

STRUCTURAL AND BIOCHEMICAL
CHARACTERIZATION OF AUTOLYSINS
FROM *STAPHYLOCOCCUS AUREUS* MU50

Sara Pintar

Doctoral Dissertation
Jožef Stefan International Postgraduate School
Ljubljana, Slovenia

Supervisor: Prof. Dr. Dušan Turk, Jožef Stefan Institute, Ljubljana, Slovenia

Evaluation Board:

Prof. Dr. Marko Anderluh, Chair, Faculty of Pharmacy, University of Ljubljana,
Ljubljana, Slovenia

Prof. Dr. Miquel Coll, Member, Institute for Research in Biomedicine, Barcelona, Spain

Prof. Dr. Nataša Poklar Ulrih, Member, Biotechnical Faculty, University of Ljubljana,
Ljubljana, Slovenia

MEDNARODNA PODIPLOMSKA ŠOLA JOŽEFA STEFANA
JOŽEF STEFAN INTERNATIONAL POSTGRADUATE SCHOOL



Sara Pintar

STRUCTURAL AND BIOCHEMICAL
CHARACTERIZATION OF AUTOLYSINS FROM
STAPHYLOCOCCUS AUREUS MU50
Doctoral Dissertation

STRUKTURNA IN BIOKEMIJSKA KARAKTERIZACIJA
AVTOLIZINOV IZ *STAPHYLOCOCCUS AUREUS* MU50

Doktorska disertacija

Supervisor: Prof. Dr. Dušan Turk

Ljubljana, Slovenia, October 2018

Acknowledgments

I thank Slovenian Research Agency for funding my doctoral studies.

I would also like to thank my supervisor, Prof. Dr. Dušan Turk, for accepting me into his research group and ensuring good working conditions that enabled me to perform all the necessary experiments. Additionally, I would like to thank him for the supervision of my dissertation and all the advice with the work.

I would like to thank the members of my evaluation board, Prof. Dr. Marko Anderluh, Prof. Dr. Miquel Coll, and Prof. Dr. Nataša Poklar Ulrih, for lending their expertise in the revision of my dissertation.

Further I need to thank my collaborators Dr. Jure Borišek and Dr. Andrej Perdih for doing all the computational work involved in the dissertation, as well as for all the help with presenting of the results and publishing.

I would like to thank Dr. Aleksandra Usenik for helping me with experiments and all the advice during the course of my work.

I thank Prof. Dr. Dušan Kordiš for helping me with bioinformatic analysis.

I am thankful to Dr. Matic Lozinšek and Dr. Blaž Alič for helping me with challenging experiments that were beyond the scope of biochemistry.

I thank Prof. Dr. Ines Mandić-Mulec for generously providing me the *B. subtilis* strain.

I would like to thank my research group, Dr. Marko Mihelič, Dr. Miha Renko, Dr. Piotr Sosnowski, Dr. Ajda Verčič-Taler, Andreja Sekirnik, Dr. Nataša Lindič and others, as well as all of the Department B1 for all their help and advice.

Special thanks goes to my family and friends for all their support during the course of the study.

Abstract

The bacteria are surrounded by a cell wall that provides them with mechanical protection. The cell wall consists of a peptidoglycan scaffold, built from polysaccharide strands cross-linked by peptides. Autolysins are enzymes that remodel the cell wall during growth and division. Among them are N-acetylglucosaminidases that cleave the glycosidic bonds between NAG and NAM monosaccharide residues. Here we focused on N-acetylglucosaminidases of the *Staphylococcus aureus*, a pathogenic microorganism that represents a major public health burden.

We have identified four peptidoglycan N-acetylglucosaminidases that are highly conserved in *S. aureus*, SagB, AtlA-gl, AtlE and ScaH-gl. Crystal structures have revealed that the enzymes have a highly similar secondary structure organization and markedly different tertiary structure. After the analysis and comparison of their substrate binding sites we proposed that for the enzymatic reaction to proceed, a substantial conformational change needs to take place. This leads to closing of the enzyme. Molecular dynamics simulations have confirmed that closed conformations of these enzymes are stable, but only when the substrate is bound. This indicates that the conformational change occurs with binding of the substrate.

Study of the catalytic mechanism of AtlE has shown that for of the reaction only one catalytic amino acid is necessary. As the energetically favorable path, water-assisted catalysis was established. Additionally, we proposed the transition state of the reaction is represented by oxocarbenium ion.

In addition, we have designed the first compounds that AtlE binds. Compounds from the structural class of (phenylureido)piperidinyl benzamides were found to exhibit promising binding affinity in the micromolar range.

The results we present here are a basis for further drug target validation of *S. aureus* N-acetylglucosaminidases as drug targets.

Povzetek

Celična stena obdaja bakterijske celice in jim daje mehansko zaščito. Sestoji iz peptidoglikanske mreže, zgrajene iz polisaharidnih verig, ki so prečno povezane s peptidi. Avtolizini jo mestoma razgradijo, da omogočijo rast in delitev celic. Med to vrsto encimov sodijo tudi N-acetilglukozaminidaze, ki cepijo glikozidno vez med NAG in NAM monosaharidnima ostankoma. Mi smo se osredotočili na patogeni organizem *Staphylococcus aureus*, ki predstavlja velik javnozdravstveni problem.

Identificirali smo štiri peptidoglikan N-acetilglukozaminidaze, ki so visoko ohranjene pri preučevani vrsti, SagB, AtlA-gl, AtlE in ScaH. Določitev struktur je razkrila, da imajo encimi zelo podobno sekundarno in obenem drugačno terciarno strukturo. Analiza in primerjava njihovih vezavnih mest za substrat je pokazala, da je za to, da reakcija lahko poteče, potrebna velika konformacijska sprememba, ki vodi v zapiranje encima. Simulacije molekularne dinamike so potrdile, da so zaprte konformacije encimov stabilne, vendar le z vezanim substratom. To kaže na to, da do konformacijske spremembe pride z vezavo substrata.

Med delom smo preučevali tudi katalitični mehanizem AtlE. Študija je pokazala, da reakcija lahko poteče z eno katalitično aminokislino. Bolj energetsko ugodno pot predstavlja kataliza z nukleofilnim napadom vode. Dodatno smo pokazali, da je najbolj verjetno prehodno stanje oksokarbenijev ion.

Glede na strukturo AtlE smo oblikovali prve spojine, ki jih ta encim veže. (fenilureido)piperidinil benazmidi so izkazali afiniteto do encima v mikromolarnem območju.

Rezultati, ki jih predstavljamo tukaj, so osnova za nadaljnjo validacijo N-acetilglukozaminidaz iz *S. aureus* kot tarč za razvoj novih zdravil.

Contents

List of Figures	xv
List of Tables	xvii
Abbreviations	xix
Symbols	xxi
Standard Amino Acid Abbreviations	xxiii
1 Introduction	1
1.1 Antibiotic Resistance	1
1.2 <i>Staphylococcus aureus</i>	2
1.3 Cell Wall.....	2
1.4 Autolysins	3
1.5 N-Acetyl- β -D-glucosaminidases.....	4
1.6 Structures of GH73 Family of Enzymes.....	4
1.6.1 Substrate binding site.....	5
1.6.2 Reaction mechanism.....	6
2 Purpose of the Work	9
3 Materials and Methods	11
3.1 Materials	11
3.1.1 Equipment.....	11
3.1.2 Chemicals, commercial kits and chromatography columns.....	11
3.1.3 Buffers, solutions and media.....	13
3.1.4 Cloning primers	17
3.1.5 Mutagenesis primers	18
3.1.6 Plasmids	20
3.1.7 Bacterial cultures.....	20
3.1.8 Recombinant enzymes	20
3.2 Methods	20
3.2.1 Protein sequence analysis	20
3.2.1 Sequence analysis of <i>S. aureus</i> GH73 enzymes.....	21
3.2.2 Cloning.....	21
3.2.2.1 Construct DNA sequence amplification from genomic DNA	21
3.2.2.2 T4 DNA polymerase treatment	22
3.2.2.3 Plasmid amplification and isolation.....	22
3.2.2.4 Plasmid restriction and T4 DNA polymerase treatment	23
3.2.2.5 Ligation	23
3.2.2.6 Construction of mutants.....	24

3.2.3	Protein Production and Purification.....	24
3.2.3.1	Testing the solubility and expression levels of different constructs.....	24
3.2.3.2	Large scale protein production and purification.....	24
3.2.3.2.1	Se-Met derivative large scale production	25
3.2.4	DSF.....	25
3.2.4.1	Screen of conditions of optimal stability for different proteins	25
3.2.4.2	Determination of the stability of mutant proteins	27
3.2.5	Crystallization.....	27
3.2.6	Structure determination.....	27
3.2.7	Substrate preparation	28
3.2.8	Activity assay	28
3.2.9	Computational simulations	28
3.2.9.1	Simulation of protein flexibility	28
3.2.9.2	Simulations involving catalytic mechanism elucidation	29
3.2.9.2.1	Molecular Dynamics of the AtIE Remodeled Complex with (NAG-NAM) ₃	29
3.2.9.2.2	Definition of the QM Region for the QM/MM Simulations	29
3.2.9.2.3	Reactants, Products, and Oxazolinium Ion Generation for QM/MM Simulations	30
3.2.9.2.4	Calculation of the Reaction A and B Pathways Using the Replica Path Method.....	30
3.2.9.3	Design of AtIE inhibitors	31
3.2.9.3.1	Surface plasmon resonance (SPR) experiments	31
3.2.9.3.2	Nuclear magnetic resonance (NMR) experiments	32
4	Results	33
4.1	Protein Sequence Analysis.....	33
4.2	Sequence Analysis of GH73 Enzymes from <i>S. aureus</i>	33
4.3	Cloning.....	34
4.3.1	Different protein constructs cloning.....	34
4.3.2	Mutant construction	35
4.4	Protein Production and Purification	35
4.4.1	Production of different constructs.....	35
4.4.2	Production of mutant proteins.....	37
4.5	DSF	37
4.5.1	Wild type proteins	37
4.5.1.1	SagB.....	37
4.5.1.2	SAV0909-gl	38
4.5.1.3	AtIA-gl.....	38
4.5.1.4	ScaH-gl.....	39
4.5.1.5	AtIE	39
4.5.2	Mutant proteins.....	40
4.5.2.1	SagB.....	40
4.5.2.2	AtIA-gl.....	40
4.5.2.3	AtIE	40
4.6	Crystallization	41
4.6.1	SagB.....	41
4.6.2	AtIA-gl.....	42
4.6.3	ScaH.....	42

4.7	Data Collection and Refinement Statistics	43
4.8	Description of the Structures	44
4.8.1	SagB	44
4.8.2	AtlA-gl	44
4.8.3	Comparison of the <i>S. aureus</i> GH73 structures	45
4.8.4	Substrate binding site.....	46
4.9	Activity Assay of the Mutant proteins	48
4.9.1	SagB	48
4.9.2	AtlA-gl	49
4.9.3	AtIE.....	50
4.10	Simulation of Flexibility	51
4.10.1	SagB	51
4.10.2	AtlA-gl	52
4.11	Catalytic Mechanism	53
4.11.1	Exploring the possibility of the oxazolinium ion-like transition state in AtIE	53
4.11.2	Calculation of the Possible Reaction Mechanisms.....	55
4.12	AtIE Inhibitors.....	57
5	Discussion	63
5.1	Structures Indicate High Conformational Flexibility of All of the Conserved Peptidoglycan <i>N</i> -acetylglucosaminidases of <i>S. aureus</i>	63
5.2	A Water-Assisted Catalytic Mechanism in Glycoside Hydrolases Demonstrated on AtIE	65
5.3	Discovery of (Phenylureido)PiperidinyI Benzamides as Prospective Inhibitors of Bacterial Autolysin E from <i>S. aureus</i>	67
6	Conclusions	69
	References	71
	Bibliography	77
	Biography	79

List of Figures

Figure 1: Illustration of the antibiotic discovery void.....	1
Figure 2: Primary structure of <i>S. aureus</i> peptidoglycan.....	3
Figure 3: The structure of AtlE.....	5
Figure 4: Comparison of the substrate binding of <i>N</i> -acetylglucosaminidase and muramidase	5
Figure 5: Enzymatic glycoside bond hydrolysis mechanisms.	6
Figure 6: Plasmid map of pMCSG7.....	20
Figure 7: Domain organization of proteins from GH73 family of <i>S. aureus</i> Mu50.....	33
Figure 8: A phylogenetic tree of all the GH73 enzymes from <i>S. aureus</i> species.	34
Figure 9: 1 % AGE gel depicting amplified DNA, coding for appropriate constructs	35
Figure 10: 1 % AGE gel depicting amplified DNA, coding for appropriate constructs.....	35
Figure 11: Purified SagB	36
Figure 12: Purified SAV0909-gl	36
Figure 13: Purified ScaH.....	36
Figure 14: Purified ScaH-gl.....	36
Figure 15: Purified ScaH-CHAP	36
Figure 16: Purified AtlE.....	36
Figure 17: Purified AtlA-gl.....	36
Figure 18: Purified AtlA-gl1	36
Figure 19: Purified AtlA-am.....	36
Figure 20: Thermal stability of SagB	37
Figure 21: Thermal stability of SAV0909-gl.....	38
Figure 22: Thermal stability of AtlA-gl+1CWB	38
Figure 23: Thermal stability of ScaH-gl.....	39
Figure 24: Thermal stability of AtlE	39
Figure 25: Thermal stability of SagB mutants relative to wild type protein.....	40
Figure 26: Thermal stability of AtlA-gl mutants relative to wild type protein.	40
Figure 27: Thermal stability of AtlE mutants relative to wild type protein.	41
Figure 28: SagB crystals.....	41
Figure 29: AtlA-gl+1CWB crystal.....	42
Figure 30: ScaH crystals.....	42
Figure 31: Structures of the GH73.....	45
Figure 32: The residues studied by point mutations.....	47
Figure 33: Binding of the substrates to the <i>S. aureus</i> enzymes of GH73 family.....	48
Figure 34: Activity of SagB mutants.....	48
Figure 35: Activity of AtlA-gl mutants	49
Figure 36: Activity of AtlE mutants	50
Figure 37: SagB flexibility simulations	52
Figure 38: AtlA-gl flexibility simulations.....	53

Figure 39: An optimized energy reaction profile for the oxazolinium ion-to-product pathway	54
Figure 40: Selected frames from the MD simulations for the QM/MM RPAth simulations	55
Figure 41: Reaction scheme of the proposed reaction mechanisms of the hydrolysis of the glycosidic bond in AtIE.	56
Figure 42: Optimized energy reaction profile of reaction path A and path B.....	57
Figure 43: Remodeled substrate in AtIE and schematic presentation of the selected NAG-NAM central unit for further virtual screening campaign.....	58
Figure 44: Proposed binding mode of the hit compound 3 in the AtIE binding site.....	59
Figure 45: STD NMR and ¹ H NMR spectra for the compounds 13 and 14.....	61
Figure 46: Hydrophilic and hydrophobic regions of the AtIE binding surface for modeled compound 10.....	62
Figure 47: Proposed interaction binding pattern between the active (phenylureido)piperidiny benzamides compound 10 and the AtIE binding site based on molecular docking calculations.....	68

List of Tables

Table 1: Table of the primers used to amplify different constructs from the genomic DNA of <i>S. aureus</i> Mu50.....	17
Table 2: Primers used to introduce mutations into AtlE	18
Table 3: Primers used to introduce mutations into SagB.....	19
Table 4: Primers used to introduce mutations into AtlA-gl.	19
Table 5: Constructs and the primers used to produce them.....	21
Table 6: Data collection and statistics table.....	43
Table 7: Compounds from the structural class of (phenylureido)piperidiny1 benzamides tested for AtlE binding affinity in SPR experiments.	60

Abbreviations

CHAP	...	Cysteine, Histidine-dependent Amidohydrolase/Peptidase
CWB	...	Cell-Wall Binding domain
DSF	...	Differential Scanning Fluorimetry
GH73	...	Glycoside Hydrolase family 73
LIC	...	Ligation Independent Cloning
MD	...	Molecular Dynamics
MM	...	Molecular Modelling
MRSA	...	Methicillin Resistant <i>Staphylococcus aureus</i>
NAG	...	<i>N</i> -acetylglucosamine
NAM	...	<i>N</i> -acetylmuramic acid
NMR	...	Nuclear Magnetic Resonance
QM	...	Quantum Mechanics
RMSD	...	Root-Mean-Square Deviation
SAR	...	Structure-Activity Relationship
SPR	...	Surface Plasmon Resonance
STD	...	Saturation-Transfer Difference
VRSA	...	Vancomycin Resistant <i>Staphylococcus aureus</i>

Symbols

ΔE . . . Energy change

ΔH . . . Enthalpy change

K_d . . . Dissociation constant

T_m . . . Melting temperature

Standard Amino Acid Abbreviations

Amino acid	Three-letter abbreviation	One-letter abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Phenylalanine	Phe	F
Glycine	Gly	G
Glutamine	Gln	Q
Glutamic acid	Glu	E
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Proline	Pro	P
Serine	Ser	S
Tyrosine	Tyr	Y
Threonine	Thr	T
Tryptophan	Trp	W
Valine	Val	V

Chapter 1

Introduction

1.1 Antibiotic Resistance

Antibiotic resistance is a growing public health threat (World Health Organization, 2014). A problem so serious that it threatens the achievements of modern medicine (World Health Organization, 2014). A post-antibiotic era in which common infections and minor injuries can kill is far from being an apocalyptic fantasy and is a very real possibility for the 21st century (World Health Organization, 2014). For over 60 years, we have relied on antibacterial drugs as the universal remedy to cure infections, whether or not their use was appropriate. Alexander Fleming, who discovered the first antibiotic, penicillin, already cautioned in his Nobel Prize speech in 1945 that bacteria could become resistant to antibiotics (Flemming, 1945). Indeed, the development of each new antibacterial drug has been followed by the detection of resistance to it (Levy, 2004). The development of resistance is a normal evolutionary process for microorganisms, but it is accelerated by the selective pressure exerted by widespread use of antibacterial drugs (Levy, 2004). Resistant strains are able to propagate and spread in the areas where there is non-compliance with infection prevention and control measures (World Health Organization, 2014).

Multiple drug-resistant bacterial pathogens are a growing public health threat, because infections they cause are challenging and expensive to treat (World Health Organization, 2014). This problem is exacerbated by the fact that few antimicrobial compounds using novel mechanisms of action are in clinical development (Figure 1) (Coates, 2011; Piddock, 2012). Some bacterial pathogens, namely *Streptococcus pneumoniae*, *Enterobacteriaceae*, *Neisseria gonorrhoeae*, *Mycobacterium tuberculosis* and *Staphylococcus aureus* develop multiple antibiotic resistance more readily than others (Chang, 2015).

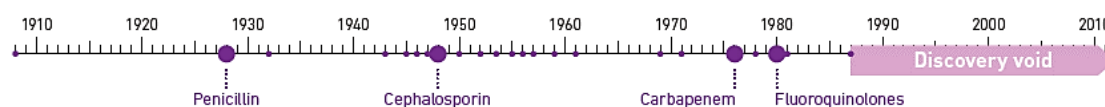


Figure 1: Illustration of the antibiotic discovery void. Years listed represent when a new class of antibiotics was discovered (World Health Organization, 2014).

1.2 *Staphylococcus aureus*

Staphylococcus aureus (*S. aureus*) is a spherical (coccal) gram-positive bacterium. It is one of the most important bacterial pathogens in clinical practice (Aryee, 2017). It colonizes up to two-thirds of the population predominantly in the anterior nares, as well as other surfaces of the human body (Aryee, 2017). Through skin breaches it can gain access deeper into the body and cause infections, which can range from superficial lesions such as wound infection, toxinoses such as food poisoning, scalded skin syndrome and toxic shock syndrome; and conditions such as endocarditis, osteomyelitis, pneumonia, brain abscesses, meningitis, and bacteremia (Aryee, 2017; Al-Mebairik, 2016). *S. aureus* can bypass the host immune system by avoiding recognition by antibodies, complement system and neutrophils, as well as actively fight the immune system by neutrophil cytolysis and altering T cell functions (Al-Mebairik, 2016). Thus the development of long-term memory can be prevented and *S. aureus* infects the same human host repeatedly throughout life (Al-Mebairik, 2016). Additionally, it has an ability to form biofilm, which effectively enables it to attach and grow on smooth surfaces. Consequently *S. aureus* is one of the most common pathogens in biofilm-related infections of implanted medical devices (Al-Mebairik, 2016). Bacterial biofilms are a complex microenvironment consisting of a single or mixed species that are attached to one another or to surfaces and encased within extracellular polymeric substances (Al-Mebairik, 2016). The key characteristic of biofilm is that it enhances the resistance of bacteria to host defenses and tolerance of antimicrobials (Al-Mebairik, 2016).

In the 1960s methicillin-resistant *S. aureus* (MRSA) has emerged in hospitals (Aryee, 2017). Since the 1990s MRSA has increasingly started appearing in community settings (David, 2010). MRSA strains are often resistant to many classes of antibiotics, besides β -lactams, and this makes MRSA infections difficult to treat (Aryee, 2017; David, 2010). Vancomycin is the current drug of choice for the treatment of MRSA infections and with the appearance of vancomycin-resistant strains (VRSA), there is no effective antibiotic class against multi antibiotic-resistant *S. aureus* infections (Al-Mebairik, 2016).

1.3 Cell Wall

Bacteria are single cell organisms that can survive in a wide variety of conditions. The structure that enables them to weather this is the cell wall. It is a rigid structure that envelops the cell and enables the survival of the cell in dilute environments by withstanding the osmotic pressure and thus preventing the cell from bursting. For cells to increase their volume and to divide, the peptidoglycan also has to enlarge. This has to happen by cutting existing bonds in peptidoglycan and inserting new elementary units. As it is an essential part of a bacterial cell, yet it is not a part of animal cells, it can be a good target for antibiotics, as demonstrated by β -lactam antibiotics.

Peptidoglycan is composed of glycan strands with *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid (NAM) 1- β 4 disaccharide as the repeating units (Kim, 2015). The disaccharide composition of glycan strands is conserved among all eubacteria, where the length of polymerized disaccharides varies among the species (Kim, 2015). The glycan strands are cross-linked by peptide moieties where amino acid composition as well as the level of the cross-linking vary among species (Kim, 2015). The pentapeptide stem attaches to 3'-OH of NAM and is followed by a bridge structure that connects two stems (Kim, 2015).

In peptidoglycan of *S. aureus* most of glycan strands are between 3 and 10 disaccharide units long, with an average length of 6 disaccharide units (Boneca, 2000). Most of the other

bacteria have much longer glycan strands. Pentapeptide stem is composed of L-Ala-D-iso-Gln-L-Lys-D-Ala-D-Ala and crosslinked by pentaglycine bridge link in between the carboxyl group of last D-Ala of one pentapeptide stem and amine group of the side chain of L-Lys of the other pentapeptide stem (Figure 2). In *S. aureus* peptidoglycan is very highly cross-linked, the estimated numbers are around 80 % (Gally, 1993).

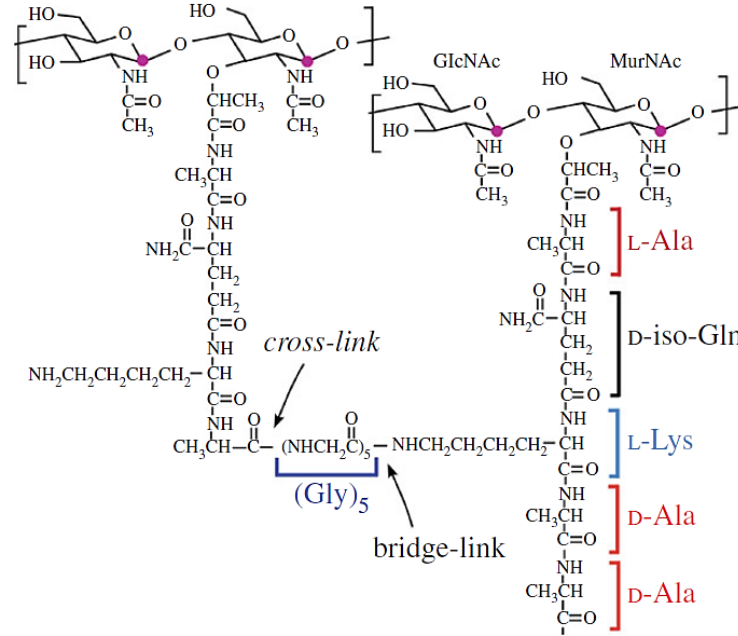


Figure 2: Primary structure of *S. aureus* peptidoglycan (Romaniuk, 2015).

1.4 Autolysins

Peptidoglycan hydrolases are an important part of fundamental bacterial physiology. They are involved in remodeling of peptidoglycan, which is a part of growth, turnover and recycling of the cell wall, as well as an important part of cell separation in the cell division, assembly of cell surface organelles, autolysis, sporulation and germination, resuscitation of dormant cells, biofilm formation, and bacterial cell lysis in interbacterial species competition (Friedrich, 2013). Their activity has to be regulated rigorously, as they can cause cell lysis, hence their name autolysins. As peptidoglycan is composed of glycan and peptide moieties. There are three different types of enzymes, glycosidases that cleave glycosidic bonds in glycan strands, peptidases that cleave peptide bonds in peptide part and *N*-acetylmuramyl-L-alanine amidases that cleave the amide bond between the NAM and L-Ala. Glycosidases are further divided into *N*-acetyl- β -D-glucosaminidases (*N*-acetylglucosaminidases) and *N*-acetyl- β -D-muramidases (muramidases), each named by the carbohydrate after which they cleave the glycosidic bond.

As the shape, cell wall structure and habitats of different bacteria vary widely, the representation of peptidoglycan hydrolases with different specificities also differs extensively. The peptidoglycan hydrolases often act on a certain type of a substrate. Hence the presence or absence of secondary modifications and the size of the substrate often determines their activity (Vollmer, 2008).

1.5 N-Acetyl- β -D-glucosaminidases

N-acetylglucosaminidases are enzymes commonly found in bacteria. Most of bacterial peptidoglycan *N*-acetylglucosaminidases belong to glycoside hydrolase family 73 (GH73) (Pfam PF01832) (Vollmer, 2008). They often have additional domains that can be either peptidoglycan binding modules or a domain with enzymatic activity (Vollmer, 2008). Exact functions of these enzymes are not yet elucidated. For example *Enterococcus faecalis* autolysin has a role in biofilm development (Guiton, 2009; Thomas, 2009). Autolysins from *Lactococcus lactis* have been implicated in cell division (Visweswaran, 2013) and it has been determined their activity can be decreased by *N*-deacetylation of peptidoglycan decreases (Meyrand, 2007).

The most studied *S. aureus* *N*-acetylglucosaminidase is AtlA. It is a protein with two catalytic domains and three cell-wall binding domains (CWB), which is proteolytically processed into a mature protein with N-terminal amidase domain with two CWBs and C-terminal *N*-acetylglucosaminidase domain with one CWB (Oshida, 1995; Biswas, 2006). AtlA inactivated strain has a rough outer surface (Foster, 1995). By smoothing out the outer cell wall surface AtlA enables *S. aureus* to avoid the recognition by innate immune system (Atilano, 2014). Inactivated strain also grows in clusters, indicating that it is involved in separation of daughter cells after cell division (Foster, 1995; Sugai, 1995). This is further substantiated by the finding that it localizes at the septum of dividing cells, forming a ring (Yamada, 1996; Schlag, 2010). It has been shown that the repeat domains of AtlA are involved in adherence to and internalization by human host cells (Hirschhausen, 2010). It was proposed that it is also involved in initial stages of biofilm development (Houston, 2011). *N*-acetylglucosaminidase domain has been shown to bind DNA in a sequence nonspecific way (Grilo, 2014). It has been proposed that *N*-acetylglucosaminidase domain of AtlA could be a good target for the development of vaccine (Varrone, 2011). Another *S. aureus* *N*-acetylglucosaminidase, SagB, has been shown to be responsible for the typical short glycan chain length of *S. aureus* (Chan, 2016; Wheeler, 2015). Its inactivation perturbs protein secretion from the cell (Chan, 2016).

1.6 Structures of GH73 Family of Enzymes

GH73 family of enzymes share lysozyme-like fold, with conserved core region that consists of 5 to 6 helices (Figure 3). The only established catalytic amino acid is positioned on top of the central helix ($\alpha 7$, AtlE numbering) (Lipski, 2015; Maruyama, 2010; Inagaki, 2009; Bublit, 2009; Herlihey, 2014; Tamai, 2017; Yokoi, 2008; Rico-Lastres, 2015). Lobe domains can vary significantly and it has been suggested the size of the left lobe domain is the determinant of the catalytic mechanism, as it can fold in a long β -hairpin that extends over the substrate binding site and house an additional carboxylic acid that can enable inverting mechanism (Lipski, 2015).

