DEPARTAMENT OF CHEMISTRY

UNRAVELLING THE REACTION MECHANISM OF GLUTAMATE AMIDATION IN *STAPHYLOCOCCUS AUREUS* PEPTIDOGLYCAN

FRANCISCO MIGUEL PIÇARRA LEISICO

Master in Molecular Biotechnology

DOCTORATE IN MOLECULAR BIOSCIENCES
NOVA University Lisbon
March, 2022

DEPARTAMENT OF CHEMISTRY

UNRAVELLING THE REACTION MECHANISM OF GLUTAMATE AMIDATION IN *STAPHYLOCOCCUS AUREUS* PEPTIDOGLYCAN

FRANCISCO MIGUEL PIÇARRA LEISICO

Master in Molecular Biotechnology

Adviser: Dr. Teresa Sacadura Santos-Silva
Assistant Professor, FCT-NOVA

Co-advisers: Dr. Maria João Lobo de Reis Madeira Crispim Romão
Full Professor, FCT-NOVA

Examination Committee:

Chair: Dr. Susana Filipe Barreiros,
Full Professor, FCT-NOVA

Rapporteurs: Dr. Isabel Maria Travassos de Almeida de Jesus Bento,
Staff Scientist, European Molecular Biology Laboratory
Dr. Adriano O. Henriques,
Associate Professor, ITQB-NOVA

Adviser: Dr. Teresa Sacadura Santos-Silva,
Assistant Professor, FCT-NOVA

Members: Dr. Annette Eva Langkilde,
Associate Professor, University of Copenhagen
Dr. Karina Bivar Xavier,
Principal Investigator, Instituto Gulbenkian de Ciência
Dr. Carlos Alberto Gomes Salgueiro,
Associate Professor, FCT-NOVA
Dr. Sérgio Joaquim Raposo Filipe,
Assistant Professor, FCT-NOVA

DOCTORATE IN MOLECULAR BIOSCIENCES

NOVA University Lisbon
March, 2022

Unravelling the reaction mechanism of glutamate amidation in *Staphylococcus aureus* peptidoglycan

À minha querida família

Acknowledgments

O trabalho desenvolvido nesta tese de doutoramento contou com a contribuição de várias pessoas e instituições. Gostaria de agradecer:

Às minhas orientadoras, Professora Teresa Santos-Silva e Professora Maria João Romão, pela oportunidade de desenvolver esta tese no grupo de Cristalografia Macromolecular da FCT-NOVA, pela formação académica em métodos estruturais e bioquímicos, desenvolvimento de competências transversais como a comunicação de ciência (escrita e oral) e também, pelo desenvolvimento pessoal que me servirá para o futuro. Um sincero obrigado por tudo!

Ao Programa Doutoral Biociências Moleculares, pela formação académica e oportunidade de desenvolver esta tese de doutoramento.

Aos membros da comissão de tese, Pedro Matias e Rita Sobral, que sempre estiveram disponíveis para acompanhar e contribuir no desenvolvimento deste trabalho.

Aos colaboradores, pelas suas contribuições científicas e discussões que enriqueceram este trabalho e a minha formação científica. A special thanks to Cecíle Morlot and André Zapun for their contribution in my scientific competence.

I would like to thank the Examination Committee for the time taken to provide a rich discussion of this thesis.

Aos meus colegas do Xtal: Márcia, Viviana, Benedita, Dianas, Raquéis, Ana Luísa, Angelina, Anjos, Cristiano, Bruno, Joana, Filipas e Muthu que foram como uma segunda família no quotidiano do laboratório, contribuindo para o meu desenvolvimento pessoal e científico. Um obrigado especial ao Marino, por toda a ajuda, preocupação e dedicação à formação académica das gerações futuras. Que Deus te conserve! Ao Hugo.

Aos Dim Sum: Rita, Filipe e Catarina, pela amizade, preocupação e dedicação com a minha vida profissional e pessoal. Obrigado pela vossa descontracção e sabedoria.

Aos meus colegas Shaking, pelo grupo de amigos que criámos e por me mostrarem novos horizontes.

Aos meus amigos de Serpa, que estão tão presentes em mim como as muralhas do nosso castelo.

À minha família, em especial os meus pais Chico e Maria Helena, e mana Carolina, a quem devo tudo. São a minha constante motivação.

To Jennyfer, for all her love and future to come.

Abstract

MurT-GatD is the bi-enzymatic complex responsible for catalysing the amidation of peptidoglycan in Gram-positive bacteria, ensuring the correct assembly of their cell wall. MurT-GatD is essential in antibiotic resistant human pathogens such as *Staphylococcus aureus*, *Streptococcus pneumoniae* and *Mycobacterium tuberculosis*, and therefore, constitutes a promising target for the development of new antimicrobial agents. Peptidoglycan amidation is achieved through the glutamine amidotransferase activity of the MurT-GatD complex, requiring the presence of glutamine, ATP and magnesium. The main goal of this thesis is to unravel the reaction mechanism of peptidoglycan amidation through an integrative approach that combines structural methods with functional assays, in order to characterize the structure-function relationship of the *S. aureus* MurT-GatD complex.

The crystal structure of isolated *S. aureus* GatD was solved at 1.9 Å of resolution (PDB 5N9M) and provided the first structural insights into the MurT-GatD complex. GatD adopts the overall fold of glutaminase proteins and shows the nucleophilic cysteine and the polarizing histidine commonly associated with glutamine hydrolysis. ^{1}H-NMR experiments showed that GatD C94A or H189A abolished the glutaminase activity of the complex. Similarly, GatD R128 also proved to be an essential residue for catalysis, likely by capturing glutamine to the active site.

The crystal structure of *S. aureus* MurT-GatD was solved at 2.9 Å of resolution (PDB 7Q8E), showing a heterodimer in an extended conformation. GatD adopts the same glutaminase-like fold as in its isolated form, while MurT displays two distinct domains: the central domain containing, a cysteine-rich insertion and the ATP binding site, and the C-terminal domain that interacts with GatD. The overall structure of *S. aureus* MurT-GatD is very different from *S. pneumoniae* MurT-GatD, which adopts a compact conformation. The two complexes also adopt different conformations in solution, as observed through Small-Angle X-ray Scattering (SAXS) studies, and showed different *in vitro* enzymatic activities, suggesting that the extended conformation of *S. aureus* MurT-GatD is catalytically less competent.

The structural data of MurT-GatD from *S. aureus* and *S. pneumoniae* were combined to identify the molecular determinants involved in the glutaminase and amidotransferase active sites of the complex, as well as, in protein-protein interactions. The work carried out in this thesis contributed to the clarification of the reaction mechanism of peptidoglycan amidation, which can be explored for the development of new drugs with antimicrobial activity.

Keywords: MurT-GatD, Peptidoglycan amidation, *Staphylococcus aureus*, *Streptococcus pneumoniae*, X-ray crystallography.

Resumo

MurT-GatD é o complexo bi-enzimático que catalisa a amidação do peptidoglicano em bactérias Gram-positivas, garantindo a correta organização da sua parece celular. MurT-GatD é essencial em bactérias resistentes a antibióticos que são patogénicas a humanos, como *Staphylococcus aureus*, *Streptococcus pneumoniae* e *Mycobacterium tuberculosis*, e, consequentemente constitui um alvo promissor a ter em conta no desenvolvimento de novos agentes antimicrobianos. A amidação do peptidoglicano é realizada pela actividade de glutamina amidotransferase do complexo MurT-GatD, requerendo a presença de glutamina, ATP e magnésio. O principal objectivo desta tese é revelar o mecanismo reaccional da amidação do peptidoglicano através de uma abordagem integrativa que combina métodos estruturais com ensaios funcionais, de modo a caracterizar a relação estrutura-função do complexo MurT-GatD de *S. aureus*.

A estrutura cristalina da proteína GatD de *S. aureus* foi resolvida a 1.9 Å de resolução (PDB 5N9M) e forneceu as primeiras evidências estruturais sobre o complexo MurT-GatD. GatD adopta uma organização 3D semelhante a glutaminases, apresentando uma cisteína nucleofílica e uma histidina polarizadora comummente associadas à hidrólise de glutamina. Experiências de ^{1}H-NMR mostraram que os mutantes C94A ou H189A da GatD aboliram a actividade de glutaminase do complexo. De forma similar, R128 da GatD também provou ser um resíduo essencial na catálise, provavelmente por capturar glutamina para o centro activo.

A estrutura cristalina do complexo MurT-GatD de *S. aureus* foi resolvida a 2.9 Å de resolução (PDB 7Q8E), mostrando um heterodímero numa conformação estendida. A GatD adopta a mesma organização estrutural característica de glutaminases, enquanto a MurT apresenta dois domínios distintos: o domínio central que contém uma inserção rica em cisteínas e o local de ligação ao ATP, e o domínio C-terminal que interage com a GatD. A estrutura global do complexo MurT-GatD de *S. aureus* é muito diferente do complexo MurT-GatD de *S. pneumoniae*, tendo em conta que este adopta uma conformação compacta. Os dois complexos também adoptam diferentes conformações em solução, como observado através de estudos de Dispersão de raios X a Baixo Ângulo (SAXS), tendo demonstrado diferentes actividades enzimática *in vitro*, sugerindo que a conformação estendida de MurT-GatD de *S. aureus* é cataliticamente menos competente.

Os dados estruturais sobre o complexo MurT-GatD de *S. aureus* e *S. pneumoniae* foram combinados para identificar os determinantes moleculares envolvidos nos centros activos de glutaminase e de amidotransferase do complexo, assim como, nas interacções proteína-proteína. O trabalho desta tese teve uma importante contribuição para a elucidação do mecanismo reaccional da amidação do peptidoglicano, o que pode vir a ser explorado no desenvolvimento de novos fármacos com actividade antimicrobiana.

Palavras-chave: MurT-GatD, Amidação do peptidoglicano, *Staphylococcus aureus*, *Streptococcus pneumoniae*, Cristalografia de raios X.

Table of Contents

List of Figures

List of Tables

Acronyms

AMP-PNP	Adenylyl-imidodiphosphate
AS	Anthranilate synthase
ATP	Adenosine 5′-triphosphate
CobB	Cobyrinic acid a,c-diamide synthase
CobQ	Cobyric acid synthase
CPS	Carbamoyl phosphate synthetase
D*max*	Maximum particle dimension
ESRF	European Synchrotron Radiation Facility
GATase	Glutamine amidotransferase
Glc*N*Ac	*N*-acetyl-glucosamine
$\mathbf{F_{calc}}$	Calculated structure factor
FGAR-AT	Formylglycinamidine ribonucleotide amidotransferase
$\mathbf{F_{obs}}$	Observed structure factor
IGPS	Imidazole Glycerol Phosphate Synthase
K_D	Binding affinity
ka	Association rate constant
kd	Dissociation rate constant
K_M	Michaelis-Menten constant
MRSA	Methicillin-resistant *Staphylococcus aureus*
MSSA	Methicillin-susceptible *Staphylococcus aureus*
Mur*N*Ac	*N*-acetyl-muramic acid
NMR	Nuclear Magnetic Resonance
PBP	Penicillin-binding proteins
PDB	Protein Data Bank

PG	Peptidoglycan
PLPS	Pyridoxal 5'-phosphate synthase
Rg	Radius of gyration
RU	Response units
[S]	Substrate concentration
SAXS	Small-angle X-ray scattering
SEC	Size-exclusion chromatography
SPR	Surface plasmon resonance
UDP	Uridine 5′-diphosphate
V_{max}	Maximum velocity
V_i	Initial velocity

CHAPTER 1

GENERAL INTRODUCTION

Table 1.1. The outbreak of a drug-resistant "bad bug" in the 80s

Adapted from [1] .

Day 1	*A 34-year-old New Hampshire expectant mother visits her doctor's office complaining of severe stomach pain, vomiting, diarrhoea, fever, and chills. She is diagnosed with an intestinal infection, given intravenous fluids and a prescription for a fluoroquinolone – an antibiotic – and is sent home.*
Day 2	*At a Massachusetts hospital's emergency room, a 2-year-old boy with a severe case of diarrhoea, vomiting, dehydration, and fever is given fluids and administered a cephalosporin, another type of antibiotic, and is admitted to the hospital.*
Day 4	*The boy's lab results come back identifying the cause of his illness as Salmonella, a common foodborne bacterial infection, but, in this instance, the "bug" is highly resistant to the antibiotics commonly used to treat such infections, including cephalosporins and fluoroquinolones.* *The baby boy dies of dehydration and bloodstream infection. As for the 34-year-old woman, the Salmonella infection results in a miscarriage of an otherwise normal baby followed by the woman's death.*
Day 5	*325 people are dead. Thousands – many of them children, the elderly, and other vulnerable individuals – jam emergency rooms across the Northeast complaining of similar symptoms. Cases have been reported in 15 states along the East Coast and in the Mid-Atlantic region. Isolated cases are reported in other states, including Texas and California. Fourteen cases are reported in Mexico and 27 cases in Canada.*
Day 6	*1,730 deaths and 220,000 illnesses have occurred in the United States. The epidemic expands in other countries.* *Canada, Mexico, and Europe close their borders to U.S. food imports, and travel initiated from the United States is banned around the globe. Economic losses to the U.S. and global economies soon reach tens of billions of dollars.* *The Food and Drug Administration and Centers for Disease Control and Prevention identify the source of the infections as a milk distribution facility located in New York state. They confirm that the Salmonella not only causes severe illness, but also is resistant to all available antibiotics. Doctors can only provide supportive care, not specific, antibiotic treatment.*
Day 7	*The number of deaths and illnesses continues to climb.*

Think it can't happen? Think again. *In 1985, milk contaminated with Salmonella typhimurium infected 200,000 people across the Midwest. What distinguishes that case from our scenario is the development of a fully antibiotic-resistant strain of the bacteria as compared to the one that is only partially drug-resistant. Such "bad bugs" are evolving. Some are already here.*

Can we avert this catastrophe? If we act now, the answer is yes.

1.1. Antibiotic Resistance as a Global Threat for Public Health

Antibiotics are medicines used to prevent and treat bacterial infections since the clinical introduction of penicillin in 1941 [2]. The discovery of these antimicrobial agents is one of the greatest health care advances in history, rapidly reducing morbidity and mortality of fatal diseases. The role of antibiotics in modern medicine and its contribution to the increase of human life expectancy goes beyond their use as bacteria killers in the treatment of minor or life-threatening infectious diseases. Their ability to prevent infections ensure ideal conditions in medical settings that allow routine procedures like complex surgeries in challenging anatomic locations, organ transplantation, joint replacements, and safer chemotherapy for patients with cancer. Chronic diseases such as diabetes, asthma and rheumatoid arthritis are also dependent on the ability to treat infections using antibiotics [3]. The importance of antibiotics in modern medicine and the diverse scope of application is well described, *e.g.* in the United States, by the following numbers [4]:

1. **Sepsis Treatment: 1.7 million** adults develop sepsis each year, in which antibiotics are used to control the extreme body's reaction to infection;
2. **Surgery: 1.2 million** women had a caesarean section in 2017, in which antibiotics are recommended to prevent infections;
3. **Chronic Conditions: 30 million** people have diabetes and antibiotics are used to treat common infections in these patients;
4. **Organ Transplants**: in 2016, more than **33,000** organ transplants were performed, in which antibiotics are crucial at the surgery level and to control the risk of infection due to the weaken immune system;
5. **Dialysis for Advanced Kidney Disease**: more than **half a million** patients received dialysis treatment in 2016, and therefore, have higher risk of infection, which is the second leading cause of death in dialysis patients;
6. **Cancer care**: **more than half a million** people receive outpatient chemotherapy each year.

The large spectrum of antibiotic applications resulted in a growing threat to public health worldwide: antibiotic resistance. Additional applications of antibiotics in animal husbandry to prevent and treat bacterial infections also contribute greatly for the antibiotic resistance phenomenon.

> Antibiotic resistance consists in the ability of bacteria to resist antibiotic doses previously used to kill them. It is, therefore, an evolutionary outcome to avoid lethal selective pressure. With these mechanisms, current antibiotics are no longer effective and infectious diseases can become untreatable.

The World Health Organization, in 2014, categorized antibiotic resistance as a major threat to public health, on its first global report on this topic adverting that "*this serious threat is no longer a prediction for the future, it is happening right now in every region of the world and has the potential to affect anyone, of any age, in any country*" [5]. More recently, antibiotic resistance figures in the list of top ten threats to global health covered by the World Health Organization's thirteenth general programme of work 2019-2023 [6].

Despite the global spread of antibiotic resistance mechanisms, the Food and Drug Administration approval of new antibiotics decreased 80% from 1980 to 2010 [7]. The approval rate during the first decade of the new millennium was inferior to 1 new antibiotic per year. The weak pipeline of antibiotic development is a major contributor for the antibiotic resistance crisis. The high cost of antibiotic R&D and the rapid evolution of bacteria under selective pressure, in developing resistance mechanisms, have resulted in a low-profit business to the pharmaceutical industry. The warning call from health organization for the threat of antibiotic resistance was materialized with an increase of public investment and strategies to facilitate clinical trials regulations, resulting in the Food and Drug Administration approval of 14 new antibacterial drugs from 2010 and 2020 [8].

The World Health Organization's report "*2019 Antibacterial agents in clinical development – an analysis of the antibacterial clinical development pipeline*" revealed that there are 50 products currently under development, however, the majority are derivatives of existing chemical classes of antibiotics and have limited benefits [9]. Currently, at least 252 new agents are in the early stages of development, from which maximum five are estimated to become available in the next 10 years [10].

The ineffectiveness of medicines is the major driving force for the global burden of antibiotic resistance, however, additional risk factors must be highlighted: lack of rapid diagnostic tests to guide antibiotic prescription (and avoid misuse); difficulty to implement good infection control practises; natural epidemic transitions; immigration; antibiotic overuse in agriculture; over prescription of broad-spectrum antibiotics that apply selective pressure on commensal bacteria; and pathogen transmission between animals and humans. Low-income countries are particularly vulnerable due to inequitable access to medicines, and ease of bacteria spreading due to poor sanitation, and inadequate healthcare services and procedures.

Despite the efforts of health organizations in fighting antibiotic resistance global threat, infectious diseases are among the top 10 causes of death worldwide, *e.g.* acute lower respiratory tract infections and diarrhoeal diseases [11]. In 2019, it was estimated that 2.8 million antibiotic-resistant infections and more than 35,000 deaths, occur each year in the United States [4]. Since more than 70% of pathogenic bacteria are resistant to at least one antibiotic, treating infectious diseases requires the combination of different antibiotics and complementary approaches in healthcare settings during extended hospitalization. Its economic burden is estimated to cost nearly $5 billion annually [1].

In 2017 [11] and 2019 [4], the World Health Organization and the Centers for Disease Control and Prevention published a ranking list of current antibiotic resistance threats, respectively. Besides morbidity and mortality, they assessed the threat level of pathogenic bacteria based on additional criteria such as transmissibility, treatability, preventability, number of potential antimicrobial molecules in the development pipeline, and burden in healthcare settings and in the community (Table 1.2). The resulting rankings comprised 12 classes of pathogens that were grouped as urgent, serious, and concerning threats, according to the priority for a call of action, as shown in Table 1.3 [4]. Considering the number of infections and deaths for each pathogen, methicillin-resistant *Staphylococcus aureus* (MRSA) plays a major role in antibiotic resistance crisis.

Table 1.2. Criteria used to assess threat level of pathogenic bacteria associated with antibiotic resistance

Reproduced from [11].

Mortality	Pooled prevalence of all-cause mortality in patients with infections caused by antibiotic-resistant bacteria
Health-care burden	Need for hospitalization and increase in length of stay in patients with infections caused by antibiotic-resistant bacteria compared to patients infected by susceptible strains
Community burden	Prevalence of resistance and type of infections in community settings
Transmissibility	Isolation and transmission among three sectors: animal-human, food-human, and human-human in community and hospital settings
Prevalence of resistance	Pooled prevalence of resistance in clinically significant isolates (blood, cerebrospinal fluid, stools or swabs), stratified by World Health Organization region
10-year trend of resistance	Linear increase in 10-year prevalence of resistance in clinically significant isolates (blood, cerebrospinal fluid, stools or swabs), stratified by World Health Organization region
Preventability in healthcare setting	Availability and effectiveness of preventive measures in health- care settings
Preventability in community	Availability and effectiveness of preventive measures in community
Treatability	Availability of effective treatments (number of antibiotic classes, residual activity of antibiotics, and oral and paediatric formulations)
Pipeline	Likelihood of future development (5-7 years) of new antibiotics based on the current drug development pipeline

Table 1.3. Ranking of pathogens on the Antibiotic Resistance Threat Report (2019) from the Centers for Disease Control and Prevention

In this ranking the most important pathogens were grouped as urgent, serious and concerning threats: urgent threat pathogens require urgent and aggressive action; serious threats require prompt and sustained action; and concerning threats require careful monitoring and prevention action. The number (*N*) of infections and deaths associated to each pathogen is also presented. Adapted from [4].

Threat Level	Pathogen	*N* infection	*N* deaths
Concerning	Erythromycin-resistant group A *Streptococcus*[VI]	5,400	450
	Clindamycin- resistant group B *Streptococcus*[VI]	13,000	720
SERIOUS	Drug-resistant *Salmonella* serotype Typhi[IV]	4,100	<5
	Drug-resistant *Shigella*[IV]	77,00	<5
	Drug-resistant Tuberculosis[V]	847	62
	Drug-resistant non-typhoidal *Salmonella*[IV]	212,500	70
	Drug-resistant *Campylobacter*[IV]	448,400	70
	Drug-resistant *Candida*	34,800	1,700
	Multidrug-resistant *Pseudomonas aeruginosa*[I]	32,600	2,700
	Drug-resistant *Streptococcus pneumoniae*[IV]	900,000	3,600
	Vancomycin-resistant *Enterococcus*[I]	54,500	5,400
	ESBL-producing Enterobacteriaceae[I]	197,400	9,100
	Methicillin-resistant *S. aureus*[I]	**323,700**	**10,600**
Urgent	Drug-resistant *Neisseria gonorrhoeae*[IV]	550,000	N/A
	Candida auris[II]	323	N/A
	Carbapenem-resistant *Actinetobacter*[I]	8,500	700
	Carbapenem-resistant Enterobacteriaceae[I]	13,100	1,100
	Clostridioides difficile[III]	223,900	12,800

I - Incident hospitalized positive clinical cultures, including hospital- & community-onset – non-hospitalized cases not considered; II - only clinical cases considered; III - Infections requiring hospitalizations or in already hospitalized patients – non-hospitalized infections not considered; IV – all infections; V- cases; VI – Invasive infections - Non-invasive infections including common upper-respiratory infections like strep throat / Non-invasive infections & asymptomatic intrapartum colonization requiring prophylaxis; N/A: Not applicable; ESBL: Extended-spectrum beta-lactamase.

1.1.1. MRSA crisis

β-lactam antibiotics, such as methicillin, have been massively used worldwide to treat bacterial infections, leading to the global spread of resistant bacteria such as MRSA. The prevalence and criteria score supporting the serious threat of MRSA and its high priority for health organizations [4,11] is presented in Fig. 1.1. The World Health Organization estimates that MRSA strains are the origin of 80% of the *S. aureus* infections in Africa, 90% in the Americas, 50% in the Eastern Mediterranean, 60% in Europe, 25% in South-East Asia and 80% in Western Pacific [11]. During 2017, only in United States it was estimated that MRSA alone had a healthcare cost of $1.7 billion, higher than the economic burden associated to all the urgent threats together (Table 1.3) [4].

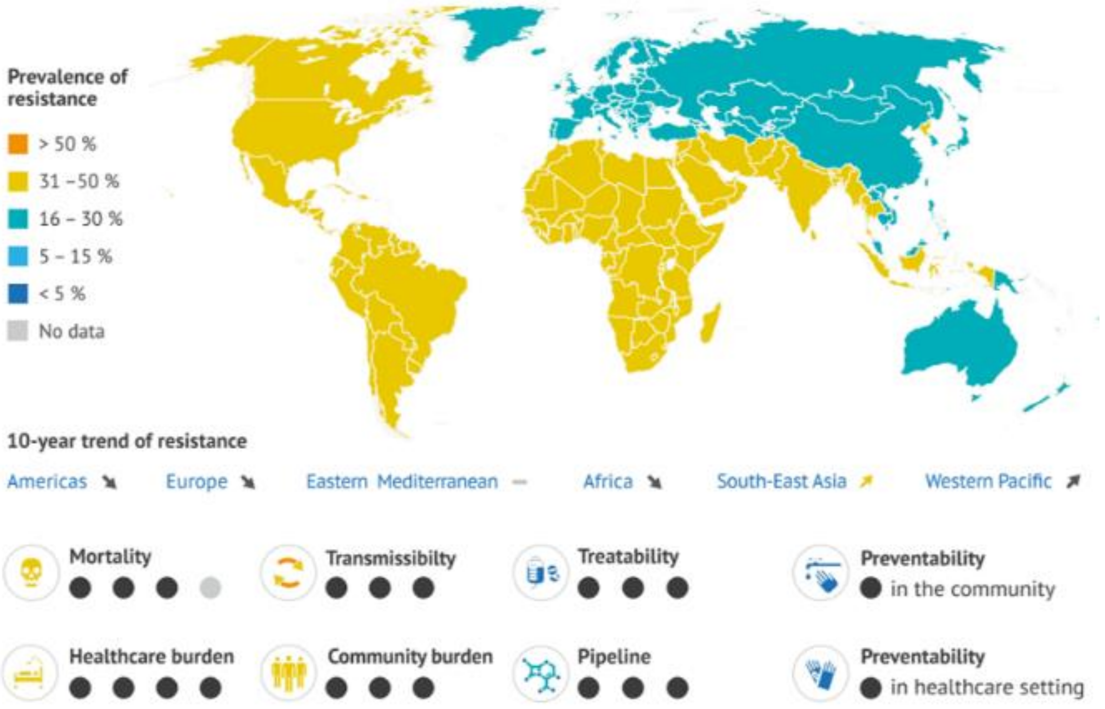


Figure 1.1. Assessment of MRSA global threat by World Health Organization.

Graphical representation of current and 10-year trend MRSA prevalence of resistance compared to colonized individuals among World Health Organization regions. Results of MRSA impact for each criterion are shown by black circles. Criteria definition is shown in Table 1.2. Reproduced from reference [11].

S. aureus is a mammalian commensal organism, usually found on the skin and in mucosal membranes of the upper part of the respiratory tract. Its presence in humans is asymptomatic, as result of the long term co-evolution within human hosts that increased the bacteria fitness throughout populations. However, in case of rupture of the colonising tissue, *S. aureus* can become an opportunistic pathogen. All *Staphylococcus* species are able to cause infectious diseases in their hosts, but *S. aureus* is considered the most prevalent pathogen of this genus of bacteria [12]. The largest reservoir of *S. aureus* is the human nares, being estimated that 20% of human population are persistent nasal carriers, 30% intermittent carriers and 50% non-carriers (age, gender and frequency of hospitalization are factors influencing the prevalence of carriage) [13]. Additionally, farm animals, with their sheer population size, are important stocks of *S. aureus* carriers. It is estimated that 3% of worldwide milking cows are infected with this bacterium [14], which is the cause of 10-12% of all clinical mastitis cases [15].

Most *S. aureus* lineages are host-specific, however swapping events between humans and other animal hosts can occur. This important transmission route through domesticated animals is critical in the case of antibiotic resistant *S. aureus* strains, like MRSA. Indeed, host-switching events of MRSA lineages between humans and cows [16], pigs [17] or chickens [18] were observed. Usually after host-switching events, bacteria acquire/lose mobile genetic elements like bacteriophages, pathogenicity islands, plasmids, transposons and staphylococcal cassette chromosomes in order to adapt and expand into new host populations. The overuse of antibiotics and utilization of contaminated animal wastes in agriculture represent another massive channels of transmission, as observed in the 1995 food-initiated outbreak of MRSA in The Netherlands [19].

Most MRSA infections in humans occur on the skin and symptoms comprise swelling, warmth, redness and pain at the infection site. However, bloodstream infections or bacteremia, which can lead to the infections of different organs such as lungs (pneumonia) and heart (endocarditis), have a mortality rate around 80% if antibiotics are not used [20,21]. Such infections are associated with breaks in the skin or mucosal membranes due to surgical procedures, use of intravascular devices or injected drugs. Implantation of medical devices is also a common source of *S. aureus* infections, since the bacterial biofilm formation can occur at the device surface. In this kind of infection, antibiotics are ineffective and the device has to be removed, followed by antibiotic treatment. *S. aureus* is also the most common bacterium causing bone and joint infections [22]. Besides the MRSA spread within healthcare facilities, it can also be disseminated through open wounds and shared equipment at the community level like in schools and social events.

The available treatments to MRSA infections comprise highly toxic and expensive secondary drugs. The severe side-effects of such treatments imply constant monitoring, administration of symptoms-related drugs and often mechanical ventilation. Compared to methicillin-susceptible *S. aureus* (MSSA), MRSA infections involve a 5 days-longer hospitalization with posterior discharge to secondary-healthcare facilities with high rate of readmissions. Despite the burden of resource consumption in treatment and long-term complications, patients infected with MRSA are estimated to be 64% more likely to die than MSSA carriers [5]. Therefore, the critical scenario around MRSA infections highlights the urgent need to develop new antibiotics.

1.1.2. Multicomponent treatments of MRSA infections

Besides the high cross resistance of MRSA to several β-lactam antibiotics [23], different *S. aureus* strains have developed resistance mechanisms against the most standard antibiotics [24,25]. The ability by which bacteria can adopt multiple resistance mechanisms challenges the efficacy of therapies based on standalone drugs to inhibit a single target. An alternative approach to treat MRSA infections is, therefore, to combine methicillin with agents that target the resistance mechanism to methicillin, potentiating its antimicrobial activity [26]. This strategy has been applied successfully in β-lactam resistant Gram-negative bacteria [27].

Methicillin, like other β-lactam antibiotics, targets the penicillin-binding proteins (PBPs) of *S. aureus*, which are responsible for the polymerization and cross-linking of the different bacterial cell wall elements. Inhibiting these enzymes leads to a weaken cell wall and consequent cell lysis. The resistance mechanism present in MRSA was acquired upon the lateral transfer of the gene *mecA*, which encodes a specific PBP (PBP2A) that has reduced affinity to methicillin. In the presence of methicillin concentrations deadly to native *S. aureus*, MRSA expressing PBP2A is able to maturate the cell wall and divide into viable cells. A multicomponent therapy combining PBP2A inhibitors and methicillin, or other β-lactam antibiotics, could be used to restore the MRSA antibiotic susceptibility [28].

Interestingly, the methicillin resistance mechanism of MRSA is not fully dependent on PBP2A since when the *mecA* gene from the homogeneously resistant MRSA COL strain was introduced into a MSSA strain, only a heterogeneous resistant phenotype was observed [29]. Further genetic studies indeed identified genes encoding factors required for full methicillin resistance. Importantly, these factors are genetically unrelated to *mecA* gene and their effect on methicillin resistance is independent on the PBP2A expression level [30]. Their complementary role demonstrated that the methicillin resistance of MRSA is multifactorial and PBP2A expression is required but not sufficient. Therefore, these factors have a great therapeutic potential as β-lactam potentiation agents.

The identification of genes that contribute for the optimal expression of methicillin resistance started in the late 70s with studies using the *S. aureus* Tn551 transposon [31]. This genetic approach allowed integrating the transposon at random locations of the chromosome from the homogeneously MRSA COL strain, assessing the restored susceptibility to methicillin of the resulting mutant strains in comparison to the parental strain and identifying the location of the transposon in the bacterial chromosome. This genetic approach only allows the identification of non-essential genes, since inhibiting the transcription of essential genes results in non-viable cells, where methicillin resistance level cannot be screened. The contribution of essential proteins to the MRSA methicillin resistance was studied using plasmids with antisense interference fragments that target specifically the genes of interest [32]. Each plasmid can then be introduced and the respective gene partially expressed in the homogeneously MRSA COL strain, resulting in viable cells that can be tested in respect to their antibiotic resistance level.

Most of the identified genes encode factors that are involved in *S. aureus* cell wall biosynthesis, while smaller groups participate in cell division, protein secretion and signal transduction systems [28,33]. Just as methicillin inhibits PBP2 function and induces bacterial lysis in the absence of PBP2A, agents that affect the production of enzymes involved in the cell wall biosynthesis will likely impair the cell structure and survival [33]. Therefore, these enzymes offer dual opportunities, as standalone therapeutic drugs, as well as β-lactam potentiation agents. Understanding the role of these factors in the biosynthesis of the cell wall is therefore critical for the rational design of new inhibitory drugs.

1.2. *S. aureus* cell wall

S. aureus was discovered in 1881 by the surgeon Alexander Ogston, while working with patients of ulcerated sores [34]. This bacterium is a member of the Firmicutes and belongs to the class Bacilli. It is a non-sporulating Gram-positive bacterium with a round shape and size up to 1 μm in diameter (Fig. 1.2). It is a facultative anaerobe characterized by low G+C DNA content. Although *S. aureus* grows optimally at 37°C and pH 7.4, it is adapted to grow at 15-45°C and at high salt concentration (up to 3.5 M NaCl). Moreover, *S. aureus* does not form flagella or spores. Tests of plasma coagulase, catalase activity, and deoxyribonuclease activities are positive in *S. aureus* and can be used to distinguish from bacteria belonging to *Streptococcus* and *Enterococcus* species [20,35].

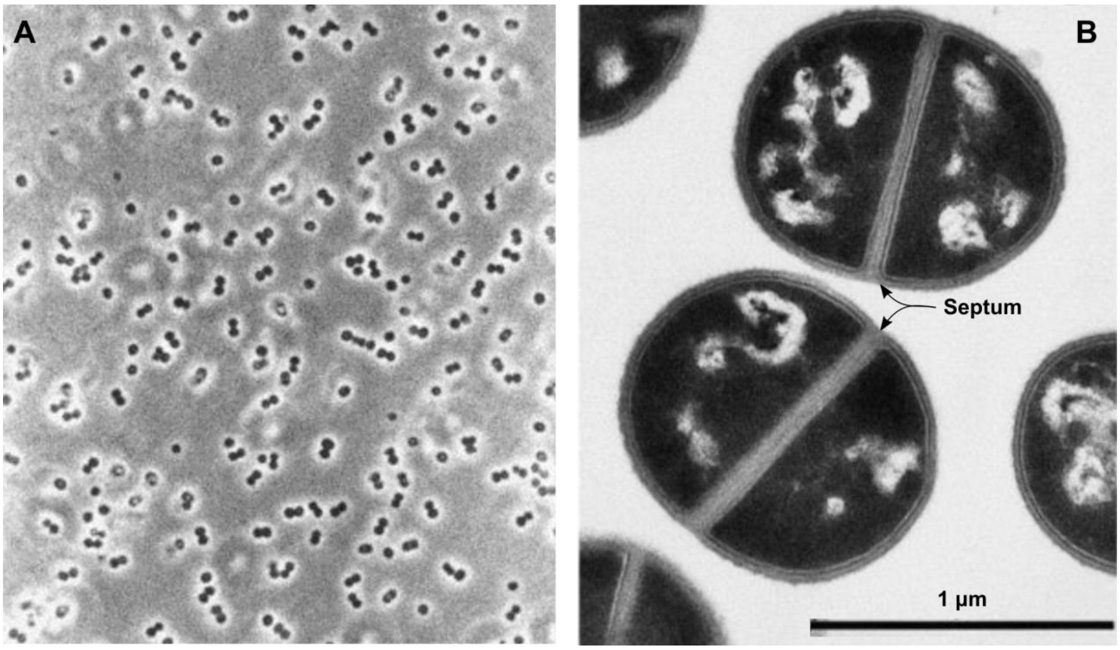


Figure 1.2. ***S. aureus*** **cell shape.**

Phase-contrast photomicrograph **(A)** and thin-section electron micrograph **(B)** of *S. aureus* cells showing its cocci shape and anatomy, respectively. In panel B, the division site or septum is highlighted. Adapted from [36].

In general, bacteria can be found in relatively dilute environments; in this case, cells experience osmotic pressure due to the higher concentration of solutes in the cell cytoplasm when compared to the extracellular space. Since the cell membrane is a sensitive and plastic structure, it must be restrained by an outer material, called cell wall. Together, both cell membrane (also called plasma membrane or cytoplasmic membrane) and cell wall form the bacterial cell envelope. Bacteria can display an additional layer of polysaccharides on top of the cell wall that function as virulence factors, helping bacteria to adhere to surfaces and often preventing phagocytosis by eukaryotic cells, such as macrophages. This outermost capsule also ensures protection against desiccation, viruses attack, and detergent-induced lysis. So far, as many as 13 chemically different capsular polysaccharides were described for *S. aureus* [37,38], ensuring a pathogenic advantage since host's immunity to one capsule type does not imply immunity to the other types.

The study of *S. aureus* cell envelope ultrastructure with cryogenic transmission electron microscopy revealed that the cell membrane is bound to a bipartite cell wall, composed by a 16-nm low-density inner wall zone, and a 19-nm high-density outer wall zone [39] (Fig. 1.3A). The new cell wall formed between two daughter cells during bacteria division, known as septum, is composed by two bipartite cell wall layers separated by an additional low-density inner wall zone where separation takes place (Fig. 1.3B) [35,39]. This study showed that the outer wall zone is mainly composed of peptidoglycan (PG; also known as murein or muropeptide) and teichoic acids. The inner wall zone lacks dense polymeric structure and is mainly composed of soluble components, being proposed to be a periplasmic space, as previously suggested to happen in other Gram-positive bacterium as *Bacillus subtilis* [40].

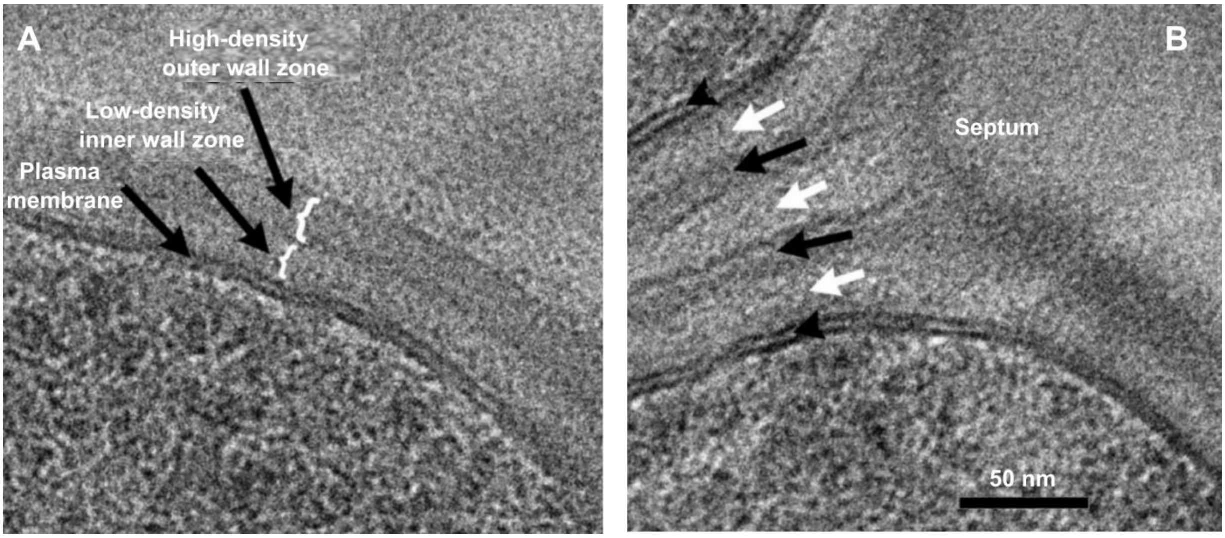


Figure 1.3. Ultrastructure *S. aureus* cell envelope.

Frozen-hydrated sections investigated using cryo-transmission electron microscopy allowed a closer look into the ultrastructure of *S. aureus* cell envelope at non-septal regions **(A)** and at the septum **(B)**. Cell envelope at non-septal regions **(A)** in *S. aureus* shows the plasma membrane bound to a bipartite wall composed by a low-density inner wall zone and a high-density outer wall zone. At the septum **(B)**, five alternated low- (white arrows) and high-density (black arrows) zones are distinguished between the plasma membranes (arrowheads) of the daughter cell. Adapted from [39].

The cell wall of Gram-positive and Gram-negative bacteria differs both in composition and in the proportion of the shared components (Fig. 1.4) [41]. One of the main differences is the presence of an outer membrane in Gram-negative bacteria. This protects the organism from the environment, stabilizes the osmotic pressure of the inner membrane, and establishes the periplasmic space that functions as a retained compartment for the extra cytoplasmic enzymes. The outer membrane also serves to anchor lipopolysaccharides and porins; both contribute for the outer cell membrane permeability to hydrophilic molecules, while the lipopolysaccharides have the additional role to promote interactions with the host [42].

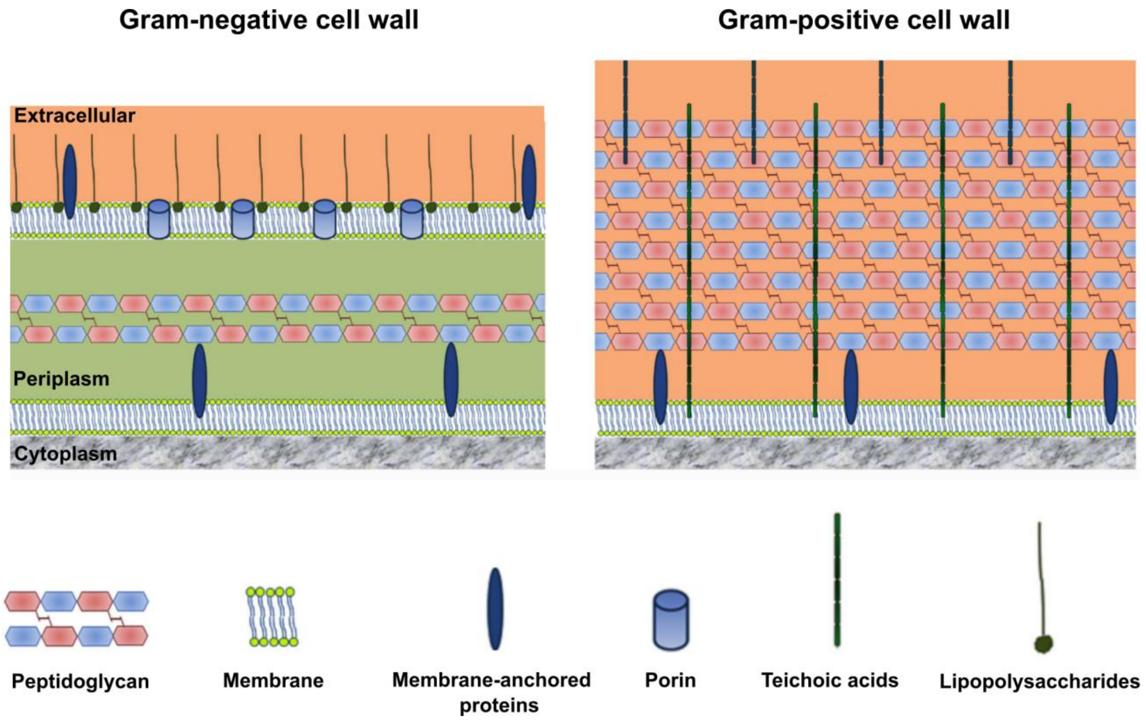


Figure 1.4. Schematic drawing of Gram-negative and Gram-positive cell walls.

Recent model of the cell wall organization in bacteria highlighting the structural role of an outer membrane in Gram-negative bacteria and PG-teichoic acids in Gram-positive bacteria. Adapted from [41].

The cell wall of Gram-positive bacteria lacks the outer membrane and relies on its major components - PG and teichoic acids - to achieve similar chemical, osmotic and mechanical protection. PG is a polymer of sugars and amino acids forming a thick mesh-like layer, while teichoic acids are repetitive polyol subunits of glycerol phosphate or ribitol phosphate. Teichoid acids are divided in lipoteichoic acids or wall teichoic acids, depending if they are anchored to the cytoplasmic cell membrane through lipid domains, or covalently bound to the PG, respectively [41,43]. Gram-negative bacteria cell walls do not possess teichoic acids and have a single PG layer, which along with the outer membrane, stabilizes the osmotic pressure of the inner membrane (Fig. 1.4).

The structural organization of the bacterial cell wall is also dependent on the presence of a large number of proteins, as shown by the identification of around 100 different examples in the proteome of the homogeneously MRSA COL strain [44]. In Gram-positive bacteria, these proteins can be found bound to the cytoplasmic membrane or directly to the cell wall PG, and are involved in a multitude of functions. The membrane-anchored proteins are mainly involved in the acquisition of ions and amino acids, but can additional function as virulence factors in pathogenic strains [44,45]. Cell wall-interacting proteins (either covalently or non-covalently) are mainly surface adhesins participating in both commensal and pathogenic functions like cell division and invasion the non-phagocytic cells and innate immune responses of the host organism, respectively [45,46].

The lack of outer membrane in Gram-positive bacteria implies that the protection against turgor pressure and mechanical stress is fully dependent on the cell wall. However, it must also allow constant exchanges of biomolecules between the cytoplasm and the surrounding environment. These characteristics of the cell wall are provided by the PG polymer, which is a ubiquitous component of bacterial cell walls. Given its key roles in cell division, shape maintenance, conservation of the internal osmotic pressure, anchoring surface-exposed virulence factors, and promoting of cell wall dynamics for constant exchange with the environmental medium, the PG biosynthetic machinery has been the preferred target for antibiotic development since the discovery of penicillin.

1.2.1. Biosynthesis of *S. aureus* PG

The monomer of *S. aureus* PG (Fig. 1.5A) is composed by a disaccharide backbone of *N*-acetyl-glucosamine (Glc*N*Ac) and *N*-acetyl-muramic acid (Mur*N*Ac) linked by a β-1,4 glycosidic bond. The latter is covalently bound to the stem peptide, corresponding to L-Ala-D-isoGln-L-Lys-D-Ala-D-Ala. The ε-amine group of the residue L-Lys additionally contains an interpeptide bridge composed of five glycines. PG monomers are assembled in glycan strands, which are cross-linked to form the mesh-like structure of PG (Fig. 1.5B). Cross-linking specifically occurs between the pentaglycine interpeptide bridge of one monomer and the penultimate D-Ala of a neighbouring monomer, which loses the terminal D-Ala of the stem peptide during the cross-linking reaction.

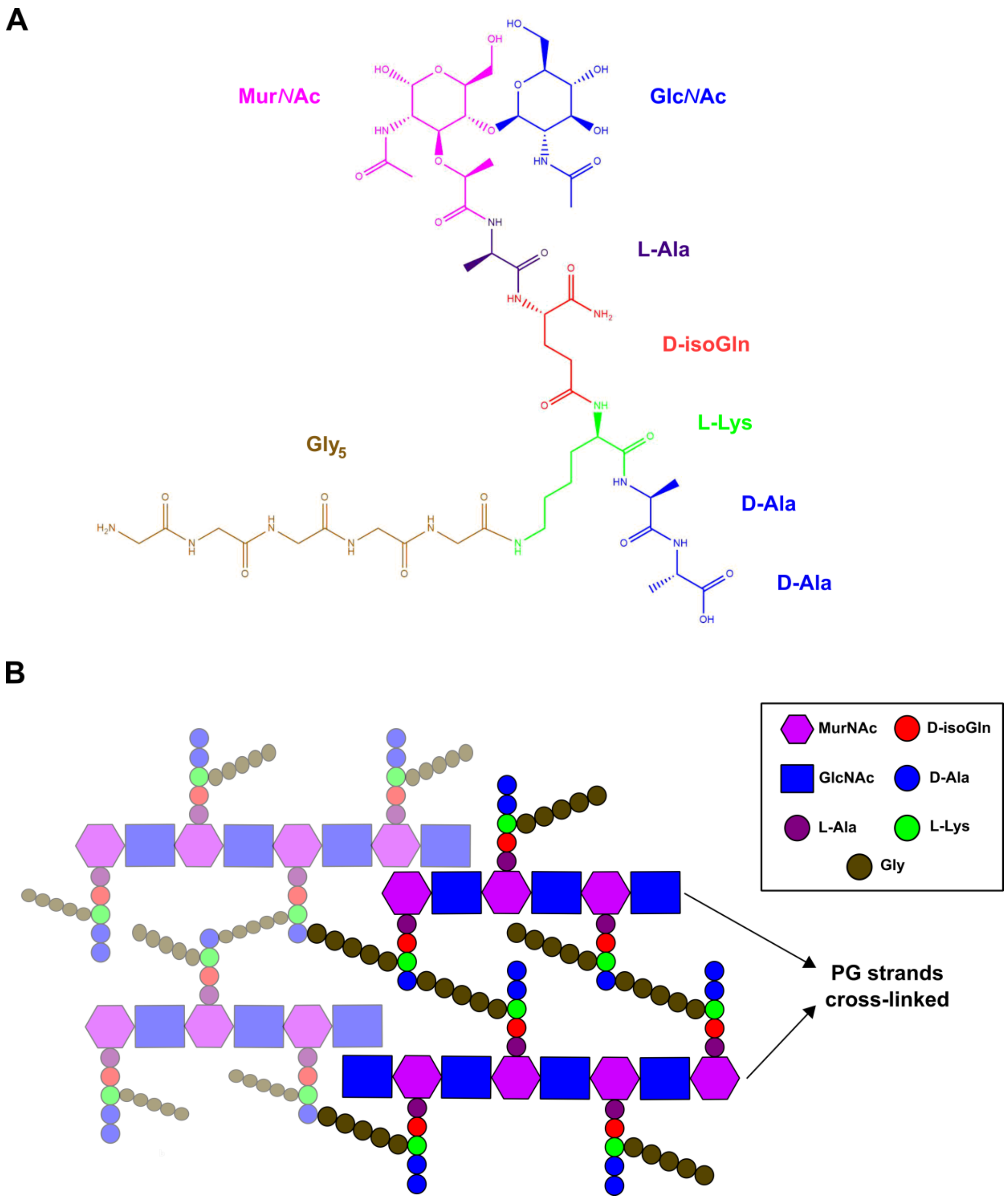


Figure 1.5. ***S. aureus*** **PG structure.**

(A) Chemical structure of *S. aureus* PG monomer, which can be divided in three parts: the disaccharide unit of Glc*N*Ac (in blue) and Mur*N*Ac (in pink) linked by a β-1,4 glycosidic bond; the stem peptide composed of L-Ala-D-isoGln-L-Lys-D-Ala-D-Ala coloured in violet, red, green and dark blue, respectively; and the interpeptide bridge (pentaglycine) moiety in brown. **(B)** Schematic representation of *S. aureus* PG mesh-like organization within the cell wall, showing cross-links between PG monomer of successive glycan strands. Glycans are represented according to the Symbol Nomenclature For Glycans guidelines [47].

S. aureus PG is characterized by short glycan strands of 5 to 26 disaccharide units [48] with a cross-linking degree of 80% [49]. PG composition and organization is species-specific and can vary greatly among bacteria: for example, *B. subtilis* PG is composed of long glycan strands (hundreds of disaccharide units) with 40% cross-linking [50]. High-resolution analysis of *S. aureus* PG treated with digestive enzymes (muramidase and/or endopeptidases) revealed a heterologous mixture of muropeptide oligomers. In fact, only 6% of PG corresponds to monomeric building blocks (disaccharide and stem peptide), while the remaining 94% show the presence of cross-linked oligomers: 20% is composed of muropeptide dimers; 40% of oligomers possessing 3-9 muropeptides; 15-25% of higher oligomeric structures with up to 26 muropeptide units [51]. The composition and cross-linking degree of *S. aureus* PG is conserved between isolates of different clonal types, either susceptible or resistant to methicillin, which is consistent with the species-specificity of bacterial PG organization [37].

PG biosynthesis is a well-studied multi-step process that can be divided in three stages, occurring at three different cellular locations (see reference [52]) (Fig. 1.6). The first step takes place at the cytoplasm, where the soluble nucleotide precursors uridine 5′-diphosphate (UDP)-Glc*N*Ac and UDP-Mur*N*Ac–stem peptide are synthesized. The following step occurs at the inner face of the cytoplasmic membrane and involves the transfer of UDP-Mur*N*Ac-stem peptide to the membrane carrier undecaprenyl phosphate, which results in the PG precursor called lipid I. The subsequent glycosidic linkage of UDP-Glc*N*Ac to lipid I lead to the formation of the lipid II. This last PG precursor is then translocated across the cytoplasmic membrane and integrated into the nascent PG. The polymerization and cross-linking of PG constitutes the last stage of its biosynthesis, which is accomplished through transglycosylation and transpeptidation reactions catalysed by PBPs.

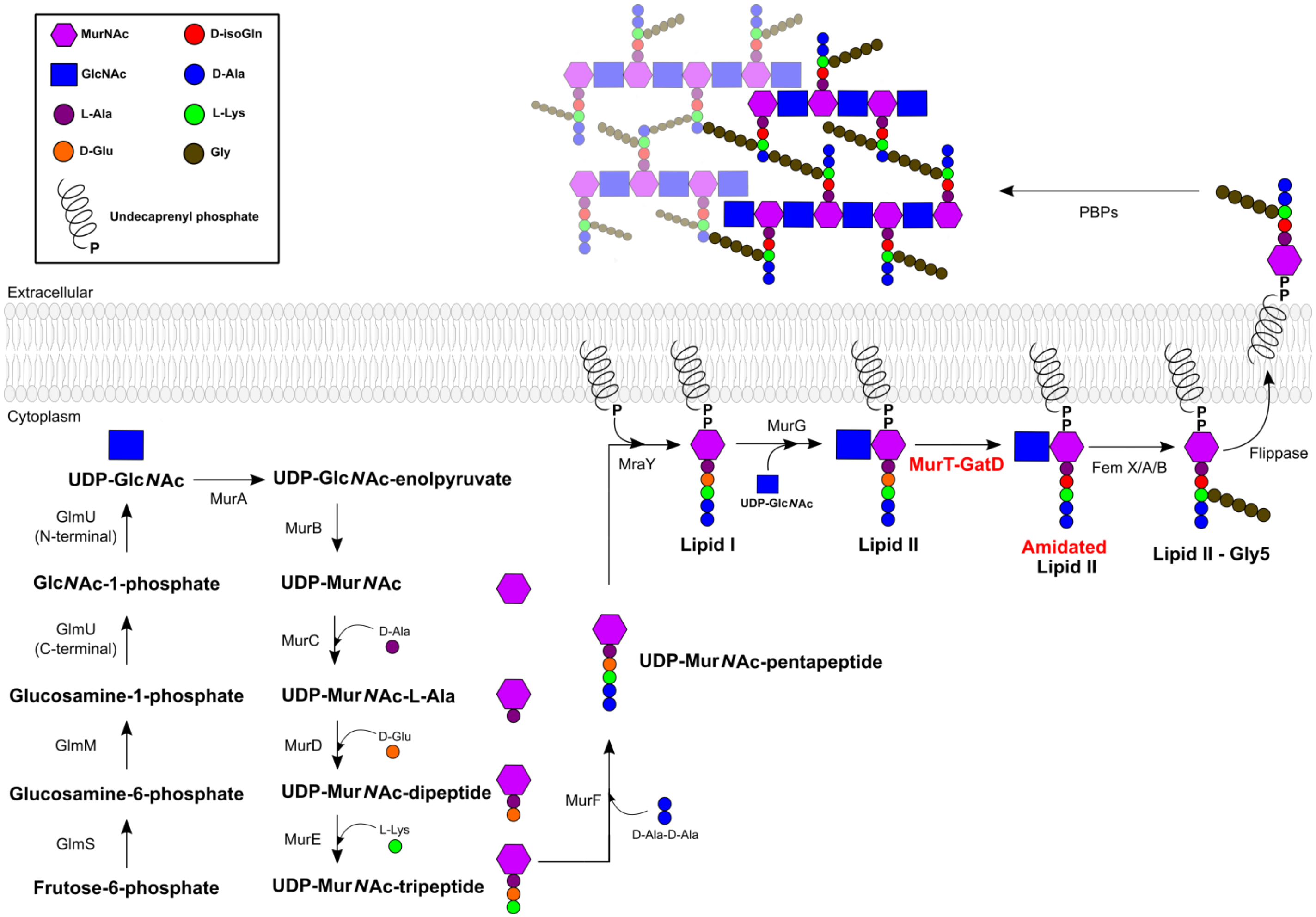


Figure 1.6. ***S. aureus*** **PG biosynthesis.**

Schematic representation of *S. aureus* PG biosynthesis and modification reactions, following the configuration of Fig. 1.5. The enzymes catalysing each reaction and key substrates for the final composition of *S. aureus* are accordingly stated.

1.2.1.1. The cytoplasmic step

The biosynthesis of UDP-Glc*N*Ac involves successive reactions catalysed by four enzymes (Fig. 1.6). In the first reaction, the enzyme aminotransferase GlmS (Uniprot Q5HE49) uses glutamine to amidate the C2 hydroxyl group of fructose-6-phosphate, which is converted into glucosamine-6-phosphate. Then, this product is converted into glucosamine-1-phosphate through an isomerization reaction catalysed by the mutase GlmM (Uniprot Q5HE43). The following enzyme acting is the bifunctional GlmU (Uniprot Q5HIH6): its C-terminal acetyltransferase domain transfers the acetyl group from acetyl-CoA to glucosamine-1-phosphate, by which results the Glc*N*Ac-1-phosphate; the GlmU N-terminal acts as uridyltransferase with the transfer of uridine-5′-monophosphate (from uridine-5′-triphosphate) to Glc*N*Ac-1-P, yielding the formation of UDP-Glc*N*Ac.

The formation of UDP-Mur*N*Ac arises from UDP-Glc*N*Ac in a two-step reaction (Fig. 1.6). The first is catalysed by the transferase MurA (Uniprot Q5HE76), which transfers the enolpyruvyl group from phosphoenolpyruvate to the C3 hydroxyl group of UDP-Glc*N*Ac, leading to the formation of UDP-Glc*N*Ac-enolpyruvate. In the last step, the enolpyruvyl group is reduced to a D-lactyl moiety by MurB (Uniprot Q5HHT2), resulting in the formation of UDP-Mur*N*Ac.

The cytoplasmic step of PG biosynthesis (Fig. 1.6) ends with the enzymatic stepwise attachment of the stem peptide onto the D-lactyl moiety of UDP-Mur*N*Ac. These reactions are adenosine 5′-triphosphate (ATP)-dependent and catalysed by Mur ligases: MurC (Uniprot Q5HF34), MurD (Uniprot Q5HGP8), MurE (Uniprot O86491) and MurF (Uniprot A0A0H2WWP1) enzymes are responsible for the successive incorporation of L-alanine, D-glutamate, L-lysine and dipeptide D-alanyl-D-alanine, respectively. The peptide bond between D-glutamate and L-lysine has the $\gamma \rightarrow \alpha$ conformation (amide bond between the δ-carboxyl of D-glutamate and α-amine of L-lysine). Previously, the Ddl-ligase enzyme (Uniprot Q5HEB7) is responsible to assebmle the dipeptide D-Ala-D-Ala. The final cytoplasmic precursor of PG is UDP-Mur*N*Ac-L-Ala-D-γ-Glu-L-Lys-D-Ala-D-Ala [52].

1.2.1.2. The cytoplasmic membrane inner face step

The second step of PG biosynthesis (Fig. 1.6) starts with the phosphotransferase activity of MraY (Uniprot Q5HGP9). This enzyme catalyses the transfer of Mur*N*Ac-stem peptide from UDP-Mur*N*Ac-stem peptide to the membrane-bound lipid carrier undecaprenyl phosphate, involving the anomeric C1 of Mur*N*Ac. This reaction yields the PG precursor called lipid I. Then, the membrane-tethered glycosyltransferase MurG (Uniprot Q5HG02) transfers the Glc*N*Ac moiety of UDP-Glc*N*Ac to lipid I through a β-1,4 glycosidic bond (between the C4 hydroxyl oxygen of Mur*N*Ac and the anomeric C1 of Glc*N*Ac), resulting in the formation of lipid II. The enzymatic activities of MurY and MurG are presumably coupled, since little amount of free lipid I is observed in the membrane bilayer [53,54].

