisks.

5.3 HHPRED search results for AHK5

Table 5-2: Top 50 HHPRED hits for full-length AHK5. The data for this table was automatically generated by HHPRED on 15/Oct/2019 and reflects that version of databases. **Hit:** PDB ID and name of identified structural homologue. **Prob, E-value, P-value:** statistical probability of hit relevance. **Score:** HHPRED score. **SS:** Secondary structure alignment score. **Cols:** No. of residues aligned. **Query-HMM:** Portion of query sequence in alignment. **Templat:** portion of target in alignment. **e HMM:** emission parameter of hidden model.

No	Hit	Prob	E-value	P-value	Score	SS	Cols	Query HMM	Templat	e HMM
1	4GCZ_A Blue-light photorecepto	100	1.50E-31	3.10E-36	292.9	43.1	357	213-618	8-377	-385
2	4I5S_A Putative histidine kina	100	3.10E-32	6.40E-37	307.4	33.6	376	188-629	66-450	-450
3	3A0R_A Sensor protein (E.C.2.7	100	9.00E-31	1.80E-35	286.6	42.1	342	213-617	2-349	-349
4	4R3A_A Blue-light-activated hi	99.9	1.10E-27	2.30E-32	262.9	38.4	329	216-619	10-345	-352
5	4E01_A [3-methyl-2-oxobutanoat	99.9	3.30E-27	6.60E-32	266.9	35.5	359	214-623	22-413	-418
6	1B3Q_A CHEMOTAXIS PROTEIN CHEA	99.9	5.00E-26	1.00E-30	253.9	18.7	287	375-833	1-341	-379
7	6DK7_D RetS (Regulator of Exop	99.9	4.00E-24	8.20E-29	214.5	27.7	225	369-615	2-227	-227
8	2C2A_A SENSOR HISTIDINE KINASE	99.9	1.70E-23	3.50E-28	215.1	31.6	250	352-623	5-258	-258
9	4Q20_A Sensor protein DivL (E.	99.9	1.40E-22	2.80E-27	211.4	33	246	347-624	19-267	-268
10	5IDJ_A Cell cycle histidine ki	99.9	3.40E-23	6.90E-28	211.3	27.2	227	358-615	1-242	-242
11	4U7O_B Histidine protein kinas	99.9	1.20E-22	2.30E-27	213.9	29.2	248	352-625	20-275	-277
12	3D36_B Sensor protein (E.C.2.7	99.9	2.80E-22	5.60E-27	205.7	28.2	235	355-626	2-238	-244
13	6NB0_A Histidine kinase; SSGCI	99.8	5.70E-22	1.10E-26	210.5	28	244	349-615	40-287	-287
14	4KP4_B Osmolarity sensor prote	99.8	9.60E-22	2.00E-26	199.4	27.7	230	359-624	4-236	-236
15	3JZ3_B Sensor protein qseC (E.	99.8	1.10E-21	2.30E-26	198.3	25.9	219	366-617	1-221	-222
16	4BIX_A SENSOR PROTEIN CPXA (E.	99.8	5.10E-21	1.00E-25	204.7	30.7	242	334-616	56-298	-298
17	4XIV_B Chemotaxis protein CheA	99.8	1.90E-20	3.90E-25	195.7	23.6	211	372-613	2-252	-252
18	4CTI_C OSMOLARITY SENSOR PROTE	99.8	1.10E-19	2.10E-24	191.4	29.1	241	334-624	33-273	-273
19	2E0A_B Pyruvate dehydrogenase	99.8	1.10E-19	2.30E-24	205.3	30.2	288	327-639	48-376	-394
20	4JPB_A Chemotaxis protein CheA	99.8	7.90E-21	1.60E-25	209.5	19.4	236	436-835	1-271	-319

No	Hit	Prob	E-value	P-value	Score	SS	Cols	Query HMM	Templat	e HMM
21	6E52_A PROTEIN; Protein histid	99.8	7.60E-19	1.50E-23	179.7	29.6	240	327-613	11-250	-250
22	4R39_A Blue-light-activated hi	99.8	2.10E-19	4.20E-24	182.7	25	228	352-626	4-232	-232
23	5J71_A [Pyruvate dehydrogenase	99.7	1.10E-18	2.20E-23	197.9	32.1	300	334-669	75-400	-400
24	4QPK_A Phosphotransferase; Chp	99.7	2.00E-19	4.10E-24	188.6	23.7	224	340-609	20-243	-243
25	3EHH_A Sensor kinase (YocF pro	99.7	8.10E-18	1.60E-22	168.2	28.2	215	341-615	2-218	-218
26	2AYX_A Sensor kinase protein r	99.7	1.60E-19	3.30E-24	190.8	16.6	213	536-922	34-246	-254
27	2Q8G_A [Pyruvate dehydrogenase	99.7	2.90E-18	5.80E-23	195.5	28	265	334-623	79-373	-407
28	4FPP_C phosphotransferase; fou	99.7	3.00E-18	6.10E-23	178.3	22.8	224	340-612	19-247	-247
29	5EPV_A Blue-light-activated hi	99.7	7.60E-18	1.50E-22	176.3	24.9	216	321-622	1-219	-233
30	1Y8O_A [Pyruvate dehydrogenase	99.7	1.40E-17	2.80E-22	191.4	29.1	257	359-644	117-402	-419
31	4EUK_A Histidine kinase 5 (E.C	99.6	1.10E-16	2.30E-21	155.3	17.9	153	770-922	1-153	-153
32	4PL9_A ETR1-CA; ETR1, Histidin	99.5	1.90E-15	3.90E-20	148.5	20	181	430-620	1-183	-183
33	3SL2_A Sensor histidine kinase	99.5	1.70E-15	3.50E-20	150.9	18.1	171	440-634	3-173	-177
34	1ID0_A PHOQ HISTIDINE KINASE (99.5	4.10E-15	8.30E-20	141.7	19.9	151	437-620	1-152	-152
35	5IDM_A Cell cycle histidine ki	99.5	2.40E-15	4.90E-20	149.6	18.7	156	435-617	1-170	-179
36	1R62_A Nitrogen regulation pro	99.5	3.10E-15	6.20E-20	143.1	18	152	433-614	1-160	-160
37	1YSR_B Sensor-type histidine k	99.4	1.20E-14	2.40E-19	139.1	18.7	149	434-614	2-150	-150
38	5LNN_A Histidine kinase CK11 (99.4	4.60E-15	9.20E-20	152.7	16.7	165	586-921	15-194	-206
39	1BXD_A PROTEIN (ENVZ); HISTIDI	99.4	7.50E-15	1.50E-19	141.5	17.2	158	434-624	4-161	-161
40	1I58_B CHEMOTAXIS PROTEIN CHEA	99.4	1.10E-14	2.30E-19	148.2	19.1	155	436-613	1-189	-189
41	6BLK_B Phage-associated cell w	99.4	2.10E-14	4.30E-19	137.7	18.8	151	434-615	3-158	-158
42	3A0Y_A Sensor protein (E.C.2.7	99.4	1.20E-13	2.50E-18	130.9	18.5	151	435-617	1-152	-152
43	4QIC_C Sensory transduction re	99.3	3.00E-14	6.00E-19	152.4	13.7	125	767-922	129-255	-276
44	1DBW_B TRANSCRIPTIONAL REGULAT	99.3	9.40E-14	1.90E-18	124.9	14.2	119	776-922	2-120	-126
45	4MYR_B CpaE2 pilus assembly pr	99.3	6.60E-14	1.30E-18	133.5	13.9	122	773-922	16-143	-147

No	Hit		Prob	E-value	P-value	Score	SS	Cols	Query HMM	Templat	e HMM
46	4LDA_E	TadZ; Response regulato	99.3	1.00E-13	2.00E-18	126.2	14	119	776-922	2-122	-131
47	4BXI_A	ACCESSORY GENE REGULATO	99.3	2.80E-13	5.60E-18	128.6	17.4	152	431-613	1-153	-153
48	4JA2_A	Response regulator; alp	99.3	1.80E-13	3.70E-18	122	15	120	777-922	2-121	-122
49	2JBA_B	PHOSPHATE REGULON TRANS	99.3	1.50E-13	3.10E-18	123.7	14.4	120	777-922	2-121	-127
50	3T6K_A	Response regulator rece	99.3	1.60E-13	3.20E-18	126.1	14.6	123	774-922	1-123	-136

Table 5-3: Top 50 HHPRED hits for residues 1 – 400 of AHK5. The data for this table was automatically generated by HHPRED on 15/Oct/2019 and reflects that version of databases. **Hit:** PDB ID and name of identified structural homologue. **Prob, E-value, P-value:** statistical probability of hit relevance. **Score:** HHPRED score. **SS:** Secondary structure alignment score. **Cols:** No. of residues aligned. **Query-HMM:** Portion of query sequence in alignment. **Templat:** portion of target in alignment. **e HMM:** emission parameter of hidden model.

No	Hit		Prob	E-value	P-value	Score	SS	Cols	Query HMM	Templat	e HMM
1	4GCZ_A	Blue-light photorecepto	99.3	3.00E-13	6.20E-18	130	20.7	165	214-394	9-179	-385
2	4HIA_A	LOV protein; LOV, PAS,	99.3	2.20E-12	4.40E-17	107.1	18.9	164	214-392	2-171	-176
3	4R3A_A	Blue-light-activated hi	99.1	1.20E-11	2.30E-16	118.2	17.9	155	216-398	10-170	-352
4	4I5S_A	Putative histidine kina	99.1	1.40E-12	2.80E-17	128.7	10.7	172	192-400	67-241	-450
5	3A0R_A	Sensor protein (E.C.2.7	99	5.30E-11	1.10E-15	113.5	18.3	159	213-399	2-164	-349
6	3P7N_B	Sensor histidine kinase	98.8	6.70E-10	1.40E-14	100	15.6	185	211-400	55-245	-258
7	5HZI_B	Intersectin-1,NPH1-2,In	98.8	5.10E-10	1.00E-14	112.1	15.7	213	159-392	11-236	-502
8	4HHD_A	Phototropin-1 (E.C.2.7.	98.8	2.00E-08	4.00E-13	81.5	20.9	159	204-380	1-165	-166
9	2MWG_B	Blue-light photorecepto	98.7	9.50E-10	1.90E-14	98.3	12.5	157	211-400	6-168	-261
10	3RTY_B	Period circadian protei	98.5	1.30E-08	2.70E-13	93.7	13.6	161	221-390	146-315	-339
11	1V9Y_B	Heme pas sensor protein	98.5	4.00E-08	8.10E-13	81.1	15.2	142	197-340	18-166	-167
12	4DJ3_B	Period circadian protei	98.4	2.50E-08	5.10E-13	91.3	13.6	138	224-361	169-312	-317
13	3KX0_X	Uncharacterized protein	98.4	2.50E-07	5.10E-12	77.4	17.9	153	206-376	24-185	-185
14	3GEC_A	Period circadian protei	98.4	6.00E-08	1.20E-12	89	14.4	149	221-382	155-312	-312
15	3UE6_E	Aureochrome1; PAS/LOV d	98.3	2.20E-07	4.40E-12	75.4	15.2	136	204-339	20-161	-166
16	2V0U_A	NPH1-1; LOV2, KINASE, T	98.3	1.10E-06	2.20E-11	68.6	18.7	138	219-356	2-145	-146
17	4E01_A	[3-methyl-2-oxobutanoat	98.3	3.80E-08	7.70E-13	96.5	11.7	152	220-400	28-184	-418
18	5A8B_A	PTAUREO1A LOV2 DOMAIN;	98.2	1.10E-06	2.10E-11	71.7	17	133	217-349	21-159	-162
19	6IDE_B	Transcriptional regulat	98.2	7.20E-08	1.50E-12	86.2	10.3	181	214-394	3-197	-256
20	5J7E_D	Potassium voltage-gated	98.2	4.70E-07	9.40E-12	70.6	13	136	204-339	1-145	-146
21	3VOL_A	Aerotaxis transducer Ae	98.2	2.00E-06	4.00E-11	75.4	18.3	174	214-399	17-205	-233
22	4DJ2_C	Period circadian protei	98.1	2.50E-07	5.20E-12	84.7	12.1	124	224-347	171-314	-320
23	3K3C_C	protein Rv1364c/MT1410;	98.1	2.80E-06	5.60E-11	67.8	16.4	144	207-350	5-155	-158

No	Hit		Prob	E-value	P-value	Score	SS	Cols	Query HMM	Templat	e HMM
24	6GG9_B	Sensory box protein; bl	98.1	2.00E-06	4.10E-11	69.9	14.9	136	214-349	20-161	-162
25	3GDI_A	Period circadian protei	98	3.30E-07	6.70E-12	83.4	10.6	124	224-347	160-303	-309
26	6I20_A	Aureochrome1-like prote	98	4.60E-06	9.20E-11	63.4	15.5	127	216-342	2-134	-134
27	5XGB_A	Uncharacterized protein	98	1.80E-06	3.70E-11	87.7	17	154	207-394	1-162	-568
28	4HH2_A	Regulatory protein PpsR	98	2.20E-06	4.50E-11	81.3	15.6	140	218-361	17-159	-384
29	4LRZ_F	PTS-dependent dihydroxy	98	3.70E-06	7.50E-11	77.9	16.4	138	192-337	176-316	-318
30	4L9E_A	Transcriptional regulat	97.9	4.50E-06	9.20E-11	65.5	14.4	133	191-326	3-139	-139
31	3MR0_B	Sensory box histidine k	97.9	4.30E-06	8.80E-11	64.4	14.1	136	211-346	1-141	-142
32	4HH3_A	Regulatory protein PpsR	97.9	7.80E-06	1.60E-10	72.5	17.5	140	218-361	17-159	-262
33	4ZP4_B	Aryl hydrocarbon recept	97.9	5.30E-07	1.10E-11	84.4	9.4	117	227-343	239-359	-360
34	2Z6D_A	Phototropin-2 (E.C.2.7.	97.9	6.90E-06	1.40E-10	61.7	13.4	124	214-337	1-130	-130
35	4HH3_A	Regulatory protein PpsR	97.8	8.10E-06	1.60E-10	72.4	15.4	133	190-326	125-261	-262
36	2PR5_B	Blue-light photorecepto	97.8	1.50E-05	3.10E-10	60.4	14.9	124	221-346	1-130	-132
37	3EWK_A	Sensor protein (E.C.2.7	97.8	8.20E-06	1.70E-10	70.1	14.7	119	204-326	100-227	-227
38	4ZPR_B	Aryl hydrocarbon recept	97.8	2.00E-06	4.00E-11	80.3	10.7	112	227-338	226-341	-345
39	2KDK_A	Aryl hydrocarbon recept	97.8	4.40E-06	9.00E-11	62.3	10.5	115	217-331	1-120	-121
40	2GJ3_B	Nitrogen fixation regul	97.7	1.80E-05	3.60E-10	58.4	13.6	117	215-331	1-120	-120
41	4R38_D	Blue-light-activated hi	97.7	1.10E-05	2.20E-10	62.1	12.5	122	217-338	11-138	-140
42	6HMJ_B	Putative PAS/PAC sensor	97.7	9.00E-06	1.80E-10	76.2	14.2	128	214-341	231-364	-373
43	5Y7Y_B	Aryl hydrocarbon recept	97.7	5.50E-06	1.10E-10	76.2	12.4	104	224-327	203-311	-311
44	5HWV_B	Sensor histidine kinase	97.7	2.70E-05	5.40E-10	59.1	14.2	120	226-345	3-130	-130
45	3MXQ_D	Sensor protein (E.C.2.7	97.7	1.20E-05	2.50E-10	64.2	12.7	127	208-335	11-152	-152
46	3EWK_A	Sensor protein (E.C.2.7	97.7	2.50E-05	5.10E-10	67.1	14.9	116	231-346	1-119	-227
47	2WKQ_A	NPH1-1, RAS-RELATED C3	97.7	2.40E-05	4.90E-10	73.6	15.9	136	221-356	13-154	-332
48	4F3L_A	BMAL1b, Circadian locom	97.7	6.20E-06	1.20E-10	77.4	11.7	112	224-335	246-361	-361

No	Hit		Prob	E-value	P-value	Score	SS	Cols	Query HMM	Templat	e HMM
49	3LYX_B	Sensory box/GGDEF domai	97.6	4.20E-05	8.60E-10	56.7	14.1	120	214-333	2-124	-124
50	3ICY_B	Sensor protein; Sensory	97.6	1.90E-05	3.80E-10	58.8	12.1	113	216-328	1-118	-118

5.4 Publications

E. R. Nöldeke, L. M. Muckenfuss, V. Niemann, A. Müller, E. Störk, G. Zocher, T. Schneider, and T. Stehle, “Structural basis of cell wall peptidoglycan amidation by the GatD/MurT complex of *Staphylococcus aureus*,” *Sci. Rep.*, vol. 8, no. May, p. 12953, **2018**. doi: [10.1038/s41598-018-31098-x](https://doi.org/10.1038/s41598-018-31098-x)

E. R. Nöldeke and T. Stehle, “Unraveling the mechanism of peptidoglycan amidation by the bifunctional enzyme complex GatD/MurT: A comparative structural approach.,” *Int. J. Med. Microbiol.*, vol. 309, no. 6, p. 151334, Sep. **2019**. doi: [10.1016/j.ijmm.2019.151334](https://doi.org/10.1016/j.ijmm.2019.151334)

5.5 Contributions

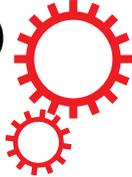
5.5.1 Contributions to Nöldeke et al. (2018)

I expressed, purified, and crystallised of *S. aureus* GatD/MurT. I generated derivate crystals containing the ATP analogue AMPPNP. I performed data collection, processing, phase determination, and model building and phase refinement for the AMPPNP-containing structure (PDB ID 6H5E). I refined and rebuilt the previously solved unliganded structure (PDB ID 6GS2). I planned and performed SAXS experiments, including beam time application and logistics. I reviewed all available data and generated the hypothetical closed conformation model. I wrote the main portion of the manuscript (results, discussion, and methods sections). I prepared all figures, including supplementary figures, with the exception of Figure S1.

5.5.2 Contributions to Nöldeke and Stehle (2019)

I performed the literature review and bioinformatic sequence-based searches. I reviewed all available data, generated all figures, and wrote the manuscript.