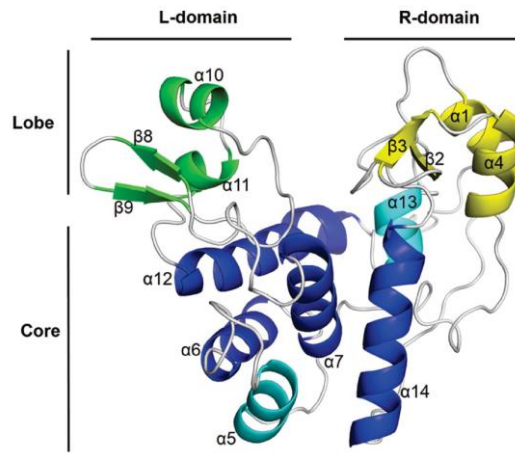


Figure 3: The structure of AtIE with helices in dark blue conserved among lysozyme family and additional two core helices in cyan conserved among GH73 family (Mihelič, 2017).

1.6.1 Substrate binding site

Determination of the structure of AtIE with substrate fragments muramyl dipeptide (4PI9) and NAG-NAM (4PI7) enabled the identification of the binding site and gave an insight into substrate binding (Mihelič, 2017). It revealed that in comparison with muramidases, glucosaminidases approach the substrate from the opposite face (Figure 4) (Mihelič, 2017).

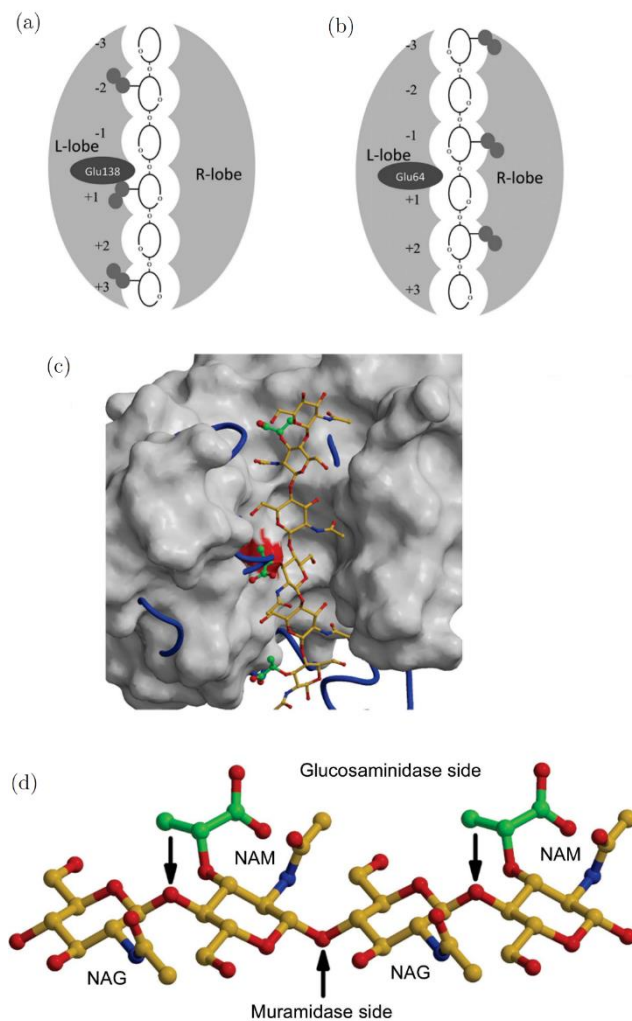


Figure 4: Comparison of the substrate binding of *N*-acetylglucosaminidase and muramidase. (a) and (b) depict a schematic representation of *N*-acetylglucosaminidase and muramidase, respectively, binding of the substrate relative to their catalytic amino acid, with *N*-acetyl group stylized as circles. (c) shows a model of substrate binding with AtIE surface in grey with the catalytic Glu138 in red and a blue chain trace of ACOD (3GXX), a muramidase, where a clear sterical hindrance is visible. (d) portrays a ball and stick representation of the substrate and the faces of the substrate from which respective enzymes approaches are labelled and cleavage sites are marked with an arrow (Mihelič, 2017).

1.6.2 Reaction mechanism

The reaction mechanism of GH73 enzymes is still not elucidated, whereas a lot of studies have been done on glycoside hydrolases (Davies, 1995; Vocadlo D. J., 2008; Withers, 2001; Vuong, 2010). Most of glycoside hydrolases possess two acidic amino acids that are involved in the reaction (Gloster, 2010). In a single step mechanism with inversion of stereochemistry catalytic acid residue donates a proton to anomeric carbon and the catalytic base removes a proton from the water molecule, thus increasing its nucleophilicity and thereby facilitating its attack on the anomeric center (Figure 5a). Retaining mechanism transpires in two steps, a general acid/base catalyst acts first as an acid by donating a proton to the glycosidic oxygen atom and facilitating the departure of the leaving group, while nucleophile forms a covalent glycosyl-enzyme intermediate. In the next step, general acid/base is deprotonated and acts as a base to deprotonate a water molecule which provides a nucleophilic species which attacks the anomeric carbon from the same face and this way leads to retention of the stereochemistry (Figure 5b). A number of glycosidase families which hydrolyze substrates containing *N*-acetylhexosamine have been shown to lack a conventional catalytic nucleophile and have been shown to utilize a substrate-assisted mechanism with retention of configuration. In this case the C2 *N*-acetyl group of the substrate takes the place of the nucleophile forming an oxazolinium intermediate, which in turn is hydrolyzed by a water molecule that was activated by a general base (Figure 5c). In most such cases a second carboxylate-containing residue is found to orient and polarize the *N*-acetyl group to increase its nucleophilicity (Gloster, 2010).

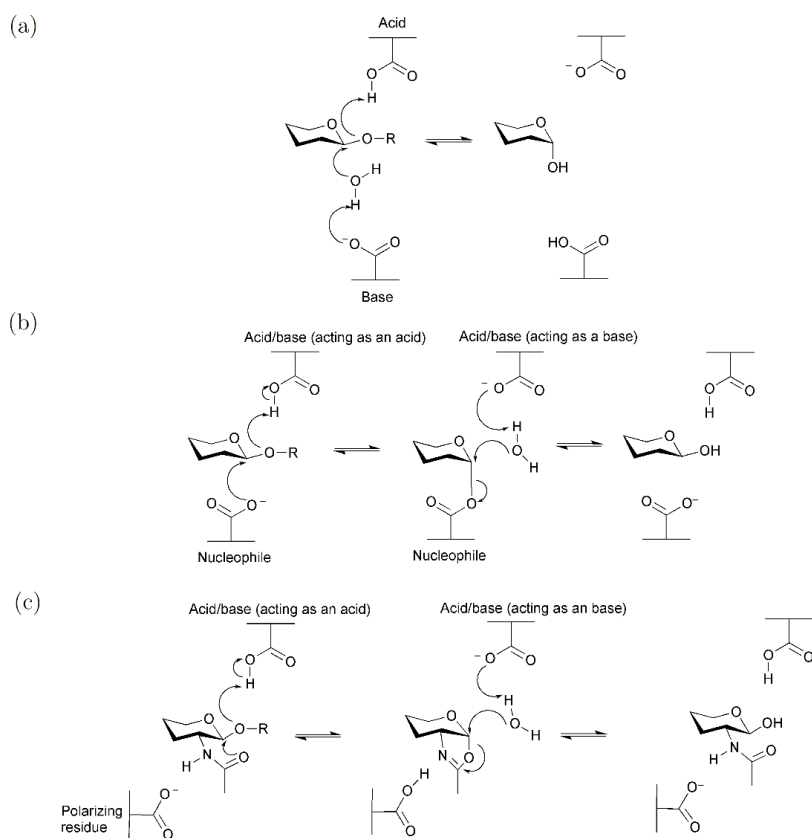


Figure 5: Enzymatic glycoside bond hydrolysis mechanisms. (a) Inverting mechanism, (b) Retaining mechanism and (c) Substrate-assisted mechanism (Gloster, 2010).

For the GH73 enzymes the second carboxylic group remains elusive (Lipski, 2015; Maruyama, 2010; Inagaki, 2009; Bublitz, 2009; Herlihey, 2014; Tamai, 2017; Yokoi, 2008; Rico-Lastres, 2015). Lipski et al. proposed that it is likely that not all GH73 enzymes share the same catalytic mechanism, as it has been shown for GH23 and GH97 families (Lipski, 2015).

Chapter 2

Purpose of the Work

Peptidoglycan cell wall is present only in bacterial cells and is crucial for their survival. For bacterial cells to grow and divide, this rigid structure has to be remodeled. This process has to involve degradation as well as synthesis. Peptidoglycan degrading enzymes exhibit the potential of novel drug targets. This prompted us to study *S. aureus* *N*-acetylglucosaminidases from the GH73 family. These enzymes are located at the cell surface of bacteria and are easily accessible to drugs.

The determination of the crystal structures of several enzymes of GH73 family from *S. aureus* Mu50 enabled us to characterize the substrate binding groove and active site. Together with mutational studies it gave us an additional insight into catalytic mechanism and this helped us to contribute towards the understanding of the mechanism of substrate cleavage of GH73 enzymes. The structural insight into the mechanism of the family of GH73 enzymes will open up new possibilities of manipulating the activity of these enzymes. With novel insight and understanding of the mechanisms underlying remodeling of the peptidoglycan cell wall of *S. aureus*, a dangerous human pathogen, we may contribute to the development of novel classes of antibiotics.

Chapter 3

Materials and Methods

3.1 Materials

3.1.1 Equipment

- Automatic liquid handling system Glyphoon (Art Robbins)
- Centrifuge 5415R (Eppendorf)
- Centrifuge 5804R (Eppendorf)
- Centrifuge Heraeus Multifuge X3R (Thermo Scientific)
- Centrifuge RC 5C (Sorval)
- Chromatographic system Akta xpress (GE)
- Crystal imaging system RI1000 (Formulatrix)
- Diffractometer Proteum X8 (Bruker)
- Digital sonifier 450D (Brandson)
- Electrophoresis power supply Power Pack 300 (Biorad)
- Horizontal Electrophoresis Systems Easycast B2-BP and B1A (Owl Scientific Inc.)
- Incubator I (Memmert)
- Laminar flow chamber Iskra PIO LFVP 12
- Laminar flow chamber Iskra PIO MC 12-2
- Magnetic stirrer Rotamix 560MM and 550MM (Tehtnica)
- Nanodrop One (Thermo Scientific)
- PCR Thermal Cyclor 2720 and 2700 (Applied Biosystems)
- pH meter Seven Easy (Mettler Toledo)
- Shaking incubator MaxQ 4000 (Thermo Scientific)
- Shaking incubator innova 44 and innova 4430 (New Brunswick Scientific)
- Shaking incubator (Labtherm)
- Spectrophotometer Ultrospec 8000 (GE)
- Spectrophotometer BioSpectrometer (Eppendorf)
- Thermomixer compact (Eppendorf)
- Vertical Electrophoresis Lab Apparatus MiniProtean Tetracell (Biorad)

3.1.2 Chemicals, commercial kits and chromatography columns

- Acrylamide-bisacrylamide ready-to-use solution (37.5:1) (Merck)
- Acetic acid (Carlo Erba)
- Ammonium chloride (Sigma)
- Ammonium sulfate (Merck)
- Amersham™ LMW Calibration Kit For SDS Electrophoresis (GE Healthcare)

- Ampicillin (Sigma-Aldrich)
- Boric acid (Sigma)
- β -mercaptoethanol (Serva)
- Calcium chloride (Sigma)
- Chromatography column HiLoad 16/600 Superdex 200 pg (GE Healthcare)
- Chromatography column HiPrep 16/60 Sephacryl S-200 HR (GE Healthcare)
- Chromatography column HisTrapTM FF crude (GE Healthcare)
- Cobalt(II) chloride (Sigma)
- Copper(II) chloride (Sigma)
- dNTP (NEB)
- D-lactose monohydrate (Sigma-Aldrich)
- D-glucose monohydrate (Acros Organics)
- DMSO (Merck)
- DTT (Serva)
- EDTA (Serva)
- Gene Ruler 100 bp Plus DNA Ladder (Thermo Scientific)
- Genomic DNA from *Staphylococcus aureus subsp. aureus* strain Mu50 (ATCC® 700699TM) (ATCC)
- Glycerol (Carlo Erba)
- HEPES (Serva)
- Hydrofluoric acid (40 %) (Sigma)
- Iron chloride (Sigma)
- Isopropanol (Sigma-Aldrich)
- L-alanine (Sigma)
- L-arginine – HCl (Sigma)
- L-asparagine (Sigma)
- L-aspartic acid (Sigma)
- L-glutamine (Sigma)
- L-glycine (Sigma)
- L-histidine – HCl (Sigma)
- L-isoleucine (Sigma)
- L-leucine (Sigma)
- L-lysine – HCl (Sigma)
- L-methionine (Sigma)
- L-phenylalanine (Sigma)
- L-proline (Sigma)
- L-selenomethionine (Sigma)
- L-serine (Sigma)
- L-threonine (Sigma)
- L-tryptophan (Sigma)
- L-valine (Sigma)
- Lithium chloride (Sigma)
- Magnesium sulfate (Sigma)
- Manganese chloride (Sigma)
- MES (Serva)
- MPD (Fluka Analytical)
- N-Z-Case[®] Plus (Sigma)
- Nickel(II) chloride (Sigma)
- NeXtal DWBlock Suites (Qiagen)
- Phast GelTM Commassie Brilliant Blue tablets (GE Healthcare)
- Potassium chloride (Sigma)

- Qiaquick Gel extraction kit (Qiagen)
- Qiaquick PCR purification kit (Qiagen)
- Remazol Brilliant Blue R (Sigma)
- SDS (Serva)
- Sodium acetate (Sigma-Aldrich)
- Sodium bromide (Sigma)
- Sodium chloride (Applichem)
- Sodium citrate (Sigma)
- Sodium fluoride (Sigma)
- Sodium L-glutamate (Sigma)
- Sodium hydroxide (Merck)
- Sodium iodide (Sigma)
- Sodium malonate (Sigma)
- Sodium molybdate (Sigma)
- Sodium nitrate (Sigma)
- Sodium selenite (Sigma)
- Sodium sulfate (Sigma)
- Sodium tartrate (Sigma)
- Sodium thiocyanate (Sigma)
- SYBR Safe DNA gel stain (Invitrogen)
- Sypro Orange protein gel stain (5000 x concentrate in DMSO) (Life Technologies)
- PEG 200 (Sigma)
- PEG 400 (Sigma)
- Tris (Serva)
- Tris-HCl (Fischer Scientific)
- Vitamin B₁₂ (Sigma)
- Wizard® Plus SV Minipreps DNA Purification System (Promega)
- Zinc chloride (Sigma)

3.1.3 Buffers, solutions and media

- Gel chromatography buffer
30 mM Tris – HCl pH 7.5
400 mM NaCl
- Lysis buffer
30 mM Tris – HCl pH 7.5
400 mM NaCl
1 mg/ml lysozyme
- Ni-NTA Binding Buffer
30 mM Tris – HCl pH 7.5
400 mM NaCl
10 mM imidazole
- Ni-NTA Elution buffer
30 mM Tris – HCl pH 7.5
400 mM NaCl
300 mM imidazole

- Default buffer for purified proteins
20 mM HEPES pH 7.5
100 mM NaCl
- Solution for SDS-PAGE stacking gel
4% acrylamide-bisacrylamide ready to use solution
0.375 M Tris-HCl pH 8.8
0.1% SDS
0.15 % TEMED
0.03 % APS
- Solution for SDS-PAGE resolving gel
15% acrylamide-bisacrylamide ready to use solution
0.125 M Tris-HCl pH 6.8
0.1% SDS
0.15 % TEMED
0.03 % APS
- SDS-PAGE loading buffer (6x)
375 mM Tris-HCl pH 6.8
6% SDS
48% glycerol
100 mM DTT
0.03 % bromophenol blue
- SDS-PAGE running buffer
25 mM Tris
200 mM glycine
0.1% SDS
- SDS-PAGE staining solution
0.1 % Coomassie Brilliant Blue
28.8 % ethanol
10 % acetic acid
- SDS-PAGE fixing solution
28.8 % ethanol
10 % acetic acid
- AGE loading buffer
0.09 % bromophenol blue
0.09 % xylene cyanol FF
60 % glycerol
60 mM EDTA
- AGE running buffer
40 mM Tris-acetate pH 8
1 mM EDTA

- Luria-Bertani liquid medium (LB)
 - 1 % N-Z-Case® Plus
 - 0.5 % Yeast extract
 - 1 % NaCl

- Luria-Bertani liquid medium with ampicillin
 - 1 % N-Z Case® Plus
 - 0.5 % Yeast extract
 - 1 % NaCl
 - 100 µg/ml ampicillin

- LB agar medium with ampicillin
 - 1 % N-Z Case® Plus
 - 0.5 % Yeast extract
 - 1 % NaCl
 - 1.5 % agar
 - 100 µg/ml ampicillin

- ZYM-5052A
 - 1% N-Z Case® Plus
 - 0.5% Yeast extract
 - 25mM Na₂HPO₄
 - 25mM KH₂PO₄
 - 50mM NH₄Cl
 - 5mM Na₂SO₄
 - 2mM MgSO₄
 - 0.5% glycerol
 - 0.05% glucose
 - 0.2% α-D-lactose
 - 10 µM FeCl₃
 - 4 µM CaCl₂
 - 2 µM MnCl₂
 - 2 µM ZnSO₄
 - 0.4 µM CoCl₂
 - 0.4 µM CuCl₂
 - 0.4 µM NiCl₂
 - 0.4 µM Na₂MoO₄
 - 0.4 µM Na₂SeO₃
 - 0.4 µM H₃BO₃
 - 100 µg/ml ampicillin

- PASM-5052A
 - 50 mM Na_2HPO_4
 - 50 mM KH_2PO_4
 - 25 mM $(\text{NH}_4)_2\text{SO}_4$
 - 2 mM MgSO_4
 - 0.5 % glycerol
 - 0.05 % glucose
 - 0.2 % α -D-lactose
 - 200 $\mu\text{g}/\text{ml}$ sodium glutamate
 - 200 $\mu\text{g}/\text{ml}$ aspartic acid
 - 200 $\mu\text{g}/\text{ml}$ lysine – HCl
 - 200 $\mu\text{g}/\text{ml}$ arginine – HCl
 - 200 $\mu\text{g}/\text{ml}$ histidine – HCl
 - 200 $\mu\text{g}/\text{ml}$ alanine
 - 200 $\mu\text{g}/\text{ml}$ proline
 - 200 $\mu\text{g}/\text{ml}$ glycine
 - 200 $\mu\text{g}/\text{ml}$ threonine
 - 200 $\mu\text{g}/\text{ml}$ serine
 - 200 $\mu\text{g}/\text{ml}$ glutamine
 - 200 $\mu\text{g}/\text{ml}$ asparagine
 - 200 $\mu\text{g}/\text{ml}$ valine
 - 200 $\mu\text{g}/\text{ml}$ leucine
 - 200 $\mu\text{g}/\text{ml}$ isoleucine
 - 200 $\mu\text{g}/\text{ml}$ phenylalanine
 - 200 $\mu\text{g}/\text{ml}$ tryptophan
 - 10 $\mu\text{g}/\text{ml}$ methionine
 - 125 $\mu\text{g}/\text{ml}$ selenomethionine
 - 100 nM vitamin B_{12}
 - 10 μM FeCl_3
 - 4 μM CaCl_2
 - 2 μM MnCl_2
 - 2 μM ZnSO_4
 - 0.4 μM CoCl_2
 - 0.4 μM CuCl_2
 - 0.4 μM NiCl_2
 - 0.4 μM Na_2MoO_4
 - 0.4 μM Na_2SeO_3
 - 0.4 μM H_3BO_3
 - 100 $\mu\text{g}/\text{ml}$ ampicillin

3.1.4 Cloning primers

Table 1: Table of the primers used to amplify different constructs from the genomic DNA of *S. aureus* Mu50. F denotes sense (5') strand primer and R anti-sense (3') strand primer. Portion of the primer in capital letters corresponds to the sequence necessary for LIC and the rest of the sequence is complementary to the amplified gene. All primers are listed 5' to 3'.

Name of the primer	Sequence of the primer
SagB-F1	TACTTCCAATCCAATGCCGCAGCAAttcaaacaatgttaaaccga
SagB-F2	TACTTCCAATCCAATGCCGCAGCAAttatcaaagtatcaaggat
SagB-R	TTATCCACTTCCAATGTTAttacttattcaaagtgttac
SAV0909-F1	TACTTCCAATCCAATGCCGCAGCAgctaaaaataaaatagctat
SAV0909-F3	TACTTCCAATCCAATGCCGCAGCAagtccttatacattcaaaca
SAV0909-R1	TTATCCACTTCCAATGTTAAAtccacctgggtccatcatgat
SAV0909-R2	TTATCCACTTCCAATGTTAttattttatatttatcccttg
ScaH-F1	TACTTCCAATCCAATGCCGCAGCAccagcaagtgacgataaagc
ScaH-F2	TACTTCCAATCCAATGCCGCAGCAtatccacatggccaatgtac
ScaH-R1	TTATCCACTTCCAATGTTAttattttacctgtaatatatg
ScaH-R2	TTATCCACTTCCAATGTTAatcgtcaaactgagtttaatt
AtlA-F1	TACTTCCAATCCAATGCCGCAGCAacgactcaaaaaatatactgtt
AtlA-F2	TACTTCCAATCCAATGCCGCAGCAgaattaattaagtataatcaa
AtlA-F3	TACTTCCAATCCAATGCCGCAGCAaatgatgttaagcatgcaatg
AtlA-F4	TACTTCCAATCCAATGCCGCAGCAgttgctgcaacacacaaaaac
AtlA-R1	TTATCCACTTCCAATGTTAacgacatcccacaataataaa
AtlA-R2	TTATCCACTTCCAATGTTAaccccatggcgccactttac
AtlE-F	TACTTCCAATCCAATGCCGCAGCAaatgatgtgaattactctttt
AtlE-R	TTATCCACTTCCAATGTTAagatattagacaaacatattataaa

3.1.5 Mutagenesis primers

Table 2: Primers used to introduce mutations into AtIE. F denotes sense (5') strand primer and R anti-sense (3') strand primer. All primers are listed 5' to 3'.

Primer name	Primer sequence
E58A-F	gaaaggtattgtacagacaaaagcagaggatggtaaatttgttgaa
E58A-R	ttcaacaaatttaccatcctctgcttttgtctgtacaataccttcc
E138A-F	gttagccatgcattagtagcaacaggtaacggcaaatca
E138A-R	tgatttgccgttacctgttgctactaatgcatggctaac
E138Q-F	gatttgccgttacctgtctgtactaatgcatggctaacaagataaatgac
E138Q-R	gtcatttatcttgttagccatgcattagtagacaggtaacggcaaatc
E138D-F	gccgttacctgtatctactaatgcatggctaacaag
E138D-R	cttgtagccatgcattagtagatacaggtaacggc
E145A-F	gaaacaggtaacggcaaatcagcattagcaaaaggcattaaagat
E145A-R	atctttaatgccttttgtaaatgctgatttgccgttacctgttcc
D167A-F	ttggtataggagcattcgttagtagtgctgttcgtag
D167A-R	ctacgaacagcactactagcgaatgctcctataccaa
Y224A-F	gaatggcagcttgatgttgccgaggattttctggatt
Y224A-R	aatccagaaaatcctgcgcaacatcaagctgagtgacattc
S226A-F	gcgcaacatcaatatgcggctgacattcgatgggcaga
S226A-R	tctgccatcgaatgtcagccgcatattgatgttgccg
S226T-F	ccatcgaatgtcagtcgcatattgatgttgccga
S226T-R	tgcgcaacatcaatatgcgactgacattcgatgg
S226V-F	atctgccatcgaatgtcaaccgcatattgatgttgccga
S226V-R	tgcgcaacatcaatatgcggttgacattcgatgggcagat
D227A-F	caacatcaatatgcgagtgccattcgatgggcagataaa
D227A-R	tttatctgccatcgaatggcactcgcatattgatgttg
D227N-F	cccatcgaatgttactcgcatattgatgttgccga
D227N-R	tgcgcaacatcaatatgcgagtaacattcgatggg
D232A-F	gagtgacattcgatgggcagctaaaattgccaaattaatgga
D232A-R	tccattaatttggcaatttttagctgccatcgaatgtcact
D232N-F	ccattaatttggcaattttatttgccatcgaatgtcactcg
D232N-R	cgagtgacattcgatgggcaataaaattgccaaattaatgg
Y201A-F	ttggtggtgcaaagttcattcgtaatgaagcttttgaaaacaatcaactgaatttatatca
Y201A-R	tgatataaattcagttgattgttttcaaaagcttcattacgaatgaactttgcaccaccaa
Y201F-F	ttggtggtgcaaagttcattcgtaatgaagcttttgaaaacaatcaactgaatttatatca
Y201F-R	tgatataaattcagttgattgttttcaaaagcttcattacgaatgaactttgcaccaccaa
W230A-F	catcaatatgcgagtgacattcgagcggcagataaaattgccaa
W230A-R	ttggcaattttatctgccgctcgaatgtcactcgcatattgatg

Table 3: Primers used to introduce mutations into SagB. F denotes sense (5') strand primer and R anti-sense (3') strand primer. All primers are listed 5' to 3'.

Primer name	Primer sequence
E121A-F	gtttatttaatttcacatgcattattagcaactggcgcagttaaaagtg
E121A-R	cacttttaactgcgccagttgctaataatgcatgtgaaattaaataaac
Y209A-F	gatggaatccaaaaaatccaggagaacatcaagctgctactgatattaagtg
Y209A-R	cacttaatatcagtagcagcttgatgttctcctggatttttggattccatc
Y209F-F	ggaatccaaaaaatccaggagaacatcaatttgctactgatattaag
Y209F-R	cttaatatcagtagcaaattgatgttctcctggatttttggattcc
T211A-F	caggagaacatcaatatgctgctgatattaagtgggcagaa
T211A-R	ttctgcccacttaatatcagcagcatattgatgttctcctg
T211S-F	caggagaacatcaatatgcttctgatattaagtgggcagaa
T211S-R	ttctgcccacttaatatcagaagcatattgatgttctcctg
D211A-F	ggagaacatcaatatgctactgctattaagtgggcagaaagtaat
D211A-R	attactttctgcccacttaatagcagtagcatattgatgttctcc
W215A-F	aacatcaatatgctactgatattaaggcggcagaaagtaatgaacaattatc
W215A-R	gataattgttgctacttcttctgcccgttaatatcagtagcatattgatgtt
F185A-F	ggtgctgatttcattcataagcacgccttatcaagcacagatcaaaatac
F185A-R	gtattttgatctgtgcttgataaggcgtgcttatgaatgaaatcagcacc

Table 4: Primers used to introduce mutations into At1A-gl. F denotes sense (5') strand primer and R anti-sense (3') strand primer. All primers are listed 5' to 3'.

Primer name	Primer sequence
E116A-F	tctcacatgccctattagcaacaggtaacgggtacttc
E116A-R	gaagtaccgttacctgttgctaataaggccatgtgaga
Y214A-F	tcctgcacatccaggaacacaccaagctgctacagatgtag
Y214A-R	ctacatctgtagcagcttggtgtgttctcctggatgtgcagga
Y214F-F	cacatccaggaacacaccaatttgctacagatgtagatt
Y214F-R	aatctacatctgtagcaaattggtgtgttctcctggatgtg
T216A-F	ggaacacaccaatatgctgcagatgtagattgggcta
T216A-R	tagcccaatctacatctgcagcatattggtgtgttcc
T216S-F	ggaacacaccaatatgcttcagatgtagattgggcta
T216S-R	tagcccaatctacatctgaagcatattggtgtgttcc
D217A-F	acacaccaatatgctacagctgtagattgggctaacatc
D217A-R	gatgttagcccaatctacagctgtagcatattggtgtgt
D217N-F	gaacacaccaatatgctacaaatgtagattgggctaacatc
D217N-R	gatgttagcccaatctacattttagcatattggtgtgttcc
W219A-F	cacaccaatatgctacagatgtagatgctggaacatcaatgctaaaat
W219A-R	atcttagcattgatgttagccgcatctacatctgtagcatattggtgtg
Y191A-F	gtggtgctaaattcatcggaactcagctgtaaaagctggtcaaa
Y191A-R	tttgaccagcttttacagctgagttgccgatgaatttagcaccac

3.1.6 Plasmids

- pMCSG7 (Figure 6).