Lipid II is further modified before translocation to the outer side of the cytoplasmic membrane (Fig. 1.6). Lipid II is amidated at the α-hydroxyl group of D-glutamate, giving rise to D-isoglutamine. This reaction is catalysed by the enzymatic complex formed by MurT (Uniprot A0A0H2WZQ7) and GatD (Uniprot A0A0H2WZ38) proteins [55,56], which is the main scope of this dissertation. Moreover, the interpeptide bridge pentaglycine is attached to the ε-amine group of the stem peptide L-Lys by peptidyltransferases that use glycyl-charged tRNAs as donors. Enzymes FemX (Uniprot Q5HDU6), FemA (Uniprot Q5HG45), and FemB (Uniprot Q5HG44) catalyse the sequential addition of glycine residues to position 1, positions 2 and 3, and positions 4 and 5, respectively [57].

The fully modified lipid II molecule is translocated across the cytoplasmic membrane and integrated into the nascent PG (Fig. 1.6). The identity of the lipid II flippase is still under debate. Conflicting evidences suggest that either MurJ [58] or FtsW [59] perform this function. Although it was recently shown that FtsW functions as PG polymerase when complexed with its cognate PBP, its secondary role as flippase was not rejected [60].

1.2.1.3. The cytoplasmic membrane outer face step

The final step of PG biosynthesis consists in the polymerization of lipid II molecules into its final mesh-like structure (Fig. 1.6). First, the disaccharide unit of successive lipid II molecules are transglycosylated, forming PG strands. Then, adjacent PG strands are cross-linked through their stem peptides by transpeptidation. The membrane-anchored PBPs on the outer leaflet of the cytoplasmic membrane orchestrate this process through their glycosyltransferase and transpeptidase activities (Fig. 1.6). The number of PBPs varies among bacterial species according to their morphology and life cycle, however their reaction mechanisms are conserved.

The transglycosylation reaction occurs between the newly available lipid II molecule after translocation from the cytoplasm and the nascent glycan strand that is still bound to the cytoplasmic membrane by the undecaprenyl moiety of the lastly added lipid II molecule. The PBP glycosyltransferase activity starts with the deprotonation of the C4 hydroxyl group of Glc*N*Ac from the new translocated lipid II (acceptor of the glycosidic bond) and stabilization of the pyrophosphate-undecaprenyl group of the Mur*N*Ac from the last lipid II added into the growing PG glycan strand (donor of the glycosidic bond). This undecaprenyl-pyrophosphate group is released, leaving the anomeric carbon of Mur*N*Ac free to bind the deprotonated C4 of Glc*N*Ac and form a β-1,4 glycosidic bond [61]. The direction of glycan strand elongation happens through its reducing end, which first is protected by the undecaprenyl-pyrophosphate moiety, but becomes available to bind freshly translocated lipid II molecules [62].

The transpeptidation reaction is the final step of PG biosynthesis and consists in the cross-linking of the nascent glycan strand into the existing PG mesh-like structure. In the first half-reaction, the transpeptidase active site of PBPs recognizes the last two alanine residues of the stem peptide from the nascent glycan strand (acyl-donor strand) and a nucleophilic serine residue attacks the amide bond between the two alanine residues. This leads to the acylation of the penultimate D-Ala and release of the last one. In the second half-reaction, the tetrapeptide acyl-PBP intermediate is deacylated by a nucleophilic attack of the stem peptide from an adjacent glycan strand within the existing PG mesh-like structure (acyl-acceptor strand). Such nucleophile corresponds to the primary amine from either the L-Lys of the stem peptide or from the interpeptide bridge (pentaglycine in case of *S. aureus*). This reaction results in the cross-linking of both glycan strands through their stem peptides.

S. aureus has four native PBPs. PBP1 displays transpeptidase activity [63] required for the last stages of cell division and separation of daughter cells [64]. PBP2 is the only bifunctional PBP in *S. aureus*, exhibiting transglycosylation and transpeptidation activities through its N- and C-terminal domains, respectively [65]. β-lactam antibiotics have a high antimicrobial activity due to their resemblance with the D-Ala-D-Ala moiety of PG pentapeptide and consequent competition to the transpeptidation active site of PBP2. As mentioned in the previous section, PBP2A has lower affinity to β-lactam antibiotics like methicillin and is able to perform the transpeptidation reaction required for PG cross-linking [65]. PBP3 is the less studied staphylococcal PBP and although it contains a transpeptidase-like domain, its function in cell wall synthesis and cell division has not be described yet [37]. PBP4 possesses transpeptidase activity that has a critical function in the maturation of PG: PBP2 catalyse the initial cross-linking of PG into dimers, trimers, tetramers and pentamers of strands, which then serve as substrates for the PBP4 activity in producing more complex cross-linked PG oligomers [66].

Once the *S. aureus* PG is synthesized and assembled, it undergoes several modifications such as O-acetylation and N-deacetylation of its MurNAc moiety, which have been mainly linked with mechanisms of evasion within the human host [37]. Moreover, the covalent binding of wall teichoic acids, proteins and polysaccharides onto the PG structure also occur at the cell wall level [41]. This PG remodelling step is essential for septum formation and separation of daughter cells during cell division [37].

The critical role of these enzymes in cell wall biosynthesis and modification offers a broad spectrum of targets and therapy combinations to treat infections caused by resistance bacterial strains, including MRSA.

1.3. Fighting MRSA through PG amidation

All the enzymes involved in PG biosynthesis were identified as factors essential for the full expression of MRSA resistance to methicillin, since their coordinated activity ensures the production of PG precursors that are suitable for the function of PBPs (including PBP2A). The therapeutic potential of targeting these enzymes led to the development of inhibitory drugs that target specifically the biosynthesis of PG precursors (GlmS, MurA, MurG and FemA), the translocation of lipid II to the extracellular leaflet of the membrane (MurJ) or both the transglycosylation and transpeptidation activities of PBPs (see review [28]). These molecules can be used as standalone drugs since they target essential processes for bacterial survival or as β-lactam potentiation agents in multicomponent strategies [26]. Combining drugs that act synergistically is particularly interesting because it allow reducing the level of each component, limiting the toxicity of the treatment.

Multicomponent therapies face specific challenges regarding co-dosing and co-formulation strategies, drug administration routes, and clinical and regulatory issues [28]. Additionally, important differences between target enzymes must be considered, such as their capacity to restore β-lactam susceptibility, localization in the cell, intrinsic druggability or conservation among pathogenic species. MRSA resistance to methicillin depends on the cumulative action of many factors, offering a broad array of target combinations and therapies. However, most of the enzymes targeted so far catalyse reactions that are ubiquitous in bacteria and therefore constitute therapies prone to the emergence of new resistance mechanisms. Importantly, enzymes that catalyse essential reactions in the biosynthesis of PG but are specific to restricted groups of pathogens constitute safer therapeutic targets.

In general, the disaccharide backbone of PG is conserved among bacteria, however the stem peptide can widely vary [52]. The first residue is mainly L-alanine, but in some species, it can be glycine or L-serine. Gram-positive bacteria mostly have D-isoglutamine as the second amino acid of the stem peptide, while the majority of Gram-negative bacteria possess D-glutamate. The highest variability is found at the third position, where diamino acids such as meso-diaminopimilic acid and L-lysine are present in most Gram-negative and Gram-positive bacteria, respectively. The fourth and fifth position are occupied by two D-alanine residues, however the last residue is replaced by D-lactate or D-serine in some drug-resistance species.

The nature of the stem peptide variability among bacterial species is related to the type of PG cross-linking. Usually it consists of 3-4 cross-linkages, in which the amino group from the third residue of one stem peptide moiety (acyl-acceptor) is covalently bound to the carboxyl group of the fourth amino acid of a neighbouring stem peptide subunit (acyl-donor). In most Gram-negative bacteria this linkage occurs directly and in most Gram-positive bacteria through an interpeptide bridge. The size of interpeptide bridges can vary between one to seven residues, comprising several arrangements of different amino acids. For example, while *S. aureus* displays a pentaglycine, *S. pneumoniae* possesses a dipeptide composed of L-serine-L-alanine or L-alanine-L-alanine.

Given the importance of cross-linking for the structure and function of PG, the enzymes involved in the biosynthesis of the variable stem peptides and interpeptide bridges constitute promising targets for the development of drugs with narrow range of application and therefore less likely to promote the emergence of new resistance mechanisms. This strategy is encouraging to treat MRSA infection given the high degree of cross-linking in *S. aureus* PG and its structural particularities. Fem proteins, which produce the *S. aureus* pentaglycine interpeptide bridge, have been extensively characterized regarding their essential role in cell survival [67,68] and MRSA methicillin resistance [69–71]. Furthermore, an inhibitor specifically targeting FemA was identified as enhancing the potency of β-lactams against MRSA [72]. Even though the amidation of the D-glutamate at the second position of PG stem peptide is also crucial for PG cross-linking, the MurT and GatD enzymes that catalyse this reaction have not been explored as a therapeutic targets.

1.3.1. PG amidation and the MurT-GatD protein complex

The first phenotypic report of a Tn551 mutant that restored MRSA methicillin susceptibility through the impairment of PG amidation corresponded to the *glutamine synthetase repressor* gene [73]. Since this gene is co-transcribed with the *glutamine synthetase* gene, inhibition of its expression results in production of the glutamine synthetase enzyme and consequent decrease of the intracellular glutamine pool. Liquid chromatography followed by mass spectrometry analysis of the PG from this mutant showed an increase in stem peptide D-glutamate in comparison with the parental strain containing D-isoglutamine [73], which was later confirmed to be dependent on the lack of glutamine production, since addition of this amino acid in the growth medium restored PG amidation [74].

The *glutamine synthetase repressor* mutant also showed a reduction in PG cross-linking and a high percentage of PG monomers with free D-Ala-D-Ala at their stem peptide terminus, suggesting that non-amidated PG precursors are poor substrates for the transpeptidation activity of PBPs [75]. These observations are in agreement with Nakel and co-workers that showed PG dimers only form in the cell wall of the Gram-positive bacteria *Aerococcus viridans* if at least one monomer is amidated [76], and Zapun *et al.* that showed the *in vitro* efficient transpeptidase activity of *S. pneumoniae* PBPs required amidated PG monomers [77]. In the last couple of years, biochemical studies demonstrated that the PG cross-linking reaction is more efficient if D-isoglutamine is present in the acyl-donor strand than in the acyl-acceptor one [78,79]. Such observation is not surprising since the PG acyl-donor strand is the one recognized by PBPs to initiate the reaction and form the acyl-intermediate, while the acyl-acceptor strand is responsible by the nucleophilic attack though its variable interpeptide bridge.

The importance of PG amidation for cell wall cross-linking and its involvement in the MRSA resistance to methicillin highlights the therapeutic potential of targeting this molecular process. Even though the *glutamine synthetase repressor* mutant was identified and characterized in the early days of the Tn551 transposon mutants, the enzymes catalysing the PG amidation reaction were only recently discovered.

In 2011, two new genes of unknown function were identified by the antisense-based genetic strategy as potentiators of the β-lactam antibacterial activity [32]. Two studies in 2012 reported these genes as *murT* and *gatD* in the MRSA strains COL [55] and N315 [56]. The PG profile of the MRSA COL *murT-gatD* conditional mutant showed two extra species when compared to the parental homogeneously MRSA COL strain: muropeptides containing D-glutamate instead of D-isoglutamine, with the stem peptide terminal either containing two D-alanine or one D-alanine linked to glycine residues [55]. These observations demonstrated that MurT and GatD proteins are important for PG amidation. Even though the PG profiles of the *murT-gatD* and *glutamine synthetase repressor* mutants are similar, it was shown that *murT* and *gatD* genes do not interfere with the transcription of the *glutamine synthetase* and *glutamine synthetase repressor* genes [55]. This observation suggests that both phenotypes result from different molecular mechanisms.

Analysis of MRSA COL genome sequence revealed that the open reading frames of *murT* and *gatD* genes are transcribed in the same direction and inserted in the same operon. The *murT* stop codon and *gatD* first methionine codon are separated only by four base pairs, suggesting that both genes might be co-transcribed by the same promoter. These genes occur as a syntenic block in bacteria, being mainly found in Gram-positive bacteria [55]. The essentiality role of MurT and GatD proteins was demonstrated for *S. aureus* [56,80,81], *S. pneumoniae* [77,82,83] and *M. tuberculosis* [84]. These proteins therefore have the important role of providing amidated PG for the efficient transpeptidation reaction in PG cross-linking.

In vitro studies showed that MurT (49 kDa) and GatD (28 kDa) form an obligatory heterodimer [56]. Interestingly, *S. aureus* MurT-GatD complex was able to amidate lipid I, lipid II and lipid II-pentaglycine but not UDP-Mur*N*Ac-pentapeptide, suggesting that the substrate of PG amidation is one of the lipid precursors [56]. These observations agree with the report that both strains, MRSA COL and *murT-gatD* conditional mutant, have their cytoplasmic non-lipid PG precursors containing only D-glutamate in the pentapeptide moiety [55]. A different *in vivo* study showed that the *murT-gatD* conditional mutant presents higher susceptibility to plectasin [56], a fungal defensin peptide that binds lipid II impairing cell wall biosynthesis [85]. Biochemical characterization of the lipid II amidation reaction catalysed by the *S. aureus* MurT-GatD complex revealed that glutamine, ATP and Mg^{2+} ions are required [56].

In comparison to the parental MRSA COL strain, the *murT-gatD* conditional mutant showed lower resistance to the β-lactam antibiotic oxacillin, confirming the previous observations that these proteins contribute to the β-lactams resistance mechanisms.

Moreover, this mutant also showed higher sensitivity to lysozyme, which suggest that PG amidation is important for the bacterial to evade host defence systems [55,86]. The direct role of MurT-GatD complex on PG amidation, cross-linking, cell survival, and antibiotic resistance mechanisms present in a narrow group of Gram-positive pathogenic bacteria highlights its therapeutic potential as an antimicrobial target with low perspectives of resistance mechanism dissemination.

1.4. Objectives of the thesis

Exploring MurT-GatD as a drug target requires the detailed description of how these proteins function together to catalyse the amidation of lipid II. The main goal of this work is to characterize *S. aureus* MurT-GatD using an integrative structure-function methodology, by combining structural methods such as X-ray Crystallography and SAXS with functional assays. The work plan aims at:

1) Optimizing protein expression and purification conditions in order to obtain pure, homogeneous and stable samples;

2) Setting up crystallization conditions for the production of reproducible protein crystals;

3) Determining the crystal structure of the proteins at medium/high resolution;

4) Exploring the structural information in order to identify the main molecular determinants involved in the enzymatic activities of the complex;

5) Developing functional assays to follow the enzymatic activities of the complex and, consequently, test the catalytic role of specific residues.

The insights into the function of MurT-GatD and the reaction mechanism of lipid II amidation obtained from this work plan can then be explored to design small molecules with antimicrobial activity. The development of crystallization conditions and functional assays also provide excellent platforms to test and understand the inhibitory action of these potential drugs.

1.5. Protein crystallography

Structural biology concerns the structure of biological macromolecules, from proteins to proteoglycans and protein-nucleic acids complexes. The 3D structure of a protein corresponds to the folding of its polypeptide chain based on the interaction established among the different amino acids. Access to the arrangement of a protein's atoms reveals how it interacts with other biological components in order to function. The substrate binding pocket of an enzyme or the protein-protein interface of a multicomponent complex, are just two examples of interactions driving physiological functions that can be observed through structural studies. Several methods are applied in structural biology, like mass spectrometry, limited proteolysis, electron paramagnetic resonance, X-ray crystallography, Small-angle X-ray scattering (SAXS), nuclear magnetic resonance (NMR) spectroscopy and cryo-electron microscopy. The method is chosen mainly based on the question to be addressed, where the protein's size can be a limiting factor. Comparing with the other techniques, X-ray crystallography allows studying molecular details with no size limit, making it the most used method to obtain structural information. It also offers the advantage of providing atomic resolution, which allows observing the position of individual atoms. This characteristic is particularly important when studying the interaction of proteins with small molecules that have either physiological or pharmaceutical relevance.

X-ray crystallography is based on the interaction of electrons present in crystallized molecules with X-rays. This type of radiation was discovered by Wilhelm Conrad Röntgen in 1895 [87] but only in 1912 Max von Laue showed that X-rays irradiated from salt crystals produce diffraction patterns that reveal the internal atomic periodicity of the crystal [88]. In the same year the father-and-son team, Henry and Lawrence Bragg successfully solved the first crystal structure of sodium chloride [89]. X-ray crystallography was then born. However, the first diffraction data from a protein (pepsine) crystal was obtained in 1934 by Bernal and Crowfoot [90], andthe first protein structure was only determined in 1958 by John Kendrew [91], marking the beginning of protein crystallography. The contribution of this field in chemistry and medicine has been recognized by more than a dozen of Nobel Prizes attributed to crystallographers [92].

Experimentally, the initial steps of a biomolecular crystallography project are shared with all the other methods of structural biology, corresponding to gene cloning and, protein production and purification. A pure, homogeneous and stable protein sample is crucial for the success of structure determination. Each method requires specific sample preparation procedures according to the respective strategy of data collection. Protein crystallography requires the production of protein crystals.

1.5.1. Protein crystals

1.5.1.1. Crystal symmetry

Protein crystals are arrays of protein molecules orderly arranged, which result from the translation of a repeating building block in the three dimensions, denominated as the unit cell. The unit cell can still be divided by motifs that are related by symmetry operations and constitute the asymmetric unit (Fig. 1.7).

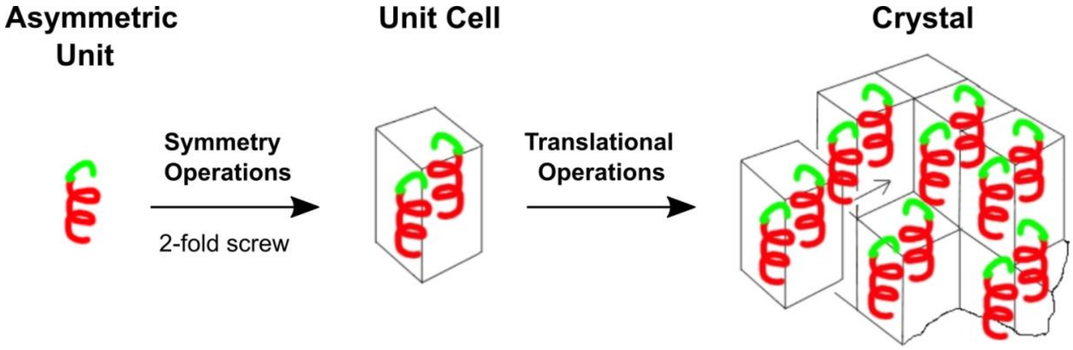
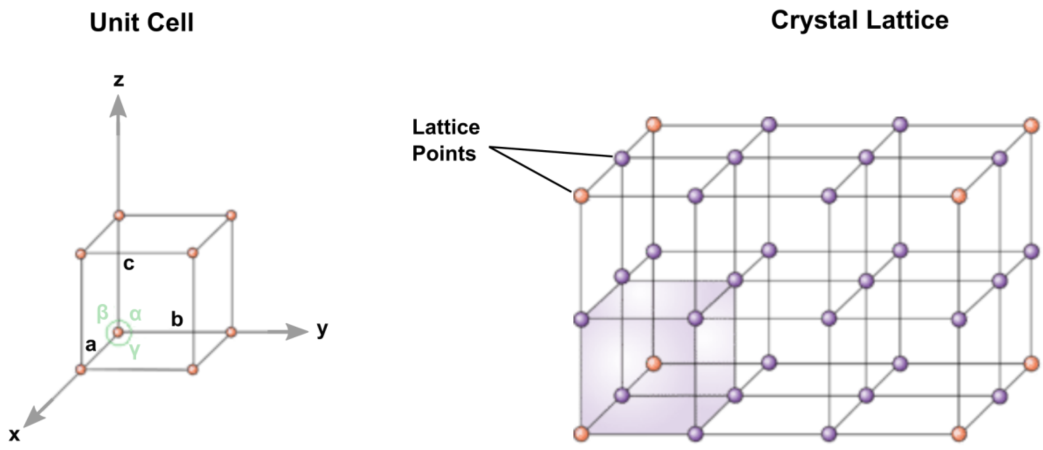


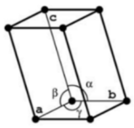
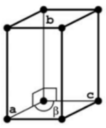
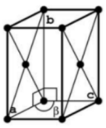
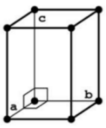
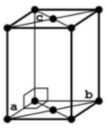
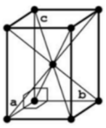
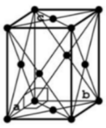
Figure 1.7. Protein crystals.
A structural motif with no internal symmetry forms the asymmetric unit of the crystal. This unit might include more than one protein molecule. Asymmetric motifs can form assemblies that are related by symmetry operations, such as a 2-fold screw (2-fold rotation + translation on the vertical plan). Such assemblies can constitute the unit cell, which establish the crystal through translational operations. Adapted from [93].

The unit cell is characterized by its axes a, b, c, and the angles (α, β, γ) between them (Fig. 1.8A). As the unit cell is translated in the three dimensions, the crystal can be described by the grid of equidistant parallel planes between unit cells, which is known as the crystal lattice. Depending on the specific combination of axes and angles of the unit cell, seven different crystal lattices can be formed: triclinic, monoclinic, trigonal, cubic, tetragonal, hexagonal or orthorhombic (Fig. 1.8A). In the three-dimensional space, a lattice point can be centered at the vertice of unit cells. The simplest unit cell type (denominated as primitive, P) contains a total of 1 lattice point (1/8 of a lattice points per vertice). Depending on the arrangement of the asymmetric units on the unit cell, some crystal lattices allow the presence of additional lattices points to be centered on one face (C), on more than one face (F) or at the center (I) of the unit cell. Auguste Bravais showed that there are 14 combinations of unit cell type and crystal lattices, known as the Bravais lattices (Fig. 1.8B).

A



B

Crystal Lattice	Triclinic	Monoclinic		Orthorhombic			
Unit Cell Dimensions	a≠b≠c α≠β≠γ	a≠b≠c α=γ=90°; β≠90°		a≠b≠c α=β=γ=90°			
Cell Type	P 	P 	C 	P 	C 	I 	F 
Space Groups	P1	P2, $P2_1$	C2	P222, $P2_12_12_1$ $P2_12_12$, $P222_1$	C222, $C222_1$	F222	I222, $I2_12_12_1$

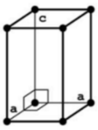
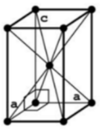
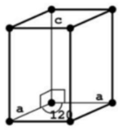
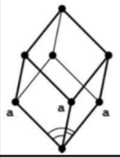
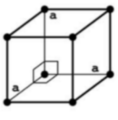
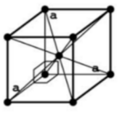
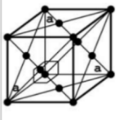
Crystal Lattice	Tetragonal		Hexagonal	Trigonal	Cubic		
Unit Cell Dimensions	a=b≠c α=β=γ=90°		a=b≠c α=β=90°; γ=120°	a=b≠c α=β=90°; γ=120°	a=b=c α=β=γ=90°		
Cell Type	P 	I 	P 	P 	P 	I 	F 
Space Groups	P4, $P4_1$, $P4_2$, $P4_3$, P422, $P42_12$, $P4_122$, $P4_12_12$, $P4_222$, $P4_22_12$, $P4_32_12$, $P4_322$	I4, $I4_1$, $I4_122$, I422	P6, $P6_5$, $P6_1$, $P6_4$, $P6_3$, $P6_2$, P622, $P6_122$, $P6_222$, $P6_322$, $P6_422$, $P6_522$	P3, $P3_1$, $P3_2$, P312, P321, $P3_121$, $P3_112$, $P3_221$, $P3_212$	P23, $P2_13$, P432, $P4_232$, $P4_132$, $P4_332$	I23, $I2_13$, I432, $I4_132$	F23, F432, $F4_132$

Figure 1.8. Symmetry of protein crystals.
(**A**) Schematic representation of a unit cell with its axes a, b and c and angles α, β and γ. Adapted from [94]. (**B**) Unit cells with specific combinations of axes and angles can form seven different crystal lattices. Their combination with the different unit cell type primitive (P), single face centered (C), multi-face centered (F) and body centered (I), sums up to the 14 Bravais lattices. Depending on the internal symmetry of the unit cell, 65 space groups are compatible with protein crystals. As of note, the trigonal crystal lattice can possess an alternative cell type setting, denominated as rhombohedral and with a=b=c and $\alpha=\beta=\gamma\neq 90°$, that includes the space groups R3 and R32, completing the total of 65 possible space groups for protein crystals. Adapter from [95].

The description of the crystal symmetry must still include the symmetry operations by which the asymmetric unit generates the unit cell. Given the chirality of amino acids (and hence, protein molecules), symmetry operations such as mirror planes and inversion centers are not possible because they would change the handedness of the symmetry-related molecule. Only rotations and screw axes are compatible in chiral molecules like proteins. The translational periodicity of the lattice restricts rotations and screw axes to 2-, 3-, 4-, and 6-fold operations. Combining translational lattice symmetry with internal unit cell symmetry, 65 different crystal symmetries are possible for proteins (Fig. 1.8B). The symmetry of the crystal is described by the space group, which is required for structure determination.

1.5.1.2. Protein crystallization

Protein crystals form when the protein concentration exceeds its solubility [96], a state denominated as supersaturated (Fig. 1.9A). Several strategies can be used to promote protein supersaturation, such as altering the temperature, pH or adding crystallization agents [97]. In the last case, salt and polymers can be added to the solution as precipitant agents to decrease protein solubility (Fig. 1.9A). In supersaturated conditions, protein molecules are more likely to interact and form aggregates that can become ordered protein assemblies and form crystal nuclei. The activation energy barrier associated to the nucleation process (Fig. 1.9B) can only be overcome when the protein concentration (and hence the supersaturation conditions) is large enough. The phase diagram showing the solubility curve of the protein in a specific crystallization condition (Fig. 1.9A) divides the supersaturation space associated to crystallization in two regions: labile and metastable. The assembly of multiple protein molecules as crystal nuclei occurs in the labile region, inducing the overall decrease of protein concentration in solution and pushing the equilibrium to the metastable region where free protein molecules can attach to the ordered nuclei and promote crystal growth.

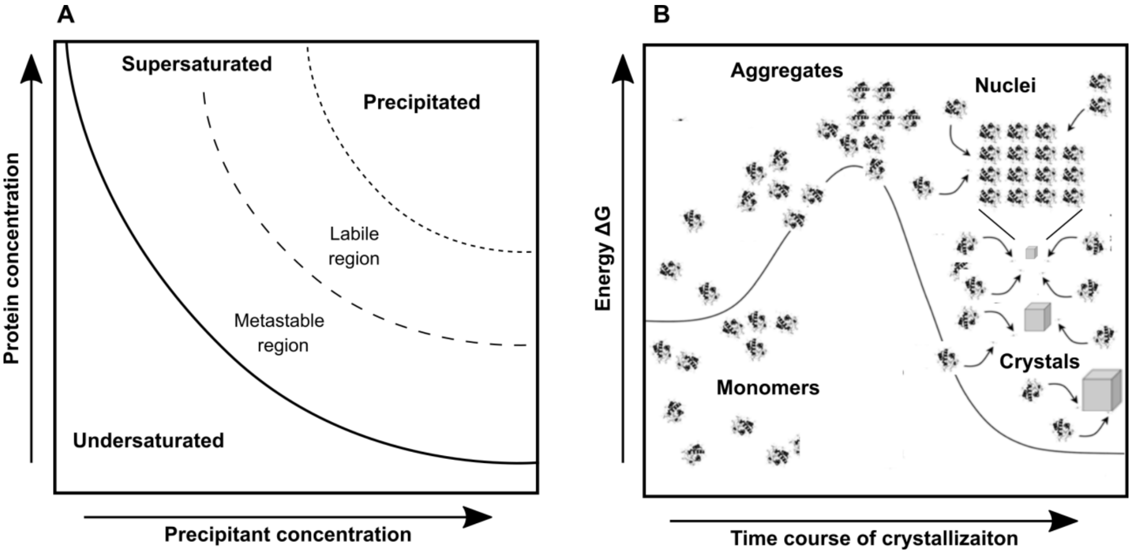


Figure 1.9. Principles of protein crystallization.
(**A**) Schematic phase diagram of protein crystallization. Supersaturation conditions of a protein, found between its undersaturated and precipitated states, can be achieved by increasing the concentration of either the protein or a precipitant agent. Protein nucleation can only occur at the high supersaturated conditions of the labile region, while crystal growth happens in both the labile and metastable regions. (**B**) Free energy of protein states occurring during crystallization. The formation of nuclei requires high free energy, which is only achieved in large protein supersaturated conditions (labile region of the phase diagram). Crystal growth requires lower free energy. Adapted from [98].

The main goal of crystallization methods is to promote the minimum protein solubility and, consequently, achieve supersaturated conditions [97]. The most used method to produce protein crystals is vapor diffusion, where precipitant agents are used to decrease protein solubility. In this method, a solution containing the precipitant agent is mixed with the protein solution and the resulting crystallization drop is allowed to equilibrate with the precipitant solution (reservoir solution) in a sealed chamber, using either the hanging or sitting drop techniques (Fig. 1.10). Since the concentration of the precipitant agent is higher in the reservoir solution than in the crystallization drop, water evaporates from the drop into the reservoir solution until the osmolarity of the precipitant is in equilibrium. The dehydration of the drop promotes concentration of both the protein and the precipitant agent, potentially reaching protein supersaturation conditions high enough for nucleation to occur (Fig. 1.9A). Alternative crystallization methods promoting protein supersaturation such as microdialysis, counter-diffusion or micro-batch under oil techniques can also be applied depending on the protein of interest, crystal size required and data collection strategy [97].

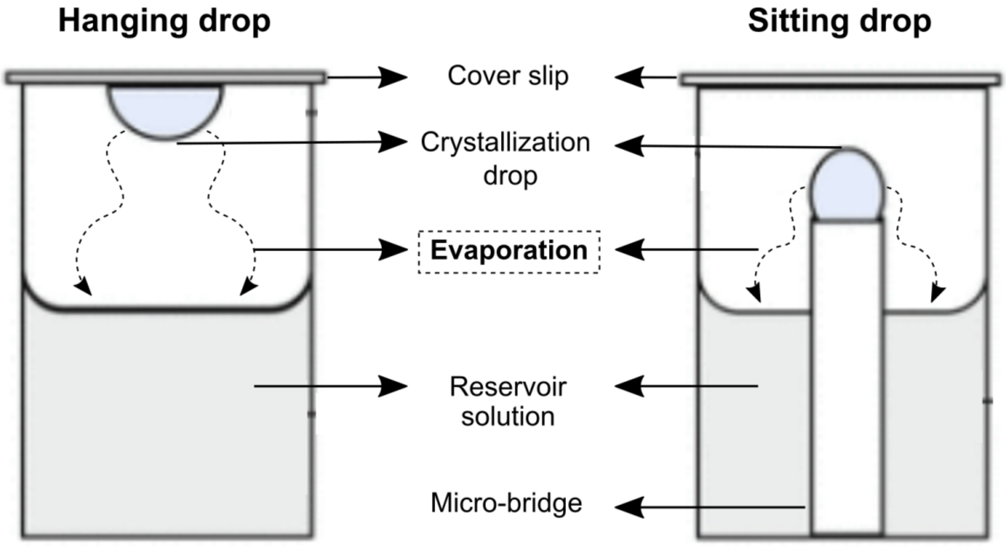


Figure 1.10. Protein crystallization by the vapor diffusion method.
A crystallization drop comprising the protein of interest and a precipitant agent is incubated in a sealed chamber with a reservoir solution that contains a higher precipitant concentration. Water evaporates from the drop to promote equilibrium in the osmolarity of the precipitant, resulting in the increases of the protein concentration and, potentially, crystallization. This method can be setup having the crystallization drop either hanging from the cover slip or sitting on a micro-bridge. Adapter from [99].

Crystallization is the bottleneck of structural biology since it is very difficult to predict the conditions that produce protein crystals [100]. Even though bioinformatic tools can be currently used to build constructs that are more likely to crystallize, protein crystallization still requires screening hundreds/thousands of conditions by varying the mixture of precipitants (most commonly salts, organic solvents and polymers as polyethylene glycol), the size of the crystallization drop, the protein:precipitant ratio, crystallization temperature, among other factors [97]. Screening is achieved by high-throughput approaches like pre-formulated crystallization solutions and liquid-handling robots that form and dispense 96 crystallization drops within 10 minutes [100]. Beside the fast setup, these approaches use fewer amounts of protein/reagents and ensure higher reproducibility of the crystallization outputs. In crystallization screens, precipitants can be combined randomly, based on conditions that have previously crystallized proteins (sparse matrix), or by refining conditions within a single group of molecules (for example, varying the concentration of a particular crystallization agent). Automation is mainly applied to vapour diffusion or batch methods due to its easy setup.

An important characteristic of protein crystals is their high solvent content that, on average, ranges from 30 to 70% of its total volume [101]. While this characteristic often represents a drawback in crystal handling since high solvent content increases crystal fragility, it has great potential for protein-ligand studies (enzyme substrate, products, inhibitors and modulators). Soaking the crystals with a solution of the ligand allows these small molecules to travel through the solvent channels of the crystal and bind to the protein molecules. Such protein-ligand crystallization technique Is favoured by binding pockets that are exposed to the solvent channels, otherwise interaction can promote a reorganization of the crystal lattice. Assemble of the protein-ligand complex in solution, followed by co-crystallization represents an alternative method to soaking [102].

After crystallization, protein crystals are harvested using µm thin cryo loops and flashed-frozen in liquid nitrogen (77 K). Freezing allows storing the crystals until data collection (performed at 100 K) and protects crystal lattice from major conformational oscillations of the protein, among other advantages [103]. Before freezing, protein crystals need to be cryoprotected, so the water present inside and around the crystal does not form ice, which can damage the crystal lattice and hinder the diffraction data. Cryoprotectant agents such as small polyols (most commonly used is glycerol), sugars, polyethylene glycol and oils, compete for water-water hydrogen bonds and, consequently, supress ice nucleation [104]. Once frozen, protein crystals can be used in X-ray diffraction experiments.

1.5.2. X-ray diffraction and data collection

X-rays are high-energy photons used for diffraction experiments in the range of 5,000-15,000 electron volts [105], corresponding to wavelengths of 0.1-1 nm (visible light photons have energies from 2 to 3 electron volts and wavelengths within 380-750 nm). Since photons behave like electromagnetic waves that oscillate upon interaction with electrons, the resulting X-ray wave can be diffracted in any direction. Crystals offer the advantage that diffraction from ordered arrays of atoms is amplified and the diffracting waves interfering constructively (in phase) are intense enough to be measured. This phenomenon of diffraction is mathematically described by the Bragg’s law [89] (Equation 1.1).

$$n\lambda = 2dsin\theta$$

(**Equation 1.1**)

The Bragg's law, schematically shown in Fig. 1.11A, dictates that constructive interference (maximum intensity) of crystal diffraction occurs when parallel planes distanced by $\boldsymbol{d}$ passing through crystal lattice points diffract waves at the same angle $\boldsymbol{\theta}$ as the incident radiation beam and the path length between them (incident and diffracted wave) is an integer multiple $\boldsymbol{n}$ of the wavelength $\boldsymbol{\lambda}$. For X-rays with angles not fulfilling Bragg's law, the scattered waves will not be in phase and interfere destructively with each other, providing no information. Reflections at high angles (large $\boldsymbol{\theta}$) are required to resolve closely spaced planes (small $\boldsymbol{d}$). Therefore, the resolution limit corresponds to the highest scattering angle at which reflections can still be measured. Given that the wavelength of X-rays is at the same magnitude as interatomic distances (e.g., a C-C bond is around 1.5 Å), individual atoms can be resolved in X-ray diffraction experiments.

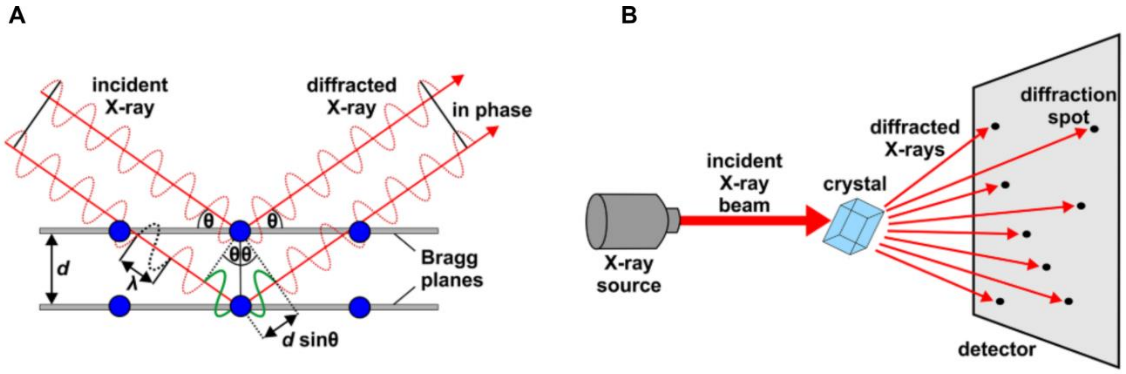


Figure 1.11. X-ray crystallography diffraction.
(**A**) Diffraction according to Bragg's law. Constructive interference (in phase) is only possible, if the path difference (**dsinθ**) between the waves (red dotted waves) is an integer multiple of the wavelength **λ**, which promote the summation of their amplitudes into a measurable signal on the detector. (**B**) Scheme of the diffraction experiment, in which an incident X-ray beam enters the crystal and the diffracted rays produce a diffraction pattern (diffraction spots) on a detector. Adapted from [106].

Under Bragg's law, each set of parallel planes of the crystal lattices can diffract a wave strong enough to be recorded by an appropriate detector as a spot (Fig. 1.11B). The resulting wave or diffraction spot has the contribution from all the individual waves diffracted by the electrons of each atom present in the unit cell. Depending on the angle that the X-rays hit the crystal, spots from different reflections are recorded in the form of diffraction pattern. Each reflection is assigned with Miller *h*, *k* and *l* indices, corresponding to integers of the intercepts that the reflecting parallel lattice plans do with the unit cell axes a, b and c, respectively.

The diffraction pattern represents the Fourier transform of the crystal structure, thus representing the reciprocal space. Here, the diffracted waves are described mathematically as vectors denominated structure factors **F(hkl)**, which represents the summation of the partial waves contributing to the reflection **hkl** by each and every **j** atom at position **x, y, z** of the unit cell with a scattering factor **fj** (Equation 1.2).

$$F(hkl) = \sum_{j=1}^{n} f_j \; e^{2\pi i(hx+ky+lz)}$$

(Equation 1.2)

The structure factor is a wavevector, thus being characterized by an amplitude dependent on the electrons contributing to the diffracted wave and a phase $\boldsymbol{\varphi}$ in relation to the incident X-rays (Equation 1.3).

$$F(hkl) = |F(hkl)| \cdot e^{i\varphi(hkl)}$$

(Equation 1.3)

Finally, the electron density throughout the unit cell of the crystal (Equation 1.4) can be calculated from the inverse integration of the Fourier transform to the structure factor.

$$\rho(xyz) = \frac{1}{V} \sum_{\substack{hkl \\ -\infty}}^{+\infty} |F(hkl)| \cdot e^{-2\pi i[hx+ky+lz-\varphi(hkl)]}$$

(Equation 1.4)

The main steps of a diffraction experiment are: the crystal is mounted on a goniometer head and placed under a nitrogen cryostream, which keeps the crystals at *ca* 100 K in order to minimize radiation damage from free radicals formed after the atoms of the protein interact with X-rays; the X-rays hit the crystal and the diffraction spots are recorded in the detector; the goniometer rotates the crystal and thousands of diffraction patterns are collected to ensure that all reflections are recorded. Different X-rays sources can be used, depending on the crystal size: diffraction of large and strongly diffracting crystals can be measured using in-house rotating anode X-ray generator, while small and weakly diffracting crystals must be measured at synchrotrons, where intense X-ray beams provide faster data collections and detection of less intense reflections such as those at higher resolutions.

Data processing using dedicated software involves estimating the crystal lattice and unit cell dimensions (as a space group) according to the organization of the spots in the diffraction patterns, assigning Miller indices to each reflection, attributing a common scale to all reflections and merging the measurements of equivalent reflections. However, determination of the electron density is not straightforward because the reflection intensities ***I(hkl)*** are only related to the amplitude of the structure factor (Equation 1.5).

$$\sqrt{I(hkl)} \propto |F(hkl)|$$

(Equation 1.5)

Electron density in the real space depends on both the amplitudes and phases of the diffracted waves (the structure factors in the reciprocal space). Experimental data does not contain information on the phases, leading to the phase problem. However, different phasing methods are available to achieve structure determination.

1.5.3. Structure determination

The methods to retrieve information of the lost phase are also denominated as structure determination methods since they provide an initial estimate of the phases, which can be used to calculate a first electron density map.

One of the most common methods to determine the structure of a protein is molecular replacement [107], which relies on the existence of a model from another protein that shares a sequence identity above 25% with the one at hand. This method calculates a Patterson map that considers the measured intensities as interatomic vector peaks rather than absolute atomic positions, thereof avoiding the determination of the phases. The Patterson map from the known model can be used as a template to align the peaks of the unknown structure. Rotation and translation functions provide the orientation and position of the unknown molecule in the unit cell. In order to avoid the introduction of bias, software such as *Chainsaw* [108] and *Sculptor* [109] are used to pre-process the template structure by for example pruning the side chains of residues that are not conserved with the protein under study. Bias from the search model can be assessed by identifying unique features in the determined structure.

Structure determination of proteins with no homology model requires the implementation of alternative methods, which mainly rely on the interaction of non-native protein atoms with X-rays.

Isomorphous replacement [91,110] rely on soaking or co-crystallization of a protein's crystal with heavy atoms (*eg*. mercury and platinum). Since these atoms contain more electrons than the native atoms of the protein, datasets from the native and heavy-atom derivative have different structure factor amplitudes. This difference represents an estimate of the heavy atom structure factor amplitude, which can be used by Patterson methods to determine its position in the unit cell. The obtained phase can then be used to estimate the phases of the atoms in the native crystal, under the condition that both crystals are isomorphous (have the same unit cell).

Phasing methods based on anomalous dispersion overcome the limitation of using several isomorphic crystals. Anomalous scattering happens when an inner electron absorbs radiation and jumps to higher energy levels, consequently affecting the amplitude and phase of its atomic scattering factor. The level of these absorption and dispersive effects depend on the radiation wavelength and has its maximum at the absorption edge of the anomalous scatterer. The anomalous difference can be used by Patterson methods to locate the anomalous scatterers and further estimate the phase for the remaining atoms of the unit cell.

These methods can use intrinsic anomalous scatterers present in the protein structure, such as the sulphur of cysteine and methionine residues or bound ions. Multiple-wavelength anomalous diffraction (MAD) [111] and single-wavelength anomalous diffraction (SAD) [112] are the two main strategies used. MAD comprises the collection of datasets at different wavelengths (close and far from the absorption edge of the scatterer) for the same crystal in order to maximize the difference in structure factor amplitudes, while SAD uses only the anomalous signal. MAD experiments provide better phase estimates but the crystal undergoes more radiation damage. The data collected in a SAD experiment requires high multiplicity and completeness in order to compensate the weak anomalous signal obtained from a single dataset. This limitation can be overcome by expressing recombinant proteins with seleno-methionines, for which *E. coli* strains lacking the ability to synthesize methionine amino acids are supplemented exclusively with the derivative seleno-methionine amino acid during protein overexpression. The higher anomalous diffraction of selenium in comparison to sulfur atoms provides better phase estimates.

Recently, the pipeline of structure determination undergone a major breakthrough with the development of AlphaFold [113], an artificial intelligence-based approach that predicts the structure of a protein based on its amino acid sequence. The high accuracy of these models can be directly used as search template for molecular replacement methods, providing solutions that require few further modifications.

All these phasing methods provide initial estimates for the phases of the different structure factors, which can be used to calculate the electron density maps.

1.5.4. Model building and refinement

The experimental phases obtained during structure determination are not accurate enough to provide a full interpretable electron density map. After structure determination, density modification methods are applied to improve the phase estimates, for example, by differentiating the electron density between solvent and protein atoms [112]. The improved electron density can then be used to assign the position of each atom present in the unit cell of the crystal. Model building can be automatically performed based on the sequence of the protein using dedicated software, such as *ARP/wARP* [114] and *Buccaneer* [115]. Molecular replacement methods have the advantage of providing an initial model of the protein based on the structure template used.

The accuracy of the experimental phases depends on the data resolution and phasing method. The working model is used to generate an additional set of structure factor amplitudes that contribute for better phase estimates and map interpretation. The electron density maps used in refinement combine structure factor amplitudes obtained from the experimental/observed data (F_{obs}) and calculated from the working model coordinates (F_{calc}). The $2F_{obs}$-F_{calc} map provides the best approximation to an experimental map and is used to evaluate model fitting. The difference F_{obs}-F_{calc} map highlights inconsistencies between the model and the experimental data, such as missing atoms (positive density usually displayed in green) or atoms wrongly modelled (negative density usually displayed in red) [116]. Potential bias introduced by using model-based structure factors in model building and refinement is minimized through maximum-likelihood σ-weighted maps (σ representing the error from the model-based structure factors). The $2F_{obs}$-F_{calc} and F_{obs}-F_{calc} maps are usually analysed as isosurfaces contoured at electron density levels above a standard deviation σ of 1 and 3, respectively. The presence of important structural features such as ligands can be evaluated using special types of maps, such as omit maps. In this methodology, the structural feature is omitted from the model and the difference map after refinement should show a clear positive density for the omitted atoms [117].

The electron density maps are manually interpreted by using computer graphics such as *coot* [118]. The usual workflow involves improving the overall fit of the model to the map filling empty electron density peaks with missing atoms, and assigning the position of ions, waters and other ligands. This refinement in the real space is then complemented with a software-based refinement in the reciprocal space. In this step, the fitting of the atomic coordinates to the electron density maps are improved, while the bond lengths and angles of the protein are restrained to values biologically relevant, commonly observed in protein structures obtained from data at high resolution. Application of this knowledge produces models biologically reasonable. Moreover, the fractional atomic occupancies of atoms in one or alternate positions, as well as, the vibration of an atom around its mean position (also denominated as atomic displacement parameter or B-factor) are also refined at this stage [116]. The parameterization of B-factors vary mainly with the resolution of the experimental data, ranging from the refinement of individual atoms within a sphere (isotropic) or ellipsoid (anisotropic) model to treating groups of atoms with the same displacement behaviour, such as a loop, a domain or an entire chain. The anisotropic refinement of rigid-body motions for groups of atoms is characteristic of the Translation, Libration, Screw (TLS) approach. When several molecules related by non-crystallography symmetry are present in the asymmetric unit, the refinement parameters of one molecule can be applied to all the others, which is advantageous for low resolution data [117].

The main goal of the reciprocal space refinement is minimizing the difference between the experimental structure factor amplitudes and those calculated from the working model, which can be assessed by the R factor (Equation 1.6). However, since during refinement the model-based structure factor amplitudes are used to improve the model, this factor is not independent. To overcome this bias, a set of 1-5% of the experimental reflections are randomly selected to be excluded from the refinement process and used to calculate an independent R factor called R_{free} (opposed to the R_{work} that is calculated with all the remaining reflections during structure refinement). Alternate cycles of real space and reciprocal space refinement are performed until the R factors converge to a minimum. Depending on the resolution of the experimental data, models associated to R factors of around 0.2 are considered of good quality. After obtaining R factor convergence and no major discrepancies between the model and the electron density maps, structure refinement is concluded.

$$R = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$$

(Equation 1.6)

1.5.5. Structure validation and publication

Refinement and model validation are intertwined. The model represents the interpretation of the electron density and must be validated regarding what is usually observed in nature. Several tools can be used to validate the determined model, like Molprobity software [119] or web services such as wwPDB validation [120]. Model quality is usually evaluated based on the geometry of the residues: the backbone atoms are analysed using the Ramachandram plot based on the torsion angles *phi* and *psi* of each residue, while the side chains usually must adopt the most common rotamers. Moreover, a good model must avoid stereochemical clashes.

Structural analysis is an important step of model validation, since comparison with structures of similar proteins might provide functional insights for the geometry errors of the model. Residues that are Ramachandran/rotamer's outliers might be present in the active site of an enzyme or surrounded by special stereochemical environments (*eg*. metal coordination), flexible loops without defined electron density can participate in substrate binding sites and adopt a defined structure only when the substrate is bound. Such analysis might also help identifying the nature of unknown electron density peaks that might correspond to ligands bound to the protein during expression, purification or crystallization.

The final model and respective electron density maps are deposited at the Protein Data Bank (PDB) for dissemination to the scientific community. An important note to have in mind is that the deposited model corresponds to the crystallographer's interpretation of the electron density, which might contain errors. The validation report is provided along with the model and the structure factors, which allows calculating the electron density maps for independent analysis. The structural data available at the PDB represents the starting point to explore the function of proteins, from ligand interactions (substrate, product, inhibitory drugs, *etc*) to protein-protein interaction of molecular machines.

1.6. References

1. *BAD BUGS, NO DRUGS: As Antibiotic Discovery Stagnates ... A Public Health Crisis Brews* (2004) Infectious Diseases Society of America.
2. Kirby WMM (1944) EXTRACTION OF A HIGHLY POTENT PENICILLIN INACTIVATOR FROM PENICILLIN RESISTANT STAPHYLOCOCCI. *Science (80-)* **99**, 452–453.
3. Watkins RR & Bonomo RA (2016) Overview: Global and Local Impact of Antibiotic Resistance. *Infect Dis Clin North Am* **30**, 313–322.
4. *Antibiotic Resistance Threats in the United States, 2019* (2019) Centers for Disease Control and Prevention.
5. *Antimicrobial Resistance: Global Report on Surveillance* (2014) World Health Organization.
6. *Thirteenth general programme of work 2019–2023* (2019) World Health Organization.
7. Shlaes DM, Sahm D, Opiela C & Spellberg B (2013) The FDA Reboot of Antibiotic Development. *Antimicrob Agents Chemother* **57**, 4605–4607.
8. Talbot GH, Jezek A, Murray BE, Jones RN, Ebright RH, Nau GJ, Rodvold KA, Newland JG & Boucher HW (2019) The Infectious Diseases Society of America's 10 × '20 Initiative (Ten New Systemic Antibacterial Agents FDA-approved by 2020): Is 20 × '20 a Possibility? *Clin Infect Dis an Off Publ Infect Dis Soc Am* **69**, 1–11.
9. *2019 ANTIBACTERIAL AGENTS IN CLINICAL DEVELOPMENT - an analysis of the antibacterial clinical development pipeline* (2019) World Health Organization.
10. *ANTIBACTERIAL AGENTS IN PRECLINICAL DEVELOPMENT - an open access database* (2019) World Health Organization.
11. *Prioritization of pathogens to guide discovery, research and development of new antibiotics for drug-resistant bacterial infections, including tuberculosis* (2017) World Health Organization.
12. Smeltzer MS, Lee CY, Harik N, Hart ME, Jefferson KK, Archer GL & Fowler VG (2009) Chapter 4 - Molecular Basis of Pathogenicity. In *Staphylococci in Human Disease -2nd edition* pp. 65–108.
13. Wertheim HFL, Melles DC, Vos MC, van Leeuwen W, van Belkum A, Verbrugh HA & Nouwen JL (2005) The role of nasal carriage in *Staphylococcus aureus* infections. *Lancet Infect Dis* **5**, 751–762.
14. Schukken YH, González RN, Tikofsky LL, Schulte HF, Santisteban CG, Welcome FL, Bennett GJ, Zurakowski MJ & Zadoks RN (2009) CNS mastitis: Nothing to worry about? *Vet Microbiol* **134**, 9–14.
15. Tenhagen B-A, Hansen I, Reinecke A & Heuwieser W (2009) Prevalence of pathogens in milk samples of dairy cows with clinical mastitis and in heifers at first parturition. *J*

Dairy Res **76**, 179–187.

16. Sakwinska O, Giddey M, Moreillon M, Morisset D, Waldvogel A & Moreillon P (2011) *Staphylococcus aureus* Host Range and Human-Bovine Host Shift. *Appl Environ Microbiol* **77**, 5908–5915.
17. Rasigade J-P, Laurent F, Hubert P, Vandenesch F & Etienne J (2010) Lethal Necrotizing Pneumonia Caused by an ST398 *Staphylococcus aureus* Strain. *Emerg Infect Dis* **16**, 1.
18. Lowder B V, Guinane CM, Zakour NL Ben, Weinert LA, Conway-Morris A, Cartwright RA, Simpson AJ, Rambaut A, Nübel U & Fitzgerald JR (2009) Recent human-to-poultry host jump, adaptation, and pandemic spread of *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* **106**, 19545–50.
19. Kluytmans J, Leeuwen W Van, Goessens W, Hollis R, Messer S, Herwaldt L, Bruining H, Heck M, Rost J, Leeuwen N Van, Belkum A Van & Verbrugh H (1995) Food-Initiated Outbreak of Methicillin-Resistant *Staphylococcus aureus* Analyzed by Pheno- and Genotyping. *J Clin Microbiol* **33**, 1121–1128.
20. Guo Y, Song G, Sun M, Wang J & Wang Y (2020) Prevalence and Therapies of Antibiotic-Resistance in *Staphylococcus aureus*. *Front Cell Infect Microbiol* **10**, 1–11.
21. Humphreys H (2012) *Staphylococcus aureus*: The enduring pathogen in surgery Hilary. *Surg* **10**, 357–360.
22. Tong SYC, Davis JS, Eichenberger E, Holland TL & Fowler VG (2015) *Staphylococcus aureus* infections: Epidemiology, pathophysiology, clinical manifestations, and management. *Clin Microbiol Rev* **28**, 603–661.
23. Styers D, Sheehan DJ, Hogan P & Sahm DF (2006) Laboratory-based surveillance of current antimicrobial resistance patterns and trends among *Staphylococcus aureus*: 2005 status in the United States. *Ann Clin Microbiol Antimicrob* **5**.
24. Flamm RK, Mendes RE, Ross JE, Sader HS & Jonesa RN (2013) Linezolid surveillance results for the united states: LEADER surveillance program 2011. *Antimicrob Agents Chemother* **57**, 1077–1081.
25. Brink AJ (2012) Does resistance in severe infections caused by methicillin-resistant *Staphylococcus aureus* give you the "creeps"? *Curr Opin Crit Care* **18**, 451–459.
26. Kalan L & Wright GD (2011) Antibiotic adjuvants: multicomponent anti-infective strategies. *Expert Rev Mol Med* **13**, 1–17.
27. Drawz SM & Bonomo RA (2010) Three Decades of Beta-Lactamase Inhibitors. *Clin Microbiol Rev* **23**, 160–201.
28. Roemer T, Schneider T & Pinho MG (2013) Auxiliary factors: A chink in the armor of MRSA resistance to β-lactam antibiotics. *Curr Opin Microbiol* **16**, 538–548.
29. Murakami K & Tomasz A (1989) Involvement of Multiple Genetic Determinants in High-Level Methicillin Resistance in *Staphylococcus aureus*. *J Bacteriol* **171**, 874–9.

30. Hartman BJ & Tomasz A (1986) Expression of methicillin resistance in heterogeneous strains of *Staphylococcus aureus*. *Antimicrob Agents Chemother* **29**, 85–92.

31. Novick RP, Edelman I, Schwesinger MD, Gruss AD, Swanson EC & Pattee PA (1979) Genetic translocation in *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* **76**, 400–404.

32. Lee SH, Jarantow LW, Wang H, Sillaots S, Cheng H, Meredith TC, Thompson J & Roemer T (2011) Antagonism of Chemical Genetic Interaction Networks Resensitize MRSA to β-Lactam Antibiotics. *Chem Biol* **18**, 1379–1389.

33. De Lencastre H, Wu SW, Pinho MG, Ludovice AM, Filipe S, Gardete S, Sobral R, Gill S, Chung M & Tomasz A (1999) Antibiotic resistance as a stress response: complete sequencing of a large number of chromosomal loci in *Staphylococcus aureus* strain COL that impact on the expression of resistance to methicillin. *Microb Drug Resist* **5**, 163–175.

34. Ogston A (1881) Report upon Micro-Organisms in Surgical diseases. *Br Med J* **1**, 369–375.

35. Matias VRF & Beveridge TJ (2007) Cryo-electron microscopy of cell division in *Staphylococcus aureus* reveals a mid-zone between nascent cross walls. *Mol Microbiol* **64**, 195–206.

36. Sieradzki K & Tomasz A (2003) Alterations of Cell Wall Structure and Metabolism Accompany Reduced Susceptibility to Vancomycin in an Isogenic Series of Clinical Isolates of *Staphylococcus aureus*. *J Bacteriol* **185**, 7103–7110.

37. Sobral R & Tomasz A (2019) The Staphylococcal Cell Wall. *Microbiol Spectr* **7**, 1–21.

38. Weidenmaier C & Lee JC (2017) Structure and Function of Surface Polysaccharides of *Staphylococcus aureus*. *Curr Top Microbiol Immunol* **409**, 57–93.

39. Matias VRF & Beveridge TJ (2006) Native Cell Wall Organization Shown by Cryo-Electron Microscopy Confirms the Existence of a Periplasmic Space in *Staphylococcus aureus* . *J Bacteriol* **188**, 1011–1021.

40. Matias VRF & Beveridge TJ (2005) Cryo-electron microscopy reveals native polymeric cell wall structure in *Bacillus subtilis* 168 and the existence of a periplasmic space. *Mol Microbiol* **56**, 240–251.

41. Rohde M (2019) The Gram-Positive Bacterial Cell Wall. *Microbiol Spectr* **7**, 1–21.

42. Bertani B & Ruiz N (2018) Function and biogenesis of lipopolysaccharides. *EcoSal Plus* **8**, 1–33.

43. Xia G, Kohler T & Peschel A (2010) The wall teichoic acid and lipoteichoic acid polymers of *Staphylococcus aureus*. *Int J Med Microbiol* **300**, 148–154.

44. Nguyen MT & Götz F (2016) Lipoproteins of Gram-Positive Bacteria: Key Players in the Immune Response and Virulence. *Microbiol Mol Biol Rev* **80**, 891–903.

45. Dreisbach A, Van Dijl JM & Buist G (2011) The cell surface proteome of *Staphylococcus aureus*. *Proteomics* **11**, 3154–68.

46. Geoghegan JA & Foster TJ (2015) Cell Wall-Anchored Surface Proteins of *Staphylococcus aureus*: Many Proteins, Multiple Functions. *Curr Top Microbiol Immunol* **409**, 95–120.

47. Neelamegham S, Aoki-Kinoshita K, Bolton E, Frank M, Lisacek F, Lütteke T, O'Boyle N, Packer N, Stanley P, Toukach P, Varki A, Woods RJ, Group & Discussion the S (2019) Updates to the Symbol Nomenclature For Glycans (SNFG) Guidelines. *Glycobiology* **29**, 620–624.

48. Boneca IG, Huang ZH, Gage DA & Tomasz A (2000) Characterization of *Staphylococcus aureus* Cell Wall Glycan Strands, Evidence for a New *β-N*-Acetylglucosaminidase Activity. *J Biol Chem* **275**, 9910–9918.

49. Gally D & Archibald AR (1993) Cell wall assembly in *Staphylococcus aureus*: proposed absence of secondary crosslinking reactions. *J Gen Microbiol* **139**, 1907–1913.

50. Kern T, Giffard M, Hediger S, Amoroso A, Giustini C, Bui NK, Joris B, Bougault C, Vollmer W & Simorre JP (2010) Dynamics Characterization of Fully Hydrated Bacterial Cell Walls by Solid-State NMR: Evidence for Cooperative Binding of Metal Ions. *J Am Chem Soc* **132**, 10911–10919.

51. Boneca IG, Huang Z, Gage DA & Tomasz A (2000) Characterization of *Staphylococcus aureus* Cell Wall Glycan Strands, Evidence for a New beta-*N*-Acetylglucosaminidase Activity. *J Biol Chem* **275**, 9910–9918.

52. Vollmer W, Blanot D & De Pedro MA (2008) Peptidoglycan structure and architecture. *FEMS Microbiol Rev* **32**, 149–167.

53. van Heijenoort Y, Gómez M, Derrien M, Ayala J & van Heijenoort J (1992) Membrane Intermediates in the Peptidoglycan Metabolism of *Escherichia coli*: Possible Roles of PBP 1b and PBP 3. *J Appl Crystallogr* **174**, 3549–3557.

54. Bouhss A, Crouvoisier M, Blanot D & Mengin-Lecreulx D (2004) Purification and Characterization of the Bacterial MraY Translocase Catalyzing the First Membrane Step of Peptidoglycan Biosynthesis. *J Biol Chem* **279**, 29974–29980.

55. Figueiredo TA, Sobral RG, Ludovice AM, de Almeida JMF, Bui NK, 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**.