SCIENTIFIC REPORTS



OPEN

Structural basis of cell wall peptidoglycan amidation by the GatD/MurT complex of *Staphylococcus aureus*

Erik R. Nöldeke¹, Lena M. Muckenfuss^{1,4}, Volker Niemann^{1,5}, Anna Müller², Elena Störk¹, Georg Zocher¹, Tanja Schneider² & Thilo Stehle^{1,3}

The peptidoglycan of *Staphylococcus aureus* is highly amidated. Amidation of α -D-isoglutamic acid in position 2 of the stem peptide plays a decisive role in the polymerization of cell wall building blocks. *S. aureus* mutants with a reduced degree of amidation are less viable and show increased susceptibility to methicillin, indicating that targeting the amidation reaction could be a useful strategy to combat this pathogen. The enzyme complex that catalyzes the formation of α -D-isoglutamine in the Lipid II stem peptide was identified recently and shown to consist of two subunits, the glutamine amidotransferase-like protein GatD and the Mur ligase homolog MurT. We have solved the crystal structure of the GatD/MurT complex at high resolution, revealing an open, boomerang-shaped conformation in which GatD is docked onto one end of MurT. Putative active site residues cluster at the interface between GatD and MurT and are contributed by both proteins, thus explaining the requirement for the assembled complex to carry out the reaction. Site-directed mutagenesis experiments confirm the validity of the observed interactions. Small-angle X-ray scattering data show that the complex has a similar conformation in solution, although some movement at domain interfaces can occur, allowing the two proteins to approach each other during catalysis. Several other Gram-positive pathogens, including *Streptococcus pneumoniae*, *Clostridium perfringens* and *Mycobacterium tuberculosis* have homologous enzyme complexes. Combined with established biochemical assays, the structure of the GatD/MurT complex provides a solid basis for inhibitor screening in *S. aureus* and other pathogens.

Staphylococcus aureus is a frequent constituent of human nasal microflora and a major cause of severe endogenous infections¹. Effective treatment of staphylococcal infections remains a worldwide challenge. In the United States alone, Staphylococci are responsible for about 19,000 deaths per year, a number that is higher than that associated with HIV². Methicillin-resistant *Staphylococcus aureus* (MRSA) strains, which are resistant to many commonly used antibiotics including methicillin, amoxicillin, penicillin, and oxacillin, represent an increasing challenge to human health worldwide³.

Species-specific cell wall modifications impact on several key aspects of the infection process, including adherence^{1,4}, immune recognition⁵, and resistance to host defenses^{6,7}. In Gram-positive bacteria such as *S. aureus*, a thick multilayered peptidoglycan (PG) layer constitutes the major component of the cell wall. The PG is essential for survival and maintenance of cell shape and is crucial to resist osmotic pressure⁸. The PG heteropolymer consists of alternating disaccharide units composed of N-acetyl-glucosamine (GlcNAc) and N-acetyl-muramic acid (MurNAc), which are cross-linked by short peptides to generate a rigid network.

Assembly of PG is a multistep process that begins in the cytoplasm and terminates on the exterior of the cell (Supplementary Fig. S1). The process is initiated by the MurA-F ligases, which catalyze the formation of the soluble PG precursor UDP-MurNAc-pentapeptide in the cytoplasm⁹. The membrane-bound enzyme MraY then links this precursor to the membrane carrier undecaprenyl phosphate to yield Lipid I

¹Interfaculty Institute of Biochemistry, University of Tübingen, D-72076, Tübingen, Germany. ²Institute for Pharmaceutical Microbiology, University of Bonn, D-53115, Bonn, Germany. ³Vanderbilt University School of Medicine, Nashville, Tennessee, 37232, USA. ⁴Present address: Department of Biochemistry, University of Zurich, CH-8057, Zurich, Switzerland. ⁵Present address: Hain Lifescience GmbH, D-72147, Nehren, Germany. Correspondence and requests for materials should be addressed to T. Stehle (email: thilo.stehle@uni-tuebingen.de)

Dataset (PDB ID)	native (6GS2)	thiomersal	AMPPNP (6H5E)
Data collection			
X-ray source	SLS X06DA (PXIII)	SLS X10SA (PXII)	SLS X06DA (PXIII)
X-ray detector	Pilatus 2M	Pilatus 6M	Pilatus 2M
Wavelength [Å]	1.0	1.0	1.0
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Unit cell axes (Å)	a = 107.10 b = 110.37 c = 116.36	a = 106.97 b = 109.27 c = 116.03	a = 109.72 b = 109.74 c = 123.30
Unit cell angles (°)	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$
Resolution [Å]	50–2.04	50–2.49	50–2.14
Reflections (unique)	641308 (88009)	1411718 (91196)	1114473 (82581)
Redundancy	7.3	13.8	13.5
Completeness [%] (last bin)	99.7 (99.1)	99.7 (98.1)	99.8 (98.6)
I/ σ (I)	14.01 (1.57)	17.33 (1.88)	19.37 (1.41)
R _{meas} [%]	11.2 (178.9)	3.5 (131.3)	11.6 (213.0)
CC _{1/2} [%]	99.9 (72.5)	99.9 (71.2)	100.0 (67.6)
Wilson B [Å ²]	40.9	56.0	51.9
SIRAS phasing			
Resolution [Å]		50–2.08	
No. of heavy atoms in ASU		13	
Phasing power (iso/ano)			
centric		0.898/–	
acentric		0.685/0.382	
FOM			
centric		0.15862	
acentric		0.14178	
R _{culis} (iso/ano)			
centric		0.856/–	
acentric		0.886/0.965	
Refinement			
Resolution included [Å]	49.07–2.04		49.08–2.14
Software	PHENIX (1.10.1)		PHENIX (1.10.1)
Non-solvent atoms	9893		9909
Solvent atoms	639		442
R _{work} /R _{free} [%]	17.5/21.6		19.2/23.5
Size of R _{free} test set [%]	1.71 (1503 reflections)		1.82 (1504 reflections)
Bond r.m.s.d.	0.011		0.011
Angle r.m.s.d.	1.097		1.424
Ramachandran [%] (favoured, allowed, outliers*)	97.43, 2.57, 0.00		96.44, 3.48, 0.08
Rotamer outliers [%]	1.97		1.79
All-atom clashscore ⁺	3.34		5.02
Average B factors [Å ²]			
protein	41.0		60.1
ligand	–		50.4
ions	42.8		59.3
solvent	45.6		63.8

Table 1. Crystallographic data and refinement statistics. *Outliers are residues C94 and G190 in each of the two copies of GatD. Both have well-defined density. ⁺PDB validation reports show that these values are comparable to or better than those of structures with similar resolution.

(undecaprenyl-phosphate-MurNAc-pentapeptide), which is then connected with UDP-GlcNAc to form Lipid II (undecaprenyl-phosphate-MurNAc-pentapeptide-GlcNAc) by the glycosyltransferase MurG¹⁰. In *S. aureus*, Lipid II is modified by a Gly₅-interpeptide bridge attached by the FemXAB peptidyltransferases^{11,12}, followed by translocation across the cytoplasmic membrane likely facilitated by the flippases FtsW, RodA and MurJ¹³. Once it has reached the exterior cell surface, the modified Lipid II is assembled into the growing PG network by penicillin-binding proteins (PBPs) catalyzing transglycosylation and transpeptidation reactions^{14,15}. In particular, transpeptidation has been proposed to require at least one amidated stem peptide^{16,17}.

Amidation of the α -carboxyl group of the D-isoglutamate residue in Lipid II, resulting in the formation of D-isoglutamine^{16,18}, is catalyzed by a recently identified enzyme complex^{19,20}. This complex consists of two

as sticks. **(d)** Tilted view of the MurT middle domain to show the central β -sheet and the bound AMPPNP and its surrounding secondary structure elements, as well as the zinc ribbon. **(e)** Topological representation of the GatD/MurT architecture. Secondary structure nomenclature of GatD was done according to Leisico *et al.*²⁴. As the short helices $\alpha 1$ and $\alpha 5$ in the isolated GatD structure do not conform to helical geometry in our complex, they were not assigned. The MurT domains were assigned separately with the prefixes *m* and *c* indicating the middle and C-terminal domains, respectively. The drawing was generated with TopDraw⁵⁴.

proteins, GatD and MurT, which assemble into a binary complex. GatD has sequence similarities to the catalytic domains of glutamine amidotransferases, while MurT is similar in sequence to the substrate-binding domains of Mur ligases. Together, these two proteins catalyze the amidation of α -D-isoglutamic acid of cell wall precursor stem peptides in an ATP-dependent reaction. Mutants that are deficient in Lipid II amidation show a reduction in PG cross-linking and are more susceptible to antibiotics^{21–23}. Thus, intervening with the amidation reaction in *S. aureus* may represent a useful strategy to combat this pathogen.

In order to provide insight into the overall organization of this complex and to facilitate an understanding of the amidation mechanism, we have determined the crystal structure of the GatD/MurT complex. We find that the two proteins assemble into a curved, boomerang-shaped structure, with GatD docking to the C-terminal domain of MurT. Together with mutagenesis data and structural analysis of a complex with an ATP analog, our data provide an excellent foundation to understand the concerted activities of both proteins. Small-angle X-ray scattering (SAXS) experiments confirm that the complex has a similar open conformation in solution, and suggest that some flexibility between the domains exist. Structure-based sequence alignments demonstrate that several other pathogenic organisms have homologous enzyme complexes that likely function in the same manner. In combination with the established *in vitro* assays, our findings provide the basis for more directed inhibitor screenings.

Results

Formation and characterization of the GatD/MurT complex. Full-length GatD and MurT were co-expressed as described¹⁹, and the complex was purified using nickel affinity chromatography. A final size exclusion chromatography step demonstrated that the two proteins elute together, forming a stable complex in solution. The elution volume in size exclusion chromatography corresponds to an estimated molecular weight of 72 kDa, which is consistent with the calculated molecular weight of 78.8 kDa for a binary GatD/MurT complex. The SAXS data (see below) also clearly indicate that one copy of GatD and one copy of MurT assemble into a stable heterodimer.

Overall structure of GatD/MurT. The native structure of the GatD/MurT complex was solved at a resolution of 2.04 Å using single isomorphous replacement with anomalous scattering (SIRAS). The refined structure has excellent statistics (Table 1) and includes all residues of the expressed proteins with the exception of MurT residues 1–35, 195–196 and 434–437. These regions are poorly visible in the electron density maps and therefore likely have multiple conformations and increased mobility. The GatD/MurT heterodimer adopts a boomerang-shaped conformation, with GatD packing against the C-terminal domain of MurT (Fig. 1). As previously postulated¹⁹ and recently shown²⁴, GatD exhibits a class-I glutamine amidotransferase-like fold. A DALI search²⁵ identifies the enzymes HisH, PdxT and PurQ from *Thermotoga maritima* as the closest structural homologs (Z-values of 17.4, 15.8 and 15.0, respectively). Superimposition of GatD with structures obtained from a secondary structure-based search using HHPRED²⁶ reveals a well-conserved core architecture, with root-mean-square deviation (r.m.s.d.) values of 2.6 Å (all C- α atom pairs) over the entire length of GatD for the closest structural homolog, HisH. However, GatD distinguishes itself from other, homologous structures through the presence of an extended C-terminal helix, termed helix $\alpha 7$ (Fig. 1c). This helix mediates many of the contacts with MurT, explaining its presence in GatD.

MurT contains the Mur ligase middle and C-terminal domains typical for the Mur ligase family. The C-terminal domain is built around a central six-stranded, predominantly parallel β -sheet (Fig. 1c,e) that is sandwiched between four α -helices on one ($\alpha c1$ –4) and two α -helices ($\alpha c5$ –6) on the other side. The middle domain is constructed around a crescent-shaped, nine-stranded predominantly parallel β -sheet, which encloses a three-helix bundle ($\alpha m1$, $\alpha m7$ –8). The ATP molecule required for catalysis, replaced here by the non-hydrolyzable analog β, γ -imidoadenosine 5'-triphosphate (AMPPNP), is bound at the base of helix $\alpha m1$ with the adenine moiety bound in a cleft formed next to helix $\alpha m1$ while the phosphates are contacted by residues from strands $\beta m1$ and $\beta m3$ (Fig. 1d). The central β -sheet is capped on the opposite side by five α -helices. A RanBP-type Zinc-ribbon²⁷ is located on the side of the middle domain as an insertion between β -strands $\beta m6$ and $\beta m9$. This feature is missing in homologous proteins. A search with DALI²⁵ yielded MurF enzymes from *Pseudomonas aeruginosa*, *Escherichia coli* and *Acinetobacter baumannii* as the closest structural homologs, with Z-values of 23.2, 22.6, and 21.7, respectively. Mur ligases typically contain a third, N-terminal domain, which is not present in MurT. Instead, MurT only contains a truncated and likely flexible N-terminus (residues 1–37), which is not visible in our electron density maps.

Overall conformation. It can sometimes be challenging to assign correct physiologic contacts from crystal structures alone as crystal packing can offer alternative solutions for possible contacts between two subunits that are often not easy to distinguish from physiologic contacts²⁸. In order to validate the observed interaction between GatD and MurT, all possible GatD/MurT interfaces in the crystal were analysed with the PISA²⁸ and EPPIC²⁹ servers. Both algorithms clearly classify only the interaction depicted in Fig. 1c as physiologic. Other

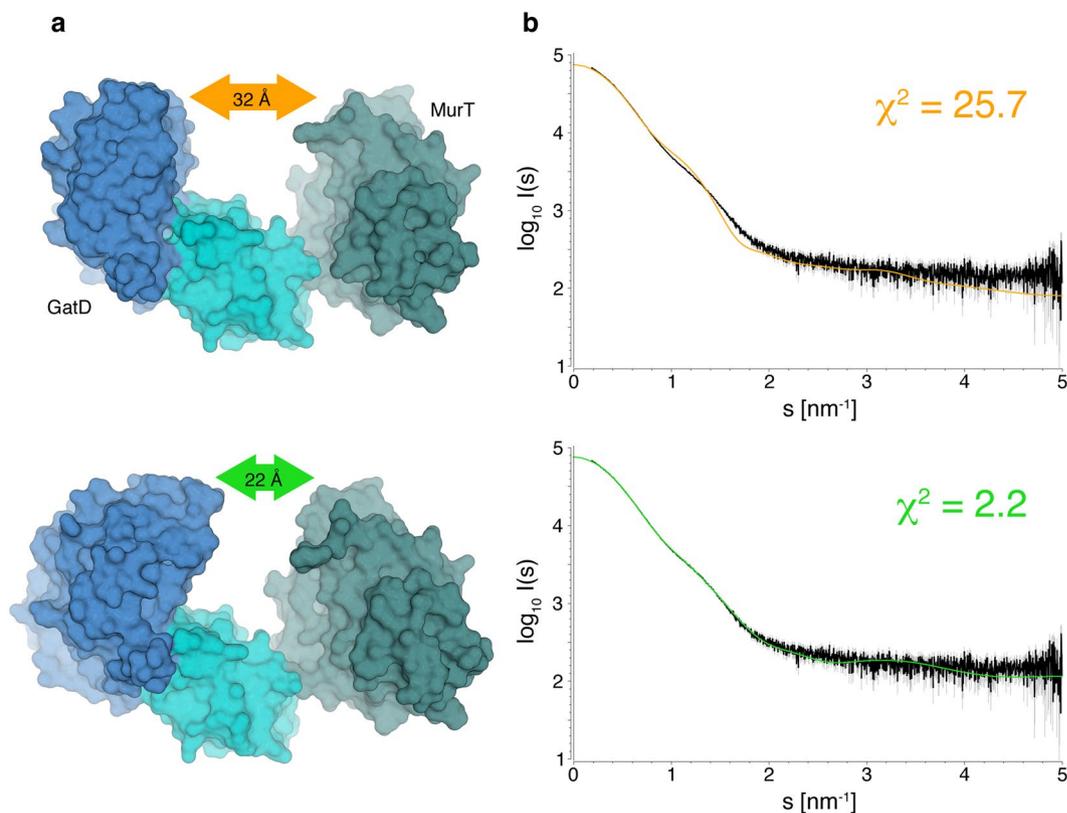


Figure 2. Conformation of the GatD/MurT complex in solution. **(a)** GatD/MurT crystal structure before (top) and after (bottom) refinement against SAXS data. While the overall organization and shape of the complex remain similar, the increased diameter of the GatD and MurT middle domain envelopes as well as the smaller gap of 22 Å instead of 32 Å between them suggest that some flexibility between domains exists in solution. **(b)** Fit of synthetic scattering profiles derived from the crystal structure (top) and refined model (bottom) against the experimental SAXS data.

possible interactions between GatD and MurT have much smaller buried surface areas and are classified as crystal packing artifacts. Moreover, they do not place critical catalytic residues (see below) in close proximity.

The overall structure of the GatD/MurT complex is also supported by solution SAXS experiments (Fig. 2). Scattering profiles derived from dilution series experiments were used to refine the crystallographic model. After rigid body refinement, the overall fit of a calculated scattering curve to the experimental data improved from an initial $\chi^2 = 25.7$ to $\chi^2 = 2.2$ (Fig. 2b). The refined model retains the overall characteristics of the crystal structure, but it has a more compact conformation with a closest distance between GatD and the MurT middle domain, measured at the top of the structure, of 22 Å instead of 32 Å (Fig. 2a). These experiments show that, while the conformation of the unliganded GatD/MurT complex is indeed open and elongated as observed in the crystal structure, the complex can breathe somewhat in solution. Unfortunately, SAXS data obtained from titrations with AMPPNP and a soluble Lipid II mimic (UDP-MurNAC-L-Ala-D-Glu- γ -L-Lys-D-Ala-D-Ala) resulted in moderate to severe aggregation of GatD/MurT (Supplementary Fig. S2) and could thus not be used for modelling. However, the observed change in protein solubility clearly suggests a considerable conformational rearrangement in the protein upon ATP and Lipid II binding.

