Bárbara Vitorino Gonçalves

Degree in Biology

A new antimicrobial target in *Staphylococcus aureus* – amidation of peptidoglycan

Dissertation presented to obtain a master degree in
Molecular Genetics and Biomedicine

Supervisor: Dr. Rita Sobral, CREM-FCT/UNL
Co-supervisor: Dr. Eurico Cabrita, Requimte-FCT/UNL
Co-supervisor:Prof. Ana Madalena, DCV-FCT/UNL; ITQB/UNL

Jury:

President: Prof. Dr. Paula Maria Theriaga Mendes Bernardes Gonçalves
Examiner: Dr. Jorge da Silva Dias
Supervisor: Dr. Rita Sobral

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro 2014

Bárbara Vitorino Gonçalves

Degree in Biology

A new antimicrobial target in *Staphylococcus aureus* – amidation of peptidoglycan

Dissertation presented to obtain a master degree in
Molecular Genetics and Biomedicine

Supervisor: Dr. Rita Sobral, CREM-FCT/UNL
Co-supervisor: Dr. Eurico Cabrita, Requimte-FCT/UNL
Co-supervisor:Prof. Ana Madalena, DCV-FCT/UNL; ITQB/UNL

Jury:

President: Prof. Dr. Paula Maria Theriaga Mendes Bernardes Gonçalves
Examiner: Dr. Jorge da Silva Dias
Supervisor: Dr. Rita Sobral

Setembro 2014

A new antimicrobial target in *Staphylococcus aureus* - amidation of peptidoglycan

Acknowledgments

I would like to start by expressing my profound gratitude to PhD Rita Sobral, PhD Eurico Cabrita and Prof. Ana Madalena Ludovice for giving me the opportunity to participate in this Master project and gain experience in environmental research with the Universidade Nova de Lisboa.

A special acknowledge to my professor and supervisor Rita Sobral for the monitoring, support, transmission of knowledge, and patience given, during the development of this thesis.

Equally I would like to mention my deep gratitude to my colleagues Raquel Portela, Lia Domingues and Lia Godinho, who were a reliable and experienced source of knowledge and support during the development of this thesis.

I must extend my gratitude to Ricardo Lobo for the monitoring given during the development of this thesis.

I have to mention my gratitude to my ex-college and college, Damien Costa, Mario Ferreira, Tiago Dias and Cynthia Barroco.

I am also immensely grateful to my friend Sara Silva, Andreia Soares, Ana Lucas, Aristides Mendes for all the emotion support and strength given during the development of the thesis

Last but not least I have to mention my extremely thankful to my family.

Resumo

Staphylococcus aureus é um perigoso agente patogénico oportunista associado a um grande número de fatores de virulência e responsável por um vasto leque de doenças. A sua importância clínica deve-se à sua elevada capacidade de acumular mecanismos de resistência, a, virtualmente, todos os antibióticos. Durante o último século, muitos esforços têm sido efetuados no estudo dos mecanismos moleculares subjacentes à resistência de *S. aureus* aos β-lactâmicos. Além do gene exógeno, *mecA*, identificado como sendo o principal responsável do mecanismo de resistência aos *β*-lactâmicos, vários genes constitutivos, também foram identificados como sendo necessários para a expressão ótima de resistência, tendo-lhes sido atribuída a designação de genes auxiliares. O uso de agentes que atuam em sinergia com os β-lactâmicos já provou ser eficaz na recuperação da atividade de β-lactâmicos contra estirpes resistentes. Entre os genes auxiliares, o operão *murt-gatD* recentemente identificado, codifica para as enzimas responsáveis pela amidação do resíduo de glutamato no peptidoglicano. Amidação do glutamato, uma modificação secundária do peptidoglicano, é essencial para a viabilidade de *S. aureus* e está envolvido nos mecanismos de resistência a *β*-lactâmicos e à lisozima, sendo um excelente alvo para o desenvolvimento de novos compostos antimicrobianos.

O objetivo desta dissertação de mestrado foi caracterizar dinamicamente a estrutura do complexo MurT-GatD por Ressonância Magnética Nuclear (RMN). A informação estrutural sobre a interação física entre as duas proteínas será uma contribuição essencial para o futuro desenvolvimento de um composto que bloqueia a associação MurT-GatD. Tal composto, agiria em sinergia com os *β*-lactâmicos, restaurando a susceptibilidade em *S. aureus*. Os resultados deste estudo identificaram o domínio DUF1727 (domínio C-terminal da proteína MurT) como responsável pela interação entre MurT e GatD bem como também mostrou, através de análise por RMN (^{1}H-^{15}N-espectros), uma forte interação entre DUF1727 e proteína GatD. Este trabalho abriu caminho para a análise global da interação de Murt-GatD.

Palavras-Chaves: Complexo MurT-GatD; amidação do peptidoglicano; *Staphylococcus aureus*; resistência aos antibióticos β-lactâmicos; determinação da estrutura por RNM.

Abstract

Staphylococcus aureus is a dangerous opportunistic pathogen with a large number of virulence factors, responsible for a wide range of diseases. Its clinical importance is mainly due to its high capacity to accumulate resistance mechanisms to virtually all antibiotics. Over the last century, huge efforts have been dedicated to the study of the molecular mechanisms underlying resistance of *S. aureus* to β-lactam antibiotics. Besides the exogenous *mecA* gene, identified as the main player of the mechanism of β-lactam resistance, several housekeeping genes were also identified to be required for the optimal expression of resistance, the so-called auxiliary genes. The use of combination agents in synergy with β-lactams has already proven to be efficient in restoring β-lactam activity against resistant strains. Among the auxiliary genes, *murT-gatD* operon was recently identified as encoding for the enzymes responsible for amidation of the glutamate residue of peptidoglycan. Glutamate amidation, a secondary modification of peptidoglycan, is essential for *S. aureus* viability and is involved in the mechanisms of resistance to β-lactams and to lysozyme, being an excellent target for the development of new antimicrobial compounds.

The purpose of this Master thesis was to dynamically characterize the structure of MurT-GatD complex by Nuclear Magnetic Resonance (NMR). Structural information on the physical interaction between the two partner proteins will be an essential contribution for the future development of a compound that blocks MurT-GatD association. Such a compound would act in synergy with β-lactams, restoring *S. aureus* susceptibility. The results of this study identified the DUF1727 domain (C-terminal domain of MurT protein) as responsible for the interaction between MurT and GatD and also showed, through ^{1}H-^{15}N-spectrum NMR analysis, a strong interaction between DUF1727 and GatD protein. This work paved the way for the comprehensive analysis of MurT-GatD interaction.

Keywords: MurT-GatD complex; Peptidoglycan amidation; Staphylococcus aureus; resistance to β-lactams; structure determination by NMR.

Table of Contents

List of Figures

Chapter 1- Introduction

Chapter 3- Results

Chapter 4 - Discussion

Chapter 6- Annex

List of Tables

Acronyms

Cm	Chloramphenicol
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
HSQC	Heteronuclear Single Quantum Coherence
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Km	Kanamycin
PMSF	Phenylmethanesulfonyl Fluoride
SDS	Sodium dodecyl sulfate

Introduction

1.1. General characteristics of the *Staphylococcus* genus

Staphylococcus is a commensal and pathogenic genus of humans and other mammals. It is composed by gram-positive coccal bacterium, facultative anaerobic, that belong to the family of *Staphylococcaceae,* order *Bacillales,* class *Bacilli.* Together with *Streptococcus, Staphylococcus* is the gram-positive genus with more clinical importance, being distinguished by the production of the catalase enzyme which converts hydrogen peroxide into water and oxygen (Wilkinson, 1997).

The name *Staphylococcus* was introduced by Ogston in 1883 for the group micrococci causing inflammation and suppuration and a formal description of this genus was provided by Rosenbach in 1884 (Ogston,1883; Rosenbach, 1884).

The classification of this genus has changed over time, according to the typing methods used, first phenotypic, then genotypic (Vandenbergh and Verbrugh, 1996). Traditionally, they were divided into two groups according to the coagulase reaction that provides the ability to clot blood plasma. The coagulase-positive *Staphylococcus* includes two species, *Staphylococcus aureus* and *Staphylococcus intermedius*, while all the other Staphylococci are coagulase-negative. Currently, the taxonomy is based on 16s rRNA sequences, dividing this genus in 11 clusters, being *S. aureus* and *S. simiae* in the same cluster.

The staphylococcus genus includes 36 species and 17 subspecies, being *S. aureus* and *S. epidermidis* the most relevant species concerning interaction with humans. For this reason, their epidemiology, virulence and antibiotic resistance mechanisms were the focus of a huge number of studies in the past decades.

1.2. *Staphylococcus aureus*

First of all, *S. aureus* is a commensal bacterium of human and several animal species (Williams, 1963) being the nose, more specifically the anterior nares, its main ecological niche. It can however, colonize other body sites such as the skin and the nasopharynx. Over 80% of the population can carry this bacterium, persistently (20%) or intermittently (60%) without any associated symptoms (asymptomatic carriers), stimulating the propagation of *S. aureus* (Peacock *et al*., 2001*;* Wertheim *et al*., 2005). Although it is primarily a colonizer, *S. aureus* can also be a pathogen, having been established an association between *S. aureus* nasal carriage and staphylococcal infection by Danbolt in 1931 (Solberg, 1965; Vandenbergh and Verbrugh, 1996; Wertheim *et al.*, 2005).

S. aureus, as a well-adapted opportunistic pathogen, is the leading human pathogen of this genus, with high associated mortality. It is not only the most common cause for nosocomial infections worldwide but is also a dangerous infectious agent among the healthy community (Grundmann *et al.*, 2006). The success of *S. aureus* as a pathogen is mainly due to their high capacity to develop resistance to virtually all classes of antimicrobials available and to their high number of virulence factors (see section 1.3 and 1.4).

1.3. *S. aureus* as a successful pathogen

The diversity of the virulence factors of *S. aureus* is an essential factor to the wide range of diseases caused by this bacterium, from skin and soft tissue infections (noninvasive infections) such as impetigo, folliculitis, cellulitis, to more serious and life-threatening infections (invasive infections), such as pneumonia, bacteremia or endocarditis. Besides that, absorption of toxins from food products colonized by *S. aureus*, can cause acute gastroenteritis (Boucher *et al*., 2010; Lee, 2003; Projan *et al*., 1997).

Most virulence factors are involved in one of the following processes: adherence of *S. aureus* to surfaces, avoidance of the host immune system and damage of the host. Most virulence factors are cell-surface-associated, such as protein A, wall teichoic acids and the polysaccharide capsule which help to avoid phagocytosis and consequently allow *S. aureus* to evade the host defenses. Beside this, invasion factors, such as proteases, toxins and superantigens, promote the bacterial attack of the host and prevent the development of a strong antibody response, compromising the immune memory and leading to repeated infections (Foster, 2005; Projan *et al*., 1997).

Another important characteristic for the success of *S. aureus* as a pathogen is its insensitivity to lysozyme, a bactericidal protein produced by the innate immune system. This enzyme is present in most human body fluids such as, saliva, sweat and tears, and its production is increased during infection. Lysozyme is a muramidase that interact and hydrolyzes the peptidoglycan (see section 1.5), resulting in cell wall degradation and cell lysis (Bera *et al.*, 2005; Levy, 2000 Levy O, 2000; Schindler, *et al.*, 1997).

1.4. Resistance to antibiotics in *S. aureus* – an historic perspective

The development of antibiotics remains one of the most significant advances in modern medicine and continues to be the basis of therapy against bacterial infections, being the β-lactams (penicillin and its derivatives) the class of antibiotics most frequently used for gram-positive infections (Fisher *et al.*, 2005).

Penicillin was the first β-lactam used to combat *S. aureus* infections and was discovered in 1928 by Alexander Fleming (Fleming, 1929), but only in 1940 Florey and Chain showed its high efficiency against gram-positive bacteria (Heatley, 1940). Until the introduction of penicillin, the mortality rate of *S. aureus* infections was 80%; with penicillin, the mortality rate was reduced drastically, however, resistant strains were reported two years after its introduction (Deurenberg *et al.*, 2007; Friedmann, 1948). Penicillin resistance resulted from the acquisition of a plasmid encoding for a β-lactamase protein (penicillinase) (Abraham and Chain, 1940) that hydrolyzes the beta-lactam ring and consequently inactivates the antibiotic (Ghuysen, 1991). To overcome the rapidly acquired resistance to this natural product, semi-synthetic antibiotics, derivatives of penicillin, modified to resist to β-lactamase action, were created and introduced in clinical practice to combat *S. aureus* penicillin resistant infections: first methicillin in 1960 (Rolinson *et al*., 1960) and then, oxacillin, cloxacillin and

others. However, only one year later, in 1961, was the first methicillin resistant *S. aureus* (MRSA) strain reported, carrying the *mecA* gene (Jevons, 1961).

Besides β-lactams antibiotics, MRSA strains also developed resistance to all other classes of antibiotics that were introduced into clinical practice, such as macrolides, chloramphenicol and tetracycline, targeting protein synthesis or fluoroquinolones and rifampicin, targeting nucleic acid synthesis (Bambeke, *et al.,* 2003).

S. aureus became a serious problem to public health due to its capacity to acquire resistance to most classes of antibiotics available. Since the 1980s, MRSA became the major cause of nosocomial infections worldwide, the hospital-acquired MRSA strains (HA-MRSA). As the MRSA frequency begun to decrease in the hospitals, around 1990, the community-acquired MRSA strains (CA-MRSA) with increased virulence, emerged, causing infections among healthy individuals with no risk factors or contact with health care facilities. Later, such community associated strains emerged also in the hospital settings (de Lencastre *et al.*, 2007; Grundman *et al.*, 2006).

For treatment of serious MRSA infections, the last resort drug used was vancomycin (a glycopeptide antibiotic), however, vancomycin resistance was first reported in 2002 and VRSA resistant strains were already reported worldwide (Pupp *et al*., 2003). Recently, new antimicrobial drugs were introduced into clinical practice with high efficiency against MRSA strains, such as linezolid (Gu *et al*., 2013), tigecycline and daptomycin (Dabul and Camargo, *et al*., 2014). However, *S. aureus* seems to be winning this battle again, as new resistant strains have already emerged. New strategies to address β-lactam resistance are urgently needed.

1.4.1. β-lactam mechanism of action

There are different types of β-lactams in terms of spectrum, pharmacokinetics and activity but all of them share a common mode of action (Fisher *et al.*, 2005). The target of the β-lactam antibiotic is the penicillin binding proteins (PBPs) that are involved in the final steps of peptidoglycan biosynthesis (se section 1.5) (Bambeke *et al*., 2003). The β-lactam ring mimics the natural substrate of PBPs, the D-alanyl-D-alanine carboxy-terminus of the peptidoglycan pentapeptide (On *et al.,* 1966). The β-lactams compete with peptidoglycan for binding to PBPs (Figure 1.1): the carboxylate or the sulfonate group of the β-lactams acetylates the PBPs to form a acyl-enzyme complex (Yocum *et al*., 1979). In contrast to the complex with the natural substrate, the β-lactam-PBP complex is stable resulting in irreversible inhibition of the PBP function (see section 1.5). β-lactams are the most used antibiotics due to their high efficiency and low toxicity, result of the fact that their target is the absent from all organisms besides bacteria.

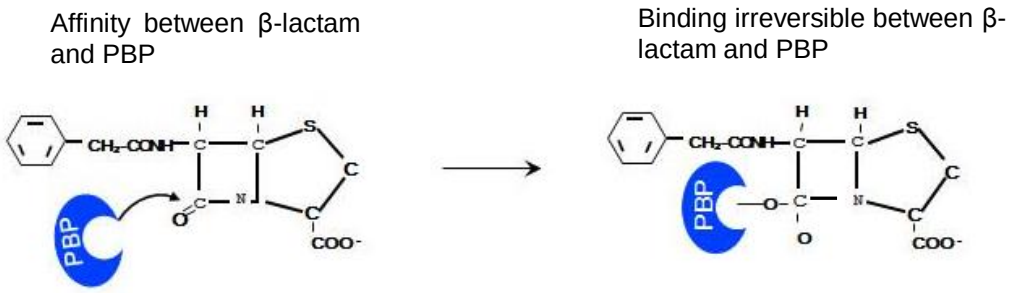

Figure 1.1. β-lactam mechanism of action. The PBP present in the cell wall of the bacteria recognize and binds to the β-lactam antibiotic, being inactivated. Adapted from Sobral, *et al*, (unpublished).

1.4.2. β-lactam resistance

Resistance to methicillin and to all β-lactams antibiotics was only recognized in the 1980s and involves an acquisition of the exogenous gene *mecA* that encodes for an extra PBP (PBP2A) (Reynolds and Brown 1985), possibly transmissible by lateral transfer from *S. sciuri* or *S. fleuretti* (Couto et al., 1996; (Tsubakishita *et al.*, 2010). The *mecA* gene is located on a large heterologous chromosomal island called the staphylococcal chromosomal cassette *mec* (SCC*mec*), that integrates into the *S. aureus* chromosome at a site-specific location, (*attBscc)* (Deurenberg *et al*., 2007).

SCC*mec* carry not only the *mecA* gene but also the genes responsible for their regulation, the repressor MecI and the trans-membrane β-lactam-sensing signal-transducer MecRI. Briefly, in the absence of a β-lactam antibiotic, MecI represses the transcription of all genes (*mecA* and *mecRI* and *mecI*), whereas in the present the antibiotic, MecRI becomes active after autocatalytic cleavage allowing transcription of *mecA* and subsequent production of PBP2A (Lowy, 2003).

Initially, PBP2A was believed to take over the biosynthesis of staphylococcal cell wall in the presence of antibiotics, as it has low affinity to β-lactam antibiotics and its activity is enough to restore the function lost by the others native PBPs (see section 1.5). However, the β-lactam resistance mechanism is more complex as it is not only *mecA*-dependent, but multifactorial. The PBP2A expression is necessary but not sufficient for the optimal resistance phenotype and needs the optimal expression of many housekeeping genes (Murakami and Tomasz, 1989).

1.4.3. The contribution of auxiliary genes for the β-lactam resistance mechanism

For optimal expression of methicillin resistance, besides *mecA* gene, a large number of native genes, called *aux* genes, are required. This was demonstrated by the construction of transposition libraries, in the background of MRSA strains, which were screened for reduced levels of methicillin resistance (Tn*551* mutants) (Berger-Bächi, 1983; de Lencastre, *et al*., 1999) to identify the inactivated genetic determinants. This strategy identified many genes directly or indirectly involved in cell wall biosynthesis, such as: *femA* and *femB* (Berger-Bächi, 1983; Berger-Bächi, *et al*., 1989; Henze, *et al.*, 1993), genes involved in the synthesis of the interpeptide glycine bridge (de Jorge *et al.*, 1992; Jonge *et al.*, 1993); *femC (glnR),* a glutamine synthesis repressor (Gustafson *et al.*, 1994); *femD*, the gene

that encoding GlmM phosphoglucosamine mutase (Jolly *et al.*, 1997); the *murE* gene that encodes the MurE ligase (Gardete *et al.*, 2004); *pbpB* gene, encoding PBP2 (Pinho *et al.*, 2001); the *murF* gene that encodes the MurF ligase protein was identified as an auxiliary gene by plasmid insertion (Sobral *et al.*, 2007). The *murJ* gene, encoding a proposed translocase, was identified as auxiliary gene through reverse genetic strategy using inducible antisense interference fragments targeting specific *S. aureus* essential genes (Huber *et al.*, 2009). Recently, two new auxiliary genes were identified, *murT and gatD*, implicated in the amidation of peptidoglycan of *S. aureus* (Figueiredo *et al.*, 2012; Münch *et al.*, 2012)*)*. Besides the auxiliary genes involved in peptidoglycan biosynthesis, genes involved in other processes such as, cell division (*ftsz* and *ftsA*), protein secretion (*spsB*), wall teichoic acid biosynthesis (*tarL*) and signal transduction systems (*pknB*), were found to also contribute to MRSA β-lactam resistance (Lee *et al.*, 2011). The characterization of these auxiliary genes, new potential antimicrobial targets, is essential to define new strategies to combat MRSA.

1.5. Peptidoglycan biosynthesis in *Staphylococcus aureus*

Peptidoglycan is the major component of the cell wall and is essential for cell survival, providing structural strength (Gally and Archibald 1993), maintenance of cell shape and counterbalance of the internal osmotic pressure (Höltje, 1998; Vollmer *et al.*, 2008). Peptidoglycan is composed of repeating disaccharide units of *N*-acetylglucosamine(GlcNAc)-(β-1,4)-*N*-acetylmuramic acid (MurNAc) (Figure 1.2), forming parallel glycan chains that are cross-linked by short peptides (Heijenoort, 2001).

The peptidoglycan biosynthesis is a complex pathway, which requires numerous enzymatic reactions and occurs in three different compartments of the cell: first in the cytoplasm, then in the inner face of the membrane and finally at the cell wall, as shown in Figure 1.3.

Briefly, the assembly of the peptidoglycan monomer unit occurs in the cytoplasm where GlmS, GlmM and GlmU enzymes sequentially convert fructose-6-phosphate to UDP-GlcNAc which is later converted to UDP-MurNAc by MurA, an enzyme that catalyzes the condensation of phosphoenolpyruvate and MurB, an enzyme that reduces this unit to a lactyl group. Then five residues (L-Ala-D-Glu-L-Lys-D-Ala-D-Ala) are sequentially added through the enzymatic activity of the Mur ligase family enzymes, MurC, MurD, MurE and MurF, respectively (Heijenoort, 2001; Smith, 2006;). To complete the monomer synthesis and facilitate export to the cell surface, the UDP-MurNAc-pentapeptide unit is attached to a membrane-bound lipid carrier molecule (C55-P) by the membrane protein MraY, forming Lipid I (C55-P-P-MurNAc-pentapeptide). Then, MurG transfers UDP-GlcNAc to lipid I, forming lipid II (C55-P-P-GlcNAc-MurNAc-pentapeptide) (van Heijenoort, 2007). In the next step of biosynthesis, five glycine residues are added to the L-lysine residue at position 3 of the pentapeptide chain by FemA, FemB and FemX (Gly-tRNA synthetase), forming the interpeptide bridge (Hegde and Shrader, 2001; Schneider *et al*., 2004). In addition, GatD and MurT enzymes amidate the D-glutamate at position 2 (Figueiredo *et al,.* 2012; Münch *et al*., 2012;) into glutamine. The precursor molecule is then translocated across the membrane by FtsW and/or RodA activity (Mohammadi *et al*., 2011), exposing the substrate to the action of PBPs. Depending on their specific enzymatic activity, PBPs incorporate the precursor into the growing peptidoglycan layer by two types of reactions:

transglycosylation, elongation of the glycan chains and transpeptidation, cross-linking of the peptidoglycan muropeptides, between D-Ala and L-Lys residues of adjacent muropeptides with the concomitant cleavage of the terminal D-alanine of the peptide chain (Scheffers and Pinho, 2005).

PBPs can be classified into two major groups: high and low-molecular-weight (HMW or LMW). HMW PBPs are two domain proteins that can be divided into class A or B, according their N-terminal domain. Both classes share a structurally common C-terminal domain, responsible for transpeptidation and for binding to β-lactams. The HMW class A-PBPs are bifunctional enzymes with transglycosylase activity at the N-terminal region while the class B HMW-PBPs are monofunctional and their N-terminal region has unknown role *(Goffin et al.,* 1998). *S. aureus* has four native PBPs (Georgopapadakou and Liu, 1980) and MRSA strains have an extra PBP, PBP2A. The native PBP1 is a class B HMW-PBP and plays a key role in cell growth and division (Pereira *et al.,* 2007). PBP2 is a class A HMW-PBP, a bifunctional enzyme that functionally cooperates with PBP2A, conferring resistance to β-lactams in MRSA strains (Pinho *et al.,* 2001). PBP3 is a class B HMW-PBP essential for maintenance of the size and shape of the cell, however its inactivation produces no change in the muropeptide composition of the cell wall (Pinho and Lencastre, 2000). PBP4 is a LMW-PBP with carboxypeptidase and β-lactamase activity, which overexpression results in an increase in β-lactam resistance and in greater cross-linking of the peptidoglycan (Navratna *et al.,* 2010).

Peptidoglycan biosynthesis has been extensively studied and the synthesis steps of the primary backbone have been known for a long time, however, the mechanisms and the proteins involved on the secondary modifications such as acetylation (Bera *et al*., 2005; Bera *et al*., 2006) and amidation (Figueiredo *et al*., 2012; Münch *et al.,* 2012), were only recently described.

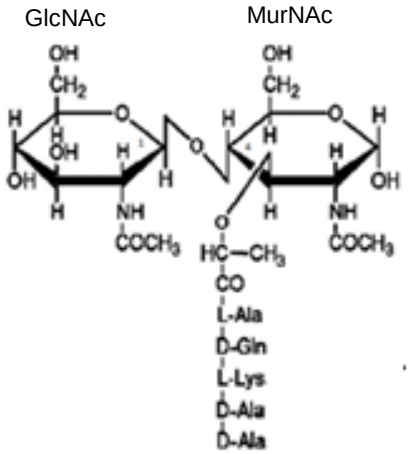

Figure 1.2. Pentapeptide monomer of peptidoglycan of *S. aureus.* Adapted from Scheffers and Pinho, 2005. GlcNAc- *N*-acetylglucosamine; MurNAc- *N*-acetylmuramic acid.

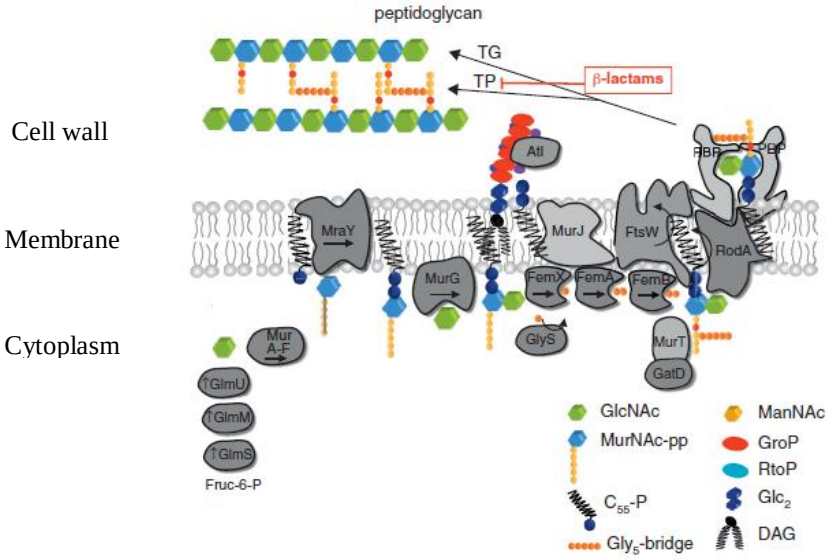

Figure 1.3. Representation of peptidoglycan biosynthesis in *S. aureus.* Adapted from Roemer*, et al.,,* Microbiology 2013.