56. Münch D, Roemer T, Lee SH, Engeser M, Sahl HG & Schneider T (2012) Identification and *in vitro* analysis of the GatD/MurT enzyme-complex catalyzing lipid II amidation in *Staphylococcus aureus*. *PLoS Pathog* **8**, 1–11.

57. Schneider T, Senn MM, Berger-Bächi B, Tossi A, Sahl H-G & Wiedemann I (2004) *In vitro* assembly of a complete, pentaglycine interpeptide bridge containing cell wall precursor (lipid II-Gly5) of *Staphylococcus aureus*. *Mol Microbiol* **53**, 675–685.

58. Sham L-T, Butler EK, Lebar MD, Kahne D, Bernhardt TG & Ruiz N (2014) MurJ is the

flippase of lipid-linked precursors for peptidoglycan biogenesis. *Science (80-)* **345**, 220–222.

59. Mohammadi T, Dam V Van, Sijbrandi R, Vernet T, Zapun A, Bouhss A, Bruin MD, Nguyen-Distèche M, Kruijff B de & Breukink E (2011) Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. *The* **30**, 1425–1432.

60. Taguchi A, Welsh MA, Marmont LS, Lee W, Sjodt M, Kruse AC, Kahne D, Bernhardt TG & Walker S (2019) FtsW is a peptidoglycan polymerase that is functional only in complex with its cognate penicillin-binding protein. *Nat Microbiol* **4**, 587–594.

61. Lovering AL, De Castro LH, Lim D & Strynadka NCJ (2007) Structural insight into the transglycosylation step of bacterial cell-wall biosynthesis. *Science (80-)* **315**, 1402–5.

62. Perlstein DL, Zhang Y, Wang T-S, Kahne DE & Walker S (2007) The direction of glycan chain elongation by peptidoglycan glycosyltransferases. *J Am Chem Soc* **129**, 12674–12675.

63. Wada A & Watanabe H (1998) Penicillin-Binding Protein 1 of Staphylococcus aureus Is Essential for Growth. *J Bacteriol* **180**, 2759–2765.

64. Pereira SFF, Henriques AO, Pinho MG, de Lencastre H & Tomasz A (2009) Evidence for a dual role of PBP1 in the cell division and cell separation of *Staphylococcus aureus*. *Mol Microbiol* **72**, 895–904.

65. Pinho MG, De Lencastre H & Tomasz A (2001) An acquired and a native penicillin-binding protein cooperate in building the cell wall of drug-resistant staphylococci. *Proc Natl Acad Sci U S A* **98**, 10886–10891.

66. da Costa TM, de Oliveira CR, Chambers HF & Chatterjee SS (2018) Pbp4: A new perspective on *Staphylococcus aureus* β-lactam resistance. *Microorganisms* **6**.

67. Monteiro JM, Covas G, Rausch D, Filipe SR, Schneider T, Sahl H-G & Pinho MG (2019) The pentaglycine bridges of *Staphylococcus aureus* peptidoglycan are essential for cell integrity. *Sci Rep* **9**, 1–10.

68. Tschierske M, Mori C, Rohrer S, Ehlert K, Shaw KJ & Berger-Bächi B (1999) Identification of three additional *femAB*-like open reading frames in *Staphylococcus aureus*. *FEMS Microbiol Lett* **171**, 97–102.

69. Henze U, Sidow T, Wecke J, Labischinski H & Berger-Bachi B (1993) Influence of femB on methicillin resistance and peptidoglycan metabolism in *Staphylococcus aureus*. *J Bacteriol* **175**, 1612–1620.

70. Strandén AM, Ehlert K, Labischinski H & Berger-BÄCHI B (1997) Cell wall monoglycine cross-bridges and methicillin hypersusceptibility in a *femAB* null mutant of methicillin-resistant *Staphylococcus aureus*. *J Bacteriol* **179**, 9–16.

71. Rohrer S, Ehlert K, Tschierske M, Labischinski H & Berger-Bachi B (1999) The essential *Staphylococcus aureus* gene *fmhB* is involved in the first step of peptidoglycan

pentaglycine interpeptide formation. *Proc Natl Acad Sci U S A* **96**, 9351–9356.

72. Koyama N, Tokura Y, Münch D, Sahl HG, Schneider T, Shibagaki Y, Ikeda H & Tomoda H (2012) The nonantibiotic small molecule cyslabdan enhances the potency of β-lactams against MRSA by inhibiting pentaglycine interpeptide bridge synthesis. *PLoS One* **7**.

73. Gustafson J, Strässle A, Hächler H, Kayser FH & Berger-Bächi B (1994) The *femC* Locus of *Staphylococcus aureus* Required for Methicillin Resistance Includes the Glutamine Synthetase Operon. *J Bacteriol* **176**, 1460–1467.

74. Strandén AM, Roos M & Berger-Bächi B (1996) Glutamine Synthetase and Heteroresistance in Methicillin-Resistant *Staphylococcus aureus*. *Microb Drug Resist* **2**, 201–207.

75. Boyle-vavra S, Labischinski H, Ebert CC, Ehlert K & Daum RS (2001) A Spectrum of Changes Occurs in Peptidoglycan Composition of Glycopeptide-Intermediate Clinical *Staphylococcus aureus* Isolates. *Antimicrob Agents Chemother* **45**, 280–287.

76. Nakel M, Ghuysen J-M & Kandler O (1971) Wall Peptidoglycan in *Aerococcus viridans* strains 201 Evans and ATCC 11563 and in *Gaffkya homari* strain ATCC 10400. *Biochemistry* **10**, 2170–2175.

77. Zapun A, Philippe J, Abrahams KA, Signor L, Roper DI, Breukink E & Vernet T (2013) *In vitro* Reconstitution of Peptidoglycan Assembly from the Gram-Positive Pathogen *Streptococcus pneumoniae*. *ACS Chem Biol* **8**, 2688–2696.

78. Pidgeon SE, Apostolos AJ, Nelson JM, Shaku M, Rimal B, Islam MN, Crick DC, Kim SJ, Pavelka MS, Kana BD & Pires MM (2019) L,D-Transpeptidase Specific Probe Reveals Spatial Activity of Peptidoglycan Cross-Linking. *ACS Chem Biol* **14**, 2185–2196.

79. Apostolos AJ, Pidgeon SE & Pires MM (2020) Remodeling of Crossbridges Controls Peptidoglycan Crosslinking Levels in Bacterial Cell Walls. *ACS Chem Biol* **15**, 1261–1267.

80. Xu HH, Trawick JD, Haselbeck RJ, Forsyth RA, Yamamoto RT, Archer R, Patterson J, Allen M, Froelich JM, Taylor I, Nakaji D, Maile R, Kedar GC, Pilcher M, Brown-Driver V, McCarthy M, Files A, Robbins D, King P, Sillaots S, Malone C, Zamudio CS, Roemer T, Wang L, Youngman PJ & Wall D (2010) *Staphylococcus aureus* TargetArray: Comprehensive differential essential gene expression as a mechanistic tool to profile antibacterials. *Antimicrob Agents Chemother* **54**, 3659–3670.

81. Chaudhuri RR, Allen AG, Owen PJ, Shalom G, Stone K, Harrison M, Burgis TA, Lockyer M, Garcia-Lara J, Foster SJ, Pleasance SJ, Peters SE, Maskell DJ & Charles IG (2009) Comprehensive identification of essential *Staphylococcus aureus* genes using Transposon-Mediated Differential Hybridisation (TMDH). *BMC Genomics* **10**, 1–18.

82. Liu X, Gallay C, Kjos M, Domenech A, Slager J, Kessel SP, Knoops K, Sorg RA, Zhang J & Veening J (2017) High-throughput CRISPRi phenotyping identifies new essential genes in *Streptococcus pneumoniae*. *Mol Syst Biol* **13**, 931.

83. Thanassi JA, Hartman-Neumann SL, Dougherty TJ, Dougherty BA & Pucci MJ (2002) Identification of 113 conserved essential genes using a high-throughput gene disruption system in *Streptococcus pneumoniae*. *Nucleic Acids Res* **30**, 3152–3162.

84. Long JE, Boutte CC, Rubin EJ, Schnappinger D, Ehrt S, Fortune SM, Sassetti CM & Ioerger R (2017) Comprehensive Essentiality Analysis of the *Mycobacterium tuberculosis* Genome. *MBio* **8**, e02133-16.

85. Schneider T, Kruse T, Wimmer R, Wiedemann I, Sass V, Pag U, Jansen A, Nielsen AK, Mygind PH, Raventós DS, Neve S, Ravn B, Bonvin AMJJ, De Maria L, Andersen AS, Gammelgaard LK, Sahl H-G & Kristensen H-H (2010) Plectasin, a fungal Defensin, Targets the Bacterial Cell Wall Precursor Lipid II. *Science (80-)* **328**, 1168–1172.

86. Figueiredo TA, Ludovice AM & Sobral RG (2014) Contribution of Peptidoglycan Amidation to Beta-Lactam and Lysozyme Resistance in Different Genetic Lineages of *Staphylococcus aureus* . *Microb Drug Resist* **20**, 1–12.

87. Röntgen WC (1896) On a new kind of rays. *Science (80-)* **3**, 227–231.

88. Friedrich W, Knipping P & Laue M (1913) Interferenzerscheinungen bei Röntgenstrahlen. *Ann Phys* **346**, 971–988.

89. Bragg PWH & Bragg WL (1913) The Reflection of X-rays by Crystals. *Proc R Soc London Ser A* **88**, 428–438.

90. Bernal JD & Crowfoot D (1934) X-Ray Photographs of Crystalline Pepsin. *Nature* **133**, 794–795.

91. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H & Phillips DC (1958) A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature* **181**, 662–8.

92. Jaskolski M, Dauter Z & Wlodawer A (2014) A brief history of macromolecular crystallography, illustrated by a family tree and its Nobel fruits. *FEBS J* **281**, 3985–4009.

93. Cristalografía y Biología Estructural CSIC (2022) Elements of a Protein Crystal. *https://www.xtal.iqfr.csic.es/Cristalografia/parte_04-en.html*.

94. CK-12 Foundation (2022) Unit cell and Crystal Lattice. *https://www.ck12.org/c/chemistry/unit-cells/lesson/Unit-Cells-CHEM/*.

95. Dauter Z & Jaskolski M (2010) How to read (and understand) Volume A of *International Tables for Crystallography*: an introduction for nonspecialists. *J Appl Crystallogr* **43**, 1150–1171.

96. Asherie N (2004) Protein crystallization and phase diagrams. *Methods* **34**, 266–272.

97. McPherson A & Gavira JA (2014) Introduction to protein crystallization. *Acta Crystallogr Sect FStructural Biol Commun* **70**, 2–20.

98. Nemcovicová I & Smatanová IK (2012) Alternative Protein Crystallization Technique:

Cross-Influence Procedure (CIP). In *Crystallization and Materials Science of Modern Artificial and Natural Crystals* pp. 249–276. InTech.

99. Ranatunga WK, Su D & Romero MF (2020) Obtaining crystals. In *Molecular Life Sciences* (Wells RD, Bond JS, Klinman J, & Masters BSS, eds), Springer, New York.
100. Shaw Stewart P & Mueller-Dieckmann J (2014) Automation in biological crystallization. *Acta Crystallogr Sect FStructural Biol Commun* **70**, 686–696.
101. Matthews BW (1968) Solvent content of protein crystals. *J Mol Biol* **33**, 491–497.
102. Rondeau JM & Schreuder H (2015) *Protein Crystallography and Drug Discovery* Elsevier Ltd.
103. Pflugrath JW (2015) Practical macromolecular cryocrystallography. *Acta Crystallogr Sect F Struct Biol Commun* **71**, 622–642.
104. Towey JJ, Soper AK & Dougan L (2013) What happens to the structure of water in cryoprotectant solutions? *Faraday Discuss* **167**, 159–176.
105. Rupp B (2010) *Biomolecular crystallography: principles, practice and applications to structural biology* Garland Science, New York.
106. Bijelic A & Rompel A (2018) Polyoxometalates: more than a phasing tool in protein crystallography. *ChemTexts* **4**, 1–27.
107. Rossmann MG & Blow DM (1962) The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallogr* **15**, 24–31.
108. Stein N (2008) *CHAINSAW*: a program for mutating pdb files used as templates in molecular replacement. *J Appl Crystallogr* **41**, 641–643.
109. Bunkóczi G & Randy J (2011) Improvement of molecular-replacement models with Sculptor. *Acta Crystallogr Sect D Biol Crystallogr* **67**, 303–312.
110. Perutz MF (1956) Isomorphous replacement and phase determination in non-centrosymmetric space groups. *Acta Crystallogr* **9**, 867–873.
111. Hendrickson WA & Ogata CM (1997) Phase determination from multiwavelength anomalous diffraction measurements. *Methods Enzymol* **276**, 494–523.
112. Wang B-C (1985) Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol* **115**, 90–112.
113. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P & Hassabis D (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.
114. Langer GG, Cohen SX, Lamzin VS & Perrakis A (2008) Automated macromolecular

model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* **3**, 1171–1179.

115. Cowtan K (2006) The *Buccaneer* software for automated model building. 1. Tracing protein chains. *Acta Crystallogr Sect D Biol Crystallogr* **62**, 1002–1011.

116. Lamb AL, Kappock TJ & Silvaggi NR (2015) You are lost without a map: Navigating the sea of protein structures. *Biochim Biophys Acta - Proteins Proteomics* **1854**, 258–268.

117. Shabalin IG, Porebski PJ & Minor W (2018) Refining the macromolecular model – achieving the best agreement with the data from X-ray diffraction experiment. *Crystallogr Rev* **24**, 236–262.

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

119. Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, III WBA, Snoeyink J, Adams PD, Lovell SC, Richardson JS & Richardson DC (2018) MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293–315.

120. Gore S, Sanz García E, Hendrickx PMS, Gutmanas A, Westbrook JD, Yang H, Feng Z, Baskaran K, Berrisford JM, Hudson BP, Ikegawa Y, Kobayashi N, Lawson CL, Mading S, Mak L, Mukhopadhyay A, Oldfield TJ, Patwardhan A, Peisach E, Sahni G, Sekharan MR, Sen S, Shao C, Smart OS, Ulrich EL, Yamashita R, Quesada M, Young JY, Nakamura H, Markley JL, Berman HM, Burley SK, Velankar S & Kleywegt GJ (2017) Validation of Structures in the Protein Data Bank. *Structure* **25**, 1916–1927.

CHAPTER 2

FIRST INSIGHTS OF PEPTIDOGLYCAN AMIDATION IN GRAM-POSITIVE BACTERIA - THE HIGH-RESOLUTION CRYSTAL STRUCTURE OF *STAPHYLOCOCCUS AUREUS* GLUTAMINE AMIDOTRANSFERASE GATD

2.1. Context

This chapter corresponds to the publication with reference Sci Rep. 2018 Mar 28;8(1):5313. doi: 10.1038/s41598-018-22986-3 that contains the results from the collaboration with the Laboratory of Molecular Microbiology of Bacterial Pathogens @ FCT-UNL led by Prof. Rita Sobral.

S. aureus GatD had been previously crystallized and its structure determined by Dr. Diana Vieira under the supervision of Dr. José Trincão and Prof. Maria João Romão from the Macromolecular Crystallography Laboratory @ FCT-UNL. This initial work was published in Acta Crystallogr F Struct Biol Commun. 2014 May;70(Pt 5):632-5. doi: 10.1107/S2053230X14007298. Epub 2014 Apr 17.

This thesis started with the refinement of the *S. aureus* GatD crystal structure. I was involved in structure refinement and analysis, which allowed me to propose the relevance of specific residues on catalysis. To confirm these hypotheses, I performed site-directed mutagenesis and produced the full length MurT-GatD complex in the wild-type and mutant versions. Initial conditions of protein production were obtained in collaboration with the *Pneumococcus* Group (Dr. André Zapun and Dr. Thierry Vernet) at the Institut de Biologie Structurale in France, funded by Instruct-ERIC. ^{1}H-NMR data collection and analysis was done by the NMR group at FCT-NOVA (Dr. Micael Silva and Prof. Eurico Cabrita).

2.2. Front page of the manuscript

First insights of peptidoglycan amidation in Gram-positive bacteria - the high-resolution crystal structure of *Staphylococcus aureus* glutamine amidotransferase GatD

Francisco Leisico[1¶], Diana Vieira[1,2¶], Teresa A. Figueiredo[1,3], Micael Silva[1], Eurico J. Cabrita[1] Rita G. Sobral[1], Ana Madalena Ludovice[1], José Trincão[4], Maria João Romão[1], Hermínia de Lencastre[3,5*] and Teresa Santos-Silva[1*]

[1]UCIBIO, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal; [2]Oxford Protein Production Facility, Research Complex at Harwell, Didcot, United Kingdom; [3]Laboratory of Molecular Genetics, Microbiology of Human Pathogens Unit, Instituto de Tecnologia Química e Biológica António Xavier da Universidade Nova de Lisboa, Oeiras, Portugal; [4]Diamond Light Source, Didcot, United Kingdom; [5]Laboratory of Microbiology and Infectious Diseases, The Rockefeller University, New York, USA

Email: tsss@fct.unl.pt and lencash@mail.rockefeller.edu

[¶]These authors contributed equally to this work.

2.3. Abstract

Gram-positive bacteria homeostasis and antibiotic resistance mechanisms are dependent on the intricate architecture of the cell wall, where amidated peptidoglycan plays an important role. The amidation reaction is carried out by the bi-enzymatic complex MurT-GatD, for which biochemical and structural information is very scarce. In this work, we report the first crystal structure of the glutamine amidotransferase member of this complex, GatD from *Staphylococcus aureus*, at 1.85 Å resolution. A glutamine molecule is found close to the active site funnel, hydrogen-bonded to the conserved R128. *In vitro* functional studies using ^{1}H-NMR spectroscopy showed that *S. aureus* MurT-GatD complex has glutaminase activity even in the absence of lipid II, the MurT substrate. In addition, we produced R128A, C94A and H189A mutants, which were totally inactive for glutamine deamidation, revealing their essential role in substrate sequestration and catalytic reaction. GatD from *S. aureus* and other pathogenic bacteria share high identity to enzymes involved in cobalamin biosynthesis, which can be grouped in a new sub-family of glutamine amidotransferases. Given the ubiquitous presence of GatD, these results provide significant insights into the molecular basis of the so far undisclosed amidation mechanism, contributing to the development of alternative therapeutics to fight infections.

2.4. Introduction

Staphylococcus aureus is considered one of the most important human Gram-positive bacterial pathogens due to its capacity to acquire and develop antibiotic resistance, causing high levels of mortality both in hospital and community settings [1]. Methicillin-resistant *S. aureus* (MRSA) are resistant to all ß-lactam antibiotics and show higher propensity to accumulate resistance to other classes of antibiotics. Fighting *S. aureus* is an urgent task that requires finding new drugs for new targets [2]. In this quest, peptidoglycan (PG) biosynthesis plays an important role since PG is essential for bacterial survival, cell shape maintenance and turgor pressure counterbalance [3].

PG biosynthesis is a complex process that takes place in several sequential enzymatic steps [2] (Fig. S2.1). The first step of PG biosynthesis occurs in the cytoplasm, where the nucleotide carbohydrate UDP-MurNAc (UDP-N-acetyl muramic acid) is covalently linked to a small pentapeptide (L-alanine-D-γ-glutamate-L-lysine-D-alanine-D-alanine) forming UDP-MurNAc-pentapeptide. At the membrane level, an undecaprenyl-phosphate lipid carrier binds to the precursor molecule, generating lipid I (Fig. S2.1). Subsequently, a GlcNAc unit is transferred from UDP-GlcNAc (UDP-N-acetyl-glucosamine) to the MurNAc carbohydrate of lipid I, leading to the formation of lipid II (Fig. S2.1). This structure suffers several modifications, such as the addition of a pentaglycine bridge at the L-lysine residue of the pentapeptide, the O-acetylation of MurNAc carbohydrate and the amidation of the D-glutamate residue into D-isoglutamine (Fig. S2.1). Finally, the PG monomer is translocated across the cytoplasmic membrane and assembled into the growing PG, through transglycosylation and transpeptidation reactions [4] (Fig. S2.1).

The PG of *S. aureus* is characterized by a high degree of crosslinking and almost complete lack of carboxyl groups, due to the amidation of the D-glutamate residue of the stem peptide [5]. This amidation of PG structure has a major impact in β-lactam antibiotic resistance, as shown with the *femC* (*glnRA*) mutant of MRSA [6]. In vancomycin-resistant *S. aureus* strains, it was recently shown that the presence of D-isoglutamine in lipid II stem peptide enhances the binding affinity to chloroeremomycin and oritavancin [7].

Recently, MurT and GatD enzymes were identified as responsible for catalyzing glutamate amidation in *S. aureus* PG stem peptides (Fig. 2.1) [8,9]. The *murT* and *gatD* gene arrangement found in *S. aureus* is conserved among Gram-positive bacteria [8,9]. Genetic studies showed that the impairment of the MurT-GatD complex reduces the bacterial growth rate, resistance to β-lactam antibiotics and to lysozyme [9,10]. Moreover, *in vitro* experiments showed that the MurT-GatD complex catalyzes glutamate amidation of lipid II using L-glutamine as a direct amine donor, when adenosine 5′-triphosphate (ATP) and Mg^{2+} were present. Maximum lipid II amidation was achieved with a MurT:GatD molar ratio of 1:1, suggesting the formation of a heteromeric complex. Site-directed mutagenesis identified GatD C94 as an important residue for catalysis [8]. Similarly, Zapun and co-workers showed that lipid II amidation in *Streptococcus pneumoniae* strain R6 is catalyzed by MurT-GatD complex, which is essential for PG cross-linking and for cell viability [11].

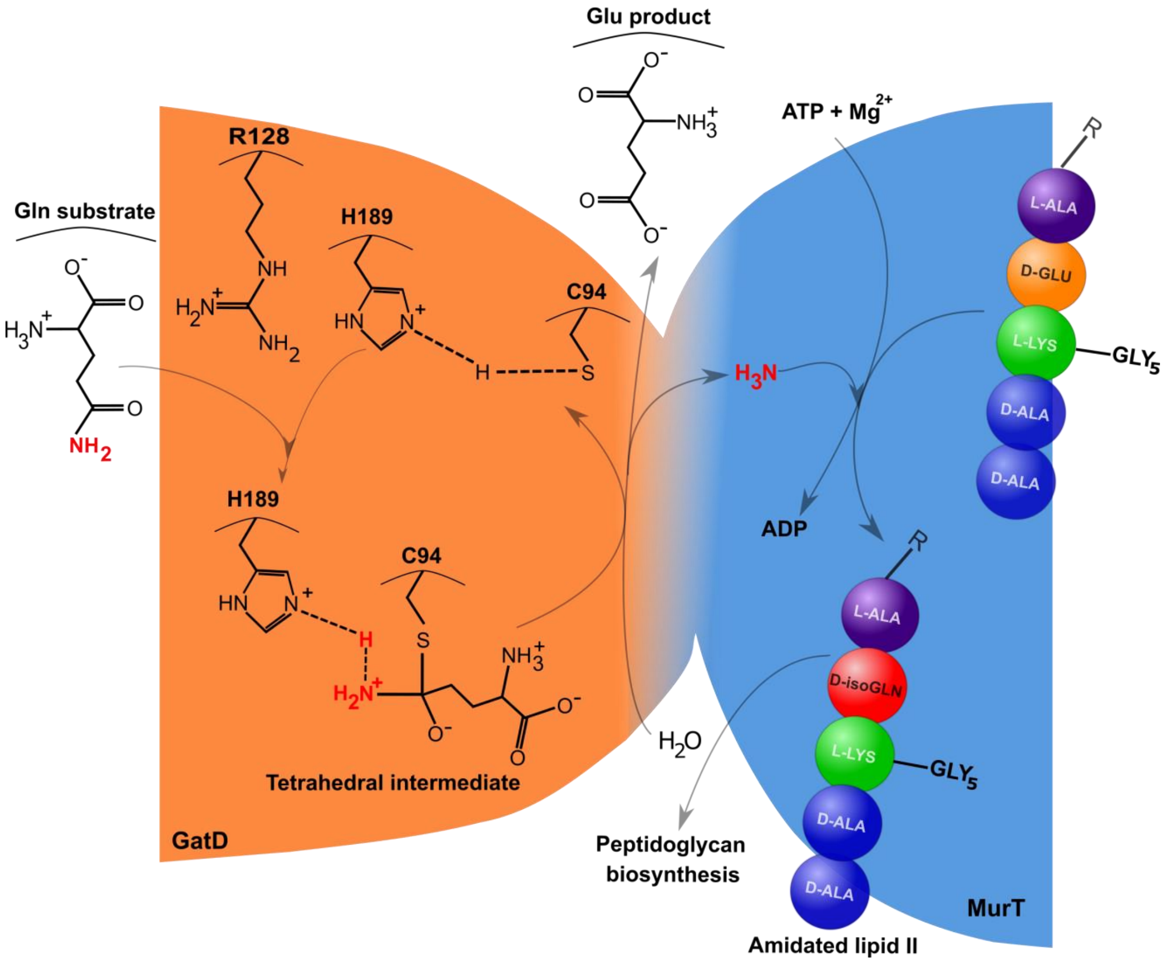


Figure 2.1. Schematic representation of lipid II amidation in *S. aureus*.

The MurT-GatD-catalysed amidation reaction in the PG biosynthetic pathway is divided in two stages: first, in GatD, deamidation of glutamine releases ammonia, which travels to MurT, where it is used to amidate the stem peptide residue D-Glu of lipid II to D-isoGln in the amidated lipid II.

MurT belongs to the family of Mur ligases, which are cytoplasmic enzymes responsible for the sequential addition of amino acid residues to the growing UDP-MurNAc precursor, containing the conserved motifs for ATP and Mg^{2+} binding. GatD shows high sequence homology to glutamine amidotransferase (GATase) enzymes, in particular to cobyric acid synthase, which is involved in cobalamin biosynthesis [9].

The amidation reactions are catalyzed by different domains of a single polypeptide chain or by different subunits of a protein complex and it can be divided into two steps. The first step is the glutaminase reaction, in which glutamine is converted into ammonia and glutamate. The NH_3 molecule is then transferred to the synthase domain/subunit through a solvent channel that prevents its protonation to ammonium. Finally, ammonia is added to a specific substrate, in a synthase reaction. Glutamine is used as an amide donor in several pathways that use amino acids, carbohydrates, nucleotides, co-enzymes and antibiotics as substrates [12].

Based on amino acid sequence analysis, GATases have been divided in two classes, with distinct protein folds and very different active sites. The enzymes that belong to class I have a conserved catalytic triad at the active site, formed by a cysteine, a histidine and a glutamate residue, and are usually called triad GATases. Structurally characterized triad GATases include: anthranilate synthase (AS), carbamoyl phosphate synthetase (CPS), cytidine triphosphate synthetase, formylglycinamidine ribonucleotide amidotransferase (FGAR-AT), guanosine monophosphate synthetase, imidazole glycerol phosphate synthase (IGPS) and pyridoxal 5'-phosphate synthase (PLPS) [13]. In class II enzymes, the main characteristics are the position of the essential catalytic cysteine at the N-terminal and no recognizable catalytic triad in the sequence [12,14]. Asparagine synthetase B, glucosamine-6-phosphate synthase, glutamate synthase and glutamine phosphoribosylpyrophosphate amidotransferase are examples of structurally characterized class II GATases [13].

A reaction mechanism has been proposed for class I triad GATases, based on structural evidences [13]. Generally, the glutamate residue is hydrogen-bonded to the side chain of the catalytic histidine, restricting its rotation and inducing its polarization. The histidine imidazole ring can then deprotonate the nucleophilic cysteine, increasing its reactivity. The catalytic cysteine, in the form of a thiolate, attacks the carbonyl carbon atom of the GATase substrate, the glutamine carboxamide group, forming a tetrahedral intermediate [15,16]. The stabilization of this negatively charged intermediate is achieved through hydrogen bonds with main chain atoms at the conserved "oxyanion hole". After the protonation of the amino group in the tetrahedral intermediate, an ammonia molecule is released via a channel to the synthase domain. Finally, the former glutamine/new glutamate is released from the glutamyl-thioester via an acid/base catalysis of a nucleophile attacking water [12].

In this work, we combined X-ray crystallography and Nuclear Magnetic Resonance (NMR) spectroscopy to study the glutaminase activity of *S. aureus* strain COL GatD. The crystal structure was solved at 1.85 Å resolution, unraveling singular features of GATase architecture and important residues at the active site. Site directed mutagenesis revealed the key amino acids involved in MurT-GatD glutaminase activity and substrate sequestration.

2.5. Methods

2.5.1. GatD expression, purification and crystallization

S. aureus GatD was cloned, expressed, purified and crystallized as previously described [17]. Briefly, the PCR amplified sequence of GatD was cloned into the pOPINF vector (Table S2.1) using the In-Fusion method [18]. The protein was expressed by the auto-induction method [19] in *E.coli* strain Lemo21(DE3) and purified in two chromatographic steps: IMAC and SEC. The crystallization screens were performed at a high-throughput facility (OPPF-UK) using a Cartesian instrument [17]. Native and Se-Met labeled protein crystals were obtained by vapor diffusion using polyethylene glycol as precipitating agent. The procedure from protein expression to crystallization is described in [17].

2.5.2. Data collection and structure determination

X-ray diffraction data of native GatD crystals were collected at temperature 100 K at beamline I02 (λ = 1.0000 Å), Diamond Light Source, UK. GatD crystallized in the space group $P2_12_12_1$ with unit cell dimensions of a = 48.61 Å, b = 93.92 Å, c = 110.08 Å, containing 2 molecules per asymmetric unit and a solvent content of 47%. The data was processed using Xia2 [20,21] through the automatic software pipeline available at Diamond Light Source. The phase problem was solved by SAD using a Se-Met derivative, as described in [17] . The experimental phases were later used to solve the structure of GatD in the native dataset (at 1.85 Å resolution). The structure was manually built using Coot [22] and refined using Phenix, considering restraint refinement with 15 TLS groups and isotropic B-factors [23]. After several rounds of manual rebuilding and refinement the R and R_{free} of the final model converged to 0.149 and 0.186. To avoid bulk solvent mask effect on the modeled glutamine molecule, an omit mFo-DFc difference map was generated for the glutamine using phenix. polder program [24], where the glutamine region within a radius of 10 Å was excluded (Fig. S2.6).

Automatic water picking followed by manual examination allowed identifying 625 water molecules, one polyethylene glycol molecule, and one glutamine molecule. The structure quality was validated using MolProbity [25]. Data collection and refinement statistics are summarized in Table 2.1. Atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB) with the accession code 5N9M.

Table 2.1. Data collection and refinement statistics.

The dataset was collected from a single crystal. *Highest resolution shell is shown in parenthesis.

	GatD-glutamine complex
Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c (Å)	48.61, 93.92, 110.08
α, β, γ (°)	90, 90, 90
Resolution (Å)	36.44–1.85 (1.89-1.85) *
R_{pim}	0.028 (0.313)
$I / \sigma I$	25.4 (5.2)
Completeness (%)	100.0 (99.9)
Redundancy	21.7 (20.1)
Refinement	
Resolution (Å)	36.44–1.85
No. reflections	42,922
$R_{\text{work}} / R_{\text{free}}$	0.149 / 0.186
No. atoms	
Protein	3,893
Ligand/ion	21
Water	625
B-factors	
Protein	27.43
Ligand/ion	43.47
Water	33.57
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.990
Protein geometry	
Poor rotamers	1 (0.24%)
Ramachandran outliers	0 (0.00%)
Ramachandran favored	467 (97.9%)
Cβ deviations > 0.25 Å	0 (0.00%)
Bad backbone bonds	0 / 3980 (0.00%)
Bad backbone angles	0 / 5369 (0.00%)

2.5.3. Sequence homology studies

The BLAST search (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was achieved using GatD sequence as query against UniProtKB/Swiss-Prot database in the standard protein blast (blastp). All multiple sequence alignments were performed using the ClustalW algorithm [26].

2.5.4. Cloning and production of *S. aureus* MurT-GatD *wt* and mutants

The coding regions for *S. aureus murT-gatD* genes were amplified by PCR from total DNA obtained from *S. aureus* strain COL (Table S2.1) and the flanking restriction sites *Nco I* and *Xho I* were introduced in the resulting DNA fragments by PCR primers. The amplified *murT-gatD* operon was cloned into the *Nco I – Xho I* sites of the expression plasmid pET28a (Table S2.1). The resulting plasmid pET28a-*murT-gatD*-His_6 was introduced into *E. coli* (DE3) CodonPlus RIPL (Stratagene).

Site-directed mutagenesis was used to produce three variants of MurT-GatD *wt*, where C94, H189 and R128 from GatD protein were altered to alanine. The base pair exchanges were introduced in *gatD* gene on plasmid pET28a-*murT-gatD*-His_6 by PCR mutagenesis using Phusion DNA Polymerase (Thermo Fisher Scientific) resulting in the plasmids pET28a-*murT-gatD*-His_6C94A, pET28a-*murT-gatD*-His_6H189A and pET28a-*murT-gatD*-His_6R128A (Table S2.1). All primers are shown in Table S2.2. Presence of the mutations was confirmed by sequencing.

2.5.5. *S. aureus* MurT-GatD *wt* and mutants expression and purification

Escherichia coli BL21-CodonPlus (DE3)-RIPL cells (Agilent) were used for heterologous expression of *S. aureus* MurT-GatD *wt* and mutants C94A, H189A and R128A, after transformation with plasmids pET28a-*murT-gatD*-His_6C94A, pET28a-*murT-gatD*-His_6H189A and pET28a-*murT-gatD*-His_6R128A, respectively. Cells were grown aerobically in LB medium at 310 K and then the media was supplemented with 0.5 mM isopropyl β-D-1-thiogalactopyranoside to induce protein expression at 293 K and 150 r.p.m. for 16 hours. The cells were harvested by centrifugation at 7500 g for 15 minutes. The cell pellet was resuspended in buffer A (100 mM Tris-HCl, pH 8.2, 500 mM NaCl, 10 mM $MgCl_2$, 5 mM 2-

mercaptoethanol) supplemented with 10 mM imidazole, 10 µg/mL DNaseI and ½ tablet of cOmplete™ ULTRA Tablets, Mini, EDTA-free, EASYpack Protease Inhibitor Cocktail (Roche). Cells lysis was completed by sonication and the resulting supernatant was applied onto a 5 mL HisTrap column (GE Healthcare). Bound protein was eluted using a linear gradient of imidazole. The fractions, where the presence of the complex MurT-GatD was confirmed by SDS-PAGE, were pooled and concentrated. The resulting sample was applied onto a Superdex 200 10/300 GL (GE Healthcare) pre-equilibrated with buffer A. The peak fractions were pooled and concentrated to a final concentration of 70 µM. Buffer exchange to 50 mM Tris-d11, 500 mM NaCl, 10 mM $MgCl_2$, 5 mM 2-mercaptoethanol in D_2O was performed using a PD MiniTrap G-25 (Ge Healthcare).

2.5.6. NMR experiments

All NMR experiments were performed at 293K using a Bruker Avance III 400 operating at 400.15 MHz for protons, equipped with a 5 mm BBO probe head. Enzyme kinetics was followed by ^1H NMR spectroscopy measuring the conversion of glutamine into glutamate by integration of their γ-CH_2 resonance over time. In a typical experiment, the MurT-GatD protein sample (*wt*, H189A, R128A or C94A) was added in a 5mm NMR tube containing a solution of glutamine and ^1H NMR spectra were recorded approximately every 12 minutes for 320 minutes. The initial glutamine concentration was 875 µM in a 99.9% D_2O buffered solution (pH meter reading pH 7.9, uncorrected for deuterium isotope effect, 50 mM Tris-d11, 500 mM NaCl, 10 mM MgCl2, 5 mM 2-mercaptoethanol, and 50 µM TSP) and the sample protein concentration was 35 µM (i.e. protein:glutamine ratio of 1:25). ^1H NMR spectra were acquired with 16 K data points and 256 scans in a spectral window of 6009.6 Hz centered at the water resonance (1880.5 Hz) with a $T_{1\rho}$ relaxation filter of 35 ms to suppress protein signals. Chemical shifts were calibrated relative to TSP as internal reference. NMR data was processed in Bruker TopSpin™ 3.5 and analysed with Bruker Dynamics Center™ 2.5. Errors were determined based on the signal to noise ratio.

2.6. Results/Discussion

2.6.1. Overall structure

GatD crystallized in $P2_12_12_1$ space group, with 2 molecules in the asymmetric unit, forming a non-physiological dimer via non-crystallographic symmetry operations (validated by visual inspection and using PISA software [27]). Molecule A (Fig. 2.2) corresponds to the full-length GatD protein with 243 amino acids, while molecule B lacks electron density for the last four residues of the polypeptide chain. A free glutamine molecule was found at the surface of GatD chain A. The GatD molecule has a globular shape with the approximate dimensions 50 x 54 x 33 Å^3.

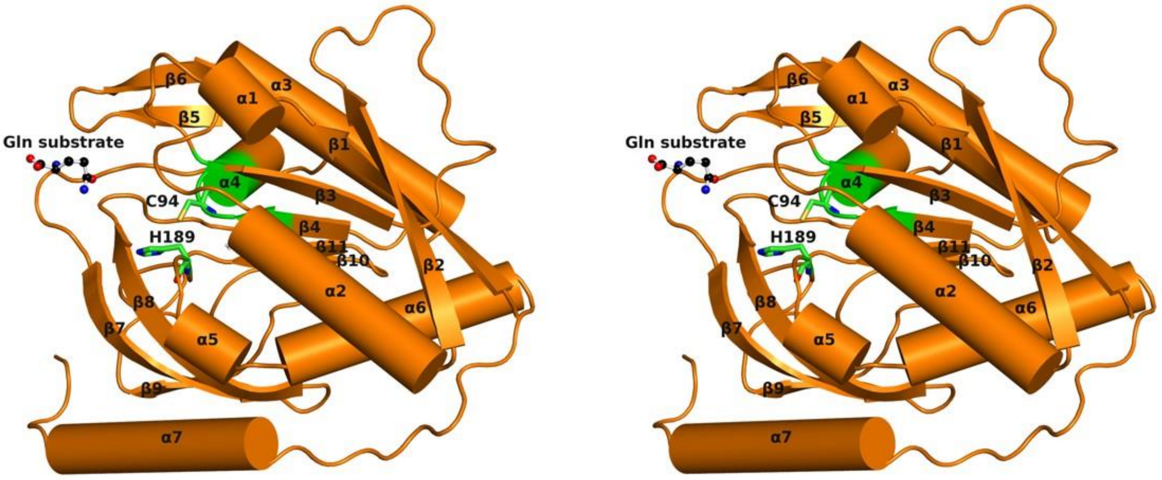


Figure 2.2. Ribbon representation of *S. aureus* GatD 3D structure.

GatD chain A with secondary structural elements labelled. The active site residues (C94 and H189) are represented as sticks with carbon atoms colored green and free glutamine as ball and stick with carbon atoms colored black. The structure is shown in wall-eyed stereo view.

The enzyme adopts an open α/β structure fold, similar to the glutaminase domain from a wide range of GATases [15]. Indeed, using the GatD model as input in the protein structure comparison service PDBeFold at European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/ssm) [28], several class I triad GATases were identified (Table 2.2). The most similar structure found is the *Thermotoga maritima* IGPS, with a RMSD of 2.58

Å (162 residues aligned) upon superposition, with an evident overall structure conservation (Fig. 2.3).

Table 2.2. Similarity of *S. aureus* GatD Structure with PDB Archive (PDBeFold).

The 3D alignment was performed in pairwise form using molecule A of *S. aureus* GatD structure. The target structures are ranked by RMSD from the 3D superposition of Cα-atoms. Protein names are abbreviated, as following: imidazole glycerol phosphate synthase, IGPS; piridoxal phosphate synthase, PLPS; formylglycinamidine ribonucleotide amidotransferase, FGAR-AT; anthranilate synthase, AS. Model 1I7Q is here included since it is used for structural comparison of triad GATases active site. [a]Identity includes the number of aligned residues with *S. aureus* GatD sequence and the percentage of identical residues in parenthesis.

Protein	Organism	Biosynthetic pathway	Protein length	Identity[a]	RMSD (Å)	PDB ID:chain	Reference
IGPS	*Thermotoga maritima*	Histidine	201	179 (16%)	2.58	3ZR4:B	[29,30]
PLPS	*Thermotoga maritima*	Vitamin B6	185	163 (18%)	2.61	2ISS:D	[31]
PLPS	*Geobacillus stearothermophilus*	Vitamin B6	202	176 (18%)	2.65	1Q7R:A	[31]
FGAR-AT	*Thermotoga maritima*	Purines	212	168 (16%)	2.67	3D54:D	[32]
PLPS	*Geobacillus kaustophilus* HTA426	Vitamin B6	188	171 (16%)	2.77	4WXY:B	[33]
PLPS	*Bacillus subtilis*	Vitamin B6	187	168 (16%)	2.80	2NV0:B	[34]
PLPS	*Plasmodium berghei*	Vitamin B6	217	179 (16%)	2.99	4ADS:I	[35]
PLPS	*Plasmodium falciparum*	Vitamin B6	216	178 (16%)	2.99	2ABW:A	[36]
AS	*Sulfolobus solfataricus*	Tryptophan	195	160 (14%)	3.13	1QDL:B	[37]
AS	*Serratia marcescens*	Tryptophan	195	193 (16%)	3.07	1I7Q:D	[38]

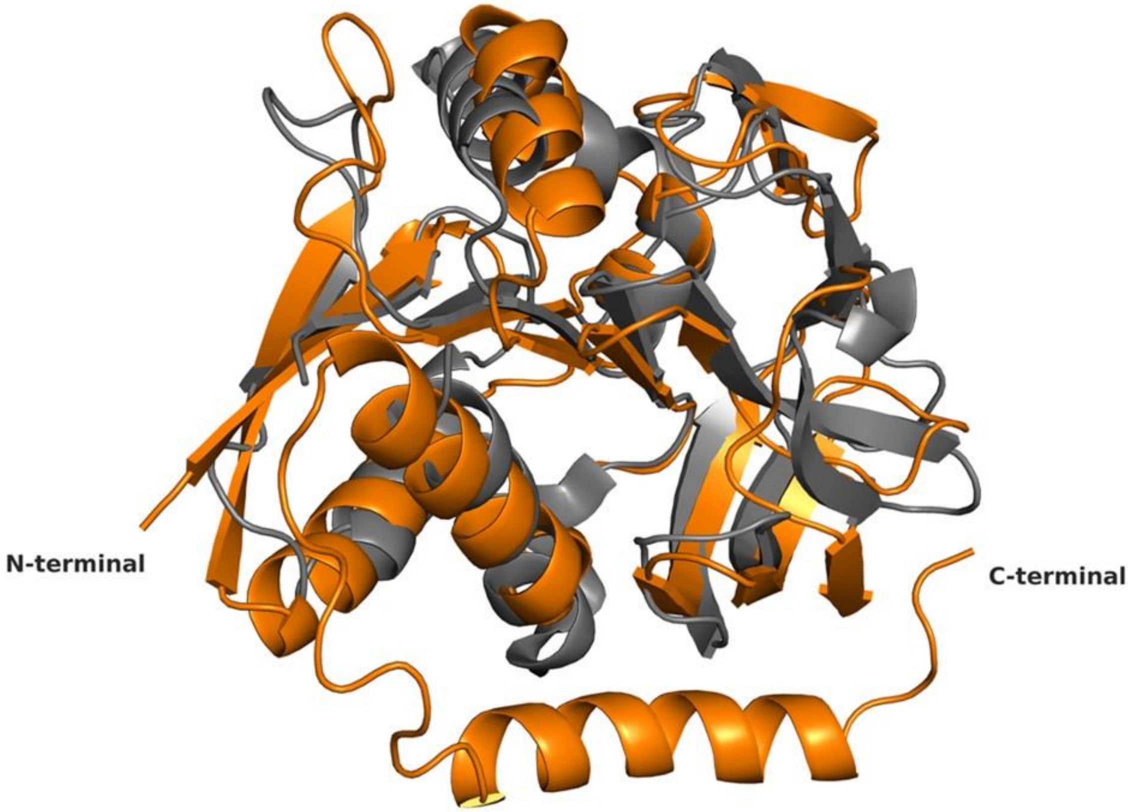


Figure 2.3. 3D structural alignment of *S. aureus* GatD and GATase domain of imidazole glycerol phosphate synthase.

Superposition of Cα-atoms of GatD chain A and IGPS from *T. maritima* (PDB ID 3ZR4) chain B. Both structures are represented as ribbons with GatD colored orange and IGPS dark grey.

The *S. aureus* GatD crystal structure shows the characteristic β-sheet core composed by six strands (β1, β2, β3, β4, β10 and β11), all parallel except for β10 (Fig. 2.2). This β-sheet core is flanked on one side by two helices (α3 and α4) and a two-stranded antiparallel β-sheets (β5 and β6), similar to what has been described for other GATase structures [39]. The other side of the β-sheet core is covered by a three-stranded antiparallel β-sheets (β7, β8 and β9) and by four α-helices (α1, α2, α5 and α6) (Fig. 2.2). Interestingly, GatD holds an extra 17 residues long α-helix (α7) at the C-terminal that is absent in all GATases characterized so far [39] (Fig. 2.3 and Fig. S2.2). This α-helix is located at the surface of the protein, as depicted in Fig. 2.2.

The 3D models of bi-enzymatic complexes selected for structural comparison show a common region for the interaction between the GATase and the corresponding synthase domain. Upon protein-protein interaction, conformational changes occur in the GATase domain that allow the substrate to reach the active site and the catalytic reaction to occur with formation of the ammonia shuttle [13]. The *S. aureus* GatD structure here reported was crystallized without its synthase domain (MurT), consequently, in an unproductive enzymatic form.

2.6.2. Active site

As expected, the active site of *S. aureus* GatD shows high structural similarity to the triad GATases previously mentioned. For comparison, the model of AS from *Serratia marcescens* (PDB ID 1I7Q) was selected since the glutamyl-thioester intermediate was captured in this crystal structure, highlighting the important residues for the reaction. The superposition (RMSD of 3.07 Å, 148 residues aligned) shows only slight differences in the locations of the catalytic residues between both models (Fig. 2.4a), probably due to conformational changes that occur upon substrate binding, in the presence of the synthase domain.

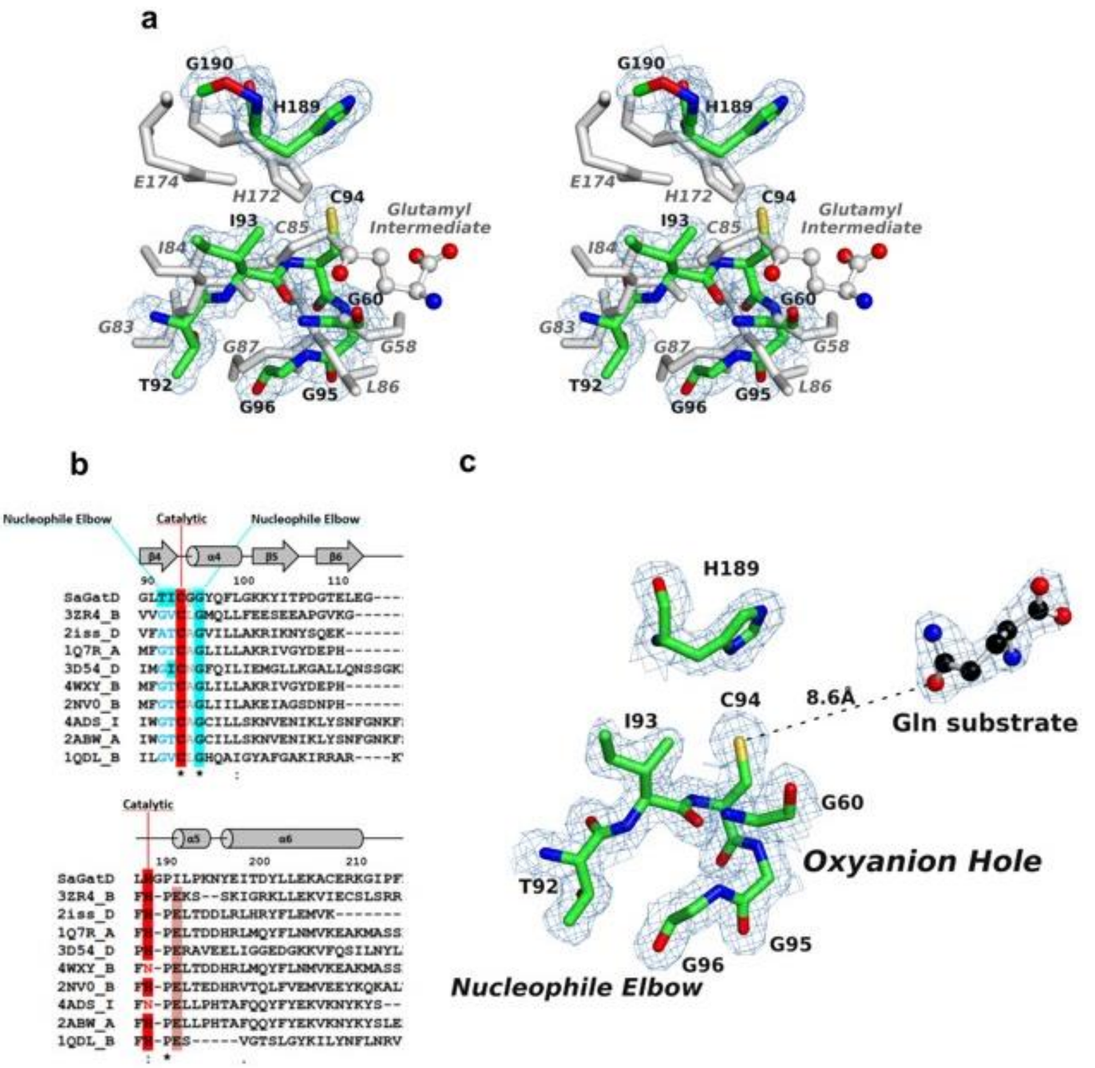


Figure 2.4. Representation of *S. aureus* GatD active site.

(**a**) Structural comparison of the catalytic sites of *S. aureus* GatD (in green) and Anthranilate Synthase (AS) from *S. marcescens* (in grey). The AS crystal structure (PDB ID 1I7Q) includes the glutamyl-thioester intermediate with C85. In AS, the G58 and L86 form the oxyanion hole; the catalytic triad is formed by the C85, H172 and E174; and the G83, I84 and G87 belong to the nucleophile elbow. The superposition is shown in wall-eyed stereo view and the 2Fo-Fc map is contoured at 1.0 σ. (**b**) Partial sequence alignment (using the ClustalW algorithm) of GatD and PDBeFold hits presented in Table 2.3 showing the sequence conservation of class I GATases active site, apart from the catalytic glutamate. The secondary structural elements from the GatD 3D structure are shown above the aligned sequences, with α-helices represented as cylinders and β-sheets as arrows. (**c**) GatD active site architecture, with the catalytic residues (C94 and H189), the nucleophile elbow and the oxyanion hole residues represented as sticks with carbon atoms colored green. The free glutamine molecule is represented as ball and stick with carbon atoms colored black. The 2Fo-Fc electron density map is contoured at 1.0 σ.

An important difference between the *S. aureus* GatD and other GATases so far characterized, is the absence of the glutamate residue that usually completes the catalytic triad (E174 in *S. marcescens* AS protein, Fig. 2.4a). In the structure here reported, glutamate is replaced by a glycine (G190), thereby eliminating any possible polarization effect upon the catalytic histidine H189. In fact, the catalytic role of the conserved glutamate in other GATases is not well understood. In CPS from *Escherichia coli*, site directed mutagenesis analysis revealed that the catalytic cysteine (C269) and histidine (H353) residues are essential for catalysis, but not the conserved glutamate (E355) [40]. Additionally, Hart, E.J., *et al.* showed that the structurally conserved catalytic triad of CPS can act as a functional dyad, since substitution of the neighboring residues of the catalytic histidine (Q310, N311, D334 and Q351) by alanine had no effect on the reaction [15]. IGPS from *T. maritima* is another example of a triad GATase, where biochemical data supported the non-catalytic role of the conserved glutamate [30]. All these evidences point towards the hypothesis that glutamine deamidation can be carried out by cysteine/histidine dyads, which is in agreement with the GatD crystal structure here reported. However, we cannot exclude the hypothesis that GatD residues at the active site may rearrange upon MurT-GatD complex formation (*in vitro* and/or *in vivo*), bringing another acidic residue into play and generating a catalytic triad. The glutamate conservation in class I triad GATase and its absence in *S. aureus* GatD could be related to the specific function of MurT-GatD in the context of PG synthesis pathway and should be further studied.

Additional features of GATases active sites, such as the oxyanion hole and the nucleophile elbow, affect glutamine deamidation, empowering the hydrolysis reaction. The oxyanion hole is formed by the amino acid residue next to the catalytic cysteine, commonly a leucine, and by a second residue adjacent to the so-called oxyanion strand [13] (G58 in *S. marcescens* AS protein, Fig. 2.4a). The backbone NHs of these residues are oriented in such a way that enable two hydrogen bond interactions with a single oxygen atom from the substrate. In *S. aureus* GatD the residues G95 and G60 of strand β3 are structurally aligned with *S. marcescens* AS, forming the oxyanion hole (Fig. 2.4a).

Generally, in triad GATases, the catalytic cysteine is located at the edge of a central strand of the β-sheet core and adjacent to an α-helix, forming a "nucleophile elbow" structural motif. The sequence of this motif is conserved among triad GATases and corresponds to G-X-C-X-G [41] (Fig. S2.2). The nucleophile elbow is flanked by the loop containing the catalytic histidine and glutamate, and by the oxyanion strand, inducing the cysteine to adopt a disallowed backbone conformation [42,43]. In the *S. aureus* GatD crystal structure here reported, the catalytic C94 is located between β4 and α4 (Fig. 2.2) with a Ramachandran disallowed conformation (Φ= 49.5° and Ψ= -114.0°), similarly to other GATase triads such as CPS [15] and AS [37]. Moreover, the GatD C94 shows very well-defined electron density. The sequence of the nucleophile elbow motif in *S. aureus* GatD, T92-I93-C94-G95-G96, is not fully conserved in comparison to other triad GATases since the first residue of the G-X-C-X-G motif (glycine) has been replaced by a threonine (Fig. 2.4a and b). However, the small size of the threonine residue will not hinder the required arrangement of the cysteine within the nucleophile elbow, for the catalytic reaction (Fig. 2.4c).

2.6.3. Glutamine sequestration

S. aureus GatD crystallized in the presence of its substrate, glutamine, which likely arose from the expression conditions of the protein, since it was not added during the purification/crystallization steps. The amino acid molecule is found at the surface of the protein, at 8.6 Å from the nucleophilic thiol group of C94 (Fig. 2.4c), only interacting with R128 from molecule A. The amide oxygen of the free glutamine is establishing hydrogen bonds with the guanidinium group of R128, with bond distances of 3.3 Å and 2.7 Å (Fig. 2.5). In this regard, R128 seems to be an important residue for sequestrating and accommodating the substrate at the surface of the protein before directing it to the catalytic site for hydrolysis. It is worth mentioning that, due to crystal packing effects, the glutamine binding site is accessible only in chain A. R128 of chain B is facing chain A, which prevents substrate binding.

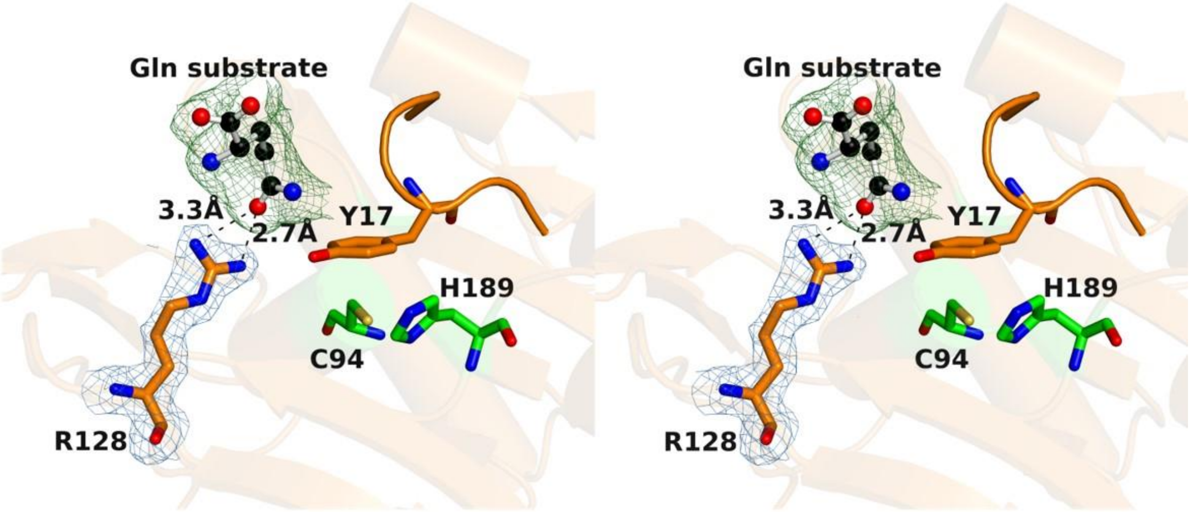


Figure 2.5. ***S. aureus*** **GatD-glutamine complex in wall-eyed stereo view.**

The glutamine substrate molecule is hydrogen-bonded to the protein surface residue R128, probably because the loop containing the Y17 amino acid is blocking its entrance to GatD active site. Y17 and R128 residues are represented as sticks with carbon atoms colored orange while glutamine is represented as ball and stick with carbon atoms in black. The catalytic dyad residues C94 and H189 are represented as sticks with carbon atoms colored green. The structure is shown in wall-eyed stereo view and the 2Fo-Fc map for R128 (in blue) contoured at 1.0 σ. Omit map mFo-DFc for glutamine is colored green and contoured at 3.0 σ.

Thoden, J.B., *et al.* [44] reported a crystal structure of the CPS enzyme from *E. coli* (PDB ID 1JDB), in which the glutamine substrate also did not enter the GATase active site, but was located at the surface, near the interface with the synthase domain. In this case, the carboxylic group of the glutamine main chain is hydrogen-bonded to two GATase arginine residues, R120 and R123. Additionally, in the structure of cytidine triphosphate synthetase from *Thermus thermophilus* (PDB ID 1VCO), the enzyme was crystallized in the presence of glutamine, which was found at the active site [39]. The authors claim that a hydrogen bond between the amine group of R472 main chain and the α-carboxylate group of glutamine is potentially important for substrate stabilization at the active site. Together, these results suggest that the presence of arginine residues can play an important role in glutamine sequestration for subsequent hydrolysis.

According to the related structures described in the literature, acceptor binding or complex assembly promote the reorganization of a loop at the surface of the glutaminase domain that allows glutamine entrance [13]. This loop may shield the GATase active site, preventing glutaminase reaction in the absence of the synthase domain. In *S. aureus* GatD, this corresponds to residues $_{14}$LNLYSD$_{19}$, which are flanked by helices $\alpha 1$ and $\alpha 2$. The GatD-glutamine complex here reported shows the Y17 side chain nearly co-planar to the H189 imidazole ring and establishing a hydrophobic interaction with C94 (Fig. 2.5). In this conformation, Y17 is obstructing the access of the free glutamine to the active site.

2.6.4. Structural insights into the *S. aureus* GatD sequence homologs

GatD has been assigned as belonging to the group of prokaryotic enzymes involved in cobalamin biosynthesis, as cobyrinic acid a,c-diamide synthase (CobB) or cobyric acid synthase (CobQ) [9]. CobB [45] and CobQ [46] have been mostly studied in *Pseudomonas denitrificans*. CobB has glutaminase activity, catalyzing the conversion of cobyrinic acid to cobyrinic acid a,c-diamide [45], while CobQ is responsible for the amidation reaction of 5′-deoxy-5′-adenosyl-cobyrinic acid a,c-diamide to cobyric acid [46]. Both proteins are organized in two domains, the N-terminal synthase domain with an ATP-binding motif, and a GATase domain at the C-terminal region.

Using *S. aureus* GatD sequence as query in a BLASTP search retrieves a set of targets, all belonging to the CobQ family of proteins, with higher homology towards a probable cobyric acid synthase protein from *Methanocella arvoryzae* (Table 2.3). Differentiation between CobQ and CobB proteins is not clear since their functional and structural information is scarce and different views about domain organization have been discussed, namely including them in the class I triad GATase [43] or in a, so far, non-classified group of GATases [12]. Sequence alignment of CobB and CobQ proteins with triad GATases shows that only the catalytic cysteine and histidine residues are conserved [47]. The lack of the catalytic glutamate suggests that these enzymes can perform the glutaminase activity as functional dyads, as *S. aureus* GatD, or might have other residue to perform the role of glutamate in a potential triad [47].