Interactions between GatD and MurT. Extensive interactions between GatD and MurT establish a large, continuous interface that buries a total surface area of 940 Å² from solvent (calculated with PISA²⁸). Residues involved in interface formation are contributed by two consecutive MurT α -helices (α c2/3 and α c4), which tightly pack against two α -helices and a loop of GatD (helices α 2 and α 7, loop β 13- α 6 Figs 1 and 3). One of these α -helices is the C-terminal helix α 7 absent in homologous class-I glutamine amidotransferases, which do not form similar complexes with MurT. The interface is entirely devoid of solvent molecules. Two large aromatic (GatD-F146, MurT-Y354) and several aliphatic side chains (I20, I24, P194, and V231 in GatD and I353, L381 and L385 in MurT) that would otherwise be solvent-exposed are buried upon formation of the GatD/MurT contact. Figure 3a shows interactions made by the MurT-Y354 and MurT-L381 side chains as an example. MurT-Y354 also contacts two prolines in the GatD β 13- α 6 loop (P191 and P194), with P191 being especially well conserved throughout GatD homologs but not across functionally unrelated glutamine amidotransferases. These more centrally located hydrophobic interactions are augmented by intermolecular salt bridges (GatD-K31/MurT-E387, GatD-E225/MurT-K384, GatD-K195/MurT-E378, GatD-R235/MurT-D342), which are located towards the edges of the interface, and by hydrogen bonds (Fig. 3b,c). Residues participating in interface formation are highly

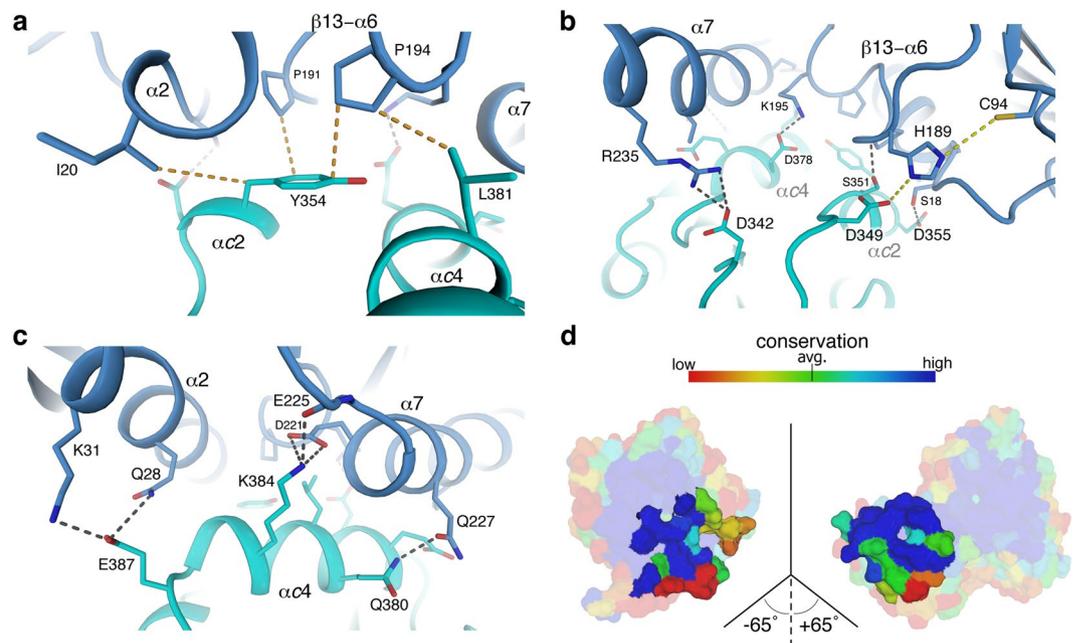


Figure 3. Architecture of the interface between GatD and MurT. **(a)** Hydrophobic interactions at the core of the GatD–MurT interface, centered around MurT–Y354. **(b)** and **(c)** Views of the rim of the interface showing mainly polar contacts or salt bridges. The C-terminal helix of GatD, which is unique to these types of proteins, is labeled helix $\alpha 7$, and the core hydrophobic loop was termed $\beta 13$ – $\alpha 6$. **(d)** Open-book view of the interface between GatD (left) and MurT (right). Sequence conservation scores were calculated with ConSurf³⁴ and mapped onto the protein surface as described in the methods section. The coloring scheme is also shown at the top of the panel. The majority of the interactions between GatD and MurT are very highly conserved, as indicated by dark and light blue colors.

conserved across species that have homologous proteins, as demonstrated by the conservation analysis shown in Fig. 3d. The interface also contains several residues with particularly small side chains (G21, A25, G190, A224, and A228 in GatD and S351 in MurT), and in each case a larger side chain would be incompatible with the observed interface.

Catalytic site of GatD. Class-I glutamine amidotransferases contain a highly conserved catalytic triad in their active site, which consists of a cysteine responsible for the initial nucleophilic attack on the substrate and nearby histidine and glutamic acid residues to form a proton relay chain. Even remotely related members of this enzyme family superimpose very well for the active site (Fig. 4). Earlier studies suggested that GatD residues C94 and H189 might form part of a catalytic triad¹⁹. In homologous GATases, the third residue of this triad is typically a glutamic acid that is located two residues downstream of the histidine in a conserved HPE tripeptide sequence³⁰. This glutamic acid is replaced with a proline (P191) in GatD and separated from H189 by a glycine, resulting in a HGP motif that is well conserved throughout putative GatD enzymes. However, as proline cannot function as a proton relay in a catalytic triad, GatD enzymes must either function with only a catalytic dyad, as recently suggested²⁴, or have arrived at a different solution to establish such a triad.

The structurally homologous enzymes PdxT from *Bacillus subtilis* (PDB ID 2nv0), two homoserine O-acetyltransferases from *Bacillus cereus* and *Thermotoga maritima* (PDB IDs 2vfj and 2h2w, respectively), and the glutamine amidotransferase Mfla_0438 from *Methylobacillus flagellatus* (PDB ID 3m3p), obtained from a HHPRED search, align well with the overall structure of GatD. A closer inspection of the active site architecture reveals that GatD–C94 superimposes well with the active site cysteine residues from these enzymes (Fig. 4a). The sulfhydryl groups in particular are located in almost identical positions. However, the side chain of GatD–H189 does not overlay well with the histidines from the related enzymes (Fig. 4a). Instead, GatD–H189 is shifted towards the heterodimer interface, where it contacts an aspartic acid (D349) of MurT, thus establishing a possible alternative catalytic triad (GatD–C94, GatD–H189, MurT–D349) involving residues from both GatD and MurT. A superimposition of our structure with the recently published structure of monomeric GatD²⁴ (PDB ID 5n9m) shows that the interaction with MurT is required to fully position the loop containing GatD–H189 (Supplementary Fig. S3). The close proximity of the side chains of GatD–C94 and GatD–H189 (3.7 Å) and GatD–His189 and MurT–D349 (3.1 Å) as well as their relative orientation suggest that these three residues might indeed function as a proton relay system in catalysis. GatD–H189 is followed by a short hydrophobic sequence framed by two proline residues, GatD–P191 and GatD–P194, that are involved in contacts with MurT (Fig. 3a). The arrangement of the putative catalytic triad is therefore dependent on correct assembly of the binary complex. As shown in Fig. 4, sequence comparison with confirmed³¹ and predicted homologs¹⁹ as well as a sequence- and structure based conservation analysis revealed the combination of GatD–H189 and the subsequent hydrophobic

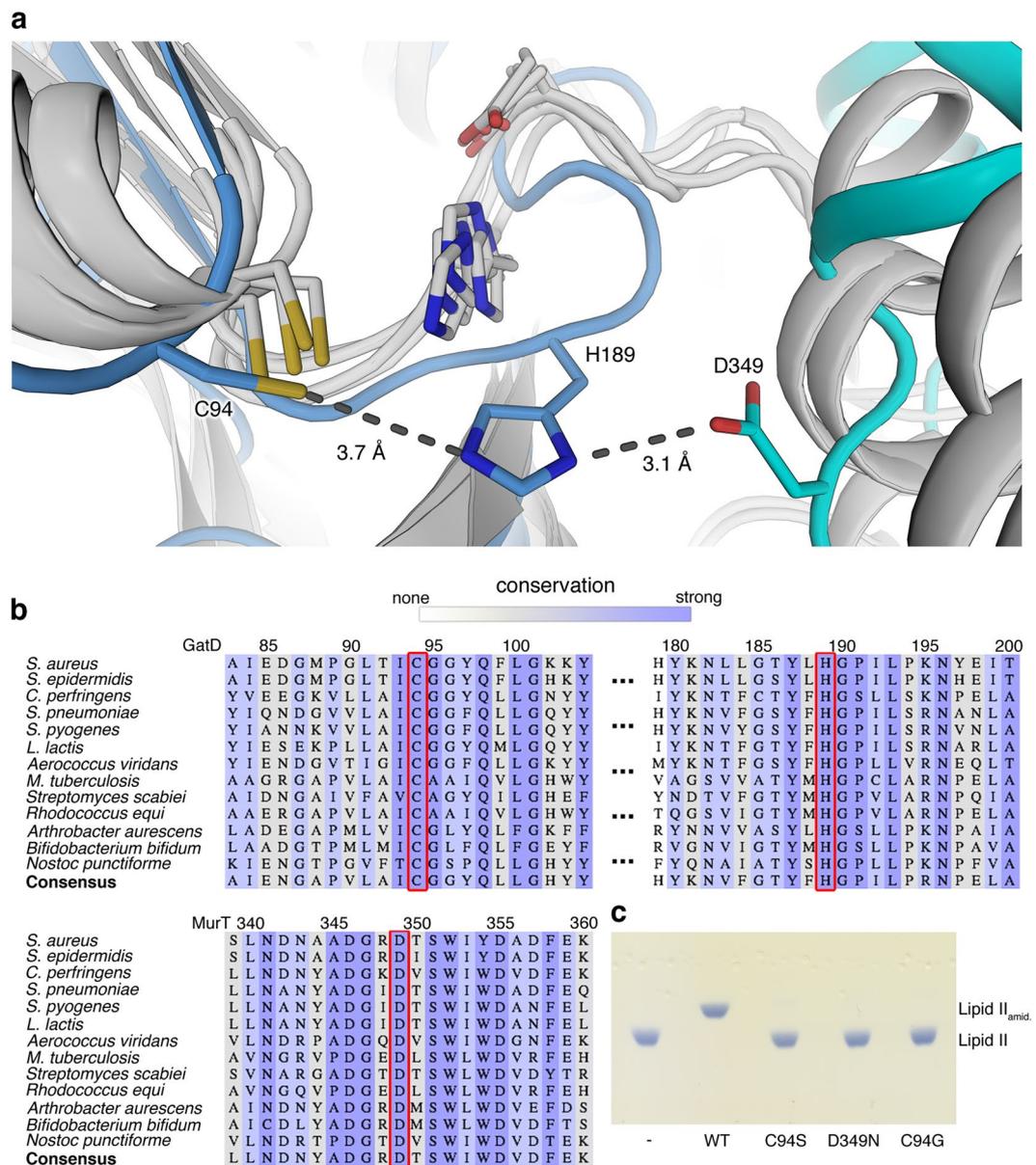


Figure 4. The putative catalytic triad of GatD/MurT. **(a)** Overlay of GatD/MurT with four glutamine amidotransferase fold-containing proteins obtained from a HHPRED²⁶ search (white cartoon and sticks). Whereas most GATases possess a conserved catalytic triad consisting of cysteine, histidine and glutamate residues, the glutamate is replaced by a proline (GatD-P191) in the GatD sequence and the conserved histidine (GatD-H189) is oriented towards an aspartic acid (MurT-D349) in MurT. **(b)** Multiple sequence alignment of putative homologous GatD/MurT enzymes. Conservation is color-coded, with white indicating low conservation, grey medium, and dark blue indicating high conservation. Residues GatD-C94, GatD-H189 and MurT-D349 are highly conserved (red box), as well as their immediate surroundings. **(c)** Thin-layer chromatography analysis of an activity assay of catalytic triad mutants. Mutation of MurT-D349 to asparagine completely abolishes catalysis *in vitro*, similarly to GatD-C94 mutations to either serine or glycine.

stretch to be highly conserved, suggesting that the GatD homologs possess similar catalytic triads and interfaces with corresponding MurT homologs. Similarly, the sequence context of MurT-D349 is conserved within MurT homologs, but not across other Mur ligases.

In order to investigate whether the spatial proximity of MurT-D349 to GatD-C94 and GatD-H189 is indicative of a possible involvement in catalysis, a MurT-D349N mutant was generated. *In vitro* amidation assays of Lipid II revealed severely reduced *in vitro* activity of this mutant compared to the wildtype GatD/MurT complex (Fig. 4, Supplementary Fig. S4). A conservative mutation of the active site cysteine (GatD-C94S) to serine completely abolished activity *in vitro*, similarly to the previously characterized, less conservative GatD-C94G mutant¹⁹. The folds and thermal stabilities of the mutants were assessed with circular dichroism (CD) spectroscopy and thermal shift assay (TSA), respectively. Both MurT-D349N and GatD-C94S possess a fold indistinguishable from that of

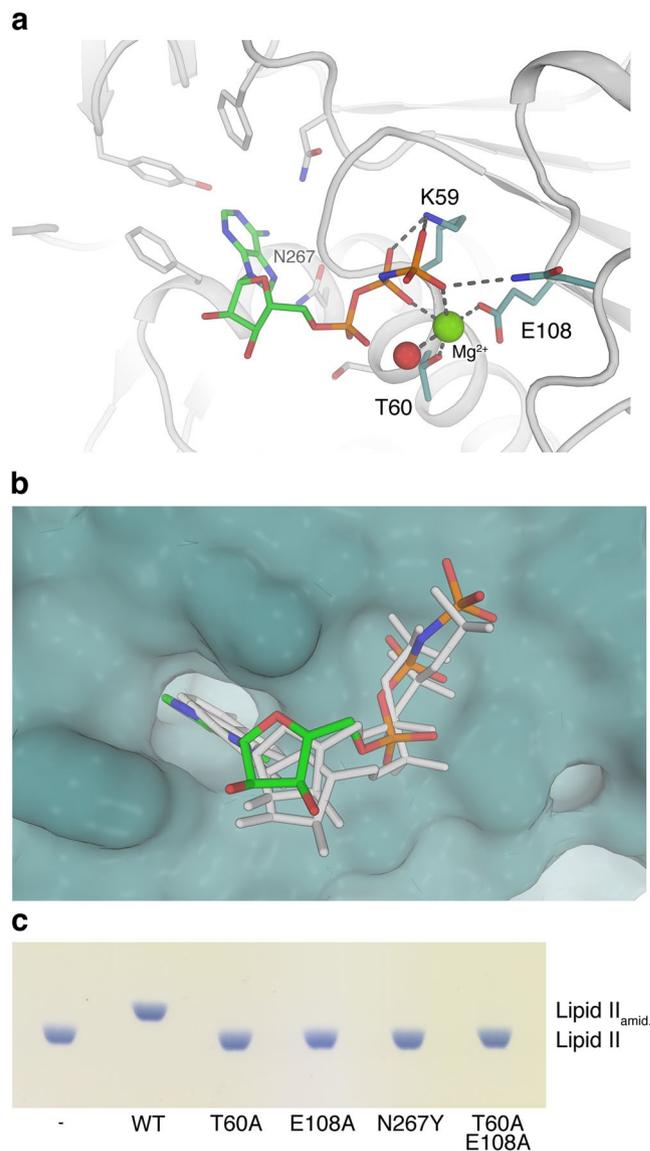


Figure 5. The AMPPNP binding site in GatD/MurT. **(a)** Catalytic center of MurT bound to the ATP analogue AMPPNP. The adenine base is inserted into a pocket composed of several aromatic residues and two asparagines, including N267, while the conserved K59, T60 and E108 residues coordinate the β and γ phosphates as well as a magnesium ion (green sphere) found in the active center of ATPases. A bound water is shown with a red sphere. **(b)** Superimposition of ATP analogues from *S. aureus* MurE (Protein Data Bank ID 4c12)³² and *P. aeruginosa* MurF (Protein Data Bank ID 4cvk) onto the MurT ATP-binding pocket in surface representation based on structural superimpositions of the entire domains. The MurT-bound AMPPNP is shown as a colored stick model, the superimposed nucleotides from the two related structures are shown as white sticks. **(c)** Thin-layer chromatography analysis of an activity assay of ATP-binding site mutants. Mutation of the magnesium-coordinating residues T60 and E108 to alanines completely abolishes catalysis. Replacement of the conserved N267 with a bulky tyrosine residue also impedes catalysis, probably by interfering with AMPPNP binding.

the wildtype protein (Supplementary Fig. S5) as well as virtually identical melting temperatures (Supplementary Fig. S6, Table S1) that lie well above the temperature used for activity assays. Taken together, the mutagenesis experiments demonstrate that both GatD-C94 and MurT-D349 are likely relevant for catalytic activity of the complex, and they therefore also provide additional support for the physiologic nature of the observed GatD/MurT interface.

The nucleotide binding site of MurT. To obtain insight into interactions of the GatD/MurT complex with ATP, we prepared a ternary complex by soaking crystals with 2.1 mM of AMPPNP, an ATP analog. The structure of this complex was solved at 2.14 Å resolution using molecular replacement (Table 1), and the corresponding electron density allowed us to unambiguously model AMPPNP into a binding site in the middle domain of MurT (Figs 1d and 5, Supplementary Fig. S7). The binding site contains the consensus sequence GTNGKT¹⁹,

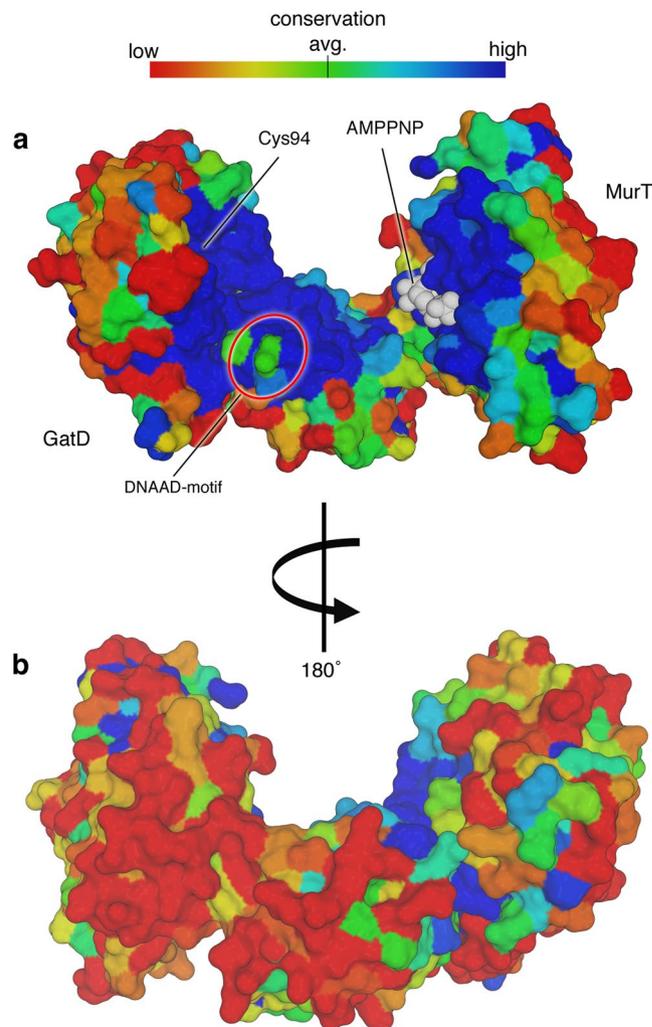


Figure 6. Conservation analysis of the GatD/MurT surface. **(a)** Inner and **(b)** outer surface of the GatD/MurT crystal structure colored by a color gradient indicating conservation, as shown on the top of the figure. Conservation scores were calculated with ConSurf³⁴ and mapped onto the protein surface as described in the methods section. Dark blue signifies a high conservation score. Most of the concave surface of GatD/MurT is highly conserved, whereas the remaining surface displays high variability. The AMPPNP ligand is shown with white spheres, and the location of the catalytic triad and the conserved DNAAD sequence that has been implicated in mucopeptide binding are indicated as well.

with residues K59 stabilizing the β and γ phosphate groups of AMPPNP and T60 coordinating a magnesium ion that is also contacted by the side chain of the conserved MurT-E108 residue. The magnesium ion helps in positioning the AMPPNP β and γ phosphates from the opposite side (Fig. 5a). The position of the observed ATP binding pocket is strictly conserved throughout Mur ligases as structural superposition of the middle domain of MurT with the middle domains of close sequence homologues MurE from *S. aureus*³² (PDB ID 4c12) and MurF from *Acinetobacter baumannii*³³ (PDB ID 4qdi) revealed nearly identical modes of nucleotide-binding (Fig. 5b). The observed binding mode is in excellent agreement with our mutagenesis data, as single or double mutants of nucleotide-binding site residues (T60A, E108A and N267Y) have completely lost their enzymatic activity (Fig. 5c, Supplementary Fig. S8). As with the previously described mutants, protein fold and stability were confirmed via CD spectroscopy and TSA (Supplementary Figs S5 and S6, Table S1).