1.6. Combination agents for antibiotic synergy strategies

Since resistant strains to all new antibiotics introduced in clinical practice were already reported, different approaches are needed and the combination of a β-lactam drug with a second agent to restore β-lactams efficiency against MRSA, can provide a viable solution. In vitro, combination agents that restore β-lactam efficacy already exist. Moenomycin, a non-clinically used inhibitor of transglycoslation, acts synergistically with β-lactams by inactivating the PBP2 transglycosylase domain that is required, together with the transpeptidase domain PBP2A, for peptidoglycan synthesis in the present of β-lactams (Pinho *et al.*, 2001). Cyslabdan, an actinomycete metabolite, was found to potentiate β-lactam activity against MRSA by inhibiting FemA and the pentaglycine bridge synthesis, however, a synergistic effect was obtained only which Carbapenem, more specifically, the imipenem (Fukumoto *et al.*, 2008).

This approached is recent and more studies are needed to obtain more effective antibiotics or combination agents, being crucial the identification of others *aux* genes and an improved structural and functional understanding of the aux genes that are already identified.

1.7. Peptidoglycan amidation

Amidation is a secondary modification to the peptidoglycan primary structure. With this modification, D-glutamate, the second amino acid of the pentapeptide chain is amidated (a group NH_3 is transferred to the amino acid) into D-iso-glutamine. This reaction is catalyzed by MurT and GatD enzymes and uses Lipid II as substrate (Münch *et al*., 2012). The respective genes are part of a small operon, *murT-gatD* and are expressed as a single transcript. MurT and GatD form a physically stable bi-enzymatic complex in a molar ratio of 1:1 and share identity and similarity with two families of proteins: the Mur ligases and the glutamine amidotransferases (GATases). More specifically, MurT

protein shares the characteristics of the Mur ligases central domain, which includes the conserved motifs for ATP and Mg^{2+} binding, however, its C-terminal domain, the DUF1727 domain exhibits no identity or similarity with any described protein domain.

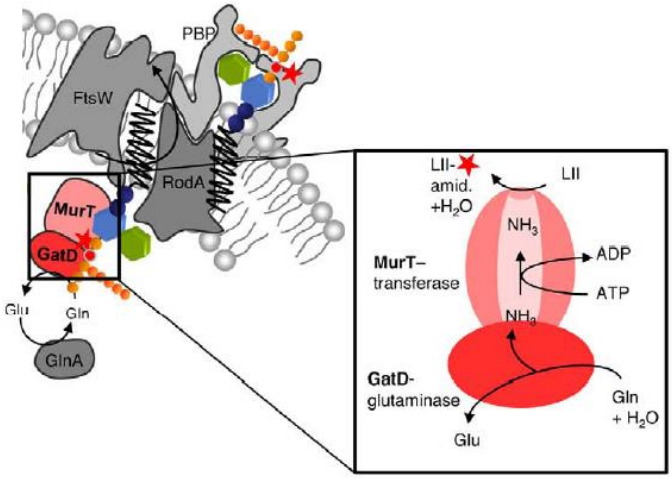

Figure 1.4. Model for the GatD/Murt amidation in cell wall of S. *aureus* adapted from Münch*, et al.,* 2012.

The GatD protein shares identity with the conserved residues important for glutaminase activity that belong to the class I-type GATases. Interestingly, the motifs for ATP binding present in GATases are not present in GatD protein. Due to these similarities, MurT was proposed to be responsible for the recognition of lipid II, the lipid-linked muropeptide and ATP, while GatD would be the catalytic subunit involved in the transfer of the amino group from free glutamine to the peptidoglycan precursor (Figueiredo *et al.*, 2012; Münch *et al*., 2012), as shown in Figure 1.4. The GatD active site has a canonical cysteine at position 94, and its substitution by mutagenesis resulted in the loss of amidation activity (Münch *et al*., 2012). The GatD protein has already been crystallized (Vieira *et al*., 2014), however its structure is not yet available.

1.7.1. Impact of PG amidation in β-lactam resistance

The amidation of peptidoglycan was shown to reduce the growth rate and to be necessary for resistance to β-lactam antibiotics, making *murT* and *gatD* auxiliary genes. Furthermore, the amidation step was found to be essential for the *S. aureus* intrinsic resistance to lysozyme (Figueiredo *et al.*, 2012), probably because amidation of D-glutamate results in a less negatively charged peptidoglycan.

Amidation of peptidoglycan is thought to play a role in the coordination of the biochemical events that occur after the precursor crosses the membrane, translocation, transglycoslation and transpeptidation (Münch *et al.*, 2012).

The amidation step has a three-fold advantage: it is an essential enzymatic step for *S. aureus*, it is essential for the optimal expression of β-lactam resistance in MRSA and is involved in staphylococcal intrinsic lysozyme resistance. The combination of a new antimicrobial agent that inhibits the amidation step together with a β-lactam drug may be an efficient solution to eliminate MRSA and other resistant

strains. For this reason, the peptidoglycan amidation is an excellent candidate as the target of a new antimicrobial compound.

1.7.2. Importance of peptidoglycan amidation for the resistance mechanisms of the most widespread clones of MRSA

The lack of peptidoglycan amidation showed a different impact in the expression of resistance to β-lactams and to lysozyme according to the genetic backgrounds of the MRSA strains tested: CA-MRSA strains showed lower levels of resistance to β-lactams then HA-MRSA strains when the operon *murT/gatD* was inhibited, and an identical behavior was observed for lysozyme resistance (Figueiredo *et al.,* 2014). Among the CA-MRSA strains, strains without any genetic relation with HA-MRSA exhibited a higher decreased resistance upon amidation inhibition, than strains genetically related with HA-MRSA strains. Inhibition of the amidation step showed more impact in strains with difficulty to maintain or stabilize the *mecA* gene. To reinforce that, *murT/gatD* mutation was transferred into *mecA*-independent resistant strains and a decrease in the level of resistance occurred, indicating that peptidoglycan amidation is also essential for a *mecA*-independent resistant strategy. In short, peptidoglycan amidation is involved through different mechanistic links in the β-lactam resistance strategies of strains from distinct backgrounds, indicating the existence of more than one physiological approach in *S. aureus* for survival to antibiotic stress (Figueiredo *et al*., 2014).

1.7.3. Peptidoglycan amidation in other species

The amidation peptidoglycan step was only described in gram-positive bacteria, not occurring in gram-negative bacteria. Sequence alignments searches showed the presence of homologs of *murT/gatD* only in other Gram-positive bacteria, in four phyla (*Dictyoglomi, Chloroflexi, Actinobacteria, Cyanobacteria*) two classes (*Clostridia, Erysipelotrichi*), seven families (*Alicyclobacillaceae, Aerococcaceae, Carnobacteriaceae, Enterococcaceae, Lactobacillaceae, Leuconostocaceae, Streptococcaceae*) and two genus (*Macrococcus, Staphylococcus*). Interestingly, it is also present in one phylum of Archaea domain (Euryarchaeota) (Figueiredo *et al.*, 2012).

The presence of homologs of this operon in *Mycobacterium tuberculosis, Streptococcus pneumonia, Clostridium perfringens* is in accordance with their peptidoglycan analysis that shows D-iso-glutamine in the peptidoglycan stem, indicating that amidation occurs in this three species (Schleifer and Kandler, 1972).

Recently, spr1443/spr1444 operon was identified in *Streptococcus pneumonia* as the orthologous of *murT/gatD.* In this species, the transpeptidation efficiency was proven to be dependent on the amidation status of the PBPs substrate. Amidated lipid II was preferred for the transpeptidation activity of PBP2a, 2b, and 2x. The PBP1a demonstrated a higher transpeptidation activity with amidated lipid II but it retained some activity with non-amidated lipid II (Zapun *et al*., 2013). These observations evidenced that amidation is required for the PBP-catalyzed polymerization of peptidoglycan in *S. pneumonia*.

1.8. Rational drug design

Rational drug design is the process by which new drugs are developed based on the knowledge of a specific target. It involves the design of small molecules that are complementary in shape and charge to a biomolecular target.

In contrast to traditional methods of drug discovery, which rely on trial-and-error testing of natural or synthetic chemical compounds on cultured cells or animals, rational drug design begins with the hypothesis that modulation of a specific biological target may have a therapeutic outcome (Mandal *et al.*, 2009).

Nuclear magnetic resonance (NMR) spectroscopy is a unique methodology as it is able to reveal the atomic structure of macromolecules in solution, provided that highly concentrated samples (approx. 1 mM) are available (Bharti and Roy, 2012). This technique depends on the fact that certain atomic nuclei are intrinsically magnetic. Only a limited number of isotopes display this property, called spin (½), namely ^{1}H, ^{15}N, ^{13}C, being they natural abundance (% By weight Nucleus of the element) 99.984%, 0.365% and 1.108%, respectively.

For the hydrogen nucleus (^{1}H), the spinning of a proton generates a magnetic moment, with two possible orientations or spin states, when an external magnetic field is applied (α–protons align with the external magnetic field and β-protons align against the external magnetic field). A spinning protein in α state can be raised to an excited state (β state) by applying a pulse electromagnetic radiation (a radio-frequency). The spin will change from α to β and resonance will be obtained. A resonance spectrum for a molecule can be obtained by varying the magnetic field at a constant frequency of electromagnetic radiation or by keeping the magnetic field constant and varying electromagnetic radiation.

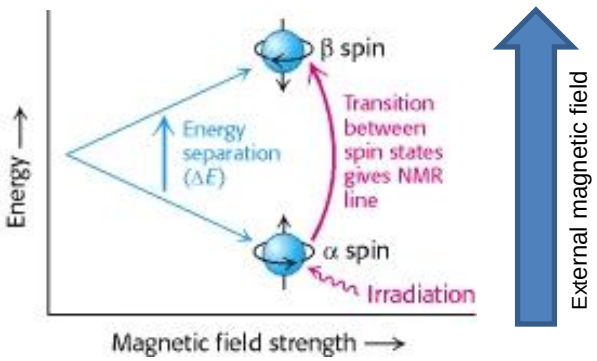

Figure 1.5- Transition between the energies of the two orientations of a nucleus of spin ½ when a magnetic field is applied. Adapted from Berg *et al.,* 2002.

These properties can be used to examine the chemical surroundings of the hydrogen nucleus.

The flow of electrons around a magnetic nucleus generates a small local magnetic field that opposes the applied field. The degree of such shielding depends on the surrounding electron density. Consequently, nuclei in different environments will change states, or resonate, at slightly different field strengths or radiation frequencies. The nuclei of the perturbed sample will absorb electromagnetic

radiation at a frequency that can be measured. The different frequencies, termed chemical shifts, are expressed in fractional units δ. The proton chemical shifts range from 0 ppm to 15 ppm) (Bharti and Roy, 2012; Jeremy *et al*., 2002).

As an example, protons of an unfolded protein will have a different chemical shifts than if in a folded state, due the different environment created by the increased proximity from others protons in the amino acid chain.

With this technique, one, two and tree-dimension spectra (^{1}H-spectrum, ^{1}H-^{15}N-spectrum and ^{1}H-^{15}N-^{13}C-spectrum) are obtained in order to determine the protein structure (Montelione *et al.*, 2000). To surpass the low natural abundance of ^{15}N and ^{13}C for obtaining two and tree-dimension spectra, incorporation of such isotopes must be forced by expressing the protein in minimal media supplemented with the isotopes.

To obtain the spectrum, the proteins must be pure (typically >95% is required), folded and at least marginally stable.

1.9. Thesis objectives

Despite the high number of antibiotics available, MRSA strains continue to represent a major threat to human health, in both hospital and community settings. Due to the constant emergence of antibiotic resistance mechanisms, the development of new strategies of treatment and new antimicrobials is essential to combat MRSA strains. Peptidoglycan amidation impacts the growth rate, the resistance mechanism to β-lactams and to lysozyme, being an essential step to the survival of *S. aureus* strains. The two proteins responsible for peptidoglycan amidation, MurT and GatD are attractive targets for the design of new antimicrobials; it is essential to well characterize the enzymatic activity of these proteins and their molecular structure. The objective of this Master project was to dynamically characterize the structure of MurT-GatD complex by Nuclear magnetic resonance (NMR).

The work plan of this Master project includes cloning *gatD* and *murT* genes of MRSA strain COL (Dyke *et al.,* 1966) to construct recombinant tagged proteins and optimize their expression in the *Escherichia coli* host strain. The production of ^{15}N, ^{13}C, uniformly labeled GatD protein and assignment of the respective ^{15}N HSQC NMR spectra will provide GatD structure and will be followed by interaction studies of ^{13}C, ^{15}N GatD with MurT by NMR spectroscopy.

Materials and Methods

2.1. Bacterial strains and plasmids

In this study, *S. aureus* strain COL (Dyke *et al.,* 1966) was grown at 37ºC with aeration in tryptic soy broth or agar (TSB/TSA) (Difco Laboratories, USA) and *Escherichia coli* strains were grown at 37ºC in Lysogeny broth or agar (LB/LA) (Difco Laboratories) or minimum medium (MM) (Weber *et al*., 1992) with the respective antibiotics for the selection and maintenance of *E. coli* transformants. These strains and their genotype are listed in Table 2.1, as well as the plasmids used. The strains used were selected due to their distinctive characteristics, namely: DH5α (Invitrogen) is frequently used for routine cloning applications because of its high level of transformation; BL21(DE3) (Novagen) is a widely used host for protein expression from T7 promoter and has the advantage of being deficient in the *lon (8)* and *ompT* proteases; BL21-CodonPlus(DE3)-RIPL (Novagen), Rosetta(DE3)pLysS (Novagen) and Rosetta(DE3) (Novagen) are used to facilitate the expression of genes that include rare *E. coli* codons. The first strain is a derivative of BL21(DE3) and is supplemented with genes that encode tRNAs for the arginine codons (AGA and AGG), the isoleucine codon (AUA), the leucine codon (CUA) and proline codon (CCC). Strain Rosetta(DE3) is a derivate from strain Tuner (Novagen) and is supplemented with genes that encode tRNAs for codons AUA, AGG, AGA, CUA, CCC and GGA (Glycine codon). Rosetta(DE3)pLys contains plasmid pLysS which carries the gene encoding T7 lysozyme, that allows a more controlled expression from T7 promoter.

The plasmid (Figure 5.1, annex) used in all the protein expression strategies was pET28a (Novagen) because it offers different cloning strategies, as shown in Table 2.1.

Table 2.1. Strains and plasmids used in this study.

Strains	Description	Source
S. aureus		
COL	Homogeneous Mc^r (MIC, 1600 µg/ml); Em^s	Rockefeller University Collection
E. coli		
DH5α	*recA endA1 gyrA96 thi-1 hsdR17 supE44 relA1 F80 DlacZDM15*	Invitrogen
DH5α+pET28a-His_6tag-GatD	DH5α with pET28a plasmid with His_6-*gatD* gene, Km^r	Figueiredo *et al.* (unpublished)
DH5α+pET28a-GatD-His_6tag	DH5α with pET28a plasmid with *gatD*-His_6 gene, Km^r	Lobo *et al.* (unpublished)

continues

Strains	Description	Source
DH5α+pET28a-MurT-GatD-His$_6$tag	DH5α with pET28a plasmid with murT-*gatD*-His$_6$ operon, Km^r	Figueiredo *et al.* (unpublished)
DH5α+pET28a-MurT-His$_6$tag	DH5α with pET28a plasmid with *murT* gene, Km^r	This study
DH5α+pET28a-DUF1727-His$_6$tag	DH5α with pET28a plasmid with *DUF1727*-His$_6$ domain, Km^r	This study
DH5α+pET28a-His$_6$tag-DUF1727	DH5α with pET28a plasmid with His$_6$-*DUF1727* domain, Km^r	This study
BL21(DE3)	*E. coli B F- dcm ompT hsdS(r_B^- m_B^-) gal λ(DE3)*	Novagen
BL21(DE3)+pET28a-His$_6$tag-GatD	BL21(DE3) with pET28a plasmid with His$_6$-*gatD* gene, Km^r	This study
BL21(DE3)+pET28a-GatD-His$_6$tag	BL21(DE3) with pET28a plasmid with *gatD*-His$_6$ gene, Km^r	This study
BL21(DE3)+pET28a-MurT-His$_6$tag	BL21(DE3) with pET28a plasmid with *murT*-His$_6$ gene, Km^r	This study
BL21(DE3)+pET28a-DUF1727-His$_6$tag	BL21(DE3) with pET28a plasmid with *DUF1727*-His$_6$ gene, Km^r	This study
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]*	Novagen
BL21-CodonPlus(DE3)-RIPL+pET28a	BL21-CodonPlus(DE3)-RIPL with pET28a plasmid	Figueiredo *et al.* (unpublished)
BL21-CodonPlus(DE3)-RIPL+pET28a–His$_6$tag-GatD	BL21-CodonPlus(DE3)-RIPL with pET28a plasmid with His$_6$-*gatD gene,* Km^r, Cm^r	Figueiredo *et al.* (unpublished)
BL21-CodonPlus(DE3)-RIPL+pET28a–GatD-His$_6$tag	BL21-CodonPlus(DE3)-RIPL with pET28a plasmid with *gatD*-His$_6$ *gene,* Km^r, Cm^r	This study
BL21-CodonPlus(DE3)-RIPL+pET28a-MurT-His$_6$tag	BL21-CodonPlus(DE3)-RIPL with pET28a plasmid with *murT*-His$_6$ *gene,* Km^r, Cm^r	This study

continues

Strains	Description	Source
BL21-CodonPlus(DE3)-RIPL-pET28a-MurT-GatD-His$_6$tag	BL21-CodonPlus(DE3)-RIPL with pET28a plasmid with *murT-gatD*-His$_6$ operon, Km^r, Cm^r	This study
BL21-CodonPlus(DE3)-RIPL-pET28a-DUF1727-His$_6$tag	BL21-CodonPlus(DE3)-RIPL with pET28a plasmid with *DUF1727*-His$_6$ *gene,* Km^r, Cm^r	This study
Rosetta(DE3)	*F$^-$ ompT hsdS$_B$(r_B^- m_B^-) gal dcm (DE3) pRARE (Cam^r)*	Novagen
Rosetta(DE3)+pET28a-DUF1727-His$_6$tag	Rosetta(DE3) with pET28a plasmid with *DUF1727*-His$_6$ *gene,* Km^r, Cm^r	This study
Rosetta(DE3)pLysS	*F$^-$ ompT hsdS$_B$(r_B^- m_B^-) gal dcm (DE3) pLysSRARE (Cam^R)*	Novagen
Rosetta(DE3)pLysS-His$_6$tag-GatD	Rosetta(DE3) with pET28a plasmid with *his$_6$-GatD gene,* Km, Cm	This study
Plasmids		
pET28a	T7 lac promoter/ T7 transcription star/ His.tag coding sequence/ T7.tag coding sequence/ Multiple cloning sites (*BamH* I – *Xho* I)/ His.tag coding sequence /T7 terminator/ *lacI* coding sequence/ pBR322 origin/ Kan coding sequence/ f1 origin	Novagen

2.2. DNA methods

Chromosomal DNA from strain COL was extracted using Wizard Genomic DNA Purification kit (Promega, USA) as suggested by the manufacturer, with some modifications, namely, an initial cell lysis step, in Tris pH=8 supplemented with 10 μg/ml of lysostaphin (AMBI PRODUCTS LLC, USA) and 30 μg/ml of RNase (SIGMA, USA). Plasmid pET28a was extracted using ZR Plasmid Miniprep ™-Classic (Zymo Research, USA) of DH5α strain. The DNA clean & concentrator™ (Zymo Research) system was used to purify PCR products and digestion products.

2.3. Cloning strategies

The purified COL DNA was used as a template for the PCR amplification of *gatD* gene (SACOL1950, 732 bps fragment), *murT* gene (SACOL1951, 1314 bps fragment), DUF1727 domain (DNA fragment including the 379 bps of the 3´terminal end of *murT* gene), and DUF1727-gatD sequence (fragment with 1111 bps). All PCR amplifications were performed using Thermocycler (T1 Biometra, Germany) and the primers are listed in Table 2.2. The proofreading *phusion* DNA polymerase (Thermo Scientific, USA) was used for cloning purposes and the Taq polymerase (NZYTech, USA) for routine purposes.

The amplification conditions used were as follows: a 98ºC step of 30 s, followed by 30 cycles, each consisting of 98ºC for 10 s, 50ºC for 30 s, 72ºC for 1 min (or 2 min for *murT* amplification), and one final extension step of 72ºC for 10 min. Different restriction sites were included in the primers used to amplify the same fragment, in this way generating C-terminal and N-terminal His_6tag fusions for some constructs.

Table 2.2. Primers used for genetic constructs.

Constructs	Sequence of primers[a]	Name of primer	Restriction enzymes	Strategy	Source
pET28a-GatD-His_6tag	Fw: 5´-GCGCCCATGGGACATGAATTGACTATTTATCATTTTATGTC-3´ Rv: 5'-GCGCCTCGAGACGAGATTTCTTCTGTCTATTTG – 3´	GatD-Fw GatD-Rv	Nco I Xho I	Cloning	This study
pET28a- His_6tag-GatD	Fw: 5´-GCGCCATATGCATGAATTGACTATTTATC-3 Rv: 5´-GCGCGCGGCCGCTTAACGAGATTTCTTCTGTC-3	His-GatD-Fw His-GatD-Rv	Nde I Not I	Cloning	Figueiredo *et al.* (unpublished)
pET28a-DUF1727-His_6tag	Fw: 5´-CGCGCCATGGCTTTAGCTAAAAATCC – 3´ Rv: 5´-GCGCGTCGACTGATTGACCTCCTTCAAACGAA- 3´	DUF-His-Fw MurT- Rv	Nco I Sal I	Cloning	This study

continues

Constructs	Sequence of primers[a]	Name of primer	Restriction enzymes	Strategy	Source
pET28a- His$_6$tag-DUF1727	Fw - 5´ - GCGCGCGCATATGTTAGCTAAAA ATCCTGCAGG - 3´ Rv - 5´ - GCGTATACGCGGCCGCTTATGAT TGACCTCCTTCAAACG - 3´	His-DUF-Fw His-MurT-Rv	Nde I Not I	Cloning	This study
pET28a-DUF1727-GatD-His$_6$tag	Fw: 5´-CGCGCCATGGCTTTAGCTAAAAA TCC – 3´ Rv: 5'-GCGCCTCGAGACGAGATTTCTTC TGTCTATTTG – 3´	DUF-His- Fw GatD-Rv	Nco I Xho I	Cloning	Lobo *et al.*,(unpublished)
pET28a-MurT-His$_6$tag	Fw: 5'-CATGCCATGGGAAGACAGTGGA CGGCAATC -3' Rv: 5´-GCGCGTCGACTGATTGACCTCCT TCAAACGAA- 3´	MurT-Fw MurT-Rv	Nco I Sal I	Cloning	Lobo *et al.*,(unpublished)
pET28a-MurT-GatD-His$_6$tag	Fw: 5´-CATGCCATGGGAAGACAGTGGA CGGCAATC-3´ Rv: 5´-GCGCCTCGAGACGAGATTTCTTC TGTCTATTTG-3	MurT-Fw GatD-Rv	Nco I Xho I	Cloning	Figueiredo *et al.* (unpublished)
pET28a-MurTwDUF1727-GatD-His$_6$tag	Fw: 5' CGAAAAGAAGCGATGATCAAT**TG A**GCTAAAAATCCTG – 3' Rv: 5´-GCATTCATTCCTGCAGGATTTTT AGC**TCA**ATTGATCA- 3'	murTwDUF-Fw murTwDUF-Rv		Site-directed mutagenesis	This study
pET28a-MurTwDUF1727-GatD-His$_6$tag	Fw: 5´-CGAAAAGAAGCGATGATCAAT**TG A**-3´ Rv: 5´-GCTAGTTATTGCTCAGCGG-3´	scmurTwDUF-Fw T7 reverse		Screening for mutagenesis	This study

Notes: [a] restriction sites are underlined; [b] codon with nucleotide exchange in bold.

Subsequently, pET28a and the PCR products were digested overnight by the respective restriction enzymes (Thermo Scientific) listed in Table 2.2, ligated by T4 DNA ligase (NZYTech) over 16 hours, as recommended by the manufacturers, and transformed into *E. coli* DH5α competent cells.

Positive clones were screened by colony PCR using the cloning primers and the following conditions: a 95ºC step of 2 min followed by 30 cycles, each consisting of 95ºC for 30 s, 50ºC for 30 s, 72ºC for 2 min, and one final extension step of 72ºC for 10 min. The PCR products were visualized using agarose gel electrophoresis (1%, TAE 1X) and sequenced at StabVida (Portugal). Sequence analysis was performed using Finch TV (Geospiza, PerkinElmer, WA, USA). Finally, the correct plasmids were transformed into the *E. coli* expression strains, listed in Table 2.1.

2.4. Site-directed mutagenesis

Plasmid pET28a-MurT-GatD was used as template for PCR amplification using the primer pair murTwDUF-Fw and murTwDUF-Rv, *Phusion* DNA polymerase and 3% DMSO (Thermo Scientific). The amplification conditions used were as follows: a 98ºC step of 1.30 min, followed by 30 cycles, each consisting of 98ºC for 10 s, 73ºC for 45 s, 72ºC for 7 min 50 s, and one final extension step of 72ºC for 12 min. Then the PCR product was digested for 5 h with DpnI (New England Biolabs, New England, USA), dialyzed against water and finally transformed to *E. coli* DH5α. Positive clones were screened by PCR using the scmurTwDUF-Fw and T7 reverse primers, and sequenced as described above. Then, they were transformed into the *E. coli* expression strains, listed in Table 2.1.

2.5. Protein expression and purification

In these assays the proteins were expressed using two different protocols, according to the objective. To obtain high amounts of protein, expression was performed in large-scale, using 500 ml to 2 L of bacterial culture, while to test the expression conditions, the assays were performed in small-scale, using 25 to 200 ml of bacterial culture. According to the protocol used, different methods of cell disruption were adopted, either mechanical disruption or chemical lysis for large-scale and small-scale, respectively.