Table 2.3. Top 10 Unique Hits from BLASTP Analysis of GatD Sequence.

[a]Identity includes the number of aligned residues with *S. aureus* GatD sequence and the percentage of identical residues.

Protein	Organism	Blast Score	Protein length	Identity	E_value	Uniprot
Probable cobyric acid synthase	*Methanocella arvoryzae*	63.5	487	202 (30%)	1.81^{-10}	Q0W1N4
Cobyric acid synthase	*Listeria monocytogenes*	62.8	511	188 (30%)	3.46^{-10}	Q8Y7R3
Cobyric acid synthase	*Listeria innocua*	60.8	511	155 (32%)	1.36^{-09}	Q92CK0
Probable cobyric acid synthase	*Methanosphaerula palustris*	60.1	487	191 (30%)	2.44^{-09}	B8GDE3
Cobyric acid synthase	*Listeria welshimeri*	60.1	511	153 (31%)	3.12^{-09}	A0AHV1
Probable cobyric acid synthase	*Methanococcus maripaludis*	57.4	492	158 (34%)	1.86^{-08}	A4FWW2
Cobyric acid synthase	*Desulfitobacterium hafniense*	57.4	514	142 (32%)	2.32^{-08}	Q24Q41
Probable cobyric acid synthase	*Methanopyrus kandleri*	56.6	494	125 (34%)	4.24^{-08}	Q8TVH5
Cobyric acid synthase	*Bacteroides fragilis*	55.8	495	187 (26%)	6.09^{-08}	Q64TD9
Probable cobyric acid synthase	*Methanothermobacter thermautotrophicus*	55.5	504	163 (32%)	9.04^{-08}	O26880

As expected, *S. aureus* GatD only aligns with the C-terminal domain of Cob proteins, which corresponds to the GATase domain of these enzymes (Fig. S2.3). Based on this sequence alignment, the important features in the GatD crystal structure here described seem to be conserved in CobB/CobQ proteins, mainly in CobQ. In addition to the catalytic dyad, the residues next to C94 and H189 (using GatD numbering) are conserved (Fig. S2.3), suggesting that the architecture of the active site of both CobQ and GatD are probably very similar. High sequence homology is also observed for the oxyanion hole and the nucleophile elbow (although, for CobQ proteins, the first residue of the motif is a glycine while in GatD it is a threonine). Interestingly, R128 of GatD aligns in most cases with positively charged lysine residues that might be responsible for substrate interaction prior to binding at the GATase active site. Both CobB and CobQ proteins show the presence of extra C-terminal amino acids that match the extra α-helix of GatD (Fig. S2.3) but are absent in all other GATases characterized so far (Fig. S2.2). According to secondary structure prediction software [48], these residues can also form an α-helix that will likely superimpose with the C-terminal region of GatD. Grounded on all these structural evidences, we suggest that the *S. aureus* GatD crystal structure might be the first representative of a third class of GATases, where CobQ and CobB enzymes are included.

2.6.5. Comparison between the *S. aureus* GatD and pathogenic Gram-positive homologs

Given the importance of PG amidation in cell viability, the *S. aureus* GatD structure can be used to derive valuable information for the MurT-GatD homologous systems in other pathogenic Gram-positive bacteria. Protein sequence alignment of *S. aureus* GatD and its homologs from *Staphylococcus epidermidis*, *Mycobacterium tuberculosis*, *S. pneumoniae* and *Streptococcus pyogenes* reveals, as expected, high homology with sequence identity of 86%, 36%, 40% and 43%, respectively (Fig. 2.6). The catalytic cysteine and histidine of *S. aureus* GatD are conserved in all the aligned proteins, for which the putative triad glutamate residue is also missing. The oxyanion hole and the nucleophile elbow are also highly conserved, suggesting that the architecture of the *S. aureus* GatD active site is preserved. Interestingly, *S. aureus* GatD R128 and extra C-terminal residues are conserved in all sequences. These results indicate that Gram-positive GatD proteins diverge from the well-known triad GATases. The conservation of the loop capping the active site in these systems, especially Y17, makes this region a good target for inhibition strategies, imprisoning the enzyme in a nonproductive conformation.

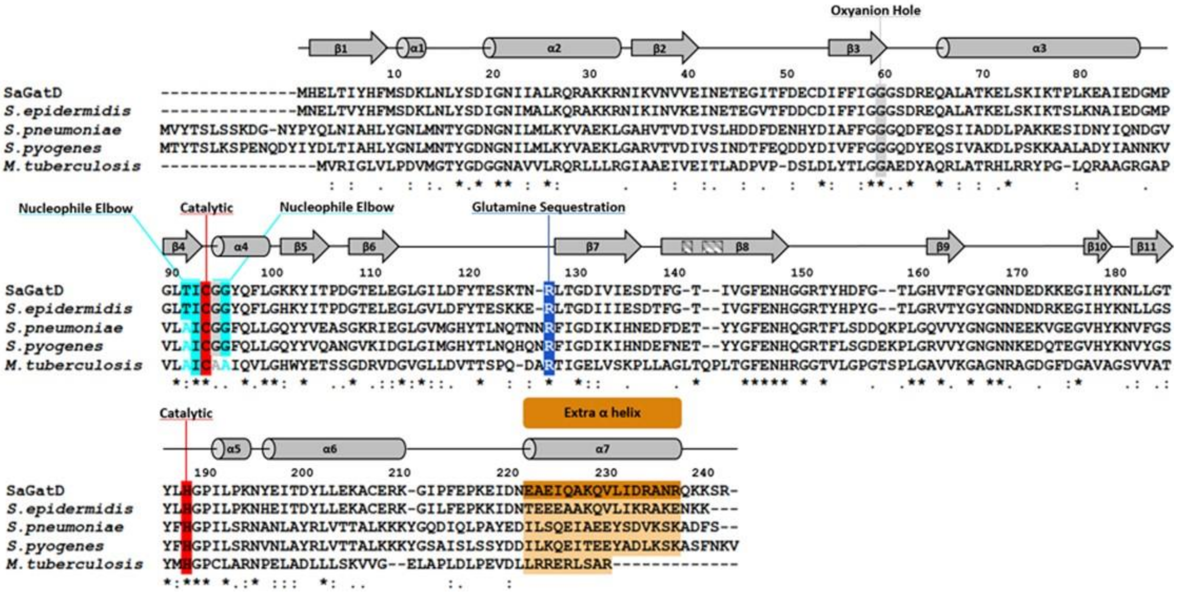


Figure 2.6. Multiple sequence alignment of GatD proteins from pathogenic Gram-positive bacteria.

S. aureus GatD (represented in the alignment as SaGatD) was aligned using the ClustalW algorithm with the amino acid sequences of GatD proteins from *S. epidermidis*, *S. pneumonia*, *S. pyogenes* and *M. tuberculosis* (with UNIPROT codes A0A0H2VH11, Q8DNZ8, Q1JH53 and I6XI14, respectively). Important GatD residues are highlighted such as the catalytic C94 and H189 dyad in red, oxyanion hole G60 and G95 in gray, nucleophile elbow T92, I93 and G96 in cyan, glutamine sequestration R128 in blue and the extra C-terminal α-helix in orange. The secondary structural elements are shown above the aligned sequences with α-helices represented as cylinders and β-sheets as arrows.

2.6.6. *In vitro* glutaminase activity of *S. aureus* MurT-GatD

The analysis of GatD crystallographic data provides strong evidence for the involvement of R128 in the glutamine sequestration, together with some important residues in the catalytic dyad, such as C94 and H189. To further validate this hypothesis, the cloning construct containing *murT* and *gatD* genes was modified to introduce point-mutations, replacing the highlighted amino acids by alanines in GatD protein. The resulting protein complexes MurT-GatD *wt* and mutants C94A, H189A, R128A were produced and their glutaminase activity evaluated using ^{1}H-NMR spectroscopy. Expression and purification of *S. aureus* MurT-GatD *wt* and the selected mutants led to similar yields of pure and homogeneous MurT-GatD protein complexes. The final samples showed a molar ratio of MurT-GatD of 1:1 (Fig. 2.7), confirming the heteromeric complex formation in solution, suggested by Munch, D., *et al*.[8]. The stability of the produced MurT-GatD variants show that the mutations imposed in GatD protein did not affect complex formation or solubility.

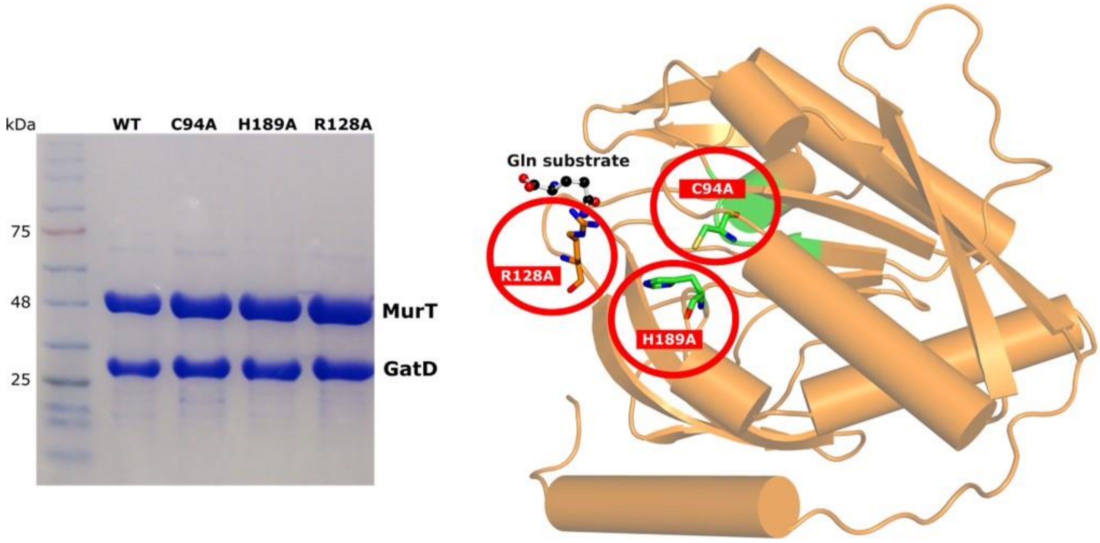


Figure 2.7. Purity of the heteromeric complex MurT-GatD.

12% SDS-PAGE gel of the final samples of *S. aureus* MurT-GatD *wt* and C94A, H189A and R128A mutants prior to enzymatic activity tests. Insert, *S. aureus* GatD cartoon highlighting the mutated residues that were prepared for further functional studies.

Glutaminase activity of *S. aureus* MurT-GatD *wt* and variants was followed by ^{1}H-NMR spectroscopy (Fig. S2.4). The NMR based activity assay is possible because glutamine and glutamate can be easily distinguished in the ^{1}H-NMR spectrum, allowing to follow and quantify the glutamine conversion into glutamate simply by recording and analyzing ^{1}H NMR spectra over time (spectra of MurT-GatD *wt* and H189A mutant in Fig. 2.8 and C94A and R128A spectra in Fig. S2.5). The clearly resolved resonances of glutamine and glutamate γ-CH$_2$ protons at 2.458 ppm and 2.348 ppm, respectively, were integrated to quantify the relative concentration of both amino acids in solution for each time point, as depicted in Fig. 2.9. These results show that MurT-GatD *wt* can convert glutamine into glutamate with a first order rate constant of $k = 0.0080 \pm 0.0001$ min^{-1}. This kinetic constant is in agreement with *T. maritima* IGPS glutaminase kinetics in the absence of the synthase substrate [30].

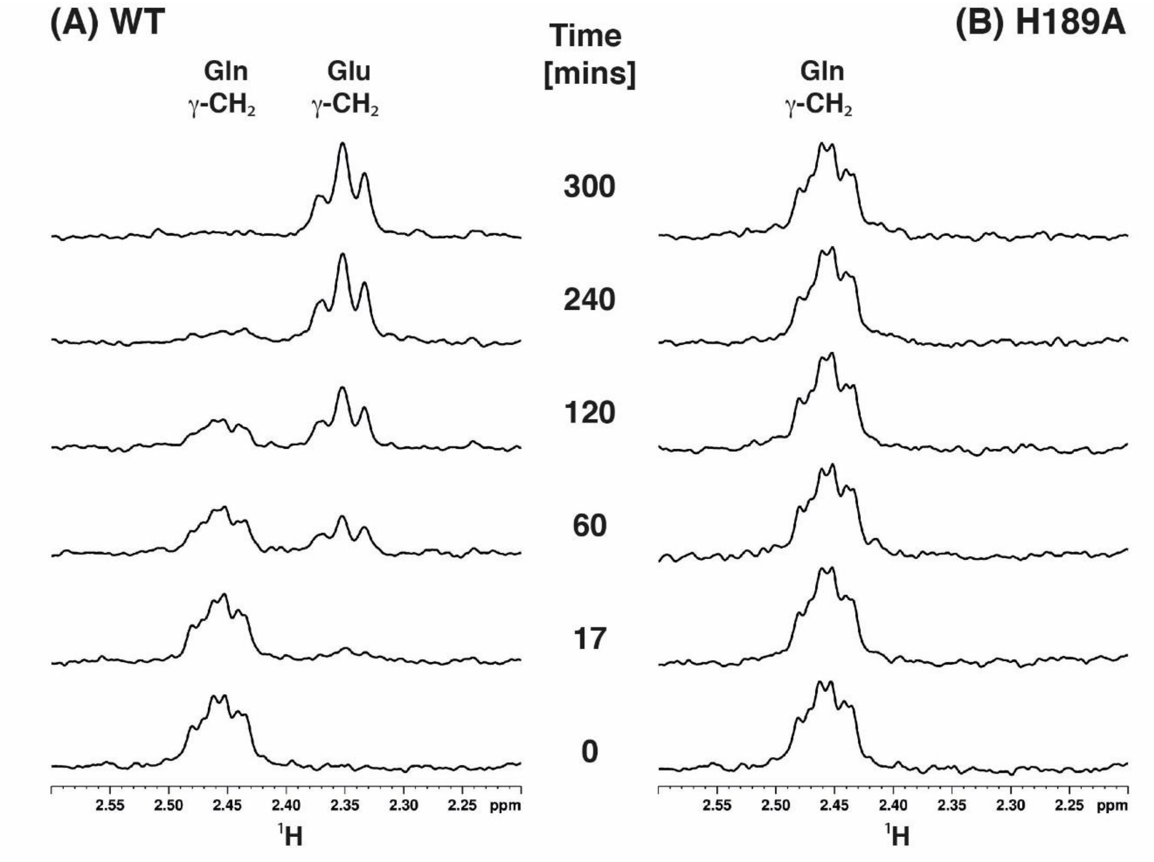


Figure 2.8. Glutaminase activity monitored by ^{1}H-NMR.

Expansions of the ^{1}H-NMR spectra of the Gln/Glu γ-CH_2 peak region with resonance assignment for different reaction times to monitor glutaminase activity for (**A**) *wt* and (**B**) H189A. Initial samples were 875 μM Gln with 35 μM MGH.

The glutaminase activity of the MurT-GatD *wt* complex shows that GatD is active in the absence of lipid II, which is the ammonia acceptor and MurT substrate. Since the ^{1}H NMR signals of glutamate can be detected, we conclude that the product of the glutaminase activity is released from the MurT-GatD complex, allowing a new glutamine molecule to be deamidated. During the catalytic cycle, the Y17 loop is likely displaced, promoting glutamine entrance to the GatD active site.

In the presence of lipid II, it is likely that the GatD active site will rearrange to a fully active state. The new position of the catalytic residues may affect the conversion rate of glutamine into glutamate. This hypothesis is supported by the 1000-fold increase in glutaminase kinetics of *T. maritima* IGPS when the synthase substrate is present [30].

The conversion of glutamine into glutamate could be followed in MurT-GatD *wt* glutaminase reaction, while the concentration of glutamine maintained constant along the reaction time for the mutant forms and no glutamate was detected in solution (Fig. 2.9). This confirms that GatD mutations in C94, H189 and R128 residues impair completely the glutaminase activity of the MurT-GatD complex. This establishes the selected residues as essential for the glutaminase activity. Similar studies performed on class I triad GATase CPS protein also showed that point mutation in the catalytic cysteine and histidine produced completely inactive protein forms [49,50]. These results reinforce the crystallographic data confirming the catalytic architecture of GatD active site, composed by these two residues. Furthermore, our results show univocally that R128 is determinant for glutaminase activity. The amine-based side chain must be detrimental to capture the substrate and drive it into the active site for catalysis.

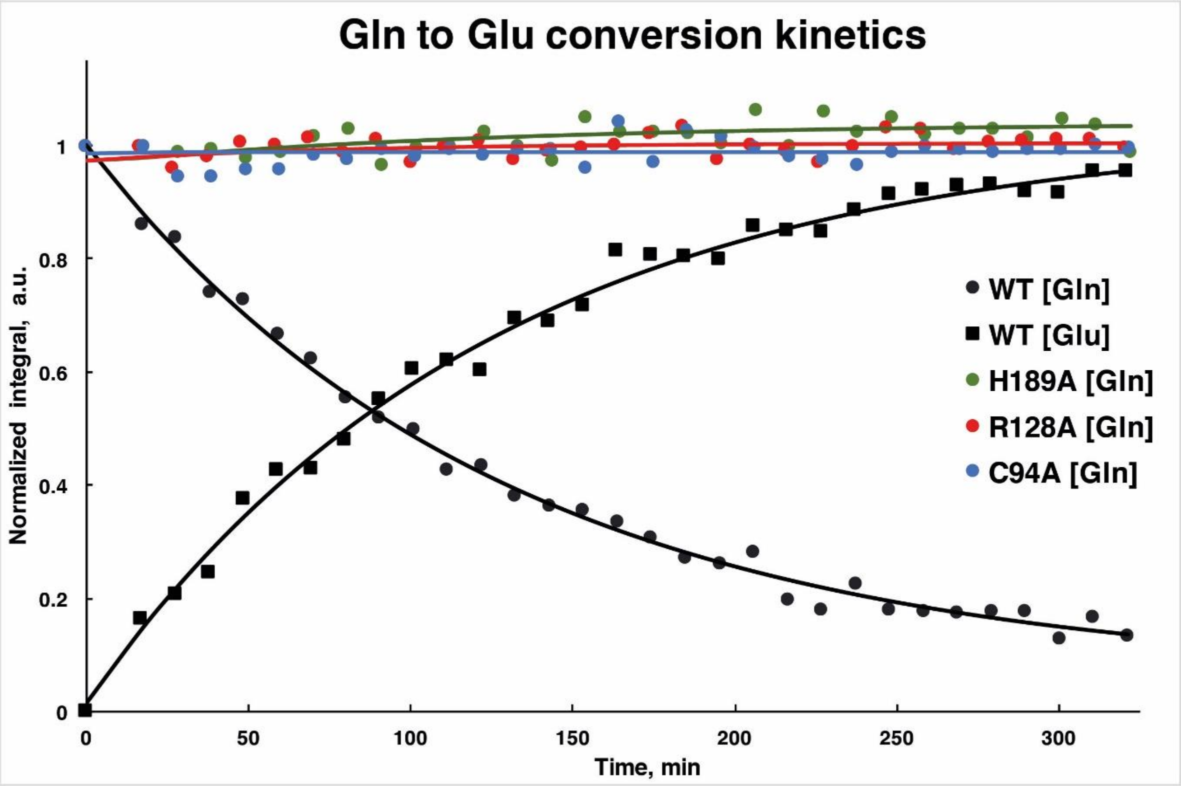


Figure 2.9. Glutaminase activity kinetics of MurT-GatD protein complexes.

Kinetics of glutamine conversion into glutamate was determined from the integration of their γ-CH_2 peaks for MurT-GatD *wt* and mutants C94A, H189A, R128A. Peak intensities were normalized to the initial concentration of glutamine at t=0. The error estimated from signal to noise ratio was 4.3% for all experiments. The curves for MurT-GatD *wt* were obtained by fitting the data to a first order rate law but the curves for the mutants are just guidelines for better visualization and comparison.

2.7. Conclusions

Multidrug resistant Gram-positive bacteria are a major threat to human health, *Staphylococcus aureus* being one of the leading causes of infection worldwide. Many efforts have been made to find new antibiotics and develop new strategies to fight antibiotic resistant bacteria. The enzymatic machinery involved in PG biosynthesis is one of the most common targets for drug development, since it is vital for shape maintenance and turgor pressure counterbalance. The amidation reaction is one of the crucial steps of this biosynthetic process and is achieved by the bi-enzymatic complex MurT-GatD. GatD is responsible for metabolizing glutamine and provide ammonia to MurT, the ligase that converts D-glutamate into D-isoglutamine. At the sequence level, *S. aureus* GatD is similar to other pathogenic Gram-positive bacteria and to CobB and CobQ-related enzymes, suggesting that these enzymes can be grouped in a new sub-family of GATase domains. The high-resolution crystal structure of *S. aureus* GatD showed unique features when compared to other GATase domain architectures. Complementary MurT-GatD glutaminase activity results revealed that GatD residues C94 and H189 are the catalytic amino acids of the active site. This arrangement suggests that GatD may rely on a catalytic dyad although, we cannot exclude putative conformational changes upon MurT/lipid II binding that may bring a third residue to participate in the enzymatic reaction. Glutamine sequestration by GatD is mediated by R128 while Y17 seems to block glutamine entrance, in the absence of MurT and lipid II. MurT binding and full complex formation seem to induce a GatD conformational change, adopting its active structural configuration. Even in the absence of lipid II, the MurT-GatD complex has glutaminase activity.

The structural features of *S. aureus* GatD identified in this work are promising hints for the development of small molecules capable to impair lipid II amidation, placing MurT-GatD as a potential new drug target to fight multidrug resistant bacteria.

2.8. References

1. Boucher HW & Corey GR (2008) Epidemiology of methicillin-resistant *Staphylococcus aureus*. *Clin Infect Dis* **46**, S344–S349.

2. Roemer T, Schneider T & Pinho MG (2013) Auxiliary factors: A chink in the armor of MRSA resistance to β-lactam antibiotics. *Curr Opin Microbiol* **16**, 538–548.

3. Holtje J V (1998) Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*. *Microbiol Mol Biol Rev* **62**, 181–203.

4. van Heijenoort J (2001) Recent advances in the formation of the bacterial peptidoglycan monomer unit. *Nat Prod Rep* **18**, 503–519.

5. Schleifer KH & Kandler O (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev* **36**, 407–477.

6. Gustafson J, Strässle A, Hächler H, Kayser FH & Berger-Bächi B (1994) The *femC* Locus of *Staphylococcus aureus* Required for Methicillin Resistance Includes the Glutamine Synthetase Operon. *J Bacteriol* **176**, 1460–1467.

7. Munch D, Engels I, Muller A, Reder-Christ K, Falkenstein-Paul H, Bierbaum G, Grein F, Bendas G, Sahl H-G & Schneider T (2015) Structural Variations of the Cell Wall Precursor Lipid II and Their Influence on Binding and Activity of the Lipoglycopeptide Antibiotic Oritavancin. *Antimicrob Agents Chemother* **59**, 772–781.

8. Münch D, Roemer T, Lee SH, Engeser M, Sahl HG & Schneider T (2012) Identification and *in vitro* analysis of the GatD/MurT enzyme-complex catalyzing lipid II amidation in Staphylococcus aureus. *PLoS Pathog* **8**, 1–11.

9. Figueiredo TA, Sobral RG, Ludovice AM, de Almeida JMF, Bui NK, 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**.

10. Figueiredo TA, Ludovice AM & Sobral RG (2014) Contribution of Peptidoglycan Amidation to Beta-Lactam and Lysozyme Resistance in Different Genetic Lineages of *Staphylococcus aureus* . *Microb Drug Resist* **20**, 1–12.

11. Zapun A, Philippe J, Abrahams KA, Signor L, Roper DI, Breukink E & Vernet T (2013) *In vitro* Reconstitution of Peptidoglycan Assembly from the Gram-Positive Pathogen *Streptococcus pneumoniae*. *ACS Chem Biol* **8**, 2688–2696.

12. Massière F & Badet-Denisot MA (1998) The mechanism of glutamine-dependent amidotransferases. *Cell Mol Life Sci* **54**, 205–222.

13. Mouilleron S & Golinelli-Pimpaneau B (2007) Conformational changes in ammonia-channeling glutamine amidotransferases. *Curr Opin Struct Biol* **17**, 653–664.

14. Raushel FM, Thoden JB & Holden HM (1999) The amidotransferase family of enzymes: molecular machines for the production and delivery of ammonia. *Biochemistry* **38**, 7891–7899.

15. Hart EJ & Powers-lee SG (2008) Mutation analysis of carbamoyl phosphate synthetase:

does the structurally conserved glutamine amidotransferase triad act as a functional dyad? *Protein Sci* **17**, 1120–1128.

16. Thoden JB, Huang X, Raushel FM & Holden HM (1999) The small subunit of carbamoyl phosphate synthetase: snapshots along the reaction pathway. *Biochemistry* **38**, 16158–16166.
17. Vieira D, Figueiredo TA, Verma A, Sobral RG, Ludovice AM, de Lencastre H & Trincao J (2014) Purification, crystallization and preliminary X-ray diffraction analysis of GatD, a glutamine amidotransferase-like protein from *Staphylococcus aureus* peptidoglycan. *Acta Crystallogr Sect F Struct Biol Commun* **70**, 1–4.
18. Bird LE (2011) High throughput construction and small scale expression screening of multi-tag vectors in *Escherichia coli*. *Methods* **55**, 29–37.
19. Studier FW (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr Purif* **41**, 207–234.
20. Winter G (2010) Xia2: An expert system for macromolecular crystallography data reduction. *J Appl Crystallogr* **43**, 186–190.
21. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 133–144.
22. Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 486–501.
23. Adams PD, Afonine P V., Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC & Zwart PH (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 213–221.
24. Liebschner D, Afonine P V., Moriarty NW, Poon BK, Sobolev O V., Terwilliger TC & Adams PD (2017) Polder maps: Improving OMIT maps by excluding bulk solvent. *Acta Crystallogr Sect D Struct Biol* **73**, 148–157.
25. Davis IW, Murray LW, Richardson JS & Richardson DC (2004) MolProbity: Structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* **32**, 615–619.
26. Thompson JD, Higgins DG & Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.
27. E. Henrick KK (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
28. Krissinel E & Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr Sect D Biol Crystallogr* **60**, 2256–2268.
29. Douangamath A, Walker M, Beismann-Driemeyer S, Vega-Fernandez MC, Sterner R & Wilmanns M (2002) Structural evidence for ammonia tunneling across the $(\beta\alpha)_8$ barrel of the imidazole glycerol phosphate synthase bienzyme complex. *Structure* **10**, 185–193.

30. List F, Vega MC, Razeto A, Häger MC, Sterner R & Wilmanns M (2012) Catalysis uncoupling in a glutamine amidotransferase bienzyme by unblocking the glutaminase active site. *Chem Biol* **19**, 1589–1599.

31. Zein F, Zhang Y, Kang YN, Burns K, Begley TP & Ealick SE (2006) Structural insights into the mechanism of the PLP synthase holoenzyme from *Thermotoga maritima*. *Biochemistry* **45**, 14609–14620.

32. Morar M, Hoskins AA, Stubbe J & Ealick SE (2008) Formylglycinamide ribonucleotide amidotransferase from *Thermotoga maritima* : structural insights into complex formation. *Biochemistry* **47**, 7816–7830.

33. Smith AM, Brown WC, Harms E & Smith JL (2015) Crystal structures capture three states in the catalytic cycle of a pyridoxal phosphate (PLP) synthase. *J Biol Chem* **290**, 5226–5239.

34. Strohmeier M, Raschle T, Mazurkiewicz J, Rippe K, Sinning I, Fitzpatrick TB & Tews I (2006) Structure of a bacterial pyridoxal 5′-phosphate synthase complex. *Proc Natl Acad Sci U S A* **103**, 19284–19289.

35. Guédez G, Hipp K, Windeisen V, Derrer B, Gengenbacher M, Böttcher B, Sinning I, Kappes B & Tews I (2012) Assembly of the eukaryotic PLP-Synthase Complex from *Plasmodium* and activation of the Pdx1 Enzyme. *Structure* **20**, 172–184.

36. Gengenbacher M, Fitzpatrick TB, Raschle T, Flicker K, Sinning I, Muller S, Macheroux P, Tews I & Kappes B (2006) Vitamin B6 biosynthesis by the malaria parasite *Plasmodium falciparum* : biochemical and structural insights. *J Biol Chem* **281**, 3633–3641.

37. Knöchel T, Ivens A, Hester G, Gonzalez A, Bauerle R, Wilmanns M, Kirschner K & Jansonius JN (1999) The crystal structure of anthranilate synthase from *Sulfolobus solfataricus*: functional implications. *Proc Natl Acad Sci U S A* **96**, 9479–9484.

38. Spraggon G, Kim C, Nguyen-Huu X, Yee MC, Yanofsky C & Mills SE (2001) The structures of anthranilate synthase of *Serratia marcescens* crystallized in the presence of (i) its substrates, chorismate and glutamine, and a product, glutamate, and (ii) its end-product inhibitor, L-tryptophan. *Proc Natl Acad Sci U S A* **98**, 6021–6026.

39. Goto M, Omi R, Nakagawa N, Miyahara I & Hirotsu K (2004) Crystal structures of CTP synthetase reveal ATP, UTP, and glutamine binding sites. *Structure* **12**, 1413–1423.

40. Huang X & Raushel FM (1999) Deconstruction of the catalytic array within the amidotransferase subunit of carbamoyl phosphate synthetase. *Biochemistry* **38**, 15909–15914.

41. Chaudhuri BN, Lange SC, Myers RS, Davisson VJ & Smith JL (2003) Toward understanding the mechanism of the complex cyclization reaction catalyzed by imidazole glycerolphosphate synthase: Crystal structures of a ternary complex and the free enzyme. *Biochemistry* **42**, 7003–7012.

42. Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J, Sussman JL, Verschueren KHG & Goldman A (1990) The α/β hydrolase fold. **5**, 197–211.

43. Zalkin H & Smith JL (1998) Enzymes Utilizing Glutamine as an Amide Donor. In

Advances In Enzymology And Related Areas Of Molecular Biology pp. 87–144. John Wiley & Sons, Inc.

44. Thoden JB, Frank M & Benning MM (1999) The structure of carbamoyl phosphate synthetase determined to 2.1 Å resolution. *Acta Crystallogr Sect D Biol Crystallogr* **D-55**, 8–24.

45. Debuscche L, Thibaut D, Cameron B, Crouzet J & Blanche F (1990) Purification and characterization of cobyrinic acid a,c-diamide synthase from *Pseudomonas denitrificans*. *J Bacteriol* **172**, 6239–6244.

46. Blanche F, Couder M, Debussche L, Thibaut D, Cameron B & Crouzet J (1991) Biosynthesis of vitamin B12: stepwise amidation of carboxyl groups b, d, e, and g of cobyrinic acid a,c-diamide is catalyzed by one enzyme in *Pseudomonas denitrificans*. *J Bacteriol* **173**, 6046–6051.

47. Galperin MY & Grishin N V (2000) The synthetase domains of cobalamin biosynthesis amidotransferases CobB and CobQ belong to a new family of ATP-dependent amidoligases, related to dethiobiotin synthetase. *Proteins* **41**, 238–247.

48. Drozdetskiy A, Cole C, Procter J & Barton GJ (2015) JPred4: A protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389–W394.

49. Miran SG, Chang SH & Raushel FM (1991) Role of the four conserved histidine residues in the amidotransferase domain of carbamoyl phosphate synthetase. *Biochemistry* **30**, 7901–7907.

50. Rubino SD, Nyunoya H & Lusty CJ (1986) Catalytic domains of carbamyl phosphate synthetase. *J Biol Chem* **261**, 11320–11327.

2.9. Acknowledgements

The authors thank the Oxford Protein Production Facility, Research Complex (Harwell, Didcot, UK) for access to high-throughput facilities and assistance during the experiments; beamline staff at I02 and I04 for assistance during data collection at Diamond Light Source (Didcot, UK). DV, TAF, FL and MS were supported by fellowships SFRH/BD/62415/2009, SFRH/BD/36843/2007, PD/BD/105737/2014 and PD/BD/128202/2016 from Fundação para a Ciência e Tecnologia (FCT), MCTES, Portugal. This project was supported by Project PTDC/BIA-MIC/3195/2012 from FCT- MCTES, Portugal, Project LISBOA-01-0145-FEDER-007660 (Microbiologia Molecular, Estrutural e Celular) funded by FEDER - COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI), Pest-OE/BIA/UI0457/2001 (CREM) and Unidade de Ciências Biomoleculares Aplicadas-UCIBIO which is financed by national funds from FCT/MEC (UID/Multi/04378/2013) and co-financed by the ERDF under the PT2020 Partnership Agreement (POCI-01-0145-FEDER-007728). The NMR spectrometers are part of the National NMR Network (PTNMR) and are supported by Infrastructure Project N° 022161 (co-financed by FEDER through COMPETE 2020, POCI, and PORL and FCT through PIDDAC).

2.10. Author contributions

The cloning assays of the project were carried out by FL, TAF, RGS, AML, and HL. DV, TAF and FL performed protein production and DV and JT carried out protein crystallization and structure determination. MJR, TS-S and FL carried out model refinement and structural analysis. NMR experiments were performed by MS under the supervision of EJC. All authors contributed to the manuscript writing. HL and TS-S directed and supervised the team members.

2.11. Additional information

First insights of peptidoglycan amidation in Gram-positive bacteria - the high-resolution crystal structure of *Staphylococcus aureus* glutamine amidotransferase GatD

Francisco Leisico[1¶], Diana Vieira[1,2¶], Teresa A. Figueiredo[1,3], Micael Silva[1], Eurico J. Cabrita[1] Rita G. Sobral[1], Ana Madalena Ludovice[1], José Trincão[4], Maria João Romão[1], Hermínia de Lencastre[3,5*] and Teresa Santos-Silva[1*]

Email: tsss@fct.unl.pt and lencash@mail.rockefeller.edu

¶These authors contributed equally to this work.

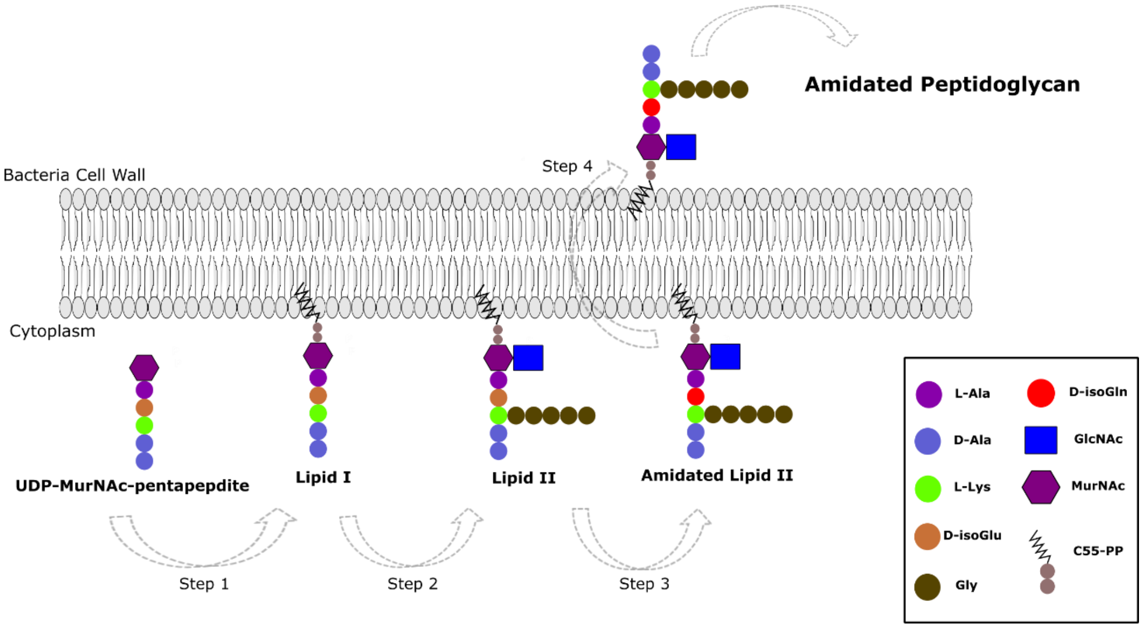


Figure S2.1. General schematic representation of *S. aureus* lipid II biosynthesis.

Figure S2.2. Multiple sequence alignment of *S. aureus* GatD and PDBeFold target sequences.

S. aureus GatD amino acid sequence (represented in the alignment as SaGatD) was aligned using the ClustalW algorithm with the amino acid sequences of PDBeFold target proteins presented in Table 2.2 (here represented with the respective PDB ID and chain). Important GatD residues (and the respective aligned amino acids) are highlighted: the potential catalytic dyad C94 and H189 in red, oxyanion hole G60 and G95 in gray, nucleophile elbow T92, I93 and G96 in cyan, glutamine sequestration R128 in blue and the extra C-terminal α helix of GatD in orange. The secondary structural elements from the *S. aureus* GatD structure are shown above the aligned sequences with α-helices represented as cylinders and β-sheets as arrows. The glutamate residue belonging to the catalytic triad of the identified PDBeFold targets are highlighted in salmon.

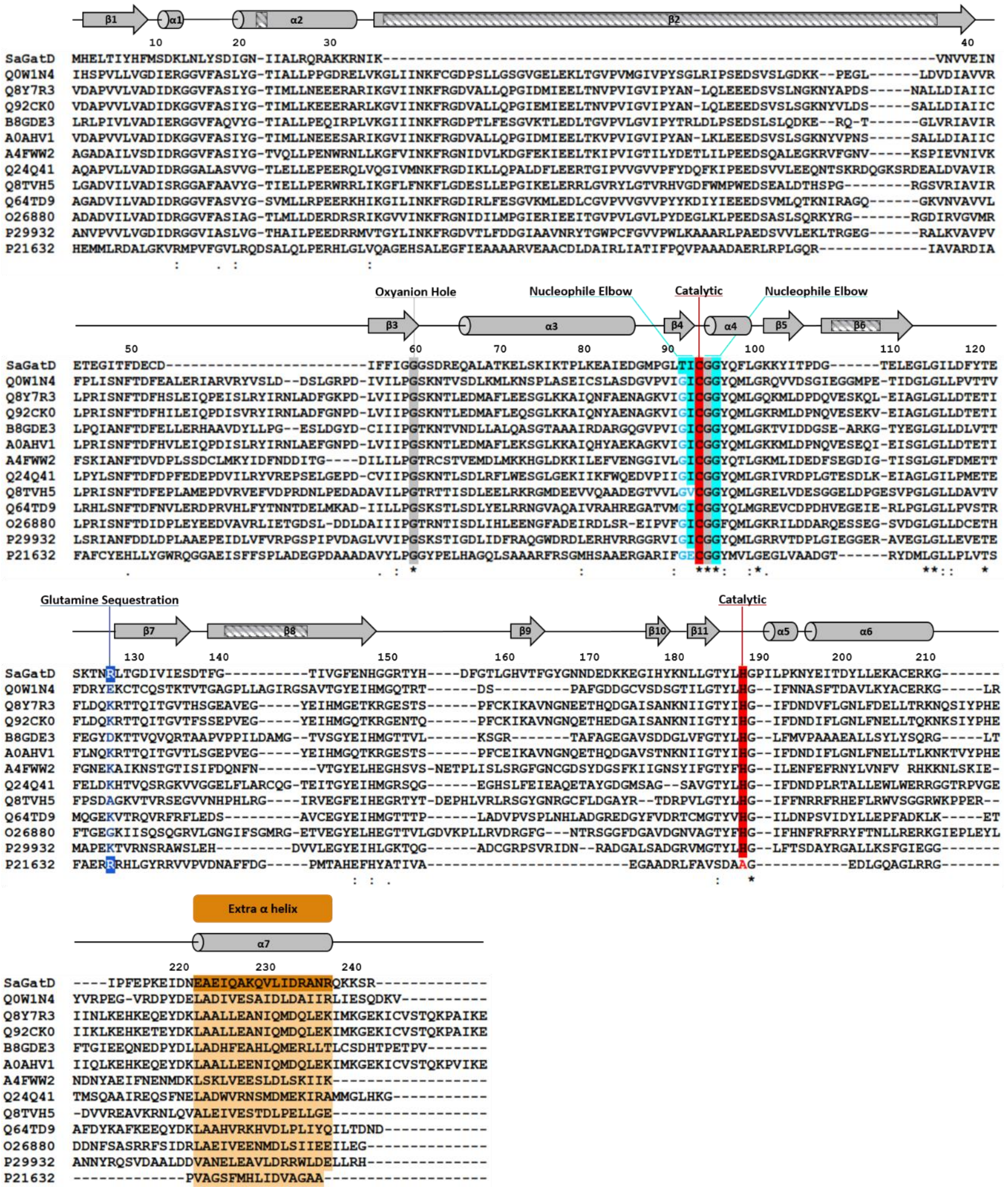


Figure S2.3. Multiple sequence alignment of *S. aureus* GatD and homologs.

S. aureus GatD (represented in the alignment as SaGatD) was aligned using the ClustalW algorithm with the amino acid sequences of the BLASTP hits (indicated in the alignment with the respective UNIPROT code) and also with CobQ and CobB from *P. denitrificans*, (UNIPROT code P29932 and P21632, respectively). Important GatD residues are highlighted such as the catalytic dyad C94 and H189 in red, oxyanion hole G60 and G95 in gray, nucleophile elbow T92, I93 and G96 in cyan, glutamine sequestration R128 in blue and the extra C-terminal α helix of GatD in orange. The secondary structural elements are shown above the aligned sequences with α-helices represented as cylinders and β-sheets as arrows.

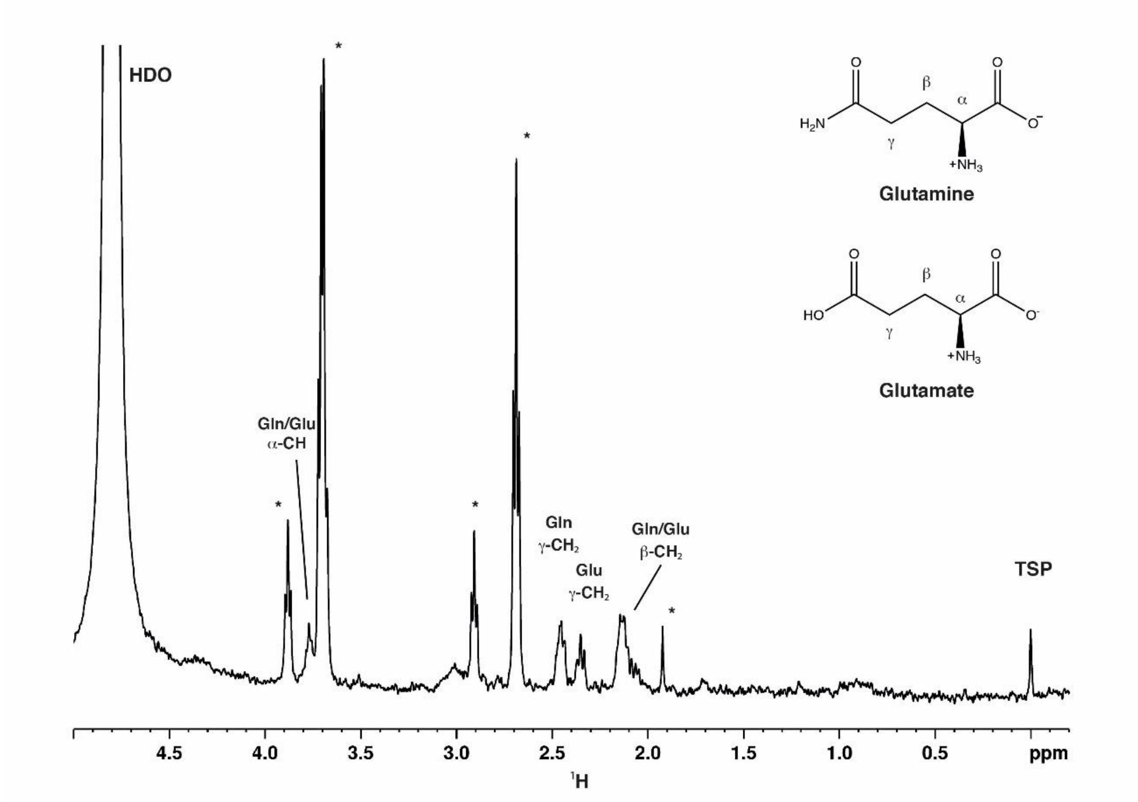


Figure S2.4. Representative ^{1}H-NMR-based glutaminase activity of MurT-GatD *wt*.

Resonance assignments show the distinct chemical shift of the γ-CH_2 of glutamine and glutamate resonances that allow the quantification of protein activity. The ^{1}H-NMR spectrum here presented shows glutamine (875 μM) conversion in glutamate catalyzed by 35 μM of MurT-GatD *wt* after 60 minutes of reaction. Asterisk (*) indicates buffer resonances.

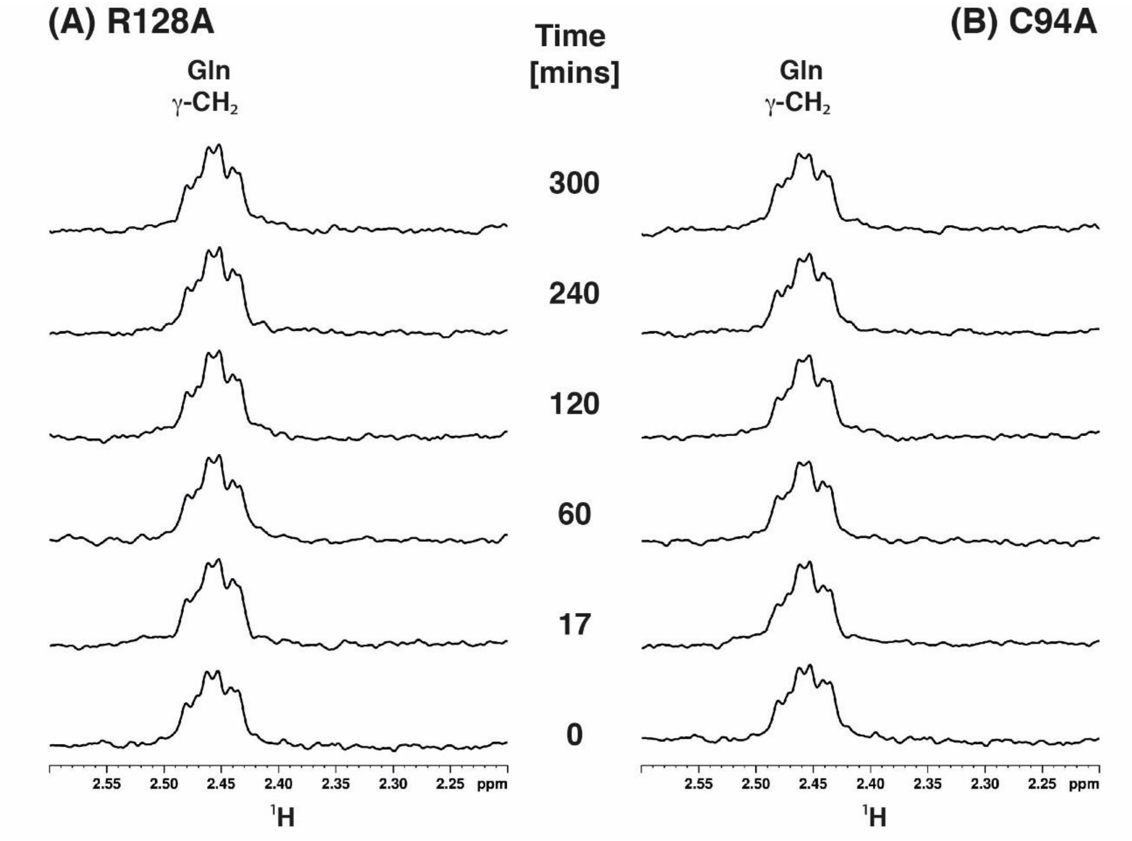


Figure S2.5 Glutaminase activity of MurT-GatD mutants monitored by ^{1}H-NMR.

Expansions of the ^{1}H-NMR spectra of the glutamine and glutamate γ-CH_2 peak region with resonance assignment for different reaction times to monitor glutaminase activity for (**A**) R128A and (**B**) C94A mutants. The reactions were performed with 875 μM of glutamine and 35 μM of protein. Glutamine concentration remains constant over time in (**A**) and (**B**) and no glutamate formation can be detected.

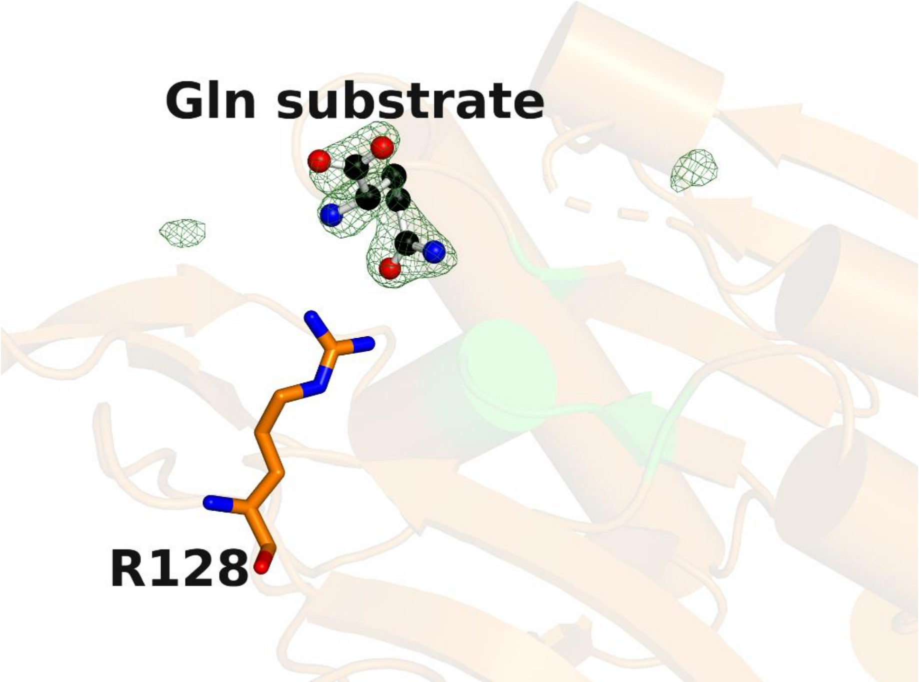


Figure S2.6. Confirmation of glutamine substrate presence by polder map.

Polder map was generated excluding the bulk solvent contribution within a 10 Å radius from glutamine substrate. The resulting polder map is contoured at 5 σ and is presented within a 10 Å radius from glutamine substrate to show that there are no other residual density peaks besides glutamine.

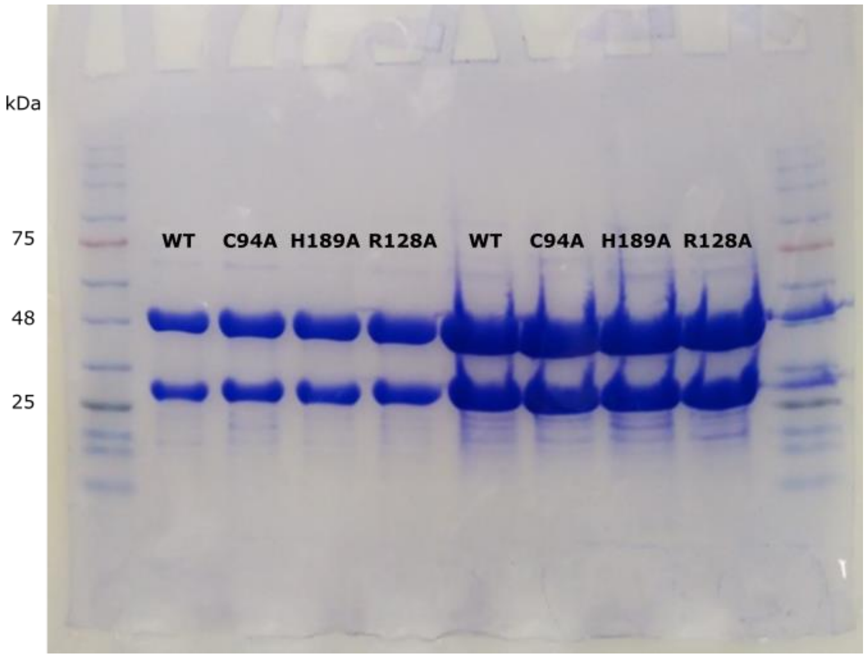


Figure S2.7. Full-length SDS-PAGE gel used to assess the heteromeric complex MurT-GatD purity. Full-length 12% SDS-PAGE gel of the final samples of *S. aureus* MurT-GatD wt and C94A, H189A and R128A mutants prior to enzymatic activity tests. Two different amounts of protein were loaded as quality check. Both MurT (49 KDa) and GatD (27 KDa) proteins are resolved in the gel.

Table S2.1. Strains and plasmids used in this study.

Strains		
S. aureus		
COL	Homogeneous Mc^r	The Rockefeller Univ. Collection
E.coli		
BL21-CodonPlus(DE3)-RIPL	*E. coli* B F$^-$ *ompT hsdS*(r$_B^-$ m$_B^-$) *dcm*$^+$ Tet^r *gal λ* (DE3) *endA* Hte [*argU proL*Cm^r] [*argU ileY leuW* Strep/Spec^r]	Stratagene.
BL21-His$_6$-*gatD*	BL21-CodonPlus(DE3)-RIPL strain with pET28a-His$_6$-*gatD*	
BL21-*murT*-*gatD*-His$_6$	BL21-CodonPlus(DE3)-RIPL strain with pET28a-*murT*-*gatD*-His$_6$	Figueiredo, T. A. *et al.* [10]
BL21-*murT*-*gatD*-His$_6$C94A	BL21-*murT*-*gatD*-His$_6$ with C94A mutation	In this study
BL21-*murT*-*gatD*-His$_6$H189A	BL21-*murT*-*gatD*-His$_6$ with H189A mutation	In this study
BL21-*murT*-*gatD*-His$_6$R128A	BL21-*murT*-*gatD*-His$_6$ with R128A mutation	In this study
Plasmids		
pOPINF	using the In-Fusion™	In this study
pET28a	Expression vector with T7/*lac* promoter, N-terminal His tag, thrombin cleavage site, C-terminal His tag; kan^r	Invitrogen
pET28a-His$_6$-*gatD*	pET28a expressing His$_6$-*gatD*	This study
pET28a-*murT*-*gatD*-His$_6$	pET28a expressing *murT*-*gatD*-His$_6$	Figueiredo, T. A. *et al.* [10]
pET28a*murT*-*gatD*-His$_6$C94A	pET28a expressing *murT*-*gatD*-His$_6$C94A	In this study
pET28a*murTgatD*His$_6$H189A	pET28a expressing *murT*-*gatD*-His$_6$H189A	In this study
pET28a*murTgatD*His$_6$R128A	pET28a expressing *murT*-*gatD*-His$_6$R128A	In this study

a) Mc^r methicillin resistant; Cm^r, chloramphenicol resistant; Spec^r, spectinomycin resistant; Kan^r, kanamycin resistant.

Table S2.2. Primers for *gatD* gene mutation in *S. aureus murT-gatD* operon.

Mutation	Forward primer (5′- 3′ sequence)	Reverse primer (5′- 3′ sequence)
C94A	GGTATGCCGGGATTAACGATT**GCT**GGAGGCTATCAATTTTTAGGG	CCCTAAAAATTGATAGCCTCC**AGC**AATCGTTATCCCGGCATACC
H189A	GGCATTCATTATAAAAATTTATTAGGTACTTATTTA**GCT**GGACCAATTTTACCTAAAAATTACG	CGTAATTTTTAGGTAAAATTGGTCC**AGC**TAATAAGTACCTAATAAATTTTTATAATGAATGCC
R128A	CTGAATCAAAGACAAAC**GCA**TTAACAGGAGATATTGTTATCG	CGATAACAATATCTCCTGTTAA**TGC**GTTTGTCTTTGATTCAG

CHAPTER 3

FUNCTIONAL INSIGHTS INTO PEPTIDOGLYCAN AMIDATION BY THE ESSENTIAL MURT-GATD COMPLEX

3.1. Context

This chapter consists in the manuscript currently under resubmission in The FASEB Journal.

This chapter addresses the structure-function relation of MurT-GatD from *S. aureus* and *S. pneumoniae*, using biochemical and biophysical techniques. I was responsible for the production of *S. aureus* GatD measured by SPR and production of *S. aureus* MurT-GatD used in the SAXS experiments and enzymatic assays. I performed SAXS data collection and preliminary analysis for *S. aureus* MurT-GatD. In collaboration with the Pneumococcus group in France, I measured and processed the enzymatic activities of *S. aureus* MurT-GatD.

It is worth mentioning that I had the opportunity to crystallize *S. pneumoniae* MurT-GatD with ligands, determine their structures and perform structural analysis. This work contributed to the publication in Nature Communication (Nat Commun. 2018 Aug 9;9(1):3180. doi: 10.1038/s41467-018-05602-w) but, since it was not the main focus of my work, it was not included in this thesis. Moreover, the crystal structure of *S. aureus* MurT-GatD was published (Sci Rep. 2018 Aug 28;8(1):12953. doi: 10.1038/s41598-018-31098-x), revealing new insights into the complex structure and function such as the potential participation of MurT in a glutaminase catalytic triad.

3.2. Front page of the manuscript (under revision)

Functional Insights into Peptidoglycan Amidation by the Essential MurT-GatD Complex

Francisco Leisico[1,2], Ana C. F. Paiva[3,4], Cécile Morlot[5], André Zapun[5], Pedro M. F. Sousa[3,4], Tiago M. Bandeiras[3,4], Haydn D. T. Mertens[6], Ana Madalena Ludovice[1,7], Maria João Romão[1,2], Rita G. Sobral[1,7*] and Teresa Santos-Silva[1,2*]

[1]Associate Laboratory i4HB - Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2819-516 Caparica, Portugal; [2]UCIBIO – Applied Molecular Biosciences Unit, Department of Chemistry, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2819-516 Caparica, Portugal; [3]iBET, Instituto de Biologia Experimental e Tecnológica, Apartado 12, 2781-901 Oeiras, Portugal; [4]Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal; [5]Univ. Grenoble Alpes, CNRS, CEA, IBS, 38000 Grenoble, France; [6]European Molecular Biology Laboratory Hamburg, Notkestrasse 85, 22607 Hamburg, Germany; [7]UCIBIO – Applied Molecular Biosciences Unit, Department of Life Sciences, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2819-516 Caparica, Portugal.

Correspondence

Rita Gonçalves Sobral; Associate Laboratory i4HB - Institute for Health and Bioeconomy, and Unidade de Ciências Biomoleculares Aplicadas UCIBIO, Department of Life Sciences, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2819-516 Caparica, Portugal; Tel: 00351-212948530; Email: rgs@fct.unl.pt.

Teresa Santos-Silva; Associate Laboratory i4HB - Institute for Health and Bioeconomy, and Unidade de Ciências Biomoleculares Aplicadas UCIBIO, Department of Chemistry, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2819-516 Caparica, Portugal; Tel: 00351-212948300 ext. 10940; Email: tsss@fct.unl.pt.

Running title: Functional Insights into MurT-GatD Heterodimer

Abbreviations: Glc*N*Ac, *N*-acetylglucosamine; Mur*N*Ac, *N*-acetylmuramic acid; PG, Peptidoglycan; RU, response units; Rg, radius of gyration; SAXS, Small-Angle X-ray Scattering; SEC, Size-Exclusion Chromatography; SPR, Surface Plasmon Resonance.

Enzymatic Activity Data: MurT-GatD (EC 6.3.5.13) - Lipid II isoglutaminyl synthase (glutamine-hydrolyzing) from *Staphylococcus aureus*

Databases: SAXS data is available in the SASBDB database with the accession codes SASDLZ9 and SASDM22.