Conservation mapping. Conservation analysis based on a redundancy-corrected alignment of 145 homologous sequences, as automatically performed on the ConSurf server³⁴ and projected onto the surface of the GatD/MurT model, revealed a high level of conservation at the inside of the crescent-shaped binary complex. While the outwards-facing surface of the protein, including that of the Zinc finger, is variable in sequence, the inner surface residues, which contain the ligand-binding regions of other Mur ligases, are strongly conserved. This includes the aforementioned components of the putative catalytic triad composed of GatD-C94, GatD-H189 and MurT-D349, as well as the MurT-GTNGKT nucleotide binding motif and MurT-E108 (Fig. 6), which coordinates the magnesium ion that lies in proximity to the bound ATP analog. The consensus sequence D(D,N)P(N,A) in the amino acid-binding pocket of MurE from *Thermotoga maritima* was found to mediate the recognition of L-lysine³⁵.

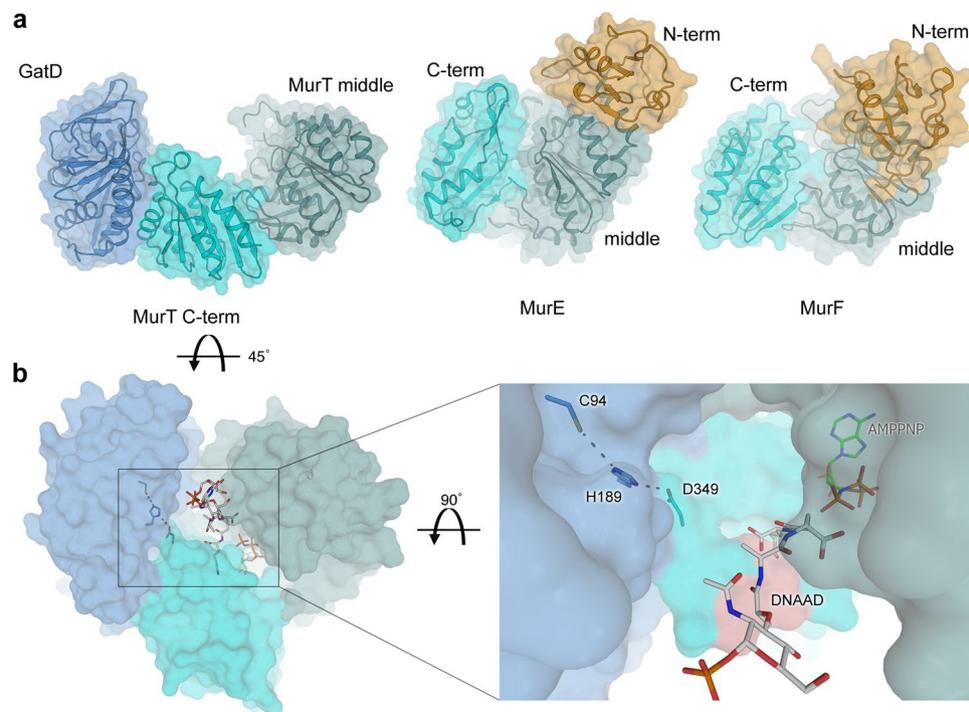


Figure 7. Comparison of GatD/MurT with Mur ligases and putative conformational changes upon ligand engagement. **(a)** Structural alignment of GatD/MurT (left) and *S. aureus* MurE³² and *A. baumannii* MurF³³ structures (right) was performed based on the Mur ligase middle domain (light green). The C-terminal domain (cyan) is rotated toward the middle domain in the MurE and MurF structures. A UDP-binding N-terminal domain (orange) is present in MurE and MurF but not in MurT. **(b)** Putative model of an active conformation of GatD/MurT with the domain movement modeled after the *S. aureus* MurE structure. The GatD catalytic triad and the bound AMPPNP molecule are shown as colored sticks, the relevant muramyltripeptide portion of the superimposed substrate of MurE (MurNAc-L-Ala- γ -D-Glu-L-Lys, but lacking UDP) is shown in grey stick representation for reference. The DNAAD motif, suggested to be involved in substrate binding is highlighted in salmon (right).

Structural work by Ruane and colleagues revealed that the *Staphylococcus aureus* MurE protein contains additional, often charged residues that are important for the interaction with L-lysine, e. g. Asp-406 and Glu-460³².

The MurT sequence DNAAD containing the conserved aspartate residue was previously suggested to be involved in interactions with the lysine residue of the Lipid II stem peptide¹⁹. Conservation mapping revealed that this motif is conserved and located in a large, concave area between the GatD catalytic triad and the ATP binding site in the middle domain of MurT (Fig. 7). Thus, it seems reasonable that this region is functionally relevant, either by engaging parts of Lipid II or by mediating conformational changes.

Discussion

The recently characterized GatD/MurT complex is responsible for the amidation of glutamic acid in the peptide stem structure of the *S. aureus* PG¹⁹. Inhibition of this reaction leads to lower growth rate, reduced resistance to beta-lactam antibiotics, and increased sensitivity to lysozyme in MRSA strains^{19,20}. Strategies to interfere with peptide stem amidation could therefore have potential for combatting *S. aureus*, including strains that are resistant to currently available antibiotics. Precise knowledge about the three-dimensional structure, organization and catalytic mechanism of the GatD/MurT complex is invaluable for fully exploiting this potential.

The crystal structure of GatD bound to MurT establishes an initial framework for probing the function of this enzyme complex and for developing strategies to intervene with amidation. The two proteins assemble into an elongated complex, with a large open space separating GatD from the middle domain of MurT. The previously identified catalytically active cysteine (GatD-C94)¹⁹ is located near the GatD/MurT interface, and is part of a catalytic triad that includes additional residues from both proteins (GatD-H189, MurT-D349). As our mutagenesis data confirm the importance of both GatD-C94 and MurT-D349 for catalysis while recent findings by Leisico and coworkers²⁴ corroborate the crucial role of GatD-H189, the location of this putative catalytic triad nicely explains the requirement of the assembled complex for activity, and it is in agreement with the observed interface between the two proteins. Complex formation with an ATP analog revealed a nucleotide-binding site in MurT that is similar to those found in other Mur ligases. It is conceivable that the large open space separating the GatD/MurT catalytic triad and ATP binding site may accommodate a portion of the bulky Lipid II substrate. The SAXS experiments demonstrate that the GatD/MurT complex has a similar open conformation in solution, and they also show that small movements of the terminal regions towards each other are possible. Such movements might be more pronounced in the presence of Lipid II, as the enzyme might partially close around it during catalysis.

The hydrophobicity of Lipid II and the lack of suitable, commercially available less hydrophobic Lipid II alternatives did not allow us to effectively pursue a complex of GatD/MurT bound to Lipid II. Soaking or co-crystallization experiments with soluble mimics of the peptide stem alone (e.g. glutamate, or commercially available peptides based on the *S. aureus* peptide stem sequence) did not result in complex formation, suggesting that additional portions of the Lipid II molecule are required for binding.

The *S. aureus* enzyme GatD/MurT likely belongs to a distinct group of enzymes that is strongly conserved throughout a subset of bacteria exhibiting a high degree of peptidoglycan amidation¹⁶, as suggested by conservation analysis¹⁹. The organisms that express GatD/MurT homologs include a number of established human pathogens, such as *Streptococcus pneumoniae*, for which the GatD/MurT enzyme complex was recently characterized³¹, and, likely, *Mycobacterium tuberculosis*. Knowledge about the structure and catalytic mechanism of GatD/MurT will therefore transcend the *S. aureus* field.

Structurally, GatD belongs to the family of glutamine amidotransferases (GATases), whose members are involved in a multitude of biosynthetic processes. While GATases exhibit a wide variety of substructure variants due to differing substrates or co-substrates, their core architecture is highly conserved, including that of GatD. The elongated C-terminal loop leading to an additional C-terminal α -helix (helix α 7) is, however, unique to GatD. This segment contributes 36% of the dimer interface with MurT, and the length of the linker connecting it to the rest of the protein suggests that it may allow for a certain degree of flexibility. Thus, the linker could help mediate breathing of the complex, which would explain the observed differences in the overall GatD/MurT conformation between the crystal structure and the solution scattering data (Fig. 2).

While MurT lacks the N-terminal domain typically found in Mur ligases, the overall folds of the middle and C-terminal MurT domains agree very well with the equivalent domains in published structures of different Mur ligases from various organisms. Although the position of the observed ATP-binding pocket is strictly conserved throughout Mur ligases (Fig. 5), the angle between the middle and C-terminal domains varies significantly across different enzymes, indicating an inherent flexibility of this feature. The almost linear arrangement between the two MurT domains appears to be unique to MurT as all other known conformations of this two-domain segment are more closed in Mur ligases. The two most striking structural differences between MurT and related enzymes are the truncated N-terminal domain and the Cys₄ Zinc-ribbon situated at the side of the middle domain. While the role of this last feature is unclear, it forms a very exposed structure protruding from the main bulk of the MurT middle domain. Zinc fingers of this class have been previously shown to provide a platform for protein-protein interactions³⁶. The zinc finger is less conserved compared with many other features of GatD/MurT, and it appears to be present only in a subset of bacteria with putative *gatD/murT* genes.

The truncated N-terminus (with respect to other Mur ligases), on the other hand, is a well-conserved feature of MurT in different organisms. The closely related MurC-F enzymes catalyze the consecutive addition of amino acids of the peptide stem to UDP-activated MurNAc in the initial phase of peptidoglycan biosynthesis. The roughly 90 residues long N-terminal domain present in these enzymes plays an important role in substrate recognition by interacting with its UDP moiety. In contrast, MurT has a much shorter N-terminus of only 35 residues, which appears to be structurally flexible and poorly ordered as it could not be observed in the crystal structure. Furthermore, the positioning of the pyrophosphate linker in the UDP-conjugated substrates of ligand-bound structures of *S. aureus* MurE (PDB ID 4c12) and *Pseudomonas aeruginosa* MurF (PDB ID 4cvk) coincides with the boundary between their N-terminal and middle domains. In the GatD/MurT substrate Lipid II, the pyrophosphate linkage connects the sugar moiety to the undecaprenyl membrane anchor, possibly marking the position at which Lipid II emerges from the membrane. Hence, a UDP-MurNAc-binding N-terminal domain is not only unnecessary, it would also likely clash with the plasma membrane.

How does Lipid II bind to GatD/MurT, and how does catalysis proceed? Prompted by the drastically altered solubility of the protein observed by SAXS upon minimal substrate mimic addition (Supplementary Fig. S2) and based on previously published data revealing that MurT is able to replace MurE *in vitro*¹⁹ as well as on alignments with the substrate-bound structures of MurE and MurF, which consistently show closed conformations, an analogous structural rearrangement upon substrate binding can be modeled for MurT. To generate this model, the C-terminal domains of MurE or MurF were superimposed with that of MurT, and the MurT middle domain was then moved to fit those of MurE or MurF, producing a more closed conformation of MurT that mimics those present in the other two enzymes (Fig. 7). If the GatD structure is then added to the modified MurT using the known GatD/MurT interface, it becomes clear that the distance between the MurT middle domain and GatD is drastically reduced. The gap between GatD-C94 and the γ -phosphate of AMPNP diminishes from over 40 Å to just over 20 Å. In such a putative more closed conformation of GatD/MurT, the peptide stem of Lipid II would be sandwiched between the MurT middle and C-terminal domains in linear continuation of the triphosphate axis of the nucleotide (Fig. 7b). It is conceivable that channeling of nascent ammonia from the active site of GatD to the mucopeptide substrate in MurT, which would then lie in close proximity to the nucleotide, may be enabled by such a closed conformation. The DNADD motif, which was previously suggested to be involved in the engagement of L-lysine in the mucopeptide sequence¹⁹, indeed comes into close proximity with the modeled MurE product, which was also extracted from PDB ID 4c12. Due to the direct backbone connection with MurT-D349, the relative position of the residues forming the putative catalytic triad may be influenced by Lipid II binding to the enzyme. Thus, hydrolysis of glutamine in the absence of an acceptor substrate and hence escape and accumulation of ammonia in the cell could be avoided.

The model presented here was obtained by domain superimposition with known substrate-bound structures of homologous proteins, and it is only meant to approximate the possible conformations of ligand-bound GatD/MurT and the Lipid II substrate. While the superimposition predicts a closed conformation of substrate-bound GatD/MurT, some residues located at the newly formed interface between the MurT C-terminal and middle domains clash both with the respective other domain and with the modeled MurE ligand, and this demonstrates that additional movements in the protein are required to properly accommodate the substrate. As the physiological

substrate of MurT, Lipid II, differs from that of MurE, it would not be surprising if the actual active conformation of GatD/MurT may be a few degrees more open to allow accommodation of the large Lipid II molecule. The precise conformation of a Lipid II-bound complex of GatD/MurT is still unknown and likely somewhat different from the one shown in Fig. 7. We therefore believe that performing in-depth molecular docking calculations would be inappropriate at this time.

Our model suggests that the mechanism of Lipid II amidation likely starts with an ATP-bound, open conformation of GatD/MurT similar to that observed in our crystal structure. A domain rearrangement between the two MurT domains to generate a more closed conformation could be triggered by the binding of the substrate peptide, and this would enable both the hydrolysis of glutamine by GatD through movement of the triad residues and the correct positioning of the free carboxylate of D-iso-glutamate in the peptide stem of Lipid II with respect to ATP, thus allowing for a concerted amidation reaction. Of course, this hypothesis needs to be tested in future experiments.

Peptidoglycan is characteristically modified in many Gram-positive pathogens and amidation is known to affect the level of methicillin resistance^{21,37} and to contribute to vancomycin susceptibility in *S. aureus*³⁸, and thus inhibition of the GatD/MurT complex may represent a strategy to combat *S. aureus* and perhaps also other related bacterial pathogens that amidate their peptide stem in a similar manner. The nucleotide-binding region of MurT may provide a good base for developing multiple-target inhibitors, owing to the similarity of this site with those of other Mur ligases. However, as most monomeric Mur ligases are also present in beneficial bacteria such as *E. coli* or other components of our gut microbiota, it may be desirable to develop drugs more specifically aimed at GatD/MurT. As the GatD/MurT interface is highly conserved and likely unique to complexes between these two proteins, interfering with the interaction between GatD and MurT could selectively inactivate the amidation reaction. It is conceivable that small molecules that mimic components of the interface could serve to block the reaction carried out by the GatD/MurT complex.

Materials and Methods

Expression and purification. A pET21b vector (Novagen) containing the *gatD/murT* operon for co-expression of untagged MurT with GatD as a C-terminal His-tag fusion protein [1] was introduced into *E. coli* BL21 (Promega) cells. The bacteria were grown in LB-medium (75 µg/mL ampicillin) at 37 °C. At an OD₆₀₀ of 0.6, the cultures were cooled to 30 °C and 0.25 mM IPTG was added for induction. Cells were harvested by centrifugation after 5 hours and stored at −20 °C until purification.

Frozen pellets were resuspended in lysis buffer (50 mM Tris pH 8.5 at 4 °C, 300 mM NaCl, 10 mM imidazole, 1 mM DTT). Next, 200 µg/mL lysozyme, 1:10000 DNase I (ThermoFisher) and 1 mg RNase A (Sigma) were added and cells were incubated on ice for 30 min. Following lysis by sonication, the lysate was clarified by centrifugation and added to a Ni-NTA-agarose slurry (Novagen). The mixture was incubated on a tilt shaker at low rpm for 18 h at 4 °C and subsequently transferred to a EconoPac gravity flow column (BioRad). Following a wash with lysis buffer, weakly bound material was removed by washing with lysis buffer containing 50 mM imidazole, and the target protein was then eluted with lysis buffer containing 350 mM imidazole. Buffer exchange to SEC I buffer (50 mM Tris pH 8.2 @ 4 °C, 500 mM NaCl, 50 mM MgCl₂, 5 mM KCl, 5 mM DTT) was carried out directly afterwards using PD-10 desalting columns (GE). Sample homogeneity was achieved by means of two subsequent iterations of preparative size exclusion chromatography using a HiLoad 26/60 column packed with Superdex 200 column material (GE Healthcare). In the second iteration, 5 mM TCEP was substituted for DTT as a reducing agent (Buffer SEC II). Throughout the purification, sample purity was assessed by SDS-PAGE. The protein was freshly prepared for each analysis.

Crystallization and structure determination. Crystals of GatD/MurT were grown in a sitting-drop vapor diffusion setup using protein concentrations of 2–5 mg/mL in a broad range of Tris-buffered conditions containing PEG8000, MgCl₂ and glycerol. The conditions from which the crystals for this study were obtained are 0.1 M Tris pH 8.6, 40% (w/v) PEG8000, 0.35 M MgCl₂ for the native data, 0.1 M Tris pH 9.0, 26.7% (w/v) PEG8000, 0.35 M MgCl₂ for the thiomersal derivative, and 0.1 M Tris pH 9.1, 18% (w/v) PEG8000, 14% glycerol, 0.35 M MgCl₂ for the AMPPNP soak, respectively. Cryoprotection was performed by adding 20% MPD or 25% glycerol to the reservoir solution, and crystals were flash-cooled in liquid nitrogen. X-ray data were collected at beamlines X10SA and X06DA of the Swiss Light Source (SLS) in Villigen (Switzerland). Data were integrated and reduced using the XDS program package³⁹.

Crystal derivatization. Crystals were derivatized with 10 mM thiomersal for phase determination and a mixture of 2.1 mM of the ATP analogue β,γ-imidoadenosine 5'-triphosphate (AMPPNP), 2.1 mM of the glutamine analogue 6-Diazo-5-oxo-L-norleucine (DON), and 2.5 mM UDP-MurNAc-L-Ala-D-Glu-γ-L-Lys-D-Ala-D-Ala for evaluating complex formation. Of the latter three compounds, only AMPPNP was found to bind to the crystals based on the inspection of difference electron density maps. Derivatization was performed by transferring native crystals into drops of otherwise equal composition containing the compound to be soaked. After soaking times of 10 to 30 min, crystals were cryoprotected and flash-cooled in liquid nitrogen.