2.5.1. Expression of GatD protein in LB medium for 1D ^{1}H NMR analysis

GatD protein was expressed in large-scale in *E. coli* BL21-CodonPlus(DE3)-RIPL (Figueiredo *et al*., unpublished) fused to a N-terminal or C-terminal poly-histidine-tag (His_6tag) and in *E. coli* BL21(DE3) fused only to a N-terminal His_6tag, using the following conditions: cells transformed with the appropriate recombinant plasmid were grown in LB supplemented with antibiotics (30 µg/ml of Km and 50 µg/ml of Cm or only 30 µg/ml of Km, according to the strain used) at 37ºC. At an OD_{600nm} of 0.8, IPTG (NZYTech) was added at a concentration of 1 mM to induce expression of the recombinant proteins. After 3 h of incubation at 30ºC, cells were harvested (6940 rpm for 15 min) (Sorvall RC-5C

Plus,SLA-150 rotor, Kendro Laboratory Products Newtown, USA) and resuspended in lysis buffer (50 mM Tris-HCl, pH 7.5, 500 mM NaCl, 1 mM PMSF (ROTH, Germany), 60 µg/ml of DNase, 0.2% tween-20). After cell disruption performed with a French Press (FA-032 (40k) standard cell at 12000 psi from Thermo Electron Corporation) and removal of cellular debris and membranes by centrifugation (12000 rpm for 1 h), (Sorvall RC-5C Plus, SS34 rotor, Kendro Laboratory Products Newtown, U.S.A) the lysate was subsequently purified as described in section 2.5.3. An aliquot was collected before and after the final centrifugation for further analysis.

2.5.2. Expression of GatD protein in MM for 2D ^{1}H-^{15}N NMR

Several strategies aimed at optimizing the expression and the solubility of GatD when the expression strain was grown in MM medium: i) different temperatures (37ºC, 30ºC and 25ºC); ii) IPTG concentrations (0.5 and 1 mM); iii) induction time (OD_{600nm} of 0.6 and 0.8); iv) induction periods (3 h, 4 h, 5 h and overnight); v) pH of the lysis buffer (pH=8 and 7); vi) host strains (*E. coli* BL21-CodonPlus(DE3)-RIPL, BL21(DE3), Rosetta(DE3) pLysS, Rosetta(DE3)); vii) addition of detergents (Triton X100 (SIGMA), Tween-20).

Besides these alterations, expression was performed as previously described for LB medium in section 2.5.1. Cell lysis was performed by enzymatic digestion with lysozyme (SIGMA) instead of disruption with a French Press, as smaller volumes of cell culture were used (under 200 ml). Harvested cells were resuspended in 10 mM Tris-HCl, pH=8, 500 mM NaCl, 1 mM EDTA and 1 mg/ml the lysozyme (stock: 10 mg/ml in 10 mM Tris-HCL, pH=8) and incubated for 30 min at 37ºC. Freeze-thaw cycles were subsequently performed to enhance cell lysis (10 min at -20ºC, 5 min at 37ºC, 5 min a -20ºC and finally 5 min at 37ºC, thoroughly mixing between all steps). To separate the soluble and insoluble compounds after disruption cells, namely, cells not disrupted or cells membranes of the cells disrupted, the cells were harvested (12000 rpm for 1 h) (Sorvall RC-5C Plus, SS34 rotor), then the clarified lysate was purified as previously described in section 2.5.3.

2.5.3. GatD purification after expression in LB and MM

Three chromatography approaches were used to purify the recombinant GatD His-tag fusion proteins (His_6tag-GatD and GatD-His_6tag): affinity chromatography, anion exchange chromatography and molecular exclusion chromatography.

Affinity chromatography

For the small-scale assays the lysate was purified using manual Ni-NTA columns (Mo Bi Tec, Germany) and for the large-scale assays, purification was performed using automatic Ni-NTA columns (HiTrap IMAC FF 5 ml, GE Healthcare, Life Sciences, USA) and system AKTA pure (GE Healthcare, Life Sciences).

Manual Ni-NTA column

The column was charged with Ni-NTA agarose (Qiagen, USA) and equilibrated with water and extraction buffer (20 mM Tris-HCL, pH=8.2; 500 mM NaCl). The lysate was supplemented with 10 mM of imidazole, loaded onto the column and an aliquot was collected (flow-through). Initially, extraction buffer with 10 mM of imidazole was used for the washing steps and elution of GatD was performed with the same buffer with 200 mM of imidazole. Later, an imidazole step gradient (10, 50, 70, 90, 100, 150, 200, 300 mM of imidazole) was adopted for the washing and elution steps. Finally, the column was washed with water and 30% ethanol and stored at 4ºC. Aliquots were collected at each step.

Automatic Ni-NTA column

Before installing the column, the AKTA pure system was equilibrated at a with maximum pressure at 0.3 MPa and flow rate at 10 ml/min, in three steps: water, followed by a mixture of 50% of buffer A and 50% buffer B to equilibrate both pump lines (buffer A- 50 mM Tris-HCl pH=7.5, 0.5 mM NaCl, 30 mM imidazole; buffer B- 50 mM Tris-HCL, pH=7.5, 0.5 mM NaCl, 500 mM Imidazole, previously degassed) and finally 100% of buffer A. The column (HiTrap IMAC FF-5 ml) was loaded with 2.5 ml of 0.1 M NiO_4S, equilibrated with water and loaded with the lysate supplemented with 30 mM imidazole. The column was washed with 100% of buffer A and 4% of buffer B and then the imidazole concentration gradient was applied (0-100% of buffer B in 35 ml) at a flow rate of 2 ml/min. Finally, the system was washed with water and 20% ethanol, and the column was washed only with water and stored at 4ºC. Aliquots were collected at each step.

Anion exchange chromatography

The AKTA pure system was used with the same settings and was equilibrated as before. The only change was the composition of the buffers used (buffer A- 20 mM Tris-HCl pH=8.2, 100 mM NaCl, 10 mM $MgCl_2$; buffer B- 20 mM Tris-HCl, pH=8.2, 1 M NaCl, 10 mM $MgCl_2$, previously degassed). The column (HiTrap Q FF-5 ml from GE Healthcare, Life Sciences) was equilibrated with water and 100% of buffer A at a flow rate of 2 ml/min. The lysate was loaded and the column was washed with 100% of buffer A. Then, a NaCl concentration gradient was applied (0-50% of buffer B in 30 ml) followed by a step of 100% of buffer B. At the end the system and column were washed with water and 20% ethanol, and the column stored at 4ºC. Aliquots were collected at each step.

Molecular exclusion chromatography

The AKTA pure system and the column (Superdex 75 10/300 GL from GE Healthcare) were equilibrated at a maximum pressure at 1.10 MPa and flow rate at 0.1 ml/min with water. Then, the flow rate was set to 0.4ml/min and the system and the column were equilibrated with 100% of buffer C (20 mM Tris-HCl pH=7.4, 200 mM NaCl, 1 mM DTT, previously degassed). Subsequently, the lysate was loaded and proteins were eluted according their molecular size with 100% of buffer C. The system and the column were washed with water and then 20% ethanol. Aliquots were collected at each step.

2.5.4. Expression of MurT protein in LB for 1D ^{1}H NMR analysis

The MurT-His_6tag protein was expressed in small-scale in *E. coli* BL21-Codonplus(DE3)-RIPL and BL21(DE3) using the following conditions: cells transformed with pET28a-MurT-His_6tag were grown in LB supplemented with antibiotics (30 µg/ml of Km and 50 µg/ml of Cm or only 30 µg/ml of Km, according to the strain used) at 37ºC. Expression was induced at an OD_{600nm} of 0.6 by addition of 1 mM IPTG. The expression assay was repeated for different incubation temperatures (25, 30, and 37ºC). After 3 h, cells were harvested (6940 rpm for 15 min) (Sorvall RC-5C Plus, SLA-1500 rotor) and resuspended in lysis buffer (20 mM Tris-HCl, pH 8.2, 500 mM NaCl, 10 mM $MgCl_2$, 1 mM PMSF and 60 µg/ml of DNase). Cell lysis was performed by enzymatic digestion, as previously described in section 2.5.2.

2.5.5. Expression of DUF1727 domain in LB for 1D ^{1}H NMR analysis

The DUF1727-His_6tag and His_6tag-DUF1727 proteins were expressed in small-scale in different conditions for optimization purposes: i) different temperatures (37ºC, 30ºC and 25ºC); ii) induction times (OD_{600nm} of 0.6 and 0.8); iii) induction periods (3h, 4h and 5h); v) host strains (*E. coli* BL21-CodonPlus(DE3)-RIPL, BL21(DE3), Rosetta(DE3)).

Besides these alterations, expression was performed as previously described in section 2.5.1. Cell lysis was performed by enzymatic digestion and purification was performed by manual affinity chromatography as previously described in sections 2.5.2 and 2.5.3.

The DUF1727-His_6tag and His_6tag-DUF1727 proteins were also expressed in large-scale in *E. coli* BL21(DE3) in the following conditions: cells were gown in LB at 37ºC with 30 µg/ml of Km. Expression was induced at an OD_{600nm} of 0.8 by addiction of 1 mM IPTG at 30ºC for 3 h. Disruption of the cells was performed as previously described in section 2.5.1 and purification was performed by automatic and manual affinity chromatography as described above in section 2.5.3.

2.5.5.1. Removal of the His_6tag from His_6tag-DUF1727 recombinant protein

After expression and purification of His_6tag-DUF1727 in BL21(DE3), buffer was exchanged to 50 mM Tris-HCl pH=8, 10 mM $CaCl_2$, 200 mM NaCl, and a total of 1 mg/ml of protein was treated with the Thrombin CleanCleave^{TM} Kit (Sigma-Aldrich, USA) to remove the His_6tag, as described by the manufacturer. Briefly, an aliquot of 100 µl of Thrombin-Agarose (50% suspension in 50% glucose and 20 mM Tris-HCl, pH=8.2) was centrifuged at 500 x g and washed with 1x and 10x Cleavage Buffer (50 mM Tris-HCl, pH=8, 10 mM $CaCl_2$). Then, 1 mg of fusion protein was added to the resin and the cleavage mixture was incubated at room temperature with gentle agitation. Two aliquots were taken at several discrete time points, one was centrifuged at 500x g to remove the resin and the samples were all analyzed by SDS-PAGE.

2.5.6. Growth of BL21(DE3)-DUF1727-His$_6$tag in LB and induction in MM for 2D ^{1}H-^{15}N NMR analysis

The DUF1727-His$_6$tag domain was expressed in large-scale in *E. coli* BL21(DE3) using the following conditions: cells were grown in LB supplemented with antibiotics (30 µg/ml of km) at 37ºC. At OD_{600nm} of 1 the cells were harvested (2400 rpm for 20 min) (Sorvall RC-5C Plus, SLA-1500 rotor) and were resuspended in MM supplemented with 18.6 µg/ml $^{15}NH_4Cl$ (Cambridge Isotope Laboratories, USA) and without antibiotics (Serber *et al.*, 2001). IPTG was added at a concentration of 1 mM at 30ºC and after 3.5 h, cells were harvested (6490 rpm for 1 h) (Sorvall RC-5C Plus, SLA-1500 rotor) and resuspended in lysis buffer (20 mM Tris-HCl, pH=8.2, 500 mM NaCl, 1 mM PMSF, 0.06 mg/ml of DNase, 0.2% Tween-20). Cell disruption was performed with a French Press, as described in 2.5.1, and the lysate was subsequently purified by automatic affinity chromatography, as previously described in 2.5.3.

2.5.7. Protein co-purification assays

i) The MurT protein without DUF1727 domain and GatD protein with a C-terminal-His$_6$tag were expressed from strain BL21-CodonPlus(DE3)-RIPL-MurTWDUF1727-GatD-His$_6$tag using the conditions described in section 2.5.1 and were purified through automatic affinity chromatography as described in section 2.5.3.

ii) The DUF1727 domain and GatD protein with a C-terminal-His$_6$tag were expressed from BL21-CodonPlus(DE3)-RIPL-DUF1727-GatD-His$_6$tag using the conditions described in section 2.5.1. Then, were purified through manual affinity chromatography as described in section 2.5.3 with the following buffers: buffer 1: 50 mM Tris-HCl, pH=7.5, 500 mM NaCl, 10 mM Imidazole; buffer 2: 50 mM Tris-HCl, pH=7.5, 1.5 M NaCl; buffer 3: 50 mM Tris-HCl, pH=7.5, 500 mM NaCl, 500 mM Imidazole.

2.6. Protein analysis by SDS-PAGE

To verify the protein purity level, the protein samples were analyzed under denaturing conditions using SDS-polyacrylamide gel electrophoresis (PAGE), with the mini-PROTEAN system (BIO-RAD, USA). Before electrophoresis, the samples were mixed in a ratio of 1:2 with 19:1 (Laemmli: beta-mercaptoethanol) solution, incubated at 95ºC for 7 min and then 5 min on ice. Electrophoresis was performed in running buffer (24 mM Tris-base, 191 mM glycine, 3.46 mM SDS) at 30 mA for 1-1.5 h. GatD and MurT proteins were analyzed in 12.5% SDS-PAGE with Low Molecular Weight Protein marker (NZYTech) whereas DUF1727 domain and DUF1727-GatD complex were analyzed in 18% SDS-PAGE with Novex® Sharp Unstained Protein Standard (Life Technologies, USA) or ColorBurstA™ Electrophoresis Marker (SIGMA). Staining was performed with a solution of 0.3% coomassie brilliant blue R-250 (AppliChem, Germany) and destaining with a solution of ethanol/acetic acid (30%/10%).

2.7. Desalting, buffer exchange and protein concentration

Desalting and buffer exchange were performed using PD-10 columns (GE Healthcare). Buffer was exchanged routinely after affinity chromatography for a second chromatographic step or only to remove imidazole. After desalting or buffer exchange steps, the samples were concentrated as needed and this was performed using Amicon® Ultra-4 (10 k, 30 k) Centrifugal Filter Units (Merck Millipore, USA). Dialysis was achieved using Tube-O-Dialyzer^{TM} (G-Biosciences, USA) with cellulose membrane (4 kDa) and was used to exchange Tris-HCl for Tris-deuterium (Tris-DCl) for further NMR analysis.

2.8. Protein quantification

The total amount of protein present in each fraction collected was estimated using 2 methods: UV absorption at 280 nm (NanoDrop ND-1000, Fisher Scientific, Madrid), using extinction coefficients and protein molecular weights calculated with the online tools ProtParam and Compute pI/MW (ExPASy, Bioinformatics Resource Portal) specific for the target proteins (Table 5.1., annex) and bicinchoninic acid (BCA) protein assay reagent (Thermo Scientific). Quantification was performed after each chromatography, buffer exchange and concentration step.

2.9. Solvent condition screening

Aliquots of 1 ml of several buffers (100 mM concentration) were pipetted into each reservoir of a Linbro 24 well tissue culture plate. Two drops of 1 µl and 2 µl of protein solution in starting buffer (10 mM sodium phosphate at pH=7.4) were pipetted onto each glass cover slip. To each drop of protein solution was added 1 µl of 100 mM reservoir buffer, homogenized by pipetting and then, the glass slips were inverted and sealed onto the wells using silicone. The plate was allowed to rest undisturbed at room temperature so that vapor diffusion could take place (Lepre and Moore, 1998). The amount of precipitate in the drops was estimated by placing the tray against a black background and observing each drop under an optical microscope A total of 0.14 mg of GatD protein was used and the buffers used are listed in Table 2.3.

Table 2.3. Buffers and pH screening for GatD protein.

Drop	Buffer, pH	Drop	Buffer, pH
1	sodium acetate, 4.6	**7**	Tris, 7.5
2	sodium acetate 6.4	**8**	Tris, 5.5
3	sodium phosphate 5.5	**9**	potassium phosphate, 5.8
4	sodium phosphate 7.4[a]	**10**	potassium phosphate 6.8
5	Hepes 6.6	**11**	sodium citrate, 7
6	Hepes 5.8	**12**	sodium citrate, 4.5

[a]-Starting buffer

1.1.2

2.10. Western Blotting

Samples were separated in 18% SDS-PAGE along with molecular weight marker (NZYColour Protein Marker II, Nzytech). Protein was transferred to a nitrocellulose membrane (Amersham Hybond ECL Nitrocellulose, 0.45 µm from GE Healthcare) using the Mini Trans-blot electrophoretic transfer cell (BIO-RAD), filter papers and fiber pads all submerged in transfer solution (24 mM Tris-base, 191 mM Glycine, 10% Methanol). Blotting was performed at 4ºC for 90 min at 100 V with agitation. After blotting, the membrane was immersed overnight in blocking solution with PBS-Tween and milk, washed with PBS-Tween and probed with the primary antibody (anti-GatD or anti-His_6tag) in a ratio of 1:500 and 1:1000 respectively for 2.5 h. The membrane was then washed, immersed in fresh blocking buffer and incubated with secondary antibody (anti-rabbit or anti-mouse) (Perkin Elmer, USA) conjugated to alkaline phosphatase in a ratio of 1:5000 for 1 h. After a final washing step, the membrane was incubated with chemiluminescence detection solution for 1 min (Pierce ECL 2 Western Blotting kit, Thermo Scientific) and exposed to Amersham Hyperfilm^{TM} ECL autoradiographic film (GE Healthcare) in a cassette (Life Science) for appropriate periods of time. The film was processed by immersion in developing and fixing reagents.

2.11. Experimental section for NMR

NMR experiments were performed at 298 K in a Bruker Avance II+ 600-MHz spectrometer equipped with 5-mm TCI cryoprobe. Proton chemical shifts were referenced against external DSS while nitrogen chemical shifts were referenced indirectly to DSS using the absolute frequency ratio. Data were processed using the Bruker Topspin 3.1 package.

2.11.1. NMR for His_6tag-GatD

An initial 0.17 mM solution of GatD not labeled in 20 mM Tris-DCl buffer (10% ^2H_2O, pH=7.5) 200 mM NaCl. One ^1H-spectrum was acquired. Spectra were recorded as matrices of 1024 x 256 complex data points with spectral widths of 7945 Hz (^1H) using 8 scans per t_1 increment.

2.11.2. NMR for DUF1727-His_6tag

An initial 0.21 mM solution of ^{15}N-labeled DUF in 20 mM Tris-DCl buffer (10% ^2H_2O, pH=7.7) 200 mM NaCl, was titrated individually with 2 µL of a diluted solution of DCl and the pH was measured after each addition. One ^1H-^{15}N HSQC spectrum was acquired for each pH (7.7, 6.7, 6.0 and 5.0). Spectra were recorded as matrices of 1024 x 256 complex data points with spectral widths of 7945 Hz (^1H) and 1845 Hz (^{15}N) using 8 scans per t_1 increment.

Note: All the common reagents were the mark SIGMA.

Results

The main objective of this Master Thesis was to characterize by NMR, the physical interaction between the two proteins responsible for the amidation of *S. aureus* peptidoglycan, GatD and MurT. Nuclear magnetic resonance (NMR) is one of the two methods that can be applied to the study of three-dimensional molecular structures of proteins at atomic resolution. It's the only method that allows the determination of the three-dimensional structure of molecules in the solution phase (Jeremy *et al.,* 2002.). NMR spectroscopy can be used for structure determination of proteins in the size range of 5 and 35 kDa (Yu, 1999)

Our primary strategy was to obtain a ^{1}H-^{15}N-spectrum for GatD protein and then determine the structure of the MurT-GatD interaction, by titrating the ^{1}H-^{15}N-labeled GatD with unlabeled MurT protein. This strategy would allow to identified the amino acid residues of GatD that interact with MurT partner by comparing the differences between these two spectra.

In order to structurally analyze proteins by NMR, high amounts (in the 100 µM range) of the protein of interest with a high purity degree are required. Thus, the initial experimental approach involved the optimization of the expression yield and purification procedures of GatD and MurT proteins using an *E. coli* heterologous expression system.

The process of protein expression and purification is case-specific as it depends on the expression level and solubility characteristics of the protein of interest, being in this way a complex methodology that requires optimization. Several *E. coli* expression host strains, with different features, in addition to a wide range of different fusion tags and purification strategies, have been developed with the aim to simplify the purification of recombinant proteins. The most common approach is to fuse the protein with a terminal Histidine-tag (His_6tag) and purify it by immobilized metal affinity chromatography (IMAC).

3.1. Construction of GatD and MurT recombinant proteins

The *S. aureus gatD*, *murT* and *murT-gatD* genes were cloned into pET28a plasmid, transformed into *E. coli* DH5α strain and then transformed into other *E. coli* expression strains (Figure 1.1). The recombinant proteins used in this study are described in Table 3.1. The expression of GatD protein has been previously optimized for expression in complex medium (Figueiredo *et al*., unpublished) and the selected conditions are described in the Materials and Methods section 2.5.1.

Table 3.1. Recombinant proteins used in this study.

Recombinant proteins	Description	Source
His_6tag-GatD	GatD recombinant protein fused to a N-terminal His_6tag.	Figueiredo *et al.* unpublished
GatD-His_6tag	GatD recombinant protein fused to a C-terminal His_6tag.	R. Lobo *et al.*, unpublished
MurT-His_6tag	MurT recombinant protein fused to a C-terminal His_6tag.	This study
MurT-GatD-His_6tag	MurT with no tags and GatD fused to a C-terminal His_6tag.	Figueiredo *et al.* unpublished

3.2. Purification of GatD recombinant protein

3.2.1. Purification of His_6tag-GatD protein using affinity chromatography (IMAC)

In immobilized metal affinity chromatography (IMAC) the protein is separated according to the interaction of the fused tag with the metal ions immobilized in the matrix of the column: the His_6tag has high affinity for the positively charged metal ions so, when the lysate is loaded into the Ni-NTA column, the target protein is retained in the matrix whereas the E. coli endogenous proteins with little or no affinity are not retained and are washed out in the flow-through or during the washing steps. To elute the protein that specifically binds to the matrix, free imidazole is added to the elution buffer, in excess amounts. Imidazole, a histidine analog with higher affinity for metal ions, is used as a competitive agent to displace the His_6tag from metal ions co-ordination freeing the His-tagged proteins (Schmitt *et al.*, 1993). A low-concentration of imidazole is applied to the washing steps that anticipate elution in order to help prevent nonspecific binding of endogenous proteins that have histidine clusters. The interaction can also be disrupted by decreasing the pH of the elution buffer to 4.5-5.3, however in this pH range the target protein may be misfolded.

Automatic affinity chromatography of His_6tag-GatD

In order to optimize the purification procedure of the target protein, automatic purification was performed as this approach allows to apply a more efficient and controlled imidazole gradient than in manual purification. The AKTA pure system (automatic protein purification system) and a Hitrap IMAC FF column were used. The HiTrap IMAC FF columns are prepacked with IMAC Sepharose 6 Fast Flow, with consists of 90 µm highly cross-linked agarose beads, associated with a covalently

immobilized chelating group. The resin must be charged with a suitable positively charged ion before being used (Ni^{2+}, Co^{2+}, Zn^{2+}, Cu^{2+}, Fe^{2+}) and stripping of the ions is performed by the addiction of 50 mM EDTA (chelating reagent). In this assay the ion chosen was Ni^{2+}.

To minimize nonspecific low affinity binding to the matrix, several strategies were adopted: a low concentration of imidazole (30 mM) was added to the lysate before being loaded into the column; a high ionic strength (500 mM NaCl) was maintained along the imidazole gradient; a nonionic detergent, Tween 20 (0.2%) was added to the gradient to minimize the protein-matrix interaction. All these conditions did not affect the specific binding of His_6tag-GatD to the matrix. For the elution step an imidazole concentration gradient of 30-500 mM was applied and the results of this chromatography are presented in Figures 3.1.A-3.1.B.

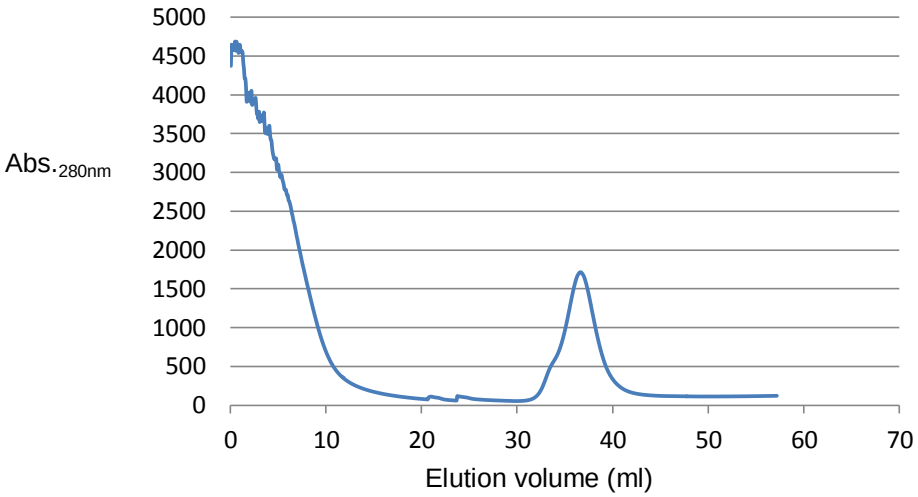

Figure 3.1.A: Chromatogram of His_6tag-GatD purification using a HiTrap IMAC FF-5 ml column. An imidazole gradient (30-500 mM) was applied. 0-10 ml: washing steps with 30 mM imidazole; 10-22 ml: washing steps with 48.8 mM imidazole; 22-57 ml: imidazole concentration gradient from 30 mM to 500 mM. Protein was detected by monitoring the absorbance at 280 nm.