Keywords: Antibiotic Resistance; Peptidoglycan Amidation; MurT-GatD; Small-Angle X-ray Scattering; Surface Plasmon Resonance

3.3. Abstract

In most Gram-positive bacteria, the peptidoglycan is amidated to ensure the efficient polymerization of the cell wall. In *Staphylococcus aureus*, this modification also contributes for the resistance to antibiotics and lysozyme. The amidation reaction is catalysed by the MurT-GatD complex, which has an essential function in critical pathogens such as *S. aureus*, *Streptococcus pneumoniae* and *Mycobacterium tuberculosis,* constituting a promising target for new antibiotics. The crystal structures of *S. aureus* and *S. pneumoniae* MurT-GatD showed extended and compact conformations, respectively. We combined structural and functional approaches to study the relevance of these conformations and their impact on the lipid II amidation reaction mechanism. Using small-angle X-ray scattering, we showed that these complexes also adopt distinct conformations in solution, each matching their respective crystal structure. The interaction between GatD and the C-terminal domain of MurT is conserved in both conformations and provides the minimal required scaffold for complex assembly in the *S. aureus* extended conformation. In this case, surface plasmon resonance assays demonstrated that MurT C-terminal domain and GatD interact with a 2.21 nM affinity, suggesting that this interface holds the complex together and grants flexibility to MurT domains. Compared to *S. pneumoniae* MurT-GatD, enzymatic activity experiments showed that the *S. aureus* complex has a lower affinity for MurT substrates, which propose that the extended conformation may be less functional. Our experimental evidence reveals differences in the structure and enzymatic activity between the MurT-GatD complexes from *S. aureus* and *S. pneumoniae,* suggesting a potential mechanistic diversity among bacterial peptidoglycan amidation reactions.

3.4. Introduction

Peptidoglycan (PG) is the major structural component of the bacterial cell wall and plays a crucial role in bacterial survival, cell shape maintenance and turgor pressure counterbalance [1]. Its conserved primary structure consists of glycan chains of alternating *N*-acetylglucosamine (Glc*N*Ac) and *N*-acetylmuramic acid (Mur*N*Ac) attached to a stem peptide, which can be cross-linked to a neighbouring peptide either directly or through a peptide bridge [2]. PG is a defining trait of bacteria and its essentiality has motivated the development of antibacterial drugs that target precursor molecules or enzymes involved in the PG biosynthesis pathway [3]. β-Lactams are one of the most clinically relevant classes of antibiotics, as they inhibit the ubiquitous penicillin-binding proteins that are responsible for PG cross-linking. Their overuse to treat a wide range of bacterial infections has contributed to the emergence of resistance mechanisms in numerous pathogens. Today, the key guideline for the next generation of antibiotics is thus to develop molecules that target a narrower range of organisms.

The stem peptide of PG presents significant structural differences among the various bacterial species, which is due to the specificity of the enzymes responsible for its biosynthesis - the Mur ligases [2]. These enzymes sequentially add to the UDP-Mur*N*Ac PG precursor the five amino acids of the stem peptide moiety (L-Ala-D-γ-Glu-L-Lys-D-Ala-D-Ala in many Gram-positive bacteria, like *S. aureus* and *S. pneumoniae*). The resulting UDP-Mur*N*Ac is then bound to the membrane-tethered undecaprenyl-phosphate, forming the PG precursor called lipid I. Further addition of Glc*N*Ac results in the formation of lipid II. This precursor can undergo secondary modifications in the stem peptide, including amidation of D-Glu into D-isoGln (Fig. 3.1A) and addition of peptide branches to L-Lys [2]. In *S. aureus*, the branches comprise a pentaglycine peptide [4], while in *S. pneumoniae* they are composed of L-Ala-L-Ala or L-Ser-L-Ala [5]. These peptides are later used by penicillin-binding proteins to bridge the glycan strands of the PG and ensure the integrity of its mesh-like structure [2].

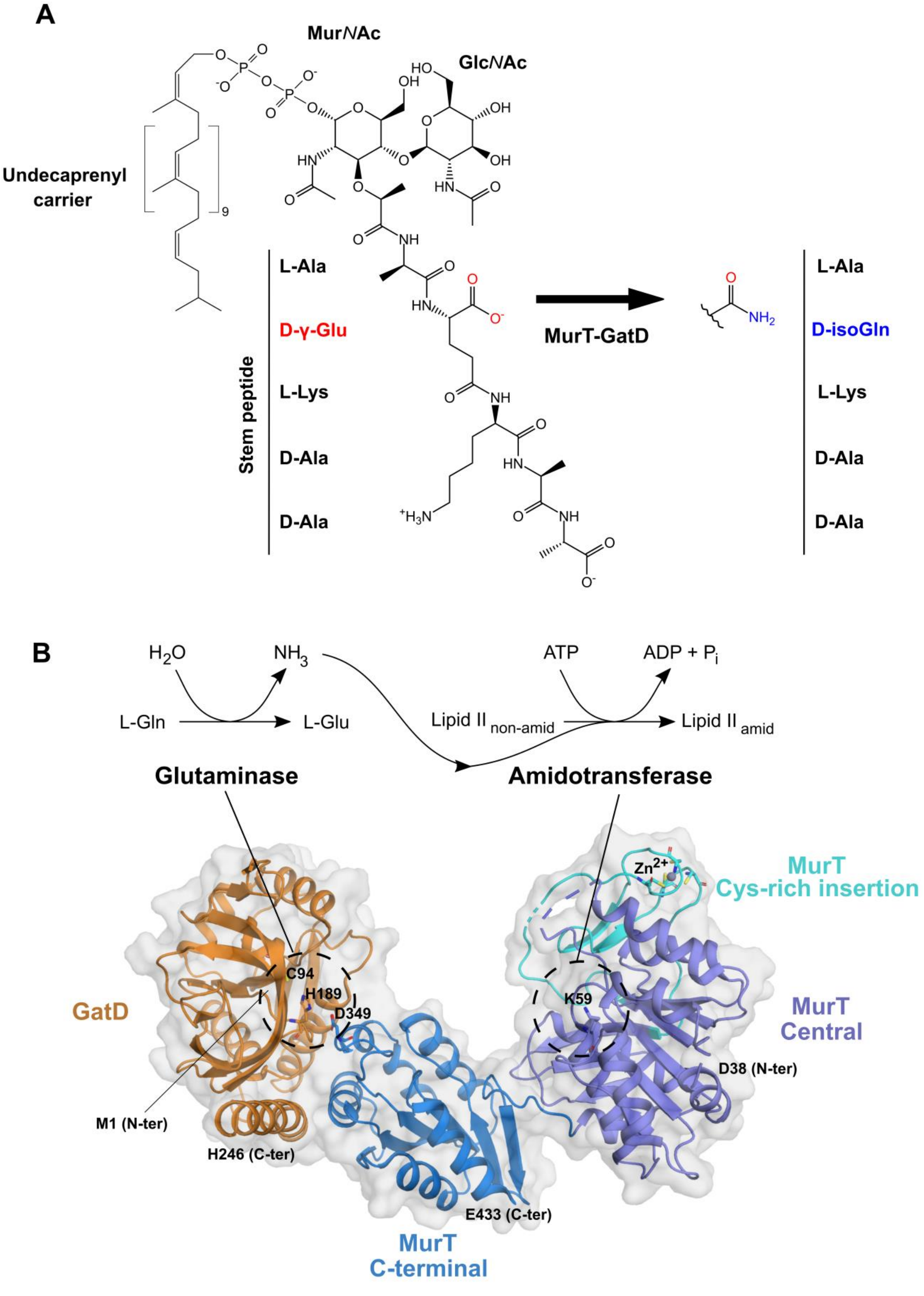


Figure 3.1. Lipid II amidation is catalysed by the MurT-GatD complex.

(**A**) Chemical structure of *S. aureus* (and most Gram-positive bacteria) lipid II before and after the amidation reaction catalysed by MurT-GatD. (**B**) Cartoon representation of the *S. aureus* MurT-GatD crystal structure

(PDB ID 6GS2). The glutaminase active site is represented by the catalytic triad C94 and H189 from GatD (orange) and D349 from MurT C-terminal domain (blue). MurT central domain (purple) possesses the ATPase active site, with the conserved lysine (K59) required for ATP hydrolysis. Complex assembly is mediated by MurT C-terminal domain and GatD. The Cys-rich insertion is coloured in cyan and the coordinated Zn^{2+} ion is shown as a grey sphere. Catalytic residues are shown as sticks.

Glutamate amidation in PG is catalysed by the heterodimeric complex MurT-GatD (Fig. 3.1) [6–8], which was shown to be essential for important pathogens such as *S. aureus* [7,9,10], *S. pneumoniae* [8,11,12] and *M. tuberculosis* [13]. Amidation of the stem peptide seems to be crucial for the optimal activity of penicillin-binding proteins in PG cross-linking [8]. Moreover, it has been demonstrated that impairment of MurT-GatD function in methicillin-resistant strains of *S. aureus* increases their susceptibility towards β-lactam antibiotics and lysozyme [6,14], making the amidotransferase enzymatic reaction a promising target for antibiotic development.

Initial *in vitro* studies on the *S. aureus* MurT-GatD complex showed that both proteins are required in a concerted manner for the amidation of lipid II [7]. It was also demonstrated that L-Gln is the amine donor of the reaction and that MurT-GatD activity is adenosine 5′-triphosphate (ATP)- and Mg^{2+}- dependent. These observations suggested that lipid II amidation is carried out by glutamine amidotransferase activity [15], which includes two catalytic steps: firstly, the glutaminase domain (GatD) converts L-Gln into ammonia and glutamate; secondly, the ammonia is channelled to the ATP-dependent synthase domain (MurT) where lipid II amidation occurs (Fig. 3.1B).

The crystal structure of the isolated *S. aureus* GatD (PDB ID 5N9M), with a glutamine molecule at the entrance of the active site, allowed identifying the residues crucial for substrate binding and *in vitro* glutaminase activity of the complex [16]. The crystal structure of the whole MurT-GatD complex from *S. aureus* [17] revealed a boomerang-like shape where the heterodimer interface is mediated by MurT C-terminal domain. Located at this interface, the conserved glutaminase catalytic triad (Cys-His-Asp) [18] involves residues from both proteins (Fig. 3.1B). The central domain of MurT is homologous to Mur-ligase proteins, sharing the same ATP binding site architecture [17]. Considering the reaction catalysed by Mur ligases [19], phosphorylation of the main chain carboxylate of lipid II D-

Glu likely serves as an activation step, prior to its amidation. The orientation of the MurT domains and GatD results in the formation of a channel that guides ammonia from the glutaminase to the amidotransferase active sites [20]. The central domain of MurT possesses a cysteine-rich insertion of roughly 50 residues that is absent in Mur ligases (Fig. 3.1B). Structurally, this insertion forms a zinc binding motif, defined by cysteines 202, 205, 224 and 226 [17]. This binding motif (C-X_2-C-X_{19}-C-$X_{1\text{-}2}$-C) is conserved in *Bacilli* class [18], but its role in the MurT-GatD complex is still unknown.

The crystal structure of *S. pneumoniae* MurT-GatD [21] revealed that the structures of GatD and MurT individual domains are similar to those of the complex from *S. aureus* [18]. However, *S. pneumoniae* MurT-GatD adopts a more compact conformation than the *S. aureus* complex; in the case of *S. pneumoniae*, GatD interacts with the C-terminal and central domains of MurT (Fig. 3.2). The different conformations of the heterodimers are due to a hinge motion of MurT domains that, in the case of the *S. pneumoniae* heterodimer, moves the glutaminase active site closer to the amidotransferase active site (*ca* 14 Å), leading to an increase of the buried surface area of the complex interface by 300 $Å^2$ (940 $Å^2$ for *S. aureus* and 1231 $Å^2$ for *S. pneumoniae*). In *S. pneumoniae*, the cysteine-rich insertion of MurT is structurally similar to *S. aureus* MurT, but the conserved cysteines are involved in disulfide bonds and no zinc is found in the crystal structure (Fig. 3.2). Thiol alkylation reaction followed by mass spectrometry analysis showed that these cysteines were reduced during purification, indicating that their thiol groups were available to coordinate zinc ions before crystals were formed and that the disulfide bonds probably formed during crystallization [21]. In *S. pneumoniae* MurT-GatD, the ATP binding site is similar to that in the *S. aureus* complex [18].

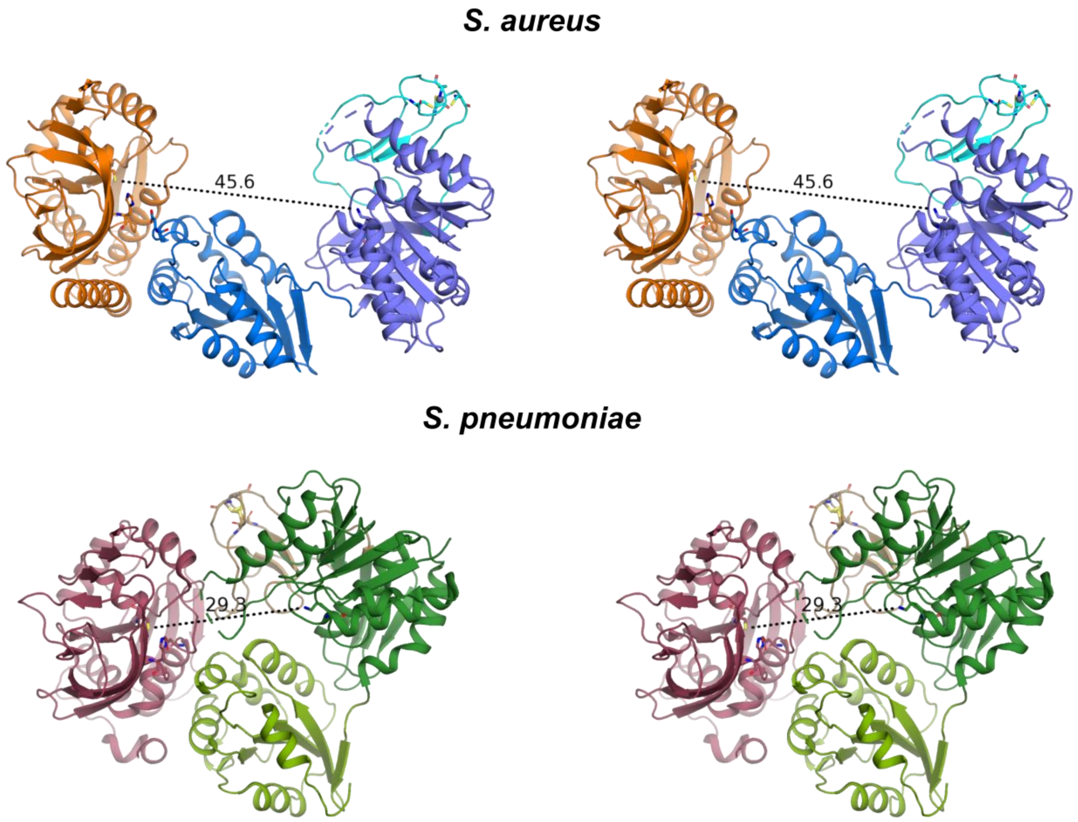


Figure 3.2. Extended and compact MurT-GatD crystal structures shown as wall-eyed stereo mode. *S. aureus* (PDB ID 6GS2) and *S. pneumoniae* (PDB ID 6FQB) MurT-GatD show different domain arrangements (distance between the glutaminase and the ATPase active sites in Å). The cysteine-rich insertion of *S. pneumoniae* complex is coloured in brown and contains two disulfide bonds involving the conserved cysteine residues 205, 208, 227 and 230 (shown in sticks). *S. aureus* GatD, MurT C-terminal and central domains are coloured orange, blue and purple, respectively. *S. pneumoniae* GatD, MurT C-terminal and central domains are coloured red, light and dark green, respectively.

As the crystal structures of MurT-GatD from *S. aureus* and *S. pneumoniae* showed MurT domains differently organized, we investigated the structure of two complexes in solution using small-angle X-ray scattering (SAXS). We also measured the affinity between *S. aureus* GatD towards MurT and the role of specific interface residues, using surface plasmon resonance (SPR). Moreover, we characterized the lipid II amidation activity of *S. aureus* MurT-GatD *in vitro*, providing an enzymatic assay to identify potential inhibitors against pathogenic Gram-positive bacteria. These structural and enzymatic data are discussed in light of the molecular mechanism of PG amidation in *S. aureus* and *S. pneumoniae*, providing new avenues for the development of narrow range antibiotics.

3.5. Methods

3.5.1. Gene cloning, expression and purification of *S. aureus* MurT-GatD

Cloning and production of *S. aureus* MurT-GatD were performed as previously described [6,16]. Briefly, *S. aureus murT-gatD* coding regions were amplified by PCR from *S. aureus* strain COL total DNA and cloned into the expression plasmid pET28a. The genetic construction was designed for the co-expression of MurT untagged and GatD tagged at the C-terminus with His_6.

Escherichia coli BL21-CodonPlus (DE3)-RIPL cells (Agilent) were used for heterologous expression of *S. aureus* MurT-GatD. Cells were grown aerobically in LB medium at 310 K until saturation and then 0.5 mM IPTG were used to induce expression at 293 K for 16 hours. Frozen or fresh pellets were resuspended in buffer A (100 mM Tris-HCl pH 8.2 at 293 K, 500 mM NaCl, 10 mM $MgCl_2$ and 3 mM DTT) supplemented with 10 mM imidazole, 10 $\mu g \cdot mL^{-1}$ DNaseI and ½ tablet of cOmplete™ ULTRA Tablets, Mini, EDTA-free, EASYpack Protease Inhibitor Cocktail (Roche). Cells were lysed by sonication and the resulting supernatant applied onto a 5 mL HisTrap™ HP column (GE Healthcare Life Sciences). Proteins retained in the column were eluted using a linear gradient of imidazole. The fractions containing MurT-GatD proteins were pooled and desalted to buffer A using 5 mL HiTrap™ Desalting columns (GE Healthcare Life Sciences). All fractions were pooled, concentrated and applied onto a Superdex 200 10/300 GL (GE Healthcare) column pre-equilibrated with buffer A.

3.5.2. Gene cloning, expression and purification of *S. pneumoniae* MurT-GatD

The constructs, strain, expression conditions and purification procedure were performed as described in Morlot *et al* [21], with the exceptions that the proteins were eluted from the 1 mL HisTrap FF column with a 25 to 500 mM linear gradient of imidazole and the anion exchange chromatography step was omitted. All solutions contained 2 mM TCEP, instead of DTT.

3.5.3. Gene cloning, expression and purification of *S. aureus* GatD

The *gatD* gene was amplified from the genome of *S. aureus* COL strain using primers His$_6$-GatD-NdeI-F and His$_6$-GatD-NotI-R and ligated into pET28a vector using T4 DNA ligase (New England Biolabs). The restriction sites of NdeI (Fermentas) and NotI (Fermentas) were used to generate a N-terminus His$_6$-GatD construction. Information regarding the strains, plasmids and primers used are shown in Table S3.1.

The recombinant GatD protein was expressed in *E. coli* BL21-CodonPlus (DE3)-RIPL cells at 310 K until saturation and induced with 0.5 mM IPTG for 3 hours at 310 K. Pellets were resuspended and protein purified as here described for *S. aureus* MurT-GatD with the following modifications: purification buffer was 100 mM Tris-HCl pH 8.0 at 293 K, 500 mM NaCl and 3 mM DTT, and resuspension was additionally supplemented with 5 mM $MgCl_2$. The final sample was directly used in the SPR experiments.

3.5.4. Gene cloning, expression and biotinylation of *S. aureus* MurT proteins for SPR measurements

In order to usebiotinylated *S. aureus* MurT as bait in SPR experiments, wild-type and mutant versions of MurT were fused to an AviTag at the C-terminus and co-expressed with BirA to accomplish biotinylation *in vivo*. The coding sequences for BirA-FLAG, His$_6$-HRV3C-MurT(2-437)-TEV-AviTag, His$_6$-HRV3C-MurT(2-437)_Y354A-TEV-AviTag and His$_6$-HRV3C-MurT(2-437)_K384A-TEV-AviTag were synthetized and cloned into pRSF-Duet™-1 by GenScript. BirA-FLAG was cloned on Multiple Cloning Site (MCS) 1, while the MurT constructs were cloned on MCS 2.

BirA-FLAG and all MurT constructs were co-expressed in *E. coli* BL21 (DE3) cells in LB medium at 310 K until saturation and induced with 0.05 mM IPTG for 3 hours at 291 K. To promote cellular biotinylation of BirA on all AviTagged MurT constructs, 5 $\mu g \cdot mL^{-1}$ of D-Biotin was supplied to the medium upon IPTG addition. Cells were then resuspended in

BugBuster™ Protein Extraction Reagent (Novagen) supplemented with 0.1 mg·mL^{-1} Lysozyme, 5 U·mL^{-1} Benzonase, 1 mM PMSF and PIC without EDTA and placed on ice for 15 min. The lysate was cleared by centrifugation at 13200 *g* and 277 K for 25 min and injected onto a PD-10 Desalting column equilibrated in 100 mM Tris-HCl pH 8.2 at 293 K, 500 mM NaCl, 10 mM $MgCl_2$, 1 mM TCEP, 1 mM PMSF and PIC without EDTA. Protein expression and biotinylation was assessed by Western Blotting using Streptavidin, Alkaline Phosphatase Conjugate antibody (Invitrogen™).

3.5.5. SAXS measurements

Synchrotron radiation X-ray scattering data coupled to an in-line chromatography system were collected for *S. pneumoniae* MurT-GatD at the EMBL P12 beamline of the storage ring PETRA III (DESY, Hamburg) using a PILATUS 6M pixel detector (DECTRIS, Switzerland) [22] and for *S. aureus* MurT-GatD at the BM29 beamline of the European Synchrotron Radiation Facility (ESRF, Grenoble) using a PILATUS 1M pixel detector (DECTRIS, Switzerland) [23]. The experimental details of the instruments and derived parameters are listed in Table S3.2 and S3.3. 35-55 µl samples of concentrated *S. pneumoniae* MurT-GatD were injected onto a Superdex 200 Increase (5/150) column (GE Healthcare) equilibrated in 50 mM HEPES pH 7.5, 10 mM $MgCl_2$ and 500 mM NaCl, with and without the addition of 0.1 mM $ZnSO_4$. Samples were exposed to X-rays while flowing at 0.4 mL·min^{-1} through a temperature-controlled capillary (1.2 mm ID) at 293 K. 40 µL of concentrated *S. aureus* MurT-GatD sample was injected onto a Bio Sec-3 300A (7.8 mm ID) column (Agilent) equilibrated in buffer A. Samples were exposed to X-rays while flowing at 0.75 mL·min^{-1} through a temperature controlled quartz capillary (1.2 mm ID) at 293 K. At both beamlines, frames of 1.0 s exposure time were collected during sample elution and data from the detectors were normalized to the transmitted beam intensity, averaged, placed on absolute scale relative to water. The scattering of buffer solutions was subtracted using CHROMIXS [24]. All data manipulations were performed using PRIMUS*qt* and the ATSAS software package [25]. The forward scattering $I(0)$ and R_g were determined from Guinier analysis [26] assuming that at very small angles ($s \leq 1.3/R_g$) the intensity is represented as $I(s)=I(0)\cdot\exp(-(s\cdot R_g)2/3))$. These parameters were also estimated from the full scattering curves using the indirect Fourier transform method implemented in the program GNOM [27], along with the distance distribution function *p(r)* and the maximum particle dimensions D_{max}. Molecular masses of solutes were estimated from the forward scattering *I(0)* by computation of partial specific volume and the contrast between the protein sequence and the chemical components of the solution using the MULCh server (http://smb-research.smb.usyd.edu.au/NCVWeb/) [28]. Protein concentrations used for the forward scattering *I(0)* normalisation were estimated from the peaks of the UV chromatograms collected during data acquisition. Computation of theoretical scattering intensities from crystal structures was performed using the program CRYSOL [29].

Hybrid Rigid body modelling—Rigid body models were computed from the experimental data using CORAL [30] and the available high-resolution structures of MurT-GatD, (PDB ID 6GS2 and 6FQB, for *S. aureus* and *S. pneumoniae*, respectively) as rigid bodies. Flexible regions and rigid bodies are as defined in Tables S3.2 and S3.3.

Ensemble analysis— Ensemble modelling analysis was conducted using the ensemble optimisation method EOM [31,32]. From a pool of 10,000 decoy models based on the *S. aureus* and *S. pneumoniae* MurT-GatD heterodimers (flexible hinge regions as defined in Tables S3.2 and S3.3), a genetic algorithm was employed to select a subset of models that best fit the experimental data.

3.5.6. Surface Plasmon Resonance experiments

The interaction between *S. aureus* GatD and the biotinylated MurT proteins (wild-type and mutant forms) was assessed by SPR at 298 K using a Biacore T200 instrument (Cytiva). The surface of a CM5 sensor chip was activated with 400 mM EDC and 100 mM NHS for 10 min. NeutrAvidin (100 $\mu g{\cdot}mL^{-1}$ in 10 mM sodium citrate pH 4.5) was immobilized onto the activated chip using the standard amine coupling procedure for 15 min at a flow rate of 10 $\mu L{\cdot}min^{-1}$ in order to reach 16000 response units (RU). The remaining activated groups were blocked with a 5 min injection of 1 M ethanolamine, pH 8.5. The biotinylated *S. aureus* $MurT^{Cter}$ wild-type (125 $\mu g{\cdot}mL^{-1}$ in HBS-N), Y354A (250 $\mu g{\cdot}mL^{-1}$ in HBS-N) and K384A (250 $\mu g{\cdot}mL^{-1}$ in HBS-N) proteins were coupled to the NeutrAvidin coated surface for 100 to 200 seconds, at a flow rate of 10 µL·min^{-1} in order to reach 100 to 300 RU. The background buffer used during immobilization was HBS-N (10 mM HEPES pH 7.4, 150 mM NaCl).

GatD (analyte) in running buffer (20 mM HEPES pH 8.2, 150 mM NaCl, 5 mM $MgCl_2$, 0.05% Tween-20, 1 mM DTT) was injected at 10 different concentrations using a 2-fold dilution series, with the highest concentration tested being 50 nM for the surfaces immobilized with MurT^{Cter} wild-type and 250 nM for the surfaces immobilized with MurT^{Cter} Y354A and K384A. Interaction analysis cycles consisted of a 220 seconds analyte injection (association phase) followed by 600 seconds of buffer injection (dissociation phase) at a flow rate of 30 µL·min^{-1}. All sensorgrams were processed by first subtracting the binding response recorded from the nearest buffer blank injection and all datasets were fitted to a simple 1:1 Langmuir interaction model to determine the interaction affinity and kinetic rate constants, using the provided Biacore T200 evaluation software.

3.5.7. *In vitro* lipid II amidation

Lipid II was a generous gift from E. Breukink (Univ. Utrecht). Glc*N*Ac-Mur*N*Ac-stem peptide was a generous gift from S. Fort (CERMAV, Grenoble). Lipid II amidation was followed by measuring MurT-GatD ATPase activity as previously described by Morlot *et al* [21]. ADP generated from MurT-GatD-catalysed reaction of lipid II amidation serves as substrate for the coupled reaction using pyruvate kinase and lactate dehydrogenase enzymes. In the coupled reaction, pyruvate kinase phosphorylates the generated ADP using phosphoenolpyruvate and the pyruvate thus formed is then reduced to lactate by lactate dehydrogenase along with the oxidation of NADH, that is measured by the decrease of the absorbance at 340 nm [33]. Absorbance was recorded with a thermostated Uvikon 943 double beam spectrophotometer (Kontron Instruments). Kinetic parameters were determined for each of the three MurT-GatD substrates (L-Gln, ATP and non-amidated lipid II) at 303 K. Depending on the experiments, freshly prepared reaction mixtures contained the following final concentrations of all components in buffer A without DTT, excluding the substrate under study: 3 µM of MurT-GatD, 10 mM L-Gln, 2 mM ATP, 10 µM non-amidated lipid II, or 75 µM or 4.7 mM Glc*N*Ac-Mur*N*Ac-stem peptide, 1 mM potassium phosphoenolpyruvate, 50 µg·mL^{-1} pyruvate kinase (rabbit, Roche Diagnostics), 25 µg·mL^{-1} L-lactate dehydrogenase (porcine, Roche Diagnostics), 250 µM NADH, 2 mM TCEP, 1% n-docecyl-β-D-maltoside (Anatrace) and 1.2 mM KCl. In each individual enzymatic reaction, 90 µL of the reaction mix without non-amidated lipid II were placed in a 1 cm quartz cuvette and equilibrated at 303 K. Once absorbance at 340 nm was constant, 10 µL of the non-amidated lipid II at 10-fold concentration in buffer A were added to the cuvette (t=0) and NADH oxidation was monitored. For each substrate, different protein preparations were used for independent measurements of subsets of concentrations.

Initial velocities (V_i) were assessed by linear fitting of the reaction progression curves at early time points of the reaction using the extinction coefficient of NADH at 340 nm ε = 6300 M^{-1}·cm^{-1}. The non-linear fitting of the V_i at different substrate concentrations [S] to the Michaelis-Menten equation $V_i = [S]\cdot V_{max}\cdot(K_M + [S])^{-1}$ was performed using the Kaleidagraph software (Synergy) in order to determine the apparent K_M for each substrate.

3.6. Results

3.6.1. Structure of the MurT-GatD heterodimer in solution

The *S. aureus* and *S. pneumoniae* MurT-GatD recombinant protein complexes were obtained by co-expression: for the *S. aureus* construct, a His-tag was placed at the C-terminus of GatD, while for the *S. pneumoniae* construct, a His-tag was fused to the N-terminus of MurT. Homogeneous and pure samples of the heterodimeric complexes were obtained through multi-step purification procedures [16,21]. Scattering data in solution were collected for both MurT-GatD orthologs using in-line size-exclusion chromatography (SEC) to improve sample homogeneity and remove potential protein aggregates formed during sample storage. Frames across the major peak of the SEC-SAXS profile showing a stable radius of gyration (*Rg*) were scaled and averaged to generate MurT-GatD experimental scattering curves.

The SEC-SAXS data of *S. pneumoniae* MurT-GatD yielded a single elution peak with an *Rg* of 3.0 nm (Fig. 3.3A,B) and a molecular weight estimate of 83 kDa (Table S3.2), which is close to the expected molecular weight of the *S. pneumoniae* MurT-GatD heterodimer (81 kDa). The experimental data in solution shows good agreement with the computed scattering profile of the compact *S. pneumoniae* MurT-GatD crystal structure and a poor fit to the more extended *S. aureus* MurT-GatD crystal structure (Fig. 3.3C,D). Altogether, these observations suggest that the *S. pneumoniae* MurT-GatD heterodimer also adopts a compact conformation in solution under reducing conditions, when the putative Zn-coordinating cysteines are not involved in disulfide bridges. In order to investigate the role of Zn^{2+} in the 3D structure of the complex, SEC-SAXS experiments under reducing conditions were carried out in the presence of $ZnSO_4$. The protein complex showed obvious aggregation (Fig. 3.3A,B and S3.1), which might have occurred due to a disturbance of the electrostatic potential at the surface of the complex, or to intermolecular Zn^{2+} coordination [34]. Even though aggregation is a common behaviour of zinc-binding proteins when supplemented with Zn^{2+} [34], further experiments are required to conclude how this ion may affect the structure of the MurT-GatD protein complex.

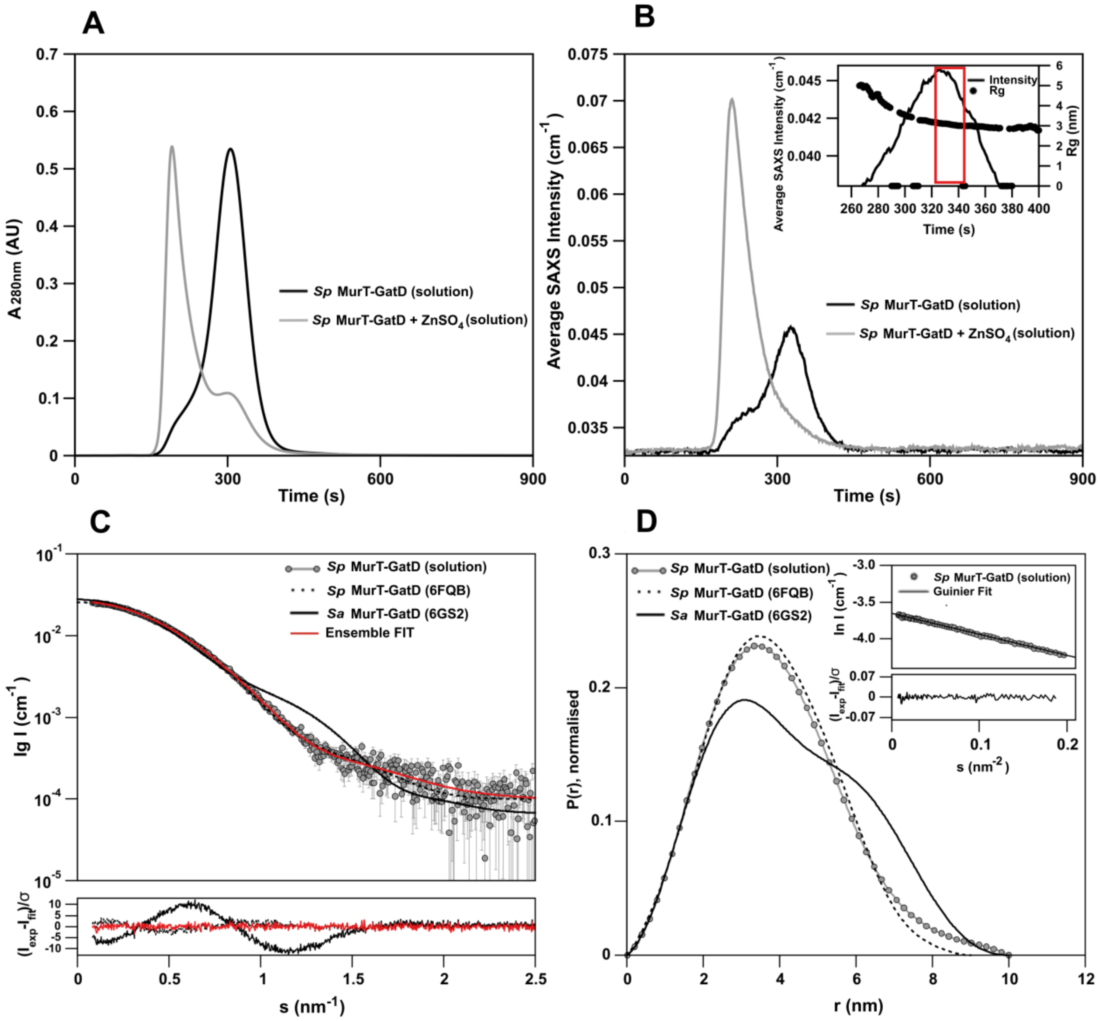


Figure 3.3. SAXS analysis of *S. pneumoniae* MurT-GatD.

(**A**) SEC profile showing the absorbance at 280 nm along the elution time for *S. pneumoniae* MurT-GatD (0.4 mL·min^{-1}, S200 5/150 column). (**B**) SEC-SAXS intensity trace for *S. pneumoniae* MurT-GatD. The insert shows the *Rg* profile across the elution peak corresponding to the heterodimeric complex, with the frames selected for SAXS analysis boxed in red. (**C**) Experimental SAXS data in solution for *S. pneumoniae* MurT-GatD (grey circles) and fits to the computed scattering profiles of the *S. aureus* (solid line, PDB ID 6GS2) and *S. pneumoniae* (dotted line, PDB ID 6FQB) MurT-GatD crystal structures. The fit to an ensemble of *S. pneumoniae* MurT-GatD complex decoy models using the EOM program is shown as a red line. Residuals corresponding to the computed model fits are included. (**D**) Real-space distance distributions calculated from the experimental data and computed fits shown in panel C. The Guinier plot for the experimental data and the residuals of the corresponding fit are shown in the insert.

The elution profile of the SEC-SAXS (Fig. 3.4A) for *S. aureus* MurT-GatD shows a single peak with stable *Rg* (Fig. 3.4B). Analysis of the SAXS data yielded an *Rg* of 3.2 nm and a molecular weight estimate of 79 kDa, consistent with the expected value (78 kDa) (Table S3.3). Due to the high background signal (Fig. 3.4B), we obtained low signal-to-noise data at low angles, which hinders the straightforward analysis of *S. aureus* MurT-GatD shape. However, it is possible to observe that the experimental scattering data shows a similar agreement with the theorical scattering profiles computed for the extended *S. aureus* MurT-GatD crystal structure (PDB ID 6GS2) and the compact *S. pneumoniae* MurT-GatD crystal structure (PDB ID 6FQB), with discrepancies in χ^2 of 1.1 and 1.2, respectively (Fig. 3.4C). A comparison of the real-space distance-distribution function of the computed heterodimer theorical profiles and the *S. aureus* MurT-GatD experimental SAXS data suggest that an extended conformation is most compatible (Fig. 3.4D). The *S. aureus* MurT-GatD maximum particle dimension (D_{max}) calculated (10.5 nm) slightly exceeds the one computed from the crystal structure (10.0 nm). It is worth to mention that modelling the missing region of *S. aureus* MurT N-terminus (38 residues) did not significantly increase the size of the heterodimer nor improve the fits to the data (Table S3.3). Altogether, these observations suggest that *S. aureus* MurT-GatD adopts in solution either an intermediate conformation between the extended and compact structures or an ensemble of extended conformations. Indeed, an intermediate conformation of this protein complex in solution has been as previously suggested [17], corroborating our conclusions.

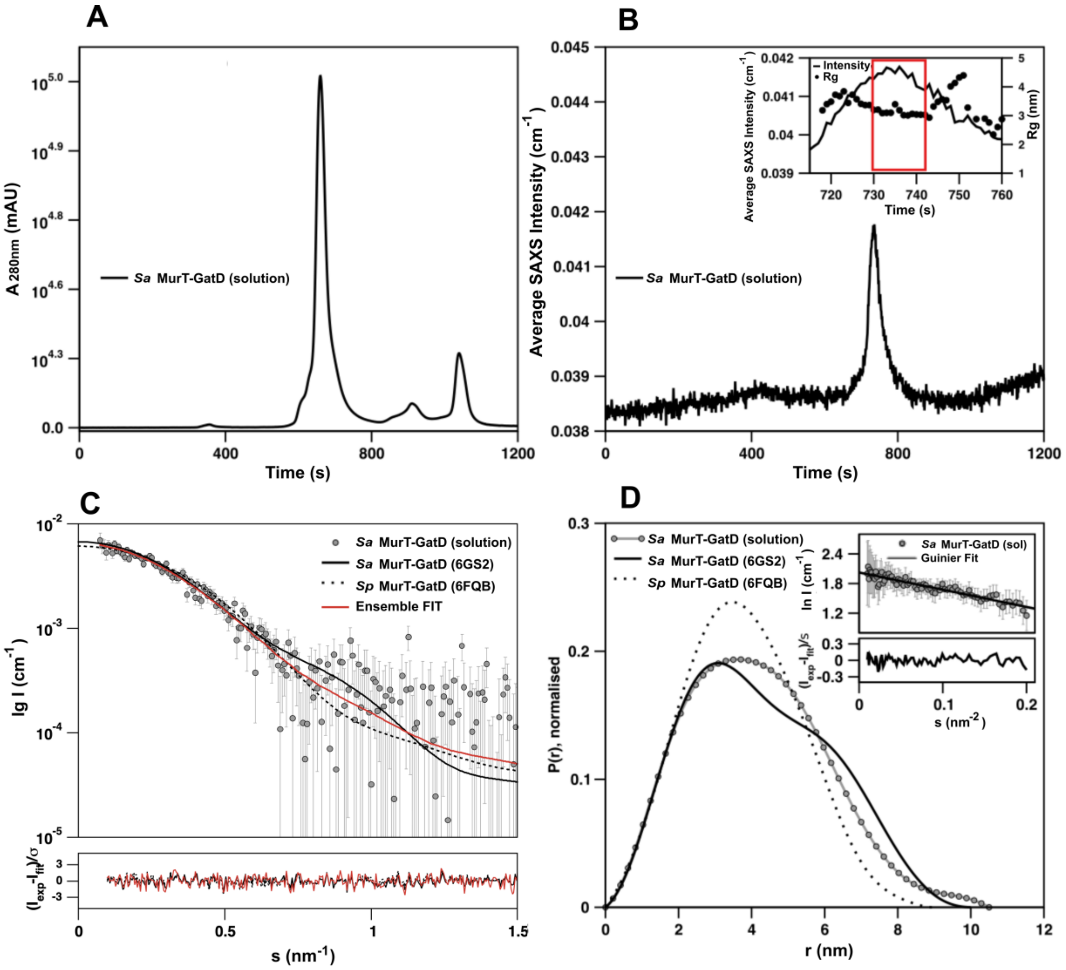


Figure 3.4. SAXS analysis of *S. aureus* MurT-GatD.

(**A**) SEC profile showing the absorbance at 280 nm along the elution time for *S. aureus* MurT-GatD (0.75 mL·min^{-1}, Bio Sec-3 300A column). (**B**) SEC-SAXS trace for *S. aureus* MurT-GatD. The insert shows the *Rg* profile across the elution peak corresponding to the heterodimeric complex, with the frames selected for SAXS analysis boxed in red. (**C**) Experimental SAXS data in solution for *S. aureus* MurT-GatD (grey circles) and fits to the computed scattering profiles of the *S. aureus* (solid line, PDB ID 6GS2) and *S. pneumoniae* (dotted line, PDB ID 6FQB) MurT-GatD crystal structures. The fit to an ensemble of *S. aureus* MurT-GatD complex decoy models using the EOM program is shown as a red line. Residuals corresponding to the computed model fits are included. (**D**) Real-space distance distributions calculated from the experimental data and computed fits shown in panel C. The Guinier plot for the experimental data and the residuals of the corresponding fit are shown in the insert.

The clear difference in conformation between the *S. aureus* and *S. pneumoniae* MurT-GatD heterodimers, shown by X-ray crystallography, suggests that the region linking the central and C-terminal domains is flexible [18]. To investigate the potential conformational flexibility of MurT in solution, an ensemble optimisation protocol using the software EOM [31,32] was employed. Considering the flexible hinge connecting the central and C-terminal domains of MurT that comprise residues T291-M299 for *S. aureus,* and R304-R306 for *S. pneumoniae*, a pool of 10,000 decoy models based on the *S. aureus* or *S. pneumoniae* MurT-GatD heterodimer structures was generated and the SAXS data was used to drive the selection of an ensemble of structures with computed scattering that best describes the experimental data. The ensemble analysis indicates that both MurT-GatD heterodimers in solution are well described by conformational equilibria (Figs. 3C, 4C and 5), with the R_{flex} flexibility metric being significant in both cases but higher for *S. aureus* (76% relative to 86% for the random pool) than for *S. pneumoniae* (68% relative to 85% for the random pool) (Fig. 3.5). Additionally, from the probability distributions that describe the conformations populating the selected ensembles (Fig. 3.5), it is clear that *S. pneumoniae* MurT-GatD predominantly exists in a confined range of compact conformations, while *S. aureus* MurT-GatD is present in an equilibrium of extended states.

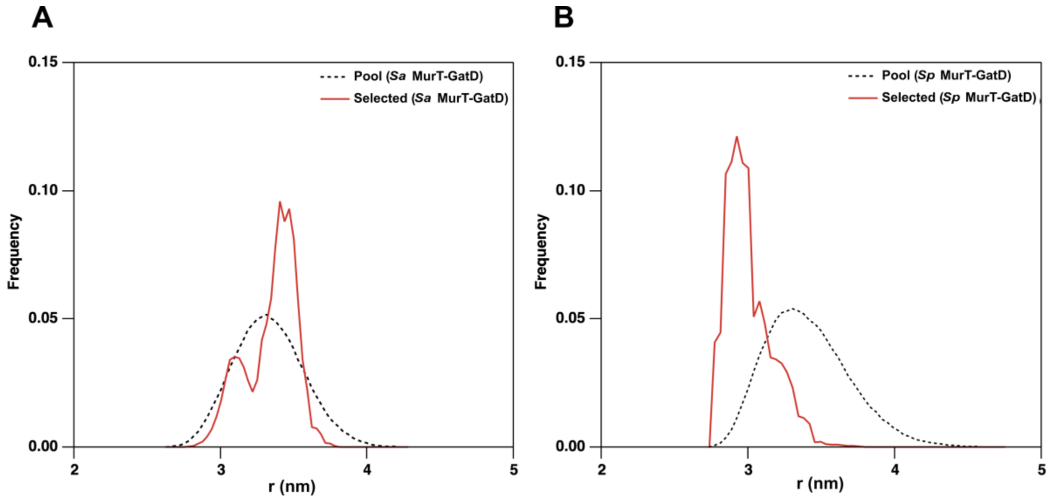


Figure 3.5. Ensemble flexibility analysis of MurT-GatD.

EOM *Rg* probability distributions of (**A**) *S. aureus* MurT-GatD and (**B**) *S. pneumoniae* MurT-GatD. These size distributions were used to quantitatively assess (R_{flex}) the flexibility of the complexes.

3.6.2. SPR analysis of the *S. aureus* MurT-GatD interface

The interaction between GatD and the C-terminal domain of MurT is conserved in the crystal structures of *S. pneumoniae* [21] and *S. aureus* [17] complexes (Fig. 3.2). In the extended conformation of the *S. aureus* complex, this interaction constitutes the full heterodimer interface (Fig. 3.6). Therefore, to quantitatively evaluate its contribution for the assembly of MurT-GatD complex, we expressed *S. aureus* GatD and MurT proteins independently and characterized their interaction by SPR. GatD from *S. aureus* is stable without its MurT partner, as shown by the determination of its crystal structure [16]. *S. aureus* MurT was fused to an AviTag at the C-terminus and biotinylated during expression in order to be used as bait in the SPR experiment. Expression and purification yielded a biotinylated product corresponding to the C-terminal domain of MurT (N-terminal sequencing indicated M257 as the starting residue), designated as MurT$^{\mathrm{Cter}}$ henceforth (Fig. S3.2). Consistent with this observation, MurT proteolysis in the absence of GatD has been observed before [21]. Nonetheless, SPR data (Fig. 3.7) show that MurT$^{\mathrm{Cter}}$ and GatD interact with nanomolar affinity (Table 3.1), indicating that MurT C-terminal domain is sufficient to form a stable complex with GatD. This high affinity is in agreement with recent pull-down and bacterial two-hybrid experiments [20]. Collectively, these observations support the extended conformation in solution of the *S. aureus* complex, where only the C-terminal domain of MurT interacts with GatD.

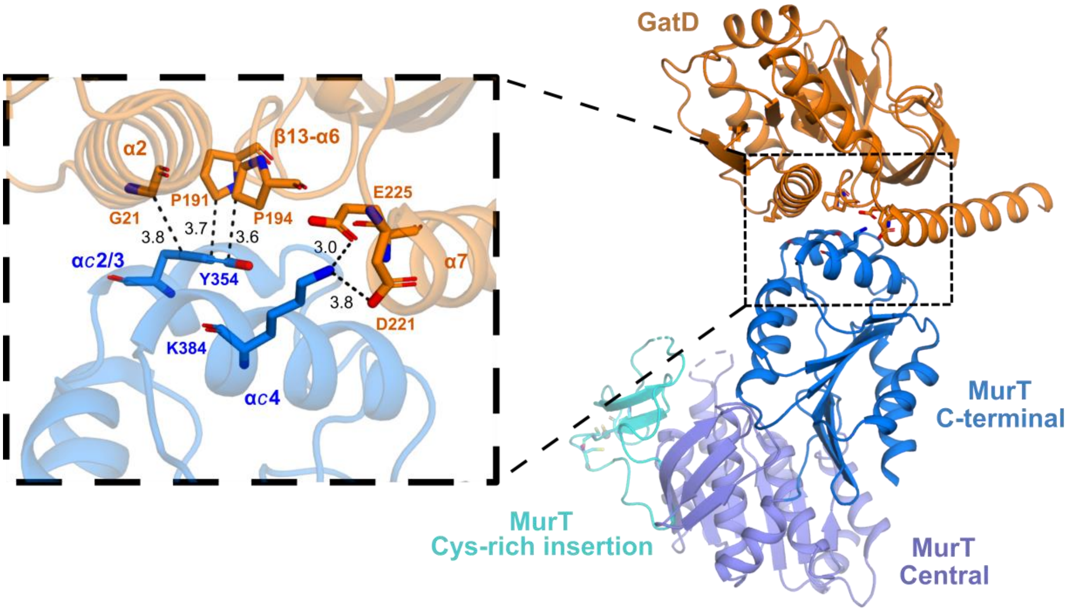


Figure 3.6. Interface of S. *aureus* MurT-GatD protein complex.

Cartoon representation of the *S. aureus* MurT-GatD crystal structure (PDB ID 6GS2), highlighting the complex interface and the intermolecular interactions (indicated in Å) between MurT residues Y354 and K384 and GatD residues, which were analysed by SPR.

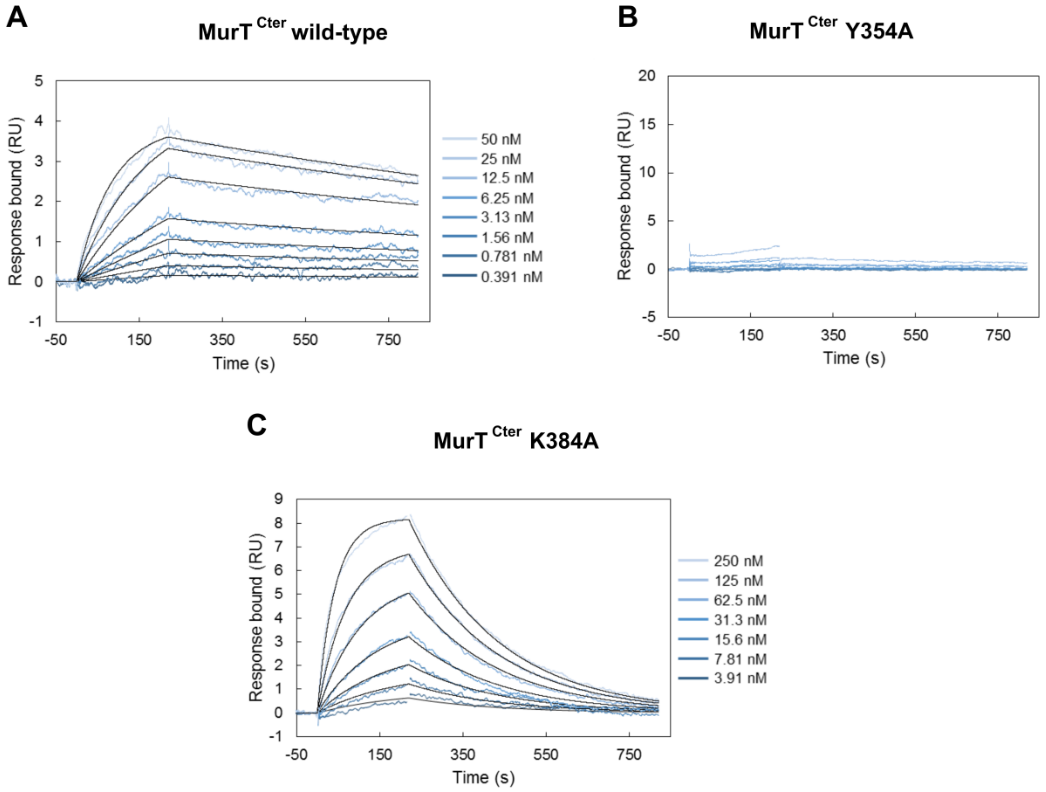


Figure 3.7. SPR binding assays between *S. aureus* MurT^{Cter} and GatD.

The sensorgrams depict the response upon injecting (t = 0 s) and washing (t = 220 s) the GatD protein at different concentrations on immobilized MurT^{Cter} proteins, including the (**A**) wild-type and (**B**) Y354A and (**C**) K384A variants.

Table 3.1. Binding affinity (K_D) and kinetic parameters (k_d, dissociation rate constant; k_a, association rate constant) of the interaction between the *S. aureus* GatD and MurT^{Cter} proteins determined by SPR.

MurT^{Cter}	**K_D (M)**	**k_d (s^{-1})**	**k_a ($M^{-1}\cdot s^{-1}$)**
wild-type	2.21×10^{-9}	5.14×10^{-4}	2.32×10^{5}
Y354A	-	-	-
K384A	6.01×10^{-8}	4.64×10^{-3}	7.72×10^{4}

The crystal structure of MurT-GatD from *S. aureus* showed that the heterodimer interface involves loop β13-α6 and α-helices α2 and α7 from GatD, and α-helices (αc2/3 and αc4) from the C-terminal domain of MurT [17]. Analysis of the molecular determinants involved in this intermolecular association revealed hydrophobic and ionic interactions (Fig. 3.6). One of the observed interactions is via *S. aureus* MurT Y354 that interacts with G21, P191 and P194 of GatD. The relevance of such hydrophobic interaction is suggested by the conservation of G21 and P191 among GatD orthologs, while MurT Y354 aligns with a conserved tryptophan in other MurT proteins [21]. In fact, *S. aureus* MurT Y354A variant was not able to interact with GatD in pull-down and bacterial two-hybrid experiments [20], and the same was observed for the equivalent *S. pneumoniae* MurT W360A [21]. As expected, the SPR sensorgrams of MurT^{Cter} Y354A with GatD showed no signal (Fig. 3.7B, Table 3.1), corroborating the idea that this mutation completely abolishes complex formation. Intermolecular salt bridges between the two *S. aureus* proteins comprise the residues MurT K384 and GatD D221 and E225 (Fig. 3.7). Although MurT K384 is conserved in most MurT orthologs, *S. pneumoniae* uses R390 to establish the same electrostatic interaction with GatD E245 [21]. As observed in Fig. 3.6C and Table 3.1, *S. aureus* MurT^{Cter} K384A variant has a 30-fold lower affinity for GatD compared with MurT^{Cter} wild-type, showing the importance of the K384 residue in the electrostatic network of the interface.

3.6.3. *In vitro* lipid II amidation by *S. aureus* MurT-GatD

We investigated the enzymatic activity of *S. aureus* MurT-GatD complex using the *in vitro* assay recently set up to study the kinetic parameters of *S. pneumoniae* MurT-GatD [21]. The activity assay is based on a pyruvate kinase/lactate dehydrogenase-coupled reaction, which ensures that the lipid II amidation reaction is rate-limiting and the ADP released from MurT-GatD activity (Fig. 3.1B) is proportional to NADH oxidation. Varying the concentration of each enzymatic substrate independently (i.e., L-Gln, ATP or lipid II) allows determining the corresponding apparent K_M by the initial velocities method (Figs. 3.8 and S3.3A,B). Due to the limited availability of lipid II, the apparent K_M of L-Gln or ATP was determined using a lipid II concentration below its calculated apparent K_M; lipid II amidation was confirmed by thin layer chromatography (Fig. S3.3C).

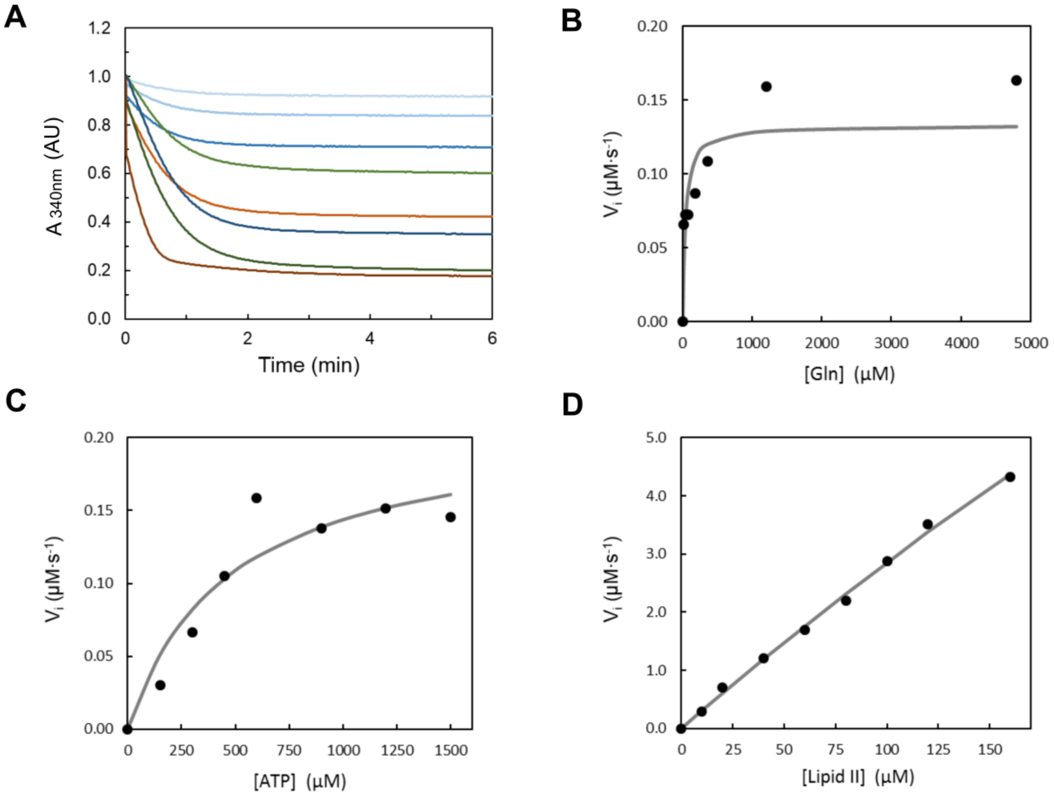


Figure 3.8. Enzymatic activity of *S. aureus* MurT-GatD.

(**A**) Measurement of ADP production (proportional to NADH consumption in the coupled reaction, which was followed by its absorbance at 340 nm) resulting from MurT-GatD activity in the presence of lipid II concentrations between 0 and 160 µM, until the reaction plateau is reached for all reactions (6 min). Blue, red and green curves are three independent measurements. The initial velocities were calculated for each reaction with various concentrations of L-Gln (**B**), ATP (**C**) or lipid II (**D**). Solid line represents the best fit of the initial velocities to the Michaelis-Menten kinetic model.

Initial enzymatic assays were performed using different concentrations of *S. aureus* MurT-GatD (0.25 and 5 µM), showing that at least 3 µM are required to reliably measure initial velocities. This result suggests that this sample is less active than the one from *S. pneumoniae* MurT-GatD, which was used at 190 nM in similar assays [21]. The apparent K_M of L-Gln determined for *S. aureus* MurT-GatD (Table 3.2) is close to the value reported for *S. pneumoniae* MurT-GatD [21], validating the experiment. The apparent K_M of ATP is in the range of values described for *S. aureus* Mur ligases (110-5400 µM for MurC-F) [35], but is, surprisingly, 30 times higher than that for *S. pneumoniae* MurT-GatD [21] (Table 3.2). Lipid II showed less affinity to *S. aureus* MurT-GatD in comparison to *S. aureus* Mur ligases and their respective UDP-Mur*N*Ac-linked substrates (41-280 μM for MurC-F) [35] and a 5-fold higher apparent K_M (Table 3.2) than *S. pneumoniae* MurT-GatD [21]. Due to limited lipid II availability, assays using higher amounts of this substrate were not carried out and, consequently, there is a large uncertainty in the estimation of the respective *Vmax* and apparent K_M. However, the shape to the curve V_i vs [lipid II] (Fig. 3.8D) suggests that the apparent K_M must be > 200 μM. Altogether, our results show that MurT-GatD has the same affinity for L-Gln in both organisms, but MurT has much lower affinity for its substrates (ATP and lipid II) in *S. aureus* than in *S. pneumoniae*.

Table 3.2. Apparent affinities of MurT-GatD complexes towards the substrates of lipid II amidation.

Apparent K_M values and respective standard errors were determined from the initial velocities.

	Apparent K_M (μM)	
	S. aureus	***S. pneumoniae* [21]**
L-Gln	56 ± 27	78 ± 8
ATP	477 ± 238	16 ± 2
Lipid II	1030 ± 642	220 ± 60

We exclude that the differences in substrate affinity between the *S. aureus* and *S. pneumoniae* complexes are due to poor sample preparation, since as reported in our previous work [16], the final *S. aureus* MurT-GatD sample contains both full-length proteins in a pure state. Moreover, both proteins form a homogeneous heterodimer in solution as observed by SAXS (Fig. 3.4A) and by size exclusion chromatography (Fig. S3.4A). Moreover, the purified complex has a melting temperature of 316K in 100 mM Tris-HCl pH 8.2 at 293 K, 500 mM NaCl, 10 mM $MgCl_2$ and 3 mM DTT (Fig. S3.4B). Altogether, these observations demonstrate the stability and overall quality of the *S. aureus* MurT-GatD heterodimer, validating the *in vitro* enzymatic activities and their differences in comparison to MurT-GatD from *S. pneumoniae*.