Phasing and initial model building. A dataset obtained from a crystal derivatized with 10 mM thiomersal was used to extract initial phase information using the single isomorphous replacement with anomalous scattering (SIRAS) approach. The derivative data were scaled against the native data using CAD and Scaleit of the CCP4 program suite⁴⁰. Using the program suite autoSHARP^{41,42}, a substructure of 13 mercury sites was determined, followed by heavy atom refinement and density modification. Initial automated model building was performed using the AutoBuild Wizard⁴³ in the PHENIX program suite prior to transfer to the native data set.

Model building and refinement. Manual model building was carried out in Coot⁴⁴ and alternated with reciprocal space refinement using the programs REFMAC5⁴⁵ and PHENIX.REFINE⁴⁶. The data set collected from crystals soaked with AMPPNP was phased by molecular replacement using the program PHASER⁴⁷. Prior to model building, simulated annealing was performed using PHENIX.REFINE in order to avoid model bias from the original dataset.

SAXS. Small-angle X-ray scattering (SAXS) data were collected at beamline P12 at the German Electron Synchrotron (DESY) in Hamburg, Germany. Triplicates of a two-fold dilution series of GatD/MurT in SEC II buffer ranging from 0.5 to 8 mg/mL were recorded and extrapolated to zero concentration to remove possible concentration effects using the PRIMUS interface of the ATSAS program suite⁴⁸. Rigid-body refinement of the native crystal structure was carried out with the program SREFLEX⁴⁹. Synthetic scattering profiles for comparison with the experimental data were generated using CRY SOL⁵⁰. SAXS experiments were repeated in the presence of increasing concentrations of the ATP analog AMPPNP (10 μ M–5 mM) and the soluble Lipid II mimic UDP-MurNAc-L-Ala-D-Glu- γ -L-Lys-D-Ala-D-Ala (8 μ M–1 mM).

Mutagenesis and *in vitro* amidation assay. Site-directed mutagenesis was performed according to the manufacturer's instructions using plasmid pET21-*murT/gatD* as the template to generate active site mutants GatD-C94S and MurT-D349N (QuikChange Lightning Site-Directed Mutagenesis Kit, Agilent) and to generate mutations in the ATP binding site (MurT mutants T60A, E108A, N267Y and double mutant T60A E108A; Q5 site-directed mutagenesis kit, New England Biolabs). Proteins were expressed and purified and impact of mutations on enzyme activity were tested as previously described¹⁹.

Circular dichroism (CD) spectroscopy. CD spectroscopy experiments were performed on a JASCO J-720 spectrometer under nitrogen flow. Protein concentration was set to 0.3 mg/mL in ten-fold diluted SEC II buffer. The resulting spectra were corrected for concentration, protein size and cuvette thickness. Mutant protein spectra were compared with the wildtype.

Thermal shift assay (TSA). TSA was performed on a QuantStudio 5 real-time PCR cycler (Applied Biosystems, Thermo Fisher Scientific). Wildtype and mutant GatD/MurT were supplied at 0.3 mg/mL and Protein Thermal Shift DyeTM (Applied Biosystems, Thermo Fisher Scientific) was added as a TSA fluorophore. The samples were equilibrated at 4 °C and gradually heated to 95 °C over 30 min while monitoring dye fluorescence. The averaged protein melting temperature for the wildtype and each mutant was derived from the melting curve inflection point using seven technical replicates and the Protein Thermal ShiftTM Software v1.3 (Applied Biosystems, Thermo Fisher Scientific).

Homology searches, conservation analysis and interface mapping. Sequence-based conservation analysis was performed using ConSurf³⁴. Herein, an initial search was automatically performed separately for GatD and MurT using BLAST (<https://blast.ncbi.nlm.nih.gov/>) and a multiple sequence alignment (MSA) of 145 sequences each was generated using MAFFT⁵¹. The resulting conservation scores were then projected onto the GatD/MurT crystal structure. Color coding of conservation was done by setting the average conservation score of the alignment to zero. More or less conserved residues were then colored according to a color ramping scheme, with dark blue and red indicating highest and lowest conservation scores, respectively. Structural homology searches were conducted at a secondary structure level using the HHPRED tool implemented in the bioinformatics toolkit of the Max-Planck-Institute for Developmental Biology in Tübingen⁵².

Additionally, previously postulated putative¹⁹ and confirmed³¹ homologous proteins were aligned separately, using the Clustal Ω algorithm⁵³.

Finally, in order to identify related proteins based on their entire structure and quantify the structural similarities, a DALI search²⁵ was performed with the final crystal structure against the entire Protein Data Bank (PDB) repository at www.rcsb.org.

In order to classify contact surfaces observed in the crystal, the crystallographic assembly was subjected to analysis by the PISA²⁸ and EPPIC²⁹ servers.

Data deposition. The GatD/MurT structures described here have been deposited with the Protein Data Bank (www.rcsb.org) with PDB IDs 6GS2 (unliganded GatD/MurT) and 6H5E (complex with AMPPNP).

References

1. Foster, T. J. The *Staphylococcus aureus* "superbug". *J Clin Invest* **114**, 1693–1696 (2004).
2. Zeller, J. L., Burke, A. E. & Glass, R. M. JAMA patient page. *MRSA infections*. *JAMA* **298**, 1826 (2007).
3. Roemer, T., Schneider, T. & Pinho, M. G. Auxiliary factors: a chink in the armor of MRSA resistance to beta-lactam antibiotics. *Curr Opin Microbiol* **16**, 538–548 (2013).
4. Weidenmaier, C. *et al.* Role of teichoic acids in *Staphylococcus aureus* nasal colonization, a major risk factor in nosocomial infections. *Nat Med* **10**, 243–245 (2004).
5. Fournier, B. & Philpott, D. J. Recognition of *Staphylococcus aureus* by the innate immune system. *Clin Microbiol Rev* **18**, 521–540 (2005).
6. Kraus, D. & Peschel, A. Molecular mechanisms of bacterial resistance to antimicrobial peptides. *Curr Top Microbiol Immunol* **306**, 231–250 (2006).
7. Peschel, A. & Sahl, H. G. The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nature reviews. Microbiology* **4**, 529–536 (2006).
8. Vollmer, W., Blanot, D. & de Pedro, M. A. Peptidoglycan structure and architecture. *FEMS Microbiol Rev* **32**, 149–167 (2008).
9. Barreateau, H. *et al.* Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiol Rev* **32**, 168–207 (2008).

10. Bouhss, A., Trunkfield, A. E., Bugg, T. D. & Mengin-Lecreux, D. The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol Rev* **32**, 208–233 (2008).
11. Hegde, S. S. & Shrader, T. E. FemABX family members are novel nonribosomal peptidyltransferases and important pathogen-specific drug targets. *J Biol Chem* **276**, 6998–7003 (2001).
12. Schneider, T. *et al.* *In vitro* assembly of a complete, pentaglycine interpeptide bridge containing cell wall precursor (lipid II-Gly5) of *Staphylococcus aureus*. *Mol Microbiol* **53**, 675–685 (2004).
13. Mohammadi, T. *et al.* Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. *Embo J* **30**, 1425–1432 (2011).
14. Scheffers, D. J. & Pinho, M. G. Bacterial cell wall synthesis: new insights from localization studies. *Microbiol Mol Biol Rev* **69**, 585–607 (2005).
15. van Heijenoort, J. Lipid intermediates in the biosynthesis of bacterial peptidoglycan. *Microbiol Mol Biol Rev* **71**, 620–635 (2007).
16. Schleifer, K. H. & Kandler, O. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev* **36**, 407–477 (1972).
17. Nakel, M., Ghuysen, J. M. & Kandler, O. Wall peptidoglycan in *Aerococcus viridans* strains 201 Evans and ATCC 11563 and in *Gaffkya homari* strain ATCC 10400. *Biochemistry* **10**, 2170–2175 (1971).
18. Gotz, F., Bannerman, T. & Schleifer, K. H. The Genera *Staphylococcus* and *Micrococcus*. *Prokaryotes: A Handbook on the Biology of Bacteria, Vol 4, Third Edition*, 5–75 (2006).
19. Munch, D. *et al.* Identification and *in vitro* analysis of the GatD/MurT enzyme-complex catalyzing lipid II amidation in *Staphylococcus aureus*. *PLoS Pathog* **8**, e1002509 (2012).
20. Figueiredo, T. A. *et al.* Identification of genetic determinants and enzymes involved with the amidation of glutamic acid residues in the peptidoglycan of *Staphylococcus aureus*. *PLoS Pathog* **8**, e1002508 (2012).
21. Gustafson, J., Strassle, A., Hachler, H., Kayser, F. H. & Berger-Bachi, B. The femC locus of *Staphylococcus aureus* required for methicillin resistance includes the glutamine synthetase operon. *J Bacteriol* **176**, 1460–1467 (1994).
22. Strandén, A. M., Ehlert, K., Labischinski, H. & Berger-Bachi, B. Cell wall monoglycine cross-bridges and methicillin hypersusceptibility in a femAB null mutant of methicillin-resistant *Staphylococcus aureus*. *J Bacteriol* **179**, 9–16 (1997).
23. Boyle-Vavra, S., Labischinski, H., Ebert, C. C., Ehlert, K. & Daum, R. S. A spectrum of changes occurs in peptidoglycan composition of glycopeptide-intermediate clinical *Staphylococcus aureus* isolates. *Antimicrob Agents Chemother* **45**, 280–287 (2001).
24. Leisico, F. *et al.* First insights of peptidoglycan amidation in Gram-positive bacteria - the high-resolution crystal structure of *Staphylococcus aureus* glutamine amidotransferase GatD. *Sci Rep* **8**, 5313 (2018).
25. Holm, L. & Laakso, L. M. Dali server update. *Nucleic Acids Res* **44**, W351–355 (2016).
26. Zimmermann, L. *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J Mol Biol* (2017).
27. Malgieri, G. *et al.* The prokaryotic zinc-finger: structure, function and comparison with the eukaryotic counterpart. *The FEBS journal* **282**, 4480–4496 (2015).
28. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797 (2007).
29. Duarte, J. M., Srebnik, A., Scharer, M. A. & Capitani, G. Protein interface classification by evolutionary analysis. *BMC Bioinformatics* **13**, 334 (2012).
30. Massiere, F. & Badet-Denisot, M. A. The mechanism of glutamine-dependent amidotransferases. *Cell Mol Life Sci* **54**, 205–222 (1998).
31. Liu, X. *et al.* High-throughput CRISPRi phenotyping identifies new essential genes in *Streptococcus pneumoniae*. *Mol Syst Biol* **13**, 931 (2017).
32. Ruane, K. M. *et al.* Specificity determinants for lysine incorporation in *Staphylococcus aureus* peptidoglycan as revealed by the structure of a MurE enzyme ternary complex. *J Biol Chem* **288**, 33439–33448 (2013).
33. Cha, S. S., An, Y. J., Jeong, C. S., Yu, J. H. & Chung, K. M. ATP-binding mode including a carbamoylated lysine and two Mg²⁺ ions, and substrate-binding mode in *Acinetobacter baumannii* MurF. *Biochemical and biophysical research communications* **450**, 1045–1050 (2014).
34. Ashkenazy, H. *et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* **44**, W344–350 (2016).
35. Boniface, A., Bouhss, A., Mengin-Lecreux, D. & Blanot, D. The MurE synthetase from *Thermotoga maritima* is endowed with an unusual D-lysine adding activity. *J Biol Chem* **281**, 15680–15686 (2006).
36. Gamsjaeger, R., Liew, C. K., Loughlin, F. E., Crossley, M. & Mackay, J. P. Sticky fingers: zinc-fingers as protein-recognition motifs. *Trends Biochem Sci* **32**, 63–70 (2007).
37. Berger-Bachi, B., Strassle, A., Gustafson, J. E. & Kayser, F. H. Mapping and characterization of multiple chromosomal factors involved in methicillin resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother* **36**, 1367–1373 (1992).
38. Cui, L., Murakami, H., Kuwahara-Arai, K., Hanaki, H. & Hiramatsu, K. Contribution of a thickened cell wall and its glutamine nonamidated component to the vancomycin resistance expressed by *Staphylococcus aureus* Mu50. *Antimicrob Agents Chemother* **44**, 2276–2285 (2000).
39. Kabsch, W. Xds. *Acta crystallographica. Section D, Biological crystallography* **66**, 125–132 (2010).
40. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallographica* **D50**, 760–763 (1994).
41. Bricogne, G., Vornrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta crystallographica. Section D, Biological crystallography* **59**, 2023–2030 (2003).
42. Vornrhein, C., Blanc, E., Roversi, P. & Bricogne, G. Automated structure solution with autoSHARP. *Methods Mol Biol* **364**, 215–230 (2007).
43. Terwilliger, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta crystallographica. Section D, Biological crystallography* **64**, 61–69 (2008).
44. Emsley, P. & Cowtan, K. Coot: model building tools for molecular graphics. *Acta Crystallographica* **D60**, 2126–2132 (2004).
45. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D* **53**, 240–255 (1997).
46. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta crystallographica. Section D, Biological crystallography* **68**, 352–367 (2012).
47. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674 (2007).
48. 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).
49. Panjkovich, A. & Svergun, D. I. Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis. *Phys Chem Chem Phys* **18**, 5707–5719 (2016).
50. Svergun, D., Barberato, C. & Koch, M. H. J. CRYSOLE - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography* **28**, 768–773 (1995).
51. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059–3066 (2002).

52. Soding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**, W244–248 (2005).
53. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
54. Bond, C. S. TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics* **19**, 311–312 (2003).

Acknowledgements

We thank members of the Stehle and Schneider laboratories for critical reading of the manuscript and for discussion. We thank Lukas Schönberger and Astrid Illigmann for technical assistance. We gratefully acknowledge the Swiss Light Source (Villigen, Switzerland) for beamtime and the staff at beamline X06DA and X06SA for assistance during data collection. We also thank members of the Deutsches Elektronen Synchrotron (DESY) in Hamburg, especially Martin Schroer, for assistance with SAXS data collection and interpretation.

Author Contributions

T.Sch. and T.St. conceived and supervised the experiments, E.R.N., L.M.F., V.N. and E.S. performed the structural analyses, G.Z. assisted with the structural analyses, A.M. generated and analyzed mutants. E.R.N., A.M., T.Sch. and T.St. analyzed the data and wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-31098-x>.

Competing Interests: The authors declare no competing interests as defined by Nature Research, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018

SCIENTIFIC REPORTS

OPEN

Author Correction: Structural basis of cell wall peptidoglycan amidation by the GatD/MurT complex of *Staphylococcus aureus*

Erik R. Nöldeke¹, Lena M. Muckenfuss^{1,4}, Volker Niemann^{1,5}, Anna Müller², Elena Störk¹, Georg Zocher¹, Tanja Schneider² & Thilo Stehle^{1,3}

Correction to: *Scientific Reports* <https://doi.org/10.1038/s41598-018-31098-x>, published online 28 August 2018

In Fig. 1A, a carbonyl group is missing in the schematic drawing of the reaction product. The correct Fig. 1 appears below.

¹Interfaculty Institute of Biochemistry, University of Tübingen, D-72076, Tübingen, Germany. ²Institute for Pharmaceutical Microbiology, University of Bonn, D-53115, Bonn, Germany. ³Vanderbilt University School of Medicine, Nashville, Tennessee, 37232, USA. ⁴Present address: Department of Biochemistry, University of Zurich, CH-8057, Zurich, Switzerland. ⁵Present address: Hain Lifescience GmbH, D-72147, Nehren, Germany. Correspondence and requests for materials should be addressed to T.Stehle (email: thilo.stehle@uni-tuebingen.de)

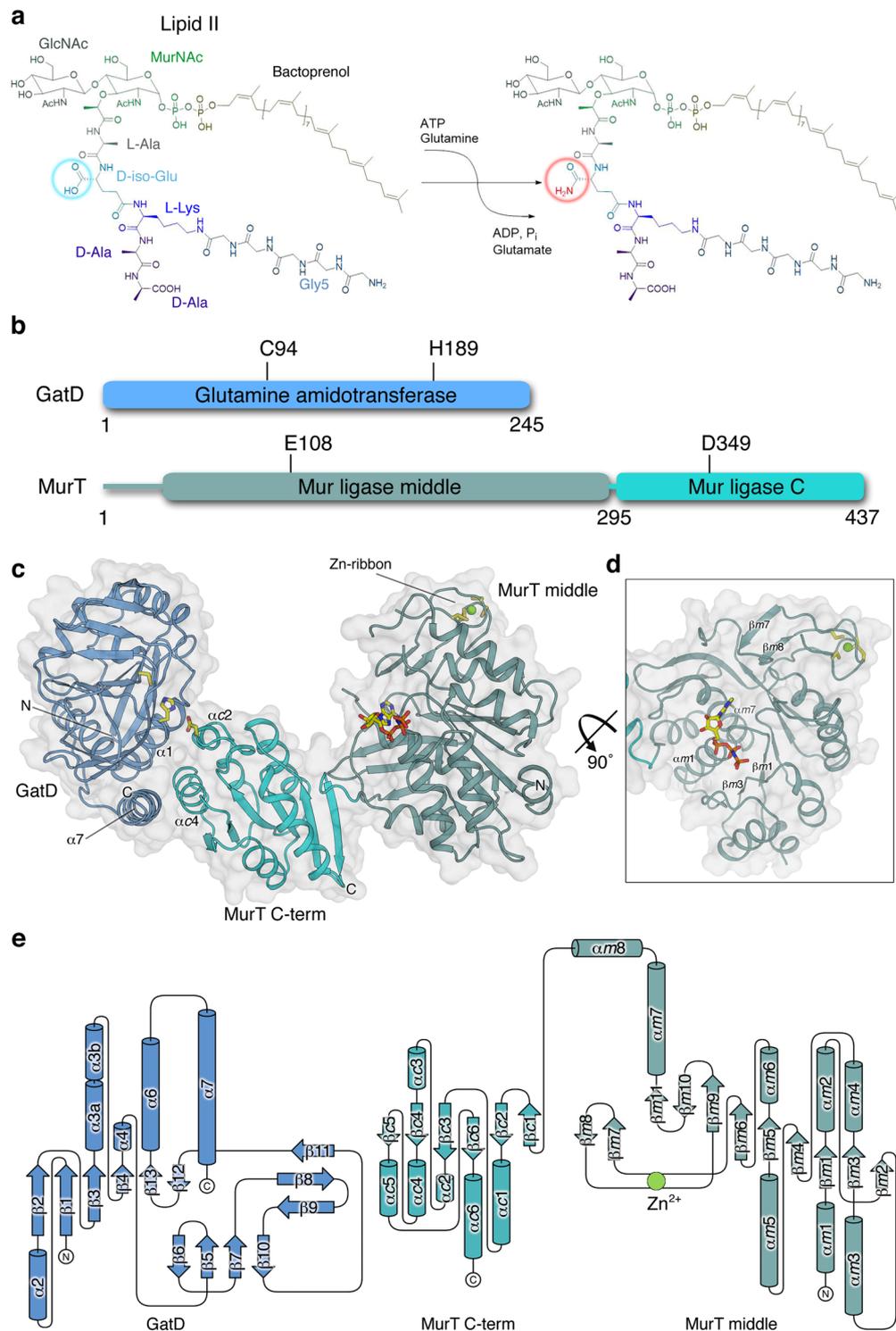


Figure 1. Overall structure and organization of the GatD/MurT complex. **(a)** Reaction catalyzed by GatD/MurT. The free α -carboxyl of D-iso-glutamate in the peptide stem is amidated in a glutamine- and ATP-dependent reaction. **(b)** Schematic overview of GatD and MurT proteins. GatD consists of a single glutamine amidotransferase (GATase) domain with a cysteine at position 94 as the active residue and a histidine at position 189 as a component of the catalytic triad¹⁹. MurT is composed of two domains: a Mur ligase middle domain (MurT middle) containing the canonical ATP binding site and, surprisingly, a ribbon-type Zinc finger, and a C-terminal Mur ligase domain (MurT C-term). MurT residue glutamate 108 participates in ATP hydrolysis, and aspartate 349 forms the third residue in the putative catalytic triad. **(c)** Overview of the GatD/MurT structure. GatD and MurT form a boomerang-shaped complex, with GatD contacting the MurT C-term domain through contacts that are in part mediated by helix α 7 of GatD. Catalytic triad residues GatD-C94, GatD-H189, MurT-D349 and the bound nucleotide AMPNP are shown in stick representation. The zinc ion in the Cys₄ zinc ribbon of MurT is shown as a green sphere, and the four cysteine residues ligating it are shown

as sticks. **(d)** Tilted view of the MurT middle domain to show the central β -sheet and the bound AMPNP and its surrounding secondary structure elements, as well as the zinc ribbon. **(e)** Topological representation of the GatD/MurT architecture. Secondary structure nomenclature of GatD was done according to Leisico *et al.*²⁴. As the short helices $\alpha 1$ and $\alpha 5$ in the isolated GatD structure do not conform to helical geometry in our complex, they were not assigned. The MurT domains were assigned separately with the prefixes *m* and *c* indicating the middle and C-terminal domains, respectively. The drawing was generated with TopDraw⁵⁴.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018