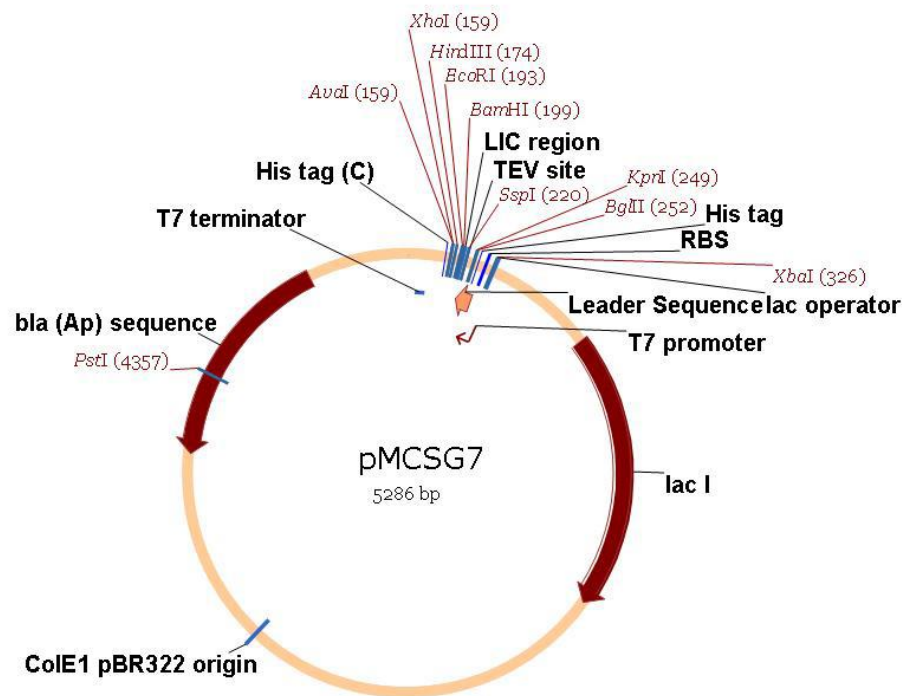


Figure 6: Plasmid map of pMCSG7.

3.1.7 Bacterial cultures

- *Escherichia coli* DH5 α
- *Escherichia coli* BL21DE3
- *Bacillus subtilis* 168 Δ trpC2

3.1.8 Recombinant enzymes

- TEV was produced “in house” (Expression and Purification of soluble His6-Tagged TEV Protease, 2009)
- Kod Hot Start polymerase (Novagene)
- SspI restrictase (NEB)
- Pronase (Roche)
- T4 DNA Polymerase (NEB)
- Taq DNA polymerase (Thermo Scientific)

3.2 Methods

3.2.1 Protein sequence analysis

GH73 proteins from *S. aureus* Mu50 were identified by CAZy (Cantarel B. L., 2009) and domain prediction was made by InterPro (Finn, 2017).

3.2.1 Sequence analysis of *S. aureus* GH73 enzymes

All non-partial *S. aureus* GH73 proteins from the CAZy database were aligned with KAlign and the conserved parts of the sequences were used for clustering with Clustal Omega. Different clusters of proteins were presented as a tree with Jalview, whole sequences of proteins of different branches of the tree were aligned with Clustal Omega.

3.2.2 Cloning

A variation of standard cloning procedures, ligation independent cloning (LIC) was used with pMSCG7 plasmid (Stols, 2002).

3.2.2.1 Construct DNA sequence amplification from genomic DNA

The primers used are listed in Table 1 and the pairs of primers used for different constructs are listed in Table 5.

Table 5: Constructs and the primers used to produce them.

Construct name	Sense (5') primer	Anti-sense (3') primer	Description
SagB	SagB-F1	SagB-R	full length protein with no transmembrane region
SagB-gl	SagB-F2	SagB-R	glucosaminidase domain
SAV0909-CHAP	SAV0909-F1	SAV0909-R1	CHAP domain
SAV0909	SAV0909-F1	SAV0909-R2	full length protein
SAV0909-gl	SAV0909-F3	SAV0909-R2	glucosaminidase domain
ScaH	ScaH-F1	ScaH-R1	full length protein
ScaH-gl	ScaH-F1	ScaH-R2	glucosaminidase domain
ScaH-CHAP	ScaH-F2	ScaH-R1	CHAP domain
AtIA-gl	AtIA-F1	AtIA-R1	glucosaminidase domain with 1 cell wall binding repeat
AtIA-gl1	AtIA-F2	AtIA-R1	glucosaminidase domain longer construct
AtIA-gl2	AtIA-F3	AtIA-R1	glucosaminidase domain shorter construct
AtIA-am	AtIA-F4	AtIA-R2	amidase domain
AtIE	AtIE-F	AtIE-R	full length protein with no transmembrane region

DNA sequences for adequate products were amplified using:

10x Buffer for KOD Hot Start DNA Polymerase	5 μ L
25 mM MgSO ₄	3 μ L
2 mM each dNTP	5 μ L
PCR grade water	22 μ L
5 M betain	10 μ L
Sense primer (10 μ M) (Table 1)	1.5 μ L
Anti-sense primer (10 μ M) (Table 1)	1.5 μ L
<i>S. aureus</i> Mu50 genomic DNA	1 μ L
KOD Hot Start Polymerase (1 U/ μ L)	1 μ L
Σ	50 μ L

and temperature protocol:

<u>2 min</u>	<u>94 °C</u>	<u>1 hold</u>
30 s	94 °C	
30 s	45 °C	5 cycles
<u>2 min</u>	<u>68 °C</u>	
30 s	94 °C	
30 s	55 °C	25 cycles
<u>2 min</u>	<u>68 °C</u>	
7 min	68 °C	1 hold
∞	4 °C	

The success of the reaction was verified by AGE on 1% gel.

3.2.2.2 T4 DNA polymerase treatment

Amplified DNA was purified using Qiaquick PCR purification kit and then treated with T4 DNA polymerase using:

Insert	10 μ L
Water	20 μ L
T4 DNA polymerase	0.5 μ L
NEB2 buffer	4 μ L
100 mM DTT	1.2 μ L
BSA	1 μ L
<u>100 mM dCTP</u>	<u>2 μL</u>
Σ	38.7 μ L

and temperature protocol:

40 min	25 °C
20 min	75 °C

3.2.2.3 Plasmid amplification and isolation

The plasmids were transformed into *E. coli* DH5 α by heat shock and grown overnight on LBA agar plates. The next day one of the successfully transformed colonies was inoculated in 10 mL of LBA liquid medium and grown overnight. The next morning the plasmids were

isolated using Wizard® Plus SV Minipreps DNA Purification System and in the final step the plasmids were eluted from the column with autoclaved MQ water.

3.2.2.4 Plasmid restriction and T4 DNA polymerase treatment

Plasmids were restricted with SspI which produces blunt ends following the protocol:

plasmid	100 μ L
SspI	6 μ L
10x SspI buffer	12 μ L
Water	2 μ L
Σ	120 μ L

and incubated for at least 4 h to overnight at 37 °C.

Restriction products were separated on 0.8 % AGE and isolated from agarose gel with Qiaquick Gel extraction kit.

Restricted linearized plasmids were treated with T4 DNA polymerase:

Plasmid	85 μ L
T4 DNA polymerase	5 μ L
NEB2 buffer	12 μ L
100 mM DTT	3.6 μ L
BSA	1.5 μ L
dGTP	6 μ L
Σ	113.1 μ L

and temperature protocol:

40 min 25 °C

20 min 75 °C

3.2.2.5 Ligation

Ligation mixture was set up as

Plasmid	3 μ L
Insert	7 μ L
Σ	10 μ L

and incubated at room temperature for 30 min. Ligation mixture was heat shock transformed into E. coli DH5 α and the success of this step was verified by growth of the resulting culture on LBA plates.

From the resulting cultures plasmids were amplified and isolated the same way as in 3.2.2.3. Successful cloning was verified by sequencing.

3.2.2.6 Construction of mutants

Mutants were prepared by QuickChange® protocole. Reaction was set up as follows:

10 x KOD Hot Start Polymerase buffer	5 μ L
Plasmid pMCSG7 with AtIE sequence	0.1 – 5 μ L
Sense primer (125 ng)	1 μ L
Anti-sense primer (125 ng)	1 μ L
2 mM each dNTP	5 μ L
MgSO ₄	3 μ L
DMSO	3 μ L
Water	30.9 – 26 μ L
KOD Hot Start Polymerase	1 μ L
Σ	50 μ L

And temperature program:

1 min	95 °C	1 hold
50 s	95 °C	
50 s	Tann	18 cycles
6 min	68 °C	
7 min	68 °C	1 hold

Each reaction was treated with 1 μ L of DpnI (10 U/ μ L) for 1 to 4 h at 37 °C. And 2 μ L of this mixture was transformed into *E. coli* DH5 α by heat shock. Plasmids were amplified and isolated as described in 3.2.2.3. Success of mutant construction was verified by sequencing.

3.2.3 Protein Production and Purification

Plasmids were transformed into *E. coli* BL21DE3 by heat shock.

3.2.3.1 Testing the solubility and expression levels of different constructs

Solubility and expression levels were determined by growing the culture in small volumes of autoinduction media. Overnight cultures were grown in LBA media. 1 mL of ZYM-5052A was inoculated with 10 μ L of overnight culture and incubated at different temperature protocols for 24 h. After that cultures were centrifuged and the media was decanted. Cells were resuspended in lysis buffer and sonicated. Lysate was cleared by centrifugation and applied to SDS-PAGE.

3.2.3.2 Large scale protein production and purification

Overnight cultures were grown in LBA media. 1 L of the ZYM-5052A was inoculated with 10 mL of the overnight culture and grown at the appropriate temperature protocol for 24 h. Cells were removed from the media by centrifugation at 6000 g at 4 °C for 15 min, resuspended in lysis buffer and sonicated for 2.5 min (10 s on, 20 s off, with amplitude 50 %). Lysates were cleared by centrifugation at 40 000 g at 4 °C for 10 min and applied to 5 mL Ni-NTA, followed by Sephacryl S-200. The success of purification was followed by SDS-PAGE. The appropriate fractions were selected to continue purification. His-tag was removed by TEV protease, which was incubated with the protein over the weekend in the presence of 10 mM β -mercaptoethanol. TEV and His-tag were removed by secondary Ni-

NTA and flow-through was collected. Protein was concentrated and the buffer was exchanged.

3.2.3.2.1 Se-Met derivative large scale production

Se-Met labelled protein was produced and purified in a very similar way to native protein. The culture was grown in PASM-5052A medium, as well as all the buffers consisted of additional 10 mM DTT.

3.2.4 DSF

3.2.4.1 Screen of conditions of optimal stability for different proteins

DSF was performed using the protocol (Niesen, 2007):

0.4 mg/ml protein	10 μ L
20 x Sypro Orange	10 μ L
Screen solution	20 μ L
Σ	40 μ L

And temperature protocol:

Starting temperature 25 °C

$\Delta T = +1$ °C 1 min 70 cycles

The conditions screened for protein stability of wild type proteins:

- 200 mM Acetate pH 4
- 200 mM Acetate pH 4.5
- 200 mM Acetate pH 5
- 200 mM Citrate pH 5
- 200 mM Citrate pH 5.5
- 200 mM Citrate pH 6
- 200 mM MES pH 6
- 200 mM MES pH 6.5
- 200 mM HEPES pH 7
- 200 mM HEPES pH 7.5
- 200 mM HEPES pH 8
- 200 mM Tris pH 7.5
- 200 mM Acetate pH 4, 100 mM NaCl
- 200 mM Acetate pH 4.5, 100 mM NaCl
- 200 mM Acetate pH 5, 100 mM NaCl
- 200 mM Citrate pH 5, 100 mM NaCl
- 200 mM Citrate pH 5.5, 100 mM NaCl
- 200 mM Citrate pH 6, 100 mM NaCl
- 200 mM MES pH 6, 100 mM NaCl
- 200 mM MES pH 6.5, 100 mM NaCl
- 200 mM HEPES pH 7, 100 mM NaCl
- 200 mM HEPES pH 7.5, 100 mM NaCl

- 200 mM HEPES pH 8, 100 mM NaCl
- 200 mM Tris pH 7.5, 100 mM NaCl
- 200 mM Acetate pH 4, 250 mM NaCl
- 200 mM Acetate pH 4.5, 250 mM NaCl
- 200 mM Acetate pH 5, 250 mM NaCl
- 200 mM Citrate pH 5, 250 mM NaCl
- 200 mM Citrate pH 5.5, 250 mM NaCl
- 200 mM Citrate pH 6, 250 mM NaCl
- 200 mM MES pH 6, 250 mM NaCl
- 200 mM MES pH 6.5, 250 mM NaCl
- 200 mM HEPES pH 7, 250 mM NaCl
- 200 mM HEPES pH 7.5, 250 mM NaCl
- 200 mM HEPES pH 8, 250 mM NaCl
- 200 mM Tris pH 7.5, 250 mM NaCl
- 200 mM Acetate pH 4, 500 mM NaCl
- 200 mM Acetate pH 4.5, 500 mM NaCl
- 200 mM Acetate pH 5, 500 mM NaCl
- 200 mM Citrate pH 5, 500 mM NaCl
- 200 mM Citrate pH 5.5, 500 mM NaCl
- 200 mM Citrate pH 6, 500 mM NaCl
- 200 mM MES pH 6, 500 mM NaCl
- 200 mM MES pH 6.5, 500 mM NaCl
- 200 mM HEPES pH 7, 500 mM NaCl
- 200 mM HEPES pH 7.5, 500 mM NaCl
- 200 mM HEPES pH 8, 500 mM NaCl
- 200 mM Tris pH 7.5, 500 mM NaCl
- 200 mM Tris pH 8
- 200 mM Tris pH 8.5
- 50 mM NaI
- 50 mM NaBr
- 50 mM NaF
- 50 mM NaNO₃
- 50 mM Na₂SO₄
- 50 mM NaSCN
- 50 mM NaMalonate
- 50 mM NaTartrate
- 200 mM Tris pH 8, 100 mM NaCl
- 200 mM Tris pH 8.5, 100 mM NaCl
- 50 mM LiCl
- 50 mM KCl
- 50 mM MgSO₄
- 50 mM CaCl₂
- 50 mM MnCl₂
- 50 mM ZnCl₂

- 50 mM AmCl
- 200 mM Tris pH 8, 250 mM NaCl
- 200 mM Tris pH 8.5, 250 mM NaCl
- 10 mM glycine
- 100 mM glycine
- 10 mM urea
- 100 mM urea
- 5 mM DTT
- 5 mM EDTA
- 10 mM arginine
- 100 mM arginine
- 1 % PEG 400
- 5 % PEG 400
- 200 mM Tris pH 8, 500 mM NaCl
- 200 mM Tris pH 8.5, 500 mM NaCl
- 1 % glycerol
- 5 % glycerol
- 10 % glycerol
- 20 % glycerol
- 5 % MPD
- 20 % MPD
- 5 % isopropanol
- 5 % t-butanol

3.2.4.2 Determination of the stability of mutant proteins

The protocol used was the same as for the screen of the conditions, but with no variation of the buffer. Buffers used were 100 mM MES pH 6.5 for SagB, 100 mM Tris pH 8 for AtIA and 100 mM MES pH 5.5 for AtIE.

3.2.5 Crystallization

For crystallization all proteins were dialyzed into appropriate buffers and concentrated to a minimum of 10 mg/mL to near maximum solubility. Three different concentrations were set up for screening for crystallization conditions using automatic liquid handling system and NeXtal DWBlock Suites. Drops of 0.1 μ L of protein and 0.1 μ L of precipitant solution were pipetted and 70 μ L of reservoir solution was used for each condition.

When crystallization conditions were determined, they were optimized, varying the pH and precipitant concentration. For this 1 μ L of protein and 1 μ L of precipitant solution were mixed and 100 to 500 μ L of precipitant solution was pipetted into the well.

3.2.6 Structure determination

Crystals were tested for diffraction and for optimization of the conditions on the home-source. All the data sets were acquired at BESSY II (Helmholtz Zentrum Berlin). For each crystal 200° were collected. Data was processed at the beamline with XDS (Kabsch, 2010).

SagB structure phases were determined by MAD with SeMet derivative. Data sets for Se-peak absorption point and in far spectrum were taken. Coordinates of Se and subsequent

phasing were determined by ShelX (Sheldrick, 2010). SagB native data set was phased using Phaser (McCoy, 2007) with a model obtained from MAD. AtlA data set was phased with Phaser as well, using SagB as a search model. Both structures were refined using MAIN (Turk, 2013).

3.2.7 Substrate preparation

B. subtilis was grown in LB overnight. LB medium was inoculated with overnight culture to OD₆₀₀ 0.03 and grown until OD₆₀₀ 0.7 and centrifuged at 14 000 g for 10 min. The pellet was resuspended in MiliQ water and snap-frozen in liquid nitrogen. Cell suspension was thawed in a boiling water bath and SDS was added to the final concentration of 4 %. The mixture was left to boil for 30 min. The material was sonicated for 2 min at 50 %. Insoluble fraction was washed with MiliQ water to remove SDS and then treated with 2 mg/mL pronase in 10 mM Tris-HCl pH 7.4 overnight at 60 °C. Proteases were removed by boiling insoluble material in 4 % SDS for 30 min. SDS was removed by washing insoluble material with MiliQ water. To remove secondary polymers peptidoglycan was treated with 40 % HF for 48 h at 4 °C. Resulting peptidoglycan was washed with water to remove HF. To dye peptidoglycan with Remazol Brilliant Blue, it was incubated in alkaline conditions in the presence of the dye for 8 h at 37 °C and for 16 h at 4 °C.

3.2.8 Activity assay

Buffers used in activity assay were 100 mM MES pH 6.5 for SagB, 100 mM Tris pH 8 for AtlA, 100 mM MES pH 5.5 for AtlE. SagB activity assay was conducted with 60 µM enzyme at 37 °C 250 rpm for 16 h. AtlA and AtlE activity assays were performed with 5 µM enzymes at 25 °C 250 rpm for 16 h. For all the conditions blank assays were conducted and deducted from the measured value.

3.2.9 Computational simulations

Computational work was performed by Jure Borišek from The National Institute of Chemistry.

3.2.9.1 Simulation of protein flexibility

The closed forms of AtlA-gl and SagB were created by independent superimposition of the R- and L-domains to the AtlE framework and energetically minimized. MD calculations were performed using the NAMD molecular modelling suite (Philips, 2005). AtlA-gl loop residues 128 to 140 from the crystal structure 6FXO that were not resolved in the experimental electron density map were constructed using Modeller (Webb, 2017). Enzymes and substrate were prepared according to our previously published protocol (Borišek, 2018). CHARMM-GUI environment (Kumar, 2007) was utilized for the protein manipulation and construction of the solvated protein-substrate complexes. CHARMM parameter and topology files (version 27) (MacKerell A. D., 1998; MacKerell A. D., 2004) were utilized to specify the force field parameters of the amino acid residues comprising the protein. CHARMM General Force Field (Vanommeslaeghe, 2010) was used to describe the atom types and partial charges of substrate. The final systems prepared for the MD simulation were composed of 39264 and 45160 atoms for AtlA and SagB, respectively. The system was minimized for 10000 steps by the steepest descent method followed by 10000 steps of the modified Adopted Basis Newton–Raphson method and an MD equilibration run of 1 ns. Production MD trajectory was generated using leapfrog integration scheme and 2 fs simulation step using SHAKE algorithm. A 500 ns long MD simulation production

run was performed for each model. Conformations were sampled every 1000th step resulting in 25,000 conformations for subsequent analysis. Visualization and analysis of the geometry parameters of the production MD trajectory was performed using VMD program (Humphrey, 1996).

3.2.9.2 Simulations involving catalytic mechanism elucidation

3.2.9.2.1 Molecular Dynamics of the AtIE Remodeled Complex with (NAG-NAM)₃

MD calculations of AtIE from the 4PIA crystal structure in complex with the remodeled substrate were performed using the CHARMM molecular modeling suite (Brooks, 2009). The remodeled substrate positioned in the complex was derived from a previous study (Mihelič, 2017). His79 and Lys80 from the 4PIA crystal structure were not resolved in the experimental electron density map and were constructed using The Discovery Studio (BIOVIA). The protonation states of all ionizable amino acid residues were assigned on the basis of the pKa estimated via PROPKA version 3.0 software (Olsson, 2011) at a pH of 5.5. Aspartate and glutamate residues were deprotonated, with the exception of the catalytic Glu138, which was protonated with the PROPKA pKa value of 6.36. Histidine residues, none of which were located near the active site, were modeled in their neutral states; their tautomeric state was assigned on the basis of hydrogen bonding using WHATIF (Vriend, 1990). The CHARMM-GUI environment (Kumar, 2007) was used for protein manipulation and construction of the solvated protein–substrate complexes. The CHARMM parameter and topology files (version 27) (MacKerell A. D.-M., 1998; MacKerell A. D., 2004) were utilized to specify the force field parameters of the amino acid residues comprising the protein. The CHARMM General Force Field (MacKerell A. D.-M., 1998; MacKerell A. D., 2004) was used to describe the atom types and partial charges of the substrate in the Discovery Studio module simulations. The protein and its (NAM-NAG)₃ unit substrate were immersed into a sphere TIP345 of water molecules with a truncated octahedral shape and an edge distance of 10 Å. A single potassium ion was added to make the system electroneutral. Ion placement was done using the Monte Carlo method. The periodic boundary conditions were applied on the basis of the shape and size of the system. Grid information for the Particle-Mesh Ewald Fast Fourier Transform was generated automatically. The final system prepared for the MD simulation was composed of 37246 atoms. To remove bad steric contacts, short steps of energy minimization were performed. The system was then minimized for 1000 steps by the steepest descent method followed by 1000 steps of the modified Adopted Basis Newton–Raphson method and an MD equilibration run of 350 ps. The production MD trajectory was generated using a leapfrog integration scheme and 2 fs simulation step using the SHAKE algorithm. A 100 ns MD simulation production run was performed. Conformations were sampled every 100th step, resulting in 25000 conformations for subsequent analysis. Visualization and analysis of the geometry parameters of the production MD trajectory were performed using the VMD program (Humphrey, 1996).

3.2.9.2.2 Definition of the QM Region for the QM/MM Simulations

The QM region for the QM/MM simulations was selected by taking into account the atoms present in the theoretical catalytic mechanisms of the glycoside hydrolase family members and consisted of a total of 43 atoms. To begin with, the Glu138 residue side chain was included in the QM region, as it is the only assumed catalytic amino acid directly involved in the catalytic process. Next, the identified potential catalytic water molecule, essential for attacking the anomeric carbon of NAG-1 in the water-assisted mechanism, was included. Furthermore, the scissile glycoside bond between the NAM+1 and NAG-1

substrates was also included as part of the QM region, since it represents the place where the reaction occurs. Finally, the whole NAG-1 unit of the substrate, including the anomeric carbon and acetamido side group essential for forming the oxocarbenium/oxazolinium ion like transition states/intermediates for both reaction pathways, directed proton transfer via Glu138, and water-mediated catalysis were included in the QM region.

3.2.9.2.3 Reactants, Products, and Oxazolinium Ion Generation for QM/MM Simulations

QM/MM calculations were performed using the CHARMM biomolecular simulation program (Eurenius, 1996). The ab initio and density functional theory (DFT) calculations were performed with General Atomic and Molecular Electronic Structure System (GAMESS) software (Schmidt, 1993) interfaced with CHARMM. The QM region was treated with the B3LYP/6-31G(d) level of theory. Molecular mechanics calculations were performed with a dielectric constant of 1 ($\epsilon = 1$) using a classical force-field shift method and a cutoff distance of 12 Å. Selected frames from the MD simulation were initially minimized using 10000 steps of an adopted-basis Newton–Raphson scheme to form the reactants for the QM/MM RPATH simulation. Next, the products were restrained from the structure of the reactants by making a bond between the Glu138-OH proton and the glycosidic oxygen, breaking the glycosidic bond between the glycosidic oxygen and the anomeric carbon atoms, and creating a bond between the OH– (water molecule) and the anomeric carbon. For the generation of the desired oxazolinium ion, the starting points were the product structures. Subsequently, the oxazolinium ion was restrained from the first product structure (units NAG-3, NAM-2, and NAG-1) on the NAG-1 unit by closing the ring via the formation of the bond between the acetylic O88 oxygen and the anomeric carbon atom and breaking the bond between the anomeric carbon and the glycosidic oxygen atoms. For the products and the oxazolinium ion generation CHARMM restrain distance (RESL) methodology, $kj = 100.0 \text{ kcal}/(\text{mol } \text{\AA}^2)$ was applied, and no weighting factor was used. RESL methodology was applied in 100 QM/MM minimization steps. Subsequently, the restraints were removed, and the system was minimized by another 10000 steps using the Adopted-Basis Newton–Raphson scheme.

3.2.9.2.4 Calculation of the Reaction A and B Pathways Using the Replica Path Method

The replica path (RPATH) method (Woodcock H. L., 2007) is an extension of the self-penalty walk method by Elber and co-workers (Czerninski, 1990; Elber, 1987). The 14 initial replicas for each of the two investigated reaction pathways, path A and path B, were generated by linearly interpolating the coordinates between all the atoms of the QM/MM geometrically optimized starting structure, A or B, and the desired corresponding product, A or B. The RPATH method uses a combined minimization involving the sum of the configurational energies coupled with the two RPATH penalty energy terms. One RPATH penalty term restrains the distances between adjacent replica points, ensuring that the generated pathway is smooth and evenly spaced. The second RPATH term restrains the angle between an adjacent and the next adjacent pathway. In this way, the RPATH method ensures that the optimized replicas represent the reaction pathway (Woodcock H. L., 2003; Woodcock H. L., 2007). QM/MM RPATH geometry optimizations of the initially generated path A and path B were performed using the Adopted Basis Newton–Raphson method (Woodcock H. L., 2007). Parameters included in the QM/MM replica penalty terms were set to the following values: $K_{\text{rms}} = 2000.0 \text{ kcal}/(\text{mol } \text{\AA}^2)$, $K_{\text{angle}} = 100.0 \text{ kcal}/(\text{mol } \text{\AA})$, and $\text{COSMAX} = 0.95 \text{ rad}$. Only the QM region was weighted in the replica path RMS calculation to avoid overestimating the contributions of the atoms not significantly

involved in the reaction (Woodcock H. L., 2007). Reaction pathways were minimized for 10000 steps or until the total pathway root mean-square gradient was less than 0.01 kcal/(molÅ) and the change in total pathway energy was less than 1.0 kcal/mol for at least 30 consecutive steps. Finally, the energy profiles were calculated for both the studied reaction mechanisms. Calculations were carried out on 128 CPU units (8 CPU/replica) within the cluster at the National Institute of Chemistry in Ljubljana.

3.2.9.3 Design of AtIE inhibitors

The starting point for the virtual screening was the remodeled AtIE substrate, derived from a previous study (Mihelič, 2017). As remodeled AtIE substrate contains 3 (NAG-NAM) units, we focused on the central NAG-NAM unit, next to the cleaved bond. LigandScout (Wolber, 2005) was used to explore interactions between the substrate and the enzyme and to derive structure-based pharmacophore for the NAG-NAM ligand. The pharmacophore model was used to screen approximately 5 million commercially available compounds, all of which were previously converted into a multifunctional format (25 conformers for each compound in the database) using the LigandScout screening module. The conformers of the molecules in the screening library were generated using the idbgen module available in LigandScout, coupled with the OMEGA software (Hawkins, 2010). The default high-throughput settings were used for the library generation: maximum number of output conformers per molecule=25; RMS threshold to duplicate conformers=0.8 Å; maximum number of all generated conformers per molecule=30,000; and maximum number of intermediate conformers per molecule=4000. Pharmacophore Fit scoring function was used to score the match of the hit molecules to the pharmacophore model. The LigandScout screening procedure retrieved approximately 10,000 hit compounds that were subsequently docked in the AtIE binding site. Molecular docking studies were performed using GOLD Suite v5.1 docking package (Verdonk, 2003). Binding site was defined around remodeled substrate coordinates of the NAG-NAM ligand in the radius of 12 Å. Validation of the docking settings was performed by redocking of the NAG-NAM ligand of the remodeled substrate into the AtIE binding site (Kirchamir, 2008). The observed RMSD of obtained docked poses versus the original NAG-NAM position were within the accepted limits and CHEMPLP scoring function gave the best results. Next, each ligand was docked into the binding site by the scoring function CHEMPLP and applying the following parameters of the GOLD genetic algorithm: population size=100, selection pressure=1.1, no. of Operations=100,000, no. of islands=5, Niche Size=2, migrate=10, mutate=95, crossover=95.42.

3.2.9.3.1 Surface plasmon resonance (SPR) experiments

All the SPR experiments were performed by Dr. Jure Borišek from the National Institute of Chemistry and Dr. Vesna Hodnik from the Biotechnical Faculty.

SPR experiments were performed using a Biacore T100 (GE Healthcare) equipped with a Series S sensor chip CM5 (GE Healthcare). PBS buffer (67 mM Na₂HPO₄, 12.5 mM KH₂PO₄, 70 mM NaCl, 0.05% P20, pH 7.4) was used for immobilisation of the protein. AtIE was attached to the surface of the chip by amine coupling. The surface of flow cells 1 and 2 was initially activated with a 10-min pulse of a 1:1 mixture of 0.4 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride and 0.1 M N-hydroxysuccinimide, according to manufacturer recommendations. AtIE was diluted into 10 mM sodium acetate buffer, pH 5.8, to a final concentration of 100 mg/mL, and injected twice for 180 s over the second flow cell. Both flow cells were blocked with a 7-min pulse of 1.0M ethanolamine. The final immobilization level was about 4100 RU. Screening was run at 25 °C using PBS

(67 mM Na₂HPO₄, 12.5 mM KH₂PO₄ pH 7.4, 70 mM NaCl, 0.05% P20) supplemented with 1% or 5% DMSO (Merck) as the running buffer. Selected compounds from the virtual screening studies were tested at two different concentrations: 40 and 400 mM. Each compound was injected for 1 min at a flowrate of 30 mL/min, and the dissociation was monitored for 30 s. Regeneration was not needed, although 30 s of buffer-flow was used to stabilize the surface after each injection. We compared the responses obtained by the injections of the buffer and the two different sample concentrations, and continued the study with the compounds that showed binding. Selected compounds were tested at eight different concentrations (depending on their solubility) in three parallel titrations. K_d values were determined by fitting the data to a 1:1 steady state binding model.