3.7. Discussion

The crystal structure of MurT-GatD heterodimer adopts an extended conformation in *S. aureus* [17] and a compact one in *S. pneumoniae* [21] (Fig. 3.2). Our SAXS data showed that while *S. pneumoniae* MurT-GatD has a compact conformation in solution, the heterodimer from *S. aureus* adopts an ensemble of extended conformations (Figs. 3.3, 3.4 and 3.5). The nanomolar affinity observed between *S. aureus* GatD and the C-terminal domain of MurT (Table 3.1) indicates this interface contains the required molecular determinants for complex formation, supporting the extended conformations of the heterodimer in solution.

The difference in conformation between *S. aureus* and *S. pneumoniae* heterodimers arises from the relative position of MurT central domain to the C-terminal domain. The extra interface patch observed in *S. pneumoniae* involves hydrophobic residues from MurT (Y198, I203, L204 and I211) and GatD (L6, N25, L26, I36, I54 and F77), which are not conserved in *S. aureus* nor among other orthologs [21]. The corresponding residues in *S. aureus* include several polar and charged residues: T195, R200, Y201 and L208 in MurT and D12, K13, I23, E41 and R64 in GatD. One can anticipate that the presence of positively and negatively charged residues might not favour the compact conformation.

The structure of MurT resembles Mur ligases [17,21], which are also multi-domain proteins that use ATP to sequentially build the PG stem peptide [19]. In this family of enzymes, substrate binding induces the conformational rearrangement (closure) of their three domains around the active sites. This binding is characterized by the interaction of the N-terminal domain with the UDP moiety of the UDP-Mur*N*Ac PG precursor, promoting the correct orientation of the growing stem peptide at the ATP binding site, located between the central and C-terminal domains [19]. A similar substrate binding mode is expected in MurT [18]. Therefore, the compact conformation of *S. pneumoniae* MurT-GatD with the three substrate binding sites closer to each other (Fig. 3.2) is more adapted for substrate binding and catalysis in comparison to the extended conformation of the *S. aureus* complex, which might explain the lower affinity of the latter towards ATP and lipid II (Table 3.2). For an extended structure, large conformational changes are expected to occur during catalysis. The discrepancy in the affinity of lipid II to *S. aureus* and *S. pneumoniae* MurT-GatD complexes here observed might also be related to differences in substrate specificity. Indeed, the physiological substrate of these complexes is still unknown, since the *S. aureus* MurT-GatD was able to amidate branched and non-branched lipid II *in vitro* [7]. If *in vivo* MurT-GatD amidates branched lipid II, which contain a pentaglycine in *S. aureus* and L-Ala-L-Ala or L-Ser-L-Ala in *S. pneumoniae* [5], substrate recognition and optimal amidation catalysis might depend on these differences.

The MurT N-terminal region is not visible in any of the available crystal structures of MurT-GatD [17,21], indicating that it is a mobile region. Modelling these residues using an ensemble approach against the scattering data support their flexibility (Fig. 3.5, Table S3.2 and S3.3) and suggest that this feature might be important for the function of the enzyme. Its high content of positively-charged residues has been implicated in the interaction with the pyrophosphate moiety of lipid II [21]. In fact, we observed that Glc*N*Ac-Mur*N*Ac-stem peptide lacking the pyrophosphate group cannot be amidated by *S. pneumoniae* MurT-GatD (Fig. S3.3D). Alternatively, these residues could promote the anchoring of the complex to the phospholipids of the plasma membrane, putatively explaining the *in vitro* lipid II amidation by purified *S. aureus* membranes [7,18]. Anchoring might also be accomplished through interactions with other proteins involved in PG biosynthesis and/or regulation [36].

In this study, we showed that *S. aureus* and *S. pneumoniae* MurT-GatD adopt different conformations in solution, which might impact substrate affinity. This observation suggests that different species found dissimilar strategies to accomplish PG biosynthesis. Such differences can be explored for the development of new antibiotics with a narrow range of action against pathogenic Gram-positive bacteria. The quantitative and complementary approaches implemented here to assess the affinity of MurT-GatD for its substrates and the affinity between the two components of the complex constitute valuable assays that may be employed in future studies for testing candidate drugs.

3.8. References

1. Holtje J V (1998) Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*. *Microbiol Mol Biol Rev* **62**, 181–203.

2. Vollmer W, Blanot D & De Pedro MA (2008) Peptidoglycan structure and architecture. *FEMS Microbiol Rev* **32**, 149–167.

3. Roemer T, Schneider T & Pinho MG (2013) Auxiliary factors: A chink in the armor of MRSA resistance to β-lactam antibiotics. *Curr Opin Microbiol* **16**, 538–548.

4. Sobral R & Tomasz A (2019) The Staphylococcal Cell Wall. *Microbiol Spectr* **7**, 1–21.

5. Vollmer W, Massidda O & Tomasz A (2019) The Cell Wall of *Streptococcus pneumoniae*. *Microbiol Spectr* **7**.

6. Figueiredo TA, Sobral RG, Ludovice AM, de Almeida JMF, Bui NK, 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**.

7. Münch D, Roemer T, Lee SH, Engeser M, Sahl HG & Schneider T (2012) Identification and *in vitro* analysis of the GatD/MurT enzyme-complex catalyzing lipid II amidation in Staphylococcus *aureus*. *PLoS Pathog* **8**, 1–11.

8. Zapun A, Philippe J, Abrahams KA, Signor L, Roper DI, Breukink E & Vernet T (2013) *In vitro* Reconstitution of Peptidoglycan Assembly from the Gram-Positive Pathogen *Streptococcus pneumoniae*. *ACS Chem Biol* **8**, 2688–2696.

9. Xu HH, Trawick JD, Haselbeck RJ, Forsyth RA, Yamamoto RT, Archer R, Patterson J, Allen M, Froelich JM, Taylor I, Nakaji D, Maile R, Kedar GC, Pilcher M, Brown-Driver V, McCarthy M, Files A, Robbins D, King P, Sillaots S, Malone C, Zamudio CS, Roemer T, Wang L, Youngman PJ & Wall D (2010) *Staphylococcus aureus* TargetArray: Comprehensive differential essential gene expression as a mechanistic tool to profile antibacterials. *Antimicrob Agents Chemother* **54**, 3659–3670.

10. Chaudhuri RR, Allen AG, Owen PJ, Shalom G, Stone K, Harrison M, Burgis TA, Lockyer M, Garcia-Lara J, Foster SJ, Pleasance SJ, Peters SE, Maskell DJ & Charles IG (2009) Comprehensive identification of essential *Staphylococcus aureus* genes using Transposon-Mediated Differential Hybridisation (TMDH). *BMC Genomics* **10**, 1–18.

11. Liu X, Gallay C, Kjos M, Domenech A, Slager J, Kessel SP, Knoops K, Sorg RA, Zhang J & Veening J (2017) High-throughput CRISPRi phenotyping identifies new essential genes in *Streptococcus pneumoniae*. *Mol Syst Biol* **13**, 931.

12. Thanassi JA, Hartman-Neumann SL, Dougherty TJ, Dougherty BA & Pucci MJ (2002) Identification of 113 conserved essential genes using a high-throughput gene disruption system in *Streptococcus pneumoniae*. *Nucleic Acids Res* **30**, 3152–3162.

13. Long JE, Boutte CC, Rubin EJ, Schnappinger D, Ehrt S, Fortune SM, Sassetti CM & Ioerger

R (2017) Comprehensive Essentiality Analysis of the *Mycobacterium tuberculosis* Genome. *MBio* **8**, e02133-16.

14. Figueiredo TA, Ludovice AM & Sobral RG (2014) Contribution of Peptidoglycan Amidation to Beta-Lactam and Lysozyme Resistance in Different Genetic Lineages of *Staphylococcus aureus* . *Microb Drug Resist* **20**, 1–12.
15. Mouilleron S & Golinelli-Pimpaneau B (2007) Conformational changes in ammonia-channeling glutamine amidotransferases. *Curr Opin Struct Biol* **17**, 653–664.
16. Leisico F, Vieira D V., Figueiredo TA, Silva M, Cabrita EJ, Sobral RG, Ludovice AM, Trincão J, Romão MJ, 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.
17. Nöldeke ER, Muckenfuss LM, 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.
18. Nöldeke ER & Stehle T (2019) Unraveling the mechanism of peptidoglycan amidation by the bifunctional enzyme complex GatD/MurT: A comparative structural approach. *Int J Med Microbiol* **309**, 151334.
19. Smith CA (2006) Structure, Function and Dynamics in the *mur* Family of Bacterial Cell Wall Ligases. *J Mol Biol* **362**, 640–655.
20. Gonçalves B V., Portela R, Lobo R, Figueiredo TA, Grilo IR, Ludovice AM, Lencastre H de, Dias JS & Sobral RG (2019) The role of MurT C-terminal domain in the amidation of *Staphylococcus aureus* peptidoglycan. *Antimicrob Agents Chemother* **63**, e00957-19.
21. Morlot C, Straume D, Peters K, Hegnar OA, Simon N, Villard A-M, Contreras-Martel C, Leisico F, Breukink E, Gravier-Pelletier C, Le Corre L, Vollmer W, Pietrancosta N, Håvarstein LS & Zapun A (2018) Structure of the essential peptidoglycan amidotransferase MurT/GatD complex from *Streptococcus pneumoniae*. *Nat Commun* **9**, 1–12.
22. Blanchet CE, Spilotros A, Schwemmer F, Graewert MA, Kikhney A, Jeffries CM, Franke D, Mark D, Zengerle R, Cipriani F, Fiedler S, Roessle M & Svergun DI (2015) Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J Appl Crystallogr* **48**, 431–443.
23. Pernot P, Round A, Barrett R, Antolinos ADM, Gobbo A, Gordon E, Huet J, Kieffer J, Lentini M, Mattenet M, Morawe C, Mueller-Dieckmann C, Ohlsson S, Schmid W, Surr J, Theveneau P, Zerrad L & McSweeney S (2013) Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J Synchrotron Radiat* **20**, 660–664.
24. Panjkovich A & Svergun DI (2018) CHROMIXS: automatic and interactive analysis of chromatography-coupled small-angle X-ray scattering data. *Bioinformatics* **34**, 1944–1946.
25. Franke D, Petoukhov M V, Konarev P V, Panjkovich A, Tuukkanen A, Mertens HDT, Kikhnery AG, Hajizadeh NR, Franklin JM, Jeffries CM & Svergun DI (2017) ATSAS 2.8:

a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr* **50**, 1212–1225.

26. Guinier A & Fournet G (1955) *Small-angle scattering of X-rays*.
27. Semenyuk A V & Svergun DI (1991) GNOM - a Program Package for Small-Angle Scattering Data Processing. *J Appl Crystallogr* **24**, 537–540.
28. Whitten AE, Cai S & Trewhella J (2008) MULCh: modules for the analysis of small-angle neutron contrast variation data from biomolecular assemblies. *J Appl Crystallogr* **41**, 222–226.
29. Svergun D, Barberato C & Koch MHJ (1995) CRYSOL-a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J Appl Crystallogr* **28**, 768–773.
30. Petoukhov M V, Franke D, Shkumatov A V, Tria G, Kikhney AG, Gajda M, Gorba C, Mertens HDT, Konarev P V & Svergun DI (2012) New developments in the *ATSAS* program package for small-angle scattering data analysis. *J Appl Crystallogr* **45**, 342–350.
31. Bernadó P, Mylonas E, Petoukhov M V., Blackledge M & Svergun DI (2007) Structural characterization of flexible proteins using SAXS. *J Am Chem Soc* **129**, 5656–5664.
32. Tria G, Mertens HDT, Kachala M & Svergun DI (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207–217.
33. Pullman ME, Penefsky HS, Datta A & Racker E (1960) Partial Resolution of the Enzymes Catalyzing Oxidative Phosphorylation: I. Purification and Properties of Soluble, Dinitrophenol-stimulated Adenosine Triphosphatase. *J Biol Chem* **235**, 3322–3329.
34. Poulson BG, Szczepski K, Lachowicz JI, Jaremko L, Emwas A-H & Jaremko M (2020) Aggregation of biologically important peptides and proteins: inhibition or acceleration depending on protein and metal ion concentrations. *RSC Adv* **10**, 215–227.
35. Patin D, Boniface A, Kova A, Hervé M, Dementin S, Barreteau H, Mengin-Lecreulx D & Blanot D (2010) Purification and biochemical characterization of Mur ligases from *Staphylococcus aureus*. *Biochimie* **92**, 1793–1800.
36. Jarick M, Bertsche U, Stahl M, Schultz D, Methling K, Lalk M, Stigloher C, Steger M, Schlosser A & Ohlsen K (2018) The serine/threonine kinase Stk and the phosphatase Stp regulate cell wall synthesis in *Staphylococcus aureus*. *Sci Rep* **8**, 1–13.
37. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD & Bairoch A (2005) Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook* pp. 571–607.

3.9. Acknowledgements

We thank Eefjan Breukink (Univ. Utrecht) for providing lipid II and Sébastien Fort (CERMAV, Grenoble) for Glc*N*Ac-Mur*N*Ac-stem peptide. We thank members of the Vernet laboratory (IBS, Grenoble) for advices and encouragements, Teresa Catarino (ITQB-AX NOVA) for critical discussions on the enzymatic data, Raquel Portela (UCIBIO) for assistance in site-directed mutagenesis and the N-terminal Sequencing Service, Research Facilities, ITQB-AX NOVA for the data provided. This work was supported by grants AstraZeneca IMMI 2014016 and from the Agence Nationale de la Recherche (ANR-20-CE07-0012) to AZ. This work used the platforms of the Grenoble Instruct-ERIC centre (ISBG; UMS 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-05-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). This work was supported by Fundação para a Ciência e a Tecnologia, I.P. (FCT-MCTES), in the scope of the grant PTDC/BIA-MIC/31645/2017 (awarded to RGS), fellowship PD/BD/105737/2014 (awarded to FL), project UIDP/04378/2020 and UIDB/04378/2020 of the Research Unit on Applied Molecular Biosciences (UCIBIO), project LA/P/0140/2020 of the Associate Laboratory Institute for Health and Bioeconomy (i4HB), and program iNOVA4Health – UIDB/04462/2020 and UIDP/04462/2020, through national funds and co-funded by FEDER under the PT2020 Partnership Agreement. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of any of the above mentioned funding agencies.

3.10. Author contribution Statement

Cloning and protein production were carried out by FL, ACFP and AZ. SAXS data collection and analysis were performed by FL and HDTM. SPR experiments were performed by ACFP under the supervision of PMFS and TMB. Enzymatic assays were executed by FL under the supervision of AZ, CM, MJR and TSS. The study design was conceived by FL, AML, MJR, RGS and TSS. All authors contributed for the writing of the manuscript.

3.11. Supporting Information

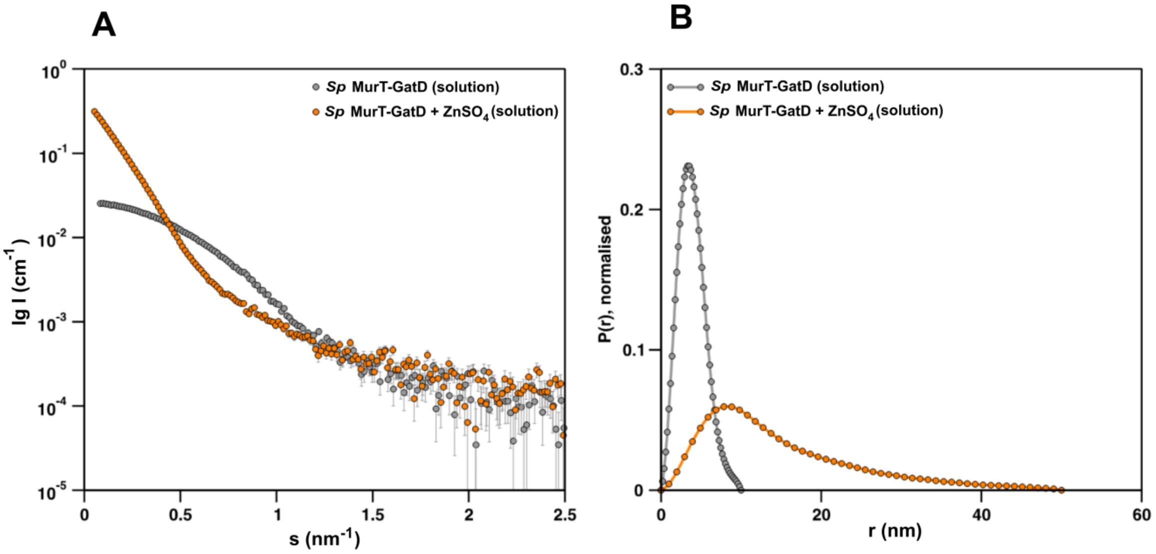


Figure S3.1. Analysis of the zinc effect on *S. pneumoniae* MurT-GatD structure in solution.

(**A**) Experimental SAXS data collected on *S. pneumoniae* MurT-GatD in the absence (grey circles) or in the presence of 0.1 mM $ZnSO_4$ (orange circles), showing a significant increase of the initial scattering intensity when zinc is added to the sample. (**B**) Real-space distance distributions calculated from the data shown in panel A. The average particle size in solution is significantly increased in the presence of zinc, indicating the presence of soluble aggregates.

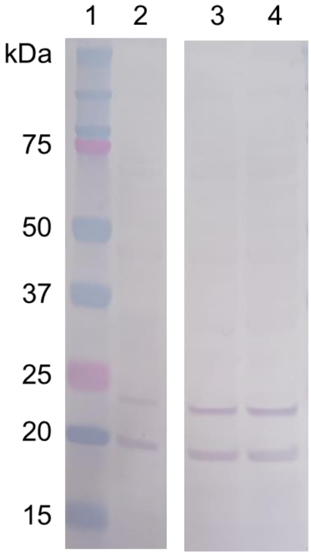


Figure S3.2. Western blotting of the biotinylated *S. aureus* MurT^{Cter} proteins used in the SPR experiments.

The biotinylated proteins were revealed using Streptavidin, Alkaline Phosphate Conjugate. Lane 1: Precision Plus Protein™ Dual Color Standards (Bio-Rad); Lane 2: Biotinylated MurT^{Cter} wild-type; Lane 3: Biotinylated MurT^{Cter} Y354A; Lane 4: Biotinylated MurT^{Cter} K384A. The bands with an estimated molecular weight of 24 and 20 kDa correspond to a truncated version of biotinylated *S. aureus* MurT (identified by N-terminal Edman Sequencing) and a biotinylated protein from the expression system (*E. coli*), respectively. Both Y354A and K384A MurT^{Cter} variants showed no further sign of proteolysis, suggesting that these mutations have no further impact on MurT^{Cter}.

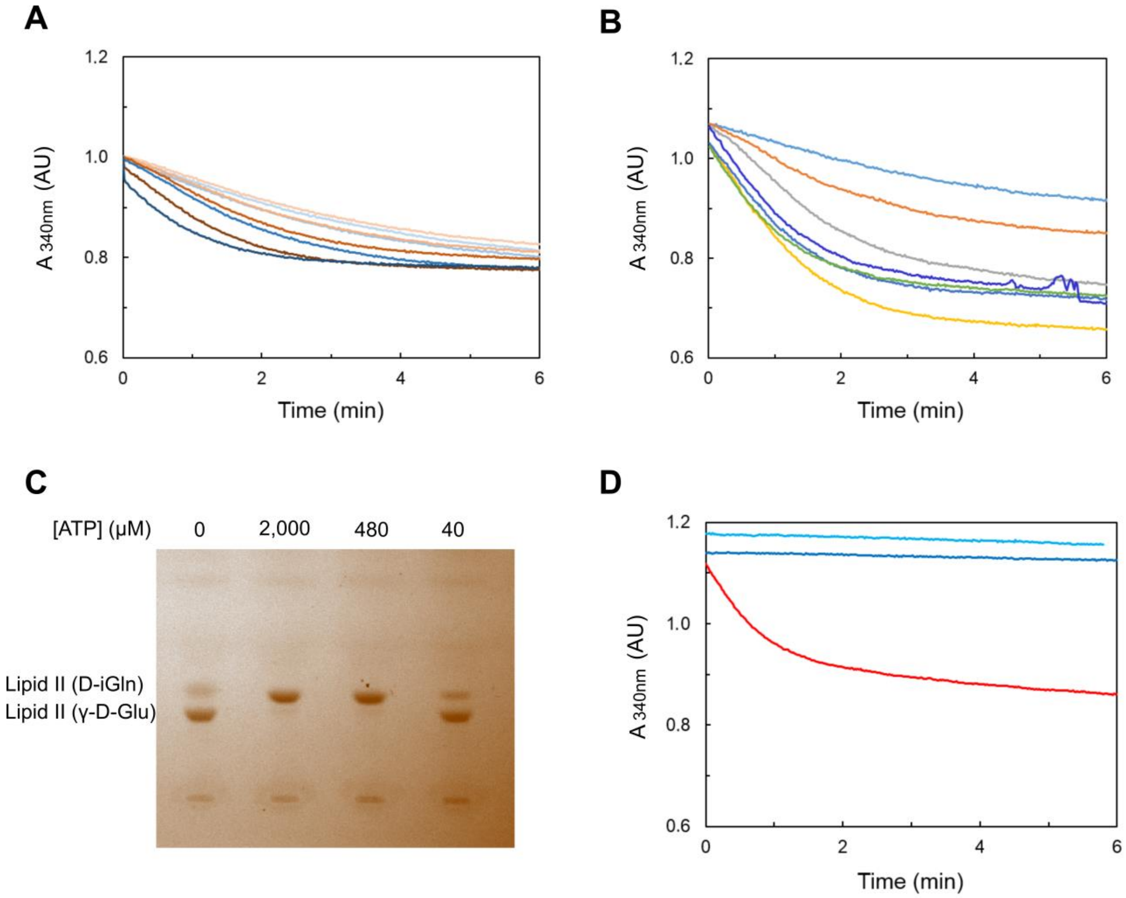


Figure S3.3. Amidation reactions catalysed by *S. aureus* and *S. pneumoniae* MurT-GatD.

Reaction curves of lipid II amidation catalysed by *S. aureus* MurT-GatD at various concentrations of (**A**) L-Gln (15-4,800 µM) blue and red curves are from two independent measurements; and (**B**) ATP (15-1,500 µM). (**C**) Lipid II amidation reactions using different concentrations of ATP were performed and the resulting products analysed by thin-layer chromatography, as described in [21]. After 10 minutes of incubation, lipid II was extracted from 100 µL aliquots with 200 µL of butanol, analysed by thin layer chromatography developed with a 88:48:10:1 chloroform/methanol/water/ammonia mixture and revealed with iodine vapour. (**D**) Amidation reaction curves catalysed by *S. pneumoniae* MurT-GatD with 50 µM lipid II (red) and two concentrations of Glc*N*Ac-Mur*N*Ac-stem peptide (75 µM in dark blue and 4.7 mM in light blue).

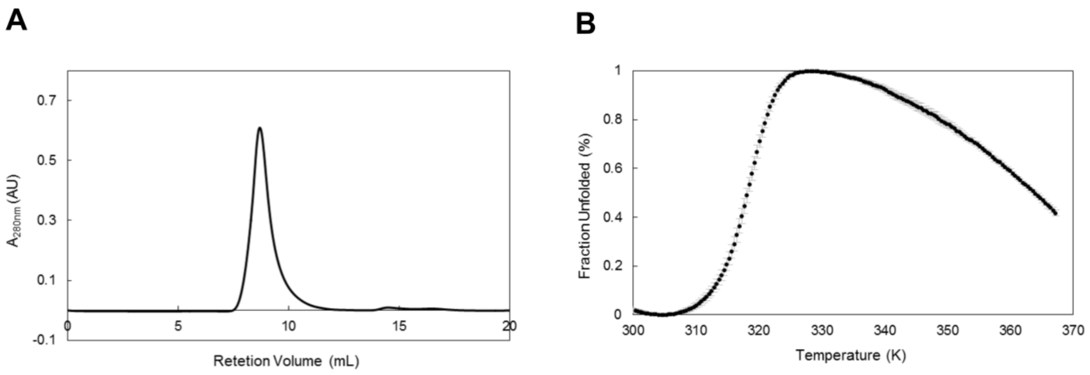


Figure S3.4. ***S. aureus*** **MurT-GatD sample quality.**

(**A**) Size-exclusion chromatography of 1.5 mg of pure heterodimer in a Superdex S200 Increase 10/300 GL (Sigma-Aldrich). (**B**) Melting curves of the purified complex at 2 μM in buffer A were measured in triplicates (standard deviation shown in grey) by Thermal Shift Assay on a StepOnePlus™ Real-Time PCR System (Applied Biosystems, Thermo Fisher Scientific) and using Protein Thermal Shift Dye (Applied Biosystems, Thermo Fisher Scientific) according to the manufacturer's specifications.

Table S3.1. Strains, plasmids and primers used in this study.

Strains and plasmids	Description^a	Source or reference
S. aureus COL	Homogeneous Mc^r	The Rockefeller Univ. Collection
E. coli BL21-CodonPlus(DE3)-RIPL	*E. coli* B F$^-$ *ompT hsdS*(r$_{B}^-$ m$_{B}^-$) *dcm*$^+$ Tet^r *gal* λ(DE3) *endA Hte* [argU proL Cam^r][argU ileY leuW Strep/Spec^r]	Agilent
pET28a	Expression vector with T7/*lac* promoter, N-terminal His tag, thrombin cleavage site, C-terminal His tag; Kan^r	Invitrogen

Primer	Sequence (5'-3')^b	Reference
His$_6$-GatD-NdeI-F	GCGCCATATGCATGAATTGACTATTTATC	This study
His$_6$-GatD-NotI-R	GCGCGCGGCCGCTTAACGAGATTTCTTCTGTC	This study

[a] Mc^r – Methicillin resistant; Tet^r – tetracycline resistance; Cam^r – Chloramphenicol resistance; Strep/Spec^r – Streptomycin/Spectinomycin resistance; Kan^r – Kanamycin resistant.

[b] Restriction sites are underlined.

Table S3.2. Small-angle X-ray scattering results for the *S. pneumoniae* MurT-GatD complex.

Sample details	MurT-GatD	MurT-GatD + 0.1 mM $ZnSO_4$
Organism	*S. pneumoniae*	*S. pneumoniae*
Source	*Escherichia coli* BL21 (DE3) Star	*Escherichia coli* BL21 (DE3) Star
UniProt sequence ID (residues in construct)	MurT: Q8DNZ9 (447); GatD: Q8DNZ8 (260)	
Extinction coefficient ε (280 nm, 0.1% w·v^{-1})	0.889	
Partial specific volume $\bar{\upsilon}$ (cm^3·g^{-1})	0.737	
Mean solute and solvent scattering length densities and mean scattering contrast $\Delta\overline{\rho}$ ($\rho_{protein}$-$\rho_{solvent}$) (10^{10}·cm^{-2})	2.709 (12.298-9.589)	2.709 (12.43-9.589)
Molecular mass *M* from chemical composition (heterodimer) (Da)	81053	81053
SEC–SAXS column, 5 × 150 mm Superdex S200 Increase (GE LifeScience)		
Loading concentration (mg·mL^{-1}) [A280nm]	6.5	5.2
Injection volume (μL)	55	55
Flow rate (mL·min^{-1})	0.4	0.4
Average concentration in combined data frames (mg·mL^{-1})	0.5	n.d.
Solvent composition (solvent blanks taken from SEC flow through prior to protein elution)	50 mM HEPES pH 7.5, 10 mM MgCl2 and 500 mM NaCl	50 mM HEPES pH 7.5, 10 mM MgCl2 and 500 mM NaCl, 0.1 mM $ZnSO_4$
SAS data collection parameters		
Instrument/Data processing	EMBL P12 (PETRA-III, DESY, Hamburg) with Pilatus6M detector [22]	
Wavelength (Å)	1.24	
Beam geometry (size, sample-to-detector distance)	0.12 × 0.25 mm^2, 3.0 m	
s-measurement range (Å^{-1})	0.002-0.7	
Absolute scaling method	Comparison with scattering from 1.2 mm pure H_2O	

Basis for normalization to constant counts	To transmitted intensity by beam-stop counter
Method for monitoring radiation damage	Frame comparison
Exposure time, number of exposures	900 s (900 × 1.00 s)
Sample temperature (K)	293

Software employed for SAS data reduction, analysis and interpretation

SAS data reduction	*I(s)* versus *s* using *RADAVER* (ATSAS 2.8.3 [30]), solvent subtraction using *PRIMUSqt* (ATSAS 2.8.3 [30])
Calculation of ε from sequence	*ProtParam* [37]
Calculation of $\Delta\overline{\rho}$ and $\bar{\upsilon}$ values from chemical composition	*MULCh web server* (http://smb-research.smb.usyd.edu.au/NCVWeb/
Basic analyses: Guinier, *P*(*r*), scattering particle volume (V_P)	*PRIMUSqt* from ATSAS 2.8.3 [30]
Atomic structure modelling	CRYSOL [29], CORAL [30]
Molecular graphics	PyMOL v2.3 MacOS 10.13.6

Structural parameters

Guinier Analysis	MurT-GatD	MurT-GatD + 0.1 mM ZnSO$_4$
I(0) (cm^{-1})	0.026 ± 0.001	0.340 ± 0.001
R_g (Å)	30.0 ± 0.1	111.5 ± 0.8
q-range (Å^{-1})	0.008-0.043	0.008-0.0092
M_r from *I*(0) (Da)	83369	n.d.
P(*r*) analysis	MurT-GatD	MurT-GatD + 0.1 mM ZnSO$_4$
I(0) (cm^{-1})	0.026 ± 0.002	0.360 ± 0.002
R_g (Å)	30.0 ± 0.1	129.5 ± 0.1
D_{max} (Å)	100 ± 5	560 ± 5
q-range (Å^{-1})	0.008-0.279	0.006-0.144
χ^2 (total estimate from *GNOM*)	1.1 (0.87)	1.1 (0.64)
M_r from *I*(0) (Da) (ratio to predicted value)	83369 (1.03)	n.d.
Volume (V_P) (Å^3)	125675	2120280

Atomistic modelling

CORAL rigid body modeling	MurT-GatD
Starting crystal structures	PDB: 6FQB
Flexible residues	-18-42, 440-447 (MurT); 248-260 (GatD)

q-range for fitting (Å)	0.008-0.25
Symmetry, anisometry assumptions	P1, none
χ^2, *CORMAP P* value	1.27, 0.00
EOM ensemble analysis	MurT-GatD
Starting crystal structures	PDB: 6FQB
Flexible residues	-18-42, 304-306 (hinge), 440-447 (MurT); 248-260 (GatD)
q-range for fitting (Å)	0.008-0.49
Symmetry, anisometry assumptions	P1, none
χ^2, *CORMAP P* value	1.06, 0.01

SASBDB IDs for data and models

S. pneumoniae MurT-GatD

SASDLZ9

Table S3.3. Small-angle X-ray scattering results for the *S. aureus* MurT-GatD complex.

Sample details

	MurT-GatD
Organism	*S. aureus*
Source	*Escherichia coli* BL21-CodonPlus (DE3)-RIPL
UniProt sequence ID (residues in construct)	MurT: A0A0H2WZQ7 (437); GatD: A0A0H2WZ38 (243)
Extinction coefficient ε (280 nm, 0.1% w·v^{-1})	0.716
Partial specific volume $\bar{\upsilon}$ (cm^{3}·g^{-1})	0.736
Mean solute and solvent scattering length densities and mean scattering contrast $\Delta\bar{\rho}$ ($\rho_{protein}$-$\rho_{solvent}$) (10^{10}·cm^{-2})	2.738 (12.326-9.588)
Molecular mass *M* from chemical composition (heterodimer) (Da)	77728
SEC–SAXS column, 7.8 × 300 mm Bio SEC-3 300A (Agilent)	
Loading concentration (mg·mL^{-1}) [A280nm]	4.1
Injection volume (μL)	40
Flow rate (mL·min^{-1})	0.75
Average concentration in combined data frames (mg·mL^{-1})	0.12
Solvent composition (solvent blanks taken from SEC flow through prior to protein elution)	100 mM Tris-HCl pH 8.2, 500 mM NaCl, 10 mM MgCl$_2$

SAS data collection parameters

Instrument/Data processing	EMBL BM29 (ESRF, Grenoble) with Pilatus1M detector [23]
Wavelength (Å)	0.99
Beam geometry (size, sample-to-detector distance)	0.7 × 0.7 mm^{2}, 2.85 m
s-measurement range (Å^{-1})	0.003-0.5
Absolute scaling method	Comparison with scattering from 1.8 mm pure H$_2$O
Basis for normalization to constant	To transmitted intensity by beam-stop counter

counts	
Method for monitoring radiation damage	Frame comparison
Exposure time, number of exposures	1200 s (1200 × 1.00 s)
Sample temperature (K)	293
Software employed for SAS data reduction, analysis and interpretation	
SAS data reduction	*I(s)* versus *s* using pipeline available at BM29 beamline, chromatogram analysis including solvent subtraction using CHROMIXS [24] and *PRIMUSqt* [30]
Calculation of ε from sequence	*ProtParam* [37]
Calculation of $\Delta\bar{\rho}$ and $\bar{\upsilon}$ values from chemical composition	*MULCh web server* (http://smb-research.smb.usyd.edu.au/NCVWeb/)
Basic analyses: Guinier, *P*(*r*), scattering particle volume (V_P)	*PRIMUSqt* from ATSAS 2.8.3 [30]
Atomic structure modelling	CRYSOL [29], CORAL [30]
Molecular graphics	PyMOL v2.3 MacOS 10.13.6
Structural parameters	
Guinier Analysis	MurT-GatD
I(0) (cm^{-1})	0.0066 ± 0.0001
R_g (Å)	31.1 ± 0.90
q-range (Å^{-1})	0.010-0.040
M_r from *I*(0) (Da)	79366
P(*r*) analysis	MurT-GatD
I(0) (cm^{-1})	0.0066 ± 0.016
R_g (Å)	32.7 ± 1.2
D_{max} (Å)	105 ± 5
q-range (Å^{-1})	0.010-0.257
χ^2 (total estimate from *GNOM*)	0.43 (0.84)
M_r from *I*(0) (Da) (ratio to predicted value)	79366 (1.02)
Volume (V_P) (Å^3)	120819
Atomistic modelling	
CORAL rigid body modeling	MurT-GatD
Starting crystal structures	PDB: 6GS2
Flexible residues	1-37, 434-437 (MurT); 247-251 (GatD)

q-range for fitting (Å)	0.008-0.20
Symmetry, anisometry assumptions	P1, none
χ^2, *CORMAP P* value	1.10, 0.09
EOM ensemble analysis	MurT-GatD
Starting crystal structures	PDB: 6GS2
Flexible residues	1-37, 291-299 (hinge), 434-437 (MurT); 247-251 (GatD)
q-range for fitting (Å)	0.008-0.20
Symmetry, anisometry assumptions	P1, none
χ^2, *CORMAP P* value	0.99, 0.32
SASBDB IDs for data and models	
S. aureus MurT-GatD	
SASDM22	

CHAPTER 4

ANALYSIS OF MULTIPLE MURT-GATD STRUCTURES DISCLOSE THE MECHANISM OF PEPTIDOGLYCAN AMIDATION

4.1. Context

In this chapter, I present the crystallization, structure determination and analysis of MurT-GatD *S. aureus* COL strain. I was responsible for protein production, crystallization and data collection and processing. A detailed comparative structural study combining the crystal structures currently available was performed, which includes the structure of MurT-GatD from *S. aureus* N315 strain, determined by competitors, and published before our structure was ready for deposition. The structure analysis here presented constitutes the basis for future *in silico*/experimental studies on structure-based drug design.

4.2. Abstract

Peptidoglycan amidation is a key reaction for cell wall assembly and, consequently, survival of Gram-positive bacteria. The bi-enzymatic complex MurT-GatD catalyzes this amidotransferase reaction using glutamine as amine donor and ATP to activate the peptidoglycan precursor lipid II. Detailed structural information of the full complex is essential to understand how it functions and, consequently, to characterize the reaction mechanism of peptidoglycan amidaiton. In this quest, the MurT-GatD complex from *S. aureus* COL strain was crystallized and the structure solved at 2.9 Å of resolution. Together with crystal structures of other MurT-GatD complexes from *S. aureus* and *S. pneumoniae*, integrated structural comparisons allowed to identify molecular determinants involved in the binding site of the different substrates, as well as, in the interface of the heterodimer. Given the essentiality of MurT-GatD complex in important human pathogens such as *S. aureus*, *S. pneumoniae* and *M. tuberculosis*, the information here gathered can guide the development of drugs with antimicrobial activity. These can be designed to target not only the function of GatD and MurT proteins, but also the assembly of the complex and its conformational dynamics. Conserved or species-specific determinants might be explored to idealize compounds with different ranges of application.

4.3. Introduction

The MurT-GatD protein complex is responsible for peptidoglycan (PG) amidation, catalysing the conversion of D-Glu, at the second position of the stem peptide, into D-isoGln [1–3]. This reaction occurs exclusively in Gram-positive bacteria [4] and is essential for cell wall assembly and, consequently, cell growth and antibiotic resistance [1–3]. These characteristics indicate that new antibiotics targeting MurT-GatD activity will act in a narrow range of microorganisms, including important pathogens such as *Staphylococcus aureus*, *Streptococcus pneumoniae* and *Mycobacterium tuberculosis*, while avoiding non-specific selective pressure on Gram-negative bacteria.

Insights into the molecular mode of action of MurT-GatD were initially obtained from protein sequence analysis [1,2]. GatD shows conservation of two catalytic residues characteristic of triad glutamine amidotransferases (GATases) [1,2]. These enzymes hydrolyse glutamine to produce ammonia, that is used by the amidotransferase domain to amidate a final acceptor [5]. On the other hand, MurT sequences contain the conserved functional motif for adenosine 5′-triphosphate (ATP) hydrolysis present in Mur ligases [1,2], the enzymes that assemble the stem peptide of PG [6]. *In vitro* reconstitution of the PG amidation reaction demonstrated that *S. aureus* and *S. pneumoniae* MurT-GatD activity is dependent on glutamine, ATP and Mg^{2+} ions [2,3], confirming its GATase activity. In addition, biochemical data showed that MurT and GatD function as an obligatory heterodimer [2], suggesting that the ammonia produced from the hydrolysis of glutamine in GatD is directly transferred to MurT and used in the amidation of the PG. The amidation reaction occurs in the cytoplasm at the lipid-bound form of PG precursors, most likely lipid II [2].

The precise and complete understanding of a reaction mechanism requires high-resolution structural information of the proteins involved. The importance of such structure-function studies was demonstrated by combining the crystal structure of *S. aureus* GatD alone and *in vitro* assessment of *S. aureus* MurT-GatD glutaminase activity [7]. It was shown that a conserved arginine is likely responsible for sequestering glutamine; and that the active site is composed by a cysteine and a histidine residue. However, no glutamate residue was found nearby to complete the characteristic triad arrangement of GATases [7]. Consequently, unravelling the complete reaction mechanism of MurT-GatD GATase activity, which includes the ATP and lipid II binding sites, the ammonia channel and dynamics between active sites, requires structural data on the full MurT-GatD heterodimer. Given the similarities between MurT and Mur ligases, the extensive data gathered on the structure and function of Mur ligases is also of great importance to shed light on the reaction mechanism of PG amidation.

The Mur ligases MurC-F catalyse the ATP-dependent stepwise ligation of five amino acids to the UDP-Mur*N*Ac precursor of PG. In *S. aureus*, they catalyse the addition of L-Ala (MurC), D-Glu (MurD), L-Lys (MurE) and D-Ala-D-Ala dipeptide (MurF) [8]. Mur ligases have a common reaction mechanism that involves three main steps (Figure 4.1): first, the activation of the acceptor carboxyl group of the UDP-Mur*N*Ac precursor (corresponding to the D-lactyl of Mur*N*Ac for MurC and the α-carboxyl of the previously added amino acid for MurD-F) by phosphorylation, leading to the formation of an acyl-phosphate intermediate; second, the nucleophilic attack of the α-amine group of the condensing amino acid or dipeptide, yielding a tetrahedral intermediate; and, finally, the breakdown of this intermediate to produce the extended UDP-Mur*N*Ac precursor, with the release of inorganic phosphate and ADP (see [9] and references therein). As shown in Figure 4.1, besides the ubiquitous Mg^{2+} ion stabilizing the γ- and β-phosphates of ATP, a second Mg^{2+} ion is required to neutralize the negative charges of the ATP γ-phosphate and the acceptor carboxyl group of the substrate (UDP-Mur*N*Ac precursor), favouring the phosphoryl-transfer reaction [10]. Enzymatic studies initially performed in MurC [11] and MurF [12] established a consistent sequential ordered kinetic mechanism in which ATP binds the free enzyme first, followed by the UDP-Mur*N*Ac precursor substrate and, finally, the condensing amino acid/dipeptide [13]. Binding of ATP and the UDP-Mur*N*Ac precursor induce the three domains of Mur ligases to move towards each other and form the active site at their junction, where the terminal carboxyl group of UDP-Mur*N*Ac precursor is accommodated close to the γ-phosphate of ATP for the phosphorylation reaction [6]. Besides their functional characterization, crystal structures in the presence of substrate and product are available for all Mur ligases (except for the MurF-product complex).

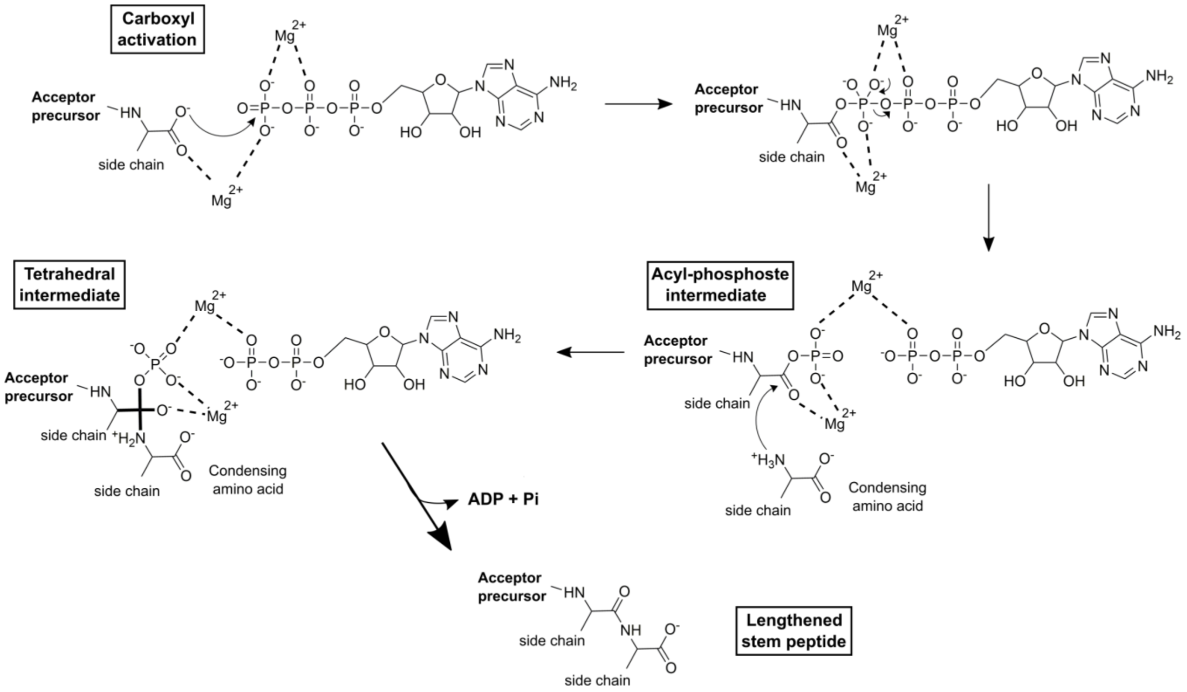


Figure 4.1. Condensation reaction catalysed by Mur ligases.

Conserved reaction mechanism catalysed by Mur ligases in the formation of the PG stem peptide. The carboxyl group of the UDP-Mur*N*Ac precursor substrate is activated by phosphorylation, forming an acyl-phosphate intermediate. This undergoes a nucleophilic attack by the condensing amino acid and forms a tetrahedral intermediate, which then leads to the release of a UDP-Mur*N*Ac precursor with a lengthened stem peptide. The reaction relies on two Mg^{2+} ions. Adapted from [14].

In order to obtain relevant structural information on the MurT-GatD heterodimer and further identify its molecular determinants, we crystallized and determined the structure of the complex from *S. aureus* COL strain. In this work we perform a detailed structure analysis of the two structures currently available for the MurT-GatD complex: the heterodimer from *S. aureus* N315 strain in its apo form and complexed with adenylyl-imidodiphosphate (AMP-PNP, a non-hydrolysable analogue of ATP) [15] and *S. pneumoniae* MurT-GatD complexed with glutamine [16]. This structural characterization provides functional insights on the reaction mechanism of lipid II, which can be used in the design of new molecules with antimicrobial activity.

4.4. Methods

4.4.1. *S. aureus* MurT-GatD production

Cloning of *S. aureus* MurT-GatD had been previously performed at the Laboratory of Molecular Microbiology of Bacterial Pathogens (UCIBIO – FCT-UNL) led by Prof. Rita Sobral as described in Leisico *et al*. [7]. The *murT-gatD* coding regions from *S. aureus* COL strain were cloned in the plasmid pET28a using the restriction sites of *NcoI* and *XhoI*, resulting in the co-expression of untagged MurT and GatD tagged at the C-terminal (Figure S4.11).

The *S. aureus* MurT-GatD heterodimer was expressed and purified according to the protocol reported in Leisico *et al*. [7]. Highly pure protein was concentrated up to 15 mg·mL^{-1} for the crystallization trials.

4.4.2. Crystallization

Initial crystallization trials were performed using the High-Throughput Crystallization platform at EMBL-Grenoble. Crystallization drops of 200 nL with a 1:1 ratio of MurT-GatD (at 15 or 7.5 mg·mL^{-1} in buffer A - 100 mM Tris-HCl pH 8.2 at 293 K, 500 mM NaCl, 10 mM $MgCl_2$ and 3 mM DTT) and crystallization solution were set up at 293 K using the sitting-drop vapor diffusion technique. The high-throughput potential of this platform allowed testing primary sparse matrix screens and secondary PEG-based refinement screens, simultaneously, summing up to 576 crystallization conditions. The first set included the screens Classics Suite (Qiagen/Nextal), JCSG-*plus* (Molecular Dimensions), Salt Grid (Hampton Research) and Wizard I&II (Rigaku Reagents), which permitted to assess the effect of a broad range of salts, polymers, organic molecules, buffer composition and pH on *S. aureus* MurT-GatD crystallization. The refinement screens PACT premier (Molecular Dimensions) and PEGs-I (Qiagen/Nextal) aimed in promoting protein-protein contacts due to molecular crowding and test different PEG/Ion environments. Within two weeks, microcrystals could be observed under light microscope in several conditions containing PEG 3,350 as precipitant.

A second round of crystallization trials were performed using the PEG-based screens JBScreen Classic 1, 2, 3 and 4 (Jena Bioscience) and JCSG-*plus* (Molecular Dimensions), accounting for a total of 192 crystallization conditions. These screens were tested immediately after protein purification to ensure the best quality of the sample. Sitting-drops were set up at 293 K using the crystallization robot Oryx 8 (Douglas Instruments Limited) by mixing 0.5 μL of MurT-GatD at 13 mg·mL^{-1} in buffer A with 0.5 μL of crystallization solution and equilibrated over 50 μL of reservoir solution. After 3 days at 293 K, protein crystals were observed in the drop with the crystallization condition containing 30 % (w/v) PEG 4,000, 100 mM Tris-HCl pH 8.5 and 200 mM $MgCl_2$. The crystals were harvested and flash-cooled in liquid nitrogen five days after the crystallization set up using the reservoir solution as cryoprotectant.

4.4.3. Data collection and structure determination

X-ray diffraction data was collected on beamline ID23-2 at the European Synchrotron Radiation Facility (ESRF), France. The *S. aureus* MurT-GatD crystals belong to space group P222 with unit cell dimensions of a = 107.09 Å, b = 111.09 Å and c = 112.42 Å, and diffract beyond 3.0 Å of resolution. Based on the Matthews coefficient [17], the crystal contained 2 molecules of the MurT-GatD heterodimer per asymmetric unit corresponding to 41 % of solvent content. The crystallographic data was integrated with *XDS* software package [18] and scaled with *Aimless* software [19] from the *CCP4* program suite [20]. The parameters of data collection and processing statistics are presented in Table 4.1..

Table 4.1. X-ray crystallography data collection and processing for *S. aureus* COL MurT-GatD.

Values in parenthesis correspond to the processing statistics at the highest resolution shell.

Data collection	
X-ray Source	ID23-2 (ESRF, France)
Wavelength (Å)	0.8731
Temperature (K)	100
Detector	Pilatus3_2M (Dectris)
Processing	
Space group	P222
a, *b*, *c* (Å)	107.09, 111.09, 112.42
Number of heterodimers per asymmetric unit	2
Matthews coefficient ($Å^3$/Da)	2.09
Mosaicity (º)	0.23
Resolution range (Å)	49.31 – 2.90 (3.06 – 2.90)
Number of observed reflections	209,578 (32,288)
Number of unique reflections	30,388 (4,366)
Completeness (%)	99.99 (100.00)
Multiplicity	6.9 (7.4)
$CC_{1/2}$	0.99 (0.58)
R_{pim}	0.113 (0.696)
Mean $(I)/\sigma(I)$	6.7 (1.7)
Wilson *B* factor ($Å^2$)	50.01

Structure determination was pursued by molecular replacement using *Phaser* software [21]. The templates used in the different attempts were modified using search model optimization protocols within *Chainsaw* [22] and *Sculptor* [23] software. The structure of MurT and GatD proteins from *S. pneumoniae* R6 strain (PDB ID 6FQB; Uniprot Q8DNZ9 and Q8DNZ8, respectively) and *S. aureus* N315 strain (PDB ID 6GS2; Uniprot A0A0H3JN63 and A0A0H3JUU7, respectively) were tested as search models. Only the latter provided a complete solution for the structure of *S. aureus* COL strain. As expected, two copies of MurT-GatD complex were found in the asymmetric unit.

Model building was carried out manually using Coot [24] and alternated with reciprocal space refinement using phenix.refine [25]. Refinement was performed considering one TLS group per chain and both chains of the same protein as one torsion non-crystallographic symmetry group. Water molecules were automatically picked using phenix.refine and manually inspected using Coot. The final model included 397 residues in each chain of MurT, 251 residues in each of chain of GatD, two zinc ions, one chloride ion and 75 water molecules. The cross-validation testing data (R_{free}) representing 5% of the total number of reflections was randomly selected and the final model shows an R_{work} and R_{free} factors of 0.24 and 0.26, and an RMSD of 0.002 Å and 0.486° for the bond lengths and bond angles, respectively. The geometry of the final model was assessed by the MolProbity software [26]. The model quality and refinement statistics are presented in Table 4.2. The final model was deposited in Protein Data Bank (PDB) with the accession code 7Q8E and its release is on hold until publication.

Table 4.2. Refinement statistics of *S. aureus* COL MurT-GatD crystal structure.

Refinement	
Resolution (Å)	49.31 – 2.90
Number of reflections	30,292
R_{work}	0.240
R_{free}	0.262
Overall Mean B Value (Å^2)	73.58
Number of atoms	
Protein	10,056
Ligand/ion	3
Water	75
RMSD from ideal values	
Bond lengths (Å)	0.002
Bond angles (°)	0.486
Protein geometry	
Rotamers outliers (%)	0.56
Favoured rotamers (%)	97.38
Ramachandran outliers (%)	0.24
Ramachandran favoured (%)	95.05
PDB accession code	**7Q8E**

4.4.4. Structure analysis

Superpositions of the crystal structures were performed by topology-based alignment in Coot [24]. The secondary structure nomenclature was considered as described in the original publications of the crystal structure of isolated *S. aureus* GatD [7], *S. aureus* MurT-GatD [15] and *S. pneumoniae* MurT-GatD [16]. Interface analysis involving the characterization of intermolecular interactions and buried surface area was performed using the PISA server [27].

4.5. Results and discussion

4.5.1. Crystal structure of MurT-GatD from *S. aureus* COL strain

The co-expression of MurT and GatD proteins from *S. aureus* COL strain leads to the formation of a heterodimer that is eluted throughout the purification without dissociation, as obtained elsewhere [2,7,15] (Chapter 3). The optimized protocol of expression and purification yields approximately 10 mg of MurT-GatD/L of culture, in a homogeneous, stable and pure form [7] (Chapter 3). The quality of the final sample allowed concentrating the complex to 15 mg/mL, which was used in crystallization experiments.

The crystals of *S. aureus* MurT-GatD complex obtained using PEG 4,000 as precipitant, diffracted at 2.9 Å (Table 4.1 and Figure S4.2S4.2). Crystals belongs to P222 space group and contain 2 molecules of the heterodimer per asymmetric unit. The initial attempts to solve the structure were performed using the crystal structure of MurT-GatD complex from *S. pneumoniae* R6 strain (PDB ID 6FQB) as search model, since these proteins have 33 and 34 % of sequence identity with the corresponding GatD and MurT orthologs, respectively. Molecular replacement was performed using the complete heterodimer model and that of each individual proteins but these strategies yielded only partial solutions for *S. aureus* GatD. Suboptimal solutions containing part of *S. aureus* MurT were explored through model building within ARP/wARP [28] without success. The automated pipeline for molecular replacement Mr. Bump [29,30], which explores the full PDB database to find the best possible search model based on sequence alignment and uses specific optimization protocols for search model processing, did not provide better solutions.

To our surprise, and during our attempts to solve the structure, the crystal structure of MurT-GatD from *S. aureus* NT315 strain (Uniprot A0A0H3JN63 and A0A0H3JUU7, respectively; PDB ID 6GS2) [15] was reported. Molecular replacement using the models of GatD and MurT separately yielded a full solution for the crystal structure of MurT-GatD complex from *S. aureus* COL. Structure refinement (Table 4.2) was carried out using low-resolution protocols [25] and the final model deposited at the PDB under the code 7Q8E.

As expected, the final model of MurT-GatD complex from *S. aureus* COL strain is very similar to that from *S. aureus* NT315 strain with a RMSD of 0.40 Å and 0.59 Å for all

MurT and GatD Cα-atoms, respectively (Figure S4**.3.** S4.3). Minor differences are found in regions at the surface of the proteins with poor electron density due to intrinsic flexibility/disorder. At the sequence level, only the residue #39 of MurT (I39 in COL and V39 in N315) differs between the proteins of the two *S. aureus* strains, however it belongs to the flexible N-terminal region with poorly defined electron density that prevents their discrimination.

The similarity between the MurT-GatD protein complex from *S. aureus* COL/N315 strains is not surprising since both were purified in similar buffers (Tris-HCl pH 8.2, NaCl and $MgCl_2$) and crystallized using almost the same conditions, in the same space group. The only difference is the temperature of the crystallization experiments, 293 K for the crystals of MurT-GatD from *S. aureus* COL and 277 K for the ones of N315 strain. This might explain the difference in resolution between the final structures, which correspond to 2.9 and 2 Å, respectively. Since the structural information at higher resolution offers more confidence for structure-function studies, only the crystal structure of MurT-GatD from *S. aureus* N315 strain will be considered in the analysis carried out in this chapter.

4.5.2. Overall structure of MurT-GatD heterodimer

The crystal structure of MurT-GatD from *S. aureus* [15] and *S. pneumoniae* [16] have different overall shapes: while *S. aureus* MurT-GatD structure is an heterodimer with a boomerang-like extended shape, the complex from *S. pneumoniae* adopts a compact globular form (Figure 4.2). The different conformations explain the difficulties in our structure determination attempts using the model of *S. pneumoniae*.

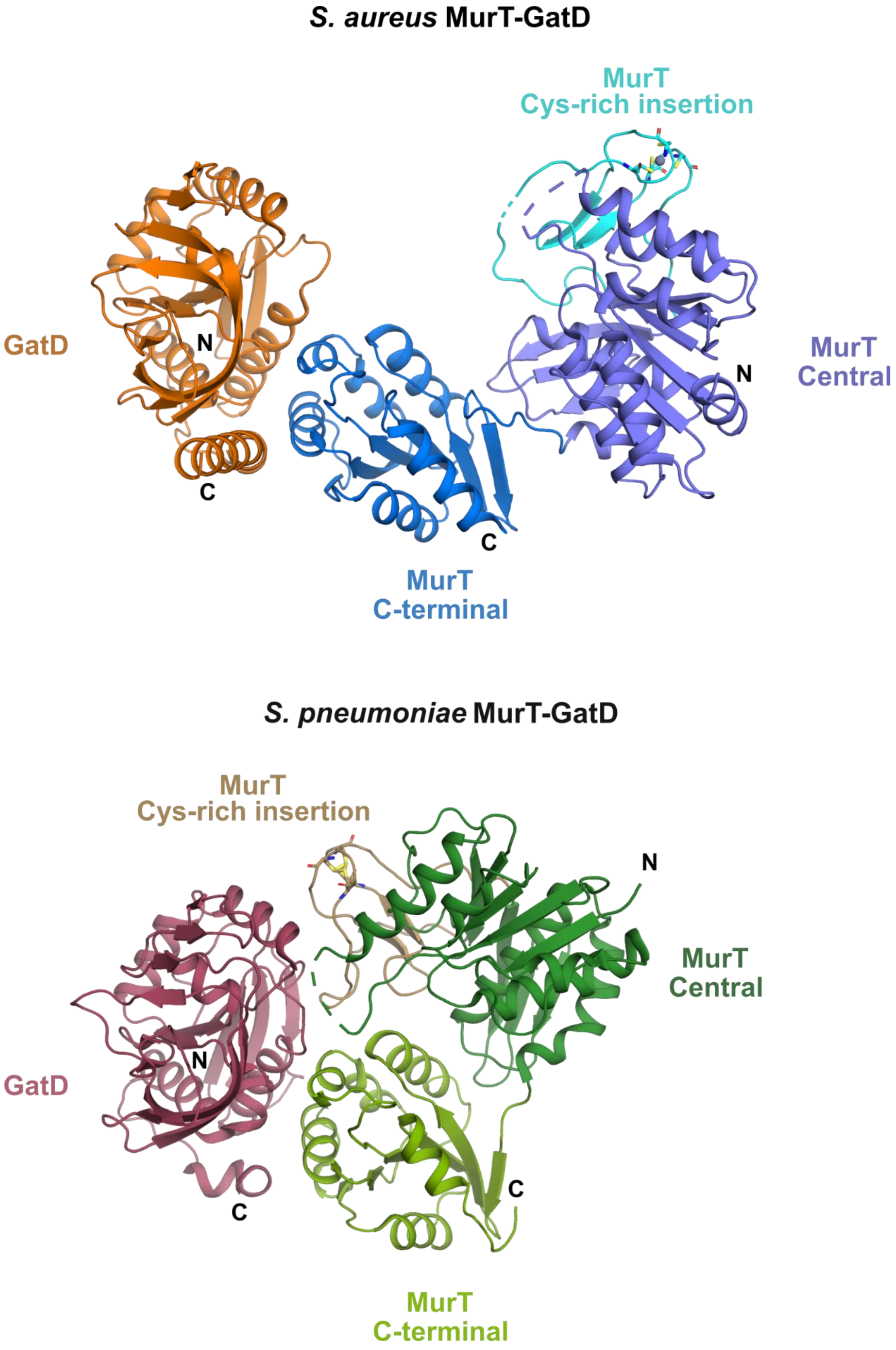


Figure 4.2. Overall structures of *S. aureus* and *S. pneumoniae* MurT-GatD heterodimers.

Cartoon representation of the *S. aureus* MurT-GatD crystal structure (PDB ID 6GS2/chains B-A, respectively): GatD (full-length of 243 residues), MurT central domain (residues 38-184+235-290), Cys-rich insertion (residues

185-234) and C-terminal domain (residues 291-433) are coloured orange, purple, cyan and blue, respectively; and *S. pneumoniae* MurT-GatD crystal structure (PDB ID 6FQB/chains A-B, respectively): GatD (residues 1-247), MurT central domain (residues 43-186+239-303), Cys-rich insertion (residues 187-238) and C-terminal domain (residues 304-439) are coloured red, dark green, brown and light green, respectively. The conserved cysteine residues of the Cys-rich insertions are shown as atom-coloured sticks and the Zn^{2+} ion present in the *S. aureus* heterodimer is shown as a grey sphere. The dashed region of the cartoon represents residues that could not be modelled due poor electron density. The N- and C-terminus of each protein are labelled with the N and C letters, respectively.