Unraveling the mechanism of peptidoglycan amidation by the bifunctional enzyme complex GatD/MurT: A comparative structural approach

Erik R. Nöldeke^a, Thilo Stehle^{a,b,*}

^a Interfaculty Institute of Biochemistry, University of Tübingen, D-72076 Tübingen, Germany

^b Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA

ARTICLE INFO

Keywords:

Crystal structure
bacterial cell wall protein

ABSTRACT

The bacterial cell wall provides structural integrity to the cell and protects the cell from internal pressure and the external environment. During the course of the twelve-year funding period of the Collaborative Research Center 766, our work has focused on conducting structure-function studies of enzymes that modify (synthesize or cleave) cell wall components of a range of bacteria including *Staphylococcus aureus*, *Staphylococcus epidermidis*, and *Nostoc punctiforme*. Several of our structures represent promising targets for interference. In this review, we highlight a recent structure-function analysis of an enzyme complex that is responsible for the amidation of Lipid II, a peptidoglycan precursor, in *S. aureus*.

1. Introduction

Ever since the discovery of penicillin in 1928 (Fleming, 1929) the bacterial cell wall and its assembly machinery have been a target for antibiotics (Schneider and Sahl, 2010). Now, nearly a century later, many bacteria have developed resistance mechanisms against a wide range of commonly used antimicrobial agents (Nikolaidis et al., 2014). Infections with multi-resistant bacterial strains such as methicillin-resistant *Staphylococcus aureus* (MRSA) can often cause severe disease and are extremely hard to treat (Foster, 2004). The situation is particularly critical in hospital environments, where spreading of such pathogens is facilitated by a high availability of susceptible hosts, for instance immunocompromised patients (CDC, 2013; Wang et al., 2017), leading to opportunistic infections. Hence, MRSA and other resistant bacteria are the cause of an increasing number of annual deaths and constitute an immense financial and logistical burden on healthcare worldwide (Chandy et al., 2014). In order to counteract this ever-rising threat, novel strategies to combat resistant bacteria are required.

The major structural component of the bacterial cell wall is the highly crosslinked peptidoglycan heteropolymer (Vollmer et al., 2008). It is composed of polysaccharide chains of alternating beta-1-4 linked N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) crosslinked by short, branched peptides that vary in composition depending on the species (Schleifer and Kandler, 1972).

Much of our initial work in the Collaborative Research Center 766 has focused on structure-function studies of enzymes that modify the

bacterial cell wall. We were able to define the structures and substrate specificities of several autolysins (Büttner et al., 2016, 2014; Lüttner et al., 2009; Zoll et al., 2012, 2010), which are needed by bacteria to lyse cell wall structures during cell division (see also review by Büttner et al., 2015). Additional studies have centered on enzymes that glycosylate teichoic acid structures in *staphylococci*, and that represent promising targets for interference (Gerlach et al., 2018; Koç et al., 2015). In recent years, we have become interested in a modification of a peptidoglycan precursor that involves amidation of peptide stem residues.

Peptidoglycan assembly is a multistep anabolic pathway starting with soluble cytosolic precursors and ending with the final heteropolymer network in the periplasmic space. Briefly, the ligases MurC–F (reviewed in Smith, 2006) catalyze the stepwise addition of the peptide stem to the UDP-activated MurNAc moiety. The resulting MurNAc-pentapeptide is then transferred to its membrane anchor undecaprenyl phosphate (bactoprenol) by MraY under UMP release (Chung et al., 2013; van Heijenoort, 2007). This first membrane-bound intermediate is termed Lipid I. Next, the second sugar, GlcNAc, is added by the transglycosylase MurG, yielding Lipid II (van Heijenoort, 2007). In some organisms, the amino acid at position 3 of the peptide stem (L-Lys or meso-diaminopimelic acid (mDAP) in most cases) is further modified by the addition of a short linker peptide, for instance Gly₅ as in *S. aureus* (Schleifer and Kandler, 1972). The final, modified Lipid II is then transferred to the periplasm, where transpeptidase and transglycosylase activities of penicillin binding proteins (PBPs) catalyze the polymerization and crosslinking into the final peptidoglycan layer (Sauvage

* Corresponding author.

E-mail address: thilo.stehle@uni-tuebingen.de (T. Stehle).

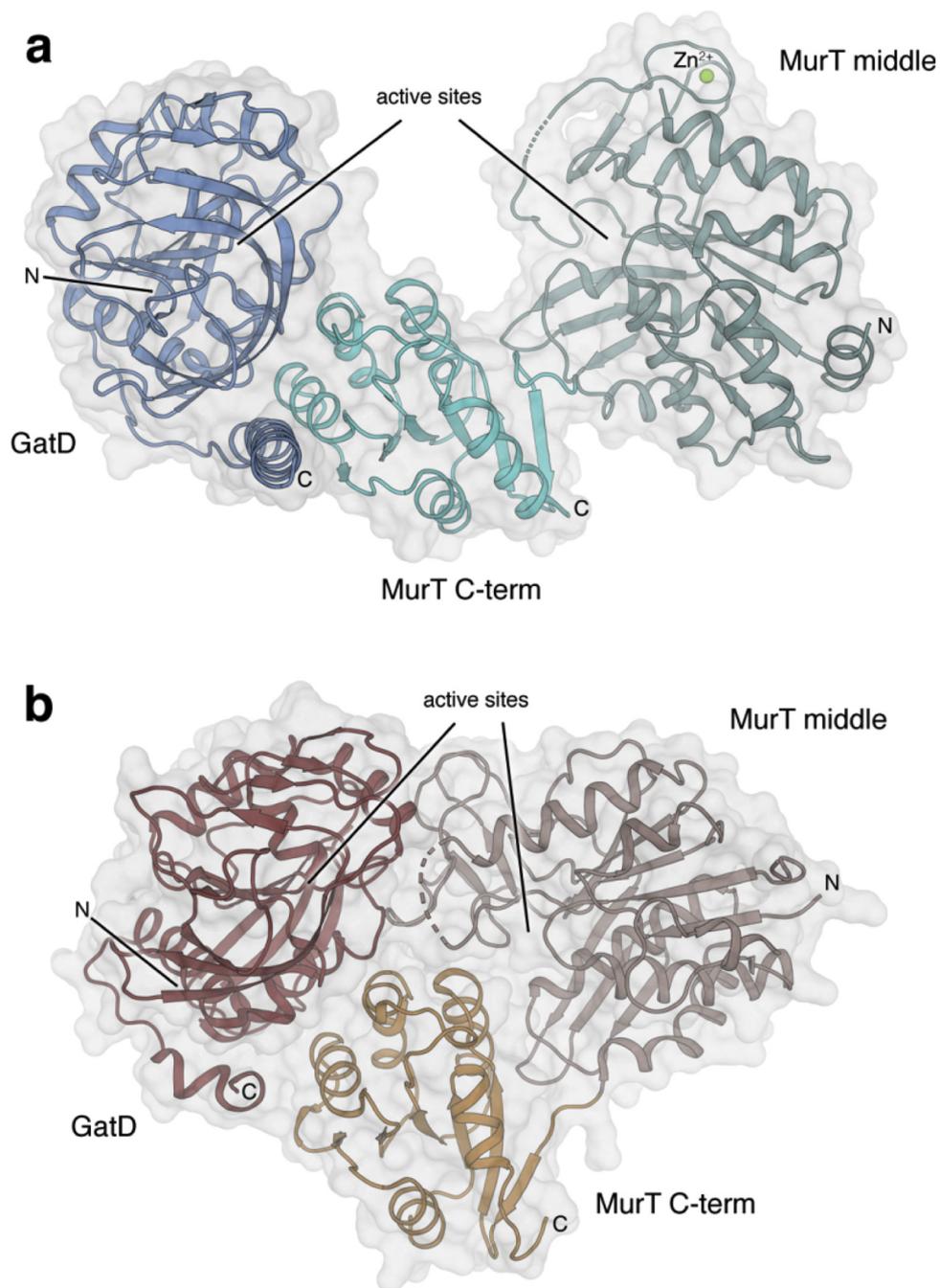


Fig. 1. Ribbon representations of the crystal structures of GatD/MurT. (a) GatD/MurT of *Staphylococcus aureus* displays an open, crescent-shaped conformation that likely represents the state adopted by the enzyme prior to substrate engagement (Nöldeke et al., 2018). (b) The homologous GatD/MurT complex of *Streptococcus pneumoniae* adopts a closed conformation (Morlot et al., 2018).

et al., 2008).

Peptidoglycan of many Gram positive bacteria, including *Staphylococci*, *Streptococci*, and even the distant genus of *Mycobacteria* show a wide variety of modifications in order to enhance its protective efficacy against antimicrobial agents such as lysozyme. The most frequent modifications are sugar moiety N-deacetylation or C₆-O-acetylation (Moynihan et al., 2014; Rajagopal and Walker, 2017), and the amidation of free carboxyl groups in the peptide. The most frequent site of amidation is the α carboxylate of D-glutamate at position 2 in the peptide stem. In peptidoglycan types containing mDAP at position 3, its ϵ -carboxylate can also be amidated (Schleifer and Kandler, 1972). Additional amidation sites exist, but they are restricted to very particular and rare peptidoglycan variations such as D-aspartate in the peptide

crossbridge present in some *Lactococci* (Veiga et al., 2009). All of these different amidation variants appear to contribute to resistance to lysozyme and other autolysins (Figueiredo et al., 2014; Levefaudes et al., 2015; Veiga et al., 2009). While the existence of peptide amidation has been known since the 1960s (Siewert and Strominger, 1968), only a few of the enzymes performing these modifications have been described to date, including LtsA and AsnB1 for mDAP amidation (Bernard et al., 2011; Levefaudes et al., 2015) and AsnH for D-aspartate amidation (Veiga et al., 2009). The most common amidation is carried out by GatD/MurT, which was identified only in 2012 (Figueiredo et al., 2012; Münch et al., 2012). GatD/MurT acts on the membrane-bound peptidoglycan precursor Lipid II and catalyzes the amidation of the second amino acid in the peptide stem from D-glutamic acid to D-isoglutamine.

The ATP-dependent reaction requires glutamine as a nitrogen donor. Interestingly, lack of peptidoglycan amidation as a result of GatD/MurT depletion drastically decreased methicillin tolerance in the context of MRSA (Figueiredo et al., 2014) and even proved lethal in the equally pathogenic *Streptococcus pneumoniae* (Liu et al., 2017; Zapun et al., 2013). These drastic depletion phenotypes have placed GatD/MurT in the spotlight as a potential drug target.

In order to understand the enzymatic mechanism and substrate specificity requirements, and to guide the design of effective inhibitor molecules, high-resolution structural information about the active complex is required. Recently, crystal structures of the GatD/MurT complexes of *Streptococcus pneumoniae* and *Staphylococcus aureus* have been published (Morlot et al., 2018; Nöldeke et al., 2018). This mini-review focuses on the comparison between the two structures and aims to ask specific questions raised by observed similarities or differences in order to help direct future research at them.

2. Structural information on GatD and MurT

Stem peptide amidation by GatD/MurT is known to be crucial for efficient peptidoglycan biosynthesis in a number of Gram positive bacteria, and its essentiality has been inferred for various members of the *Bacilli* family as well as a number of more loosely related *Actinobacteria* (DeJesus et al., 2017; Figueiredo et al., 2012; Münch et al., 2012; Zapun et al., 2013). A homologous operon was found even in the cyanobacterium *Nostoc punctiforme* (Münch et al., 2012). However, mechanistic insights into the enzyme's mode of action are still very limited. Detailed structural information is available only for the complexes from two organisms of the *Bacilli* class, *Staphylococcus aureus* and *Streptococcus pneumoniae*. The structure of isolated, unliganded GatD (Leisico et al., 2018) and two subsequently determined structures of GatD/MurT complexes (Morlot et al., 2018; Nöldeke et al., 2018) have established the fold of the enzymes and identified residues important for ligand binding and catalysis. However, none of the crystal structures contains the full set of substrates (glutamine, ATP, and the peptide stem of Lipid I/II), and the two GatD/MurT complexes also exhibit interesting, profound differences in overall structure and conformation, in the assignment of catalytic residues, and in a small domain that harbors a zinc ion in one case. Therefore, reviewing and comparing the essential structural features of the GatD/MurT complexes provides insights into how they might function in Lipid II amidation.

3. Overall structure of GatD/MurT

In both organisms the two enzymes assemble into a heterodimer, with GatD contacting the C-terminal domain of MurT (Fig. 1). GatD exhibits a mixed alpha/beta hydrolase fold (Ollis et al., 1992) typical for class I glutamine amidotransferases (GATases) (Massière and Badet-Denisot, 1998) and the GatD active site, centered around the catalytic cysteine, faces towards MurT. MurT is composed of two domains, namely a Mur ligase catalytic or middle domain (Mur middle) and a Mur ligase C-terminal domain (Mur C-term). Like in other Mur ligases, MurT contains the canonical ATP binding pocket and catalytic residues required for ATPase activity (reviewed in (Smith, 2006)). Other Mur ligases have been described to use their C-terminal domain to bind part of their respective substrates (Ruane et al., 2013; Smith, 2006), hence it could be inferred that MurT likely employs residues from both domains to coordinate the substrate peptide. However, detailed experimental evidence for the mode of substrate binding is still not available.

While the two available crystal structures of GatD/MurT from *S. aureus* and *S. pneumoniae* adopt largely different conformations overall, the single domains are similar and superimpose very well (Table 1). Hence, a detailed analysis of the similarities and differences between the two structures will help to identify yet undiscovered important features and residues as well as potential organism-specific differences.

Table 1

Root mean square deviations (R.M.S.D.s) of superimpositions of *S. aureus* and *S. pneumoniae* GatD/MurT single domains.

	R.M.S.D. [Å]	aligned residues
GatD	2.14	224
MurT middle	3.21	232
MurT C-term	1.90	128

Such information will paint a more complete picture of the mode of action of this enzyme complex, thus providing a solid foundation for future research aimed at unravelling the reaction mechanism and targeting the active site.

4. Conformational dynamics

The canonical Mur ligases MurC, MurD, MurE, and MurF, which catalyze the formation of the peptidoglycan peptide stem in the cytosolic steps of biosynthesis, have been extensively characterized and crystal structures are available for representatives of all four enzymes (Bertrand et al., 1999; Cha et al., 2014; Mol et al., 2003; Ruane et al., 2013). The middle and C-terminal domains of the Mur ligases are connected via a short flexible linker. Mur ligases C, D and F, for which both unliganded and substrate-bound structures are available, display a large degree in domain movement upon ligand binding (reviewed in (Smith, 2006)). The strong structural similarities in this region between MurT and the other Mur ligases suggest that such a high degree of mobility would indeed be possible. In support of this hypothesis, the conformations observed in the crystal structures of GatD/MurT from *S. aureus* (SaGatD/MurT) and *S. pneumoniae* (SpGatD/MurT) differ substantially in terms of domain orientation. While SpGatD/MurT adopts a closed conformation similar to that reported for many ligand-bound structures of other Mur ligases, the *S. aureus* structure exhibits a much more open, crescent-shaped conformation that is closer to the apo forms of the aforementioned enzymes. Small angle X-ray scattering experiments performed on SaGatD/MurT show that a similarly open conformation of the enzyme also exists in solution, and indicate that the complex has a considerable amount of flexibility among its domains (Nöldeke et al., 2018).

Based on these observations, it is highly likely that the different conformations adopted by *S. aureus* and *S. pneumoniae* GatD/MurT in their respective crystal structures are not indicative of radical cross-species differences but rather represent two snapshots of a range of possible conformations that can be adopted by both complexes during catalysis. The open conformation observed for SaGatD/MurT likely represents the structure of the complex prior to substrate engagement. As the catalytically active sites of GatD and MurT are separated by about 40 Å in the open conformation, any movement that decreases this distance would seem to be required for catalysis.

Although the evidence available to date clearly points towards a domain rearrangement being a critical step to achieve an active, catalysis-competent complex, the detailed mechanics of catalysis remain elusive. Even the reduced distance between active sites in the closed conformation of around 20 Å is too large to be bridged by nascent ammonia generated in GatD without significant effusive loss. Such a loss would not only hamper MurT catalysis, but if enough ammonia were to escape this could even prove toxic to the cell. Hence, additional domain movements may be required for full and efficient activity.

Is an as of yet undetected ammonia channel formed upon ligand binding to successfully perform the amidation reaction? Does catalysis involve substrate relocation after phosphorylation by MurT to the site of ammonia generation in GatD? High-resolution data on substrate peptide recognition are required in order to answer these questions and to paint a complete picture of GatD/MurT action.