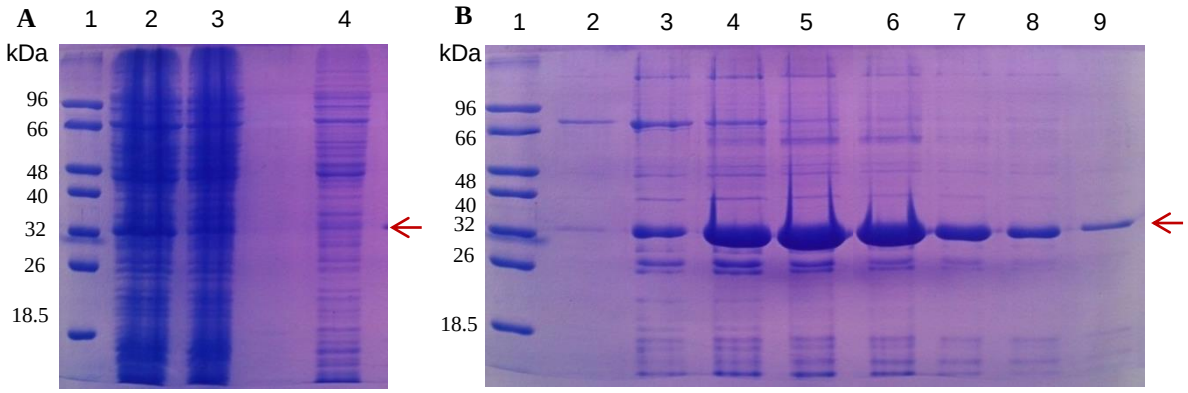

Figure 3.1.B- Expression and purification of His_6tag–GatD (29.09 kDa) from BL21(DE3)+pET28a-His_6tag-GatD grown in 750 ml LB and induced with 1 mM IPTG at OD_{600nm}=0.8 at 25ºC. Purification was performed through IMAC using AKTA pure system and HiTrap IMAC FF-5 ml column and an imidazole concentration gradient (30-500 mM). Fractions of 1.5 ml were collected. Panel **A**- Lane 1: protein marker (Low Molecular Weight Protein marker); Lane 2: soluble protein fraction; Lane 3: flow-through; Lane 4: washing fraction at 7.5 ml; Panel **B**- Lane 1: protein marker; Lane 2: washing fraction at 30 ml; Lanes 3-9: elution fractions from 35.5 ml to 44.5 ml (imidazole concentration from 211 and 330 mM). Red arrow: His_6tag-GatD.

The efficiency of the purification strategy was assessed by the analysis of some washing and elution fractions (chosen after analysis of the chromatogram in Figure 3.1.A), through 12.5% SDS-PAGE (Figure 3.1.B). Most *E. coli* endogenous proteins, without affinity for the nickel ions were washed out immediately from the matrix, in the flow-through fraction (Figure 3.1.B, panel A, lane 3). Elution of the target protein (His_6tag-GatD) started at 171 mM of imidazole and the maximum elution rate occurred for an imidazole concentration range of 220-272 mM (Figure 3.1.B, panel B, lanes 4-6).

The protein concentration of each fraction (Figure 3.1.B, panel B, lanes 4-9) was quantified by UV absorption or BCA, using the molecular weight and extinction coefficient of His_6tag-GatD (Table 5.1, annex). Overall, an approximate final yield of 13.3 mg/ml of the target protein was obtained for an initial volume of bacterial culture of 1 L, corresponding to a sample concentration approx. ***80*** µM of the GatD protein.

The elution fractions (Figure 3.1.B, panel B, lanes 3-8) include, beside His_6tag-GatD protein, many *E. coli* endogenous proteins**.** One possible drawback of His_6tag affinity purification is the co-purification, together with the target protein, of naturally histidine-rich host proteins that act as contaminants (Andersen *et al.*, 2013). The two most frequent *E. coli* contaminants are ArnA (MW=74.28 kDa), an enzyme involved in the modification of lipid A phosphate (Williams *et al.*, 2005) and SlyD (MW= 20.8 kDa), a peptidyl-prolyl cis/trans-isomerase (William *et al.*, 1994).

To obtain a ^{1}H-spectrum, the protein under analysis must have a high degree of purity; for this reason, a second purification step was performed.

3.2.2. Anion exchange chromatography of His_6tag-GatD protein

To further purify the His_6tag-GatD protein obtained from IMAC, a second purification step was performed, anion exchange chromatography, with the aim of reducing the amount of *E. coli* endogenous proteins which co-eluted with His_6tag-GatD in the affinity chromatography.

In an anion exchange chromatography, proteins are separated according to ionic interactions between the protein and the positively charged solid support. These interactions depend on the isoelectric point (pI) of each protein and the pH of buffers used. For a pH value similar to the pI of the protein of interest, the overall protein charge is neutral; for a pH value above the pI of the protein of interest, the overall protein charge is negative; for a pH value below the pI of the protein of interest, the overall protein charge is positive. Thus, a protein with a pI value below the pH of the buffer will be negatively charged and will interact with the matrix, being retained, while a protein with a pI value above the pH of the buffer will not interact with the matrix, being released in the flow-through or during the washing steps. As the surface charge of the protein becomes more negative, the stronger the interaction becomes.

As the pI of the target protein is 6.01, a buffer with a higher pH value (8.2) was used, decreasing His_6tag-GatD surface charge and allowing its interaction with the matrix.

The HiTrap Q FF column used was prepacked with Q Sepharose Fast Flow with 6% highly cross-linked beaded agarose matrix, a strong anion exchanger for small-scale protein purifications. A salt concentration gradient (100 mM to 1 M NaCl) was applied, resulting in a gradual increase of the ionic strength, with allows for protein release by competing with the matrix.

In order to appropriately bind all charged proteins to the matrix, all the buffers, including the sample must have the same pH and the sample must have the same ionic strength as the starting buffer. The results of this chromatography are presented in Figures 3.2.A-3.2.B.

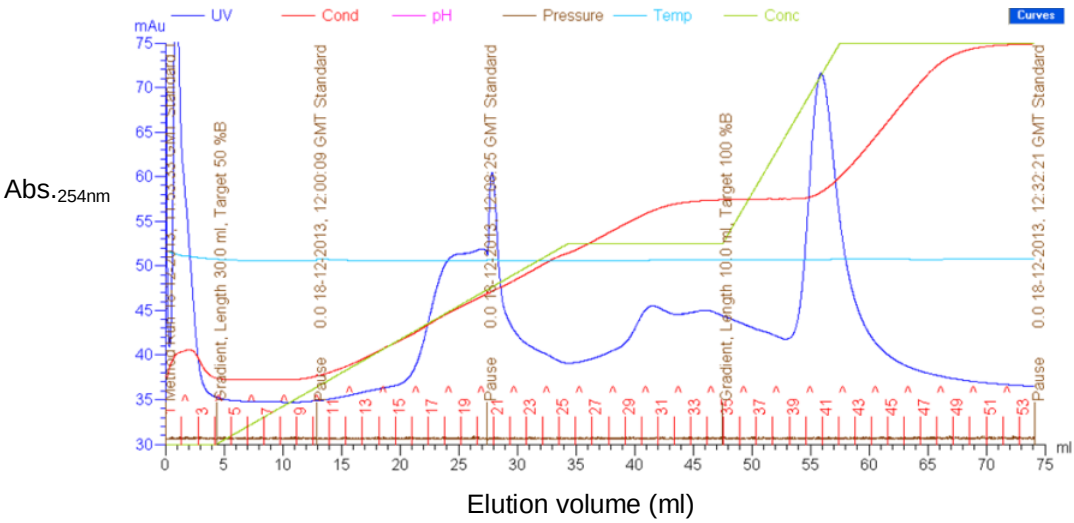

Figure 3.2.A: Chromatogram of His_6tag-GatD purification using HiTrap Q FF-5 ml column. A salt concentration gradient (100mM to 1 M NaCl) was applied. 0-5 ml: 100 mM NaCl; 5-75 ml: 100 mM to 1 mM NaCl. Protein was detected by monitoring the absorbance at 254 nm.

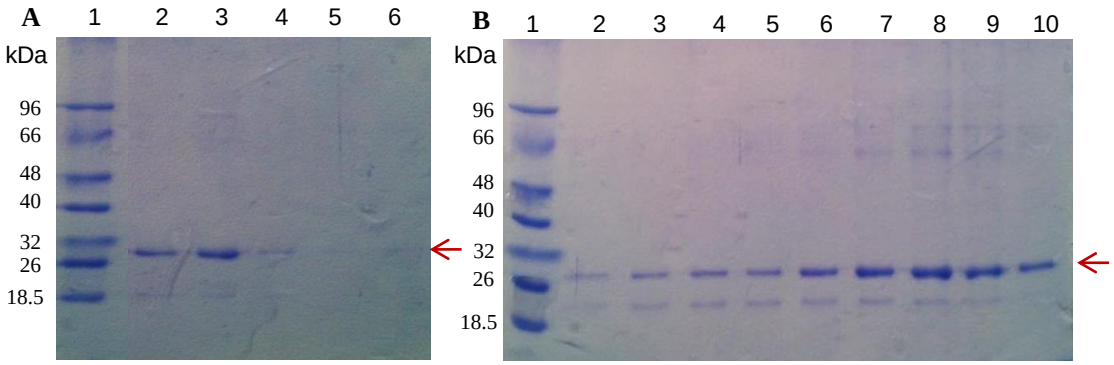

Figure 3.2.B- Purification of His_6tag-GatD was performed through anion exchange chromatography using AKTA pure system and HiTrap Q FF-5 ml and a salt concentration gradient from 100 mM to 1 M NaCl. Fractions of 1.5ml were collected. Panel **A**- Lane 1: protein marker (Low Molecular Weight Protein marker); Lanes 2-4: elution fractions from 1.5 to 4.5 ml; Lanes 5-6: washing fractions from 6 to 7.5 ml. Panel **B**- Lane 1: protein marker; Lanes 2-10: elution fractions from 28.5 to 40.5 ml. Red arrow: His_6tag-GatD.

In this assay, the protein sample and the starting buffer did not have the same ionic strength (500 mM and 100 mM respectively). This lapse affected the interaction between the protein and the matrix and some target protein was released during the first washing steps (Figure 3.2.B, panel A, lanes 2-4). Elution of the target protein occurred (Figure 3.2.A-3.2.B) at different NaCl concentrations (415-640 mM and 730 mM), although the maximum elution rate occurred for a salt concentration range of 550-570 mM (Figure 3.2.B, panel B, lanes 7-9).

The elution fractions (Figure 3.2.B, panel A, lanes 2-9) include, besides His_6tag-GatD protein, *E. coli* endogenous proteins**.** In fact, the purity degree of His_6tag-GatD was observed to decrease as the protein concentration of elution fraction increased. Because the salt composition of the protein sample affected the interaction, this assay was repeated with the correct sample ionic strength. In the second assay, elution of the target protein did not occur in the first washing steps, however, the overall results were analogous.

This chromatography did not allow a complete separation of the target protein from the *E. coli* endogenous proteins. A possible explanation would be the pI of the *E. coli* endogenous proteins being close to the one of the target protein (pI of His_6tag-GatD=6.01). In fact, the AnrA protein has a pI= 6.39 and SlyD protein has a pI=5.03.

Taken together, anion exchange chromatography resulted in the elution of His_6tag-GatD at different NaCl concentrations and consequently in the dilution of the target protein over several elution fractions. Furthermore, this method was not able to remove the contaminant proteins to completion, evidencing the low efficiency of this chromatography in the purification of His_6tag-GatD. Due to the low efficiency of anionic chromatography to further purity the His_6tag-GatD protein obtained from IMAC, an alternative second purification step was needed.

Because one of the major contaminants had a higher molecular weight (≃70 kDa), the use of molecular exclusion chromatography, emerged as an alternative.

3.2.3. Molecular exclusion chromatography of His_6tag-GatD protein

Purification of His_6tag-GatD by anion exchange chromatography resulted in the co-elution of several contaminants, with different molecular sizes (≃70, 60 and 20 kDa). To dispose of these unwanted contaminants, a molecular exclusion chromatography was performed as a second purification step, in order to increase the His_6tag-GatD purity degree.

In molecular exclusion chromatography, proteins are separated by the retention time in the matrix, according to their molecular size. Proteins with a molecular size larger than the pore of the matrix are released prior to the proteins with a smaller molecular size, which are retained in the pores of the matrix. The estimated MW of the target protein is 29.09 kDa.

Superdex 75 10/300 GL is a prepacked gel filtration column composed of cross-linked agarose and dextran for high-resolution separation of proteins and peptides according to their size. The separation range of this column includes molecules with molecular weights from 3 to 70 kDa. The maximum protein amount allowed per assay is 10 mg. Approximately 4 mg of the protein was applied and the proteins were eluted with elution buffer (20 mM Tris-HCl pH=7.4, 200 mM NaCl, 1 mM DTT). The results of this assay are presented in Figures 3.3.A-3.3.B.

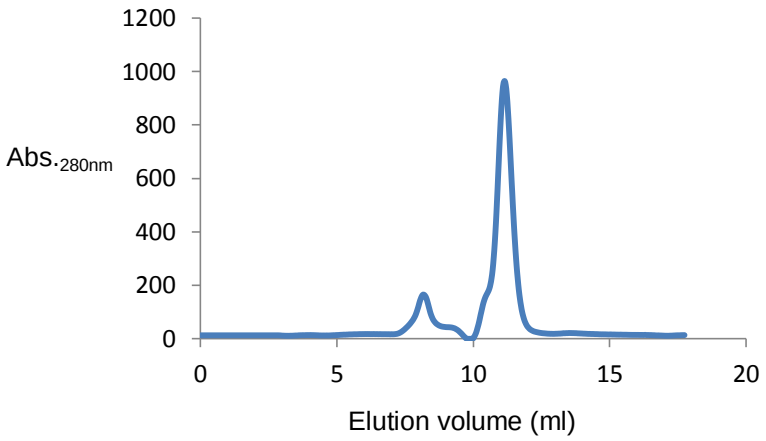

Figure 3.3.A: Chromatogram of His_6tag -GatD purification using Superdex 75 10/300 GL column. Protein was detected by monitoring the absorbance at 280 nm.

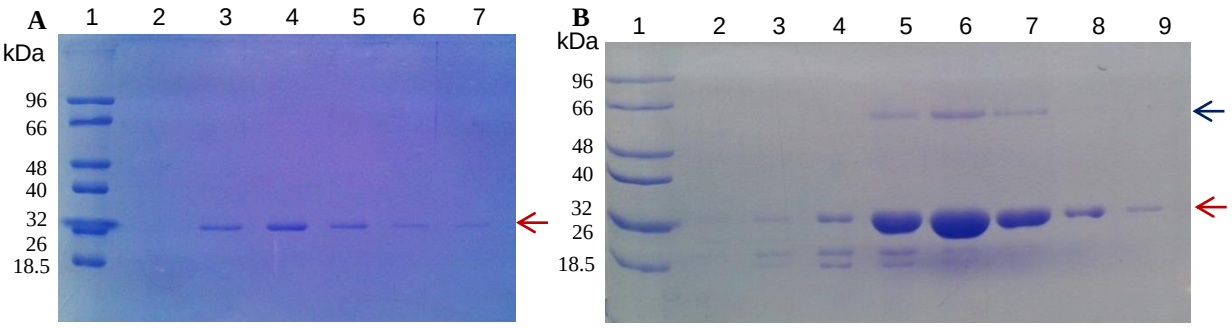

Figure 3.3.B: Purification of His_6tag-GatD was performed through molecular exclusion chromatography using AKTA pure system and Superdex 75 10/300 GL. Fractions of 0.5 ml were collected. Panel **A**- Lane 1: protein marker (Low Molecular Weight Protein marker); Lane 2: washing fractions at 6.5 ml; Lanes 3-7: eluted fractions from 7 to 9 ml; Panel **B**- Lane 1: protein marker; Lane 2-3: washing fractions from 9.5 to 10 ml; Lanes 4-9: elution fractions from 10.5 to 13 ml. Red arrow: His_6tag-GatD; Blue arrow: protein band with molecular size of approx.60 kDa.

This assay showed that His_6tag-GatD was eluted at two different elution volumes, 7-9 ml and 11-13.5 ml (Figure 3.3.A-3.3.B, panel A, lanes 3-5; panel B, lanes 4-9), most of the protein was eluted from 11.5 ml to 12.5 ml with high purity degree.

The first and the second eluted peaks (7-9 to 11-13.5 ml) correspond to the elution of proteins with a molecular size of 67-43 kDa and 43-13 kDa (Figure 5.2, annex), respectively. Unexpectedly, elution of the target protein occurred for both elution volumes, suggesting more than one conformation for His_6tag-GatD protein – as a monomer, with an estimated molecular size of 29.09 kDa and as a dimer, with an estimated molecular size of 58.18 kDa. If this was the case, the contaminant protein of higher molecular size could correspond to His_6tag-GatD dimer (Figure 3.3.B, panel B, lanes 5-7, blue arrow). To test this hypothesis, the quaternary structure of His_6tag-GatD was assessed by analyzing the protein under several unfolding conditions and by western blotting.

3.3. Assessment of the quaternary structure of His_6tag-GatD

3.3.1. Denaturing conditions applied to the His_6tag-GatD

Under complete denaturing conditions, proteins migrate through an electric filed according to their molecular weight. If a protein has not been completely denatured and spontaneously forms a dimeric structure, an extra band will appear in the SDS-PAGE, with the double value of its molecular weight. As the protein is completely denatured, this extra band will disappear. If the protein does not dimerize, no band alteration is expected to occur.

Routinely, after expression of His_6tag-GatD protein and purification by affinity chromatography, the elution samples were always analysed by SDS-PAGE. The denaturing conditions applied to the protein samples were altered in order to test several levels of protein unfolding. This also allowed to test if the conditions normally used result in full protein denaturation.

The protein sample was prepared in the default conditions (a ratio of 1:2 with 19:1 Laemmli: beta-mercaptoethanol solution, incubated at 95ºC for 7 min), and the following alterations were

independently tested: i) a longer incubation time (14 min), ii) several concentrations of the denaturing reagent DTT were tested, 0.05, 0.1 and 0.5 M. The results are presented in Figure 3.4.

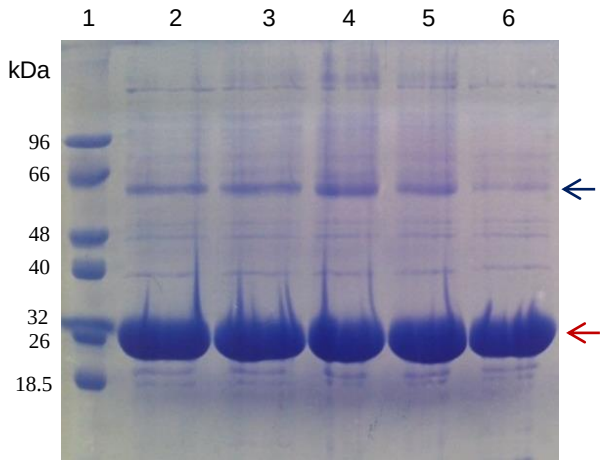

Figure 3.4- Alteration of the denaturing conditions of the His_6tag-GatD samples analysed by SDS-PAGE. Proteins samples used were from the elution step 3 after expression of His_6tag-GatD from BL21(DE3)+pET28a-His_6tag-GatD and purification by IMAC. Lane 1: molecular marker (Low molecular weight protein marker); Lanes 2-3: The samples were mixed in a ratio of 1:2 with 19:1 Laemmli: beta-mercaptoethanol solution, incubated at 95ºC for 7 and 14 min, respectively; Lane 4: The sample was mixed in a ratio of 1:2 with 19:1 Laemmli: 0.05 M DTT solution, incubated at 95ºC for 7 min; Lane 5: The sample was mixed in a ratio of 1:2 with 19:1 Laemmli:0.1 M DTT solution, incubated at 95ºC for 7 min; Lane 6: The sample was mixed a ratio of 1:2 with 19:1 Laemmli :0.5 M DTT solution, incubated at 95ºC for 7 min. Red arrow: His_6tag-GatD. Blue arrow: protein band with molecular size of approx. 60 kDa.

The results showed a decrease of the intensity of one protein band with an approximate molecular weight of 60 kDa, when the denaturing reagent was added, DTT at 0.5 M (Figure 3.4, lane 6, blue arrow). This suggests that the normal denaturing conditions used for SDS-PAGE do not completely denature the target protein, and the protein band of ≃60 kDa probably corresponds to the dimeric form of His_6tag-GatD. These results support the hypothesis of GatD protein spontaneously forming monomeric and dimeric arrangements.

3.3.2. Western blotting

Western blotting methodology detects specific proteins immobilized onto a membrane solid support, using specific antibodies (Ab) against the target protein. In order to further support the previously-raised hypothesis, that His_6tag-GatD quaternary structure includes monomeric and dimeric forms, western blotting for GatD (anti-GatD Ab) and His_6tag (anti-His_6tag Ab) was performed. The protein samples tested were elution fractions of the expression of His_6tag-GatD from BL21(DE3)+pET28a-His_6tag-GatD and purification through IMAC. The expression of pET28a from BL21-CodonPlus(DE3)-RIPL+pET28a was used as negative control. The results are presented in the Figure 3.5.

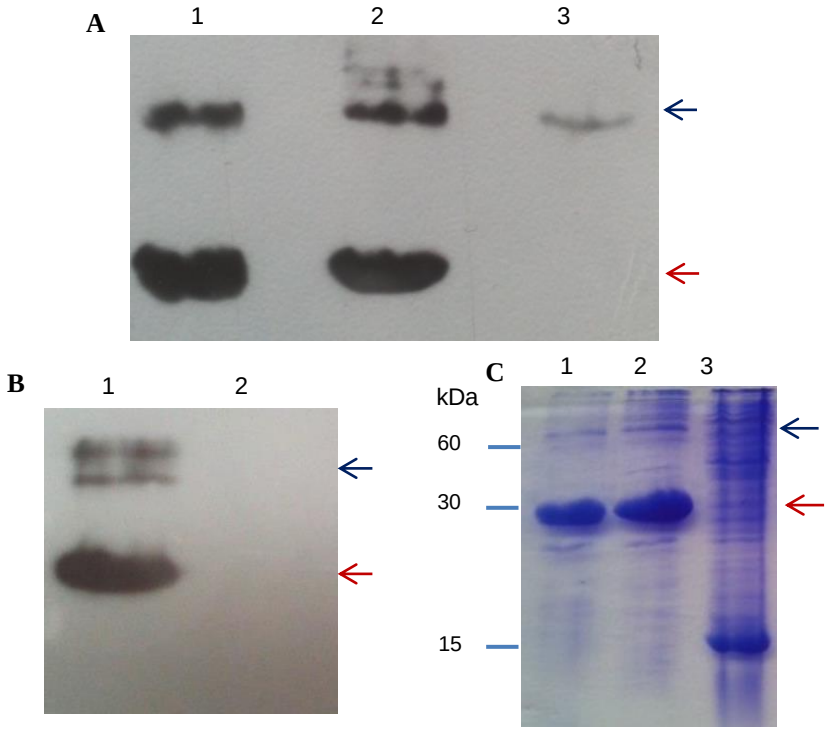

Figure 3.5- Western blotting performed for His_6tag-GatD expressed from BL21(DE3)+pET28a-His_6tag-GatD and purified by IMAC. Proteins extracts from BL21-CodonPlus(DE3)-RIPL+pET28a was used as a negative control. The samples were resolved in a 18% SDS-PAGE. Panel **A**- Western blotting against GatD protein: Lanes 1-2: elution steps 1 and 2 of His_6tag-GatD; Lane 3: soluble protein fraction of the negative control. Panel **B**- Western blotting against His_6tag: Lane 1: elution step 2 of His_6tag-GatD; Lane 2: soluble protein fraction of negative control. Panel **C-** 18% SDS-PAGE and Lanes 1-2: elution steps 1 and 2 of His_6tag-GatD; Lane 3: soluble protein fraction of the negative control. Red arrow: His_6tag-GatD; Blue arrow: protein band with molecular size of approx.60 kDa.

Western blotting using anti-GatD Ab resulted in hybridization signal in all the samples tested (Figure 3.5, panel A, lanes 1-3), including the negative control extract. For the GatD purification samples, two bands were detected, corresponding to a protein with approximately 60 kDa and to a protein with 29 kDa, corresponding to the expected molecular sizes for the dimer and monomer forms of GatD (Figure 3.5, panel A, lanes 1-2). For the negative control, only a protein with approximately 60 kDa was detected, with a low intensity signal, probably corresponding to an amidotransferase protein endogenous to *E. coli.*

Western blotting using anti-His_6tag Ab, only resulted in hybridization signal for the GatD purification samples and not in the negative control extract (Figure 3.5, panel B, lanes 1-2). Also, two bands with approximately 60 and 29 kDa were detected, corresponding to the expected molecular sizes for the dimer and monomer forms of GatD.

These results further supported the existence of more than one conformation for GatD protein: GatD monomer (≃29 kDa) and GatD dimer (≃58 kDa).

3.4. Stability test for recombinant GatD protein with N-terminal His_6tag

The optimization of the His_6tag-GatD purification process resulted in the adoption of an IMAC step followed by a size exclusion chromatography. However, the purified His_6tag-GatD obtained was observed to frequently precipitate from the solution at a high rate. In this way, optimization of the stability conditions was needed.

The His_6tag-GatD protein precipitated upon dialysis against H_2O, Tris-HCl or other buffers without NaCl. Thus, 100 mM or 200 mM NaCl concentration were added to all buffers used from this point on. However, despite the presence of NaCl, precipitation occurred as a result of protein concentration procedures or upon addition of specific compounds, namely 1 mM DMSO. To overcome this problem, screenings for the appropriate buffer conditions were performed (Table 2.3). A 1 µl and 2 µl sample of GatD protein in starting buffer (sodium phosphate at pH=7.4) at 10 mM was independently mixed with a 1µl drop of several test buffers at 100 mM and then equilibrated with a reservoir containing a higher test buffer concentration (100 mM). The buffer diffused out of the drop until its ionic strength was the same as the test buffer in the reservoir, in this way increasing the protein concentration in the drop. The GatD protein precipitated from solution according to its stability in each test buffer, precipitating first for the buffers in which it is less stable (Lepre and Moore, 1998). The amount of protein precipitation in each drop was visible after 7 days on the optical microscope. The best solvents for GatD protein were sodium acetate at pH=4.6, sodium phosphate at pH=5.5 and Tris-HCl at pH=7.5. In order to avoid buffer interference in the NMR analysis, available Tris-DCl at pH=7.5 with 200mM NaCl was used.

The stability of the protein in aqueous solution is associated with its solubility in water (polar solvent). The presence of hydrophobic regions on the surface of the protein stimulates precipitation, due to the poor interaction with water. Glycerol preferentially interacts with large patches of continuous hydrophobicity, acting as an amphiphilic interface between the hydrophobic surface and the polar solvent (Vagenende *et al.*, 2009). DTT is a strong reducing agent, frequently used to prevent intermolecular disulfide bonds of from forming between cysteine residues. The GatD protein has three cysteines, being the active site a cysteine at position 94. In order to avoid formation of disulfide bonds between molecules of GatD, DTT was tested.

To overcome the problem of precipitation in the concentration step, 1% and 5% glycerol and 1 mM DTT were added to the sample buffer, in order stabilize the His_6tag-GatD protein. After this optimization step, the protein was analysed by 1D ^{1}H-NMR.