3.2.9.3.2 Nuclear magnetic resonance (NMR) experiments

All the NMR experiments were performed by Dr. Simona Golič Grdadolnik from the National Institute of Chemistry.

The high-resolution NMR spectra were recorded on a Varian DirectDrive 800 MHz spectrometer at 25 °C. All data were collected using pulse sequences and phase-cycling routines provided in Varian libraries of pulse programs. The cryogenic triple-resonance NMR probe was used. NMR samples were prepared in a buffer containing 20 mM Tris-D11, 10% glycerol-D8, and 100 mM NaCl in D₂O, pD 7.2. DSS (0.1 mM) was used as an internal standard. All the spectra were recorded at a protein/ligand ratio of 1:100, the protein concentration was 1 mM and the ligand concentration was 0.1 mM. The STD experiments were performed with a 8802.8 Hz spectral width, 16384 data points, a saturation time of 2 s, a relaxation delay of 1.3 s, and 3800 scans. Selective saturation was achieved by a train of 50 ms long Gauss-shaped pulses separated by 1 ms delay. Water was suppressed via excitation sculpting. The on-resonance selective saturation of AtIE was applied at -0.9 ppm, which is well separated from the methyl group of compounds at 1.1 ppm. The off resonance irradiation was applied at 30 ppm for the reference spectrum. Subtraction of the on and off-resonance spectra was performed internally via phase cycling. Spectra were zero-filled twice and apodised by an exponential line-broadening function of 1 Hz.

Chapter 4

Results

4.1 Protein Sequence Analysis

In CAZy five proteins from GH73 family from *S. aureus* Mu50 have been identified, AtlE (SAV2307), SagB (SAV1775), ScaH (SAV2644), AtlA (SAV1052) and SAV0909. Their sequences were analyzed and their predicted domain organization is presented in Figure 7. AtlE and SagB have predicted transmembrane regions on N-terminal followed by a glucosaminidase domain. This is consistent with the findings of Chan et al., who have shown that SagB is membrane bound (Chan, 2016). ScaH, AtlA and SAV0909 have a N-terminal signal peptide that directs them into the extra cellular milieu followed by two domains with different enzymatic activities. ScaH has a glucosaminidase domain followed by cysteine, histidine-dependent amidohydrolase/peptidase (CHAP) domain and AtlA consists of N-terminal amidase domain and C-terminal glucosaminidase domain. AtlA is posttranslationally cleaved between second and third cell wall binding module (Oshida, 1995). SAV0909 has an N-terminal CHAP domain and C-terminal glucosaminidase domain.

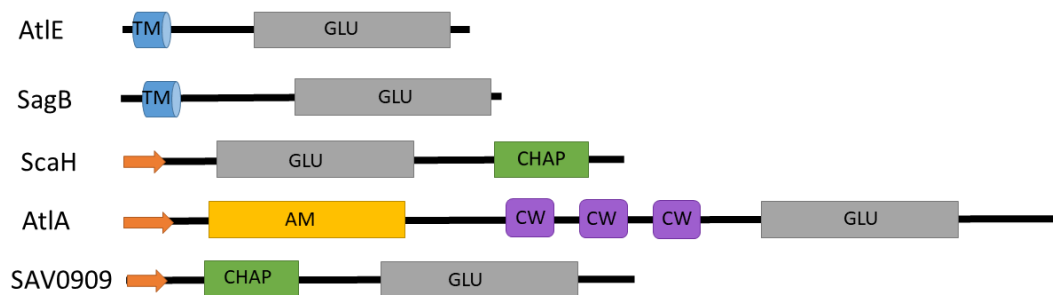


Figure 7: Domain organization of proteins from GH73 family of *S. aureus* Mu50.

4.2 Sequence Analysis of GH73 Enzymes from *S. aureus*

All *S. aureus* GH73 sequences were harvested from CAZy database (Cantarel B. L., 2009) and their sequence analysis revealed they encompass seven different proteins (Figure 8). These proteins are AtlE, SagB, AtlA, ScaH, SACOL1576, LytP_{NM} and LytP_{JB}. The first four enzymes are present in almost all strains, only SagB was absent in one of the 32 analyzed strains in Aureowiki, and seem to represent the core peptidoglycan *N*-acetylglucosaminidases of *S. aureus*. The remaining three enzymes are less widely represented across different strains, SACOL1576 in 28%, LytP_{NM} in 19 % (although

SAV0909 was not found in Mu50 in Aureowiki) of the strains and LytP_{JB} could not be searched in Aureowiki. LytP_{NM} and LytP_{JB} were identified as phage lysins, from phage ϕNM and phage ϕJB , respectively. Phage proteins are not expected to be important for normal bacterial cell processes, on the contrary, they are thought as destructive forces used by phages to lyse their host (Rodriguez-Rubio, 2013). SACOL1576 is identified as TraG, bacterial conjugation protein, by some BLAST hits, but InterPro does not identify domains present in this sequence as part of TraG\VirD4 secretion system. In our analysis of GH73 proteins from CAZy we identified 192 AtIA, 190 ScaH, 189 SagB, 161 AtIE, 53 LytP_{NM} , 43 SACOL1576 and 40 LytP_{JB} sequences.

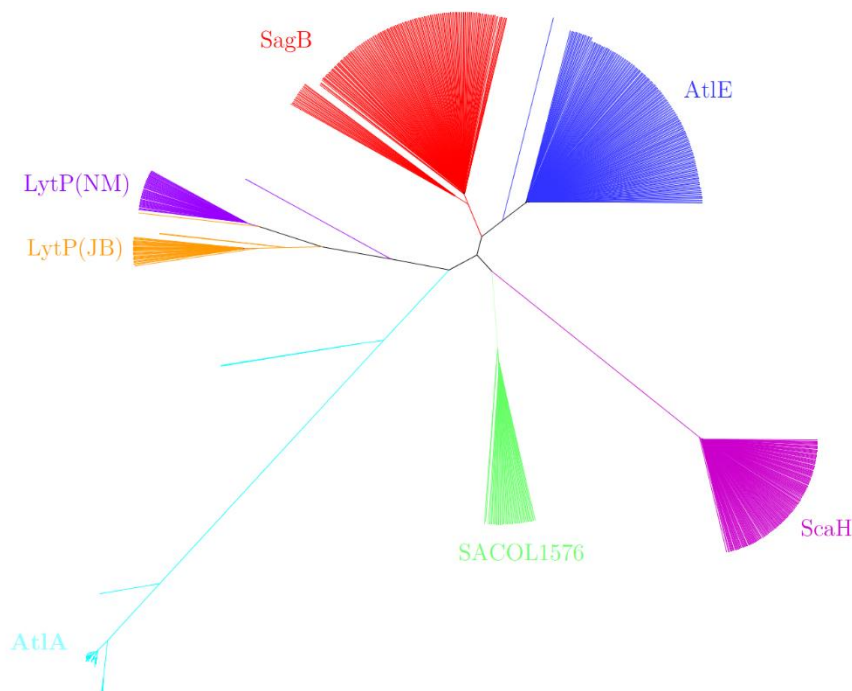


Figure 8: A phylogenetic tree of all the GH73 enzymes from *S. aureus* species.

Sequence conservation of most of the individual proteins among different strains is high, only 0.78 % of amino acid sequence in AtIE, 1.94 % in ScaH, 4.75 % in SagB and 5.83 % in AtIA exhibits less than 95 % conservation. 11 % of SagB sequences carry N-terminal insertion of 11 amino acids, which accounts for lower sequence conservation compared to other enzymes. With this region left out of analysis, the sequence conservations of less than 95 % fall to 1.06 %. Similarly, 8% of AtIA sequences have a 7 amino acid N-terminal insertion, as well as lower conservation region between Thr89 and Thr219, preceding amidase domain, where the sequence conservations of less than 95 % are 23.66%, but average sequence conservation for this region is still 92.38%. This analysis indicates the importance of four of the enzymes from *S. aureus*.

4.3 Cloning

4.3.1 Different protein constructs cloning

Constructs AtIA-gl+1CWB, AtIA-gl1, AtIA-gl2 and AtIA-am were already cloned and provided in vectors.

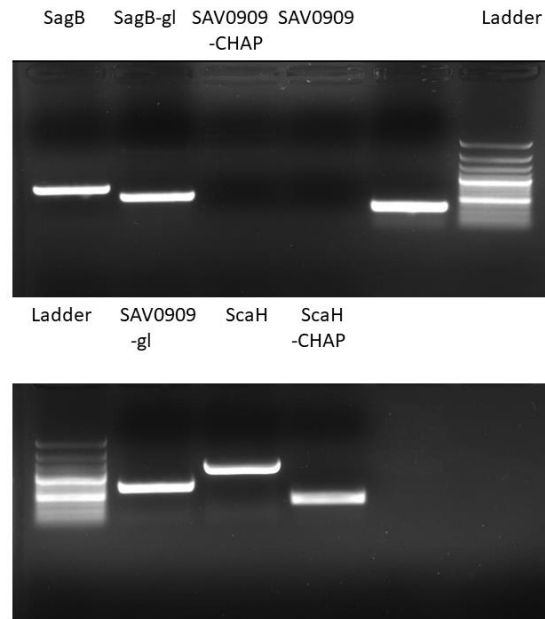


Figure 9: 1 % AGE gel depicting amplified DNA, coding for appropriate constructs, with 100 bp DNA Ladder.

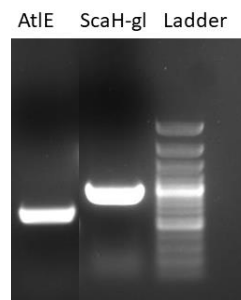


Figure 10: 1 % AGE gel depicting amplified DNA, coding for appropriate constructs, with 100 bp DNA Ladder.

Most of the constructs were successfully cloned (Figure 9 and Figure 10), as confirmed by sequencing, except SAV0909-CHAP domain and full length SAV0909, which were not successfully amplified by PCR in the first step of cloning.

4.3.2 Mutant construction

Sequencing confirmed that all mutants were successfully constructed.

4.4 Protein Production and Purification

4.4.1 Production of different constructs

All the proteins were produced and soluble, except SagB-gl. All the proteins, except AtIA-gl2, where the purification failed, purified successfully (Figure 11-19).

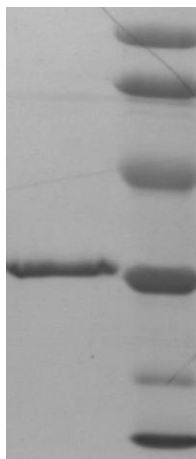


Figure 11: Purified SagB.

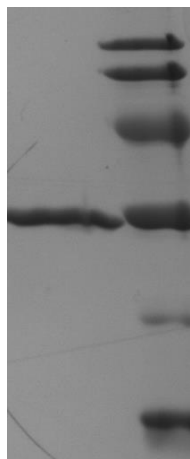


Figure 12 Purified SAV0909-gl.



Figure 13: Purified ScaH.

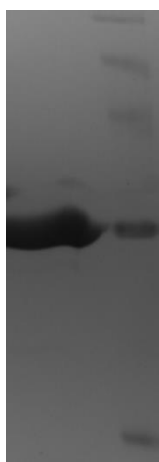


Figure 14: Purified ScaH-gl.

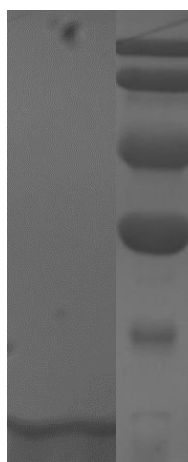


Figure 15: Purified ScaH-CHAP.

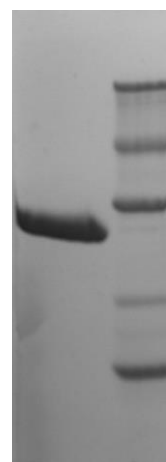


Figure 16: Purified AtlE.

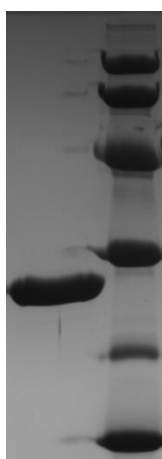


Figure 17: Purified AtlA-gl.

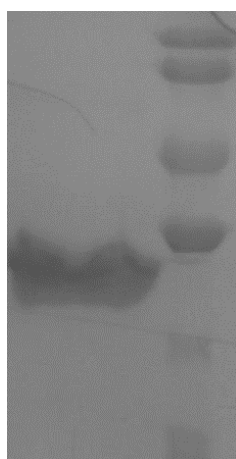


Figure 18: Purified AtlA-gl1.

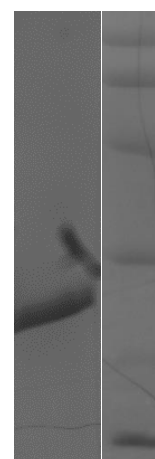


Figure 19: Purified AtlA-am.

ScaH-CHAP and AtlA-gl1 have proven themselves difficult to concentrate. Concentrating ScaH-CHAP beyond 4 mg/mL resulted in precipitation. AtlA-gl was chosen to continue to work with instead of AtlA-gl1.

4.4.2 Production of mutant proteins

All the mutant proteins, except At1A-gl D215A, were successfully produced and purified. We were able to obtain a soluble D215N At1A-gl protein. CD spectra were taken for to ensure that all mutants were correctly folded.

4.5 DSF

4.5.1 Wild type proteins

4.5.1.1 SagB

SagB produced well shaped melting curves for most of the conditions (Figure 20).

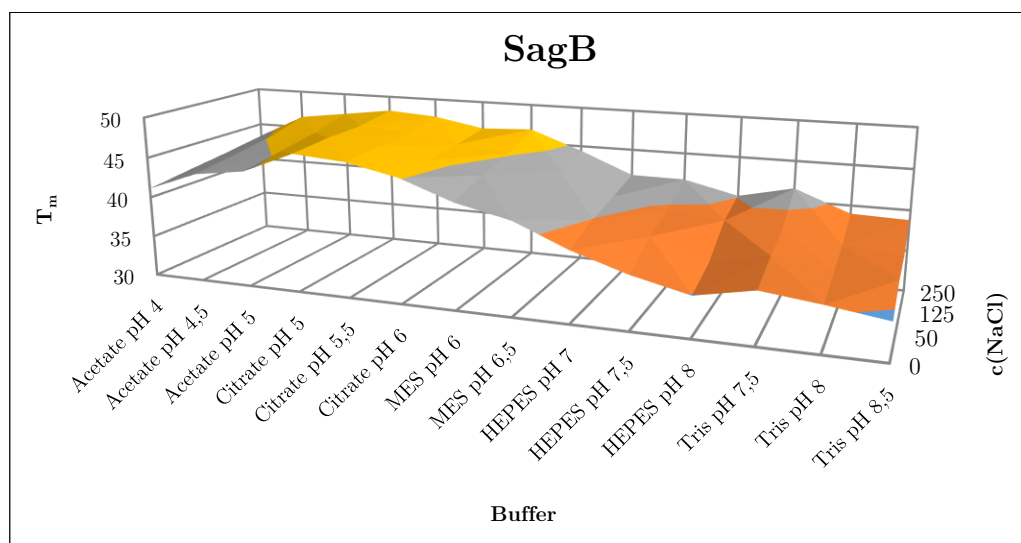


Figure 20: Thermal stability of SagB in different buffers, with different pH and ionic strengths. The buffer for its optimal stability has been determined as citrate pH 5 250 mM NaCl.

4.5.1.2 SAV0909-gl

SAV0909-gl produced well shaped curves from pH 4 to 6.5, beyond that the curves were deformed, as it seems the protein denatured even before the first measurement was taken (Figure 21).

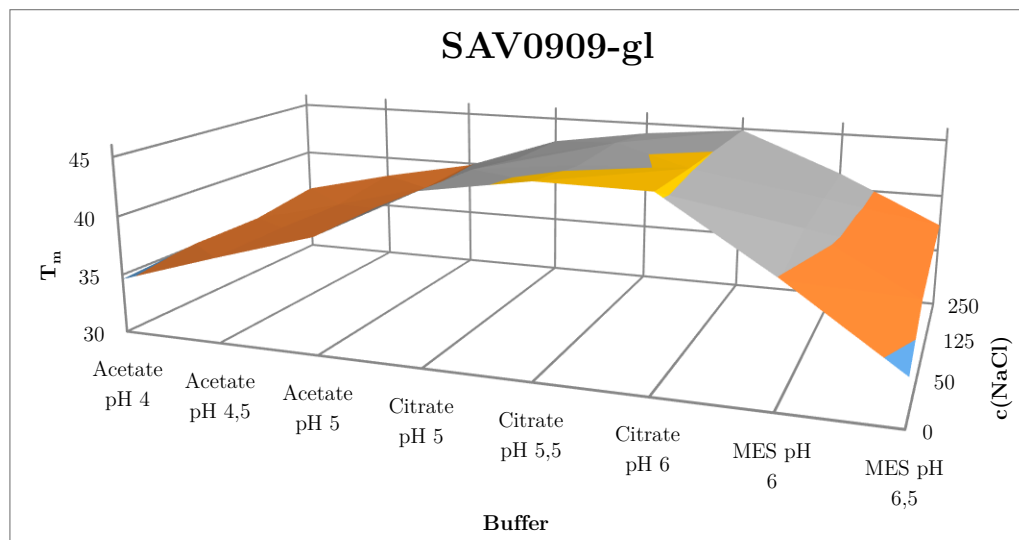


Figure 21: Thermal stability of SAV0909-gl in different buffers, with different pH and ionic strengths. It is the most thermally stable in citrate pH 5.5 with T_m 45.5 °C.

Full length ScaH is not appropriate for DSF analysis, as it has two domains and thus produces a melting curve with a two-step transition, which is not appropriate for interpretation.

4.5.1.3 AtIA-gl

AtIA-gl produced very nicely shaped curves, as it is very thermostable with T_m above 50 °C across the range (Figure 22).

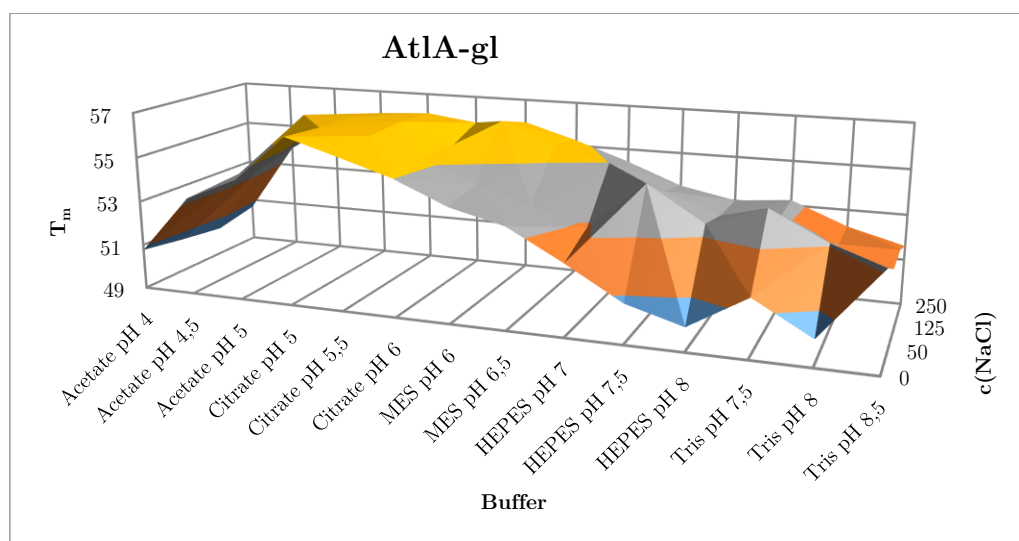


Figure 22: Thermal stability of AtIA-gl+1CWB in different buffers, with different pH and ionic strengths. It was most stable in MES pH 6 125 mM with T_m 55.6 °C.

4.5.1.4 ScaH-gl

ScaH-gl mostly produced curves that were easily interpreted (Figure 23).

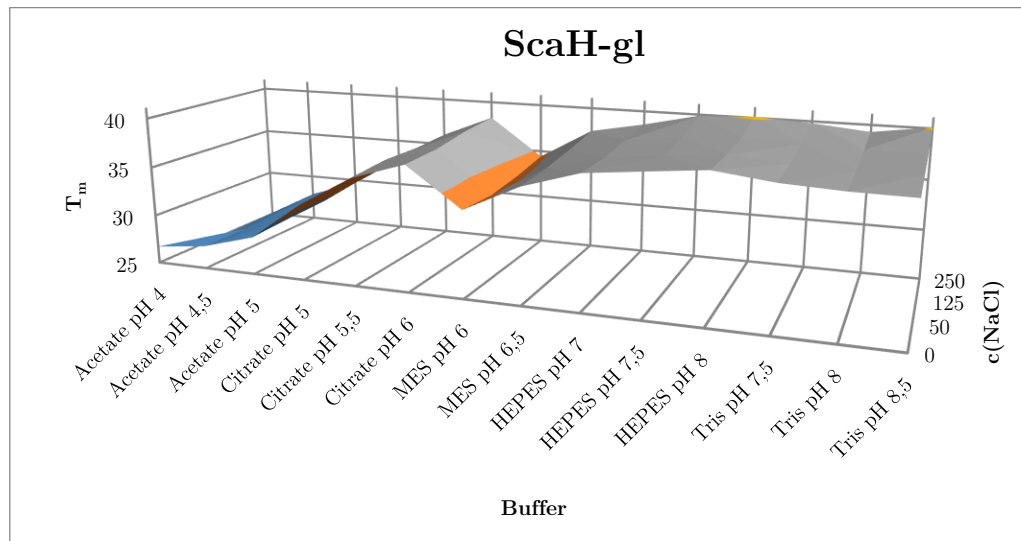


Figure 23: Thermal stability of ScaH-gl in different buffers, with different pH and ionic strengths. Its optimal stability was in Tris pH 8.5 250 mM NaCl with T_m 40.1 °C.

ScaH-CHAP did not produce a melting curve, as it seemed to be already denatured at 25 °C and thus no other assays were performed with it.

4.5.1.5 AtIE

AtIE mostly produced interpretable curves (Figure 24).

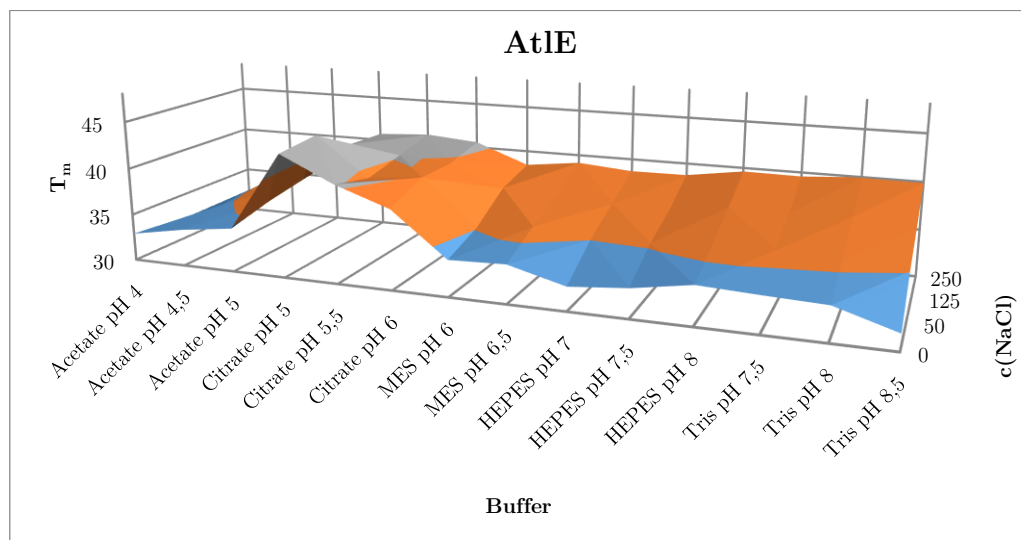


Figure 24: Thermal stability of AtIE in different buffers, with different pH and ionic strengths. It was determined to be most stable in citrate pH 5 either 50 mM NaCl with T_m 48.7 °C.

4.5.2 Mutant proteins

4.5.2.1 SagB

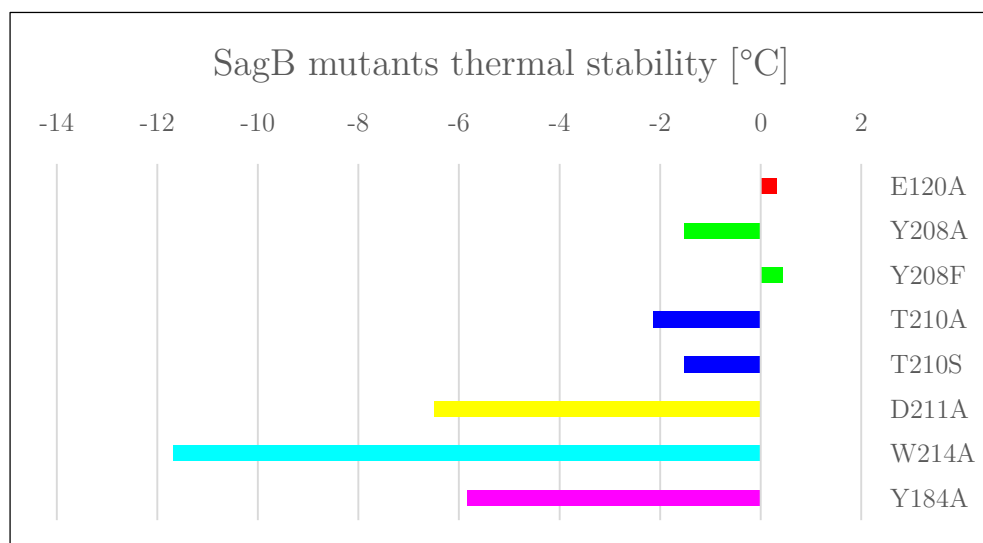


Figure 25: Thermal stability of SagB mutants relative to wild type protein.

As expected, the least stable proteins are those with aromatic residues located in helices mutated to alanine, W214A and Y184A (Figure 25). D211A mutation greatly destabilized the protein. All other mutants were not significantly affected by mutations.

4.5.2.2 AtIA-gl

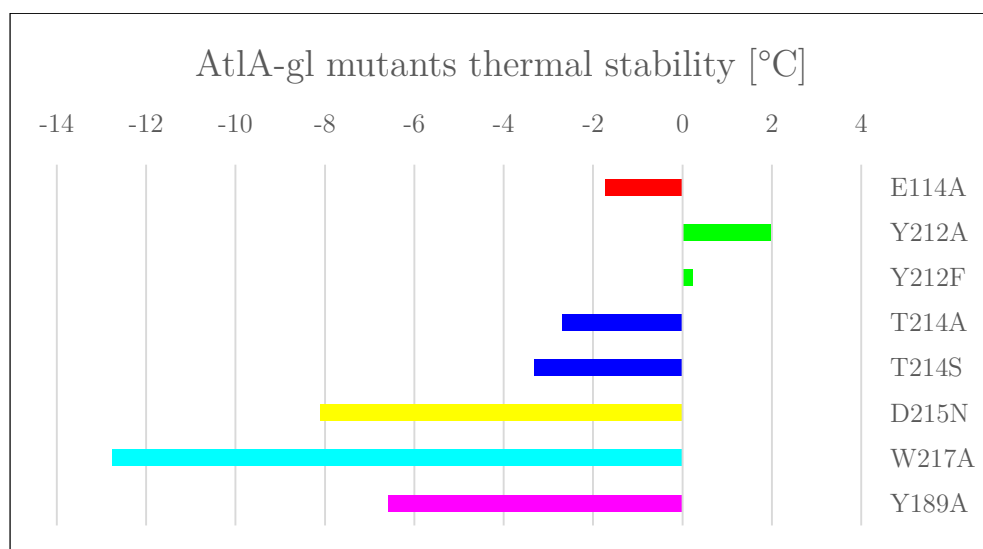


Figure 26: Thermal stability of AtIA-gl mutants relative to wild type protein.

Thermal stability of AtIA mutants, similarly to SagB, W217A, D215N and Y189A was afflicted most significantly (Figure 26).