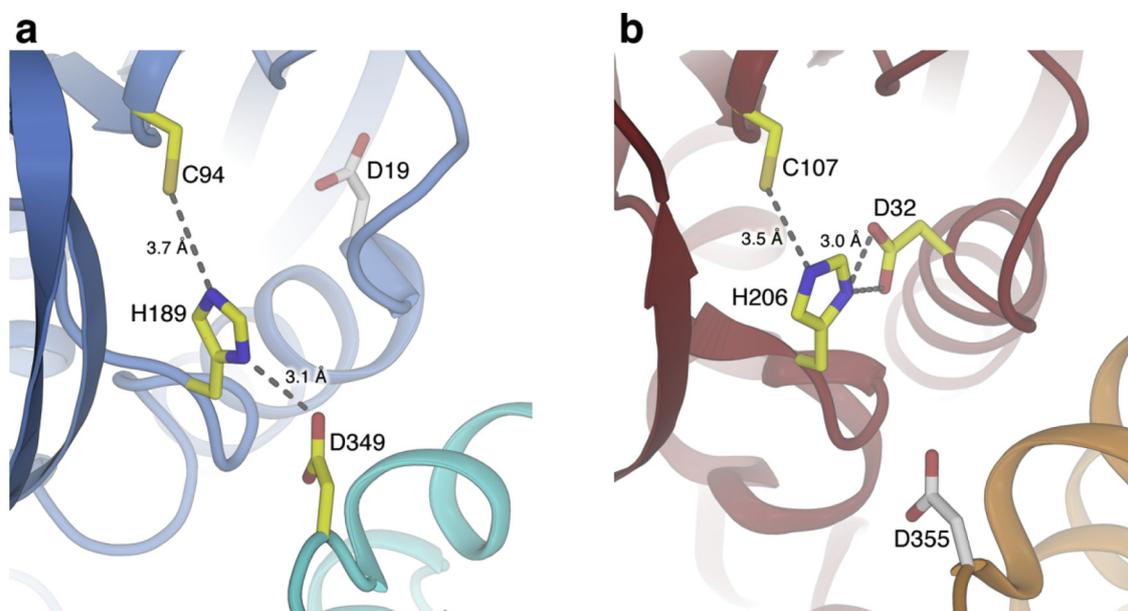


Fig. 2. Comparison of the putative catalytic triads of the GatD/MurT complexes in *S. aureus* and *S. pneumoniae*. (a) Cartoon and stick representation of the intermolecular catalytic triad of SaGatD/MurT. D349 is contributed to the triad by MurT across the dimer interface. (b) The intramolecular triad of SpGatD in the same orientation as in (a). White sticks represent the aspartates homologous to the third triad residue in the respective opposite species. (Morlot et al., 2018; Nöldeke et al., 2018).

5. GatD catalytic triad

Class I or triad type GATases contain a canonical Cys-His-Glu catalytic triad as a means to generate a nucleophile strong enough to initiate glutamine hydrolysis to glutamate (Massière and Badet-Denisot, 1998; Mouilleron and Golinelli-Pimpaneau, 2007). The central nucleophilic cysteine is located at a conserved turn from a β -strand to an α -helix, termed the nucleophile elbow (Ollis et al., 1992). The histidine and glutamate that complete the triad are located on a separate strand, and they are usually separated by a proline. The catalytic triad residues are located nearly in a straight line in order to allow for concerted proton transfer.

While the cysteine and histidine residues are conserved in GatD, the glutamate as the final proton sink is missing. Two alternative aspartic acid side chains have been proposed to substitute for it in that function, namely SaMurT-D349 (SpMurT-D355) and SpGatD-D32 (SaGatD-D19) (Fig. 2). In the first case, observed in the *S. aureus* enzyme, the aspartic acid is contributed by MurT across the heterodimer interface (Fig. 2a). The intermolecular triad generated in this manner adopts a linear conformation, which would be amenable to concerted proton transfer. In the second case, as found in the structure of *S. pneumoniae* GatD/MurT, the proposed triad is intramolecular, with the aspartic acid being contributed from a flanking helix in GatD (Fig. 2b). While its interatomic distances are smaller than those observed in SaGatD/MurT, the angle of approximately 90° formed by the triad is unusual for such motifs and probably requires slight repositioning of the sidechains involved in order to function in a proton relay reaction.

Intriguingly, both aspartates are conserved in all confirmed and inferred instances of GatD/MurT (Morlot et al., 2018; Nöldeke et al., 2018). However, their relative positions differ largely between the staphylococcal and streptococcal structures. Residue D19 of SaGatD (Fig. 2a) is turned away from the triad site by a twist of helix $\alpha 2$ (secondary structure nomenclature according to (Nöldeke et al., 2018)) when compared to SpGatD, thus precluding it from taking the same place and role as SpGatD-D32. On the other hand, SpGatD/MurT cannot form a linear triad with SpMurT-D355, as it is separated from either of the two histidine nitrogens by over 5 \AA . It however still plays an important role in SpGatD/MurT function, possibly in complex stability, as

its mutation severely impaired activity *in vitro* and even proved lethal *in vivo* (Morlot et al., 2018).

The putative catalytic triads represent a fascinating example of how two different organisms apparently found separate ways to the same end, using exactly the same available tools in terms of conserved residues. It remains to be investigated to which extent the triad conformation may be influenced by the overall conformation of the complex. Thus, a different arrangement of SpGatD/MurT may bring the residues for the intermolecular triad into closer proximity, while a movement in the *S. aureus* complex may lead to a repositioning of SaGatD-D19. Hence, the GatD catalytic triad should remain a focus of research. Structures of the complex from other species may help to fully understand the roles played by the proposed triad residues.

6. The MurT middle domain insertion

One unexpected feature that was revealed by the recent structural studies conducted on the GatD/MurT enzyme complexes from *S. aureus* and *S. pneumoniae* is an insertion of roughly 50 residues between the strands $\beta m6$ and $\beta m9$ in the middle domain of MurT. This insertion, which to date has not been observed in any other known Mur ligases, bears a strong resemblance to a type of Zinc Finger (ZnF) termed RanBP Zinc ribbon (Gamsjaeger et al., 2007). ZnF motifs exist in various shapes but share the common feature of a Zn^{2+} ion coordinated by four cysteine or histidine sidechains contributed by different portions of the motif, thus conferring stability to the motif (Laitý et al., 2001). Most canonical ZnFs are involved in nucleic acid interactions, but several examples of such motifs mediating protein-protein contacts have been reported (Gamsjaeger et al., 2007; Malgieri et al., 2015).

The Zinc ribbon insertion observed in GatD-MurT is well conserved in Bacilli (Fig. 3c) and adopts very similar structures in SaMurT and SpMurT. It is composed of two adjacent turns of the protein backbone that contain two cysteines each with one or two linking residues in between. These turns, termed "knuckles", form the beginning and end of a two-stranded antiparallel β -sheet. The cysteines from the knuckles face one another to coordinate a zinc ion that further stabilizes the fold in SaMurT. In the crystal structure of SpMurT, an array of disulfides was modelled instead of the zinc ion. Biochemical analyses however

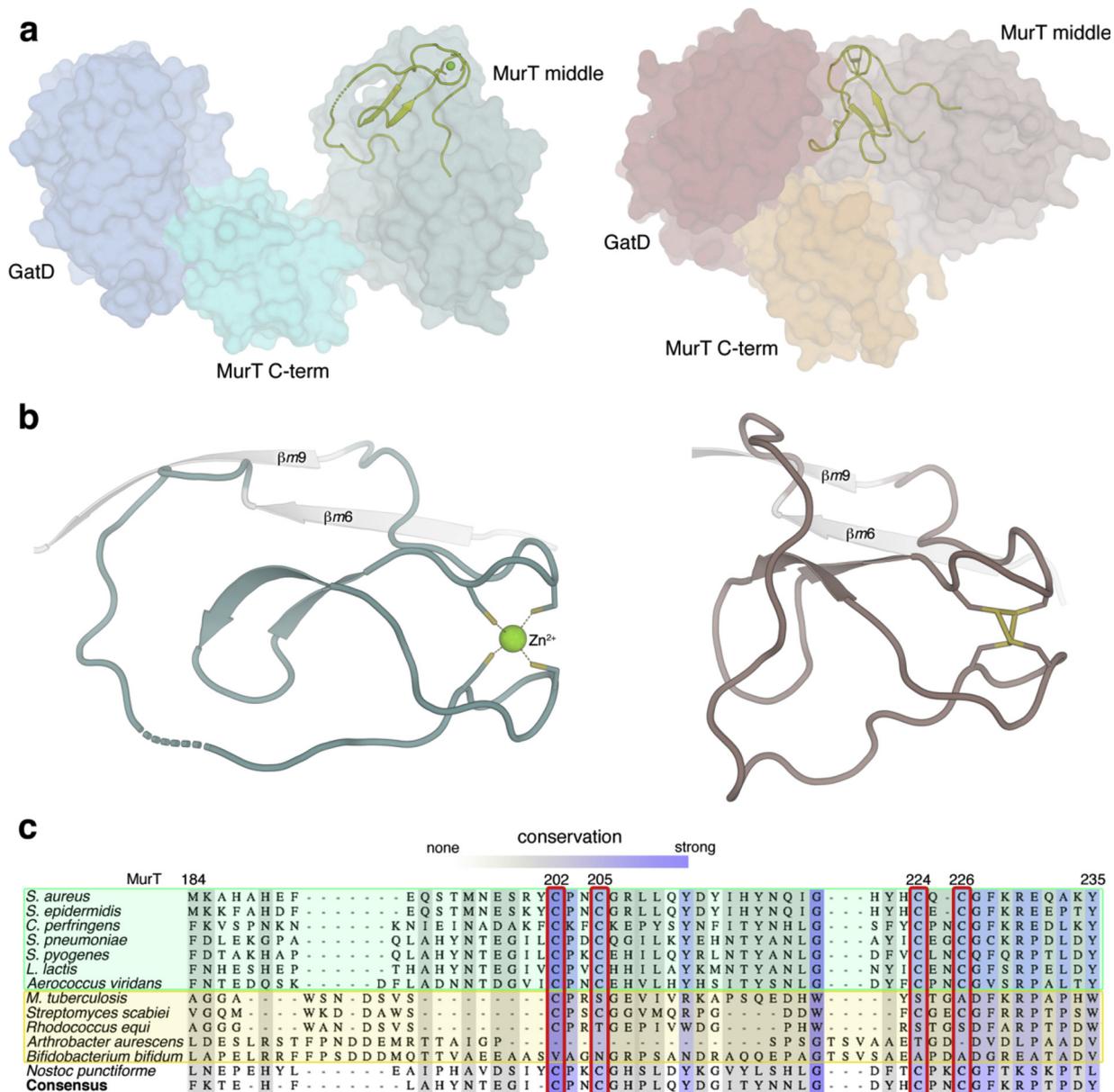


Fig. 3. The MurT insertion: a Zinc Finger. (a) Side by side comparison of the *S. aureus* (left) and *S. pneumoniae* (right) crystal structures (Morlot et al., 2018; Nöldeke et al., 2018). The ZnF is shown as ribbon. (b) Detailed topology of the ZnF highlighting the conserved conformation of the central two-stranded β -sheet and the cysteine-containing knuckle. The large leading loop adopts a largely different conformation due to the contact with GatD, as seen in (a). (c) Multiple sequence alignment of the insert region of MurT from different species. *Bacilli* (top, green box) share a relatively similar sequence with completely conserved knuckle cysteines (red). In contrast, *Actinobacteria* (yellow box) show a higher interspecies variability in the region.

indicated that the cysteines are initially reduced and the physiological presence of the coordinated metal ion was inferred (Morlot et al., 2018). The segment is located on the side of the MurT middle domain (Fig. 3a), behind the ATP binding site and putative substrate peptide binding region. It forms an extensive contact surface between SpMurT and SpGatD, resulting in a compression of the loop connecting MurT strand $\beta m6$ and the first Cys-knuckle (Fig. 3b). In contrast, the same loop is more relaxed and partially flexible in the open form of SaMurT, which has no direct interactions between its middle domain and SaGatD.

Amino acid sequence comparison of known and putative instances of MurT from different species reveal that conservation of the ZnF is much lower than that of several functionally relevant features. While it probably exists in topologically similar forms across *Bacilli* based on the locations of the cysteines, actinobacterial sequences differ from this pattern (Fig. 3c). *Streptomyces* (represented here by *S. scabiei*) have a

severely shortened insertion while still retaining the characteristic cysteines. However, it cannot be predicted whether this is sufficient to form the characteristic fold. In contrast, most other *Actinobacteria* with GatD/MurT homologues lack the cysteines altogether, and therefore the insertion is unlikely to adopt a zinc-Finger fold.

Together with the lack of detailed knowledge about the precise function of the MurT insertion, this observation raises questions about functional diversity. Is the MurT insertion involved in GatD/MurT activity? If so, have different species evolved different mechanisms to achieve the same goal while relying more or less heavily on a particular sequence, in this case the ZnF, or is catalytic activity completely uncoupled from the nature of the insertion?

RanBP-type Zn ribbons have been associated with protein-protein interactions in the past, such as the complex between Npl4 and Ubiquitin (Alam et al., 2004). The location and conformation of the bacillal ZnF leaves it accessible for an analogous interaction. It is thus

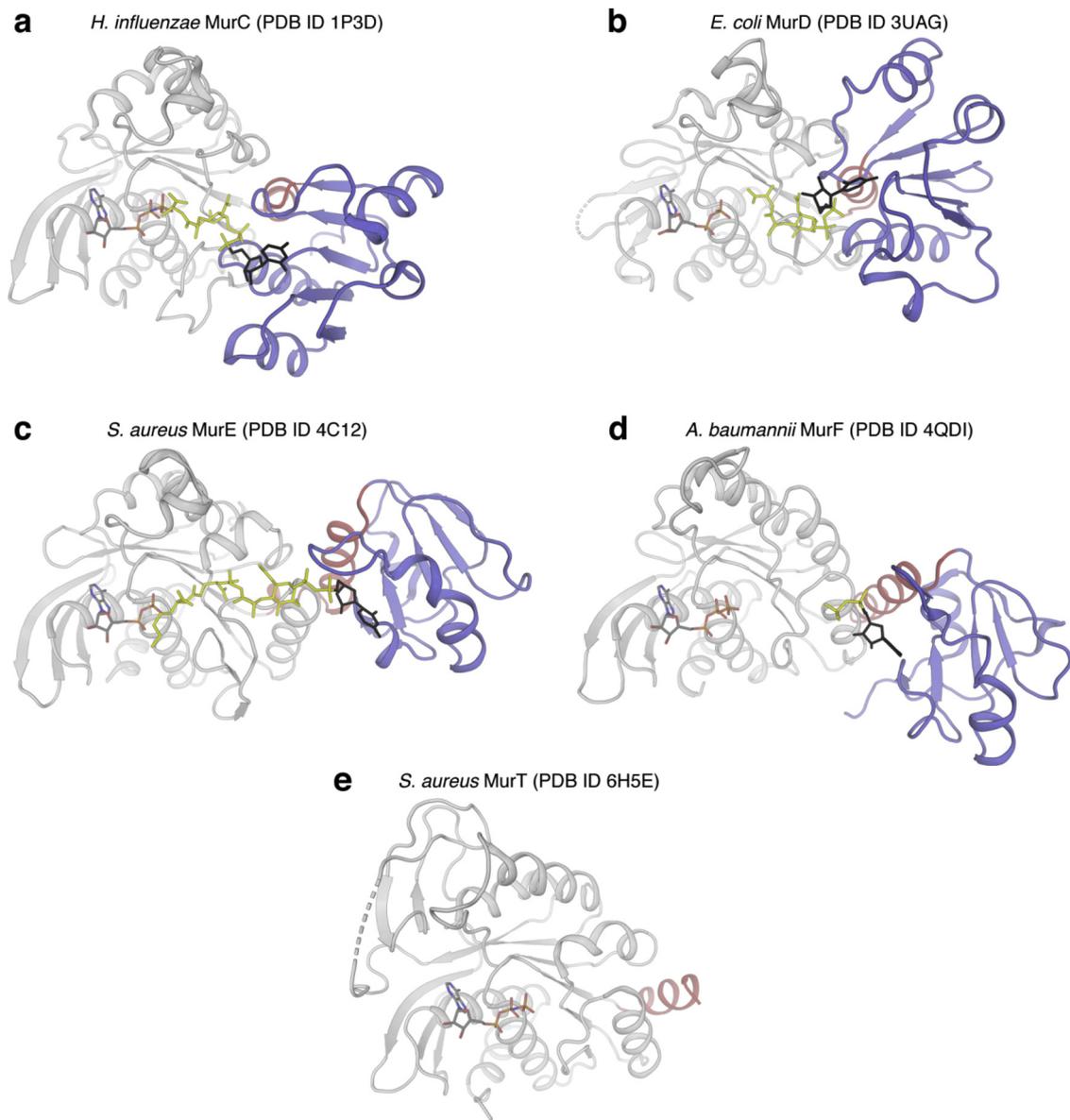


Fig. 4. The N-termini of different Mur ligases. A side by side comparison of the N-terminal and middle domains of MurC-D (Bertrand et al., 1999; Cha et al., 2014; Mol et al., 2003; Ruane et al., 2013) and MurT (Nöldeke et al., 2018) (Panels (a–d) and (e), respectively), aligned for their middle domains, highlight the differences in N-terminus orientation. The N-terminal domain (slate) coordinates the uridine moiety of the substrate (black sticks) thus positioning the substrate peptides of different lengths with respect to the invariable position of ATP in the middle domain (grey cartoon and sticks). The orientation of the N-terminal domain is partially guided by the orientation of a linking helix homologous to MurT helix $\beta m1$ (red).

conceivable that the ZnF may be involved in a scaffolding interaction with other components of the membrane-bound phase of peptidoglycan biosynthesis. This could for instance help to bring GatD/MurT into close proximity of its substrate.

7. The short MurT N-terminus

S. aureus GatD/MurT amidates peptidoglycan precursors exclusively in the membrane-bound steps of the biosynthetic pathway (Figueiredo et al., 2012) but shows no preference for either Lipid I, Lipid II or Lipid II - Gly₅ *in vitro* (Münch et al., 2012). The sole difference between the last soluble precursor, UDP-MurNAc-pentapeptide, and Lipid I lies in the exchange of a uridine moiety for a bactoprenol membrane anchor. As the peptide and MurNAc segments are unaltered, the preference of GatD/MurT for this substrate form has to be associated with the membrane anchor. However, neither GatD nor MurT contain exposed hydrophobic segments that could mediate membrane insertion and thus

recognition of the isoprenoid tail of the substrate. The N-terminus of MurT, however, has not been resolved in either of the two crystal structures available to date. The unresolved portion spans 30–35 amino acids and contains a high portion of positively charged side chains, totaling 9 and 6 net positive charges for *S. aureus* and *S. pneumoniae*, respectively.