3.5. 1D ^{1}H NMR of His_6tag-GatD

In structural biology it is necessary to efficiently screen for the presence of stable structures in proteins that can subjected to subsequent structure determination. The ^{1}H chemical distribution as detected by ^{1}H NMR spectroscopy can be an efficient tool to probe protein stability of proteins in solution. In a protein structure the individual residues are packed into chemical environments, which are very different from the random coil situation (chemical shift reference).

The spectrum of the unfolded protein corresponds to a spectrum, which in essence is the sum of the random coil spectra of the amino acid residues in the proteins. The dispersion of signals in the spectrum of the folded protein is far beyond the envelope of signals seen in the spectrum of the unfolded protein. This clearly reflects that nuclei in the folded form are subject to many different types of microenvironments of chemical screens.

Based on the data collected for well-structured proteins and proteins that exist in a molten globule state or a partially folded α-helical state, a threshold exists that can be used as a qualitative benchmark for protein structural stability (e.g., foldedness) in solution.

After optimization of purification and stability of the target protein, the His_6tag-GatD was expressed from BL21(DE3)+pET28a-His_6tag-GatD and purified using two automatic chromatographic steps, the IMAC followed by molecular exclusion. The elution fractions were concentrated with 5% glycerol, and dialyzed overnight to exchange Tris-HCl for Tris-DCl at pH=7.5. The sample, with a final concentration of 174 µM, was then analyzed by NMR.

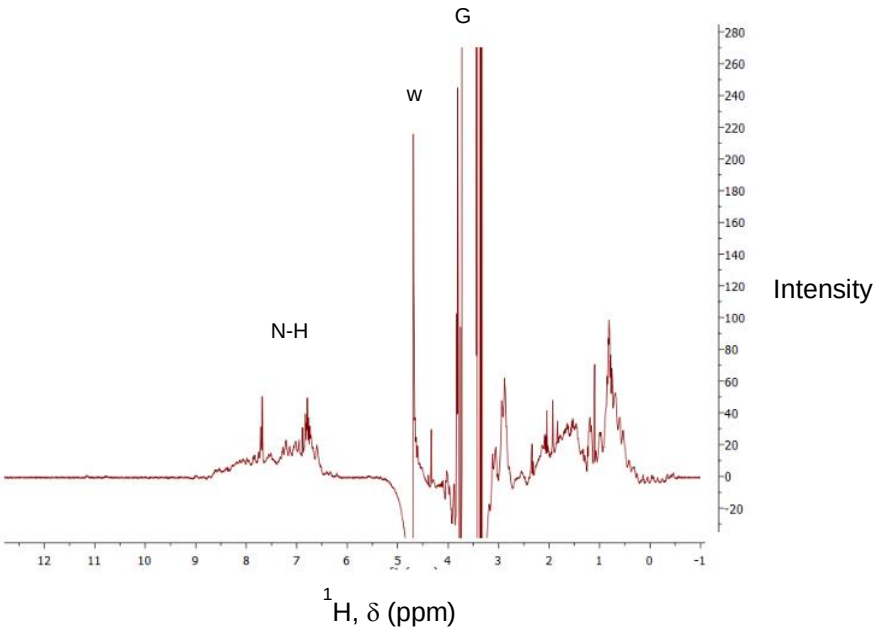

Figure 3.6- ^{1}H-spectrum 1D of His_6tag-GatD at 174 µM in 20 mM Tris-HCL pH=7.5, 200 mM NaCl, 1 mM DTT and 5% glycerol. His_6tag-GatD was expressed in 750 ml LB from BL21(DE3)+pET28a-GatD-His_6tag and purified through IMAC and molecular exclusion chromatography. N-H: N-H region; W-water; G- glycerol.

^{1}H proton signals of the target protein were detected in the range of -1-12 ppm (Figure 3.6) although the dispersion of signals in the N-H region (5-12 ppm) is not high, it is indicative of a folded protein. In oder to confirm this results, a ^{1}H-^{15}N-spectrum must be performed.

3.7. Expression of His_6tag-gatD protein in MM

The ^{1}H-spectrum of His_6tag-GatD protein indicated that the protein was correctly folded. With this information in hand, the expression and purification conditions were adapted with aim to obtain a ^{15}N-NMR spectrum of His_6tag-GatD.

To obtain the ^{1}H-spectrum, the target protein was expressed in complex medium (LB); however, in order to obtain ^{15}N and ^{13}C NMR spectra the recombinant protein must be expressed in minimal medium (MM), to ensure that the only sources of nitrogen and carbon are $^{15}NH_4Cl$ and ^{13}C-labeled glucose. Frequently, the expression yield of recombinant proteins is much lower for induction using minimal media. For this reason, the expression of His_6tag-GatD protein in MM was first optimized in small scale (100-200 ml).

In order to optimize the protein expression yield, different *E. coli* expression strains were tested. BL21-CodonPlus(DE3)-RIPL and Rosetta(DE3) were selected to express MurT and GatD because the sequence of both proteins includes codons which are rare in *E. coli,* such as the isoleucine codon (AUA), the arginine codon (AGA) and the glycine codon (GGA). BL21-CodonPlus(DE3)-RIPL and Rosetta(DE3) have some extra tRNA genes which may improve the expression of these proteins.

Additionally, different temperatures (37ºC, 30ºC and 25ºC), IPTG concentrations (0.5 and 1 mM), induction time (OD_{600nm}=0.6 and 0.8), induction periods (3 h, 4 h, 5 h and overnight), pH of the lysis buffer (pH= 8 and 7), host strains (BL21(DE3), BL21-CondonPlus(DE3)-RIPL, Rosetta (DE3)pLys, Rosetta(DE3)) and addition of detergents (Triton X-100) were tested. The results for different temperatures (30ºC and 25ºC), IPTG concentration (0.5 and 1mM) as also the strains tested (BL21(DE3), BL21-CondonPlus(DE3)-RIPL, Rosetta (DE3)pLys) are presented in Figures 3.7-3.8. The results for the others conditions are not shown.

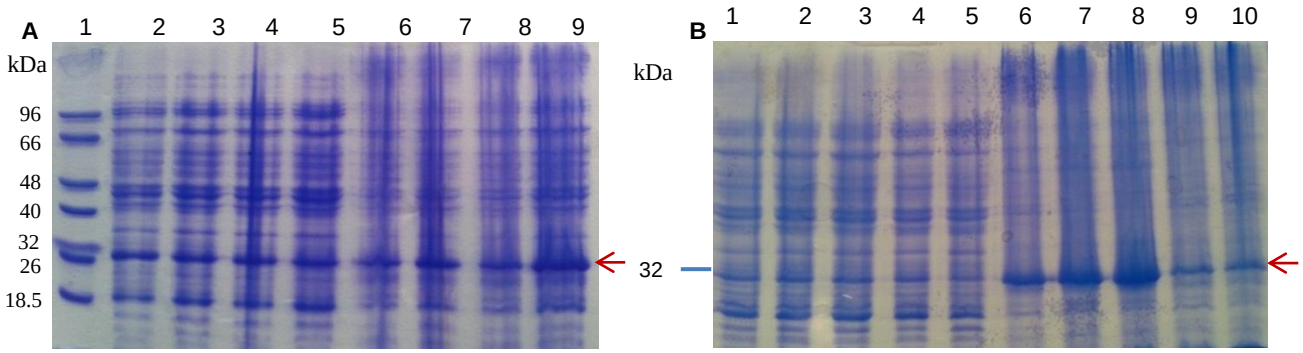

Figure 3.7- Optimization of the selected expression conditions of His_6tag-GatD from BL21-CodonPlus(DE3)-RIPL+pET28a-His_6tag-GatD grow in MM. Panel **A**. Lane 1: molecular marker (Low Molecular Weight Protein marker); Lanes 2-3: soluble protein fractions induced at OD_{600nm}=0.6 with 1mM IPTG for 3 h, at 25ºC and 30ºC, respectively; Lanes 4-5: soluble protein fractions induced at OD_{600nm}=0.8 with 1mM IPTG for 3 h, at 25ºC and 30ºC, respectively. Lanes 6-9: Insoluble protein fractions of the conditions described in lanes 2-5. Panel **B**- expression induced at OD_{600nm}=0.8. Lanes 1-2: soluble protein fractions induced at 30ºC with 0.5 mM and 1 mM IPTG for 3 h, respectively; Lane 3: soluble protein fraction induced at 30ºC with 1 mM IPTG overnight; Lane 4: soluble protein fraction induced at 20ºC with 1 mM IPTG for 3 h; Lane 5: soluble protein fraction expressed from Rosetta(DE3)pLysS His_6tag-GatD at 30ºC with 1 mM IPTG for 3 h; Lanes 6-10: Insoluble protein fractions of the conditions described in lanes 1-5. Red arrow: His_6tag-GatD.

The expression of the target protein was higher at 30ºC, when the culture was induced at an OD_{600nm}=0.8 with 1 mM IPTG overnight and then at 0.8 (Figure 3.7, panel B, lanes 3 and 8; panel A, lanes 5 and 9). However, in these conditions the solubility of His_6tag-GatD decreased, suggesting that higher expression levels result in loss of protein solubility. On the other hand, lower levels of the target protein were obtained for expression from Rosetta(DE3)pLysS and for expression at low temperature (20ºC) (Figure 3.7, panel B, lanes 4-5 and 9-10).

The expression levels of His_6tag-GatD in BL21(DE3)+pET28a and BL21-CodonPlus(DE3)-RIPL+pET28a were similar, however the last stain expressed an additional endogenous protein (Figure 3.8, lanes 3-4). In this way, the expression strain chosen was BL21(DE3) and the induction conditions were as follows: 30ºC, at with OD_{600nm}=0.8 for 3 h.

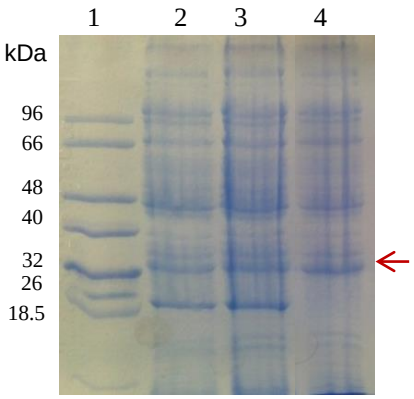

Figure 3.8- Optimization of expression conditions of His_6tag-GatD in MM induced at OD_{600nm}=0.8 with 1 mM IPTG. Lane 1: molecular marker (Low Molecular Weight Protein marker); Lanes 2-3: soluble protein fractions expressed from BL21-CodonPlus(DE3)-RIPL induced at 37ºC and 30ºC, respectively; Lane 4: soluble protein fraction expressed from BL21(DE3) and induced at 30ºC. Red arrow: His_6tag-GatD.

All the expression tests were performed in small scale, for low culture volumes (100-200 ml).

3.8. Expression and purification of His_6tag-GatD and GatD-His_6tag in BL21(DE3) in MM

After optimization of His_6tag-GatD induction conditions in MM, scale up of the expression process was performed for large-scale volume of culture (2 L). Expression was followed by purification by IMAC using AKTA pure system and Hitrap IMAC FF (Figure 3.9).

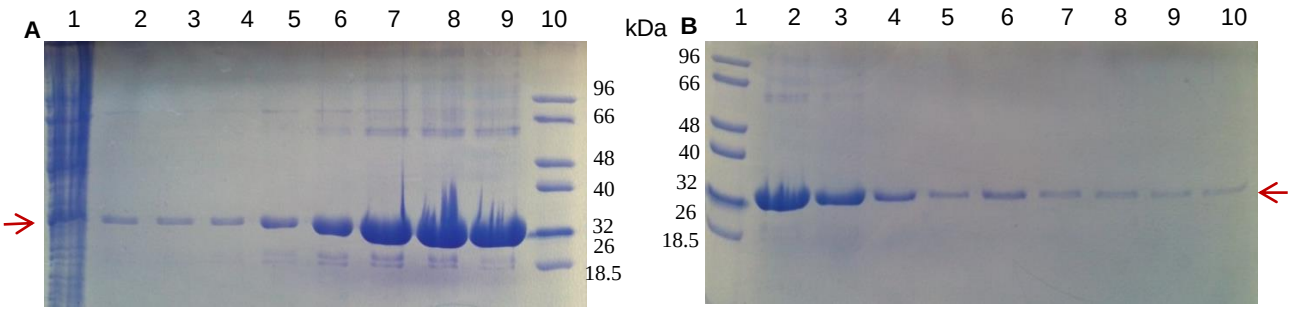

Figure 3.9- Expression and purification of His_6tag-GatD from BL21(DE3)+pET28a-His_6tag-GatD grown in 2 L MM and induced with 1 mM IPTG at OD_{600nm}=0.8 at 30ºC. Purification was performed through IMAC using AKTA pure system and HiTrap IMAC FF-5 ml column with an imidazole concentration gradient (30-500 mM). Fractions of 1.5 ml were collected. Panel **A**- Lane 1: soluble protein fraction; Lanes 2-9: elution fractions from 140-278 mM imidazole; Lane 10- molecular Marker (Low Molecular Weight Protein marker); Panel **B**- Lane 1: molecular marker; Lanes 2-10: elution fractions from 293-434 mM imidazole. Red arrow: His_6tag-GatD.

In this assay, the total amount of His_6tag-GatD obtained was approx. ***60*** µM, an acceptable value to proceed to NMR analysis. However, this yield was obtained using an initial MM culture volume of 2 L. In fact, the ratio between the amount of protein obtained and the volume of MM culture used in this assay was too low, preventing this strategy to be implemented.

Facing the insufficient low yield of His_6tag-GatD in MM, a new strategy was adopted: GatD-His_6tag was constructed and expressed. The results are presented above in Figure 3.10.

.

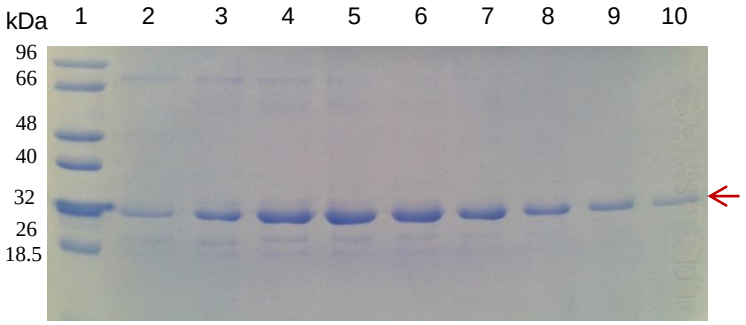

Figure 3.10: Expression of GatD-His_6tag from BL21-CodonPlus(DE3)-RIPL+pET28a-GatD-His_6tag grown in MM and induced with 1 mM IPTG at OD_{600nm}=0.8 at 30ºC. Purification was performed by IMAC using AKTA pure system and Hitrap IMAC FF-5 ml with an imidazole concentration gradient (30-500 mM). Fractions of 1.5 ml were collected. Lane 1: molecular marker (Low molecular weight protein marker); Lanes 2-10: Elution fractions from 150-287 mM imidazole. Red arrow: GatD-His_6tag.

Expression and purification of GatD-His_6tag showed no yield improvement in comparison to His_6tag-GatD.

3.9 Expression of MurT-His$_6$tag protein

3.9.1 Optimization of expression of MurT-His$_6$tag protein in LB

To study the interaction between GatD and MurT proteins, expression and purification of MurT protein must also be optimized in order to obtain high amounts of protein with a high degree of purity.

For this reason, different temperatures (37ºC, 30ºC and 25ºC) and host strains (BL21(DE3) and BL21-CodonPlus(DE3)-RIPL) were tested and the results are presented in Figure 3.11.

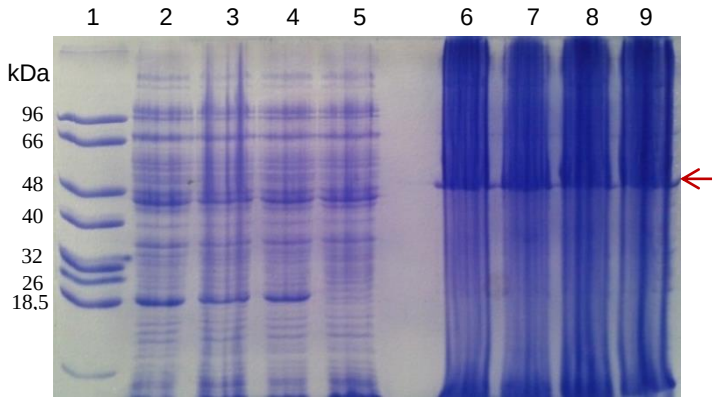

Figure 3.11- Optimization of the expression conditions of MurT-His$_6$tag (50.72kDa). Lane 1: M- molecular marker (Low Molecular Weight Protein marker); Lanes 2-4: soluble protein fractions from BL21-CodonPlus(DE3)-RIPL+pET28a-MurT-His$_6$tag induced at 37ºC, 30ºC and 25ºC, respectively; Lane 5: soluble protein fraction from BL21(DE3)+pET28a-MurT-His$_6$tag induced at 25ºC; Lanes 6-9: insoluble protein fractions of the conditions described in lanes 2-5. Red arrow: MurT-His$_6$tag.

In all the conditions tested, the MurT protein (50.72kDa) was present mainly in insoluble form (Figure 3.11, lanes 6-9) indicating that the protein is associated with the cell membrane or misfolded. Either way, proteins in insoluble form are usually devoid of biological activity and methods to purify it involve unfolding and refolding the protein. During this process the protein may not acquire the correct conformation.

At this point of the work, the primary strategy needed to be changed. In one hand, the poor levels of expression obtained for His$_6$tag-GatD and GatD-His$_6$tag in MM (see section 3.8), prevented the use of GatD protein as the partner to be labeled with ^{15}N and ^{13}C. In the other hand, minimum amounts of soluble MurT protein were retrieved, making this protein an unsuitable alterative for labeling. Furthermore, the MurT protein size surpasses the size limit of the NMR technique (approx.35 kDa).

A new strategy was adopted: GatD protein would be the unlabeled interaction partner and a domain of the MurT protein, responsible for the interaction with GatD, would be labeled. MurT protein is composed of two domains, the N-terminal domain with homology to Mur ligases and a C-terminal domain, DUF1727. While the N-terminal domain is most probably involved in substrate recognition, the DUF1727 has no function attributed, although being present in most MurT-GatD architectures, suggesting an important role for MurT-GatD complex. The role of DUF1727 domain in the interaction between GatD and MurT proteins was tested.

3.10. Co-purification of MurTWDUF1727 and GatD-His$_6$tag

In order to understand the role of DUF1727 domain in the MurT-GatD interaction, a co-purification assay between the DUF1727 domain and GatD-His$_6$tag was performed. An already available genetic construction, pET28a-MurT-GatD-His$_6$tag, which allows the co-expression of MurT and GatD-His$_6$tag was used to generate pET28a-MurTWDUF1727-GatD-His$_6$tag, which allows the co-expression of MurT protein without the DUF1727 domain (MurTWDUF1727) and GatD-His$_6$tag. Site-directed mutagenesis was performed to insert a stop codon in the MurT-GatD-His$_6$tag amino acid sequence. The stop codon was inserted before the first codon of DUF1727 domain by substituting a Leu (TTA) residue for a stop codon (TGA) (change the thiamine 938 of *murT* gene). Co-expression and co-purification of MurT without DUF1727 and GatD-His$_6$tag protein was performed and the results are presented in Figure 3.12.

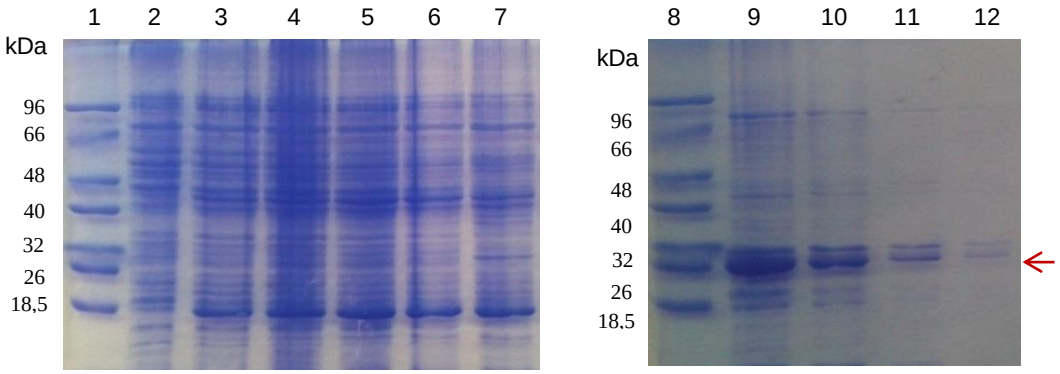

Figure 3.12: Expression of MurTWDUF1727-GatD-His$_6$tag from BL21-CodonPlus(DE3)-RIPL+pET28a-MurTWDuf-GatD-His$_6$tag grown in MM and induced with IPTG at OD_{600nm} =0.8 at 25ºC. Purification was performed through IMAC using AKTA pure system and HiTrap IMAC FF-5 ml column, with 30 and 500mM imidazole. Fractions of 1.5 ml were collected; Lanes 1 and 8: molecular marker (Low molecular weight protein marker); Lane 2: flow-through; Lanes 3-7: washing fractions at 30 mM imidazole; Lanes 9-12: elution fractions at 500 mM imidazole. Red arrow: GatD-His$_6$tag.

Analysis of the target proteins through SDS-PAGE showed that GatD protein, with an estimated molecular size of 28.25 kDa, was eluted without MurTwDUF1727 protein (35.41 kDa), suggesting that MurT protein does not interact with GatD, when it lacks the DUF1727 domain (Figure 3.11, lanes 9-11).

This result suggests that DUF1727 domain is important for the interaction of the MurT-GatD complex.

Co-purification of DUF1727 and GatD-His$_6$tag from BL21-CodonPlus(DE3)-RIPL+pET28a-DUF1727-GatD-His$_6$tag (R. Lobo, *et al.*, unpublished), showed that DUF1727 domain and GatD protein are eluted together. These two complementary co-purification assays demonstrated the importance of the DUF1727 domain for the interaction between GatD and MurT proteins. For this reason, and because MurT was not obtained in soluble form, DUF1727 domain was chosen as the interaction partner of GatD for NMR analysis instead of MurT protein.

Furthermore, due to the low expression yield of GatD protein in MM, the DUF1727 domain was chosen for the assignment of the ^{15}N and ^{13}C NMR spectra instead of GatD protein. The interactions studies were performed using ^{15}N, ^{13}C DUF1727 and unlabeled GatD.

3.11. Expression and purification of DUF1727 domain

3.11.1. Optimization of the expression conditions of DUF1727-His$_6$tag domain

Taking together the technical limitations and the information on DUF1727 role, this small domain became a crucial player in the NMR study of GatD and MurT interaction. For this reason, optimization of the expression conditions for DUF1727-His$_6$tag was performed. In order to verify its solubility and the yield of expression, different temperatures (37ºC, 30ºC and 25ºC), induction periods (3 h, 4 h and 5 h) and host strains (*E. coli* BL21-CodonPlus(DE3)-RIPL, BL21(DE3), Rosetta(DE3)) were tested. The results of different temperatures (37ºC, 30ºC and 25ºC), induction periods (3 h) and host strains (BL21(DE3), BL21-CodonPlus(DE3)-RIPL) tested are presented in Figure 3.13 and the results for the others conditions are not shown.

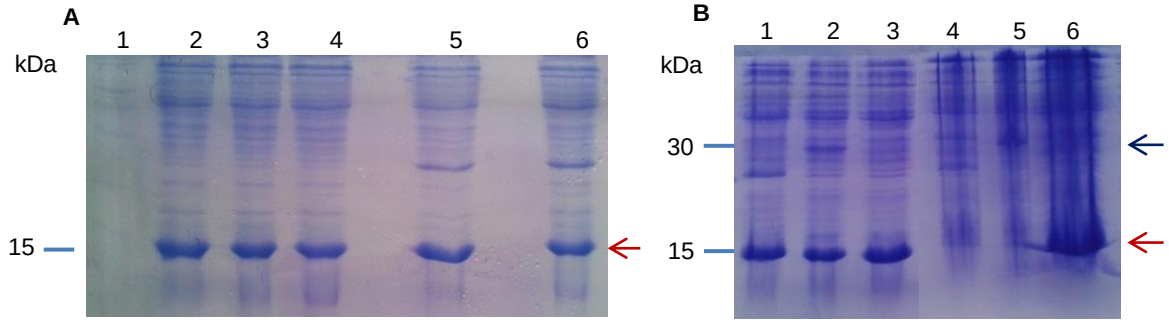

Figure 3.13- Optimization of the selected expression conditions of DUF1727-His$_6$tag (≃15.33 kDa) in LB, using as controls, the expression of His$_6$tag-GatD and pET28a. Panel **A-** Lane 1: molecular marker (Novex® Sharp Unstained Protein Standard); Lanes 2-4: soluble protein fractions from expression from BL21(DE3)+pET28a-DUF1727-His$_6$tag induced at 37ºC, 30ºC and 25ºC, respectively; Lanes 5-6: soluble protein fractions from BL21-CodonPlus(DE3)-RIPL+pET28a-DUF1727-His$_6$tag induced at 30ºC and 25º, respectively. Panel **B**- Lane 1: soluble protein fraction from BL21-Codonplus(DE3)-RIPL+pET28a induced at 30ºC; Lane 2: soluble protein fraction from BL21(DE3)+pET28a-His$_6$tag-GatD induced at 30ºC; Lane 3: soluble protein fraction from BL21(DE3)+pET28a-DUF1727-His$_6$tag induced at 30ºC; Lanes 4-6: insoluble protein fractions of the conditions in lanes 1-3. Red arrow: DUF1727-His$_6$tag; Blue arrow: His$_6$tag-GatD.

It not was possible to correctly estimate the expression level of DUF1727-His$_6$tag domain (approx.15.33 kDa) in the soluble fraction, due to the masking effect of an *E. coli* endogenous protein with a similar molecular weight. In fact, the band corresponding to such protein was observed in the soluble fraction of the negative control assays (Figure 3.13, panel B, lanes 1-2), however was not detected in the insoluble fraction (Figure 3.13, panel B, lanes 4-5). In this way, the optimization tests performed to optimize the expression of DUF1727-His$_6$tag were not informative.

To determine the best expression conditions for the DUF1727 domain in the soluble fraction, IMAC purification was performed in order to separate DUF1727-His_6tag from the endogenous protein with the same molecular size. DUF1727-His_6tag was expressed in large-scale and purification was performed by manual affinity chromatography using Ni-NTA matrix (Figure 3.14).