4.5.2.3 AtIE

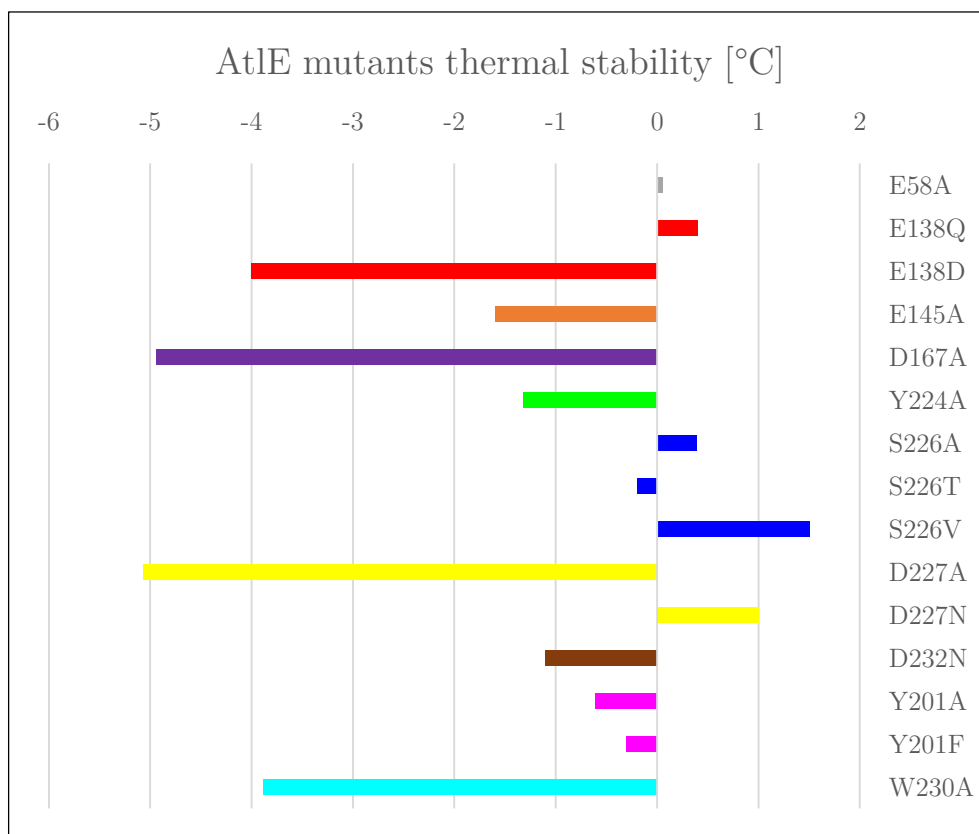


Figure 27: Thermal stability of AtIE mutants relative to wild type protein.

Most of the AtIE mutant proteins were not significantly affected by mutations, but E138A, D167A, E138D, 227A and D232A were significantly less thermally stable than the wt (Figure 27). E138A and D232A were so unstable that they were denatured at 25 °C and did not produce a melting curve.

4.6 Crystallization

4.6.1 SagB

With buffer optimization for stability of the enzyme, SagB produced hexagonal crystals. Its crystallization conditions were optimized (Figure 28).

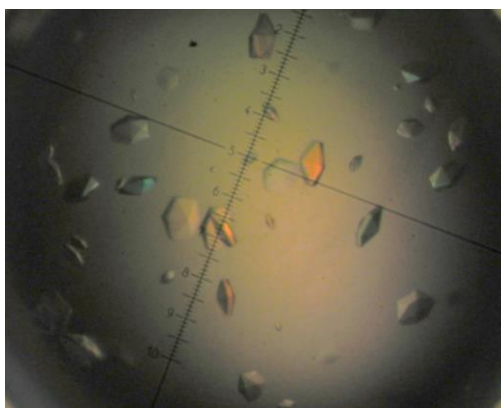


Figure 28: SagB crystals, formed in 50 mM trisodium citrate pH 4.75, 22.5 % PEG 200.

4.6.2 AtlA-gl

AtlA-gl+1CWB formed crystals from precipitate which were irreproducible and grew almost needle-like crystals from a single point (Figure 29). They were successfully broken apart and one of the bigger 3D parts of the crystal was caught on the loop.

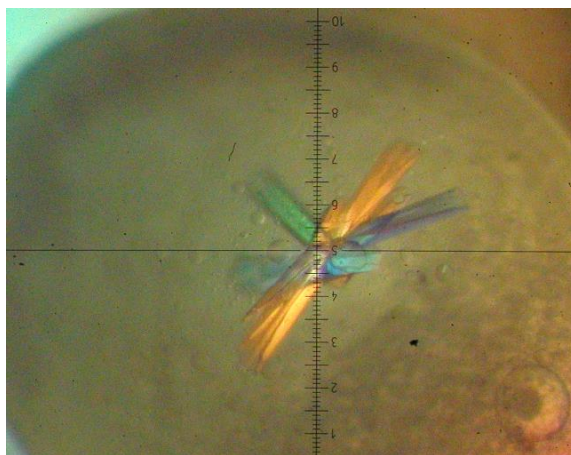


Figure 29: AtlA-gl+1CWB crystal.

4.6.3 ScaH

ScaH formed needle-like crystals (Figure 30). Optimization to obtain 3D crystals was unsuccessful. The best crystal diffracted to 5 Å. In the end the structure was determined by Minasov, G., et al. (PDB code: 5T1Q) from Center for Structural Genomics of Infectious Diseases from Chicago, USA.



Figure 30: ScaH crystals.

4.7 Data Collection and Refinement Statistics

SagB and AtIA-gl data sets were successfully taken and processed (Table 6). Intensities of the reflexes for AtIA-gl were found to be low and this fact is also reflected in the final statistics.

Table 6: Data collection and statistics table.

Structure	AtIA	SagB	SagB_SeMet_remote	SagB_SeMet_peak
PDB code	6FXO	6FXP		
Crystal parameters				
Resolution range (Å)	41.68-2.28, (2.41-2.28)	47.66-2.03 (2.15-2.03)	48.1-2.19 (2.32-2.19)	48.12-2.35 (2.50- 2.35)
Space group	P 64 2 2 (181)	P 61 2 2 (178)	P 61 2 2 (178)	
Unit-cell parameters				
a (Å)	110.17	151.77	152.78	152.82
b (Å)	110.17	151.77	152.78	152.82
c (Å)	92.39	122.51	123.83	123.82
(°)	90	90	90	
(°)	90	90	90	
(°)	120	120	120	
Data collection				
Beamline	BESSY 14.1.	BESSY 14.3.	BESSY 14.1.	
Wavelength (Å)	0.9184	0.894290	0.977031	0.979346
Total reflections (unique)	331263 (15697)	988582 (53891)	1741733 (82853)	1401570 (66990)
Multiplicity	21.1	18.34	21.02	20.92
Completeness (%)	99.8 (99.3)	99.7 (98.5)	99.8 (99.0)	99.9 (99.6)
Mean I/ (I)	25.31 (1.96)	27.07 (2.02)	17.06 (1.83)	15.67 (1.77)
Wilson B factor (Å ²)	61.38	37.76	46.18	50.51
Mosaicity (°)	0.103	0.068	0.081	0.087
CC(1/2)	100.0 (63.1)	99.9 (70.4)	99.9 (62.8)	99.9 (58.6)
R _{merge}	9.1 (162.6)	9.1 (105.0)	17.5 (192.7)	20.0 (205.5)
Refinement statistics				
R _{work}	0.2983	0.2120		
R _{free}	0.3309	0.2197		

4.8 Description of the Structures

4.8.1 SagB

SagB crystallized in the space group $P6_122$ with two molecules per asymmetric unit. The structure of SagB is well resolved by the electron density maps along the whole chain with the exception of N36 to A41 region in both molecules as indicated by Ramachandran outliers and higher atomic B-values as well as a few other side chains. The crystal structure also contains eight PEG molecules (PG4), two chloride and one sodium ion. The latter is positioned at a crystallographic two-fold axis. The two molecules superimpose with root-mean-square-deviation (RMSD) over all CA atoms of 0.46 Å. The superimposition parameters slightly deviate from a proper non-crystallographic symmetry dimer. The structure of SagB has a lysozyme-like fold and is similar to the heart-like shape of AtlE (4PIA) (Mihelič, 2017). It is composed of the right (R) and left (L) domain that can be divided in core and lobe regions. The catalytic residue E121 is positioned at the C-terminus of helix $\alpha 10$. The R-domain comprises residues from A1 to D71 and S187 to K249, whereas the L-domain from L72 to S186. The two domains of molecule A superimpose to their counterparts in molecule B with similar RMSDs (0.46 Å, 0.38 Å) as the two molecules on each other indicating equivalent domain packing.

4.8.2 AtlA-gl

AtlA-gl crystallized in space group $P6_122$ with one molecule per asymmetric unit. Its structure is well resolved by the electron density maps with the exception of the two N-terminal amino acid residues and the loop G127-K142 in spite of weak diffraction data. The electron density maps show the position of the loop region, however, density maps were not interpretable in any of the space groups with lower symmetry tested. The crystallographic R-factor of 0.30 reflects the irreproducible crystal, its small size and low intensity of diffraction data. AtlA-gl is similar to the SagB structure and shares the topological layout of AtlE in the core and lobe regions.

4.8.3 Comparison of the *S. aureus* GH73 structures

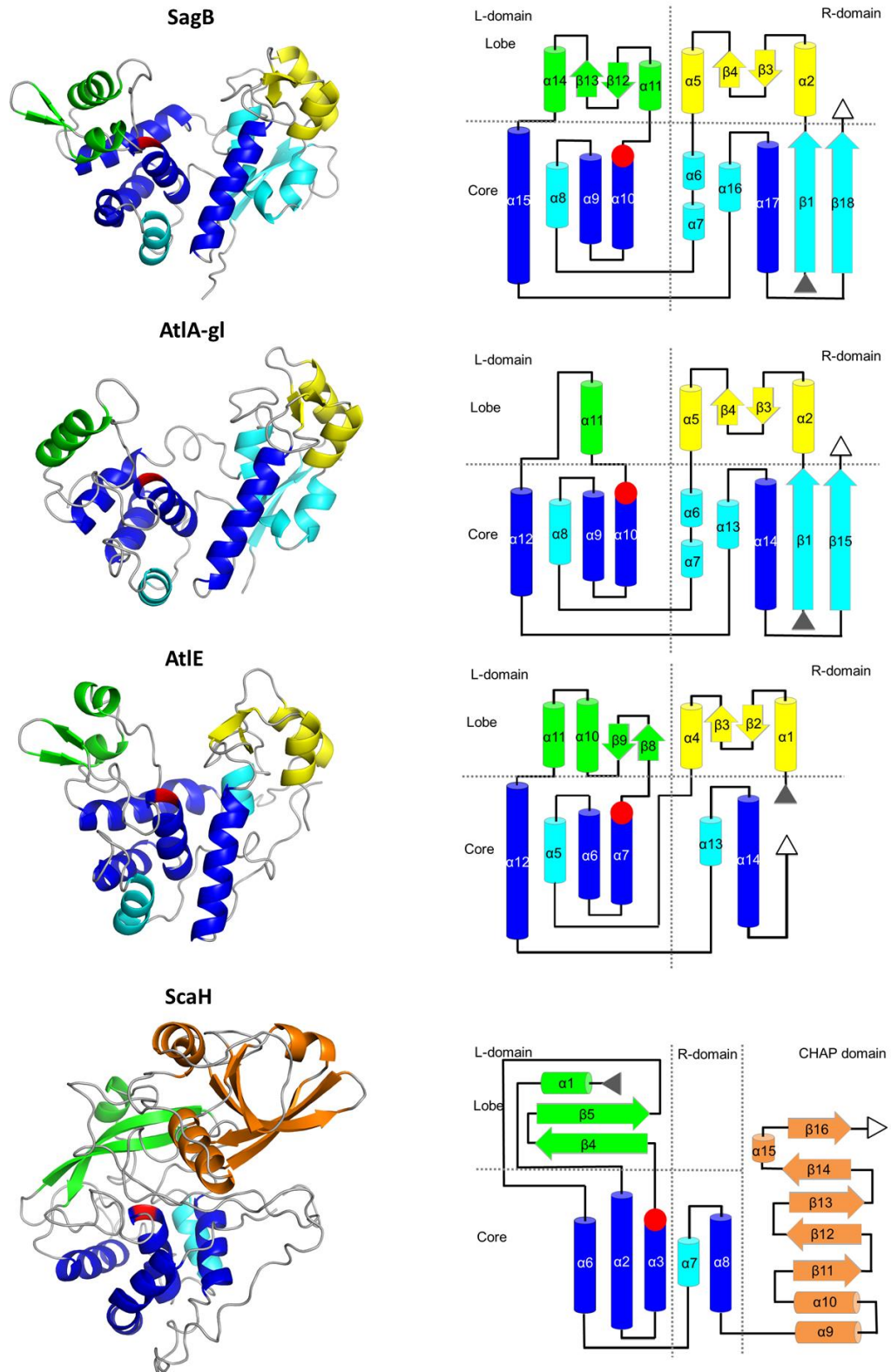


Figure 31: Structures of the GH73 enzymes in the ribbon representation (on the left) and the schematic representation of their secondary structure organization (on the right).

AtlE, SagB and AtlA-gl share secondary structure arrangement as well, however the L- and R-domains are positioned differently relative to each other (Figure 31). To quantify this the structures were superimposed as a whole and in parts with FatCat²⁸ interface imbedded in MAIN. Superimposition of SagB and AtlE of whole structures yielded 3.4 Å RMSD over 195 residues with 28% of sequence identity, whereas the superimposition of the two parts, corresponding to R- and L- domains, yielded RMSD of 2.2 Å over 104 matched residues with 28 % identity, and RMSD of 1.7 Å over 98 matching residues with 36 % identity for the R- and L-domain, correspondingly. Moreover, the angle of rotation of the SagB structure superimposed to AtlE R-domain with the SagB structure superimposed to AtlE L-domain is 31° with 3.5 Å screw translation component. This comparison clearly demonstrates that SagB and AtlE structures differ significantly.

A similar comparison with the AtlE structure as described above for the SagB structure yielded RMSD of 4.0 Å over 199 matching residues with 14% identity, whereas the superimposition of R- and L-domains yielded 2.2 Å and 1.7 Å, correspondingly. The rotation angle between the two different parts of AtlA-gl is 39°, screw translation is 4.6 Å. The packing of R- and L-domains in AtlA-gl structure thus differs from the AtlE structure even more than the SagB structure.

The major difference among the three structures is in the helix connecting both lobes. In the AtlA-gl structure it is possible to observe that the 4th turn of the helix α 13 is stretched out compared to SagB, where this helix is slightly longer and in AtlE it is even longer. The equivalent helix, α 12, in AtlE distorts slightly after the fourth turn with Y201 pulled slightly out at the center of the helix perturbation. The α 15 of SagB is almost a whole turn shorter and shares the same minor widening of, in this case, last turn of the helix with F184 drawn out.

The positioning of the core helices, which are conserved in this family (Mihelič, 2017), is equivalent to that of AtlE in 7 other known crystal GH73 structures. This includes ScaH, where the rest of the topological layout is completely different and much more akin to *StFlgJ* (5dn4), *Auto* (3fi7) and *SpFlgJ* (2zyc).

4.8.4 Substrate binding site

As a consequence of different positioning of L- and R-domains substrate binding sites in SagB and AtlA-gl are much wider in comparison to AtlE. This is reflected in the distances between the C α atoms of the catalytic Glu residue (138 in AtlE, 124 in SagB and 118 in AtlA) from the L-domain and the conserved YA(S/T)D region (224-227 in AtlE, 208-211 in SagB and 121-215 in AtlA) positioned in the R-domain at the opposite side of the active site cleft (Figure 32). The shortest distances were measured in the AtlE structure in the range from 10.8 Å to 12.8 Å, the SagB structure in the range from 17.6 Å to 20.5 Å, and the AtlA-gl structure from 17.7 Å to 21.8 Å. As indicated by the approximately 8 Å difference in the distance between the catalytic Glu from the YA(S/T)D region, the substrate binding cleft in the AtlA-gl and SagB structures is significantly wider than in the AtlE structure (Figure 33). We called SagB and AtlA-gl crystal structures open, whereas AtlE structure closed.

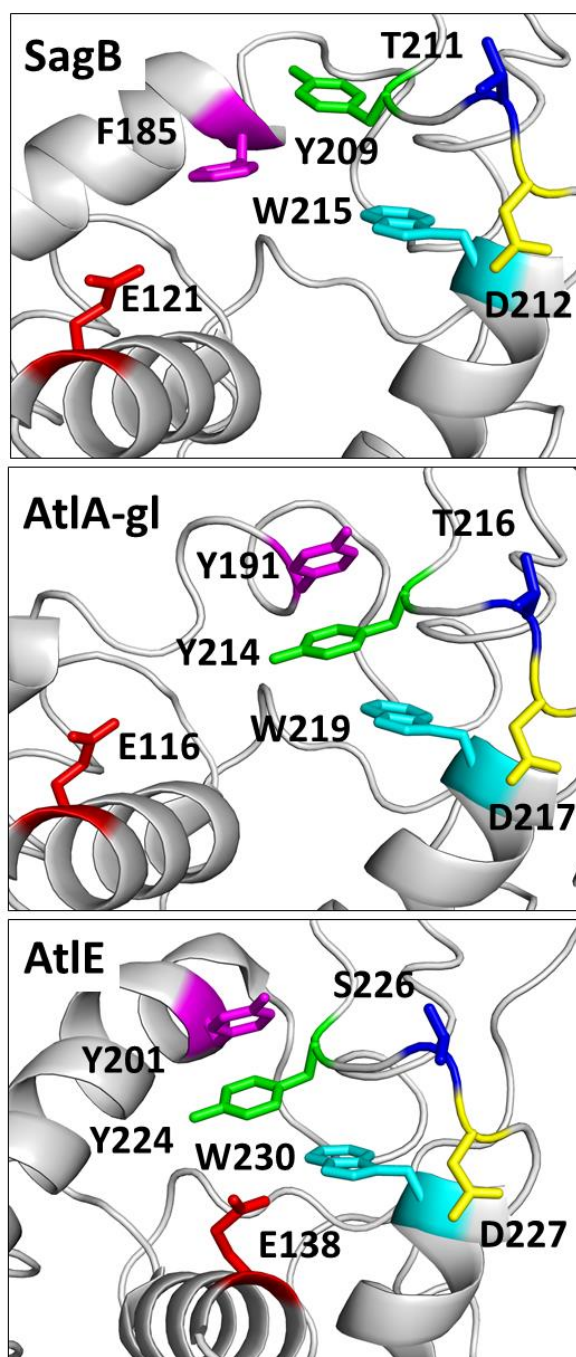


Figure 32: The residues studied by point mutations. Equivalent residues are coloured with the same colour. Substrate binding site residues identified in AtlE are conserved among the investigated enzymes, with minor variation, where Y201 of AtlE is F184 in SagB. Note the fact that S226 has a double conformation in AtlE crystal.

To gain insight in the possible (NAG-NAM)₃ substrate binding in the crystal structures of AtlA-gl and SagB, we used the binding geometry established in our AtlE studies (Mihelič, 2017; Borišek, 2018). The difference in separation between the catalytic residue and YA(S/T)D regions, positioned on the left and right hand side of the active site cleft, suggested to superimpose the AtlE substrate complex to R- or L-domains of SagB and AtlA-gl structures independently. This resulted in two substrate binding geometries of each enzyme (Figure 33).

When L-domains were superimposed, the substrate retained position over the catalytic glutamate, and when R-domains were superimposed, the substrate retained the contacts with the identified substrate binding regions. The NAM(-2) and NAG(-1) residues of the pair of substrate models positioned on the left clash with the features of the SagB and

AtlA structure indicating that binding at the catalytic site is impossible without a conformational change of the structure of the enzymes.

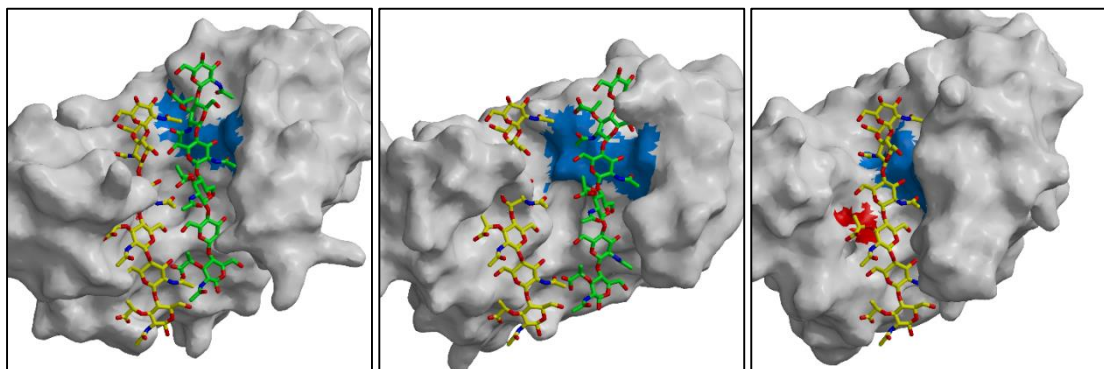


Figure 33: Binding of the substrates to the *S. aureus* enzymes of GH73 family with SagB on the left, AtlA-gl in the middle and AtlE on the right. The binding was modelled by the model of AtlE substrate binding.

4.9 Activity Assay of the Mutant proteins

4.9.1 SagB

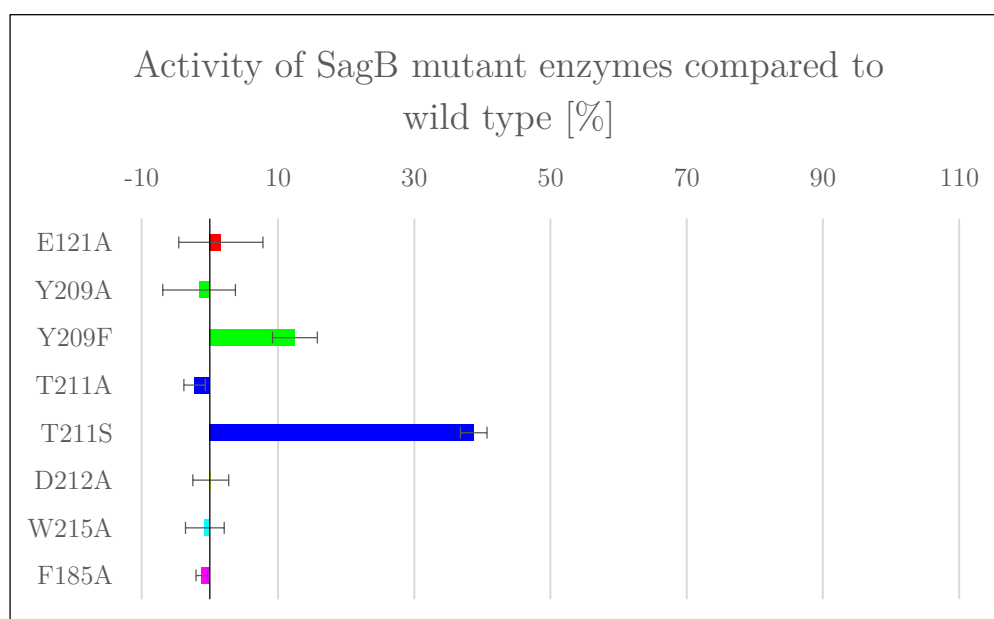


Figure 34: Activity of SagB mutants expressed as a percentage of activity of the wild type protein.

SagB has turned out to be a very sensitive enzyme, as all mutations had a significant effect on its activity (Figure 34). The assay is less sensitive for SagB than for the other two enzymes, as demonstrated by larger standard deviations. Analysis shows that E121 is the active site residue, as for other GH73 enzymes. The conserved YASD region was significantly affected by mutations and mutations to similar residues, as Y209F and T211S has recovered the activity of the enzymes to some extent. The two aromatic residues that

are adjacent to YASD region, W215 and F185, also appear to be important for the binding of the substrate.

4.9.2 AtIA-gl

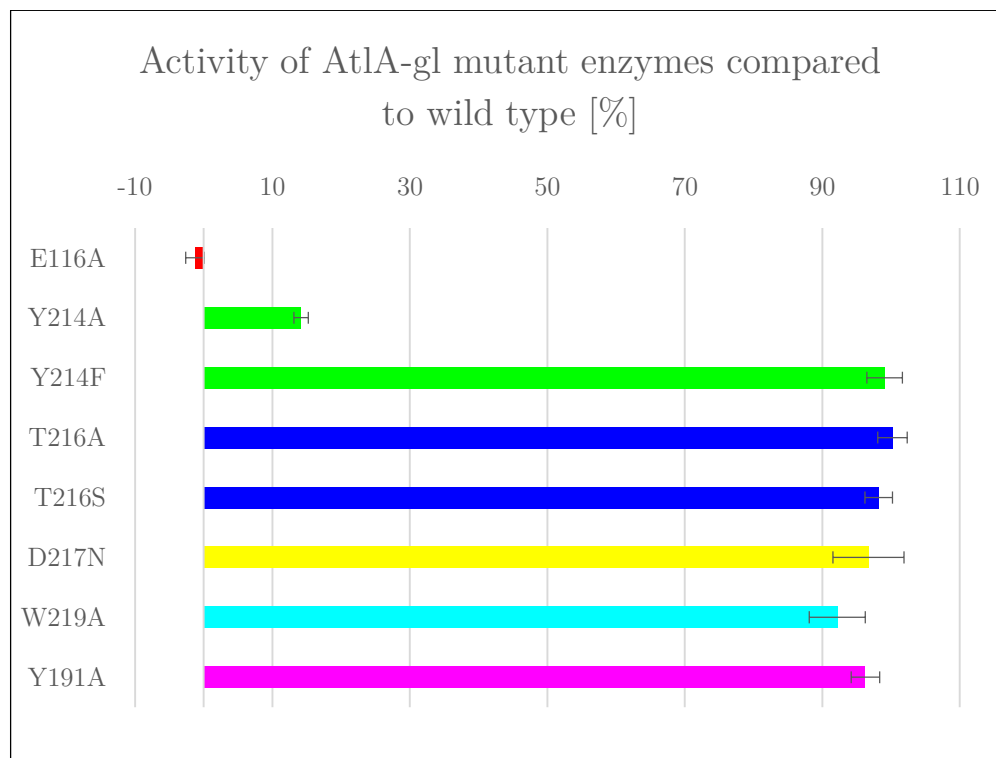


Figure 35: Activity of AtIA-gl mutants expressed as a percentage of activity of the wild type protein.

In complete contrast to SagB, AtIA-gl is a very robust enzyme, as only two single point mutations affected its activity (Figure 35). Consistently E116A was determined to be the catalytic amino acid. The only other amino acid affected by mutation was Y214 from the YATD region.

4.9.3 AtIE

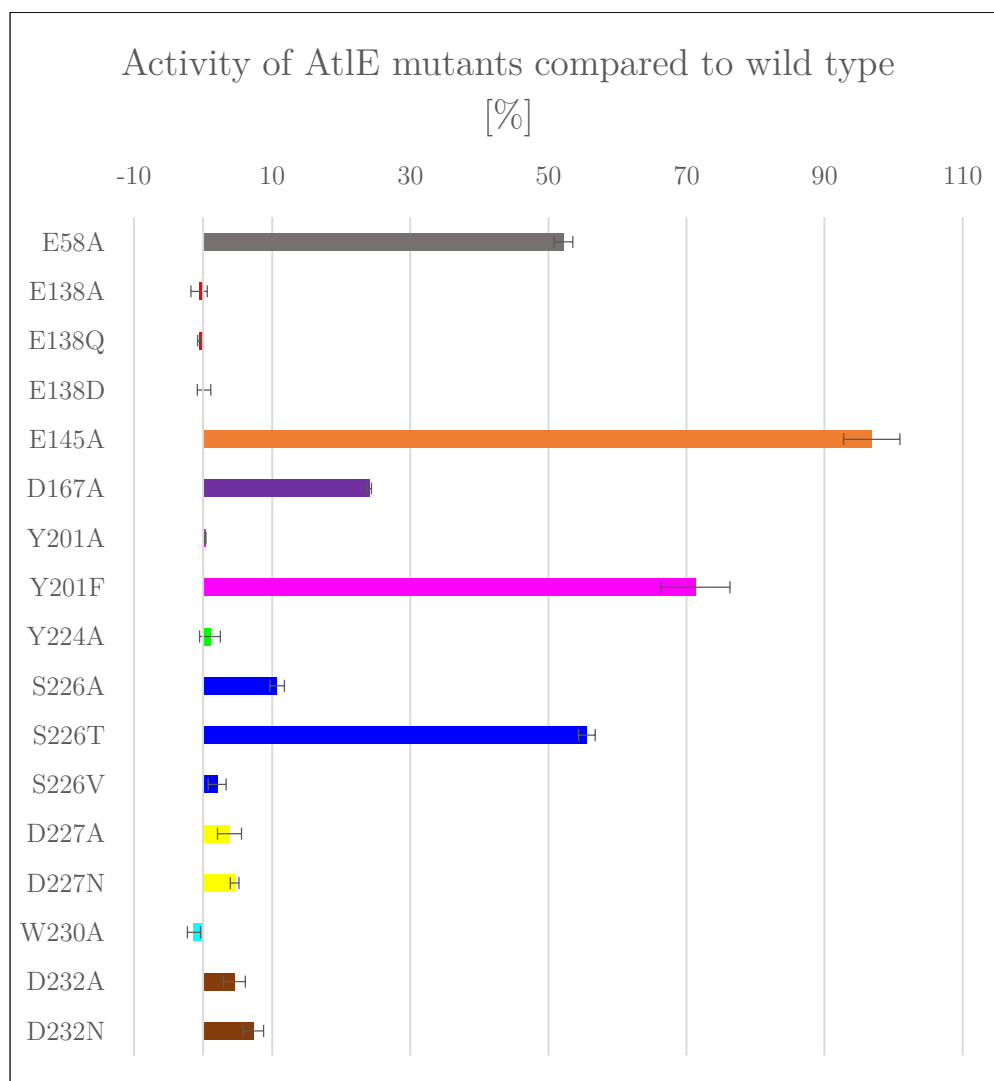


Figure 36: Activity of AtIE mutants expressed as a percentage of activity of the wild type protein.