In contrast, the canonical Mur ligases C–F have an additional N-terminal domain of 90–100 amino acids in length. Typically, the N-terminal domains of these Mur ligases adopt a canonical Rossmann fold (Smith, 2006) and coordinate the uridine moiety of the UDP-activated MurNAc-peptide precursor (Fig. 4). Interestingly, the contacts between the Mur middle domain and the growing peptide stem are not particularly extensive. The orientation of the peptide with respect to the unvarying position of ATP in the middle domain is largely controlled by the orientation of the N-terminal domain and, consequently, of the UDP moiety (Smith, 2006). This positioning appears to be mediated at least in part by the orientation of a hinge helix between the N-terminal and

middle domains, homologous to helix $\beta m1$ in MurT.

The different nature of the substrate and lack of a comparable N-terminal domain call for a different solution for substrate orientation in MurT. As the orientation of the substrate is not chiefly mediated by the middle and C-terminal domains in other Mur ligases, it can be speculated that MurT may harness external geometrical restraints in order to contact, recognize, and position its substrate peptide for catalysis. The high net positive charge of the unresolved MurT N-terminus suggests such an approach, as it could interact with the negative charges of the phospholipids in the plasma membrane, thus bringing GatD/MurT into the proximity of its substrate. Indeed, Münch et al. showed that *S. aureus* membrane but not cytosolic extracts contain the enzymatic activity required for Lipid II amidation *in vitro* in an ATP and glutamine-dependent fashion, strongly suggesting GatD/MurT to be associated with the plasma membrane (Münch et al., 2012). MurT helix $\beta m1$ points in a direction that could orient the enzyme with its active sites facing the membrane. Additionally, cationic terminal stretches have been associated in the past with peripheral membrane association (Whited and Johs, 2015). In such a hypothetical model, spatial sequestration and orientation relative to the membrane rather than direct interactions with the substrate would provide specificity. This hypothesis could provide an explanation for the observed exclusivity in GatD/MurT targeting peptidoglycan precursors at the lipidic stage of biosynthesis.

These hypotheses need to be thoroughly investigated. No experimental evidence is available to date on the role of the MurT N-terminus or the enzyme's subcellular localization.

8. Conclusion and applications

GatD/MurT is highly relevant to worldwide healthcare as a potential drug target to combat infections by multi-resistant bacterial pathogens. By comparing the available GatD/MurT crystal structures of the enzymes of *S. aureus* and *S. pneumoniae*, insights into their mode of action beyond observations on only one protein can be derived. Since the two structures differ in critical ways, the comparison identifies a number of questions that need to be addressed through follow-up studies in order to establish GatD/MurT as a viable platform for the development of structure-guided inhibitors with medical application.

Differences in GatD/MurT across species, such as the MurT middle domain insertion, need to be functionally characterized in order to evaluate their potential as targets for the design of narrow spectrum antimicrobials. Such compounds with a precisely defined target range are more difficult to use due to the necessity to identify the pathogen prior to treatment, but they offer the advantage of a lower impact on beneficial commensal bacteria such as *E. coli*. The observed similar features of the two enzymes, on the other hand, could serve as viable targets for the design of antibiotics with a broader range of action against a wide range of pathogenic bacteria such as *S. aureus*, *S. pneumoniae*, and *M. tuberculosis* while still maintaining its specificity for organisms that express GatD/MurT.

Finally, differences in the crystal structures that cannot be definitively attributed to the difference in species help to paint a detailed picture of the still only rudimentarily understood mechanism of action of GatD/MurT. In this instance, the differences in conformation provide evidence for a major domain rearrangement as an essential step for catalytic activity. Clearly, additional structural and functional studies directed at the mode of substrate engagement by GatD/MurT are needed.

Acknowledgments

We thank members of the Stehle laboratory for discussions. The authors gratefully acknowledge support from the German Research Foundation, Bonn, Germany (Collaborative Research Centers 766 and TR261).

References

- Alam, S.L., Sun, J., Payne, M., Welch, B.D., Blake, B.K., Davis, D.R., Meyer, H.H., Emr, S.D., Sundquist, W.L., 2004. Ubiquitin interactions of NZF zinc fingers. *EMBO J.* 23, 1411–1421. <https://doi.org/10.1038/sj.emboj.7600114>.
- Bernard, E., Rolain, T., Courtin, P., Hols, P., Chapot-Chartier, M.-P., 2011. Identification of the amidotransferase AsnB1 as being responsible for meso-diaminopimelic acid amidation in *Lactobacillus plantarum* peptidoglycan. *J. Bacteriol.* 193, 6323–6330. <https://doi.org/10.1128/JB.05060-11>.
- Bertrand, J.A., Auger, G., Martin, L., Fanchon, E., Blanot, D., Le Beller, D., Van Heijenoort, J., Dideberg, O., 1999. Determination of the MurD mechanism through crystallographic analysis of enzyme complexes. *J. Mol. Biol.* 289, 579–590. <https://doi.org/10.1006/jmbi.1999.2800>.
- Büttner, F.M., Faulhaber, K., Forchhammer, K., Maldener, I., Stehle, T., 2016. Enabling cell-cell communication via nanopore formation: structure, function and localization of the unique cell wall amidase AmiC2 of *Nostoc punctiforme*. *FEBS J.* 283, 1336–1350. <https://doi.org/10.1111/febs.13673>.
- Büttner, F.M., Renner-Schneck, M., Stehle, T., 2015. X-ray crystallography and its impact on understanding bacterial cell wall remodeling processes. *Int. J. Med. Microbiol.* 305, 209–216. <https://doi.org/10.1016/j.ijmm.2014.12.018>.
- Büttner, F.M., Zoll, S., Nega, M., Götz, F., Stehle, T., 2014. Structure-function analysis of *Staphylococcus aureus* amidase reveals the determinants of peptidoglycan recognition and cleavage. *J. Biol. Chem.* 289, 11083–11094. <https://doi.org/10.1074/jbc.M114.557306>.
- CDC, 2013. Antibiotic Resistance Threats in the United States [WWW Document]. Centers Dis. Control Prev. Website. URL <http://www.cdc.gov/drugresistance/threat-report-2013> (accessed 12.19.18).
- Cha, S.S., An, Y.J., Jeong, C.S., Yu, J.H., Chung, K.M., 2014. ATP-binding mode including a carbamoylated lysine and two Mg²⁺ ions, and substrate-binding mode in *Acinetobacter baumannii* MurF. *Biochem. Biophys. Res. Commun.* 450, 1045–1050. <https://doi.org/10.1016/j.bbrc.2014.06.108>.
- Chandy, S.J., Naik, G.S., Balaji, V., Jayaseelan, V., Thomas, K., Lundborg, C.S., 2014. High cost burden and health consequences of antibiotic resistance: the price to pay. *J. Infect. Dev.* 8, 1096–1102. <https://doi.org/10.3855/jidc.4745>.
- Chung, B.C., Zhao, J., Gillespie, R.A., Kwon, D.-Y., Guan, Z., Hong, J., Zhou, P., Lee, S.-Y., 2013. Crystal structure of MraY, an essential membrane enzyme for bacterial cell wall synthesis. *Science* 341 (80), 1012–1016. <https://doi.org/10.1126/science.1236501>.
- DeJesus, M.A., Gerrick, E.R., Xu, W., Park, S.W., Long, J.E., Boutte, C.C., Rubin, E.J., Schnappinger, D., Ehart, S., Fortune, S.M., Sasseti, C.M., Iroeger, T.R., 2017. Comprehensive essentiality analysis of the *Mycobacterium tuberculosis* genome via saturating transposon mutagenesis. *MBio* 8, 1–17. <https://doi.org/10.1128/mBio.02133-16>.
- Figueiredo, Ta, Ludovice, A.M., Sobral, R.G., 2014. Contribution of peptidoglycan amidation to beta-lactam and lysozyme resistance in different genetic lineages of *Staphylococcus aureus*. *Microb. Drug Resist.* 20, 1–12. <https://doi.org/10.1089/mdr.2014.0042>.
- Figueiredo, T.A., Sobral, R.G., Ludovice, A.M., de Almeida, J.M.F., Bui, N.K., Vollmer, W., de Lencastre, H., Tomasz, A., 2012. Identification of genetic determinants and enzymes involved with the amidation of glutamic acid residues in the peptidoglycan of *Staphylococcus aureus*. *PLoS Pathog.* 8. <https://doi.org/10.1371/journal.ppat.1002508>.
- Fleming, A., 1929. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *Br. J. Exp. Pathol.* 10, 226–236.
- Foster, T.J., 2004. The *Staphylococcus aureus* “superbug”. *J. Clin. Invest.* 114, 1693–1696. <https://doi.org/10.1172/JCI2004221237>.
- Gamsjaeger, R., Liew, C.K., Loughlin, F.E., Crossley, M., Mackay, J.P., 2007. Sticky fingers: zinc-fingers as protein-recognition motifs. *Trends Biochem. Sci.* 32, 63–70. <https://doi.org/10.1016/j.tibs.2006.12.007>.
- Gerlach, D., Guo, Y., De Castro, C., Kim, S.-H., Schlatterer, K., Xu, F.-F., Pearsa, C., Seeberger, P.H., Ali, S., Codée, J., Sirisarn, W., Schulte, B., Wolz, C., Larsen, J., Molinaro, A., Lee, B.L., Xia, G., Stehle, T., Peschel, A., 2018. Methicillin-resistant *Staphylococcus aureus* alters cell wall glycosylation to evade immunity. *Nature* 563, 705–709. <https://doi.org/10.1038/s41586-018-0730-x>.
- Koç, C., Gerlach, D., Beck, S., Peschel, A., Xia, G., Stehle, T., 2015. Structural and enzymatic analysis of TarM glycosyltransferase from *Staphylococcus aureus* reveals an oligomeric protein specific for the glycosylation of wall teichoic acid. *J. Biol. Chem.* 290, 9874–9885. <https://doi.org/10.1074/jbc.M114.619924>.
- Laity, J.H., Lee, B.M., Wright, P.E., 2001. Zinc finger proteins: new insights into structural and functional diversity. *Curr. Opin. Struct. Biol.* 11, 39–46. [https://doi.org/10.1016/S0959-440X\(00\)00167-6](https://doi.org/10.1016/S0959-440X(00)00167-6).
- Leisico, F., Vieira, D.V., Figueiredo, T.A., Silva, M., Cabrita, E.J., Sobral, R.G., Ludovice, A.M., Trincão, J., Romão, M.J., De Lencastre, H., Santos-Silva, T., 2018. First insights of peptidoglycan amidation in Gram-positive bacteria—The high-resolution crystal structure of *Staphylococcus aureus* glutamine amidotransferase GatD. *Sci. Rep.* 8, 1–13. <https://doi.org/10.1038/s41598-018-22986-3>.
- Levefaudes, M., Patin, D., de Sousa-d'Auria, C., Chami, M., Blanot, D., Hervé, M., Arthur, M., Houssin, C., Mengin-Lecreux, D., 2015. Diaminopimelic acid amidation in corynebacteriales: NEW INSIGHTS INTO THE ROLE OF LtaA IN PEPTIDOGLYCAN MODIFICATION. *J. Biol. Chem.* 290, 13079–13094. <https://doi.org/10.1074/jbc.M115.642843>.
- Liu, X., Gallay, C., Kjos, M., Domenech, A., Slager, J., van Kessel, S.P., Knoops, K., Sorg, R.A., Zhang, J., Veening, J., 2017. High-throughput CRISPRi phenotyping identifies new essential genes in *Streptococcus pneumoniae*. *Mol. Syst. Biol.* 13, 931. <https://doi.org/10.15252/msb.20167449>.
- Lützner, N., Pätzold, B., Zoll, S., Stehle, T., Kalbacher, H., 2009. Development of a novel

- fluorescent substrate for Autolysin E, a bacterial type II amidase. *Biochem. Biophys. Res. Commun.* 380, 554–558. <https://doi.org/10.1016/j.bbrc.2009.01.140>.
- Malgieri, G., Palmieri, M., Russo, L., Fattorusso, R., Pedone, P.V., Isernia, C., 2015. The prokaryotic zinc-finger: structure, function and comparison with the eukaryotic counterpart. *FEBS J.* 282, 4480–4496. <https://doi.org/10.1111/febs.13503>.
- Massière, F., Badet-Denisot, M.A., 1998. The mechanism of glutamine-dependent amidotransferases. *Cell. Mol. Life Sci.* 54, 205–222. <https://doi.org/10.1007/s000180050145>.
- Mol, C.D., Brooun, A., Dougan, D.R., Hilgers, M.T., Tari, L.W., Wijnands, R.A., Knuth, M.W., McRee, D.E., Swanson, R.V., 2003. Crystal structures of active fully assembled substrate- and product-bound complexes of UDP-N-acetylmuramic acid:L-alanine ligase (MurC) from *Haemophilus influenzae*. *J. Bacteriol.* 185, 4152–4162. <https://doi.org/10.1128/JB.185.14.4152-4162.2003>.
- Morlot, C., Straume, D., Peters, K., Hegnar, O.A., Simon, N., Villard, A.M., Contreras-Martel, C., Leisico, F., Breukink, E., Gravier-Pelletier, C., Le Corre, L., Vollmer, W., Pietrancosta, N., Hävarstein, L.S., Zapun, A., 2018. Structure of the essential peptidoglycan amidotransferase MurT/GatD complex from *Streptococcus pneumoniae*. *Nat. Commun.* 9, 1–12. <https://doi.org/10.1038/s41467-018-05602-w>.
- Mouilleron, S., Golinelli-Pimpaneau, B., 2007. Conformational changes in ammonia-channeling glutamine amidotransferases. *Curr. Opin. Struct. Biol.* 17, 653–664. <https://doi.org/10.1016/j.sbi.2007.09.003>.
- Moynihan, P.J., Sychantha, D., Clarke, A.J., 2014. Chemical biology of peptidoglycan acetylation and deacetylation. *Bioorg. Chem.* 54, 44–50. <https://doi.org/10.1016/j.bioorg.2014.03.010>.
- Münch, D., Roemer, T., Lee, S.H., Engeser, M., Sahl, H.G., Schneider, T., 2012. Identification and in vitro analysis of the GatD/MurT enzyme-complex catalyzing lipid II amidation in *Staphylococcus aureus*. *PLoS Pathog.* 8. <https://doi.org/10.1371/journal.ppat.1002509>.
- Nikolaïdis, I., Favini-Stabile, S., Dessen, A., 2014. Resistance to antibiotics targeted to the bacterial cell wall. *Protein Sci.* 23, 243–259. <https://doi.org/10.1002/pro.2414>.
- Nöldeke, E.R., Muckenfuss, L.M., Niemann, V., Müller, A., Störk, E., Zocher, G., Schneider, T., Stehle, T., 2018. Structural basis of cell wall peptidoglycan amidation by the GatD/MurT complex of *Staphylococcus aureus*. *Sci. Rep.* 8, 12953. <https://doi.org/10.1038/s41598-018-31098-x>.
- Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., 1992. The alpha/beta hydrolase fold. *Protein Eng.* 5, 197–211. <https://doi.org/10.1093/protein/5.3.197>.
- Rajagopal, M., Walker, S., 2017. Envelope structures of gram-positive Bacteria. *Curr. Top. Microbiol. Immunol.* 404, 1–44. https://doi.org/10.1007/82_2015_5021.
- Ruane, K.M., Lloyd, A.J., Fülöp, V., Dowson, C.G., Barreteau, H., Boniface, A., Dementin, S., Blanot, D., Mengin-Lecreulx, D., Gobec, S., Dessen, A., Roper, D.I., 2013. Specificity determinants for lysine incorporation in *Staphylococcus aureus* peptidoglycan as revealed by the structure of a MurE enzyme ternary complex. *J. Biol. Chem.* 288, 33439–33448. <https://doi.org/10.1074/jbc.M113.508135>.
- Sauvage, E., Kerff, F., Terrak, M., Ayala, J.A., Charlier, P., 2008. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol. Rev.* 32, 234–258. <https://doi.org/10.1111/j.1574-6976.2008.00105.x>.
- Schleifer, K.H., Kandler, O., 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol. Rev.* 36, 407–477.
- Schneider, T., Sahl, H.G., 2010. An oldie but a goodie - cell wall biosynthesis as antibiotic target pathway. *Int. J. Med. Microbiol.* 300, 161–169. <https://doi.org/10.1016/j.ijmm.2009.10.005>.
- Siewert, G., Strominger, J.L., 1968. Biosynthesis of the peptidoglycan of bacterial cell walls. XI. Formation of the isoglutamine amide group in the cell walls of *Staphylococcus aureus*. *J. Biol. Chem.* 243, 783–790.
- Smith, C.A., 2006. Structure, function and dynamics in the mur family of bacterial cell wall ligases. *J. Mol. Biol.* 362, 640–655. <https://doi.org/10.1016/j.jmb.2006.07.066>.
- van Heijenoort, J., 2007. Lipid intermediates in the biosynthesis of bacterial peptidoglycan. *Microbiol. Mol. Biol. Rev.* 71, 620–635. <https://doi.org/10.1128/MMBR.00016-07>.
- Veiga, P., Erkelenz, M., Bernard, E., Courtin, P., Kulakauskas, S., Chapot-Chartier, M.-P., 2009. Identification of the asparagine synthase responsible for D-Asp amidation in the *Lactococcus lactis* peptidoglycan interpeptide crossbridge. *J. Bacteriol.* 191, 3752–3757. <https://doi.org/10.1128/JB.00126-09>.
- Vollmer, W., Blanot, D., De Pedro, M.A., 2008. Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* 32, 149–167. <https://doi.org/10.1111/j.1574-6976.2007.00094.x>.
- Wang, A., Daneman, N., Tan, C., Brownstein, J.S., MacFadden, D.R., 2017. Evaluating the relationship between hospital antibiotic use and antibiotic resistance in common nosocomial pathogens. *Infect. Control Hosp. Epidemiol.* 38, 1457–1463. <https://doi.org/10.1017/ice.2017.222>.
- Whited, A.M., Johs, A., 2015. The interactions of peripheral membrane proteins with biological membranes. *Chem. Phys. Lipids* 192, 51–59. <https://doi.org/10.1016/j.chemphyslip.2015.07.015>.
- Zapun, A., Philippe, J., Abrahams, K.A., Signor, L., Roper, D.I., Breukink, E., Vernet, T., 2013. In vitro reconstitution of peptidoglycan assembly from the gram-positive pathogen *Streptococcus pneumoniae*. *ACS Chem. Biol.* 8, 2688–2696. <https://doi.org/10.1021/cb400575t>.
- Zoll, S., Pätzold, B., Schlag, M., Götz, F., Kalbacher, H., Stehle, T., 2010. Structural basis of cell wall cleavage by a staphylococcal autolysin. *PLoS Pathog.* 6. <https://doi.org/10.1371/journal.ppat.1000807>.
- Zoll, S., Schlag, M., Shkumatov, A.V., Rautenberg, M., Svergun, D.I., Götz, F., Stehle, T., 2012. Ligand-binding properties and conformational dynamics of autolysin repeat domains in staphylococcal cell wall recognition. *J. Bacteriol.* 194, 3789–3802. <https://doi.org/10.1128/JB.00331-12>.