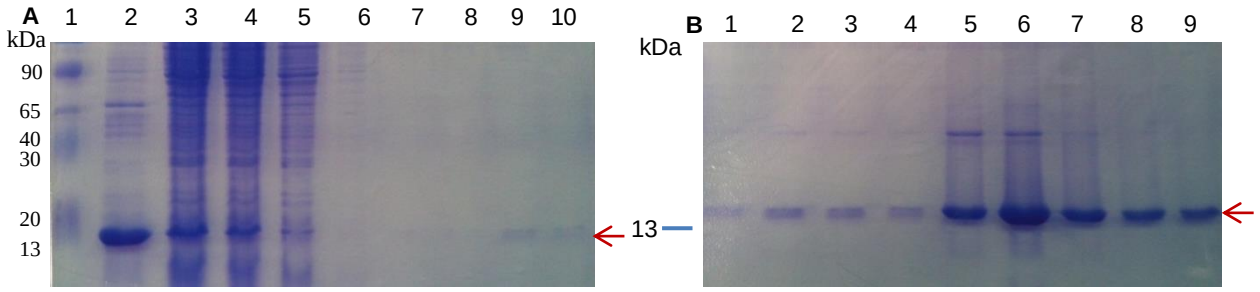

Figure 3.14- Expression of DUF1727-His_6tag (15.33 kDa) from BL21(DE3)+pET28a-DUF1727-His_6tag grown in 600 ml LB and induced with 1 mM IPTG at OD_{600nm}=0.8 at 30ºC. Purification was performed through manual affinity chromatography using Ni-NTA matrix and 10, 30, 50, and 200 mM imidazole. Panel **A**- Lane 1: molecular marker (ColorBurstA™ Electrophoresis Marker); Lane 2: insoluble protein fraction; Lane 3: soluble protein fraction; Lane 4: flow-through; Lanes 5-7: washing fractions with 10 mM imidazole; Lanes 8-10: washing fractions with 30 mM imidazole; Panel **B-**Lanes 1-4: washing fractions with 50 mM imidazole; Lanes 5-9: elution fractions with 200 mM imidazole. Red arrow: DUF1727-His_6tag.

Elution of DUF1727-His_6tag started at a 50 mM concentration of imidazole and the maximum elution rate occurred for 200 mM of imidazole (Figure 3.14, panel B, lanes 1-9). The *E. coli* endogenous protein with similar molecular weight to DUF1727 domain was eluted in the flow-through and in the washing steps with 10 mM and 30 mM of imidazole (Figure 3.14, panel A, lanes 2-9). Overall, a final concentration of approx. 133 µM of the target protein was achieved for 1 L of culture.

DUF1727 was expressed as a fusion protein with a C-terminal His_6tag. As the His_6tag can interfere with the physical interaction between two proteins, especially if a protein with low molecular size is involved, it is advisable to remove the His_6tag prior to NMR analysis. However the DUF1727-His_6tag amino acid sequence does not include a cleavage site that allows for the removal of the C-terminal His_6tag.

To surpass this, two strategies were performed: 1- Construct and expression of DUF1727 fusion with a N-terminal His_6tag, followed by cleavage of the His_6tag using thrombin; 2- Expression of DUF1727-GatD-His_6tag complex and subsequent isolation of DUF1727 domain upon purification using affinity chromatography with a high ionic strength elution buffer, in order to separate the two proteins.

3.11.2. Expression and purification of His_6tag-DUF1727

The His_6tag-DUF1727 was expressed from BL21(DE3)+pET28a-His_6atg-DUF1727 and purification was performed through manual affinity chromatography using the Ni-NTA matrix. The results are presented in Figure 3.15.

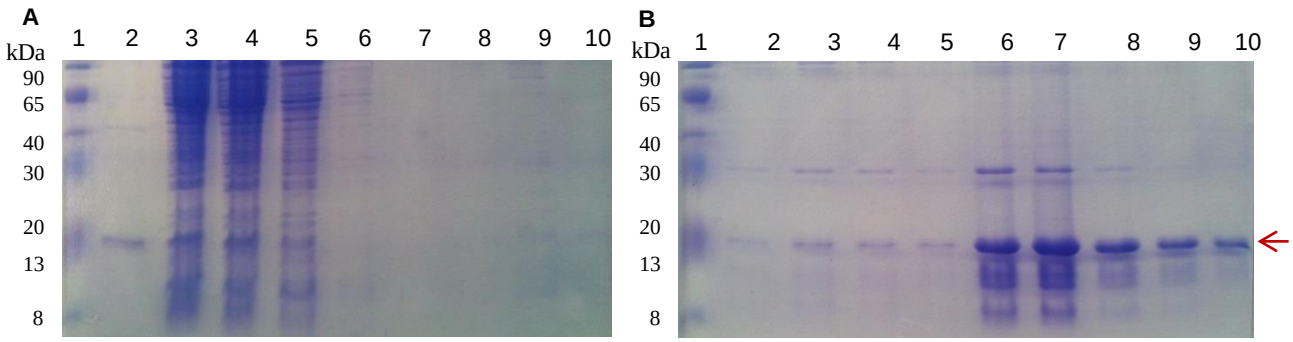

Figure 3.15- Expression of His_6tag-DUF1727 (15.47 kDa) from BL21(DE3)+pET28a-His_6tag-DUF1727 grown in 800 ml LB and induced with 1 mM IPTG at OD_{600nm}=0.8 at 30ºC. Purification was performed through manual affinity chromatography using Ni-NTA matrix and 10, 30, 50 and 200 mM imidazole. Panel **A**- Lane 1: molecular marker (ColorBurstA™ Electrophoresis Marker); Lane 2: insoluble protein fraction; Lane 3: soluble protein fraction; Lane 4: flow-through; Lanes 5-7: washing fractions at 10mM imidazole; Lanes 8-10: washing fractions at 30 mM imidazole. Panel **B**- Lane 1: molecular marker; Lanes 2-5: washing fractions at 50 mM imidazole; Lanes 6-10: elution fractions at 200 mM imidazole. Red arrow: His_6tag-DUF1727.

Elution of His_6tag-DUF1727 started a 50 mM imidazole concentration and the maximum elution rate occurred at 200 mM imidazole, (Figure 3.15, panel B, lanes 5-10). A final yield of approx. 90 µM of protein was achieved for 1L of culture.

The protein yield obtained for expression and purification of His_6tag-DUF1727 was significantly lower than the yield obtained for DUF1727-His_6tag (approx.90 µM and 133µM, respectively, using 1L of culture). Furthermore, the purity level of His_6tag-DUF1727 was lower than for DUF1727-His_6tag. This could be due to a higher amount of eluting endogenous proteins of *E .coli* or to degradation of His_6tag-DUF1727 domain, as the contaminant proteins have lower molecular size than DUF1727 (Figure 3.14, panel A, lanes 5-9).

Finally, DUF1727 domain with a N-terminal His_6tag proved to be much more unstable than if the His_6tag is placed at the C-terminal, as it precipitated almost immediately after purification.

3.11.2.1. Cleavage of the hexa-histidine tag of His_6tag-DUF1727

Thrombin is a serine protease that recognizes the sequence Leu-Val-Pro-Arg-Gly-Ser, and cleaves the peptide bond between Arg and Gly. A thrombin cleavage sequence is present between the N-terminal His_6tag and DUF1727 amino acid sequence (Figure 5.1, annex).

After the His_6tag-DUF1727 was expressed and purified through affinity chromatography, the elution fractions were concentrated to 1 mg/ml and the His_6tag was removed by digestion with thrombin (Figure 3.16).

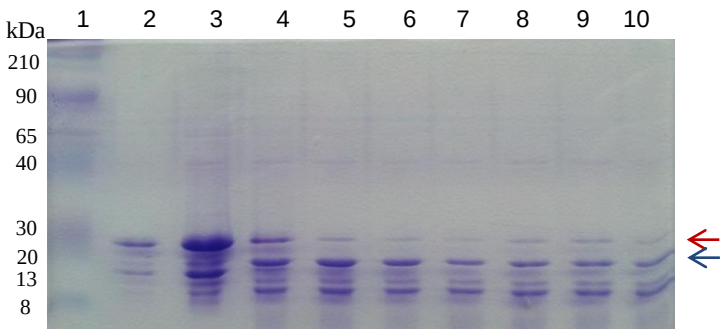

Figure 3.16- Cleavage of His_6tag after expression of His_6tag-DUF1727 from BL21(DE3)+pET28a-His_6tag-DUF1727 grown in LB and purification through manual affinity chromatography. Lane 1: molecular marker (ColorBurstA^{TM} Electrophoresis Marker); Lane 2: sample before concentration step; Lane 3: sample after concentration step; Lanes 4-10: sample after 1h18, 2h24, 3h50, 16h, 18h40, 21h and 24h digestion. Red arrow: DUF1727-His_6tag; Blue arrow: DUF1727 domain without His_6tag..

Although thrombin was able to cleave the N-terminal His_6tag of His_6tag-DUF1727, as observed by the presence of a second protein band with lower molecular size, it proved to be poorly effective, as the digestion was not complete even after 24 h. In fact, after 2h24 of digestion, the enzymatic reaction reaches a point where no further cleavage occurs (Figure 3.16, lanes 5 and 10). In comparison with the initial amount of protein used, the amount of DUF1727 without His_6tag (13.81 kDa) obtained in the soluble fraction, decreases considerably (Figure 3.16, lanes 3-10). In fact, the protein precipitated during the digestion and perhaps the aggregation of the insoluble protein did not allow for complete digestion of the sample.

In this way, the cleavage of the His_6tag, resulted in poor protein yield and in mixture of digested and undigested protein. Thus, this strategy was excluded.

3.11.3. Expression of DUF1727-GatD-His_6tag complex

In order to purify DUF1727 domain from the DUF1727-GatD-His_6tag complex, the complex was expressed and purified through manual affinity chromatography. A high ionic strength elution buffer was used, to attempt to separate the two proteins (Figure 3.17).

Three buffers were used sequentially in the same assay: buffer 1- Tris-HCl pH=7.5, 10 mM imidazole and moderate ionic strength (0.5 M NaCl), buffer 2- Tris-HCl pH=7.5 with no imidazole and with high ionic strength (1.5 M NaCl) and buffer 3- Tris-HCL pH=7.5, with 500 mM imidazole and moderate ionic strength.

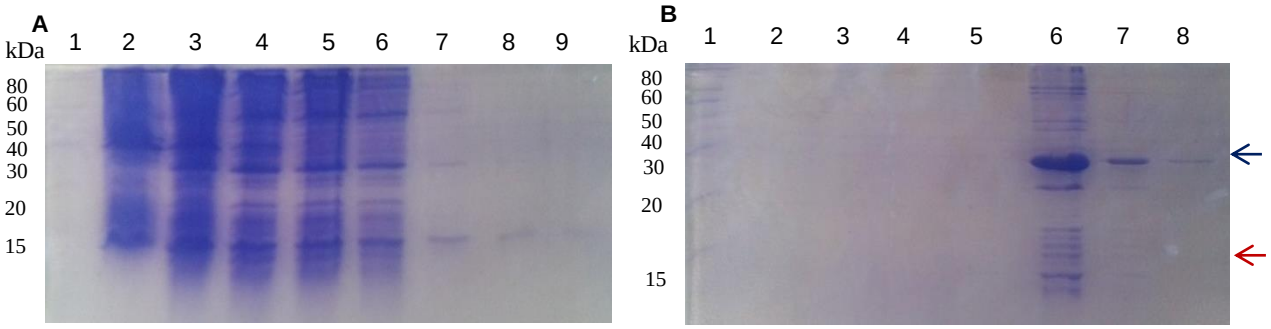

Figure 3.17- Expression of DUF1727-GatD-His_6tag from BL21-CodonPlus(DE3)-RIPL+pET28a-DUF1727-GatD-His_6tag grown in LB and induced with 1 mM IPTG at OD_{600nm}=0.8 at 30ºC. Purification was performed through manual affinity chromatography. Panel **A**- Lane 1: molecular marker (Novex® Sharp Unstained Protein Standard); Lane 2: insoluble protein fraction; Lane 3: soluble protein fraction; Lane 4: flow-through; Lanes 5-8: washing fractions 1-4 at 10 mM imidazole and 500 mM NaCl; Lane 9: washing fraction 1 without imidazole and 1.5 M NaCl. Panel **B**- Lane 1: molecular marker; Lanes 2-5: washing fractions 2-5 without imidazole and 1.5 M NaCl; Lanes 6-8: elution fractions 1-3 with 200 mM imidazole and 500 mM NaCl. Red arrow: DUF1727; Blue arrow: GatD-His_6tag.

The results showed that DUF1727 domain and GatD protein were co-eluted at a 200 mM imidazole concentration (Figure 3.17, panel B, lanes 6-8), strengthening the hypothesis that DUF1727 domain has the function to recognize and bind to GatD protein. However, the DUF1727 was not eluted with a high salt concentration (1.5 M NaCl), suggesting that DUF1727-GatD interact strongly (Figure 3.17, panel B, lanes 2-5).

The strategy to obtain DUF1727 domain through the complex did not result. Thus, to obtain a ^{1}H-^{15}N-spectrum of DUF1727 the only available option was to express the DUF1727-His_6tag.

3.11.4. Expression of ^{15}N labeled DUF1727-His_6tag

The DUF1727-His_6tag was expressed from BL21(DE3)+pET28a-DUF1727-His_6tag grown in LB to an OD=0.8. Cells were harvested and resuspended in MM supplemented with $^{15}NH_4Cl$ and 1 mM IPTG. Purification was performed through IMAC using AKTA pure system and Hitrap IMAC FF (Figure 3.18).

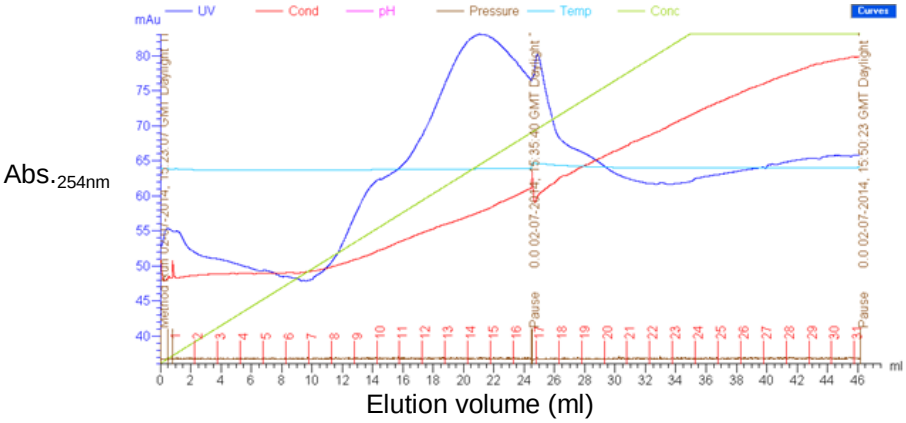

Figure 3.18.A: Chromatogram of DUF1727-His_6tag purification using HiTrap IMAC FF-5 ml column. An imidazole concentration gradient was applied. 0-35 ml: imidazole gradient from 30-500 mM; 35-46: 500 mM imidazole. Protein was detected by monitoring absorbance at 254 nm.

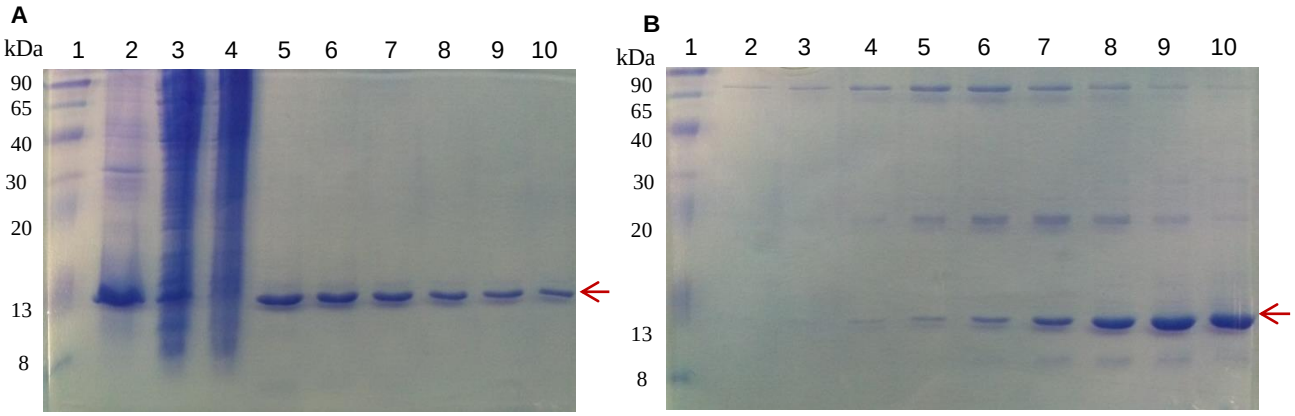

Figure 3.18.B- DUF1727-His_6tag expressed from BL21(DE3)+pET28a-DUF1727-His_6tag grown in 750 ml LB and induced with 1 mM IPTG at OD_{600nm}=0.8 in 750 ml MM supplement with $^{15}NH_4Cl$ labeled. Purification was performed through IMAC using AKTA pure system and Hitrap IMAC FF-5 ml and an imidazole concentration gradient (30-500 mM). Fractions of 1.5 ml were collected. Panel **A**-Lane 1: molecular marker (ColorBurstA™ Electrophoresis Marker); Lane 2: insoluble protein fraction; Lane 3: soluble protein fraction; Lane 4: flow-through; Lanes 5-10: elution fractions from 21.5 ml to 29 ml (312-405 mM imidazole); Panel **B**- Lane 1: molecular marker; Lanes 2-5: washing fractions from 8 ml to 12.5 ml; Lanes 6-10: elution fractions from 14 ml to 20 ml (220-296 mM imidazole). Red arrow: DUF1727-His_6tag.

Elution of DUF1727-His_6tag started at a 220 mM imidazole concentration and the maximum elution rate occurred for an imidazole concentration range of 251-331 mM (Figure 3.18.B, panel A, lanes 6-10 and panel B, lanes 5-10). The final yield of this assay was approx. 66 µM for 1L of culture. As DUF1727-His_6tag showed a considerable purity degree, no more purification steps were performed, prior to NMR analysis. As expected, expression of DUF1727-His_6tag in labeled MM showed approximately half the yield of expression in LB (Figure 3.14) (66µM and 133 µM, using 1 L of culture). After purification, DUF1727-His_6tag ^{1}H-^{15}N-spectrum was obtained.

3.11.4.1. 1D ^{1}H NMR and ^{1}H,^{15}N-HSQC of ^{15}N labeled DUF1727-His$_6$tag

Because the His$_6$tag may or not may affect protein/protein interaction and because the two strategies to obtain DUF1727 domain purified without His$_6$tag failed, due the poor yield of the domain cleaved and the co-expression of GatD with DUF1727, DUF1727-His$_6$tag was expressed and purification was performed through IMAC.

After the purification step, the buffer of the eluted fractions was exchanged, then concentrated and dialyzed overnight to exchange Tris-HCL to Tris-DCL pH=7.2.

The sample with 210µM was then analyzed by NMR. Several ^{1}H-NMR spectra and ^{1}H, ^{15}N-HSQC spectra were obtained at different pHs in order to improve the dispersion of the signals in the NH region (6.5 to 10 ppm). In Figure 3.19 the best ^{1}H-NMR spectrum obtained with the target protein corresponding to pH≃6 is presented.

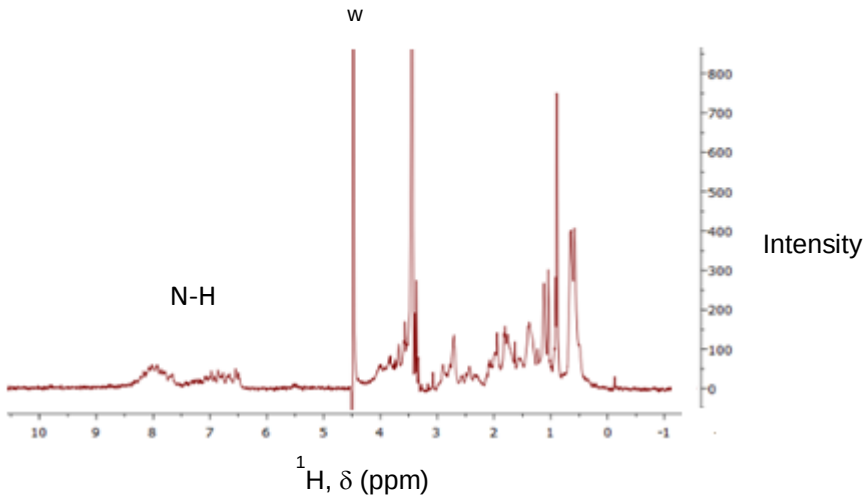

Figure 3.19- ^{1}H-NMR spectrum 1D of DUF1727-His$_6$tag at 210 µM in 20mM Tris-DCL pH=6.7, 200mM NaCl. DUF1727-His$_6$tag was expressed from BL21(DE3)+pET28a-DUF1727-His$_6$tag grown in LB to an OD=0.8. Cells were harvested and resuspended in MM supplement with $^{15}NH_4Cl$ labeled. N-H: N-H region; W-water.

The results of the 2D-^{1}H,^{15}N-HSQC spectra obtained at different pH values are presented in Figure 3.20.

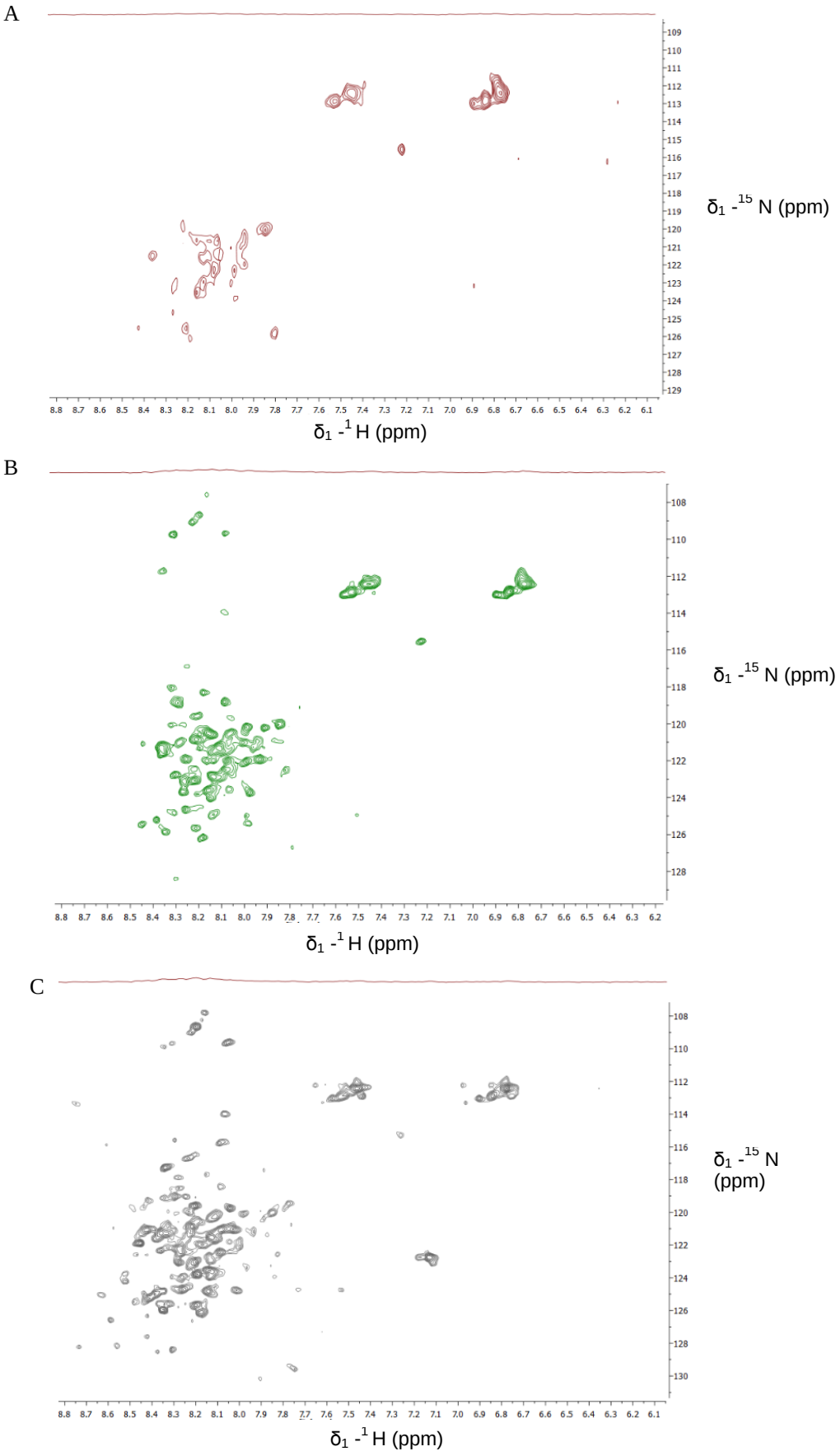

Figure 3.20: ^{1}H-^{15}N-spectrum 2D of DUF1727-His$_6$tag in 20mM Tris-DCL and 200mM NaCl. The DUF1727-His$_6$tag was expressed from BL21(DE3)+pET28a-DUF1727-His$_6$tag grown in LB to an OD=0.8. Cells were harvested and resuspended in MM supplement with $^{15}NH_4Cl$ labeled and purified through IMAC. Panel A: sample at pH=7.7; Panel B: sample at pH=6.7; Panel C: sample at pH≃6.

As can be seen on the sequence of spectra in Figure 3.20, when the pH of the sample was decreased, the dispersion of hydrogen and nitrogen resonances increased. At pH 7.7 the signals are ill defined and present a low dispersion, characteristic of a partially folded or unfolded state. With decreasing pH the signals become more intense and dispersed, with the hightest disperson obtained at pH≃6 (Figure 3.20, panel A-C). These results indicate that the domain becomes more structured with decreasing of pH, until complete unfolding at pH≃5.

3.1.4.2. H^{1}-N^{15}-spectrum 2D of DUF1727-GatD

One possible reason for the structural flexibility detected for the DUF1727 domain minght be that it only adopts a stable structure when it interacts with GatD. In order to test this hypothesis attempts were made to titrate DUF1727 with GatD and follow by $^{1}H,^{15}N$-HSQC. However, when GatD was added to DUF1727 the proteins co-precipitated and for this reason it was not possible to acquire the NMR spectrum. This result is however indicative of a strong interaction between the two proteins.