AtIE was studied more extensively than the other enzymes, as a part of the study of the catalytic mechanism. All acidic amino acids in the range of 15 Å from the catalytic E138 were mutated. As well as the YASD region and the two adjacent aromatic amino acid residues were shown to be important for the activity of AtIE (Figure 36). Additionally, E58 and E145 had a minor effect on the activity of the enzyme and together with their position in the enzyme we postulate they are not involved in catalysis. D167A mutant exhibited 24 % activity of the wild type, alas we contribute the lower activity to its lowered stability. D232 mutants displayed a very low activity. It is positioned in the middle of α -helix facing away from the active site and a conformational change to bring it closer seems extremely unlikely.

4.10 Simulation of Flexibility

4.10.1 SagB

Mutations of the AtlA-gl and SagB have clearly demonstrated that residues from both domains are important for the activity of the enzymes. The AtlE substrate binding data (Mihelič, 2017) and the evidence presented here indicate that the (NAG-NAM)₃ substrate has to interact with the catalytic glutamate on the left as well as the YA(S/T)D region on the right simultaneously. This suggests that a substantial conformational change is necessary in order for the reaction to transpire.

Crystal structure of the open form of SagB did not display significant tendencies of adopting the proposed closed form with calculated average distance between the C α atoms of the catalytic E121 and A210 of YATD region of 14.9 ± 2.0 Å and RMSD of all atoms 1.9 ± 0.4 Å (Figure 37). Introduction of the (NAG-NAM)₃ substrate onto the open form did not induce significant closing motion of the enzyme as these distance values remained at comparable values 16.2 ± 2.4 Å and RMSD 1.6 ± 0.3 Å. Longer simulation times are required for such closing motions to occur, then offered by 500 ns MD simulations. The generated closed forms of these enzymes gradually adopted a form similar to the open enzyme indicated by high average RMSD value relative to the starting structure of 3.5 ± 0.9 Å. Only the generated closed conformation with (NAG-NAM)₃ substrate yielded a stable closed structure with distance values of 14.2 ± 1.1 and RMSD 2.1 ± 0.4 Å. Only slight rearrangements of the substrate in the SagB binding pocket were observed likely reflecting only approximate positioning of the substrate.

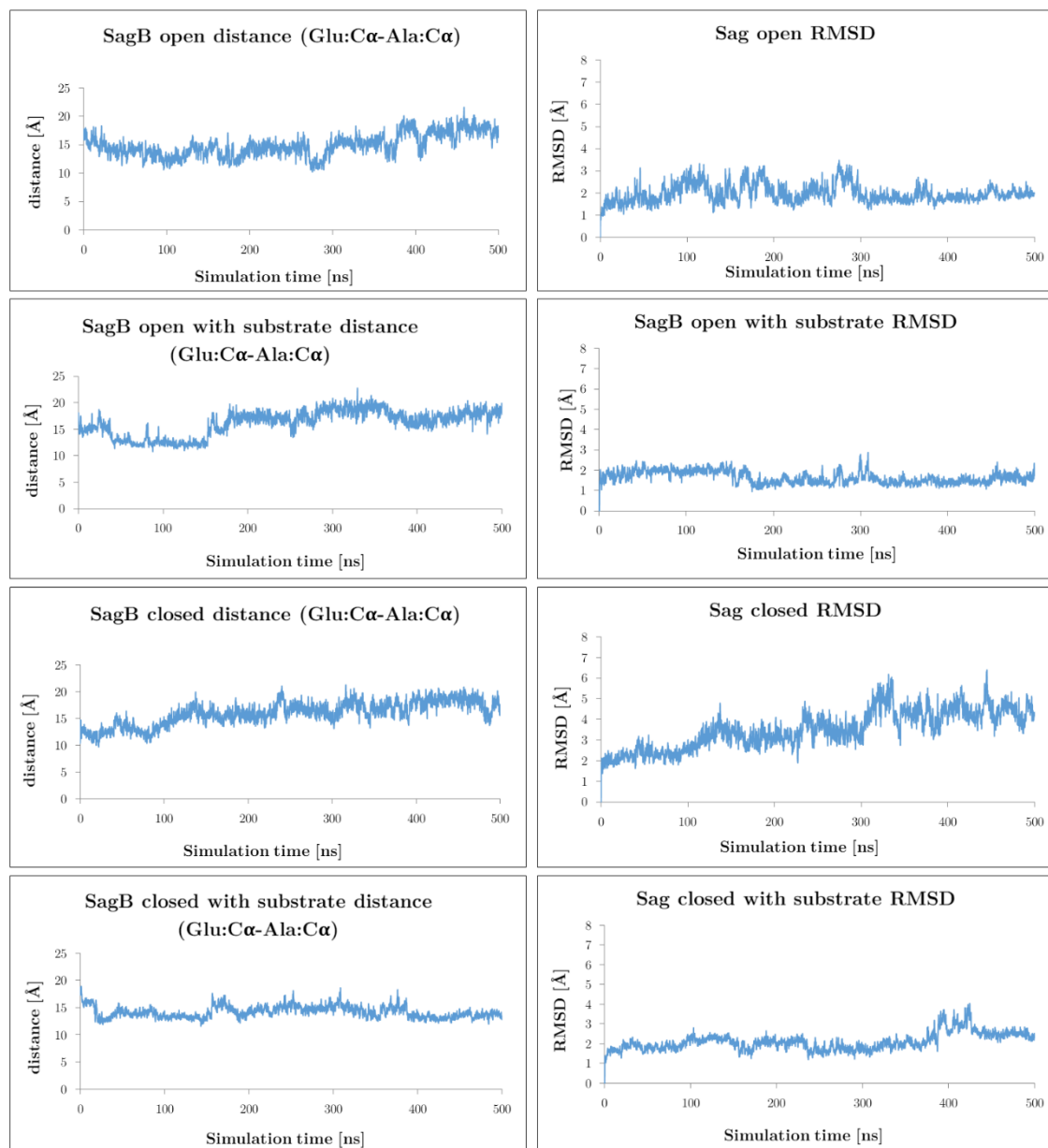


Figure 37: SagB flexibility simulations. The distances between the C α atoms of the catalytic E121 and A210 of YATD region during the MD simulations are depicted on the left and RMSD of all the atoms are shown relative to the starting structure on the right.

4.10.2 AtIA-gl

Simulations performed with AtIA-gl produced similar results as for SagB. Open form of AtIA-gl did not exhibit the propensity of adoption of the closed form with calculated average distance between the C α atoms of the catalytic E161 and A215 of YATD region of 20.3 ± 0.9 Å and RMSD of all atoms 2.5 ± 0.3 Å (Figure 38). The same as with SagB, addition of the (NAG-NAM)₃ substrate to the open form did not facilitate a shift towards the closed form of the enzyme as these distances remained at 20.5 ± 0.9 Å and RMSD at 2.7 ± 0.3 Å. The modelled closed structure without substrate displayed opening of the enzyme during simulation, which is reflected in RMSD relative to the starting structure of 4.7 ± 0.7 Å. Just as for SagB the introduction of the substrate upon the closed

conformation of AtIA-gl produces a stable enzyme with a distance value of 10.5 ± 0.4 Å and RMSD of 2.3 ± 0.2 Å.

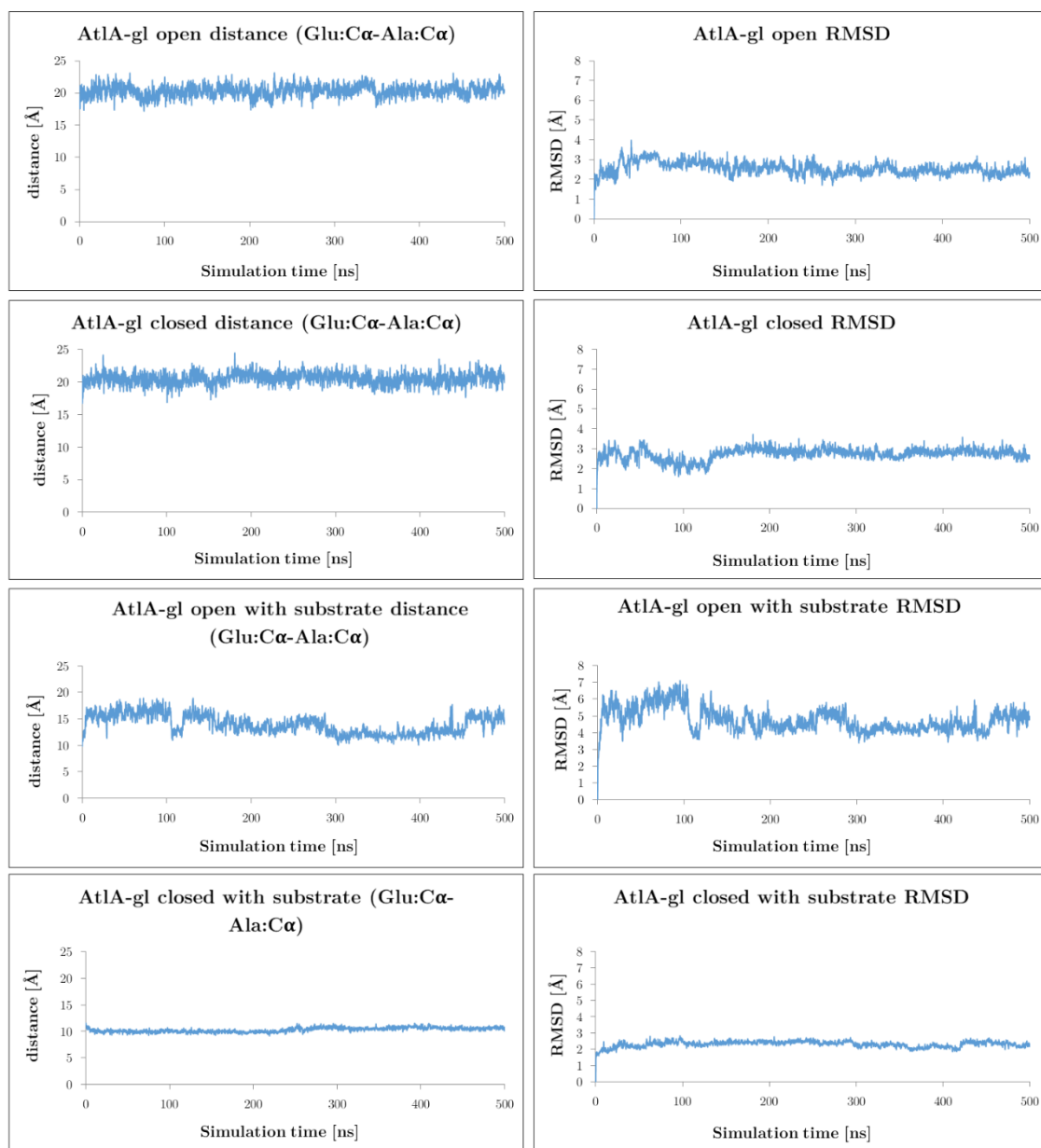


Figure 38: AtIA-gl flexibility simulations. Distances between the Cα atoms of the catalytic E116 and A215 of YATD region during the MD simulations are depicted on the left and RMSD of all the atoms relative to the starting structure are shown on the right.

4.11 Catalytic Mechanism

4.11.1 Exploring the possibility of the oxazolinium ion-like transition state in AtIE

With only one carboxylic group residue identified as catalytic in the GH73 family, several authors have proposed a substrate-assisted mechanism despite the fact that many enzymes

using this mechanism have two and not one conserved acidic amino acid residues. To probe for the possible existence of this mechanism, the theoretical/computational existence of the stable oxazolinium ion intermediate/transition state was explored.

Reactant structure was modified using appropriate RESD distance restraints in CHARMM until they formed the desired intermediate structure. These structures were minimized using a non-restrained QM/MM minimization. After the release of the restraints, the system immediately returned to its initial reactant structure. In addition, the plausibility of the presence of an oxazolinium ion-like transition state was tested. This time starting from the product structures and the distances between the reacting atoms were decreased until they formed the desired oxazolinium ion structure. After the restraints were released, again the system immediately returned to its initial starting structure.

Nevertheless, we decided to use the geometry of the oxazolinium ion structure obtained before the release of the restraints along with the previously optimized product structure, to generate the reaction pathways between them using the replica path method. The reaction geometries and energy profile (Figure 39) obtained clearly show that the boundary oxazolinium ion produced with restraints had a truly, substantially, higher relative energy value on the reaction coordinate, reflecting its previously determined instability (a relative energy of 67 kcal/mol). In addition, the boundary oxazolinium ion transforms into the product structure via an oxocarbenium ion-like transition state bearing a significantly lower relative energy of 10 kcal/mol on the reaction coordinate. Although this experiment does not correspond to realistic reaction pathway modeling, it serves to computationally rule out the existence of the oxazolinium-like transition state, which does not occur even if the oxazolinium ion is constructed as intermediate in a starting boundary structure for the replica path.

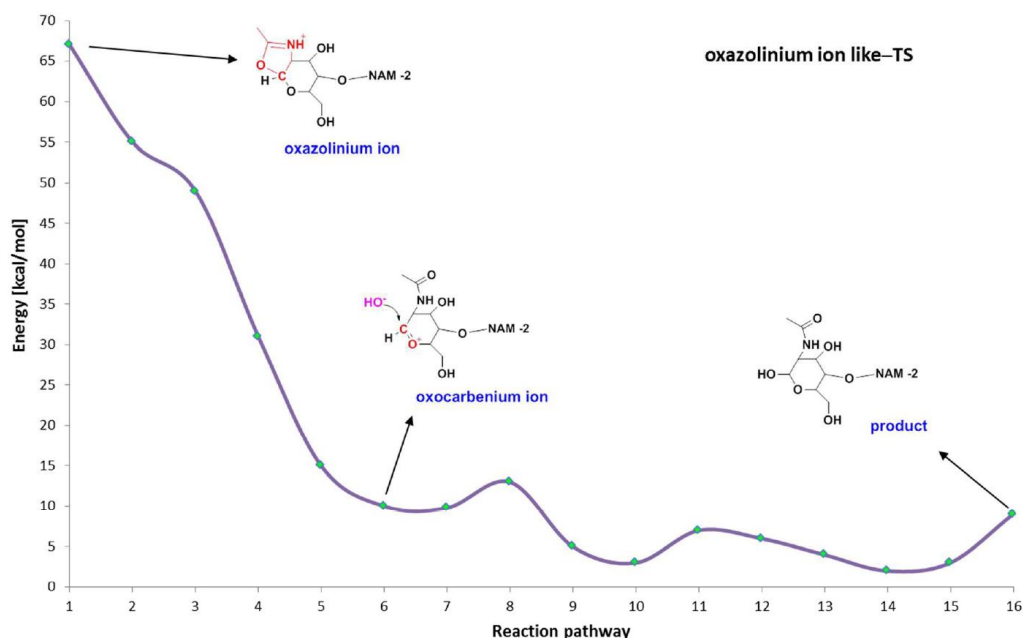


Figure 39: An optimized energy reaction profile for the oxazolinium ion-to-product pathway determined using the replica method. Replica frame 1 represents the RESD-generated oxazolinium ion, replica 6 represents the observed oxocarbenium ion-like transition state and frames 10-16 represent products (Borišek, 2018).

4.11.2 Calculation of the Possible Reaction Mechanisms

To obtain structural insight into the dynamics of substrate binding, AtIE, in complex with the substrate fragment (NAG-NAM)₃, was prepared and subjected to an MD equilibration run followed by a 100 ns MD simulation. The structure of the complex was determined to be stable and well equilibrated. In the MD trajectory, two different positions of the catalytic E138 (H ϵ 2 atom) with respect to the scissile glycosidic bond oxygen atom were observed: from 0 to 14 ns and from 60 to 100 ns of the MD simulation, the average distance between them was 2.2 Å, whereas from 14 to 60 ns of the MD simulation, their average distance was 4 Å. In the 14–60 ns interval structural frames a water molecule positioned between E138 and the glycosidic oxygen atom was noticed, prompting an investigation into the role of this water molecule in the catalysis (Figure 40).

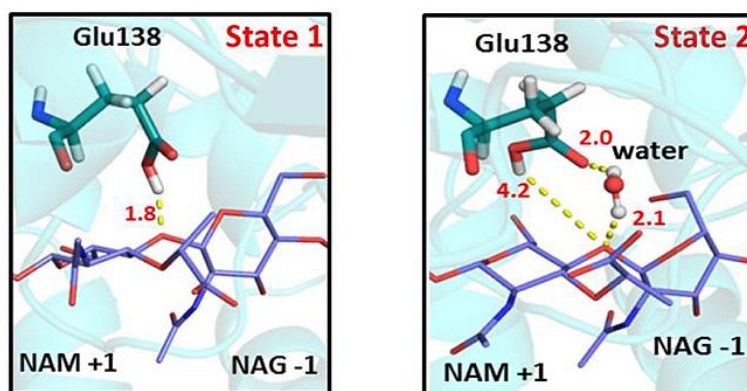


Figure 40: Selected frames from the MD simulations for the QM/MM RPATH simulations. State 1, at approximately 76 ns, corresponds to the direct involvement of E138, and state 2, at 52 ns, corresponds to water-mediated proton transfer (Borišek, 2018).

MD simulations suggested two possible mechanisms of the hydrolysis of the glycosidic bond with respect to proton transfer to the scissile glycoside bond oxygen atom. To evaluate which of the two reaction pathways more likely corresponds to the mechanism of the AtIE-catalyzed reaction, a pair of corresponding MD frames was selected and analyzed using the replica path method. Path A, where the E138 proton is transferred directly, and path B, where proton transfer is assisted by a water molecule (Figure 41). In the second step of both reactions, the glycosidic bond breaks, while in the third step, the hydroxyl anion attacks the anomeric carbon to form products in the final step. The oxocarbenium intermediate/transition states form and lead to the products. The difference between the paths is in how the glycosidic bond breaks, either by transfer of a proton from E138 anion (path A) or a water molecule (path B). Next, in path A, E138 anion attacks the water molecule to form the hydroxyl anion, which then attacks the anomeric carbon positioned on the NAG-1 and forms the products. Similarly, in path B, the hydroxyl anion attacks the anomeric carbon and forms the products. The replica path method was used to QM/MM optimize the geometries of both reaction pathways.

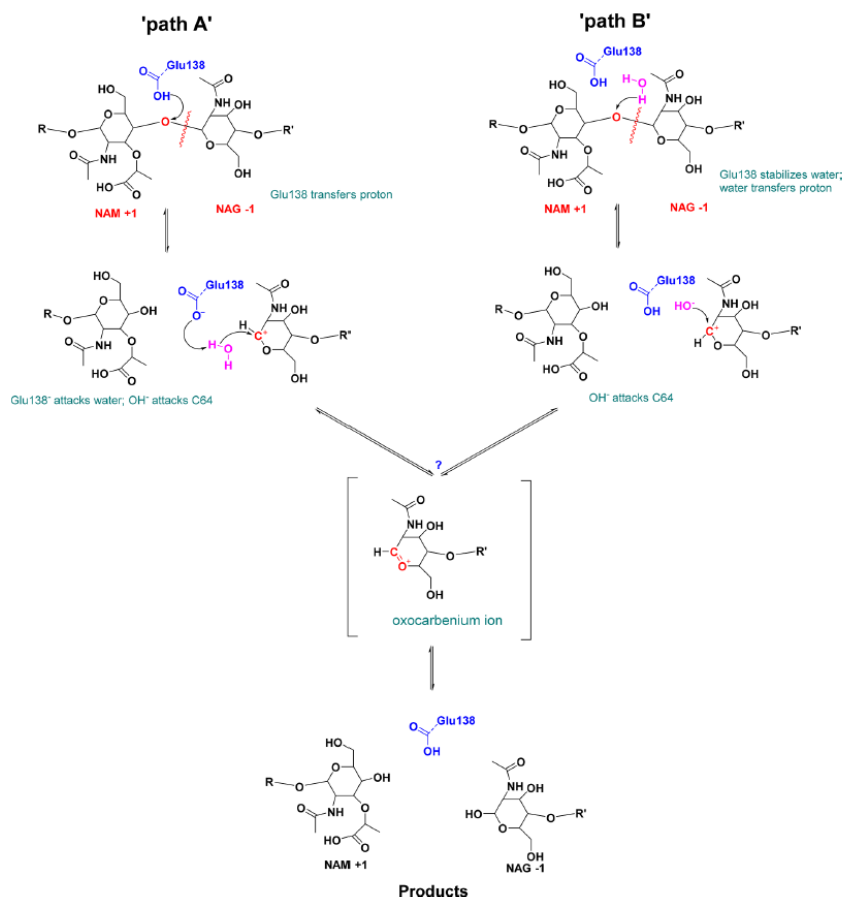


Figure 41: Reaction scheme of the proposed reaction mechanisms of the hydrolysis of the glycosidic bond in AtIE.

To evaluate which of the two reaction pathways more likely corresponds to the mechanism of the AtIE-catalyzed reaction, we selected a pair of corresponding MD frames and analyzed them both using the replica path method. The frames selected from the MD simulation were minimized to obtain the starting conformations for the QM/MM study. Substrate structures for paths A and B are very similar, as their RMSD is just 0.1 Å after the structural protein alignment. Both are stabilized by almost the same residues, the only difference being the substrate hydrogen bond with Glu138 that is unique to path A. All the AtIE binding site residues and their corresponding distances to the (NAG-NAM)₃ substrate were comparable to the distances observed in the MD simulations. These results suggest that the presence of a water molecule does not significantly influence the overall position of the (NAG-NAM)₃ substrate in the AtIE binding site. The interaction pattern did not change considerably in the product structures in comparison with the QM/MM reactant structures, indicating its stability in both the substrate and product configurations.

For both path A and path B, 14 replicas were generated by linearly interpolating the QM/MM optimized coordinates between the QM/MM optimized corresponding pairs of AtIE structures containing reactants and products to generate the initial structures on the reaction coordinate (Figure 42). The replica path method was used to QM/MM optimize the geometries of the reaction pathways. After the optimization, the pathways were visualized and the crucial distances between atoms of the interacting partners along the replica pathway were determined. The reaction in path A first involves a conformational

change of the hexopyranose NAG-1 unit from the chair ${}^4C_1$ to an intermediate half-chair ${}^4H_3$ /envelope E_3 conformation of the transition state. Further reaction steps reveal a boat conformation that ends in a skew-boat state close to the chair conformation. In path B, the reaction commenced from a ${}^4C_1$ NAG-1 hexopyranose state that proceeds to a midway half-chair ${}^2H_3$ /envelope E_3 NAG-1 conformation with a final development of the NAG-1 ring conformational changes similar to those described for path A.

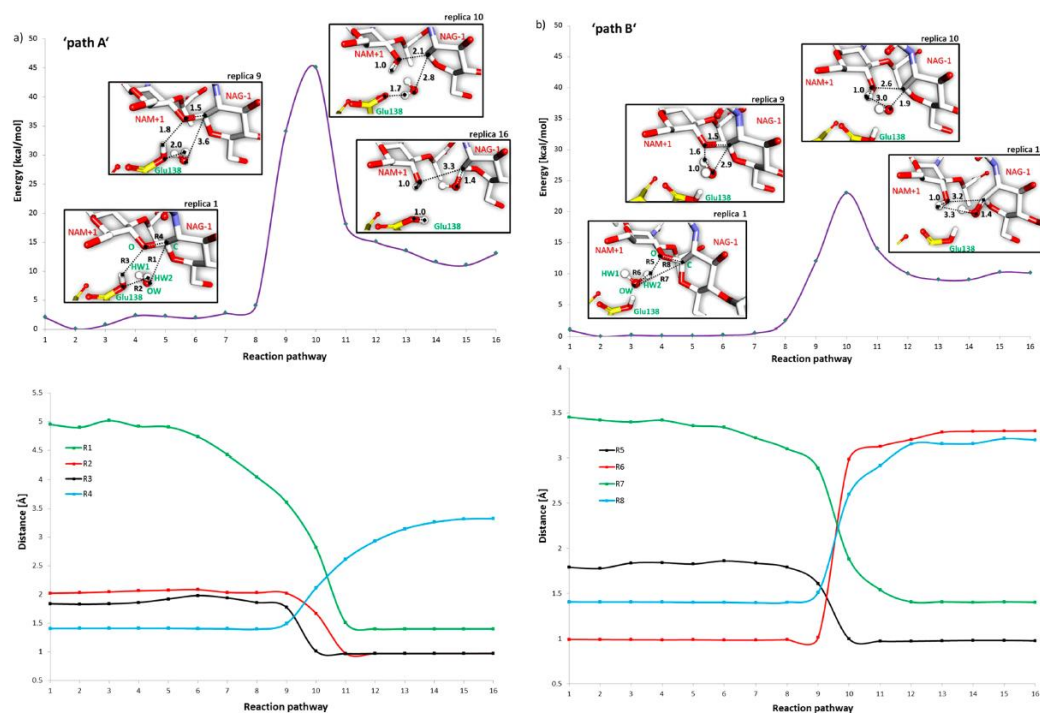


Figure 42: Optimized energy reaction profile of reaction path A (a) and path B (b) as determined by the replica path method with the distances that monitor the respective reaction mechanisms catalyzed by AtIE.

The calculated energy change (ΔE) suggests an endothermic reaction, as the ΔH value for both path A and path B was slightly above +10 kcal/mol. The lower energy of the transition state of 23 kcal/mol for path B in comparison with 45 kcal/mol for path A indicates that this pathway has a reasonably lower potential energy barrier and appears to be more favorable.

4.12 AtIE Inhibitors

Design starting point to identify AtIE inhibitors was the model of the substrate in the active site cleft of the enzyme. The central unit NAG-1–NAM+1, representing the cleavage site positioned next to the catalytic Glu138, was selected to derive structure-based pharmacophore for the initiation of the virtual screening campaign. The identified hydrogen bonding interaction pattern was with Gly162, Ser226, Asp227 and Lys233 (Figure 43a). A structure-based pharmacophore model mimicking the interactions of the selected central NAG–NAM unit with the AtIE binding site was generated (Figure 43b). Pharmacophores consisted of the following features: hydrogen bond acceptor (describing the interaction with the Gly162) and three hydrogen bond donors (observed with Ser226, Ser227 and Lys233). Exclusion volume spheres were also derived, mimicking the sterical boundary conditions of the AtIE-binding site. Interaction between scissile glycosidic substrate bond and Glu138

was not identified, likely due to the geometry of the complex not being optimal. In order to broaden the chemical space of the resulted virtual hits, we introduced an additional criterion that three out of the initial four pharmacophore features had to be satisfied for a compound to be considered as a hit. This pharmacophore model was used in a large-scale virtual screening campaign and resulted in approximately 10,000 virtual hits from different structural classes.

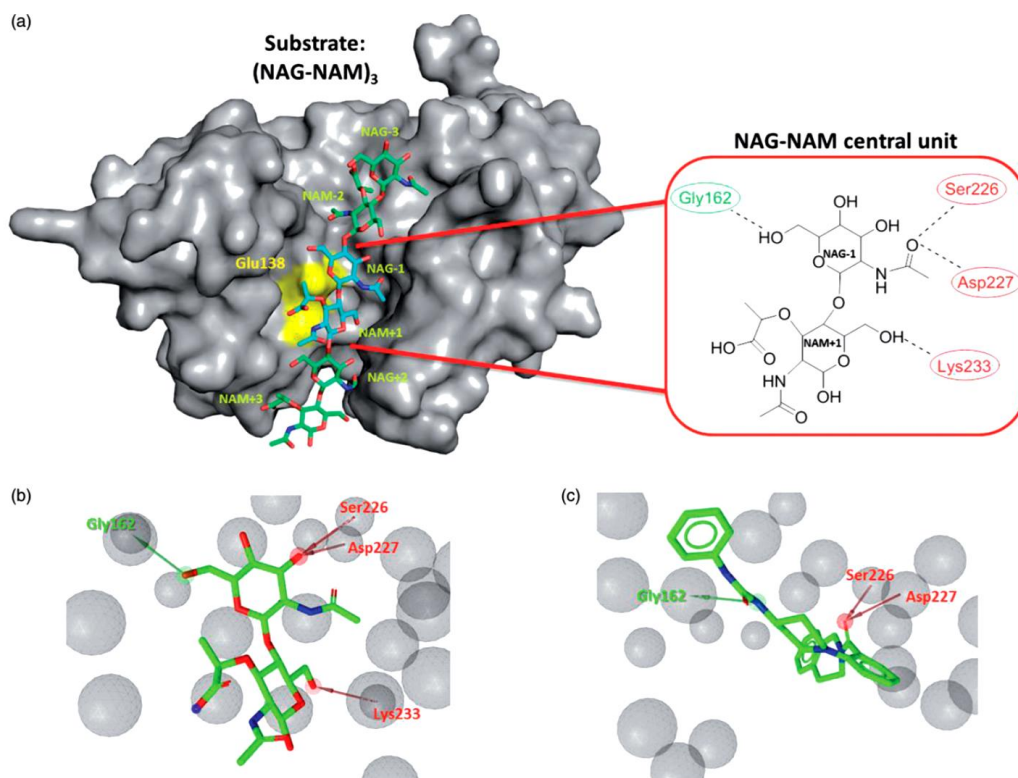


Figure 43: (a) Remodeled substrate in AtIE and schematic presentation of the selected NAG-NAM central unit for further virtual screening campaign. Yellow region represents catalytic amino acid Glu138. Dotted lines represent hydrogen bonds between the central substrate unit and the protein active site residues. Green and red residues represent hydrogen bond acceptors and donors, respectively. (b) LigandScout structure-based pharmacophore for the modeled central NAG-NAM ligand within the AtIE binding site and (c) micromolar hit 3 obtained from the first screening campaign aligned to the pharmacophore model. Green arrows represent hydrogen bond donors, red arrows hydrogen bond acceptors and gray spheres exclusion volume spheres. Pharmacophore model comprised one hydrogen bond donor (interaction with Gly162) and three hydrogen bond acceptors, describing possible interactions with Ser226, Asp227 and Lys233.