In spite of the global differences of the MurT-GatD complex from *S. aureus* and *S. pneumoniae*, the superposition of the different proteins and domains highlight their similarities: GatD displays the characteristic open α/β fold of the glutaminase domains of triad GATases [31], with an extended C-terminus [7,15,16] (Figure 4.3A); MurT exhibits the central and the C-terminal domains characteristics of Mur ligases [15,16] (Figure 4.3B,C). The MurT proteins from *S. aureus* and *S. pneumoniae* display in their central domains a 50-residue insertion that is absent in Mur ligases (Figure 4.3B). This insertion is denominated as Cys-rich insertion [16] since it comprises four cysteines highly conserved in other MurT orthologs (Figure S4.44). The N-terminal region of MurT in both crystal structures proteins could not be modelled due to flexibility/disorder.

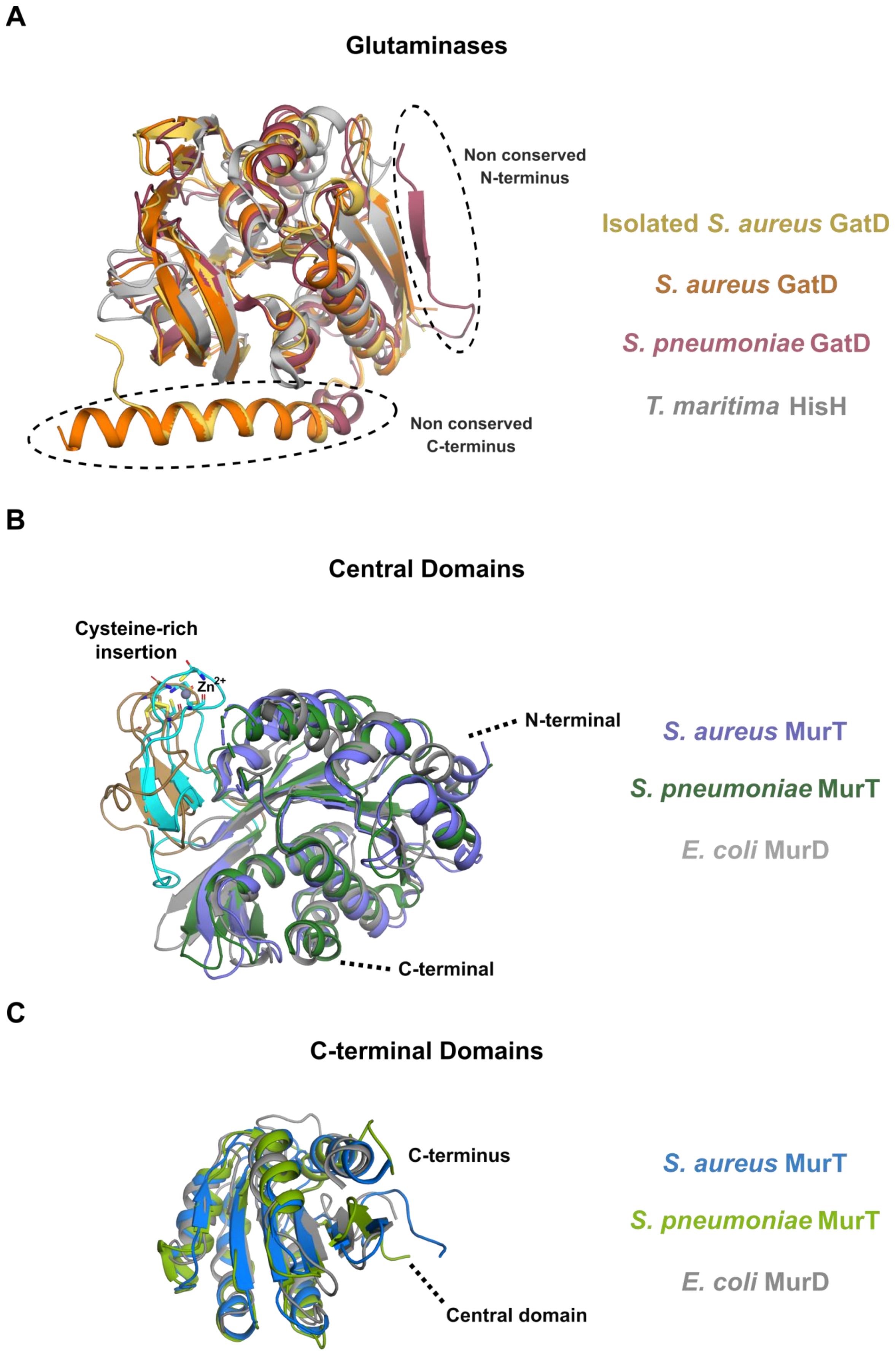

A
Glutaminases
Non conserved
N-terminus
Isolated S. aureus GatD
S. aureus GatD
S. pneumoniae GatD
T. maritima HisH
Non conserved
C-terminus
B
Central Domains
Cysteine-rich
insertion
Zn2+
N-terminal
S. aureus MurT
S. pneumoniae MurT
E. coli MurD
C-terminal
C
C-terminal Domains
C-terminus
S. aureus MurT
S. pneumoniae MurT
E. coli MurD
Central domain

Figure 4.3. Structural conservation of GatD proteins and MurT central and C-terminal domains.

(**A**) Superposition of GatD from *S. aureus* MurT-GatD complex (PDB ID 6GS2/chain A in orange), isolated *S. aureus* GatD (PDB ID 5N9M/chain A in yellow), GatD from *S. pneumoniae* MurT-GatD complex (PDB ID 6FQB/chain B in red) and glutaminase HisH protein from *T. maritima* Imidazole Glycerol Phosphate Synthase complex (PDB ID 3ZR4/chain B in grey). (**B**) Superposition of the central domains of MurT from *S. aureus* MurT-GatD complex (PDB ID 6GS2/chain B in purple), MurT from *S. pneumoniae* MurT-GatD complex (PDB ID 6FQB/chain A in dark green) and *Escherichia coli* MurD (PDB ID 1E0D/chain A in grey). The Cys-rich insertions, which are exclusive of MurT proteins, are coloured cyan in *S. aureus* and brown in *S. pneumoniae*; their conserved cysteine residues are shown as atom-coloured sticks, while the Zn^{2+} ion coordinated by *S. aureus* MurT is shown as grey sphere. (**C**) Superposition of the C-terminal domains of MurT from *S. aureus* MurT-GatD complex (PDB ID 6GS2/chain B in blue), MurT from *S. pneumoniae* MurT-GatD crystal structure (PDB ID 6FQB/chain A in light green) and *E. coli* MurD crystal structure (PDB ID 1E0D/chain A in grey). In all panels, the superposed structures are represented as cartoon, at the same size scale, with indication of the connections to their respective binding regions.

Comparison of the GatD structures from *S. aureus* and *S. pneumoniae* heterodimers indicated a RMSD of 1.64 Å for 208 out of 248 residues. The *S. aureus* GatD full-length sequence was modelled in the electron density maps, while the C-terminal region of *S. pneumoniae* GatD (248-260) could not be modelled, likely due to disorder (Figure 4.3A). *S. pneumoniae* GatD possesses an extended N-terminal β-sheet when compared to other GatD proteins (Figure 4.3.3A). The overall fold of GatD in the heterodimer is conserved with the crystal structure of isolated *S. aureus* GatD (Figure 4.3A), suggesting that only a few conformational rearrangements occur upon MurT binding.

The central and the C-terminal domains of MurT are conserved between *S. aureus* and *S. pneumoniae* heterodimers (Figure 4.3B,C), with RMSDs of 1.88 Å for 157 out of 247 residues and 1.68 Å for 127 out of 143 residues, respectively. The similarity between the individual MurT domains indicates that the different conformations of *S. aureus* and *S. pneumoniae* MurT-GatD heterodimers are related by rigid body movements of the domains (Figure 4.2). Indeed, SAXS experiments showed that the *S. aureus* MurT-GatD complex in solution adopts an intermediate conformation between the extended and compact crystal structures of *S. aureus* and *S. pneumoniae* MurT-GatD heterodimers, respectively, confirming the flexibility of MurT domains [15] (Chapter 3).

The inter-domain flexibility in Mur ligases is provided by hinge loops connecting the different domains [6]. It has been suggested that a conserved arginine at the hinge between the central and C-terminal domains is responsible for stabilizing the compact conformation of Mur ligases through hydrogen-bonds with neighbouring atoms of both domains [32]. Interestingly, that arginine residue is conserved in MurT proteins - *S. aureus* R298 and *S. pneumoniae* R304 - and is localized at the loop connecting the central and C-terminal domains. However, it establishes the same number of hydrogen-bonds in both crystal structures, suggesting that a different molecular mechanism is required to stabilize the compact conformation of *S. pneumoniae* MurT protein.

4.5.3. Cys-rich insertion

The Cys-rich insertion is a unique feature of MurT proteins. Sequence alignments have shown that a tetrad of Cys residues with the motif C-X_2-C-X_{18}-C-X_{1-2}-C (X is any amino acid) is conserved in Bacilli, but is replaced by serine and aspartate residues in other taxonomic classes of Firmicutes [16]. Gram-positive bacteria from the *Actinobacteria* phylum possess insertions with variable length and some completely lack the cysteine residues [33]. This variability suggests that the role of the Cys-rich insertion is not conserved among MurT proteins as other functional motifs that possess higher sequence conservation, such as the ATP binding site [33].

The crystal structures of *S. aureus* and *S. pneumoniae* MurT-GatD showed that although the two heterodimers crystallized in different conformations, their Cys-rich insertions share the same structure (Figure 4.4). Only the initial loop, which is located at the interface with GatD in the *S. pneumoniae* compact heterodimer, adopts a different orientation (Figure 4.2). The Cys-rich insertion is structured as a β-hairpin, with the conserved cysteines located at each β-turn (Figure 4.4). While in the crystal structure of *S. aureus* MurT-GatD, these cysteine residues coordinate a Zn^{2+} ion, in the *S. pneumoniae* heterodimer they form disulfide bonds. Nevertheless, it was shown that the thiol group of the cysteines in *S. pneumoniae* MurT are reduced prior to crystallization, which implies that before crystals were formed, these residues were free to bind zinc ions and that the disulfide bonds likely formed during crystal growth [16]. The structural similarity of both insertions and conservation of the metal ligand cysteine residues support that *S. pneumoniae* MurT is also able to coordinate Zn^{2+} ions.

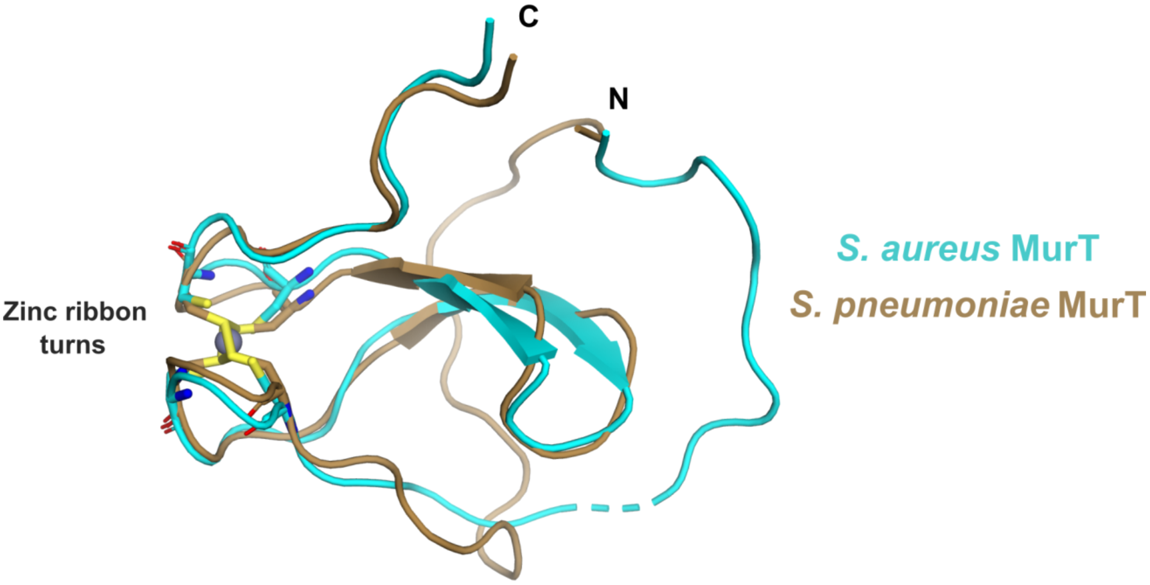


Figure 4.4. The zinc ribbon fold of MurT Cys-rich insertion.

Superposition of the Cys-rich insertion of *S. aureus* (residues 185-234 from PDB ID 6GS2/chain B in cyan) and *S. pneumoniae* (residues 187-238 from PDB ID 6GS2/chain A in brown) MurT proteins. Structures are represented as cartoon, the conserved cysteine residues are shown as atom-coloured sticks and the coordinated Zn^{2+} ion in the insertion of the *S. aureus* heterodimer is shown as a grey sphere. The directions for the N- and C-terminal of MurT proteins are labelled with the N and C letters, respectively. The dashed region of the cartoon represents residues that could not be modelled due poor electron density.

The Cys-rich insertion β-hairpin of MurT corresponds to the classic zinc ribbon fold [15,16], where the two pairs of metal-ligands are located in two turns forming structurally analogous zinc-binding sub-sites [34] (Figure 4.4). Zinc fingers with this fold are mostly present in multi-protein complexes, serving as interaction modules to bind nucleic acids, proteins and small molecules [34].

The quadruple mutation of the conserved cysteine in *S. pneumoniae* MurT-GatD by serine residues that prevent zinc coordination provided interesting insights into the function of the Cys-rich insertion [16]. This mutant was not tolerated *in vivo* [16], indicating that zinc coordination is essential for cell survival. *In vitro*, it did not abolish lipid II amidation nor impaired complex formation, and gave rise to a less stable MurT-GatD complex than the wild-type [16]. The *in vivo* intolerability might, therefore, be caused by the heterodimer's destabilization and, consequent, reduced abundance [16]. Although never reported before, the interaction of MurT with other proteins through this polypeptide segment is possible, since it is located at the surface of the protein, highly exposed to the solvent.

4.5.4. MurT-GatD interface

Lipid II amidation requires the synchronized activity of MurT and GatD proteins in the form of a heterodimer, placing protein-protein interface as a promising target for the development of antimicrobial drugs. As expected, the extended and compact conformations of *S. aureus* and *S. pneumoniae* MurT-GatD heterodimers, respectively, showed different intermolecular interactions (Figure 4.5). Although the corresponding interfaces were described individually by Nöldeke *et al.* [15] and Morlot *et al.* [16], an integrated structural comparison as the one herein allows identifying common molecular determinants. In brief, the overall MurT-GatD interface can be divided in two main patches: i) GatD-MurT C-terminal domain and ii) GatD-MurT central domain.

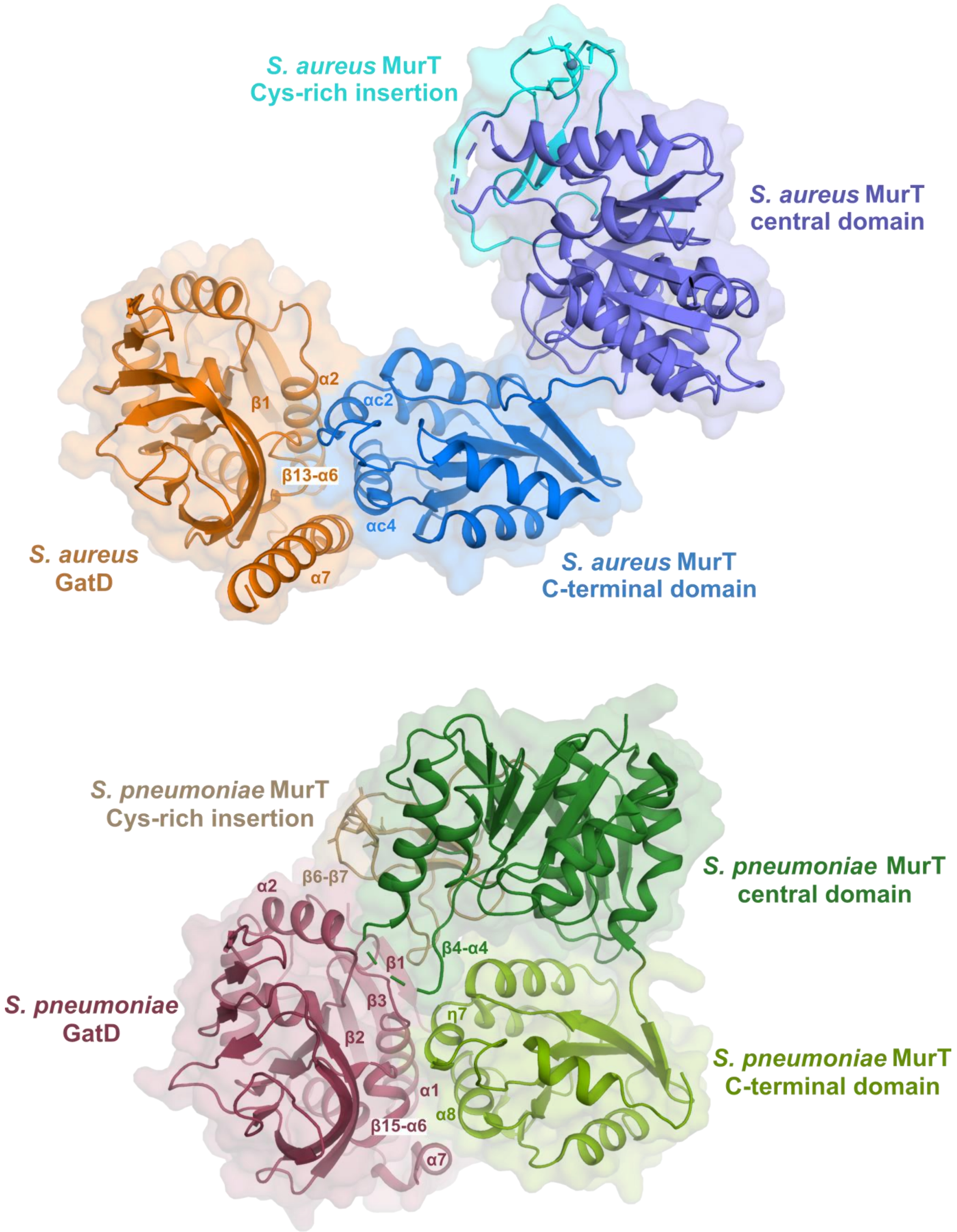


Figure 4.5. Interfaces of *S. aureus* and *S. pneumoniae* MurT-GatD complexes.

Cartoon and surface representation of MurT-GatD complex from *S. aureus* (PDB ID 6GS2) and *S. pneumoniae* (PDB ID 6FQB) with the secondary structure elements involved in the heterodimer interface labelled.

4.5.4.1. Interface GatD-MurT C-terminal domain

The interaction between GatD and the C-terminal domain of MurT is structurally organized in a conserved manner in both heterodimers (Figure 4.5), thus suggesting that its interface likely comprises the amino acids required for MurT-GatD complex assembly. This interaction covers the full interface of the *S. aureus* heterodimer, which further supports its importance.

The conserved architecture of the interface between GatD and the C-terminal domain contains a core formed by two MurT α-helices (αc2 and αc4 in *S. aureus* MurT; η7 and α8 in *S. pneumoniae* MurT) and a loop from GatD (loop β13-α6 in *S. aureus* GatD; loop β15-α6 in *S. pneumoniae* GatD), as shown in Figure 4.5. While in the *S. aureus* heterodimer this core interacts with the N-terminal α2 and the long C-terminal α7 of GatD, in the *S. pneumoniae* heterodimer it only interacts with the N-terminal α1 of GatD since its C-terminal region is flexible/disordered in the crystal structure. Hence, this interface patch is larger in *S. aureus* than in *S. pneumoniae* MurT-GatD, as revealed by the buried surface area that corresponds to 940 and 630 Å^2, respectively. Interestingly, the crystal structure of isolated *S. aureus* GatD protein shows the long C-terminal α7 helix perfectly ordered (Figure 4.3), suggesting that the secondary structure arrangement of this region is species-specific and independent of MurT interaction.

The core of the interface between GatD and the C-terminal domain of MurT comprises mainly hydrophobic contacts. The *S. aureus*/*S. pneumoniae* residues of MurT Y354/W360 and L385/V391, and GatD G21/G34, I24/L37, Q28/Y41 and P191/P208 are equivalent and occupy similar positions in the crystal structures of both heterodimers (Figure 4.6.A,B). Pull-down experiments and bacterial two-hybrid interaction assays showed that *S. aureus* MurT Y354A [35] and *S. pneumoniae* MurT W360A [16] completely impaired complex assembly with the respective GatD proteins, indicating their importance in complex assembly. The bulky side chains are likely crucial to establish the large network of intermolecular interactions observed in the crystal structures (Figure 4.6.A,B). *S. aureus* MurT L385A did not affect complex formation [35], however its essential role in protein-protein interaction cannot be excluded because the alanine mutant might establish equivalent hydrophobic interactions (Figure 4.6.A). These MurT and GatD residues are conserved among ortholog proteins (Fig. S4.4 and [7]) and likely form ubiquitous intermolecular interactions within MurT-GatD heterodimers.

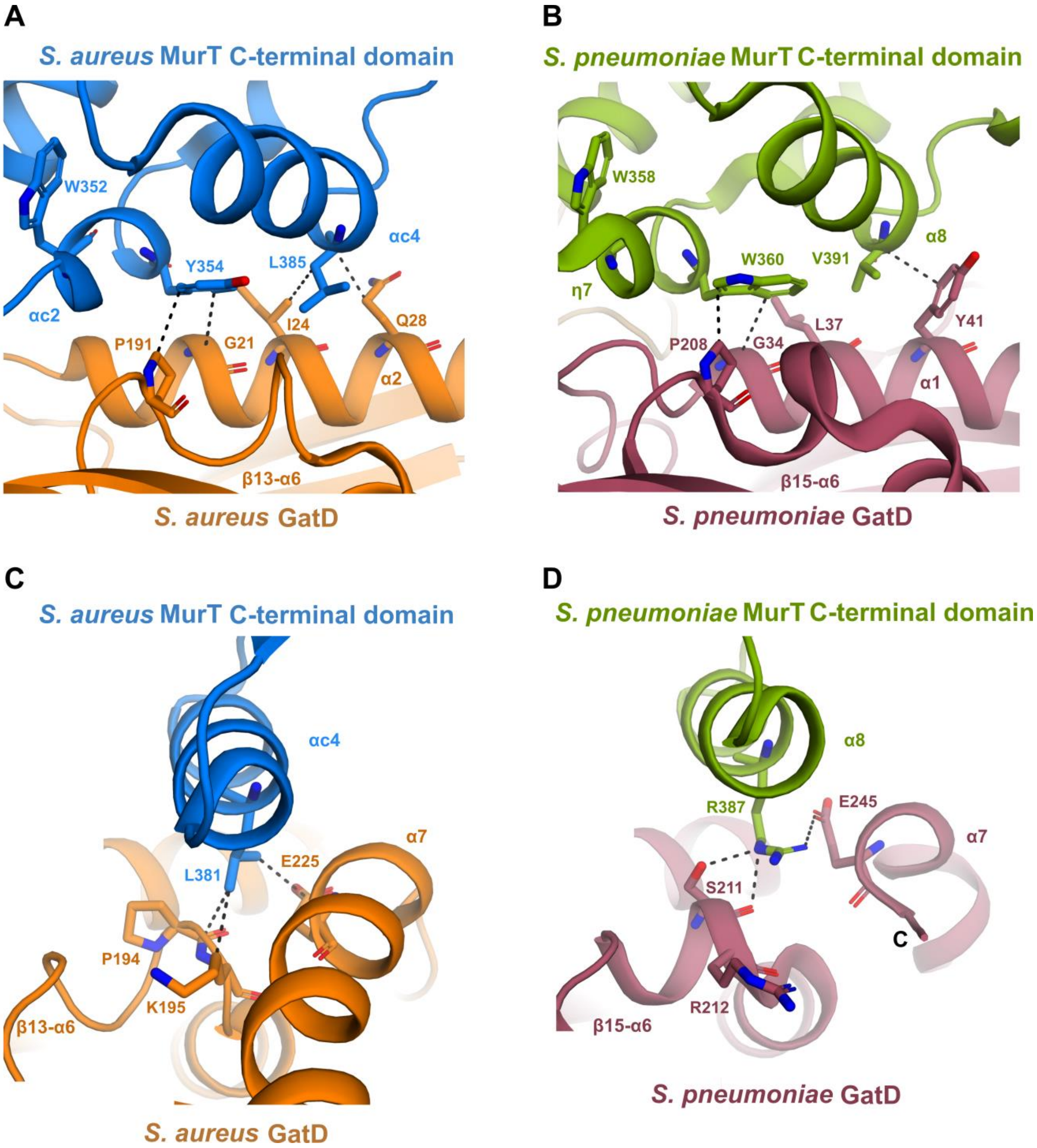


Figure 4.6. Core of the interface between MurT C-terminal domain and GatD.

Hydrophobic interactions between MurT C-terminal domain and GatD in the *S. aureus* (PDB ID 6GS2) and *S. pneumoniae* (PDB ID 6FQB) MurT-GatD complexes involving conserved **(A and B)** and non-conserved residues **(C and D).** The structures are represented as cartoon, residues involved in intermolecular interactions (dashed lines) as atom-coloured sticks and labelled, and the respective secondary structures are labelled.

S. aureus MurT W352 and L381 also proved to be important for the interaction with GatD [35]. In the crystal structure of the *S. aureus* heterodimer, MurT W352 does not interact directly with GatD. However, it is located at the centre of the helix αc2, which contains Y354

(Figure 4.6.A), and likely plays a structural role that ensures the correct orientation of the tyrosine residue at the interface. *S. aureus* MurT W352 is conserved in MurT orthologs (Figure S4.44) and the equivalent residue in *S. pneumoniae* MurT (W358) adopts a similar position at the interface with GatD (Figure 4.6.B), suggesting an ubiquitous structural role within MurT-GatD complexes. On the other hand, *S. aureus* MurT L381 performs several hydrophobic interactions with GatD (Figure 4.6.C) that are not observed in the *S. pneumoniae* heterodimer interface (Figure 4.6.C,D). This residue is not conserved in MurT proteins (Figure S4.44) and it is replaced in *S. pneumoniae* MurT by R387, which instead establishes intermolecular hydrogen-bonds and a salt bridge (Figure 4.6.C,D). This lack of conservation in essential intermolecular interactions indicates that the mechanisms promoting MurT-GatD heterodimer assembly vary among bacteria.

The hydrophobic core of the *S. aureus* MurT-GatD interface is complemented by a network of polar contacts close to MurT αc4 and GatD α7, such as the ones involving the three pairs [MurT D342 - GatD R235], [MurT E378 - GatD K195] and [MurT K384 – GatD D221 and E225] (Figure 4.7A). Interestingly, MurT D342 and E378 form salt bridges with MurT R375, which have been shown to be important for complex assembly [35]. However, the substitution of this arginine by alanine only resulted in a pull-down ratio of MurT with GatD lower than 1:1 and not in the total abolishment of heterodimer formation, indicating that its presence is not crucial for complex assembly. SPR experiments also showed that *S. aureus* MurT K384 has a non-essential role in the interaction with GatD (Chapter 3). Although the corresponding pairs (Fig. S4.4 and [7]) *S. pneumoniae* [MurT E384 - GatD R212] and [MurT R390 – GatD D240 and E245] are structurally conserved, no interactions are observed (Figure 4.7B). Taken together, the complementary role of these highly conserved residues [7] might contribute for the stabilization of the extended boomerang-like conformation of the MurT-GatD complex.

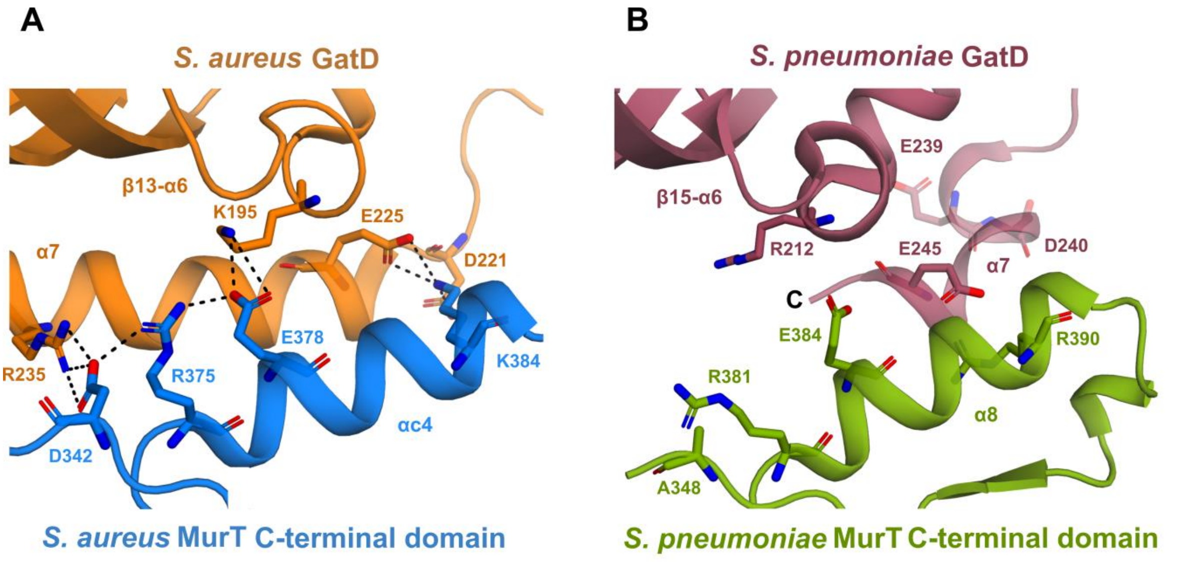


Figure 4.7. Polar interactions in the interface between MurT C-terminal domain and GatD.

(**A**) Polar interactions at the interface of *S. aureus* MurT-GatD (PDB ID 6GS2) involving residues from αc4 of MurT C-terminal domain and GatD α7. (**B**) The equivalent region in *S. pneumoniae* MurT-GatD (PDB ID 6FQB) contains similar residues but no interactions. The structures are represented as cartoon, residues involved in intermolecular interactions (dashed lines) as atom-coloured sticks and labelled, and the respective secondary structures also labelled.

4.5.4.2. Interface GatD-MurT central domain

The crystallographic model of *S. pneumoniae* MurT-GatD shows the MurT central domain docked on the N-terminal region of GatD, with the loops β4-α4 and β6-β7 of MurT central domain mediating an extensive network of contacts with β1, loop β2-α1, β3, loop β3-β4 and α2 of GatD (Figure 4.5).

The contribution of MurT loop β4-α4 for the heterodimer interface involves a hydrophobic [MurT Q137 - GatD Y30] and a hydrogen-bond [MurT G142 - GatD Q75] (Figure 4.8A). In ortholog proteins, these residues are replaced by amino acids with analogous characteristics (Figure S4.44 and [7]), supporting that similar contacts might be formed in other MurT-GatD complexes. Interestingly, the full MurT loop β4-α4 (T131-E143) is very well conserved among MurT proteins (Figure S4.44) and pull-down experiments showed that alanine mutagenesis of D136 and D139 had a moderate effect on the interaction with GatD [16], confirming its involvement in complex formation. The stretch M138-F141 is not observed in the electron density maps probably due to its mobile nature (Figure 4.8A) and might become ordered upon conformation rearrangement during catalysis.

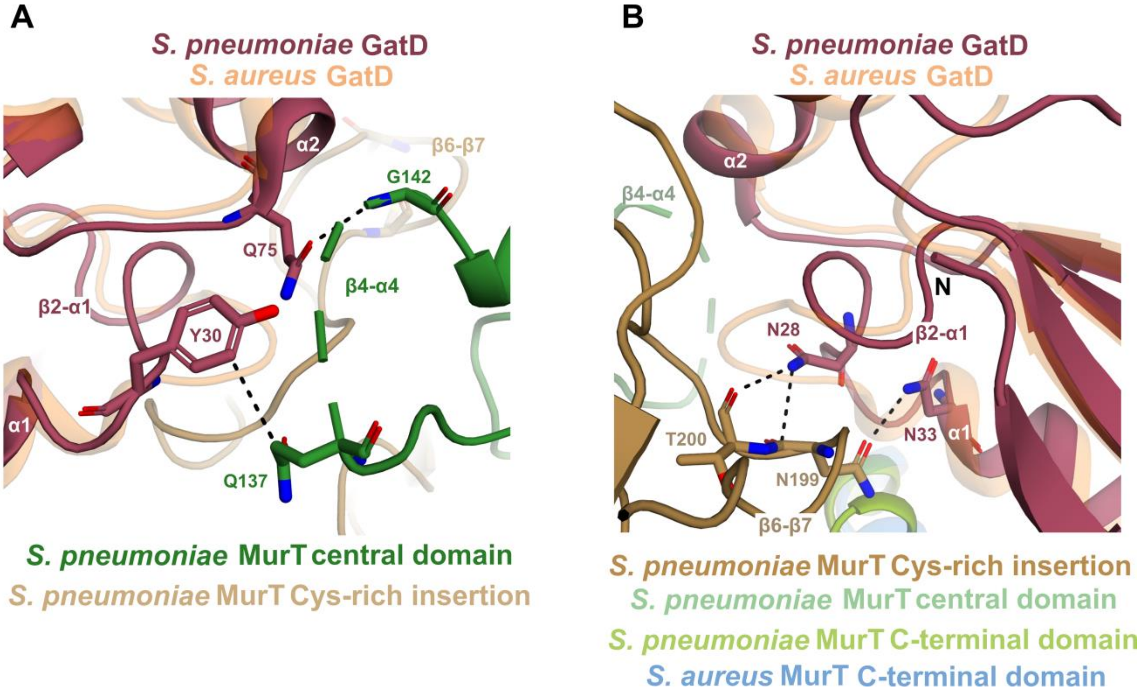


Figure 4.8. Interface of MurT central domain and GatD in the *S. pneumoniae* MurT-GatD heterodimer.

(A) Intermolecular interactions between *S. pneumoniae* GatD (PDB ID 6FQB/chain B in red) and loop β4-α4 of MurT (PDB ID 6FQB/chain A in dark green). **(B)** Intermolecular interactions between *S. pneumoniae* GatD (PDB ID 6FQB/chain B in red) and loop β6-β7 of the MurT Cys-rich insertion (PDB ID 6FQB/chain A in brown). In both panels, the structures are shown as cartoon and the residues involved in intermolecular interactions (dashed lines) as atom-coloured sticks and labelled, and the respective secondary structures also labelled. The GatD from *S. aureus* MurT-GatD (PDB ID 6GS2/chain A in orange) is superposed with *S. pneumoniae* GatD and shown as transparent cartoon. The dashed region of the loop β4-α4 represents the residues M138-F141 that could not be modelled due poor electron density.

The loop β6-β7 of the *S. pneumoniae* MurT Cys-rich insertion establishes an extensive network of interactions with GatD (Figure 4.8B). It directly faces the loop β2-α1 of GatD, promoting the hydrogen-bonds [MurT N199 and T200 – GatD N28] and [MurT N199 – GatD N33]. These residues are conserved among orthologs proteins (Figure S4.44 and [7]) and the respective interactions might be ubiquitous among MurT-GatD complexes. Interestingly, in the crystal structure of *S. aureus* MurT-GatD, the equivalent GatD loop is flexible, suggesting it might be stabilized by similar intermolecular contacts upon conformational rearrangements. The contribution of *S. pneumoniae* MurT loop β6-β7 in the interface with GatD suggests the potential role of the cysteine-rich insertion in stabilizing the compact conformation of MurT-GatD complexes. Nevertheless, the involvement of *S. pneumoniae*

GatD β1, which is absent in other GatD orthologs, in the interface with MurT suggest that complex assembly relies also on species-specific mechanisms.

4.5.5. Glutamine active site

The crystal structure of *S. pneumoniae* MurT-GatD showed an electron density fitting a glutamine molecule at the glutaminase active of the complex [16], which shares a similar architecture as the glutaminase active site of triad GATases such as the IGPS from *T. maritima* [36] (Figure 4.9.A). Most of the molecular determinants are preserved in GatD orthologs [7], suggesting they possess ubiquitous roles in the glutaminase activity of the MurT-GatD complexes. Interestingly, these residues occupy similar positions in the crystal structures of isolated *S. aureus* GatD and glutamine-free *S. aureus* MurT-GatD (Figure 4.9.B), which indicates that glutamine binding does not induce conformational rearrangements on the complex.

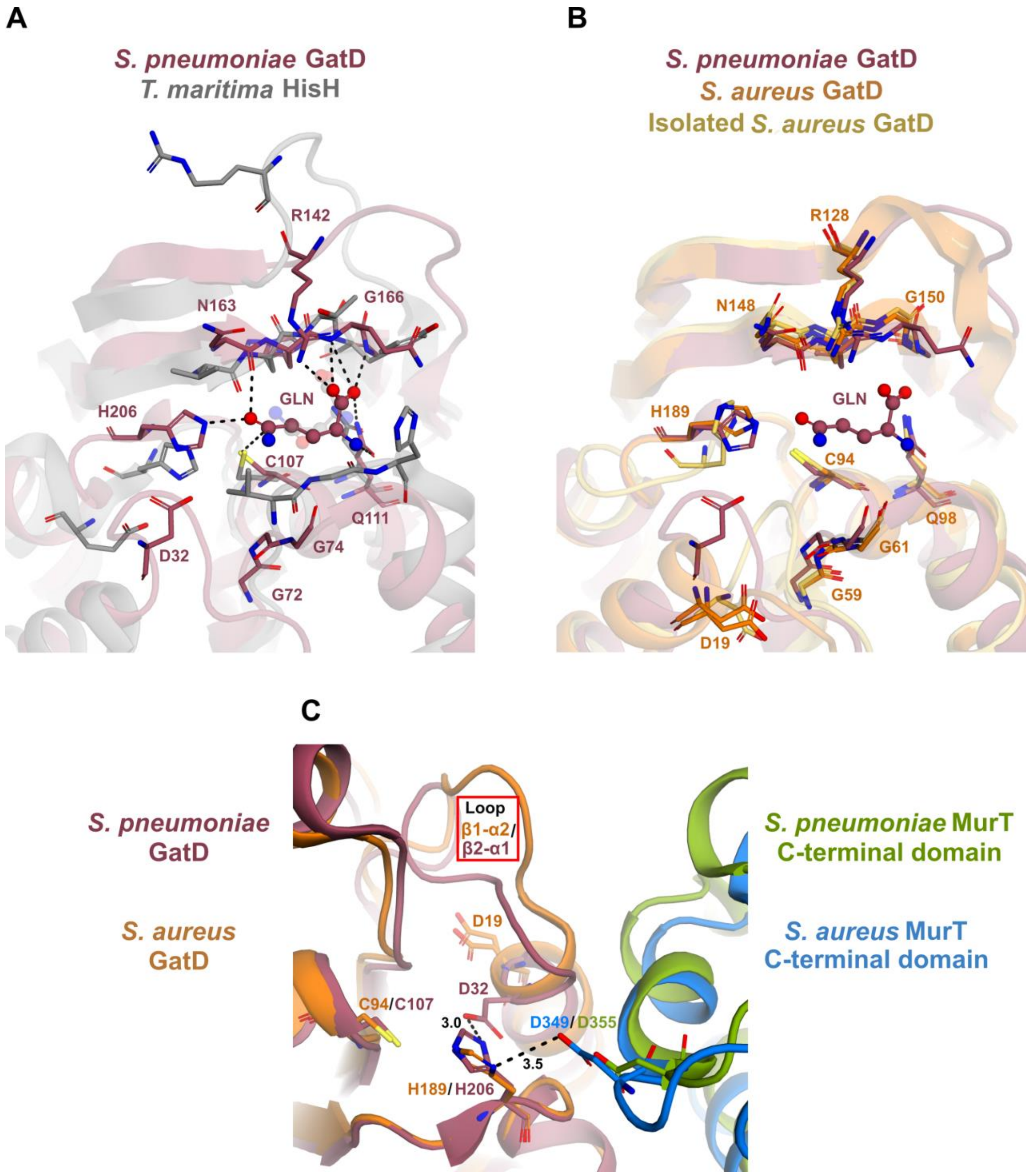


Figure 4.9. Glutaminase active site of Murt-GatD complex.

(A) Superposition of GatD from *S. pneumoniae* MurT-GatD (PDB ID 6FQB/chain B in red) and HisH protein from *T. maritima* IGPS (PDB ID 3ZR4/chain B in grey). The glutamine molecule modelled in *S. pneumoniae* GatD is labelled and represented as atom-coloured ball and sticks, while the glutamine molecule in HisH is shown as transparent atom-coloured ball and sticks. The *S. pneumoniae* GatD residues constituting the glutamine binding site are represented as atom-coloured sticks and labelled, with the relevant interactions to the substrate shown as dashed lines. **(B)** Superposition of GatD from *S. pneumoniae* MurT-GatD (PDB ID 6FQB/chain B in red) with isolated *S. aureus* GatD (PDB ID 5N9M/chain A in yellow) and GatD from *S. aureus* MurT-GatD (PDB ID 6GS2/chain A in orange). The *S. aureus* GatD residues constituting the glutamine binding site are labelled. **(C)** Superposition of GatD from *S. pneumoniae* MurT-GatD (PDB ID 6FQB/chain B in red) and GatD from *S. aureus* MurT-GatD (PDB ID 6GS2/chain A in orange). The residues constituting the potential glutaminase catalytic triads are labelled. The intramolecular and intermolecular hydrogen-bond established by the catalytic His in the *S. pneumoniae* and *S. aureus* MurT-GatD heterodimers, respectively, are shown as dashed lines. The C-terminal domain of MurT from *S. pneumoniae* (PDB ID 6FQB/chain A) and *S. aureus* (PDB ID 6GS2/chain B) are coloured light green and blue, respectively.

The glutamine molecule modelled in the structure of *S. pneumoniae* MurT-GatD is at the vicinity of GatD C107, with its thiol group at 3.3 Å from the reactive C_δ of glutamine (Figure 4.9.A). The relative position of GatD residues C107 and H206 (Figure 4.9.A), as well as their orientation towards glutamine match the characteristic pre-catalytic conformation of the glutaminase active site of triad GATases [36]. The catalytic role of these conserved amino acids was previously shown in *S. aureus* MurT-GatD, in which the mutations of GatD C94 and H189 abolished both the glutaminase and the amidotransferase activities of the complex [2,7,15]. GatD H189 and the respective loop is in a different orientation in the structure of isolated *S. aureus* GatD when compared to MurT-GatD (Figure 4.9.B), suggesting that the interaction of GatD with the C-terminal domain of MurT promotes the rearrangement of this region. This might represent a regulation mechanism of catalysis in order to avoid the production and release of ammonia in the bacterial cell when MurT is absent.

The structure of the *S. pneumoniae* heterodimer shows the characteristic catalytic triad of GATases (Figure 4.9.A), completed by the residue GatD D32 that is 3 Å apart from the catalytic H206 (Figure 4.9.C). However, the role of the acidic residue in the catalytic triad of GATases is not consensual since its substitution by non-acidic amino acids had a negligible effect in the glutaminase activity of triad GATases such as *E. coli* carbamoyl phosphate synthase [37] or *T. maritima* IGPS [36]. Therefore, even though *S. pneumoniae* GatD D32 is conserved among orthologs, its functional role needs to be further investigated. The equivalent GatD D19 in the crystal structure of *S. aureus* MurT-GatD is oriented away from the active site (Figure 4.9.C), which might be related to structural differences such as the lack of glutamine at the active site and/or its extended conformation.

The crystal structure of *S. aureus* MurT-GatD shows the catalytic GatD H189 hydrogen-bonded to MurT D349 (Figure 4.9.C), suggesting an intermolecular triad. This residue is conserved in MurT orthologs (Figure S4.44) and the equivalent MurT D355 occupies a similar position in the structure of *S. pneumoniae* MurT-GatD, slightly displaced away from the catalytic GatD H206 (Figure 4.9.C). Such intermolecular arrangement is also observed in the structure of triad GATase *T. maritima* IGPS complex, for which the acidic residue was shown to be essential for the glutaminase activity of the complex [36]. Likewise, replacement of *S. aureus* MurT D349 by an asparagine completely supressed the *in vitro* amidation of lipid II [15]. On the other hand, substitution of *S. pneumoniae* MurT D355 by an alanine residue reduced but did not abolish the *in vitro* amidation of lipid II [16]. Moreover, this MurT mutant was lethal for *S. pneumoniae in vivo* [16]. These different results suggest that either alanine and asparagine residues affect differently the *in vitro* enzymatic activity or that the aspartate residue has a more important role in *S. aureus* MurT-GatD.

The glutaminase activity can be explored based on the well-described enzymatic reaction catalysed by GATases [31,38]. The side chain amide of glutamine is expected to bind at the active site with its amino group facing the catalytic histidine in a position prone for hydrolysis, and its carbonyl group oriented towards the oxyanion hole residues. Of note, these groups were modelled in opposite positions in the glutamine molecule present in the structure of *S. pneumoniae* MurT-GatD (Figure 4.9.A). The residues comprising the oxyanion hole of GATases [38] are conserved in GatD proteins [7] and correspond to *S. pneumoniae* GatD 72-GGG-74 and *S. aureus* GatD 59-GGG-61. However, in both MurT-GatD crystal structures these GatD residues are displaced from the active site (Figure 4.9.A,B) and cannot form the hydrogen-bonds with the carbonyl group of glutamine side chain amide that are required to stabilize the intermediate species of the glutaminase reaction. Therefore, structural rearrangements are required for the enzyme to adopt an active conformation.

The crystal structure of *S. pneumoniae* MurT-GatD also shows stabilization of the glutamine α-groups similarly to triad GATases (Figure 4.9.A). Residue Q111, which is conserved in GATases [31], interacts with both α-groups of glutamine (Figure 4.9.A). The α-carboxylic group of glutamine is additionally hydrogen-bonded to the backbone α-amino groups of Q165 and G166, and with the side chain of R142 (Figure 4.9.A). Interestingly, the crystal structure of isolated *S. aureus* GatD showed the equivalent arginine, R128, interacting with the side chain carbonyl group of glutamine at the surface of the protein [7], suggesting that the substrate is rearranged at the active site.

In the crystal structure of isolated *S. aureus* GatD [7], the glutamine molecule found at the surface of the protein seems to have the access to the active blocked by Y17 (Figure 4.10.).

While in the structure of *S. aureus* MurT-GatD this residue is flexible and modelled with alternate conformations, the structure of *S. pneumoniae* MurT-GatD shows the equivalent GatD Y30 interacting with the β4-α4 loop of MurT central domain (Figure 4.8A andFigure **4.10.**). A potential reorganization of the glutaminase active site architecture upon the interaction between GatD and MurT central domain in a compact conformation could allow the entrance of glutamine into the active site. Interestingly, the tyrosine residue of IGPS glutaminase domain, that was found to function as a plug of the ammonia channel, occupies a similar position in *S. pneumoniae* GatD Y30 (Figure 4.10.). Therefore, its location between the glutaminase active site and MurT (Figure 4.10.) suggests that this residue might also gate the ammonia channel towards the amidotransferase active site. Indeed, replacing *S. pneumoniae* GatD Y30 by alanine impaired the *in vitro* amidation of lipid II, confirming its functional role.

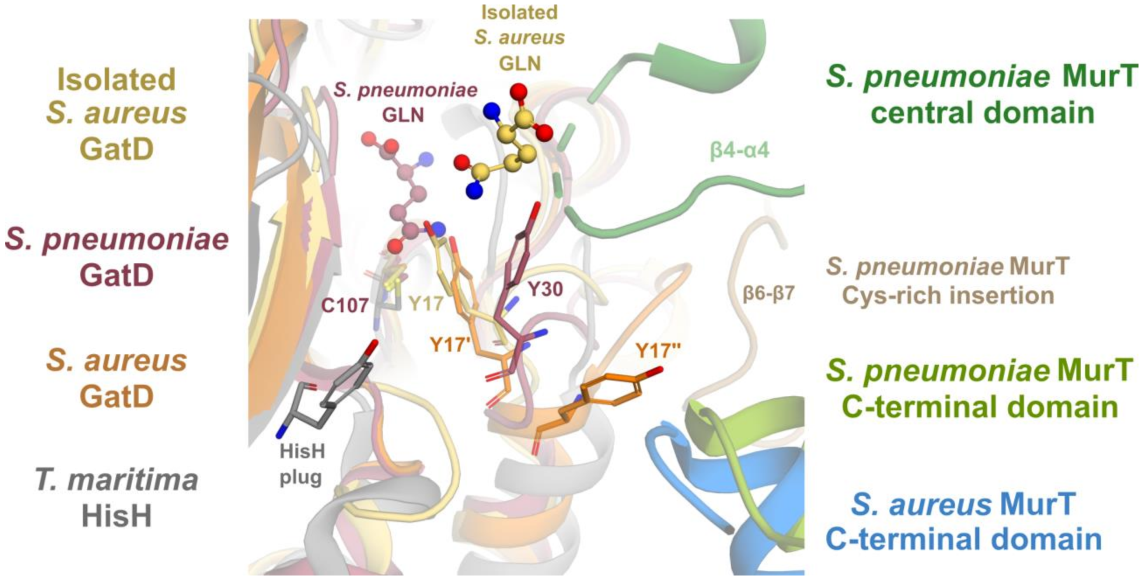


Figure 4.10. Plug of MurT-GatD glutaminase active site.

Superposition of GatD from *S. pneumoniae* MurT-GatD (PDB ID 6FQB/chain B in red) with isolated *S. aureus* GatD (PDB ID 5N9M/chain A in yellow), GatD from *S. aureus* MurT-GatD (PDB ID 6GS2/chain A in orange) and HisH protein (glutaminase) from *T. maritima* IGPS (PDB ID 3ZR4/chain B in grey). The C-terminal domain of *S. pneumoniae* MurT (PDB ID 6FQB/chain A) and *S. aureus* MurT (PDB ID 6GS2/chain B) are coloured light green and blue, respectively. The labelled loops β4-α4 and β6-β7 from MurT central domain (PDB ID 6FQB/chain A) are coloured dark green and brown, respectively. The conserved tyrosine of *S. aureus* and *S. pneumoniae*, as well as the residue serving as plug of the ammonia channel in *T. maritima* IGPS GATase are shown as atom-coloured sticks and labelled. The catalytic cysteine of *S. pneumoniae* GatD is shown as a reference of the glutaminase active site. The synthase protein of IGPS was omitted for clarity.

4.5.6. ATP binding site

The crystal structure of *S. aureus* MurT-GatD bound to AMP-PNP [15] revealed that the nucleotide binding site is located at the central domain of MurT protein with the residues 56-TNGKT-60 forming the characteristic P-loop of ATPases (Figure 4.11A). The triphosphate moiety of the AMP-PNP is hydrogen-bonded to the P-loop residues through their α-amine groups and the side chain of K59; additionally, its γ-phosphate is interacting with the side chain of N85 (Figure 4.11B). The γ- and β-phosphates of AMP-PNP are stabilized by a Mg^{2+} ion, as observed in Mur ligases [6], which is coordinated by the side chains of E108 and T60 residues. While the adenine moiety of AMP-PNP is inserted in a narrow pocket (Figure 4.11A) formed by several aromatic residues (F132, Y216, F263 and Y266) and three asparagine residues (N57, N130 and N267), the ribose moiety is mainly interacting with the side chain of N57 (Figure 4.11A). Interestingly, MurT Y216 belongs to the Cys-rich domain of the protein that is absent in Mur ligases (Figure S4.44).

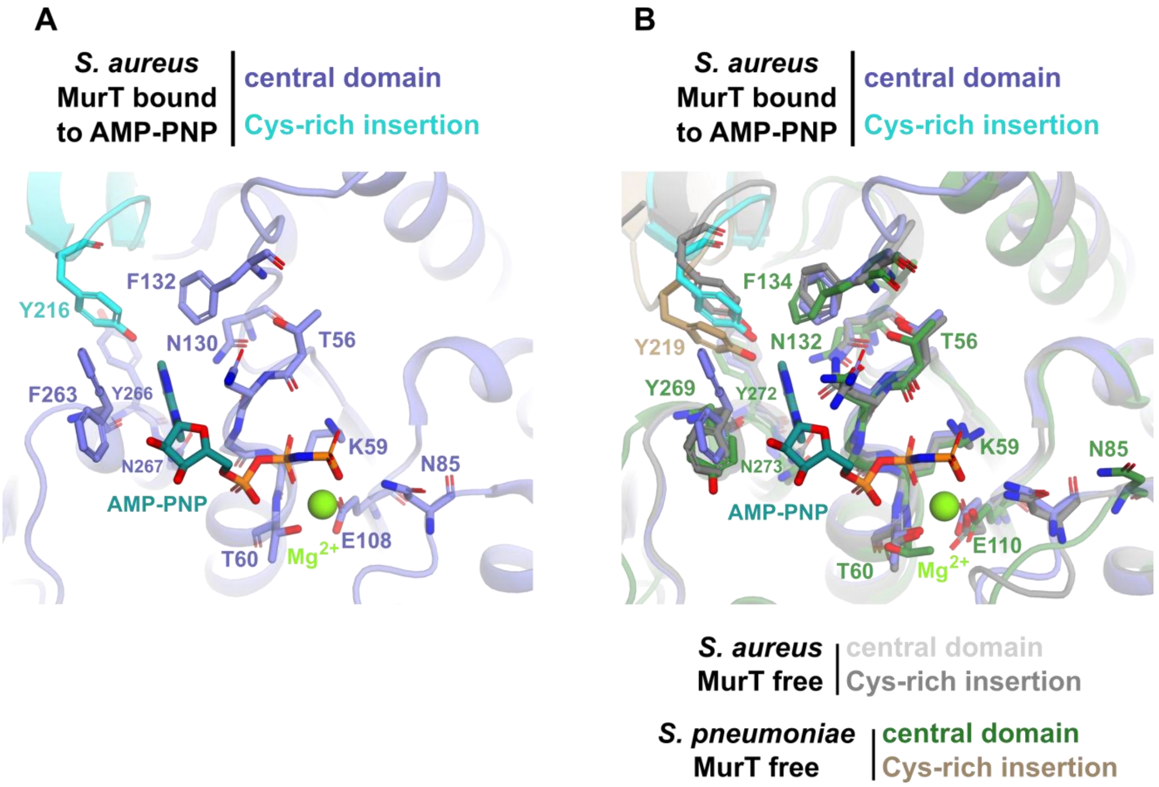


Figure 4.11. The binding site of AMP-PNP in *S. aureus* MurT-GatD complex.

(A) The residues of *S. aureus* MurT (PDB ID 6H5E/chains B) within 4 Å of the AMP-PNP ligand are labelled and presented as atom-coloured sticks. The central domain and Cys-rich insertion of MurT are represented as cartoon and coloured purple and cyan, respectively. **(B)** Superposition of the MurT central domain (including Cys-rich insertion) from *S. aureus* MurT-GatD structure complexed with AMP-PNP (PDB ID 6H5E/chains B) with *S. aureus* MurT-GatD structure with no ligands (PDB ID 6GS2/chains B, in grey; Cys-rich insertion in grey) and *S. pneumoniae* MurT-GatD structure (PDB ID 6FQB/chain A in dark green; Cys-rich insertion in brown). The structures are represented as cartoon. The residues constituting the AMP-PNP binding site are shown as atom-coloured sticks and the ones from *S. pneumoniae* MurT are labelled.

The residues interacting with AMP-PNP are structurally conserved in the crystal structure of *S. aureus* MurT-GatD free of ligands (Figure 4.11B), suggesting that ATP binding does not induce conformational rearrangements in GatD nor in MurT proteins. Also in the crystal structure of *S. pneumoniae* MurT-GatD, this region displays a similar architecture, with the exception of N85 that is 5.5 Å displaced from the active site (Figure 4.11B). The physiological importance of these residues was confirmed by site-directed mutagenesis: *S. aureus* MurT T60A, E108A or N267Y and *S. pneumoniae* MurT K59A mutants abolished lipid II amidation [15,16]. The ATP binding site is expected to be ubiquitous in MurT-GatD complexes since these residues are conserved in MurT orthologs (Figure S4.44).

4.5.7. Putative lipid II binding site

The crystal structures of MurT-GatD currently available do not possess lipid II or any related analogue interacting with the polypeptide chain and, therefore, the molecular determinants involved in the amidotransferase activity of the complex remain unknown. Lipid II amidation reaction is dependent on ATP and the binding site architecture of this substrate in MurT is similar to Mur ligases, suggesting that lipid II is activated by phosphorylation before the amidation reaction. This implies that the D-Glu carboxyl group of lipid II is likely located close to the ATP γ-phosphate, with a Mg^{2+} ion neutralizing both negatively-charged groups (Figure 4.13).

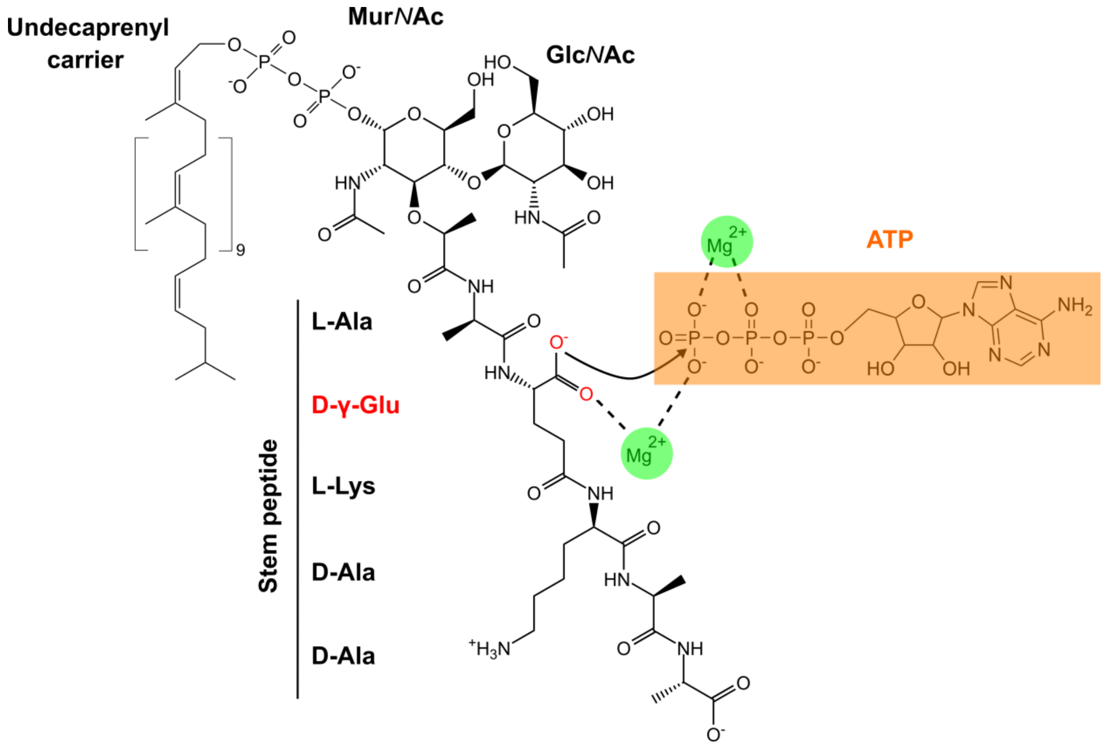


Figure 4.12. Expected activation of lipid II through phosphorylation.

Based on the reaction mechanism catalysed by Mur ligases and their similarities with MurT, lipid II is expected to be phosphorylated before the transfer of ammonia. In this activation mechanism, a Mg^{2+} ion is likely to stabilize the negatively-charged groups of lipid II D-Glu carboxyl and ATP γ-phosphate.