Discussion

In Gram-positive bacteria peptidoglycan constitutes the major component of the cell wall, is essential for survival, maintenance of cell shape and counterbalance of turgor pressure (Höltje, 1998). Its structure is well conserved, however, a certain degree of variation exists either in the stem peptide of peptidoglycan (Vollmer *et al.*, 2008), namely the third amino acid which is usually L-Lys in Gram-positive bacteria and *m*-DAP (meso-diaminopimelic acid) in Gram-negatives, secondary modifications in the glycan strands, or the position and composition of the interpeptide bridge, a pentaglycine in the case of *S. aureus*. Amidation modifications to stem peptide have been only reported in Gram-positive bacteria and have been described for D-glutamate (the second amino acid of UDP-Mur-NAc-pentapetide) in *S. aureus* (Figueiredo *et al.*, 2012 ; Münch *et al.*, 2012) and *Streptococcus pneumonia* (Zapun *et al.*, 2013), for *m*-DAP in *Bacillus subtilis* (Lebar *et al.*, 2014) and *Lactobacillus plantarum* (Bernard *et al.*, 2011) and for D-Asp (aspartic acid) in *Lactococcus lactis* (Veiga *et al.*, 2009).

Peptidoglycan biosynthesis, the target of β-lactam antibiotics, has been extensively studied; the enzymatic steps of the primary synthesis pathway have been known for a long time, however, the mechanisms and the proteins involved on the secondary modifications such as amidation, were only recently described. Amidation in *S. aureus* occurs at position 2 of the stem peptide, resulting in the conversion of D-glutamic acid into D-iso-glutamine. This secondary modification is catalyzed by two enzymes, MurT and GatD proteins, being lipid II (C55-P-P-GlcNAc-MurNAc-pentapeptide) the reaction substrate. The respective genes are part of a small operon, *murT-gatD* and are expressed as a single transcript. MurT and GatD form a physically stable bi-enzymatic complex in a molar ratio of 1:1 and share identity and similarity with two families of proteins: the Mur ligases and the glutamine amidotransferases (GATases) (Figueiredo *et al.*, 2012; Münch *et al.*, 2012).

The amidation of peptidoglycan was shown to reduce the growth rate and to be necessary for resistance to β-lactam antibiotics, making *murT* and *gatD* auxiliary genes. Furthermore, the amidation step was found to be essential for *S. aureus* intrinsic resistance to lysozyme (Figueiredo *et al.*, 2012).

Due to the impact on this modification on *S. aureus* resistance to β-lactams, the MurT and GatD proteins have become attractive targets for the design of new antimicrobial compounds. In this context, it is essential to thoroughly characterize the enzymatic activity of these proteins and their molecular structure. The objective of this Master project was to dynamically characterize the structure of MurT-GatD complex by Nuclear magnetic resonance (NMR). The first step consisted on cloning the *gatD* and *murT* genes of *S. aureus* in order to construct recombinant proteins and optimize their expression in *E. coli* host strains. The second step, included the production of ^{1}H and ^{15}N, ^{13}C uniformly labeled GatD protein and assignment of the respective ^{15}N, ^{13}C HSQC NMR spectra. Once the GatD structure is provided, interaction studies of ^{13}C, ^{15}N GatD with MurT would be performed by NMR spectroscopy.

During this project, this primary strategy had to be changed due to method limitations and a second strategy was adopted: DUF1727 domain (C-terminal domain of MurT) was chosen for the assignment of the ^{1}H, ^{15}N and ^{13}C NMR spectra instead of GatD protein. The interactions studies were performed using ^{1}H,^{15}N, ^{13}C-spectrum DUF1727 and unlabeled GatD.

4.1. Expression and purification of GatD recombinant protein

Expression and purification of GatD protein had already been reported in the literature for growth in complex medium: GatD fusion to a C-terminal His_6tag (Münch *et al.*, 2012) and to a N-terminal His_6tag (Figueiredo, *et al*, unpublished; Vieira, *et al.*, 2014), using different expression vectors and strains. In this study, the conditions for expression of the His_6tag-GatD construct (Table 2.1) (Figueiredo *et al*, unpublished) were optimized for growth in minimal medium.

The target protein was fused to a His_6tag, in order to enable purification through affinity chromatography. However, to obtain a degree of purity which is compatible with the NMR technique, two chromatographic steps (affinity and molecular exclusion chromatography) were performed for GatD recombinant protein. This extra step was needed due the co-elution of *E. coli* endogenous proteins with GatD in affinity chromatography (Figure 3.1.A-3.1.B) and anionic chromatography (Figure 3.2.A-3.2.B). These approaches were also used for GatD purification for crystallography studies (Vieira, *et al.*, 2014), which also needs high purity levels. For protein activity tests only one purification step (affinity chromatography) was performed (Münch *et al.*, 2012b) as a high degree of protein purity is not needed.

The co-elution of *E. coli* endogenous proteins in our GatD samples can be explained by the presence of proteins that interact with the matrix or directly with the target protein. Naturally histidine-rich proteins have affinity to metal ions and can interact with the matrix. The ArnA protein (74.28kDa, pI=6.39), an enzyme involved in the modification of lipid A phosphate, has several exposed non-consecutive histidine residues (Gatzeva-topalova *et al.*, 2005) and SlyD protein (20.8kDa, pI=5.03), a peptidyl-prolyl cis/trans-isomerase that contains 15 histidines residues at the C-terminal (Weininger *et al.*, 2009), are the most probable *E. coli* endogenous contaminants of the affinity chromatography step (Figure 3.1.A-3.1.B). The isoelectric point of these two proteins is similar to the His_6tag-GatD (pI=6.01), in this way suggesting their co-elution also upon anionic chromatography (Figure 3.2.A-3.2.B). These *E. coli* endogenous proteins with affinity for nickel ions only represent a problem when the expression of the recombinant protein of interest shows a low yield of expression because they are not out-competed by the sheer amount of the protein of interest (Andersen *et al.*, 2013).

Another possibility for the co-elution of endogenous proteins is the direct interaction between the *E. coli* proteins and the recombinant GatD. To address this hypothesis, purification of GatD was performed under conditions of high ionic strength and also in the presence of a denaturing reagent (β-mercaptoethanol), in order to minimize protein-protein interactions. As this strategy proved to be inefficient, the hypothesis of the co-elution with *E. coli* histidine-rich proteins that bind to the nickel matrix, was the most viable one. In this way, the purification of His_6tag-GatD protein with the required purity level was achieved by performing affinity chromatography followed by size exclusion chromatography (Figure 3.3.A-3.3.B).

After the optimization of the purification of His_6tag-GatD, several stability screenings were performed (Table 2.3). The His_6tag-GatD was stable only after the addition of glycerol, at neutral pH. After these optimizations, a ^{1}H-spectrum for His_6tag-GatD was obtained (Figure 3.6) and although the dispersion of signals in the N-H region (5-12 ppm) is not high, it is indicative of a structured molecule,

meaning that the protein is folded. Thus, suitable conditions for the analysis of GatD by NMR were determined in this project.

4.2. Native conformations of GatD recombinant protein

Several assays performed in this study showed that GatD recombinant protein naturally occurs in two conformation states: as a monomer and as a dimer, being the monomeric form the predominant one (Figure 3.3.A-3.3.B). In fact, the chromatogram of the size exclusion assay showed the elution of GatD recombinant protein at two distinct retention times, corresponding to the elution of proteins with approximately 60 kDa and 30 kDa. Furthermore, by strengthening the denaturing conditions used to separate the protein elution fractions in SDS-PAGE (Figure 3.4), the protein band with higher molecular size (approx. 60 kDa) virtually disappeared, indicating that an oligomeric structure suffered denaturation. Finally, western blotting assays performed with anti-GatD Abs or with anti-His_6tag Abs (Figure 3.5), detected two signals (approx. 60 kDa and 30 kDa bands) in the elution fraction of the GatD recombinant protein. For the negative control sample (expression from BL21(DE3)-CodonPlus-RIPL+pET28a), signal was also obtained (approx. 60 kDa), although only when using anti-GatD Abs. This result can be explained by the presence of *E. coli* glutamate amidotransferases in the protein extracts (Massière and Badet-Denisot, 1998). The absence of signal using anti-His_6tag Abs strengthens this hypothesis. In fact, this Ab only detects six sequential N-terminal histidine residues and does not detect non-consecutive histidine residues or four consecutive C-terminal histidines present in the *E. coli* histidine-rich proteins.

The GatD dimeric form observed in this study is probably an artifact from in vitro high-level overexpression, which stimulates protein self-assembly. This observation is in accordance with the results obtained for GatD crystallography analysis, as the protein crystallized as a dimer (Vieira, *et al*, 2014.).

4.3. Optimization of expression of GatD recombinant protein in minimal medium (MM)

To confirm and obtain more structural information by NMR analysis, the protein under study must be isotopically labeled, usually with ^{15}N or ^{13}C. For this purpose, the protein must be expressed in MM supplemented with the appropriate labeled reagent, in order to ensure that the only nitrogen or carbon sources present is the medium are labeled, guaranteeing their incorporation in the protein amino acid sequence. A lower expression yield of a target protein is expected for growth in minimal medium (MM) when compared to complex medium, due to the limited nutrient availability. For this reason, optimization of the expression conditions in MM was performed (Figure 3.7).

Firstly, other host expression strains were tested. Efficient expression of heterologous proteins in *E. coli* is sometimes limited by the shortage of specific tRNAs abundant in the source organism. Forced over-expression of some heterologous proteins can deplete the pool of rare tRNAs, preventing translation. *E. coli* strains have been engineered to contain extra copies of genes that encode the tRNAs that are most frequently depleted. The use of such expression strains did not result in an

increase of GatD expression, indicating that the rare codons present in GatD amino acid sequence do not interfere with expression.

Another strategy to increase protein expression was the use of Rosetta(DE3)pLysS as expression host. This strain provides a more tight control over of the basal expression level, provided by plasmid pLysS, which is an advantageous characteristic when the target protein is toxic for the cells. Moreover, the induction temperature used was lower, 20ºC, in order to slow down the transcription and translation rates, in an attempt to enhance folding. However, this approach resulted in lower protein yields.

Secondly, several induction conditions were tested in order to identify the ones which result in more GatD protein production. Higher expression yields were obtained by inducing expression overnight at 30ºC. However this higher expression was followed by a decrease in solubility, indicating that most probably the target protein becomes misfolded for long induction times.

The low-expression of GatD recombinant protein can be associated to the function of the target protein: GatD is a class I glutamine amidotransferase (Gn-AT) (Figure 4.1). The Gn-AT proteins usually have two domains (synthase and glutaminase domains) that catalyze the removal of the ammonia group from glutamine and then transfer this group to a substrate molecule by forming a carbon-nitrogen group, in a ATP-dependent process called amidation (Massière and Badet-Denisot, 1998). GatD protein has only the glutaminase domain, in this way lacking the synthase domain that includes the motif for ATP-binding. Thus, only when the GatD interacts with MurT protein, does the complete reaction occur. However, GatD alone can hydrolyze free glutamine, an ATP independent reaction, and its over-expression may result in depletion of cellular glutamine, becoming toxic to the host cell.

As all the attempts to increase the expression of GatD protein in MM, by choosing another expression host or altering the expression conditions, were unsuccessful, the higher yield obtained was not enough to proceed for NMR analysis.

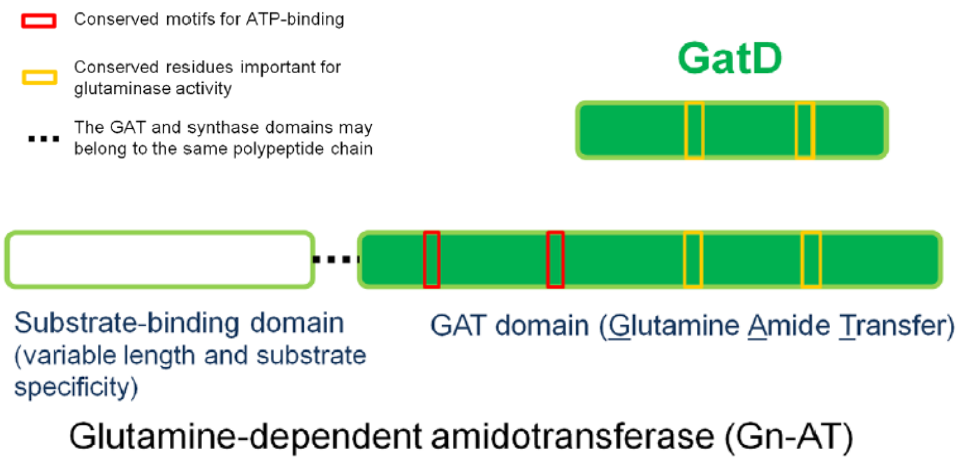

Figure 4.1- Schematic representation of the domains comparison between the GatD protein domains and the glutamine-dependent amidotransferase (GN-AT) domains. Adapted from Figueiredo *et al*., 2012

4.4. Expression of MurT protein

The MurT recombinant protein was constructed as a C-terminal His_6tag fusion and expressed in complex medium. MurT-His_6tag was almost completely obtained in the insoluble fraction, in contrast with previous reports that describe the expression of the soluble form of an analogous recombinant MurT-His_6tag protein (Munch, *et al*, 2012). Although the His_6tag was placed at the same terminus of the protein, in our construct, eight additional amino acids were incorporated between the tag and the initiation codon. These extra residues may interfere with the correct folding of MurT, being responsible for its poor solubility.

Different *E. coli* strains and induction conditions were tested, however for all, the MurT protein was obtained in the insoluble fraction. Moreover, solubilization attempts were made using detergents and reducing agents, but without success (R. Lobo *et al.,* unpublished).

The insolubility of MurT protein can be due to its association with the membrane fraction. This is a hypothesis supported by the fact that the substrate of the amidation reaction is the membrane-linked muropeptide. However, the specific sites that MurT recognizes and binds to, are not described. The MurT protein has homology with the central domain of Mur ligases, including the conserved ATP binding motif. It also exhibits a short sequence, **DN**A**A**D**D,** with similarly to the L-Lys binding site, **DN**P**A**N**D,** present in MurE sequence (Munch, *et al*., 2012). This characteristic suggests that, for MurT to recognize the lipid-linked substrate, one of structural elements that MurT associates with, is the L-lysine residue of the stem peptide.

Nevertheless, the third position of the stem peptide of *S. aureus* is occupied by a L-lysine residue, while the *E. coli* stem peptide carries meso-diaminopimelic acid (*m*-DAP), thus hampering the association of MurT with the lipid linked muropeptide. In this way, it is rather unlikely that the reason behind MurT insolubility would be binding to the *E. coli* lipid II.

Another hypothesis for the low-solubility of MurT protein could be the direct result of the expression strategy adopted. By promoting the production of high levels of recombinant protein, accumulation of the soluble protein product will occur inevitably. However, a high rate of expression may result in the incorrect folding of the target protein, as no sufficient time is provided for chaperone proteins to assist in the folding process. The accumulation of misfolded protein will result in formation of aggregates of insoluble target protein, designated as inclusion bodies. This hypothesis is strengthened by the fact that MurT protein expressed in the soluble form was reported previously (Münch *et al.*, 2012b) and the IPTG concentration used for such assays was 0.75mM, lower than the concentration of 1mM IPTG used in this study. Lower concentrations of inducer will lead to a decrease in the rate of expression and consequently will favor the correct folding of the protein. However, our attempts to increase MurT solubility, by using an expression temperature of 30ºC, described as optimal for the expression and activity of a number of *E. coli* chaperones (Sorensen and Mortensen, 2005), were not successful.

The inclusion bodies are cytoplasmic aggregates devoid of biological activity, composed of densely packed misfolded proteins and peptides. However, the correct folding and consequently the enzymatic activity of the proteins can be recuperated by applying some sequential steps: i) solubilization of the aggregates by adding harsh denaturing agents, such urea and guanidine

hydrochloride along with reducing agents such as β-mercaptoethanol; (ii) refolding by removing the denaturing agents through dialysis; (iii) finally, a purification step (Singh and Panda, 2004). This strategy was not pursued in this study due to the poor recovery yields of protein with the correct folding usually obtained and the unavailability of an enzymatic assay to confirm protein activity.

4.5. Interaction between the MurT and GatD proteins and the role of DUF1727

In previous assays, the co-elution of MurT and GatD was observed, showing a physical interaction between these proteins (Munch, *et al*., 2012; Figueiredo et al, unpublished). In order to verify if the DUF1727 domain, the C-terminal region of MurT protein, is involved in the interaction between MurT and GatD, new recombinant systems were constructed and expressed: MurTWDUF1727-GatD-His_6tag, in which DUF1727 is absent and DUF1727-GatD-His_6tag. Co-elution of the two partner proteins occurred for DUF1727-GatD-His_6tag and not for MurTWDUF1727-GatD-His_6tag, indicating that DUF1727 is the domain of MurT responsible for the physical interaction with GatD. The fact that the proteins were co-eluted in conditions of high ionic strength suggested a strong interaction between the two proteins.

Furthermore, the DUF1727 domain seems to contribute to an increase in GatD expression. The amount of GatD protein obtained in the three co-expression constructs, was comparatively higher in DUF1727-GatD-His_6tag (Figure 3.17), intermediate in MurT-GatD-His_6tag (Figueiredo *et al*, unpublished, Figure 5.3, annex) and lower in MurTWDUF1727-GatD-His_6tag (Figure 3.12), indicating that DUF1727 may contribute to the correct folding of GatD protein.

4.6. Expression and purification of DUF1727

The DUF1727 domain is the C-terminal domain of MurT protein, does not share similarity with the Mur ligases and is present in all MurT-GatD architectures identified so far; its role for the amidation reaction was until now unknown. Due not only to the fact that DUF1727 is able, by itself, to interact with GatD, but also due to its small molecular size (15.43 kDa), this domain emerged as an advantageous alternative to MurT protein, as a partner for NMR interaction studies with GatD.

Optimization of the expression of DUF1727 domain was hampered by the presence of an *E. coli* endogenous protein with a molecular size similar to DUF1727, which prevented an expedite comparison of the expression level of DUF1727 under different conditions. Furthermore, since DUF1727 has a small molecular size, the His_6tag could have a disturbing effect on the interaction with GatD. In fact, the two recombinant proteins constructed, DUF1727 fusion with N- and C-terminal His_6tag (His_6tag-DUF1727 and DUF1727-His_6tag), showed different expression yields, being the C-terminal fusion more efficiently expressed and more stable, suggesting that the His6tag interferes with the folding of the domain. Although attempts were made to remove the tag by cleavage, these approaches resulted in a too significant protein yield loss.

Despite these setbacks, enough amount of DUF1727-His_6tag was produced for NMR analysis.

4.7. ^{1}H-^{15}N-NMR analysis of DUF1727

The expression of DUF1727-His_6tag allowed to perform a ^{1}H-spectrum (Figure 3.19) and a ^{1}H-^{15}N-spectrum (Figure 3.20). The initial spectra were obtained at pH=7.7 and showed low dispersion of signals in the N-H region (6.5-10 ppm), characteristic of a partially folded or unfolded protein (Figure 3.20, panel A). After pH adjustment, in order to improve the dispersion of the signals in the NH region (6.5 to 10 ppm), other ^{1}H-NMR spectra and ^{1}H, ^{15}N-HSQC spectra were obtained. The results showed that, as the pH of the samples decreased, the signals became more intense and dispersed, with the highest dispersion obtained at pH=6. The complete unfolding occurred for pH≃5, in accordance with the fact that GatD and MurT at pH=5.5,completely lost the enzymatic activity (Münch *et al.*, 2012b).

Every protein has an optimal pH for correct folding; in the case of DUF1727 domain, at pH=6 the protein becomes more folded, however this can be an artifact and can result from the fact that the central MurT domain is lacking or from interference of the His_6tag. Several studies showed that stability folding of the protein was strongly coupled to the ionization state of amino acids (protonated and deprotonated), being pH-dependent (Jamin, *et al*., 2001).

Although the pH decrease enhances the folding of DUF1727, the domain did not acquire the optimal folding state; this may have occurred due to lack of specific regions of MurT protein or to the absence of the partner, the GatD protein, which may be needed for the acquisition of the correct folding. The construct used to express DUF1727 domain includes the last 379 bps of the 3´terminal end of *murT* gene. Alignment between the DUF1727 amino acid sequence and the sequence of other Mur ligases identified a region of 45 amino acids residues between MurT central domain and DUF1727 (Figure 4.2) This small region may be essential for the complete folding of the DUF1727 domain. Interestingly, results from structural motifs search tool, COIL (ExPASy), a coiled coil motif was found encompassing the region from amino acid 290 to 320, which match the last half of the inter-region (Figure 5.4, annex). A coiled coil is a structural motif in proteins in which 2-7 alpha-helices are coiled together. Despite its simplicity, it is a highly versatile folding motif: coiled-coil-containing proteins exhibit a broad range of different functions related to the specific 'design' of their coiled-coil domains (Burkhard *et al.*, 2001).

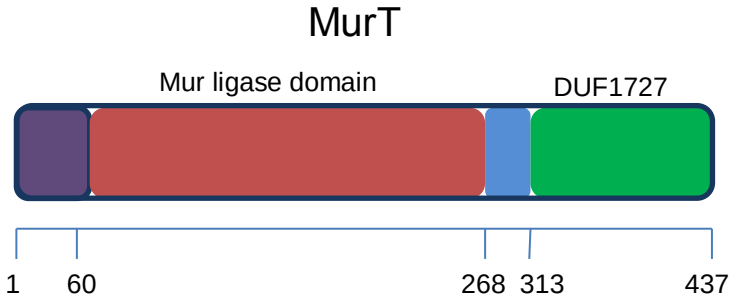

Figure 4.2 Domains of MurT protein. The only domain showing homology with Mur ligases is the central domain involved in nucleotide binding, were the conserved motifs for ATP and Mg^{2+}-binding are present. The C-terminal domain, DUF1727 was not yet characterized. A region of 45 amino acids (inter-region) separates de central domain and DUF1727.

4.8. Interaction studies of ^{15}N-DUF1727 with GatD by NMR

When GatD was added to DUF1727 domain in solution, the two proteins co-precipitated, indicating a strong protein-protein interaction. The co-precipitation was not expected and due to this event, the ^{1}H-^{13}N-spectrum for the interaction of DUF1727 with GatD was not obtained. A possible explanation is the fact that GatD protein is less stable when the pH decreases. In order to improve the folding of DUF1727 domain, the pH was lowered to 6.6. The decrease in GatD stability may have led to the co-precipitation of DUF1727, as the interaction occurred.

4.9. Conclusions and future perspectives

The main objective of this study was not fully reached, as the structure of the MurT-GatD complex was not elucidated. However, besides the insights into the role of the DUF1727 domain, this work provided important contributions for the production of the partner proteins for NMR analysis.

The initial strategy consisted in the production of uniformly ^{15}N,^{13}C, labeled sample of GatD protein and assignment of GatD ^{15}N,^{13}C-HSQC NMR spectra. The following step would be the analysis by NMR of the addition of MurT unlabeled protein to^{15}N, ^{13}C GatD. This step would provide information on which amino acid residues participate directly in the protein-protein interaction. The work developed in this Thesis allowed to re-design this initial strategy. In the new approach, the DUF1727, a smaller sized and more stable protein was used instead of GatD for the production of uniformly ^{15}N, ^{13}C-labeled partner. The MurT/GatD interaction will in this way be analysed using ^{15}N, ^{13}C DUF1727 and unlabeled GatD by NMR spectroscopy. Assignment of the ^{15}N, C^{13} HSQC NMR spectrum of labeled DUF1727 domain with GatD protein, will allow to identify the amino acid residues of DUF1727 that interact with GatD protein.

In order to apply the new workplan, some strategies are proposed to maximize the yield, purity level and stability of the partner proteins:

- Stability tests will be performed for GatD-His_6tag and DUF1727-His_6tag mixtures in order to identify the optimal buffer conditions for the interaction of the two proteins;
- Construction of DUF1727 and GatD fusion proteins with different tags used to promote solubilization, namely MBP (maltose binding protein);
- Expression of DUF1727 and GatD recombinant proteins using strains which express chaperone proteins to help the folding process;
- Construction of a new recombinant DUF1727 domain including the 45 amino acid residues located between the central domain of MurT and DUF1727. These extra residues may improve the stability and the folding of DUF1727;
- Construction of DUF1727 fusion proteins fused to a N-terminal His_6tag and removal with 3C protease.

References

Abraham, E. P., Chain, E. 1940. An enzyme from bacteria able to destroy penicillin. Nature. 146.

Andersen, K.R., Leksa, N.C., Schwartz, T.U. 2013. Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. Proteins. 81(11): 1857–61.

Bera, A., Biswas, R., Herbert, S., Götz, F. 2006. The presence of peptidoglycan O-acetyltransferase in various staphylococcal species correlates with lysozyme resistance and pathogenicity. Infect. Immun. 74(8): 4598–604.

Bera, A., Herbert, S., Jakob, A., Vollmer, W., Götz, F. 2005. Why are pathogenic staphylococci so lysozyme resistant? The peptidoglycan O-acetyltransferase OatA is the major determinant for lysozyme resistance of Staphylococcus aureus. Mol. Microbiol. 55(3): 778–87.

Berg, M. J., Tymozcko, L. J., Stryer, L. 2002. Three-Dimensional Protein Structure Can Be Determined by NMR Spectroscopy and X-Ray Crystallography. In Biochemistry (Freeman, W. H., eds). 5ª ed., chapter 4, International Edition, New York.

Berger-Bächi, B. 1983. Insertional inactivation of staphylococcal methicillin resistance by Tn551. J. Bacteriol. 154(1): 479–87.

Bernard, E., Rolain, T., Courtin, P., Hols, P., Chapot-Chartier, M.-P. 2011. Identification of the amidotransferase AsnB1 as being responsible for meso-diaminopimelic acid amidation in Lactobacillus plantarum peptidoglycan. J. Bacteriol. 193(22): 6323–30.

Bharti, S.K., Roy, R. 2012. Quantitative H NMR spectroscopy. Trends Anal. Chem. 35: 5–26.

Boucher, H., Miller, L.G., Razonable, R.R. 2010. Serious infections caused by methicillin-resistant Staphylococcus aureus. Clin. Infect. Dis. 51 Suppl 2(Suppl 2): S183–97.

Burkhard, P., Stetefeld, J., Strelkov, S. V, Burkhard, P. 2001. Coiled coils : a highly versatile protein folding motif.Trends Cell Biol. 11(2):82-8.