To explore the proposed orientations of these molecules, the obtained hits were docked into the AtIE-binding site. A successful validation of the docking model was made by redocking of the NAG-NAM molecule in the substrate binding groove. The docking procedure was performed and top 200 ranked docking solutions by scoring function were visually inspected for the fitness and orientation in the binding site.

41 compounds from different structural classes were selected by applying previously derived criteria of fitness, orientation and scoring function, and their binding affinities to AtIE were determined using SPR. In the first design phase, several compounds from the diverse structural classes exhibited binding affinity in millimolar range ($K_d > 1000$ mM).

Compound **3** from the structural class of (phenylureido)piperidinyl benzamides was found to exhibit promising binding affinity to At1E in the micromolar range (Figure 43c). To gain an insight into the molecular recognition of the compound **3** by At1E molecular docking was applied. A binding mode suggested hydrogen bonds with residues Gly162, Tyr224, Ser226 and Asp227, hydrophobic interactions with Phe63 and Tyr201, and electrostatic cation- π interaction with Lys233 (Figure 44a and b).

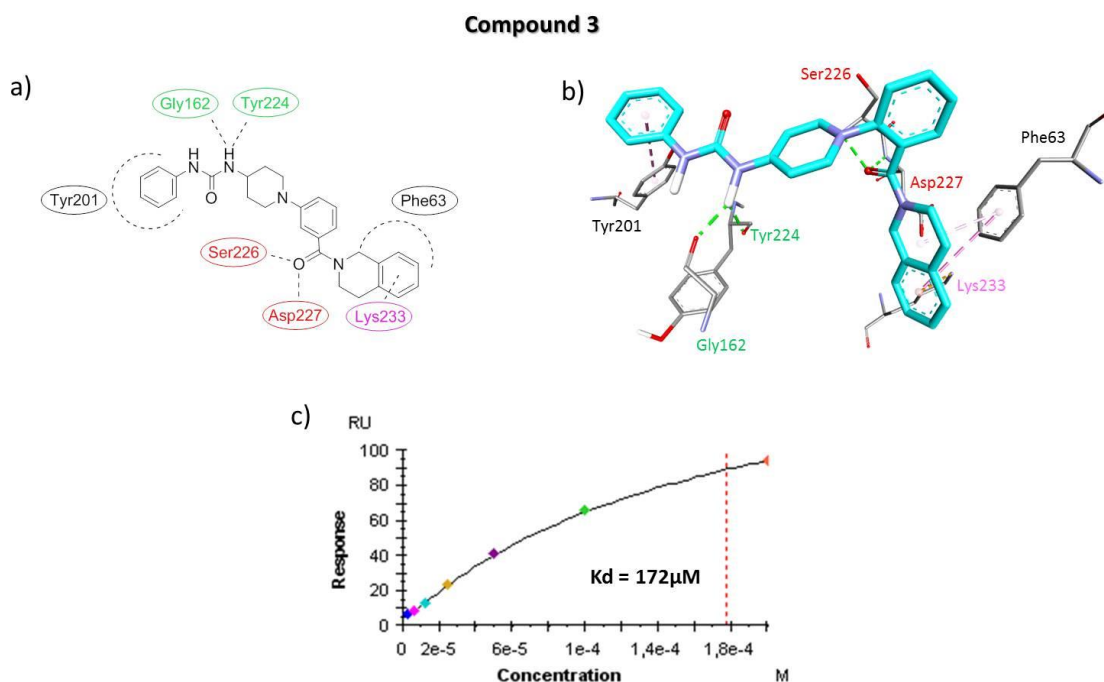
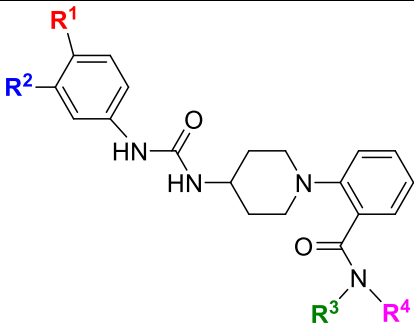
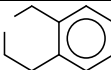
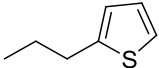
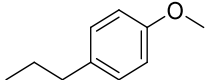
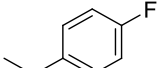
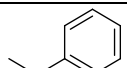
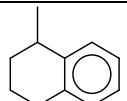
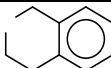
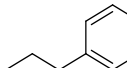
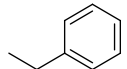
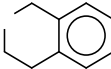
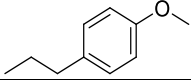
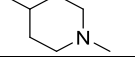
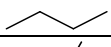
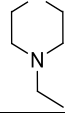


Figure 44: Proposed binding mode of the hit compound **3** in the At1E binding site with interacting residues depicted in a) 2D, b) 3D and c) obtained SPR sensorgram. Gly162, Tyr224, Ser226 and Asp227 form hydrogen bonds, Phe63 and Tyr201 hydrophobic interactions, and Lys233 electrostatic cation- π interaction. Green residue represents hydrogen bond acceptor, red residues hydrogen bond donors, black residues form hydrophobic interactions and magenta residue forms electrostatic cation- π interaction.

In order to provide initial SAR, a focused library of all commercially available, (phenylureido)piperidinyl benzamides was extracted from the eMolecules library. It consisted of 16 molecules, which were tested for their binding affinity to At1E by SPR. None of these molecules were present in our initial library of screened compounds. SPR experiments revealed fast and tight binding to the At1E at low micromolar concentrations. Additional nine compounds were found to have the binding affinity to At1E in the low micromolar range from the same structural class of (phenylureido) piperidinyl benzamides (Table 7).

Table 7: Compounds from the structural class of (phenylureido)piperidinyll benzamides tested for AtIE binding affinity in SPR experiments.

					
Comp.	R ¹	R ²	R ³	R ⁴	K _d (μM)
10	-H	-Cl			1.9
11	-Me	-H		-H	2.4
12	-H	-Cl		-H	2.5
13	-Me	-H		-H	6.1
14	-Me	-H		-H	11.3
15	-Me	-H		-H	11.7
16	-Me	-H			18.2
17	-H	-H		-H	21.1
18	-Me	-H		-Et	38.8
3	-H	-H			177
19	-Me	-H		-H	>1000
20	-Me	-H		-Me	>1000
21	-H	-H			>1000
22	-H	-H		-Et	>1000
23	-Me	-H		-H	>1000
24	-H	-H		-H	>1000
25	-Me	-H			>1000

For the STD NMR experiments, we selected the most soluble compounds among the active ones, i.e. two (phenylureido)piperidinyl benzamides, 13 and 14 (Figure 45). Due to overall low solubility of the compounds from this class their binding to AtLE was probed at experimental conditions for qualitative STD detection along with a long selective saturation time of AtLE. We obtained weak STD signals in the aromatic region of compounds, which indicates that the compounds 13 and 14 bind to AtLE. In addition to the poor solubility of the compounds (up to 100 mM), their low binding affinities ($K_d=10$ mM) near the limit of detection of the method (Mayer, 1999) may provide explanation for weak STD signals. Furthermore, due to the presence of methyl groups in ligands, we could not irradiate signals of protein in positive ppm range to improve the transfer of saturation along protein.

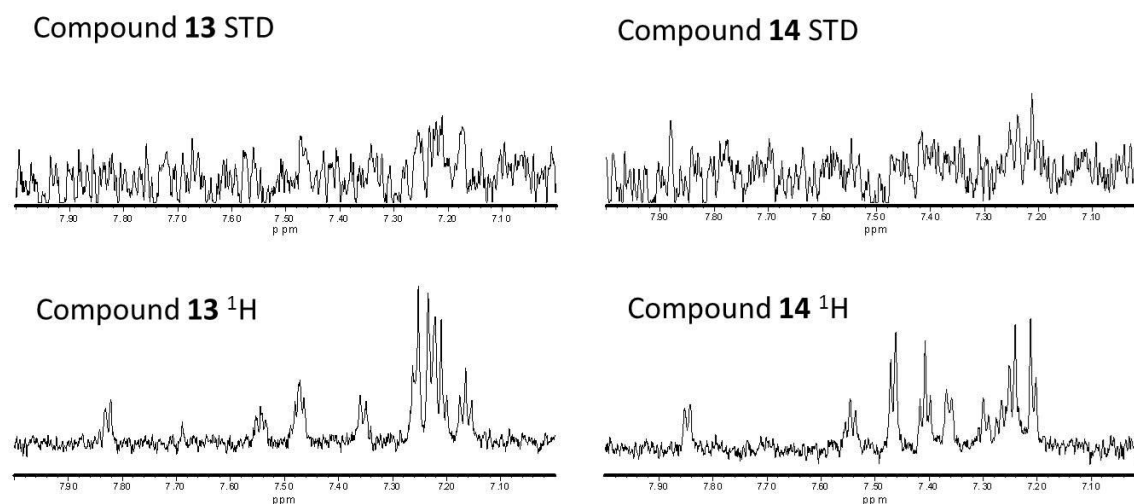


Figure 45: STD NMR and ^1H NMR spectra for the most soluble compounds 13 and 14. Around 7.2 ppm in aromatic region STD effect could be observed.

Based on the binding affinities and the molecular docking calculations of compounds 10, 16 and 3, we concluded that the chlorine substituent at the R2 position is the most important for binding affinity as the R1 methyl containing 16 has 10 times lower affinity ($K_d=18.2$ mM) and R1, R2 non-substituted 3 ($K_d=177$ mM) has 100 times lower affinity than 10 ($K_d=1.9$ mM), respectively. We could hypothesize that R1 and R2 substituents bind into the hydrophobic region of the AtLE-binding site comprising I163, Y177, F196 and Y201 residues (Figure 46). Compounds with cyclic lipophilic/aromatic fragments on the R3 and/or the R4 display a higher affinity ($K_d=\text{low mM range}$) than compounds 24, 23, 22 ($K_d>1000$ mM) with alkyl fragments on the R3 and/or R4 positions. The models suggested that the substituents on the R3 and/or R4 positions bind into the elongated hydrophobic region, comprised by F63, V137, and W230 residues and might also form additional electrostatic interactions with E138 and K233. Most of the active (phenylureido)piperidinyl benzamides compounds form hydrogen bonds with G162, G164, Y224, S226 and D227.

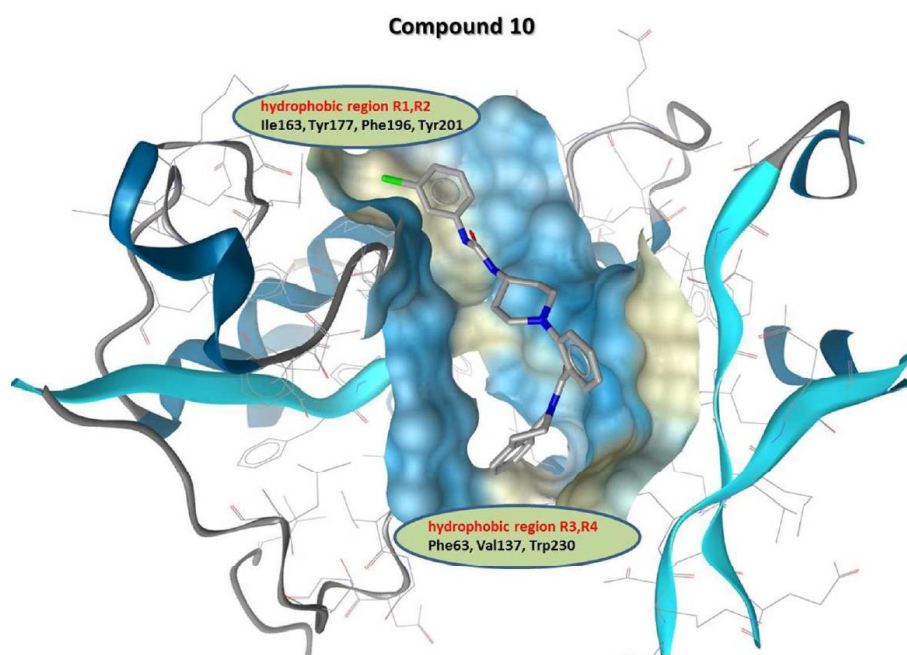


Figure 46: Hydrophilic and hydrophobic regions of the AtLE binding surface for modeled compound 10 are depicted in blue and yellow, respectively. Hydrophobic region R1, R2 and R3, R4 represent the area where substituents of (phenylureido)piperidinyl benzamides R1 and R2, and R3, R4 bind, respectively.

Chapter 5

Discussion

We have shown that SagB, AtlA-gl, AtlE and ScaH are the conserved peptidoglycan *N*-acetylglucosaminidases among *S. aureus* species. Their sequences are highly conserved. This indicated the importance for the survival of *S. aureus*. The crystal structures have revealed that the enzymes are very similar and yet the substrate binding site exists in open and closed form. This has led us to explore their substrate binding indicating conformational flexibility. The absence of the second catalytic residue prompted us to study the possible reaction transition states and catalytic mechanism of AtlE. Furthermore, we have identified first molecules bound by AtlE that represent possible inhibitors.

5.1 Structures Indicate High Conformational Flexibility of All of the Conserved Peptidoglycan *N*-acetylglucosaminidases of *S. aureus*

There are several theories explaining why conformational changes occur in the course of substrate binding and catalysis. A well-known theory is the induced fit theory, which postulates that binding of the substrate causes a conformational change of the enzyme, which brings catalytic groups into a precise orientation, required for the reaction to occur (Koshland, 1958). As the enzyme is more flexible, than the substrate, its distortion can be larger, thus in the enzyme-substrate complex the enzyme is physically distorted while the substrate is stressed, subjected to forces, yet not distorted, and the formation of the transition state is propelled by reduction of the strain in the enzyme (Fersht, 2017). This may apply particularly well to *S. aureus* peptidoglycan glycoside hydrolases, as its cell wall was found to be more rigid than those of *E. coli* or *B. subtilis* (Kern T. G., 2010) and in the glycan strands are less flexible than peptide moieties of peptidoglycan (Kern T. H., 2008). Thus, for the cleavage within the glycan strand region, the enzyme must be more flexible than an enzyme dealing with less rigid substrates. Another theory that has gained attention in recent years is conformational selection, which accounts for flexibility of the protein in solution, where it adopts a range of conformations and proposes a pre-existing subpopulation of conformations that can bind the substrate and this event subsequently shifts the equilibrium in favor of the bound conformation (Ma, 1999). Furthermore, it has been suggested that both theories are not mutually exclusive, as there has been an increasing number of cases, where conformational selection was shown to be followed by conformational adjustment (Csermely, 2010). With the fact that three different conformations have been captured by crystallization we can say at least these three conformations are present in the solution. The binding site is formed by the R-domain in such a way that it can bind the substrate, whilst L-domain displays steric hindrance for

substrate binding in SagB crystal structure. A combination of conformational selection and induced fit could be a plausible explanation of what happens upon binding of the substrate in these three enzymes.

Currently the knowledge of protein conformational flexibility is far from extensive with the non-redundant database of protein domain movements DynDom at 2035 entries with 203 of them ligand-induced (Qi, 2005). Even further, as very aptly pointed out by Gutteridge et al., there is an inherent bias towards identification of small motions, as, at this time, when the majority of the protein structures available have been solved by crystallography, the analysis of conformational flexibility is hindered by the restraints that this method imposes (Guttridge, 2005). The majority of the currently described conformational flexibility is clearly underestimated and this area undoubtedly still has a lot of development to come. Even the methods for describing such movements are still in development (Taylor, 2014). Amemiya et al. have classified large motions as the displacement of C_α for more than 1.0 Å (Amemiya, 2011). In the analysis of domain movements, the most commonly observed rotation angle was 15° with angles above 30° rarely observed (Taylor, 2014). In our case with rotation angles of 31° and 39° and translations of 3.5 Å and 4.6 Å observed for SagB and AtlA-gl, respectively, the conformational change is significant compared to a majority of the ones observed to date.

We present an exceptional example, with very similar structures caught in the crystal, as a snapshot, at different conformational states that allowed us to observe and explain a considerable flexibility among two domains with preserved fold. We propose conformational flexibility is associated with the nature of their substrate as well as with regulation of peptidoglycan degrading enzymes, since if they are not tightly regulated, cell lysis can occur.

SagB, AtlA-gl and AtlE crystal structures have revealed differences compared to other known GH73 structures. Their secondary structure organization is almost identical and yet their tertiary structures differ in the most surprising area, the substrate binding site. The two conserved regions in SagB and AtlA are positioned further apart than in the AtlE structure. Our previous work demonstrated how the shape of the substrate binding site of the AtlE determines the orientation of the cleaved glycoside bond relative to the active site residue (Mihelič, 2017). For the glycoside bond to occur substrate has to be positioned over the catalytic glutamate. The positioning is facilitated by the features of the right side of the active site cleft, whereas the catalytic amino acid is positioned on its left side. The substrate binding models of the SagB and AtlA-gl indicate that they must undergo transition from open to closed form to be able to interact with the NAG-NAM polymer on both sides of the active cleft simultaneously: with YA(S/T)D motif and two aromatic residues on the R-domain responsible for the binding and recognizing the substrate NAG(-3), NAM(-2), and NAG(-1) residues and the catalytic Glu on the L-domain. Point mutations have shown that features of both lobes are important for a productive reaction to occur. Furthermore, molecular dynamics simulations have confirmed that closed structures with bound substrate are stable and hence support the conformational change upon substrate binding. The proposed conformational change involves a change in the helix connecting both domains, $\alpha 15$ in SagB, $\alpha 13$ in AtlA-gl and $\alpha 12$ in AtlE. It is possible that upon substrate binding it assumes a more ordered structure, by extension of this helix and adopting a conformation similar to that of the AtlE crystal structure. Another possibility is further unfolding of this part of the helix. One of the amino acids in this region, F185 in SagB, Y191 in AtlA-gl and Y201 in AtlE, has been investigated for its role in the substrate binding. Its mutation in SagB and AtlE had a profound effect on the activity of the enzymes, whereas in AtlA-gl this was not the case. Thus it seems that has a role in substrate binding, but likely it is not the drive behind the conformational change, as suggested by AtlE and SagB structures, where these amino acids are slightly pulled out of

the helix. Conformational changes in helices are not an unprecedented observation, as kinking of the helices is a well-established regulating mechanism for gates of ion channels (Zhou, 2010). Simulations indicate that the SagB and AtIA-gl crystal structures are not in the widest possible state, whereas AtIE structure represents the closed state (Mihelič, 2017; Borišek, 2018).

5.2 A Water-Assisted Catalytic Mechanism in Glycoside Hydrolases Demonstrated on AtIE

For the retention and inversion acid/base mechanisms, the average distances between the proton donor and general base are approximately 5.5 and 10 Å, respectively (Davies, 1995). The importance of E224, E223, E223, E156, and E65 as putative secondary catalytic residues in substrate degradation was indicated by the structural and mutational studies of FlgJ from *Sphingomonas* sp. A1 (Maruyama, 2010), *Salmonella enterica* (Herlihey, 2014), *Salmonella typhimurium* (Rico-Lastres, 2015), Auto from *Listeria monocytogenes* (Bublitz, 2009), and TM0633 from *Thermotoga maritima* (Lipski, 2015). However, for the enzymes of the GH73 family studied here, the distances between the carboxylic groups are all larger than 10 Å. This suggests that the putative second catalytic residue is imbedded in the hairpin region above the active site. By a large conformational change, it could move close enough to the substrate during catalysis. This suggests that these proposals lack the support of structural evidence.

There is no experimental evidence that GH73 family enzymes require a cofactor or exogenous base. Our analysis of the active site, as well as sequence conservation and other point-mutation studies of this family, gave no indication of the possible alternative (nonacidic) amino acid involved in the catalysis. The retaining and inverting mechanisms provided utilizing two catalytic residues were excluded by the mutational studies, structural data, and conservation analysis reported in the literature. The remaining substrate-assisted and the potential water-assisted mechanism were simulated along with the reaction pathway to probe for potential stability of the oxocarbenium ion like transition state.

MD simulation of the NAG-NAM polymer bound to the active site of the crystal structure of AtIE indicated that the modeled structure remained stable during the simulation. There is neither unfolding of the lobe structure nor significant flexibility of the substrate model. Moreover, in the AtIE structure, the L-lobe is structured and appears to adopt a stable form that is not a hairpin, and the two potentially similar Glu residues (Glu179 and Glu181) are positioned away from the active site cleft. Furthermore, the equivalence in the position of the proposed secondary Glu residue involved in the catalysis is not clearly exposed by a conserved position in the sequence alignment throughout the GH73 family (Lipski, 2015).

The next argument for the absence of the secondary catalytic residue was based on the comparison of *N*-acetylglucosaminidase substrate binding to the c- and g-type lysozyme mechanisms (Vocadlo D. J., 2001; Helland, 2009). The secondary catalytic carboxylate group in lysozyme approaches the C1 atom from the same side of the carbohydrate ring as the primary catalytic residue. In contrast, in our recent work (Mihelič, 2017), we showed that the *N*-acetylglucosaminidases bind their substrates from the other face as muramidases (lysozymes). This suggests that the second acidic residue is to be found on the right side of the active site cleft and not on the left, as the residues responsible for substrate orientation display mirror selectivity in lysozymes relative to glucosaminidases. On the right-hand side of the active site cleft of AtIE, there is a highly conserved residue, D227, positioned at a similar distance of 8.5 Å to the catalytic glutamate as the Asp101 in the

Atlantic cod lysozyme structure (Helland, 2009). However, this residue did not attract much attention as the candidate for the second catalytic residue in spite of its high conservation. This Asp is not strictly conserved throughout the entire GH73 family (Lipski, 2015). The D227Q/A mutants showed that D227 is a sensitive residue affecting the rate of hydrolysis; however, the D227N mutant retained low activity, and in an MD simulation, a solvent molecule is not inserted between the D227 and the substrate, analogous to the g-type lysozyme (Helland, 2009). Therefore, we concluded that the role of D227 is to provide a positioning surface for the substrate using a hydrogen bond with the NAM(+1) residue hydroxyl group. In addition, the mutants of the equivalent position in *Streptococcus pneumoniae* LytB (Rico-Lastres, 2015) and *Staphylococcus warneri* (Yokoi, 2008) yielded an almost fully active protein. Altogether, this conundrum can be explained by the absence of the second catalytic carboxylic residue and the fact that the spread of the important positions of acidic residues reflects the differences in specificity of various enzymes. Hence, the water-assisted reaction mechanism we proposed is a plausible explanation for the catalytic mechanism of the GH73 family of enzymes.

Our reactant structures were based on two clusters/states derived from the 100 ns MD simulation using the crystal structure of AltE. This MD simulation allowed sufficient sampling of the conformational space of the protein–substrate complex to ensure stability of the complex (Ardevol, 2015). During MD simulation also the modeled structure and the substrate remained stable and there was neither any observable unfolding of the lobe structure nor significant flexibility of the substrate model. The analysis of the MD trajectory showed that the water molecule can be present between Glu138 and the AltE substrate. This suggested exposed water-assisted and substrate-assisted catalytic mechanisms as two possible alternatives.

In search of a possible catalytic mechanism of AltE, the replica path method was used to model paths A and B. Unlike popularly used restrained coordinate driving methods, the replica path method removes the need for the predetermination of the reaction coordinate. Thus, it circumvents the biased procedure of manually defining reaction steps and transition state-like structures and independently provides the sequence of reaction steps (Woodcock H. L., 2003). MD simulations identified two possible mechanisms of proton transfer to the scissile glycoside bond oxygen atom, via E138 or via a water molecule. Both mechanisms do not require the secondary catalytic carboxylic group, and both retain their configuration at the anomeric centers. Both calculated pathways proceeded via the oxocarbenium ion like transition state, and no stable intermediates were observed. The difference between the paths is the role of the catalytic E138 in the mechanism. In path A, E138 directly transfers the proton to the glycosidic oxygen, which leads to the cleavage of the glycoside bond followed by a nucleophilic attack of the polarized water molecule on the C1 atom of the NAG(-1) residue and, finally, the regeneration of the proton to Glu138 by the remaining water proton. In path B, the proton transfer to the glycosidic oxygen was facilitated by a water molecule, which leads to the cleavage of the glycoside bond followed by nucleophilic attack of the C1 atom of the NAG(-1) residue by the formed hydroxyl ion.

The calculated energy change (ΔE) suggests an endothermic reaction, similar as reported (Jitonnom, 2015). The lower energy of the transition state of path B, in comparison to path A, indicated that this pathway was more favorable. The energy difference observed between the path A and path B oxocarbenium ion like transition state structures could be attributed to a more energetically favorable proton transfer to the glycosidic oxygen via a water molecule and the nucleophilic attack of a nearly ideally positioned water hydroxyl ion on the anomeric carbon. Energetic favorability of path B could further be rationalized by the fact that at its transition state the hydroxyl anion has already been formed and stabilized the nascent positive charge at the anomeric carbon. It

should, however, be noted that minimal energy pathways like that produced by the replica path do not incorporate entropic effects.

We unsuccessfully attempted to obtain a stable oxocarbenium ion and oxazolinium ion intermediates, ruling out substrate-assisted mechanisms. The results suggested that formation of the stable intermediates is not possible in our system. Conformation behavior of hexopyranose was in accordance with observations reported from similar GH enzymes (Ardevol, 2015).

In this study we considered five different reaction mechanisms to obtain insight into the AtlE enzymatic reaction. In addition, unusual and rare GH mechanisms require cofactors, exogenous bases, or alternative nucleophiles (Jongkees, 2014), and are therefore not realistic explanations of the catalytic mechanism of the GH73 family, including AtlE.

We expect that the insights presented here will assist in setting up the experimental verification of the suggested mechanism and aid in drug discovery projects targeting the GH73 family of N-acetylglucosaminidases. The understanding of the reaction geometries, especially those close to the transition state structures for each GH, GH family, or type of substrate, are essential for rational development of TS mimetics, and thus, understanding the AtlE mechanism could be a step toward developing active agents that target this enzyme (Gloster, 2010).

5.3 Discovery of (Phenylureido)Piperidinyl Benzamides as Prospective Inhibitors of Bacterial Autolysin E from *S. aureus*

A combination of pharmacophore modeling, similarity search, and molecular docking has enabled us to identify a series of (phenylureido)piperidinyl benzamides as potential binders. (phenylureido)piperidinyl benzamides are the first reported non-substrate-like compounds that interact with this enzyme and enable a further study of the interaction of small molecules with bacterial AtlE as potential inhibitors of this target.

Binding model of the ligands in the AtlE binding site based on the molecular docking calculations was similar for all the (phenylureido)piperidinyl benzamides compounds (Figure 47). The AtlE substrate binding site is quite extensive as it can accommodate the substrate with six aminosugar NAG-NAM units. Thus alternative binding modes seemed also possible for the relatively small molecule as the 10. Nevertheless, the interaction binding pattern of compound 10 was well defined. It comprised of two hydrogen bond acceptors (G162 and Y224), two hydrogen bond donors (S226 and D227), and several hydrophobic interactions with F63, V137, I163, and F196 and K233 electrostatic cation- π interaction. Compared to the initial structure-based pharmacophore for the NAG-NAM, 10 retained the binding pattern of one hydrogen bond donor (G162) and two hydrogen bond acceptors (S226 and D227).