In Mur ligases, the Mg^{2+} ion is coordinated by the negatively-charged groups of ATP γ-phosphate, the acceptor carboxyl of the UDP-Mur*N*Ac PG precursor, a conserved histidine residue and three water molecules [10] (Figure 4.13A). These waters are hydrogen-bonded to both the central and the C-terminal domains of Mur ligases (MurC uses negatively-charged residues (Figure 4.13A), while MurD/E/F combine a negatively-charged residue with a carbamoylated lysine) [6]. Comparison with the crystal structure of MurC (PDB ID 1P3D) in complex with AMP-PNP, its UDP-Mur*N*Ac product and Mn^{2+} (instead of the native Mg^{2+}) ions suggests that *S. aureus* MurT Q135 is responsible for the direct coordination of the Mg^{2+} ion, while D110, E111, D134 and K315 participate in the binding site through water-mediated hydrogen-bonds (Figure 4.13). Several lines of evidence support the functional relevance of these MurT residues: they are all conserved among orthologs (Figure S4.44); glutamine residues can indeed act as O-donor in the coordination of Mg^{2+} ions [39]; D134 and Q135 are located in a flexible loop that might be stabilized after ligand binding (lipid II and Mg^{2+}); the architecture of the potential binding site is conserved in the structure of *S. pneumoniae* MurT-GatD, including the flexible loop (β4-α4) harbouring the conserved glutamine (Figure 4.8A); in *S. pneumoniae* MurT-GatD complex, the single-mutant MurT D136A (equivalent to *S. aureus* MurT D134) and the double-mutant D112S/D113S (equivalent to *S. aureus* MurT D110 and E111) eliminated lipid II amidation *in vitro* [16]. Altogether, these results suggest that MurT-GatD coordinates a Mg^{2+} ion at the amidotransferase active site in a similar way that Mur ligases. Moreover, the conservation of residues between the C-terminal domains of MurT and Mur ligases supports that the amidotransferase active site of MurT likely lies between the two MurT domains, which implies a rigid body rotation towards each other upon lipid II binding.

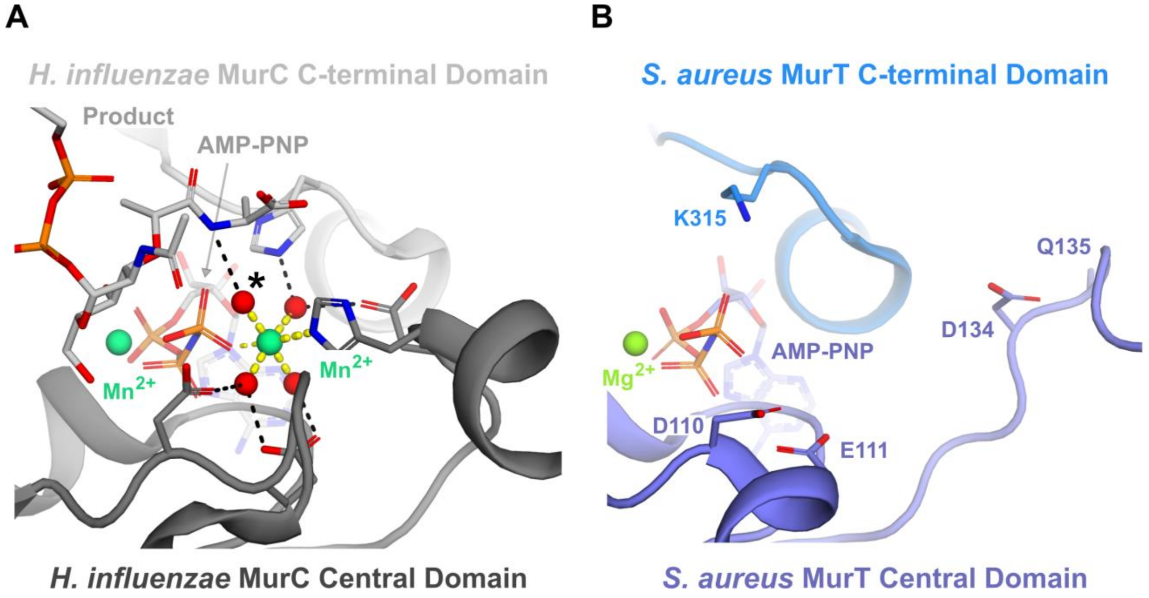


Figure 4.13. Putative Mg^{2+} binding site involved in lipid II phosphorylation.

Comparison between the (**A**) MurC from *H. influenzae* in complex with AMP-PNP, the UDP-Mur*N*Ac product UDP-Mur*N*Ac-L-Ala and Mn^{2+} (instead of the native Mg^{2+}) ions (PDB ID 1P3D/ chain A; cartoon of the central and the C-terminal domains are coloured dark and light grey, respectively; AMP-PNP and the product UDP-Mur*N*Ac-L-Ala are shown as atom-coloured sticks; the two Mn^{2+} ions and four water molecules are represented as aquamarine and red spheres, respectively) and (**B**) the MurT central and the C-terminal domains from *S. aureus* MurT-GatD heterodimer complexed with AMP-PNP (PDB ID 6H5E/chain B; cartoon of the central and the C-terminal domains are coloured purple and blue, respectively; AMP-PNP is shown as atom-coloured sticks; and the Mg^{2+} ion as green sphere). The central and C-terminal domains of MurT were superposed individually with the respective domains of MurC to provide the conformation of a bound form. The MurC residues involved in the binding site and the corresponding *S. aureus* MurT amino acids are shown as atom-coloured sticks, with the latter being labelled. The coordination of the Mn^{2+} ion between AMP-PNP and UDP-Mur*N*Ac-L-Ala is shown as yellow dashed lines, while the hydrogen-bonds stabilizing the conserved water molecules are shown as black dashed lines. The water replacing the acceptor carboxyl group of the UDP-Mur*N*Ac precursor substrate before the condensation reaction is marked with an asterisk (*).

The binding site of the UDP-Mur*N*Ac precursor substrate in Mur ligases is less informative in the quest to investigate lipid II binding to MurT due to the differences between the two substrates. Mur ligases recognize the UDP moiety of the UDP-Mur*N*Ac precursor substrate through their N-terminal domain [6]. Lipid II, however, lacks the uridine unit and contains the pyrophosphate group directly bound to the undecaprenyl carrier (Fig. 4.12). Due to its hydrophobic nature, the undecaprenyl carrier is likely inserted in the cytoplasmic membrane and not available for binding. No transmembrane domains are predicted in the N-terminal region of MurT, excluding its direct interaction with the undecaprenyl moiety of lipid II. Therefore, only the pyrophosphate group of lipid II might be recognized by the N-terminal region of MurT, explaining its short size in comparison to the N-terminal domain of Mur ligases (Figure S4.44). Identifying common molecular determinants in these regions is difficult since they share low sequence identity and the N-terminal region of MurT is mobile in the available crystal structures of MurT-GatD complexes. Nevertheless, the high content of positively-charged amino acids in this region of MurT has been proposed to bind the pyrophosphate group of lipid II [15]. Regarding the Mur*N*Ac moiety of the UDP-Mur*N*Ac precursor substrates, Mur ligases do not establish conserved interactions [6] that can be mapped on MurT. Lipid II contains an additional GlcNAc group, which is absent in the substrates of Mur ligases. Therefore, insights on how these sugar moieties bind to MurT require structural data in the presence of substrates or analogues.

The substrates of Mur ligases have their stem peptide moiety located between the central and C-terminal domains [6]. Along the PG biosynthesis pathway, the substrates of consecutives Mur ligases contain a growing version of the stem peptide, resulting in slightly different binding sites. MurD and MurF proteins possess a conserved asparagine residue responsible for hydrogen-bonding to the terminal carboxyl group of the UDP-Mur*N*Ac precursor substrate, which corresponds to MurT N85 in the crystal structure of *S. aureus* MurT-GatD [14] (Figure 4.14A). Indeed, this residue is conserved among orthologs (Figure S4.44) and mutation of the equivalent *S. pneumoniae* MurT N85 by alanine abolished the *in vitro* amidation of lipid II, suggesting its interaction with the carboxyl group of lipid II D-Glu. *S. aureus* MurT R116, also conserved among orthologs (Figure S4.44), might be involved in the lipid II binding site since it superposes with the arginine residue that establishes an hydrogen-bond with the L-Ala carbonyl group of the UDP-Mur*N*Ac precursor substrate in MurE [40](Figure 4.14B).

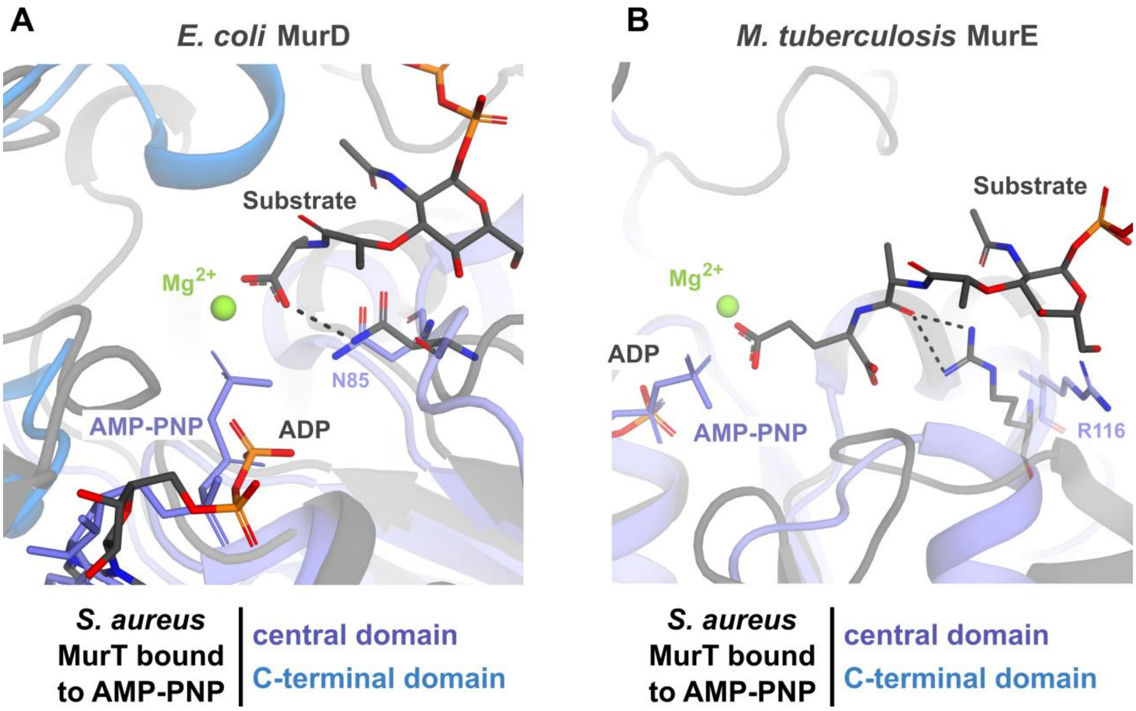


Figure 4.14. Potential role of MurT central domain in lipid II binding site.

Superposition of the MurT central and the C-terminal domains of *S. aureus* MurT-GatD heterodimer complexed with AMP-PNP (PDB ID 6H5E/chain B; cartoons coloured purple and blue, respectively) and (**A**) MurD from *E. coli* complexed with ADP and the substrate UDP-Mur*N*Ac-L-Ala (PDB ID 2UAG/ chain A; cartoon coloured grey) or (**B**) MurE from *M. tuberculosis* complexed with ADP and the substrate UDP-Mur*N*Ac-L-Ala-D-Glu (PDB ID 2XJA/ chain A; cartoon coloured grey). The N-terminal domains of Mur ligases were omitted for clarity. The AMP-PNP ligand from the structure of *S. aureus* MurT-GatD is shown as purple sticks, while the substrates and Mg^{2+} ion in each Mur ligase structure are atom-coloured and labelled. The relevant residues are shown as atom-coloured sticks, with the ones from *S. aureus* MurT being labelled.

The C-terminal domain of Mur ligases is not directly involved in the binding site of the UDP-Mur*N*Ac precursor substrate, although it establishes important interactions with the condensed amino acid of the final product [6]. Structural alignment of *S. aureus* MurT C-terminal domain with that of Mur ligases revealed that some features are conserved. *S. aureus* MurT L313 superposes with a conserved tyrosine of MurC [10] and MurE [41] that establish hydrophobic interactions with the product (Figure 4.15.A,B). This leucine residue is conserved among orthologs (Figure S4.44) and might perform hydrophobic contacts with the stem peptide of lipid II (Figure 4.15.A). The conserved *S. aureus* MurT D346 (Figure S4.44) is located close to the aspartate residue of MurE that performs electrostatic interactions with the condensing L-Lys of the product (Figure 4.15.B). Interestingly, a similar interaction between D346 and the stem peptide L-Lys of lipid II would imply that MurT-GatD amidates lipid II before the pentaglycine moiety is attached to the stem peptide.

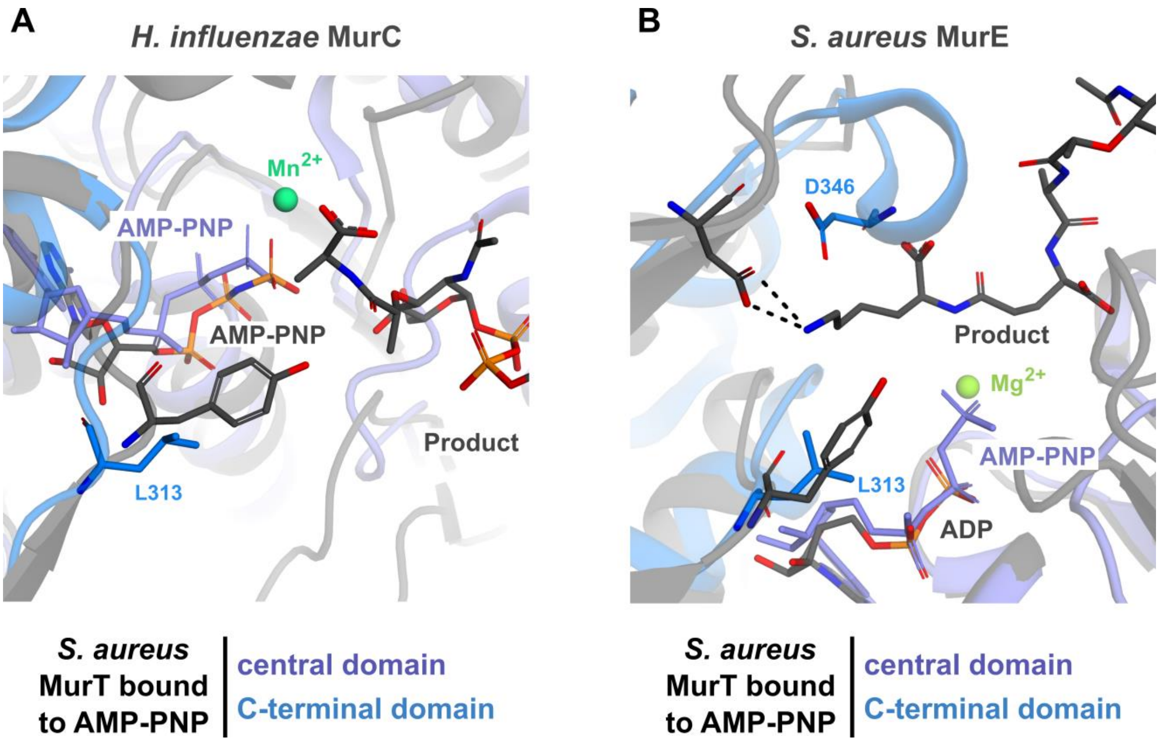


Figure 4.15. Potential role of MurT C-terminal domain in lipid II binding site.

Superposition of the MurT central and the C-terminal domains from *S. aureus* MurT-GatD heterodimer complexed with AMP-PNP (PDB ID 6H5E/chain B; cartoons coloured purple and blue, respectively) and (**A**) MurC from *H. influenzae* complexed with AMP-PNP and the product UDP-Mur*N*Ac-L-Ala (PDB ID 1P3D/ chain A; cartoon coloured grey) or (**B**) MurE from *S. aureus* complexed with ADP and the product UDP-Mur*N*Ac-L-Ala-D-Glu (PDB ID 4C12/ chain A; cartoon coloured grey). The N-terminal domains of Mur ligases were omitted for the clarity of the figure. The AMP-PNP ligand from the structure of *S. aureus* MurT-GatD is shown as purple sticks, while the AMP-PNP/ADP, the product and Mg^{2+}/Mn^{2+} ion in each Mur ligase structure are atom-coloured and labelled. The relevant residues are shown as atom-coloured sticks, with the ones from *S. aureus* MurT being labelled.

Although both MurT-GatD crystal structures represent non-competent catalytic forms of the complex, the observations here gathered allow proposing a model for the reaction mechanism of lipid II amidation. The active conformation is only completed upon the binding of all substrates. The expected interactions between MurT central domain and lipid II likely stabilize the loop (containing the conserved *S. aureus* MurT Q135 and *S. pneumoniae* MurT Q137) located between the glutaminase and amidotransferase active sites. This proximity also contributes for the rearrangements of the active sites towards an active conformation. Activation of the glutaminase active site must occur along with the displacement of the tyrosine plug in order to allow transferring ammonia towards the amidotransferase active site. The chain of events triggered by lipid II binding represents an expected regulation mechanism that allows glutamine hydrolysis and ammonia production only when the acceptor lipid II is bound to the complex, most likely in its phosphorylated form.

4.6. References

1. Figueiredo TA, Sobral RG, Ludovice AM, de Almeida JMF, Bui NK, 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**.
2. Münch D, Roemer T, Lee SH, Engeser M, Sahl HG & Schneider T (2012) Identification and *in vitro* analysis of the GatD/MurT enzyme-complex catalyzing lipid II amidation in Staphylococcus *aureus*. *PLoS Pathog* **8**, 1–11.
3. Zapun A, Philippe J, Abrahams KA, Signor L, Roper DI, Breukink E & Vernet T (2013) *In vitro* Reconstitution of Peptidoglycan Assembly from the Gram-Positive Pathogen *Streptococcus pneumoniae*. *ACS Chem Biol* **8**, 2688–2696.
4. Vollmer W, Blanot D & De Pedro MA (2008) Peptidoglycan structure and architecture. *FEMS Microbiol Rev* **32**, 149–167.
5. Zalkin H & Smith JL (1998) Enzymes Utilizing Glutamine as an Amide Donor. In *Advances In Enzymology And Related Areas Of Molecular Biology* pp. 87–144. John Wiley & Sons, Inc.
6. Smith CA (2006) Structure, Function and Dynamics in the *mur* Family of Bacterial Cell Wall Ligases. *J Mol Biol* **362**, 640–655.
7. Leisico F, Vieira D V., Figueiredo TA, Silva M, Cabrita EJ, Sobral RG, Ludovice AM, Trincão J, Romão MJ, 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.
8. Patin D, Boniface A, Kova A, Hervé M, Dementin S, Barreteau H, Mengin-Lecreulx D & Blanot D (2010) Purification and biochemical characterization of Mur ligases from *Staphylococcus aureus*. *Biochimie* **92**, 1793–1800.
9. El Zoeiby A, Sanschagrin F & Levesque RC (2003) Structure and function of the Mur enzymes: development of novel inhibitors. *Mol Microbiol* **47**, 1–12.
10. Mol CD, Brooun A, Dougan DR, Hilgers MT, Tari LW, Wijnands RA, Knuth MW, McRee DE & Swanson R V. (2003) Crystal Structures of Active Fully Assembled Substrate- and Product-Bound Complexes of UDP-N-Acetylmuramic Acid:L-Aalanine Ligase (MurC) from *Haemophilus influenzae*. *J Bacteriol* **185**, 4152–4162.
11. Emanuele Jr. JJ, Jin H, Yanchunas Jr. J & Villafranca JJ (1997) Evaluation of the kinetic mechanism of *Escherichia coli* uridine diphosphate-*N*-acetylmuramate:L-alanine ligase. *Biochemistry* **36**, 7264–7271.
12. Anderson MS, Eveland SS, Onishi HR & Pompliano DL (1996) Kinetic mechanism of the *Escherichia coli* UDPMurNAc-tripeptide D-alanyl-D-alanine-adding enzyme: Use of a glutathione S-transferase fusion. *Biochemistry* **35**, 16264–16269.
13. Šink R, Barreteau H, Patin D, Mengin-Lecreulx D, Gobec S & Blanot D (2013) MurD enzymes: some recent developments. *Biomol Concepts* **4**, 539–556.
14. Bertrand JA, Auger Á, Martin L, Fanchon E, Blanot D, Beller D Le, Heijenoort J Van,

Dideberg O & Grenoble F- (1999) Determination of the MurD Mechanism Through Crystallographic Analysis of Enzyme Complexes. *J Mol Biol* **289**, 579–590.

15. Nöldeke ER, Muckenfuss LM, 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.
16. Morlot C, Straume D, Peters K, Hegnar OA, Simon N, Villard A-M, Contreras-Martel C, Leisico F, Breukink E, Gravier-Pelletier C, Le Corre L, Vollmer W, Pietrancosta N, Håvarstein LS & Zapun A (2018) Structure of the essential peptidoglycan amidotransferase MurT/GatD complex from *Streptococcus pneumoniae*. *Nat Commun* **9**, 1–12.
17. Kantardjieff KA & Rupp B (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals. *Protein Sci* **12**, 1865–1871.
18. Kabsch W (2010) *XDS* . *Acta Crystallogr Sect D Biol Crystallogr* **D66**, 125–132.
19. Evans PR & Garib N (2013) How good are my data and what is the resolution? *Acta Crystallogr Sect D Biol Crystallogr* **D69**, 1204–1214.
20. Collaborative Computational Project N 4 (1994) The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr Sect D* **50**, 760–763.
21. Mccoy AJ, Grosse-kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) *Phaser* crystallographic software. *J Appl Crystallogr* **40**, 658–674.
22. Stein N (2008) *CHAINSAW*: a program for mutating pdb files used as templates in molecular replacement. *J Appl Crystallogr* **41**, 641–643.
23. Bunkóczi G & Randy J (2011) Improvement of molecular-replacement models with Sculptor. *Acta Crystallogr Sect D Biol Crystallogr* **67**, 303–312.
24. Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr* **66**, 486–501.
25. Afonine P V., Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH & Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr Sect D Biol Crystallogr* **68**, 352–367.
26. Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, III WBA, Snoeyink J, Adams PD, Lovell SC, Richardson JS & Richardson DC (2018) MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293–315.
27. Krissinel E & Henrick K (2007) Inference of Macromolecular Assemblies from Crystalline State. *J Mol Biol* **372**, 774–797.
28. Langer GG, Cohen SX, Lamzin VS & Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* **3**, 1171–1179.
29. Keegan RM & Winn MD (2008) *MrBUMP*: an automated pipeline for molecular replacement. *Acta Crystallogr Sect D Biol Crystallogr* **64**, 119–124.

30. Keegan RM, McNicholas SJ, Thomas JMH, Simpkin AJ, Simkovic F, Uski V, Ballard CC, Winn MD, Wilson KS & Rigden DJ (2018) Recent developments in *MrBUMP*: better search-model preparation, graphical interaction with search models, and solution improvement and assessment. *Acta Crystallogr Sect D Struct Biol* **74**, 167–182.

31. Mouilleron S & Golinelli-Pimpaneau B (2007) Conformational changes in ammonia-channeling glutamine amidotransferases. *Curr Opin Struct Biol* **17**, 653–664.

32. Šink R, Kotnik M, Zega A, Barreteau H, Gobec S, Blanot D, Dessen A & Contreras-Martel C (2016) Crystallographic Study of Peptidoglycan Biosynthesis Enzyme MurD: Domain Movement Revisited. *PLoS One* **11**, 1–17.

33. Nöldeke ER & Stehle T (2019) Unraveling the mechanism of peptidoglycan amidation by the bifunctional enzyme complex GatD/MurT: A comparative structural approach. *Int J Med Microbiol* **309**, 151334.

34. Krishna SS, Majumdar I & Grishin N V. (2003) SURVEY AND SUMMARY: Structural classification of zinc fingers. *Nucleic Acids Res* **31**, 532–550.

35. Gonçalves B V., Portela R, Lobo R, Figueiredo TA, Grilo IR, Ludovice AM, Lencastre H de, Dias JS & Sobral RG (2019) The role of MurT C-terminal domain in the amidation of *Staphylococcus aureus* peptidoglycan. *Antimicrob Agents Chemother* **63**, e00957-19.

36. List F, Vega MC, Razeto A, Häger MC, Sterner R & Wilmanns M (2012) Catalysis uncoupling in a glutamine amidotransferase bienzyme by unblocking the glutaminase active site. *Chem Biol* **19**, 1589–1599.

37. Hart EJ & Powers-lee SG (2008) Mutation analysis of carbamoyl phosphate synthetase: does the structurally conserved glutamine amidotransferase triad act as a functional dyad? *Protein Sci* **17**, 1120–1128.

38. Massière F & Badet-Denisot MA (1998) The mechanism of glutamine-dependent amidotransferases. *Cell Mol Life Sci* **54**, 205–222.

39. Rutkowska-Zbik D, Witko M & Fiedor L (2013) Ligation of water to magnesium chelates of biological importance. *J Mol Model* **19**, 4661–4667.

40. Basavannacharya C, Moody PR, Munshi T, Cronin N, Keep NH & Bhakta S (2010) Essential residues for the enzyme activity of ATP-dependent MurE ligase from *Mycobacterium tuberculosis*. *Protein Cell* **1**, 1011–1022.

41. Ruane KM, Lloyd AJ, Fülöp V, Dowson CG, Barreteau H, Boniface A, Dementin S, Blanot D, Mengin-lecreulx D, Gobec S, Dessen A & Roper DI (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.

42. Pei J, Kim B-H & Grishin N V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res* **36**, 2295–2300.

4.7. Supplementary material

***S. aureus* COL MurT**

```
        10         20         30         40         50         60
MRQWTAIHLA KLARKASRAV GKRGTDLPGQ IARKVDTDVL RKLAEQVDDI VFISGTNGKT

        70         80         90        100        110        120
TTSNLIGHTL KANNIQIIHN NEGANMAAGI TSAFIMQSTP KTKIAVIEID EGSIPRVLKE

       130        140        150        160        170        180
VTPSMMVFTN FFRDQMDRFG EIDIMVNNIA ETISNKGIKL LLNADDPFVS RLKIASDTIV

       190        200        210        220        230        240
YYGMKAHAHE FEQSTMNESR YCPNCGRLLQ YDYIHYNQIG HYHCQCGFKR EQAKYEISSF

       250        260        270        280        290        300
DVAPFLYLNI NDEKYDMKIA GDFNAYNALA AYTVLRELGL NEQTIKNGFE TYTSDNGRMQ

       310        320        330        340        350        360
YFKKERKEAM INLAKNPAGM NASLSVGEQL EGEKVYVISL NDNAADGRDT SWIYDADFEK

       370        380        390        400        410        420
LSKQQIEAII VTGTRAEELQ LRLKLAEVEV PIIVERDIYK ATAKTMDYKG FTVAIPNYTS

       430
LAPMLEQLNR SFEGGQS
```

***S. aureus* COL GatD-His$_6$**

```
        10         20         30         40         50         60
MHELTIYHFM SDKLNLYSDI GNIIALRQRA KKRNIKVNVV EINETEGITF DECDIFFIGG

        70         80         90        100        110        120
GSDREQALAT KELSKIKTPL KEAIEDGMPG LTICGGYQFL GKKYITPDGT ELEGLGILDF

       130        140        150        160        170        180
YTESKTNRLT GDIVIESDTF GTIVGFENHG GRTYHDFGTL GHVTFGYGNN DEDKKEGIHY

       190        200        210        220        230        240
KNLLGTYLHG PILPKNYEIT DYLLEKACER KGIPFEPKEI DNEAEIQAKQ VLIDRANRQK

       250
KSRLEHHHHH H
```

Figure S4.1. MurT and GatD constructs.

Amino acid sequences of the proteins used in this work, which correspond to *S. aureus* COL MurT untagged (full-length of 437 residues – 49,206 Da) and GatD tagged at the C-terminal (full-length of 243 residues + His6 tag that is underlined in the figure – 28,495 Da).

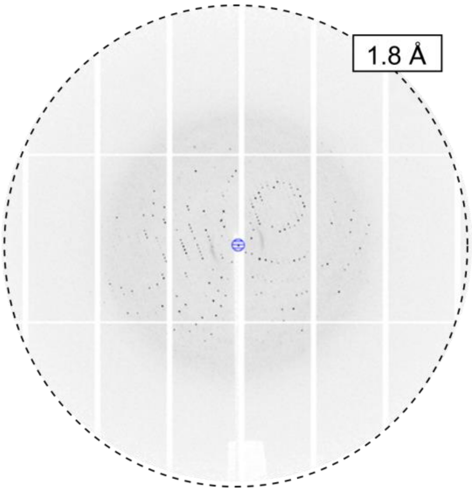


Figure S4.2. Diffraction pattern of *S. aureus* MurT-GatD crystals.

Representative diffraction pattern of the MurT-GatD crystal, where the resolution boundary is at 1.8 Å.

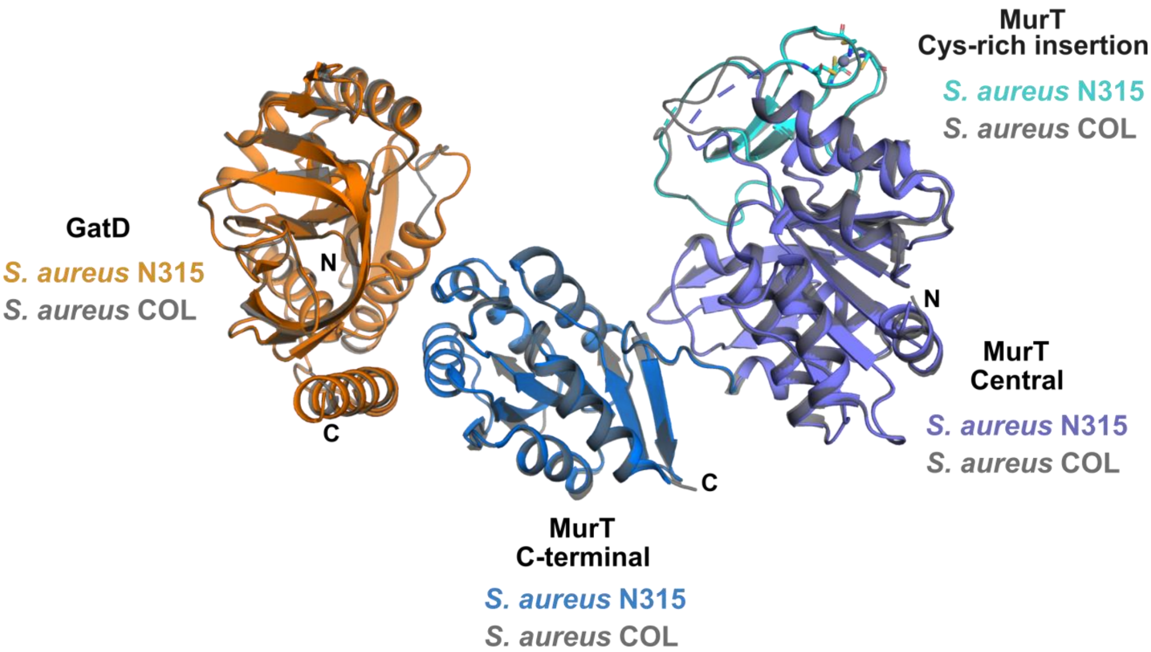


Figure S4.3. Crystal structures of MurT-GatD from *S. aureus* COL and N315 strains.

Superposition of the MurT-GatD heterodimers from *S. aureus* COL strain (coloured grey) and *S. aureus* N315 strain coloured according to Fig. 4.3. GatD, MurT central domain, MurT C-terminal domain and MurT Cys-rich insertion are labelled. The two heterodimers are aligned through GatD proteins. The structures are represented as cartoon; the N- and C-terminus of each protein are labelled with the letters N and C, respectively.

```
MurT S. aureus        -----------------------------------MRQ------------
MurT S. pneumoniae    MN---------------------------------LKT------------
MurT E. faecalis      MG---------------------------------IRS------------
MurT M. tuberculosis  MV------------------------------------------------
MurC S. aureus        -----------------MTHYHFVGIKGS--GMSSLAQIMHDLGH-----
MurC H. influenzae    MKHSHEEIRKIIPEMRRVQQIHFIGIGGA--GMSGIAEILLNEGY-----
MurD S. aureus        ML---------------------------------NYTGL----------
MurD E. coli          MA------------------------------------------------
MurE S. aureus        MD---------------------------------ASTLFKKVKV-----
MurE M. tuberculosis  MSSLARGISRRRTEVATQVEAAPTGLRPNAVVGVRLAALADQVGAALAEG
MurF S. aureus        MIN------------------------------VTLKQIQSWIPC-----

                                                    10            20
                                                     -             -
MurT S. aureus        --------------------------WTAIHLAK---LARKASRAVGKRG
MurT S. pneumoniae    --------------------------TLGLLAGR---SSHFVLSRLGRG-
MurT E. faecalis      --------------------------HLAIAAGK---TSQWVLQTFFKGG
MurT M. tuberculosis  ----------------------TTRARLALAAGA---GARWASRVTGRGA
MurC S. aureus        -----------------------------EVQGSDIENYVFTEVALRNKG
MurC H. influenzae    -----------------------------QISGSDIADGVVTQRLAQAGA
MurD S. aureus        ------------------------ENKNVLVVGLAKSGYEAAKLLSKLGA
MurD E. coli          ----------------------DYQGKNVVIIGLGLTGLSCVDFFLARGV
MurE S. aureus        -KRVLGSLEQQIDDITTDSRTAREGSIFVASVGYTVDSHKFCQNVADQGC
MurE M. tuberculosis  PAQRAVTEDRTVTGVTLRAQDVSPGDLFAALTGSTTHGARHVGDAIARGA
MurF S. aureus        -EIEDQFLNQEINGVTIDSRAISKNMLFIPFKGENVDGHRFVSKALQDGA


MurT S. aureus        -TDL----------------------------------------------
MurT S. pneumoniae    -STL----------------------------------------------
MurT E. faecalis      -SSY----------------------------------------------
MurT M. tuberculosis  GAMI----------------------------------------------
MurC S. aureus        IKILPFDAN-------------------------NIKEDMVVIQGNAFAS
MurC H. influenzae    KIYIGHAEE-------------------------HIEGASVVVVSSAIKD
MurD S. aureus        NVTVNDGKDLSQDAHAKDLESMGISVVSGSHPLTLLDNNPIIVKNPGIPY
MurD E. coli          TPRVMDTRMTPPGLDKLPEAV---ERHTGSLNDEWLMAADLIVASPGIAL
MurE S. aureus        KLVVVNKEQ---------------------------------------SL
MurE M. tuberculosis  VAVLTDPAGVA-----------------------------------EIAG
MurF S. aureus        GAAFYQRGTPI-------------------------------------DE

                                  30        40        50        60
                                   -         -         -         -
MurT S. aureus        -----------PGQIARKVDTDILRKLAEQVDDIVFISGTNGKTTTSNLI
MurT S. pneumoniae    -----------PGKVALQFDKDILQSLAKNY-EIVVVTGTNGKTLTTALT
MurT E. faecalis      -----------PGKLALKIDPKILDTLAKDY-EIVVVTGTNGKTLTTALT
MurT M. tuberculosis  -----------GGLVAMTLDRSILRQLGMGR-RTVVVTGTNGKSTTTRMT
MurC S. aureus        SHEEIVRAHQ--LKLDVVSYNDFLGQIIDQY-TSVAVTGAHGKTSTTGLL
MurC H. influenzae    DNPELVTSKQ--KRIPVIQRAQMLAEIMRFR-HGIAVAGTHGKTTTTAMI
MurD S. aureus        TVSIIDEAVKRGLKILTEVE--LSYL-ISEA-PIIAVTGTNGKTTVTSLI
MurD E. coli          AHPSLSAAADAGIEIVGDIELFCR---EAQA-PIVAITGSNGKSTVTTLV
MurE S. aureus        PANVTQVVVPDTLRVASILAHTLYDYPSHQL-VTFGVTGTNGKTSIATMI
MurE M. tuberculosis  RAAVPVLVHPAPRGVLGGLAATVYGHPSERL-TVIGITGTSGKTTTTYLV
MurF S. aureus        NVSGPIIWVEDTLTALQQLAQAYLR--HVNP-KVIAVTGSNGKTTTKDMI
```

```
                        70        80                        90
                         -         -                         -
MurT S. aureus        GHTLKANNIQIIHN--NEG-------------ANM-AAGITSAFI--MQS
MurT S. pneumoniae    VGILKEVYGQVLTN--PS--------------GAN--MITGIATTFLTAK
MurT E. faecalis      VNILRQEFDEVLTN--PT--------------GAN--MEQGIVSTFLNAK
MurT M. tuberculosis  AAALGTLGAVATNAEGANMDAG-------------------LVAA--LAA
MurC S. aureus        SHVMNGDKKTSFLI--GDGT---------------------------GMG
MurC H. influenzae    SMIYTQAKLDPTFV-NGGLV--------------------KSAGK--NAH
MurD S. aureus        GDMFKKSRLTGRLS--GNIGY-------------------VASKV--AQE
MurD E. coli          GEMAKAAGVNVGVG--GNIG---------------------LPAL--MLL
MurE S. aureus        HLIQRKLQKNSAYL--GTNGFQINETKTKGANTTPETVSLTKKIK--EAV
MurE M. tuberculosis  EAGLRAAGRVAGLI--GTIGIRVGGADLPSALTTPEAPTLQAMLA--AMV
MurF S. aureus        ESVLHTEFKVKKTQ--GNYNN-----------------EIGLPLT--ILE

                       100       110          120       130       140
                         -         -            -         -         -
MurT S. aureus        --TPKTKIAVIEIDEGS---IPRVLKEVTPSMMVFTNFFRDQMDRFG-EI
MurT S. pneumoniae    SSKTGKNIAVLEIDEAS---LSRICDYIQPSLFVITNIFRDQMDRFG-EI
MurT E. faecalis      AKKGQKKFAVLEIDEAS---LSRVTKYIQPKLFVFTNIFRDQMDRYG-EI
MurT M. tuberculosis  --HRDAELAVLEVDEMH---VPHISDAVDPAVVVLLNLSRDQLDRVG-EI
MurC S. aureus        --LPESDYFAFEACEYR-----RHFLSYKPDYAIMTNIDFDHPDYFK-DI
MurC H. influenzae    --LGASRYLIAEADESD-----ASFLHLQPMVSVVTNMEPDHMDTYEGDF
MurD S. aureus        --VKPTDYLVTELSSFQ----LLGIEKYKPHIAIITNIYSAHLDYHE-NL
MurD E. coli          --DDECELYVLELSSFQ----LETTSSLQAVAATILNVTEDHMDRYPFGL
MurE S. aureus        --DAGAESMTLEVSSH--GLVLGRLRGVEFDVAIFSNLTQDHLDFHG-TM
MurE M. tuberculosis  --ERGVDTVVMEVSSH--ALALGRVDGTRFAVGAFTNLSRDHLDFHP-SM
MurF S. aureus        -LDNDTEISILEMGMSGFHEIEFLSNLAQPDIAVITNIGESHMQDLG-SR

                           150             160       170         180
                             -               -         -           -
MurT S. aureus        DIMVNNIAETISN------KGIKLLLNADDPFVSRLKIASDT--IVYYGM
MurT S. pneumoniae    YTTYNMILDAIRKV-----PTATVLLNGDSPLFYKP-TIPNP--IEYFGF
MurT E. faecalis      YTTYRLIMEGAAAA-----PEATILCNGDSPIFNSK-ETVNP--RKYYGF
MurT M. tuberculosis  NVIERTLRAGLARHP-----DAVVVANCDDVLMTSA-AYDSP--NVVWVA
MurC S. aureus        NDVFDAFQEMAHNVK------KGIIAWGDDEHLRKI-EADV---PIYYYG
MurC H. influenzae    EKMKATYVKFLHNLP----FYGLAVMCADDPVLMEL-VPKVG-RQVITYG
MurD S. aureus        ENYQNAKKQIYKNQT----EEDYLICNYHQRQVIES-EELKA--KTLYFS
MurD E. coli          QQYRAAKLRIYEN-------AKVCVVNADDALTMPI-RGADE--RCVSFG
MurE S. aureus        EAYGHAKSLLFSQLGEDLSKEKYVVLNNDDSFSEYL-RTVTP-YEVFSYG
MurE M. tuberculosis  ADYFEAKASLFDPDS--ALRARTAVVCIDDDAGRAM-AARAA-DAITVSA
MurF S. aureus        EGIAKAKSEITIGLK----DNGTFIYDGDEPLLKPH-VKEVENAKCISIG

                          190       200       210       220        230
                            -         -         -         -          -
MurT S. aureus        KAHA-HEFEQSTMNESRYCPNCGRLLQYDYIHYNQIGHYHC-QCGFKREQ
MurT S. pneumoniae    DLEKGPAQLAHYNTEGILCPDCQGILKYEHNTYANLGAYICEGCGCKRPD
MurT E. faecalis      NHLPPKEQLAHYNTDGVLCPKCNHILHYKMITYANLGDYYCPNCGFKRPE
MurT M. tuberculosis  AGG----------------------------------AWSNDSVSCPRSG
MurC S. aureus        FKD----------------------------------------------S
MurC H. influenzae    FSE----------------------------------------------Q
MurD S. aureus        TQQ----------------------------------------------E
MurD E. coli          VNM-----------------------------------------------
MurE S. aureus        IDE----------------------------------------------E
MurE M. tuberculosis  ADR----------------------------------------------P
MurF S. aureus        VAT----------------------------------------------D
```

```
                           240                        250            260
                             -                          -              -
MurT S. aureus        AKYEISSF-DVA----------------PFLYLNI-----NDEKYDMKIA
MurT S. pneumoniae    LDYRLTKLVELTN---------------NRSRFVI-----DGQEYGIQIG
MurT E. faecalis      LDVQLTEMVRMDN---------------TSADFVI-----DGEEYGIAVG
MurT M. tuberculosis  EVIVRKAPSQEDHWYSTGADFKRPAPHWWFDDATLYGPDGLALPMRLALP
MurC S. aureus        DDIYAQNIQITDK----------------GTAFDVYVDGEFYDHFLSPQY
MurC H. influenzae    ADYRIEDYEQTGF----------------QGHYTVICPNNERINVLLNVP
MurD S. aureus        VDGIYIKD-----------------------GFIVYKGVRIINTEDLVLP
MurD E. coli          GDYHLNHQQG--------------------ETWLRVKGEKVLNVKEMKLS
MurE S. aureus        AQFMAKNIQESL----------------QGVSFDFVT-PFGTYPVKSPYV
MurE M. tuberculosis  AHWRATDVAPTDA---------------GGQQFTAIDPAGVGHHIGIRLP
MurF S. aureus        NALVCSVDDRDTT----------------GISFTI----NNKEHYDLPIL

                             270       280       290       300
                               -         -         -         -
MurT S. aureus        GDFNAYNALAAYTVLRELGLNEQTIKNGFETYTSDNGRMQYFKKE----R
MurT S. pneumoniae    GLYNIYNALAAVAIARFLGADSQLIKQGFDKSRAVFGRQETFHIG----D
MurT E. faecalis      GMYNVYNALAATAVAEYYQVAPDKIRAGLAYDEKVFGRQETIKVD----D
MurT M. tuberculosis  GSVNRGNAAQAVAAAVALGADPAVAVAAVCQVDEVAGRYRTVRIG-----
MurC S. aureus        GDHTVLNALAVIAISYLEKLDVTNIKEALETFGGVKRRFNETTIA-----
MurC H. influenzae    GKHNALNATAALAVAKEEGIANEAILEALADFQGAGRRFDQLGEFIRPNG
MurD S. aureus        GEHNLENILAAVLACILAGVPIKAIIDSLTTFSGIEHRLQYVGTN----R
MurD E. coli          GQHNYTNALAALALADAAGLPRASSLKALTTFTGLPHRFEVVLEH----N
MurE S. aureus        GKFNISNIMAAMIAVWSKGTSLETIIKAVENLEPVEGRLEVLDPS----L
MurE M. tuberculosis  GRYNVANCLVALAILDTVGVSPEQAVPGLREIRVPG-RLEQIDRG----Q
MurF S. aureus        GKHNMKNATIAIAVGHELGLTYNTIYQNLKNVSLTGMRMEQHTLE----N

                       310         320       330       340        350
                         -           -         -         -          -
MurT S. aureus        KEAMINL--AKNPAGMNASLSVGEQLEGEKVYVISLND-NAADGRDTSWI
MurT S. pneumoniae    KECTLVL--IKNPVGATQAIEMIKLAPYPFSLSVLLNA-NYADGIDTSWI
MurT E. faecalis      KECTLVL--VKNPVGLNQVIDMMGLAPYSFSLVSLLNA-NYADGIDVSWI
MurT M. tuberculosis  AHQARILL-AKNPAGWQEALAMVDKHA-DGVVIAVNGRVP-DGEDLSWLW
MurC S. aureus        NQVIVDDY-AHHPREISATIETARKKYPHKEVVAVFQP-HTF--SRTQAF
MurC H. influenzae    KVRLVDDY-GHHPTEVGVTIKAAREGWGDKRIVMIFQP-HRY--SRTRDL
MurD S. aureus        TNKYYNDSKATNTLATQFALNSFN-----QPIIWLCGGLD-RGNEF-DEL
MurD E. coli          GVRWINDSKATNVGSTEAALNGLHVD---GTLHLLLGGDG-KSADF-SPL
MurE S. aureus        PIDLIIDY-AHTADGMNKLIDAVQPFVK-QKLIFLVGMAGERDLTKTPEM
MurE M. tuberculosis  GFLALVDY-AHKPEALRSVLTTLAHPD--RRLAVVFGAGGDRDPGKRAPM
MurF S. aureus        DITVINDAYNASPTSMRAAIDTLSTLT--GRRILILGD-VLELGENSKEM

                         360        370                   380        390
                           -          -                     -          -
MurT S. aureus        YDADFEKLSKQ-QIEAIIVTGT-----------RAEELQLRLKLAE-VEV
MurT S. pneumoniae    WDADFEQITDM-DIPEINAGGV-----------RHSEIARRLRVTGYPAE
MurT E. faecalis      WDGNHEAFADM-DIPKVIAGGD-----------RHEDMALRLKVAGIPEE
MurT M. tuberculosis  -----DVRFEHFEKTRVVAAGE-----------RGTDLAVRLGYAG----
MurC S. aureus        LNEFAESLS---KADRVFLCEIFGSIRENTGALTIQDLIDKIEGA-----
MurC H. influenzae    FDDFVQVLS---QVDALIMLDVYAAGEAPIVGADSKSLCRSIRNLGKV--
MurD S. aureus        -----IPYME--NVRAMVVFGQ-----------TKAKFAKLGNSQGK---
MurD E. coli          -----ARYLNG-DNVRLYCFGR-----------DGAQLAALRPE------
MurE S. aureus        -----GRVAC--RADYVIFTPDNPAND------DPKMLTAELAKGATHQ-
MurE M. tuberculosis  -----GRIAAQ-LADLVVVTDDNPRDEDPT--AIRREILAGAAEVGGDA-
MurF S. aureus        HIGVGNYLEEK-HIDVLYTFGN-----------EAKYIYDSGQQHVE---
```

```
                             400         410
                               -           -
MurT S. aureus        PIIVERDIYKATAKTM--DYKGFTVAI------------------PNY--
MurT S. pneumoniae    KITETSNLEQVLKTIEN-QDCKHAYILATYTA-----------MLEFR--
MurT E. faecalis      KLLEIADLEQVISEIKT-LPTDHVYILATYTA-----------VLQLR--
MurT M. tuberculosis  ---VEHTLVHDTVAAIASCPPGRVEVVANYTA------------------
MurC S. aureus        ------SLINEDSINVLEQFDNAVILFMGAGD------------------
MurC H. influenzae    DPILVSDTSQLGDVLDQIIQDGDLILAQGAGS------------------
MurD S. aureus        SVIEANNVEDAVDKVQDIIEPNDVVLLSPACAS-WDQYS-----TFEERG
MurD E. coli          VAEQTETMEQAMRLLAPRVQPGDMVLLSPACAS-LDQFK-----NFEQRG
MurE S. aureus        NYIEFDDRAEGIKHAIDIAEPGDTVVLASKGREPYQIMPGHIKVPHRDD-
MurE M. tuberculosis  QVVEIADRRDAIRHAVAWARPGDVVLIAGKGHETGQRGG-GRVRPFDDR-
MurF S. aureus        KAQHFNSKDDMIEVLINDLKAHDRVLVKGSRG------------------

                     420       430
                       -         -
MurT S. aureus        TSLAPMLEQLNRSFE--GGQS--
MurT S. pneumoniae    ELLASR-------QIVRKEMN--
MurT E. faecalis      KSLTAQ-------GYIQGGMNGV
MurT M. tuberculosis  --FLQLQRALARRG---------
MurC S. aureus        --IQKLQNAYLDKLGMKNAF---
MurC H. influenzae    --VSKISRGLAESWKN-------
MurD S. aureus        EKFIERFRAHLPSY---------
MurD E. coli          NEFARLAKELG------------
MurE S. aureus        LIGLEAAYKKFGGGPVD------
MurE M. tuberculosis  VELAAALEALERRA---------
MurF S. aureus        MKLEEVVNALIS-----------
```

Figure S4.4. Sequence alignment of MurT and Mur ligases.

Multiple sequence and structure alignment using the PROMALS3D server [42]. The crystal structures used in this work were considered to generate the 3D constrains for the alignment: *S. aureus* MurT-GatD (PDB 6H5E), *S. pneumoniae* MurT-GatD (PDB 6QFB), *H. influenzae* MurC (PDB 1P3D), *E. coli* MurD (PDB 2UAG), *S. aureus* MurE (PDB 4C12) and *M. tuberculosis* MurE (PDB 2XJA). The residue numbering provided corresponds to *S. aureus* MurT.

Chapter 5

CONCLUSIONS AND FUTURE PERSPECTIVES

5.1. General conclusions

A major threat to public health worldwide comes in the form of antibiotic resistant bacteria. These pathogens have developed mechanisms to overcome the antimicrobial activity of existing antibiotics, causing hard to treat infections. Currently, there is a lack of drugs that act either as standalone strategy or in a combinatorial approach to restore the bacteria's susceptibility to existing drugs. Antibiotics target mainly the peptidoglycan (PG) component of bacterial cell wall, which forms a mesh-like structure with protective roles against osmotic, chemical and mechanical stresses. Recently discovered enzymes involved in the biosynthesis pathway of PG constitute promising targets for the development of new antibiotics. One example is the bi-enzymatic complex MurT-GatD that catalyses the reaction of PG amidation in Gram-positive bacteria, an essential modification to accomplish the efficient mesh-like structure of a mature cell wall. Indeed, MurT-GatD was shown to be critical for the survival of important pathogens such as *S. aureus*, *S. pneumoniae* and *M. tuberculosis*, confirming its potential therapeutic role.

PG amidation occurs in the cytoplasm, on the PG precursor lipid II, which comprises the disaccharide Mur*N*Ac-Glc*N*Ac bound to the undecaprenyl-phosphate lipid carrier and to a stem peptide composed of L-Ala-D-γ-Glu-L-Lys-D-Ala-D-Ala. The amidation reaction involves the conversion of D-Glu into D-isoGln. Preliminary sequence analysis showed that GatD resembles the glutaminase domain of glutamine amidotransferases, which is responsible for producing the ammonia used by the synthase domain of the complex to amidate the final substrate. MurT resembles Mur ligases, the enzymes responsible for assembling the PG stem peptide. MurT and GatD proteins form an obligatory complex and catalyse the reaction of lipid II amidation using glutamine as amide donor, along with ATP and Mg^{2+}. These observations confirmed the glutamine amidotransferase activity of the complex. Nevertheless, molecular insights on the reaction mechanism of the reaction were missing.

In order to characterize the activity of MurT-GatD and disclose the reaction mechanism of PG amidation, a structure-function study was developed integrating structural data obtained from X-ray crystallography and SAXS, with complementary functional experiments following the glutaminase and amidotransferase activity of the complex.

The crystal structure of the isolated *S. aureus* GatD provided the first structural insights on the MurT-GatD complex (**Chapter 2**). GatD showed a structure similar to the glutaminase domains of glutamine amidotransferases, including the characteristic nucleophilic cysteine and polarizing histidine of catalytic triads. Moreover, a glutamine molecule was found at the surface of the protein interacting with a conserved arginine residue, suggesting a potential mechanism by which the substrate is captured from the environment.

In order to complete the structure-function studies on MurT-GatD, the full complex from *S. aureus* COL strain was crystallized and the structure solved at 2.9 Å of resolution (**Chapter 4**). As a heterodimer, GatD adopts the same glutaminase fold as in its isolated form, while MurT shows central and C-terminal domains resembling Mur ligases. The MurT central domain contains a Cys-rich insertion that is missing in Mur ligases. The heterodimer possesses a boomerang shape, where the C-terminal domain of MurT interacts with GatD. The structure of MurT-GatD from *S. aureus* N315 strain was recently published, supporting our observations. Also the crystal structure of *S. pneumoniae* MurT-GatD became available, providing the opportunity to develop comparative studies with insights on the function of the complex in organisms other than Staphylococci. Interestingly, the heterodimer from *S. pneumoniae* showed a compact conformation, resulting from the rigid body rotation of MurT central domain towards GatD when compared to the complex of *S. aureus*. These striking differences were also observed in solution using SAXS studies (**Chapter 3**). SPR experiments showed that *S. aureus* GatD and MurT C-terminal domain interact with nanomolar affinity (**Chapter 3**), confirming the stability of the heterodimer in an extended conformation. Interestingly, the complex from *S. aureus* showed lower affinity towards ATP and lipid II than the complex from *S. pneumoniae*, likely due to their different conformations.

So far, four crystal structures of the MurT-GatD system have been reported: *S. aureus* GatD alone, *S. aureus* MurT-GatD with and without AMP-PNP, and *S. pneumoniae* MurT-GatD bound to glutamine. Together, they provide a high-resolution structural framework to explore their conformational landscape and the functional insights governing the reaction mechanism of PG amidation (**Chapter 4**).

Both crystal structures of *S. aureus* and *S. pneumoniae* MurT-GatD showed the interaction between GatD and MurT C-terminal domain. This interaction is characterized by a conserved hydrophobic core, which agrees with their nanomolar affinity measured for the complex of *S. aureus*. While the hydrophobic core in *S. aureus* MurT-GatD is complemented by polar interactions between GatD and the C-terminal domain of MurT, the complex from *S. pneumoniae* possesses an exclusive interface patch between GatD and the central domain of MurT (**Chapter 4**). These differences at the MurT-GatD interface might contribute for the stabilization of their respective extended and compact conformations. Interestingly, some of the residues involved in the different interface patches are conserved in orthologs, suggesting that the two conformations can be part of the MurT-GatD conformational landscape.

The MurT-GatD structures confirmed the similarity of the glutaminase active site with triad GATases but showed two different acidic residues interacting with the catalytic histidine. In *S. aureus* MurT-GatD the triad is completed with a conserved aspartate belonging to MurT, while in *S. pneumoniae* complex the triad is composed exclusively by GatD residues. However, the overall glutaminase active site adopts a non-competent configuration in both structures due to the displacement of the oxyanion hole residues from their catalytic position. The conserved Y30 from *S. pneumoniae* GatD deserves special attention due to its location at the channel connecting the glutaminase and the amidotransferase active sites, where it potentially regulates the transfer of ammonia. Compared to *S. aureus* Murt-GatD, the structure of *S. pneumoniae* MurT-GatD shows the stabilization of this Tyr and other residues at the active site by interactions with the MurT central domain, suggesting that its compact conformation and catalytic triad might be closer to the catalytic form.

The amidotransferase active site of MurT-GatD comprises the binding sites of ATP and lipid II. The architecture of the ATP binding site is highly conserved with that in Mur ligases. Interestingly, the residues of Mur ligases interacting with the Mg^{2+} ion and the carboxyl group of the substrate before phosphorylation are conserved in MurT proteins and show similar locations in the crystal structures of the MurT-GatD complex. Therefore, it is expected that the D-Glu carboxyl group of lipid II is activated by phosphorylation by a similar mechanism. The MurT residues that are conserved but do not superimpose with Mur ligases are located in the C-terminal domain, supporting the formation of an active conformation in which the central and C-terminal domains sandwich the ATP and the lipid II molecules, forming the aminotransferase active site. In this respect, the compact conformation of *S. pneumoniae* MurT-GatD is likely to be closer to a catalytically active form than the extended *S. aureus* complex, which agrees with their respective affinities to ATP and lipid II.

The knowledge gathered so far allows proposing a model for the reaction mechanism of lipid II amidation: (i) both MurT-GatD crystal structures represent non-competent catalytic forms of the complex and binding of all substrates probably induce conformational changes towards an active conformation like the stabilization of the loop located between the glutaminase and amidotransferase active sites; (ii) activation of the glutaminase active site must occur along with the displacement of the tyrosine plug in order to allow transferring ammonia towards the amidotransferase active site; (iii) activation must occur upon lipid II binding, serving as a regulation mechanism that allows glutamine hydrolysis and ammonia production only when the acceptor lipid II is bound to the complex.

5.2. Future perspectives

Even though the observations collected in this work depict molecular mechanisms involved in the reaction of PG amidation, many interesting questions are still open to research.

One big question still to be answered focuses on why the *S. aureus* and *S. pneumoniae* MurT-GatD complexes adopt two distinct conformations *in vitro*. On one hand, the glutamine molecule found at the glutaminase active site of *S. pneumoniae* MurT-GatD might have promoted a conformational transition to a compact structure. Replacing *S. pneumoniae* GatD R142 (equivalent to the essential *S. aureus* GatD R128) by an alanine residue is likely to impair the ability of the complex in capturing the substrate from the environment. SAXS and X-ray crystallography could be used to evaluate changes in the overall conformation of the complex. On the other hand, the role of the interactions between GatD and MurT central domain of *S. pneumoniae* that involve residues specific of this organism could also be addressed regarding their contribution for the compact conformations.

Understanding the function of MurT-GatD requires the structure of the complex in the presence of glutamine, ATP and lipid-II. This form could be trapped by replacing ATP by its non-hydrolysable analogue AMP-PNP, as observed for crystal structures of Mur ligases. Non-hydrolizable analogs of glutamine such as 6-Diazo-5-oxo-L-norleucine could also be used. Such structure would provide crucial insights: (i) the overall active conformation of the complex; (ii) the catalytic triad of the glutaminase active site, sorting out the roles of the different acidic residues identified in the crystal structures of *S. aureus* and *S. pneumoniae* MurT-GatD; (iii) the lipid II binding site; and (iv) the regulation mechanisms associated to lipid II binding, like the conformational rearrangements required for the activation of the glutaminase active site and establishment of the ammonia channel.

An interesting topic concerns the presence of MurT-GatD at the cell membrane, which brings the cellular context to the function of the complex. The localization of MurT-GatD could be tackled *in vivo* using fluorescence microscopy. A truncated version of MurT lacking the N-terminal domain could be used to study its potential role either in directly interacting with the membrane or promoting protein-protein interactions. The role of MurT Cys-rich insertion in contacts with membrane proteins is also a venue to be explored. These interactions could be assessed by yeast-two hybrid experiments, with special attention to other enzymes involved in the PG biosynthesis pathway. Regulation of biosynthesis at the membrane level by supramolecular complexes is a strong hypothesis to be tested. Moreover, studies at the cellular context in thriving and stress conditions are mandatory to understand bacteria biosynthesis regulation during bacteria adaptation.

Even though, many biological questions will continue to fuel research in this topic, the insights gathered in this thesis and in the work of others can be readily applied. Specifically, medicinal chemistry endeavours can profit from the structural data available to design small molecules with inhibitory activity against MurT-GatD. These can target both the glutaminase and the amidotransferase active sites of the complex, as well, as the interface of the heterodimer in order to prevent complex assembly or trap the complex in an incompetent conformation. The functional assays optimized in our work can be used to assess these two routes of inhibition. Antimicrobials against MurT-GatD are already specific for Gram-positive bacteria, however exploring the conformations landscape of MurT-GatD complexes from different organisms might allow the design of species-specific drugs. This ideal scenario would allow reducing even more the perspectives of resistance mechanism dissemination.

The collaboration developed during the execution of this work between the groups of Macromolecular Crystallography and Molecular Microbiology of Bacterial Pathogens at FCT-NOVA, along with the Pneumococcus Group at IBS (France) allows further developing this research project. A new partner from the University of Tübingen (Germany) with expertise in *M. tuberculosis* joined our international efforts in the quest for new drugs with potential to treat infections caused by the important pathogens *S. aureus*, *S. pneumoniae* and *M. tuberculosis*, which can either be used in a standalone strategy or in combination with current antibiotics.

UNRAVELLING THE REACTION MECHANISM OF AMIDATION
IN *STAPHYLOCOCCUS AUREUS* PEPTIDOGLYCAN

FRANCISCO LEISICO

2022