Dabul, A. N., Camargo, I. L. 2014. Molecular characterization of methicillin-resistant *Staphylococcus aureus* resistant to tigecycline and daptomycin isolated in a hospital in Brazil. Epidemiol. Infect. 142: 479-83.

Dyke, K.G., Jevons, M. P., and Parker, M. T. 1966. Penicillinase production and intrinsic resistance to penicillins in *Staphylococcus aureus*. Lancet. 1:835-8.

de Lencastre, H., Wu, S. W., Pinho, M. G., Ludovice, A. M., Filipe, S., Gardete, S., Sobral, R., Gill, S., Chung, M., and 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. 5:163-75.

de Lencastre, H., Oliveira, D., Tomasz, A. 2007. Antibiotic resistant Staphylococcus aureus: a paradigm of adaptive power. Curr. Opin. Microbiol. 10(5): 428–35.

Deurenberg, R.H., Vink, C., Kalenic, S., Friedrich, W., Bruggeman, C., Stobberingh, E.E. 2007. The molecular evolution of methicillin-resistant Staphylococcus aureus. Clin. Microbiol. Infect. 13(3): 222–35.

Höltje, J. 1998. Growth of the Stress-Bearing and Shape-Maintaining Murein Sacculus of Escherichia coli.Microbiol. Mol. Biol. Rev. 62(1): 181–203.

Figueiredo, T. A., Ludovice, A.M., Sobral, R.G. 2014. Contribution of peptidoglycan amidation to beta-lactam and lysozyme resistance in different genetic lineages of Staphylococcus aureus. Microb. Drug Resist. 20(3): 238–49.

Figueiredo, T. A, Sobral, R.G., Ludovice, A.M., Almeida, J.M.F. De, Bui, N.K., Vollmer, W., de Lencastre, H., Tomasz, A. 2012. Identification of genetic determinants and enzymes involved with the amidation of glutamic acid residues in the peptidoglycan of Staphylococcus aureus. PLoS Pathog. 8(1): e1002508.

Fisher, J.F., Meroueh, S.O., Mobashery, S. 2005. Bacterial resistance to beta-lactam antibiotics: compelling opportunism, compelling opportunity. Chem. Rev. 105(2): 395–424.

Fukumoto, A., Kim, Y., Hanaki, H., Shiomi, K., Tomoda, H. 2008. Cyslabdan, a new potentiator of Imipenem activity against Methicillin-resistant Staphylococcus aureus, produced by Streptomyces sp. Biological activities. J. Antibiot (Tokyo). 61(1): 7–10.

Fleming, A. 1929. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. B. J. Exp. Path. 10: 226-236.

Foster, T.J. 2005. Immune evasion by staphylococci. Nat. Rev. Microbiol. 3(12): 948–58.

Gally, D., Archibald, R. 1993. Cell wall assembly in Staphylococcus aureus: proposed absence of secondary crosslinking reactions. J. Gen. Microbiol. 139(8): 1907–13.

Gatzeva-topalova, P.Z., May, A.P., Sousa, M.C. 2005. Structure and Mechanism of ArnA : Conformational Change Implies Ordered Dehydrogenase Mechanism in Key Enzyme for Polymyxin Resistance. Structure. 13(6):929-42.

Georgopapadakou, N.H., Liu, F.Y. 1980. Binding of beta-lactam antibiotics to penicillin-binding proteins of Staphylococcus aureus and Streptococcus faecalis: relation to antibacterial activity. Antimicrob. Agents Chemother. 18(5): 834–6.

Ghuysen, J. M. 1968. Serine β-lactamases and penicillin-binding proteins. Annu. Rev. Microbiol. 45:37-67

Goffin, C., Ghuysen, J. 1998. Multimodular Penicillin-Binding Proteins : An Enigmatic Family of Orthologs and Paralogs. Microbiol. Mol. Biol. Rev. 62(4): 1079–1093.

Grundmann, H., Aires-de-Sousa, M., Boyce, J., Tiemersma, E. 2006. Emergence and resurgence of meticillin-resistant Staphylococcus aureus as a public-health threat. Lancet. 368(9538): 874–85.

Gu, B., Kelesidis, T., Tsiodras, S., Hindler, J., Humphries, R.M. 2013. The emerging problem of linezolid-resistant Staphylococcus. J. Antimicrob. Chemother. 68(1): 4–11.

Gustafson, J., Strassle, A., Hachler, H., Kayser, F.H., Berger-bachi, B. (1994). The femC locus ofStaphylococcus aureus required for methicillin resistance includes the Glutamine Synthetase Operon. J. Bacteriol. 176(5): 1460–1467.

Heatley, N.G. (1940). Symptoms relating. 226–228.

Hegde, S.S., Shrader, T.E. 2001. FemABX family members are novel nonribosomal peptidyltransferases and important pathogen-specific drug targets. J. Biol. Chem. 276(10): 6998–7003.

Heijenoort, J. Van. 2001. Recent advances in the formation of the bacterial peptidoglycan monomer unit (1985 to 2000). Nat. Prod. Rep. 18(5): 503–519.

Höltje, J. 1998. Growth of the Stress-Bearing and Shape-Maintaining Murein Sacculus of Escherichia coli.Microbiol. Mol. Biol. Rev. 62(1): 181–203.

Huber, J., Donald, R.G.K., Lee, S.H., Jarantow, L.W., Salvatore, M.J., Meng, X., Painter, R., Onishi, R.H., Occi, J., Dorso, K., Young, K., Park, Y.W., Skwish, S., Szymonifka, M.J., Waddell, T.S., Miesel, L., Phillips, J.W., Roemer, T. 2009. Chemical genetic identification of peptidoglycan inhibitors potentiating carbapenem activity against methicillin-resistant Staphylococcus aureus. Chem. Biol. 16(8): 837–48.

Jamin, M. , Geierstanger B., Baldwin RL. 2001. The pKa of His-24 in the folding transition state of apomyoglobin. PNAS. 98(11):6127-31.

Jevons, M.P. 1961. Celbenin-resistant staphylococci. Br. Med. J. 2:124-33.

Jolly, L., Wu, S., Heijenoort, J.V.A.N., de Lencastre, H. 1997. The femR315 gene from Staphylococcus aureus, the interruption of which results in reduced methicillin resistance, encodes a phosphoglucosamine mutase. J. Bacteriol. 179(17): 5321–5325.

Jonge, B.L., Sidow, T., Chang, Y., Labischinski, H., Berger-bachi, B., Gage, D.A., Tomasz, A. 1993. Altered Muropeptide Composition in Staphylococcus aureus Strains with an Inactivated femA Locus. J. Bacteriol. 175(9): 2779–2782.

Jonge, B.L., Chang, Y., Gage, D. 1992. Peptidoglycan composition in heterogeneous Tn551 mutants of a methicillin-resistant Staphylococcus aureus strain. J. Biol.Chem. 267(16): 11255–11259.

Lebar, M.D., May, J.M., Meeske, A.J., Leiman, S. a, Lupoli, T.J., Tsukamoto, H., Losick, R., Rudner, D.Z., Walker, S., Kahne, D. 2014. Reconstitution of peptidoglycan cross-linking leads to improved fluorescent probes of cell wall synthesis. J. Am. Chem. Soc. 136(31): 10874–7.

Lee, J.H. 2003. Methicillin (Oxacillin)-resistant *Staphylococcus aureus* strains isolated from major food animals and their potential transmission to humans. Appl. Environ. Microbiol. 69(11):6489-94

Lee, S.H., Jarantow, L.W., Wang, H., Sillaots, S., Cheng, H., Meredith, T.C., Thompson, J., Roemer, T. 2011. Antagonism of chemical genetic interaction networks resensitize MRSA to β-lactam antibiotics. Chem. Biol. 18(11): 1379–89.

Lepre, C. A, Moore, J.M. 1998. Microdrop screening: a rapid method to optimize solvent conditions for NMR spectroscopy of proteins. J. Biomol. NMR. 12(4): 493–9.

Levy, O. 2000. Review article Antimicrobial proteins and peptides of blood : templates for novel antimicrobial agents. Blood. 96(8): 2664-72.

Lowy, F.D. 2003. Antimicrobial resistance : the example of Staphylococcus aureus. J. Clin. Invest. 111(9): 1265–1273.

Mandal, S., Moudgil, M., Mandal, S.K. 2009. Rational drug design. Eur. J. Pharmacol. 625(1-3): 90–100.

Massière, F., Badet-Denisot, M.A. 1998. The mechanism of glutamine-dependent amidotransferases. Cell Mol. Life Sci. 54(3):205–222.

Mohammadi, T., van Dam, V., Sijbrandi, R., Vernet, T., Zapun, A., Bouhss, A., Diepeveen-de Bruin, M., Nguyen-Distèche, M., de Kruijff, B., Breukink, E. 2011. Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. EMBO J. 30(8): 1425–32.

Montelione, G.T., Zheng, D., Huang, Y.J., Gunsalus, K.C., Szyperski, T. 2000. Protein NMR spectroscopy in structural genomics. Nat. Struct. Biol. 7: 982-5.

Münch, D., Roemer, T., Lee, S.H., Engeser, M., Sahl, H.G., Schneider, T. (2012a). Identification and in vitro analysis of the GatD/MurT enzyme-complex catalyzing lipid II amidation in Staphylococcus aureus. PLoS Pathog. 8(1): e1002509.

Münch, D., Roemer, T., Lee, S.H., Engeser, M., Sahl, H.G., Schneider, T. (2012b). Identification and in vitro analysis of the GatD/MurT enzyme-complex catalyzing lipid II amidation in Staphylococcus aureus. PLoS Pathog. 8(1): e1002509.

Murakami, K., Tomasz, A. 1989. Involvement of multiple genetic determinants in high-level methicillin resistance in Staphylococcus aureus. J. Bacteriol. 171(2): 874–9.

Navratna, V., Nadig, S., Sood, V., Prasad, K., Arakere, G., Gopal, B. 2010. Molecular basis for the role of Staphylococcus aureus penicillin binding protein 4 in antimicrobial resistance. J. Bacteriol. 192(1): 134–44.

Ogston, A. 1883. Micrococcus poisoning. J. Anat. Physiol. 17: 24–58.

On, B., Structural, T., Tipper, B.Y.D.J., Strominger, J.L. (1966). BASED ON THEIR STRUCTURAL SIMILARITY TO ACYL-D-ALANYL-D-ALANINE *. 54(1964).

Peacock, S.J., de Silva, I., Lowy, F.D. (2001). What determines nasal carriage of Staphylococcus aureus? Trends Microbiol. 9(12): 605–10.

Pereira, S.F.F., Henriques, a O., Pinho, M.G., de Lencastre, H., Tomasz, A. 2007. Role of PBP1 in cell division of Staphylococcus aureus. J. Bacteriol. 189(9): 3525–31.

Pinho, M.G., 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(19): 10886–91.

Pinho, M.G., Lencastre, H. de, Tomasz, A. 2000. Characterization and Inactivation of the gene pbpC, encoding Penicillin-Binding Protein 3 of *Staphylococcus aureus*. J.Bacteriol. 182(4):1074-9.

Projan, S. J., Novick, R.P. 1997. The molecular basis of pathogenicity, p. 55-81. In K. B. Crossley and G. L. Archer (ed.), The staphylococci in human disease. Churchill Livingstone Inc., New York.

Chang, S., Sievert, D.M., Hageman, J.C,, Boulton, M.L., Tenover, F.C., Downes, F.P., Shah, S., Rudrik, J.T., Pupp, G.R., Brown, W.J., Cardo, D., Fridkin, S.K. 2003. Infection with vancomycin-resistant *Staphylococcus aureus* containing the *vanA* resistance gene. N Engl J Med. 3;348(14):1342-7.

Reynolds, P.E., Brown, D.F. (1985). Penicillin-binding proteins of beta-lactam-resistant strains of Staphylococcus aureus. Effect of growth conditions. FEBS Lett. 192(1): 28–32.

Rolinson, G. N., Stevens, S., Batchelor, F. R., Wood, J. C., and Berger-Bachi, E. B. 1960. Bacteriological studies on a new penicillin-BRL.1241. Lancet. 2:564-7

Rosenbach. 1884. Microorganisms of wound infections. Wiesbaden. Quoted by Fleming, A. in “A system of bacteriology”.1929. Medical Research Council. 2: 11-28.

Serber, Z., Ledwidge, R., Miller, S. M., Do, V. 2001. Evaluation of Parameters Critical to Observing Proteins Inside Living *Escherichia coli* by In-Cell NMR Spectroscopy. J. Am. Chem. Soc. 19;123(37):8895-901.

Scheffers, D., Pinho, M.G. 2005. Bacterial Cell Wall Synthesis : New Insights from Localization Studies.Microbiol. Mol. Biol. Rev. 69(4):585-607.

Schindler, M., Assaf, Y., Sharon, N., Chipman, D.M. 1997. Mechanism of lysozyme catalysis: Role of ground-state strain in subsite D in hen egg-white and human lysozymes. Biochemistry. 16: 423-431.

Schleifer, K.H., Kandler, O. 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol. Rev. 36(4): 407–77.

Schmitt, J., Hess, H., Stunnenberg, H.G. 1993. Affinity purification of histidine-tagged proteins. Mol. Biol. Rep.18(3):223-30.

Schneider, T., Senn, M.M., 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(2): 675–85.

Serber, Z., Ledwidge, R., Miller, S.M., Do, V. 2001. Evaluation of Parameters Critical to Observing Proteins Inside Living *Escherichia coli* by In-Cell NMR Spectroscopy. (10): 8895–8901.

Smith, C. A. 2006. Structure, function and dynamics in the mur family of bacterial cell wall ligases. J. Mol. Biol. 362(4): 640–55.

Sobral, R.G., Jones, A.E., Des Etages, S.G., Dougherty, T.J., Peitzsch, R.M., Gaasterland, T., Ludovice, A.M., de Lencastre, H., Tomasz, A. 2007. Extensive and genome-wide changes in the transcription profile of Staphylococcus aureus induced by modulating the transcription of the cell wall synthesis gene murF. J. Bacteriol. 189(6): 2376–91.

Solberg, C. O. 1965. A study of carriers of *Staphylococcus aureus* with special regard to quantitative bacterial estimations. Acta Med Scand Suppl. 436: 1–96.

Tomasz, A., Gardete, S., Ludovice, A.M., Sobral, R.G., Filipe, S.R., Lencastre, H. de. 2004. Role of *murE* in the Expression of β -Lactam Antibiotic Resistance in *Staphylococcus aureus*. J. Bacteriol. 186(6):1705-13.

Tsubakishita, S., Kuwahara-Arai, K., Sasaki, T., Hiramatsu, K. 2010. Origin and molecular evolution of the determinant of methicillin resistance in staphylococci. Antimicrob. Agents Chemother. 54(10): 4352–9.

Vagenende, V., Yap, M.G.S., Trout, B.L. 2009. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. Biochemistry. 48(46): 11084–96.

Van Heijenoort, J. 2007. Lipid intermediates in the biosynthesis of bacterial peptidoglycan. Microbiol. Mol. Biol. Rev. 71(4): 620–35.

Van Bambeke, F., Lambert, D. M., Mingeot-Leclercq, M. P., Tulkens, P. M. 2003. Anti-infective therapy: Mechanism of action. *In* Infectious Diseases (Armstrong, J Cohen ed.) 7.1.1-7.1.14 London, United Kingdom.

Vandenbergh, M.F.Q., Verbrugh, H. A. 1996. From the chicago meetings. 525–534.

Veiga, P., Erkelenz, M., Bernard, E., Kulakauskas, S., Crossbridge, I., Courtin, P. 2009. Identification of the Asparagine Synthase Responsible for d-Asp Amidation in the *Lactococcus lactis* Peptidoglycan Interpeptide Crossbridge. J. Bacteriol. 191(11):3752-7.

Vollmer, W., Blanot, D., de Pedro, M. A. 2008. Peptidoglycan structure and architecture. FEMS Microbiol. Rev. 32(2): 149–67.

Vieira, V., Figueiredo, T. A. Verma, A., Sobral, R.G., Ludovice, A. M., de Lencastre, H. and Trincão, J. 2014. Purification, crystallization and preliminary X-ray diffraction analysis of GatD, a glutamine

amidotransferase-like protein from *Staphylococcus aureus* peptidoglycan. Acta Crystallographica Section F. 70:632-5.

Weber, D.J., Gittis, A.G., Mullen, G.P., Abeygunawardana, C., Lattman, E.E., and Mildvan, A.S. 1992. NMR docking of a substrate into the X-ray structure of staphylococcal nuclease. Proteins. 4, 275-287

Weininger, U., Haupt, C., Schweimer, K., Graubner, W., Kovermann, M., Brüser, T., Scholz, C., Schaarschmidt, P., Zoldak, G., Schmid, F.X., Balbach, J. 2009. NMR Solution Structure of SlyD from *Escherichia coli* : Spatial Separation of Prolyl Isomerase and Chaperone Function. J. Mol. Biol. 387(2): 295–305.

Wertheim, H.F.L., Melles, D.C., Vos, M.C., van Leeuwen, W., van Belkum, A., Verbrugh, H. a, Nouwen, J.L. 2005. The role of nasal carriage in *Staphylococcus aureus* infections. Lancet. Infect. Dis. 5(12): 751–62.

Wilkinson, B.J.1997. Biology, p.3 and 35. *In* K. B. Crossley and G. L. Archer ed., The staphylococci in human disease. Churchill Livingstone Inc., New York.

William, D., Home, S.M., Young, K.D., Younh, R. 1994. sZyD , a Host Gene Required for 4x174 Lysis , Is Related to the FKBO6-binding Protein Family of Peptidyl-prolyl. 269(4): 2902–2910.

Williams, G.J., Breazeale, S.D., Raetz, C.R.H., Naismith, J.H. 2005. Structure and function of both domains of ArnA, a dual function decarboxylase and a formyltransferase, involved in 4-amino-4-deoxy-L-arabinose biosynthesis. *J. Biol. Chem.* 280(24): 23000–8.

Williams, R.E. 1963. Healthy carriage of *Staphylococcus aureus*: its prevalence and importance. Bacteriol. Rev.27:56-71.

Yocum, R.R., Waxman, D.J., Rasmussent, J.R., Stromincser, J.L. 1979. Mechanism of penicillin action : Penicillin and substrate bind covalently to the same active site serine in two bacterial D-alanine carboxypeptidases. Biochemistry : 76(6): 2730–2734.

Yu, H. 1999. Extending the size limit of protein nuclear magnetic resonance. Proc. Natl. Acad. Sci. USA.96(2):332-4.

Zapun, A., Philippe, J., Abrahams, K. A, Signor, L., Roper, D.I., Breukink, E., Vernet, T. 2013. *In vitro* reconstitution of peptidoglycan assembly from the Gram-positive pathogen *Streptococcus pneumoniae*. ACS Chem. Biol. 8(12): 2688–96.

Annex

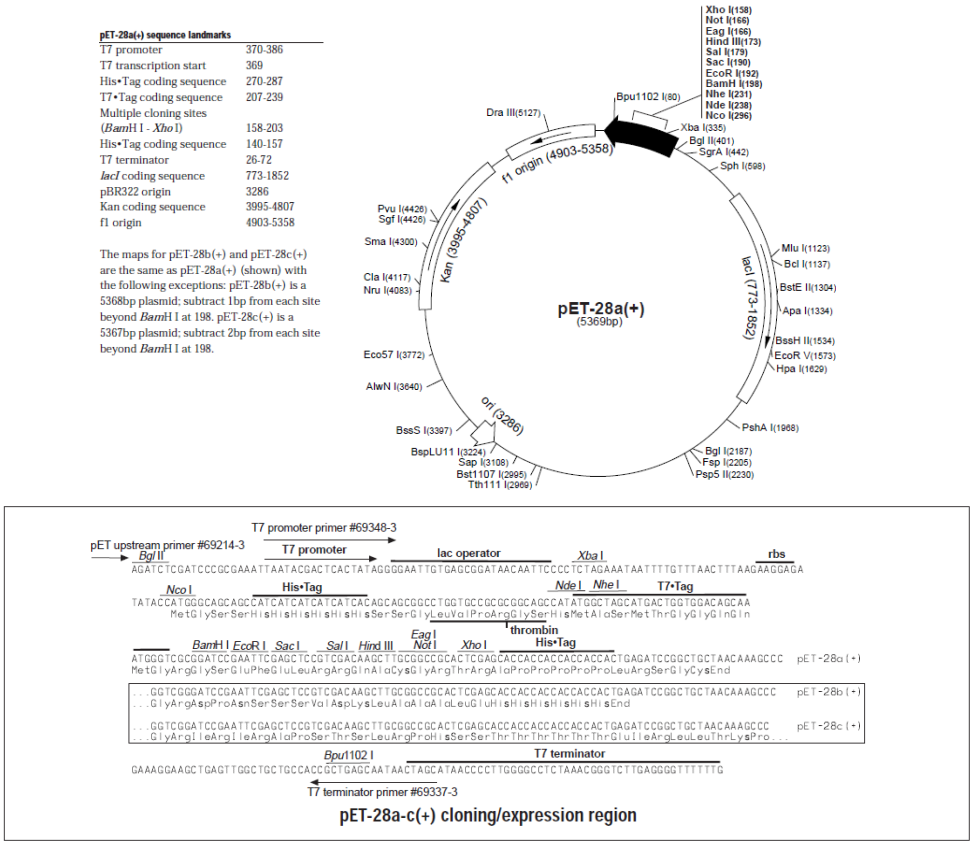

Figure 5.1. Plasmid map of pET28a. The sequence of multiple cloning sites was amplified.

Table 5.1. Isoelectric point and molecular weight of the different construct used, as also the protein without tag.

Protein	MW	pI	Ex. Coefficient (M^{-1} cm^{-1})
Histag-GatD	29.09	6.01	16390
GatD	27.43	5.46	-
GatD-His$_6$tag	28.25	5.90	-
MurTWDUF1727	35.41	8.20	-
MurT-His$_6$tag	50.72	6.41	-
DUF1727-His$_6$tag	15.33	5.47	12950
His$_6$tag-DUF1727	15.47	5.64	12950
DUF1727	13.81	4.66	

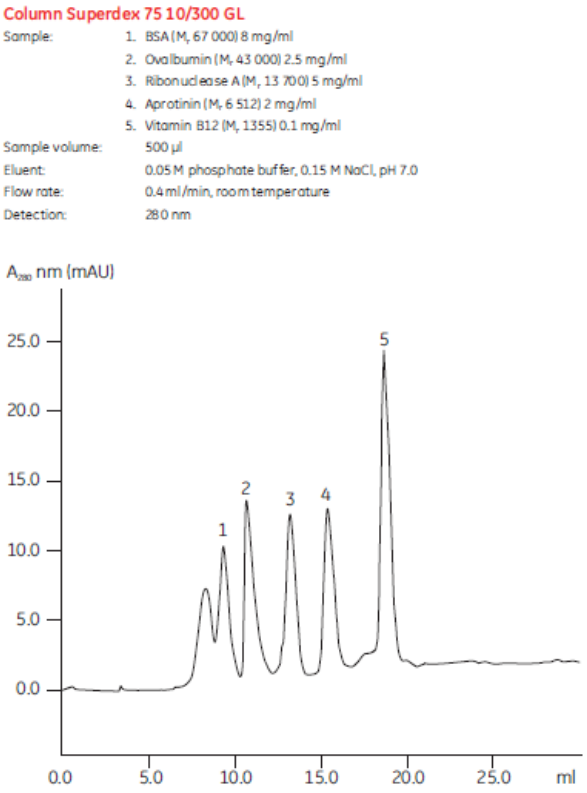

Figure 5.2. Typical chromatogram from a function test of Superdex 75 10/300 GL. Each number correspond a one protein eluted with different molecular size. 1- BSA (M_r 67000) 6mg/ml; 2- Ovalbumin (M_r 43 000) 2.5 mg/ml; 3- Ribonuclease A (M_r 13 700) 5 mg/ml; 4- Aprotinin (M_r 6 512) 2 mg/ml; 5- Vitamin B12 (M_r 1355) 0.1 mg/ml.

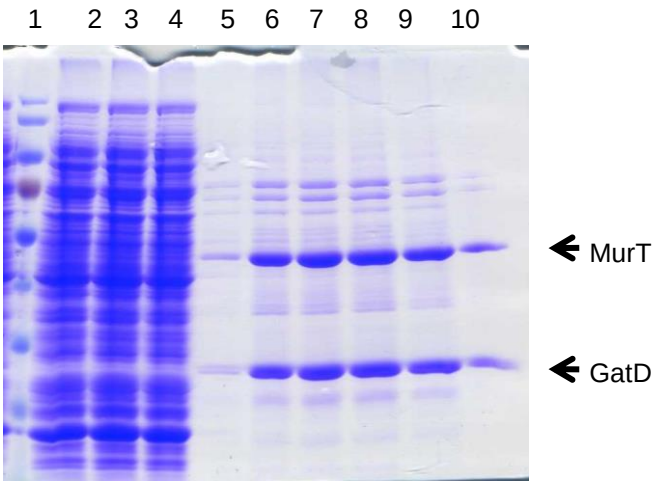

Figure 5.3. Affinity purification of MurT and GatD protein from MurT-GatD-His_6tag. Lane 1: protein marker; Lane 2: applied sample; Lane 3: flow through; Lane 4: Wash 1; Lane 5: Wash 10; Lane 6: elution 1; Lane 7: elution 2; Lane 8: elution 3; Lane 9: elution 4; Lane 10: elution 5. Figueiredo*, et al.,* unpublished.

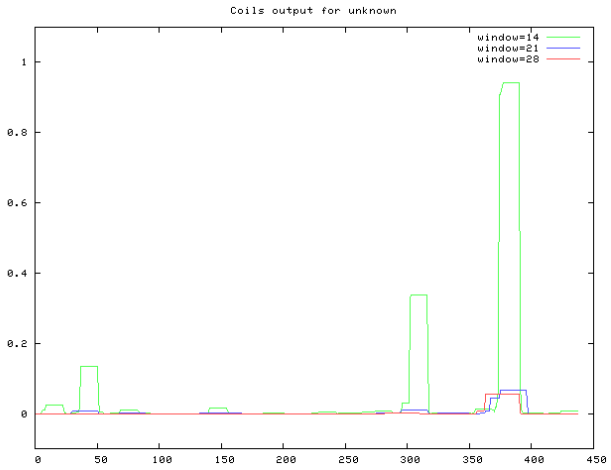

Figure 5.4. Coiled coils found in MurT sequence, through tool Coil (ExPASy), using windows of 14, 21 and 28 amino acid. In this prediction, 3 coils are found.