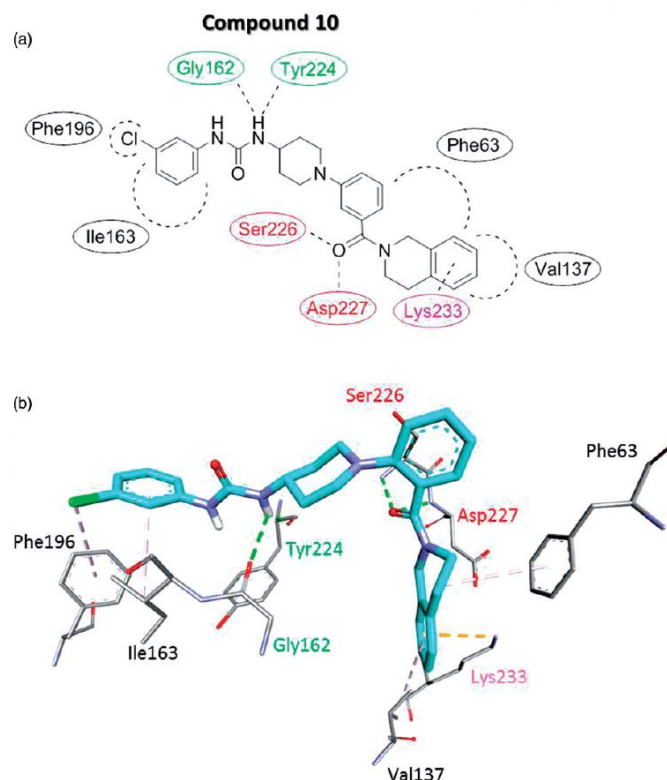


Figure 47: Proposed interaction binding pattern between the active (phenylureido)piperidinyl benzamides compound 10 and the AtIE binding site based on molecular docking calculations, depicted in (a) 2D and (b) 3D. It comprises two hydrogen bond acceptors (G162 and Y224), two hydrogen bond donors (S226 and D227), several hydrophobic interactions with F63, V137, I163, and F196 and K233 electrostatic cation-Pi interaction. Green residue represents hydrogen bond acceptor, red residues hydrogen bond donors, black residues form hydrophobic interactions and magenta residue forms electrostatic cation-Pi interaction.

Surface plasmon resonance (SPR) and saturation transfer difference (STD) NMR experiments revealed that discovered compounds bind to AtIE in a lower micromolar range. We unsuccessfully attempted to perform enzymatic assays which failed, due to poor solubility of compounds. In order for this assays to be successful, the concentration of the inhibitor needs to be above its solubility in established assay settings. However, (phenylureido)piperidinyl benzamides are the first known series of nonsubstrate-like small molecules with confirmed binding to AtIE and are as such a good starting point for subsequent optimization studies. (phenylureido)piperidinyl benzamides are the first reported non-substrate-like compounds that interact with this enzyme and enable a further study of the interaction of small molecules with bacterial AtIE as potential inhibitors of this target.

Chapter 6

Conclusions

We have gained a novel insight into the structure and mechanism of *S. aureus* GH73 enzymes. Crystal structures of SagB and AtlA-gl have given us an insight into the flexibility of these enzymes, as well as an indication of how they behave upon the binding of the substrate. Point mutations and simulations done with AtlE have furthered our understanding of the catalytic mechanism these enzymes use to cleave the glycosidic bond.

The findings of this work will enable further research into peptidoglycan hydrolases and contribute to elucidation of their role as well as overall peptidoglycan homeostasis and remodeling. Additionally, we have designed the first compounds that bind to AtlE, which are poorly soluble in aqueous solutions, but nonetheless represent a good starting point for further development of inhibitors. Furthermore, the proposed oxazolinium ion-like transition state gives another suggestion for what a successful inhibitor, namely transition state analogue, could look like. Good inhibitors would also enable further studies of the role of this family of enzymes, very likely in other species as well.

References

- Alanis, A. J. (2005). Resistance to Antibiotics: Are We in the Post-Antibiotic Era? *Archives of Medical Research*, 697-705.
- Al-Mebairik, N. F.-K.-S. (2016). A review of virulence factors, pathogenesis, and. *Reviews in Medical Microbiology*, 50-56.
- Amemiya, T. K. (2011). Classification and annotation of the relationship between protein structural change and ligand binding. *Journal of Molecular Biology*, 568-584.
- Ardevol, A. &. (2015). Crystal structures of g-type lysozyme from Atlantic cod shed new light on substrate binding and the catalytic mechanism. *Journal of American Chemical Society*, 7528-7547.
- Armstrong, G. L. (1999). Trends in Infectious Disease Mortality in the United States During the 20th Century. *JAMA*, 61-66.
- Aryee, A. &. (2017). Carriage, Clinical Microbiology and Transimission of *Staphylococcus aureus*. In F. R. Bagnoli, *Staphylococcus aureus: Microbiology, Pathology, Immunology, Therapy and Prophylaxis* (pp. 1-20). Trento: Springer International Publishing .
- Atilano, M. L. (2014). Bacterial autolysins trim cell surface peptidoglycan to prevent detection by the *Drosophila* innate immune system. *eLife*.
- Biswas, R. V. (2006). Activity of the major staphylococcal autolysin Atl. *FEMS Microbiology Letters*, 260-268.
- Boneca, I. G. (2000). Characterization of *Staphylococcus aureus* cell wall glycan strands, evidence for a new beta-N-acetylglucosaminidase activity. *Journal of Biological Chemistry*, 9910-8.
- Brooks, B. R. (2009). M. CHARMM: The Biomolecular Simulation Program. *Journal of Computational Chemistry*, 1545-1614.
- Bublitz, M. P. (2009). Structural basis for autoinhibition and activation of Auto, a virulence-associated peptidoglycan hydrolase of *Listeria monocytogenes*. *Molecular Microbiology*, 1509-1522.
- Cantarel B. L., C. P. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acid Research*, D233-D238.
- Chan, Y. G. (2016). SagB Glucosaminidase Is a Determinant of *Staphylococcus aureus* Glycan Chain Length, Antibiotic Susceptibility, and Protein Secretion. *Journal of Bacteriology*, 1123-1136.
- Chang, H. C. (2015). Origin and Proliferation of Multiple-Drug Resistance in Bacterial. *Microbiology and Molecular Biology Reviews*, 101-116.
- Coates, A. R. (2011). Novel classes of antibiotics or more of the same? *British Journal of Pharmacology*, 184-194.
- Csermely, P. P. (2010). Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends in Biochemical Sciences*, 539-546.

- Czerminski, R. &. (1990). Reaction-Path Study of Conformational Transitions in Flexible Systems - Applications to Peptides. *Journal of Physical Chemistry*, 5580-5601.
- David, M. Z. (2010). Community-Associated Methicillin-Resistant *Staphylococcus aureus*: Epidemiology and Clinical Consequences of an Emerging Epidemic. *Clinical Microbiology Reviews*, 616-687.
- Davies, G. &. (1995). Structures and Mechanisms of Glycosyl Hydrolases. *Structure*, 853-859.
- Dodds, D. R. (2017). Antibiotic resistance: A current epilogue. *Biochemical Pharmacology*, 139-146.
- Elber, R. &. (1987). Method for Determining Reaction Paths in Large Molecules - Application to Myoglobin. *Chemical Physics Letters*, 375-380.
- Eurenius, K. C. (1996). Enzyme mechanisms with hybrid quantum and molecular mechanical potentials. I. Theoretical considerations. *International Journal of Quantum Chemistry*, 1189-1200.
- Expression and Purification of soluble His6-Tagged TEV Protease. (2009). In J. E. Tropea, *Methods in Molecular Biology: High Throughput Protein Expression and Purification*, vol. 498. Totowa: Humana Press.
- Fersht, A. (2017). *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*. New Jersey: World Scientific.
- Finn, R. D.-G. (2017). InterPro in 2017 — beyond protein family and domain annotations. *Nucleic Acid Research*.
- Flemming, A. (1945, December 10). Penicillin. Stockholm, Sweden: The Nobel Foundation.
- Foster, S. J. (1995). Molecular Characterization and Functional Analysis of the Major Autolysin of *Staphylococcus aureus* 8325/4. *Journal of Bacteriology*, 5723-5725.
- Firdich, E. &. (2013). Peptidoglycan hydrolases, bacterial shape, and pathogenesis. *Current Opinion in Microbiology*, 767-778.
- Gally, D. &. (1993). Cell wall assembly in *Staphylococcus aureus*: proposed absence of secondary crosslinking reactions. *Journal of General Microbiology*, 1907-13.
- Gloster, T. M. (2010). Glycosidase inhibition: assessing mimicry of the transition state. *Organic & Biomolecular Chemistry*, 305-320.
- Grilo, I. R. (2014). The glucosaminidase domain of Atl - the major *Staphylococcus aureus* autolysin - has DNA-binding activity. *Microbiology Open*, 247-256.
- Guiton, P. S. (2009). Contribution of Autolysin and Sortase A during *Enterococcus faecalis* DNA-dependent Biofilm Development. *Infection and Immunity*, 3626-3638.
- Guttridge, A. &. (2005). Conformational changes observed in enzyme crystal structures upon substrate binding. *Journal of Molecular Biology*, 21-28.
- Hawkings, P. S. (2010). Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database. *Journal of chemical information and modeling*, 572-584.
- Helland, R. L. (2009). Crystal structures of g-type lysozyme from Atlantic cod shed new light on substrate binding and the catalytic mechanism. *Cellular and molecular life sciences*, 2585-2598.
- Herlihey, F. A. (2014). The Essential Protein for Bacterial Flagella Formation FlgJ Functions as a beta-N-Acetylglucosaminidase. The Essential Protein for Bacterial Flagella Formation FlgJ Functions as a beta-N-Acetylglucosaminidase. . *The journal of biological chemistry*, 31029-42.
- Hirschhausen, N. S. (2010). A novel staphylococcal internalization mechanism involves the major autolysin Atl and heat shock cognate protein Hsc70 as host cell receptor. *Cellular Microbiology*, 1746-1764.

- Houston, P. R. (2011). Essential Role for the Major Autolysin in the Fibronectin-Binding Protein-Mediated Staphylococcus aureus Biofilm Phenotype. *Infection and Immunity*, 1153-1165.
- Humphrey, W. D. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics and Modelling*, 33-38.
- Inagaki, N. I. (2009). Molecular properties of the glucosaminidase AcmA from Lactococcus lactis MG1363: Mutational and biochemical analyses. *Gene*, 61-71.
- Jitonnom, J. S. (2015). A DFT study of the unusual substrate-assisted mechanism of Serratia marcescens Chitinase B reveals the role of solvent and mutational effect on catalysis. *Journal of molecular graphics & modelling*, 53-59.
- Jongkees, S. &. (2014). Unusual enzymatic glycoside cleavage mechanisms. *Accounts of chemical research*, 226-235.
- Kabsch, W. (2010). XDS. *Acta Crystallographica Section D Biological Crystallography*, 125-132.
- Kern, T. G. (2010). Dynamics Characterization of Fully Hydrated Bacterial Cell Walls by Solid-State NMR: Evidence for Cooperative Binding of Metal Ions. *JACS*, 10911-10919.
- Kern, T. H. (2008). Toward the Characterization of Peptidoglycan Structure and Protein–Peptidoglycan Interactions by Solid-State NMR Spectroscopy. *JACS*, 5618-5619.
- Kim, S. J. (2015). Peptidoglycan architecture of Gram-positive bacteria by solid-state NMR. *Biochim. Biophys. Acta, Biomembranes*, 350-362.
- Kirchamir, J. M. (2008). Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection--what can we learn from earlier mistakes? *Journal of computer-aided molecular design*, 213-228.
- Koshland, D. E. (1958). Application of a Theory of Enzyme Specificity to Protein Synthesis. *PNAS*, 98-104.
- Kumar, R. I. (2007). CHARMM-GUI: A graphical user interface for the CHARMM users. *Abstracts of Papers of the American Chemical Society*, 273-273.
- Levy, S. B. (2004). Antibacterial resistance worldwide: causes, challenges and responses. *Nature Medicine*, S122-S129.
- Lipski, A. H.-L. (2015). Structural and biochemical characterization of the β -N-acetylglucosaminidase from Thermotoga maritima: toward rationalization of mechanistic knowledge in the GH73 family. *Glycobiology*, 319-30.
- Ma, B. K. (1999). Folding funnels and binding mechanisms. *Protein Engineering*, 713-720.
- MacKerell, A. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B*, 3586-3616.
- MacKerell, A. B.-M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B*, 3586-3616.
- MacKerell, A. F. (2004). Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry*, 1400-1415.
- Maruyama, Y. O. (2010). Mutational studies of the peptidoglycan hydrolase FlgJ of Sphingomonas sp. strain A1. *Journal of Basic Microbiology*, 311-317.
- Mayer, M. &. (1999). Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angewandte Chemie International Edition*, 1784-1788.
- McCoy, A. J.-K. (2007). Phaser crystallographic software. *Journal of Applied Crystallography*, 658-674.

- Meyrand, M. B.-C. (2007). Peptidoglycan N-acetylglucosamine deacetylation decreases autolysis in *Lactococcus lactis*. *Microbiology*, 3275-3285.
- Mihelič, M. V.-K.-V. (2017). The mechanism behind the selection of two different cleavage sites in NAG-NAM polymers. *IUCrJ*, 185-198.
- Niesen, F. H. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols*, 2212-2221.
- Olsson, M. S. (2011). PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK(a) Predictions. *Journal of Chemical Theory and Computation*, 525-537.
- Oshida, T. S. (1995). A *Staphylococcus aureus* autolysin that has an N-acetylmuramoyl-L-alanine amidase domain and an endo-B-N-acetylglucosaminidase domain: Cloning, sequence analysis, and characterization. *PNAS*, 285-289.
- Pham, T. (2012, April 5). *Drug Use Review*. Retrieved from U.S. Food and Drug Administration:
<https://www.fda.gov/downloads/Drugs/DrugSafety/InformationbyDrugClass/UCM319435.pdf>
- Philips, J. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 1781-1802.
- Piddock, L. J. (2012). The crisis of no new antibiotics--what is the way forward? *The Lancet Infectious Diseases*, 249-253.
- Qi, G. L. (2005). A comprehensive and non-redundant database of protein domain movements. *Bioinformatics*, 2832-2838.
- Rashid, M. H. (1993). *Bacillus subtilis* mutant deficient in the major autolytic amidase and glucosaminidase is impaired in motility. *FEMS Microbiology Letters*, 135-140.
- Rico-Lastres, P. D.-M.-B. (2015). Substrate recognition and catalysis by LytB, a pneumococcal peptidoglycan hydrolase involved in virulence. *Scientific reports*, 16198.
- Rodriguez-Rubio, L. M. (2013). Bacteriophage virion-associated peptidoglycan hydrolases: potential new enzybiotics. *Critical Reviews in Microbiology*, 427-434.
- Romaniuk, J. A. (2015). Bacterial cell wall composition and the influence of antibiotics by cell-wall and whole-cell NMR. *Philosophical Transactions of The Royal Society B*.
- Schlag, M. B. (2010). Role of staphylococcal wall teichoic acid in targeting the major autolysin Atl. *Molecular Microbiology*, 864-873.
- Schmidt, M. B. (1993). General Atomic and Molecular Electronic-Structure System. *Journal of Computational Chemistry*, 1347-1363.
- Sheldrick, G. M. (2010). Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallographica Section D Biological Crystallography*, 479-485.
- Stols, L. G. (2002). A New Vector for High-Throughput, Ligation-Independent Cloning Encoding a Tobacco Etch Virus Protease Cleavage Site. *Protein Expression and Purification*, 8-15.
- Sugai, M. K. (1995). Identification of Endo-B-N-Acetylglucosaminidase and N-Acetylmuramyl-L-Alanine Amidase as Cluster-Dispersing Enzymes in *Staphylococcus aureus*. *Journal of Bacteriology*, 1491-1496.
- Tamai, E. S. (2017). Structural and biochemical characterization of the *Clostridium perfringens* autolysin catalytic domain. *FEBS Letters*, 231-239.
- Taylor, D. C. (2014). Quantitative method for the assignment of hinge and shear mechanism in protein domain movements. *Bioinformatics*, 3189-3196.

- Thomas, V. C. (2009). A fratricidal mechanism is responsible for eDNA release and contributes to biofilm development of *Enterococcus faecalis*. *Molecular Microbiology*, 1022-1036.
- Turk, D. (2013). MAIN software for density averaging, model building, structure refinement and validation. *Acta Crystallographica Section D Biological Crystallography*, 658-674.
- Vanommeslaeghe, K. (2010). CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *Journal of Computational Chemistry*, 671-690.
- Varrone, J. J. (2011). Anti-Glucosaminidase Monoclonal Antibodies as a Passive immunization for methicilin-Resistant *Staphylococcus aureus* (MRSA) Ortopaedic Infections. *IBMS BoneKEy*, 187-194.
- Verdonk, M. C. (2003). Improved protein-ligand docking using GOLD. *Proteins*, 609-623.
- Visweswaran, G. R.-M. (2013). AcmD, a Homolog of the Major Autolysin AcmA of *Lactococcus lactis*, Binds to the Cell Wall and Contributes to Cell Separation and Autolysis. *PLOS One*, e72167.
- Vocadlo, D. D. (2001). Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature*, 835-838.
- Vocadlo, D. J. (2008). Mechanistic insights into glycosidase chemistry. *Current Opinion in Chemical Biology*, 539-555.
- Vollmer, W. J. (2008). Bacterial peptidoglycan (murein) hydrolases. *FEMS Microbiology Reviews*, 259-286.
- Vriend, G. (1990). What If - a Molecular Modeling and Drug Design Program. *Journal of molecular graphics*, 52-56.
- Vuong, T. V. (2010). Glycoside Hydrolases: Catalytic Base/Nucleophile Diversity. *Biotechnology and Bioengineering*, 195-205.
- Webb, B. &. (2017). Protein Structure Modeling with MODELLER. In M. K. Kaufmann, *Functional Genomics: Methods and Protocols* (pp. 39-54). Humana Press.
- Wheeler, R. T. (2015). Bacterial Cell Enlargement Requires Control of Cell Wall Stiffness Mediated by Peptidoglycan Hydrolases. *mBio*, e00660-15.
- Withers, S. G. (2001). Mechanisms of glycosyl transferases and hydrolases. *Carbohydrate Polymers*, 325-337.
- Wolber, G. &. (2005). Ligandscout: 3-d pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *Journal of chemical information and modelling*, 160-169.
- Woodcock, H. H. (2003). Exploring the quantum mechanical/molecular mechanical replica path method: a pathway optimization of the chorismate to prephenate Claisen rearrangement catalyzed by chorismate mutase. *Theoretical Chemistry Accounts*, 140-148.
- Woodcock, H. H. (2007). Exploring SCC-DFTB paths for mapping QM/MM reaction mechanisms. *Journal of Physical Chemistry*, 5720-5728.
- Woodcock, H. H. (2007). Interfacing Q-chem and CHARMM to perform QM/MM reaction path calculations. *Journal of Computational Chemistry*, 1486-1502.
- World Health Organization. (2014). *Antimicrobial resistance: global report on surveillance*. Geneva: World Health Organization.
- Yamada, S. S. (1996). An Autolysin Ring Associated with Cell Separation of *Staphylococcus aureus*. *Journal of Bacteriology*, 1565-1571.
- Yokoi, K. J. (2008). Molecular properties of the putative autolysin AtIWM encoded by *Staphylococcus warneri* M: Mutational and biochemical analyses of the amidase and glucosaminidase domains. *Gene*, 66-76.

Zhou, H. X. (2010). The gates of ion channels and enzymes. *Trends in Biochemical Sciences*, 179-185.

Bibliography

Publications Related to the Thesis

- Borišek, J., Pintar, S. (shared first authorship), Ogrizek, M., Turk, D. & Novič, M. (2018). A Water-Assisted Mechanism in Glycoside Hydrolases Demonstrated on the *Staphylococcus aureus* Autolysin E. *ACS Catalysis*, 4334-4345.
- Borišek, J., Pintar, S., Ogrizek, M., Golič Grdadolnik, S., Hodnik, V., Turk, D., Perdih, A. & Novič, M. (2018). Discovery of (phenylureido)piperidinyl benzamides as prospective inhibitors of bacterial autolysin E from *Staphylococcus aureus*. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 1239-1247.
- Pintar, S., Usenik, A., Borišek, J., Perdih, A. & Turk, D. (Manuscript in preparation). High conformational flexibility within GH73 domains of peptidoglycan N-acetylglucosaminidases of *Staphylococcus aureus*.

Other Publications

Conference Paper

- Tibaut, T., Tomašič, T., Hodnik, V., Anderluh, M., Pintar, S., Novič, M., Turk, D. (2018). Application of fragment based virtual screening towards inhibition of bacterial N-acetylglucosaminidase. *Book of abstracts, 18th International Conference on QSAR in Environmental and Health Sciences*. Ljubljana, Slovenia: National Institute of Chemistry.
- Taler-Verčič, A., Usenik, A., Lindič, N., Karničar, K., Pintar, S., Loboda, J., Arnolj, M., Pretnar, G., Tomaž, Š. & Turk, D. (2018). From genes to protein crystals : platform for high throughput protein production and structure determination. *Book of abstracts, 12th Meeting of the Slovenian Biochemical Society with International Participation*. Ljubljana: Slovenian Biochemical Society.

Invited lecture

- Mihelič, M., Vlahoviček-Kahlina, K., Renko, M., Mesnage, S., Doberšek, A., Pintar, S., Taler-Verčič, A., Jakas, A., Turk, D. (2017). The mechanism behind the selection of two different cleavage sites in NAG-NAM polymers. *Book of abstracts, 12th Meeting of the Slovenian Biochemical Society with International Participation*. Ljubljana, Slovenia: Slovenian Biochemical Society.

Conference Paper without print

- Pintar, Sara. (2015). Izražanje in čiščenje miristoiliranega proteina. *Mini simpozij Nef, izločen z eksosomi iz celic okuženih z virusom HIV, pomembno prispeva k razvoju nevroloških simptomov AIDS – a*. Ljubljana, 6. marec 2015.

Biography

Sara Pintar graduated from the University Study Programme Biochemistry of Faculty at the Chemistry and Chemical Technology, University of Ljubljana. In October 2013, she enrolled in the Nanosciences and nanotechnologies programme at the Jožef Stefan International Postgraduate School under the supervision of Prof. Dr. Dušan Turk. At the same time, she started working as a young researcher at the Jožef Stefan Institute at the Department of Biochemistry, Molecular and Structural Biology.

A Water-Assisted Catalytic Mechanism in Glycoside Hydrolases Demonstrated on the *Staphylococcus aureus* Autolysin E

Jure Borišek,^{*,†,‡} Sara Pintar,^{‡,§,||} Mitja Ogrizek,[†] Dušan Turk,^{‡,||} Andrej Perdih,^{*,†} and Marjana Novič^{*,†}

[†]National Institute of Chemistry, Hajdrihova 19, SI-1001 Ljubljana, Slovenia

[‡]Department of Biochemistry, Molecular and Structural Biology, Jozef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia

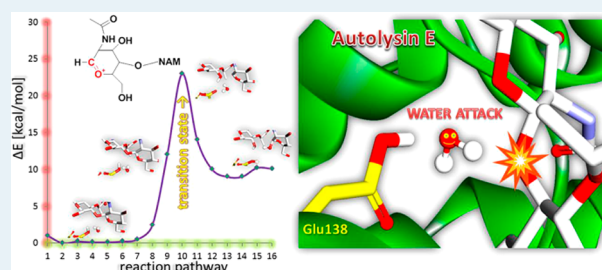
[§]Jozef Stefan International Postgraduate School, Jamova cesta 39, SI-1000 Ljubljana, Slovenia

^{||}Centre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins, Jamova cesta 39, SI-1000 Ljubljana, Slovenia

S Supporting Information

ABSTRACT: Autolysin E (AtIE), from *Staphylococcus aureus*, is a cell-wall-degrading enzyme that is a potential drug target. It is a member of the glycoside hydrolase (GH) class, enzymes that commonly have either two catalytic residues and hydrolyze their substrates by inverting or retaining mechanisms or one catalytic residue and undergo retaining, substrate-assisted catalysis. Here, we address the catalytic mechanism of AtIE. Site-directed mutagenesis studies identified Glu138 as the only catalytic residue. Quantum mechanics/molecular mechanics (QM/MM) simulations of the possible reaction pathways suggest that hydrolysis proceeds via a retaining, water-assisted mechanism and an oxocarbenium ion like transition state. These results, on the basis of data from a member of the hydrolase GH73 family, support the hypothesis of the presence of an alternative catalytic mechanism in glycoside hydrolases, which can be considered in the design of future AtIE inhibitors.

KEYWORDS: enzymatic catalysis, glycoside hydrolases, autolysin E, QM/MM, transition state, water-assisted catalysis



INTRODUCTION

Glycoside hydrolases (GH) are widespread enzymes with numerous roles from the degradation of biomass, antibacterial defense, and pathogenic mechanisms to physiological cellular functions. They catalyze the hydrolysis of the glycosidic bond in sugars, resulting in a sugar hemiacetal or hemiketal and the corresponding free aglycon products. Our particular interest lies in the GH73 family of enzymes in the pathogenic bacterium *Staphylococcus aureus*. With the increasing emergence of antibiotic-resistant strains such as MRSA and VRSA, it is evident that new classes of antibiotics need to be discovered. Therefore, additional pathways essential for the survival of bacteria need to be explored. Peptidoglycan cell wall remodeling and hydrolysis are as important for bacteria as their synthesis. We believe that understanding the chemical mechanisms underlying bacterial peptidoglycan hydrolysis has the potential to lay the foundation for the design of principles of novel chemical compounds for drug discovery.^{1,2}

Enzymatic hydrolysis of the glycosidic bond usually proceeds by the general acid/base catalysis proposed by Koshland back in the 1950s.³ Hydrolysis of the glycosidic bond results in either the retention or inversion of the configuration of the anomeric atom (Scheme S1).^{4,5} The mechanisms require a proton donor and a nucleophile/base. In both, the origin of the proton donor is an identical acidic residue, whereas the nucleophiles differ. In the retaining mechanism, a covalent intermediate with the second carboxylate is formed, as shown by the crystal structure

of the c-type lysozyme (Scheme S1a).⁶ A general acid/base catalyst works first as an acid and facilitates departure of the leaving group by donating a proton to the glycosidic oxygen atom, while the nucleophile forms an enzyme-sequestered covalent intermediate. In the next step, it is deprotonated, and it functions as a general base to activate a water molecule that carries out a nucleophilic attack on the glycosyl-enzyme intermediate, leading to retention of the stereochemistry at the anomeric center. The acid/base mechanism by inversion of the configuration at the anomeric center requires two carboxylates: one that acts as an acid, and one that acts as a base. The catalytic acid residue donates a proton to the anomeric carbon, whereas the catalytic base removes a proton from a water molecule, which increases its nucleophilicity and thereby facilitates its attack on the anomeric center (Scheme S1b). It was proposed that, for the g-type lysozymes of the GH23 family of enzymes, the second carboxylate would stabilize the transition state via a water molecule.⁷ A variation of this mechanism is substrate-assisted catalysis or neighboring group mechanism. The C2 acetamido group of the substrate acts as an intramolecular nucleophile and forms an oxazolinium ion intermediate (Scheme S1c). Rigorous chemical and structural analyses of the catalytic mechanism of the GH73 family of enzymes at the level of the c-type lysozyme studies⁶

Received: March 16, 2018

Published: April 10, 2018

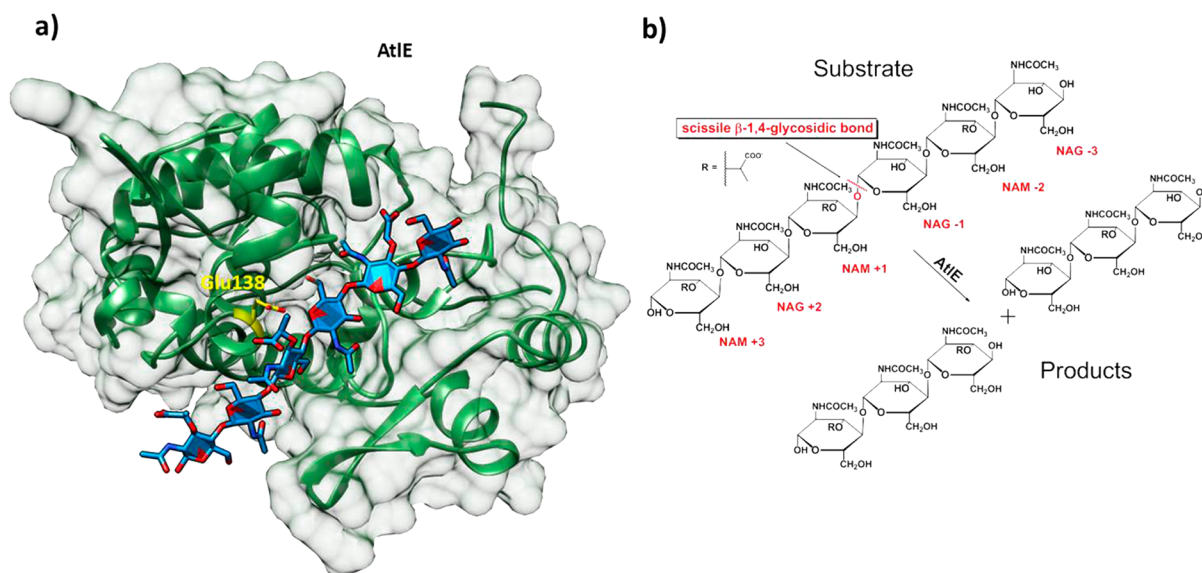


Figure 1. (a) Peptidoglycan substrate consisting of the three (NAG-NAM) units within the AtIE binding site with the proposed catalytic Glu138 depicted in yellow (PDB: 4PI7). (b) Substrate and products of the AtIE-catalyzed hydrolysis reaction.

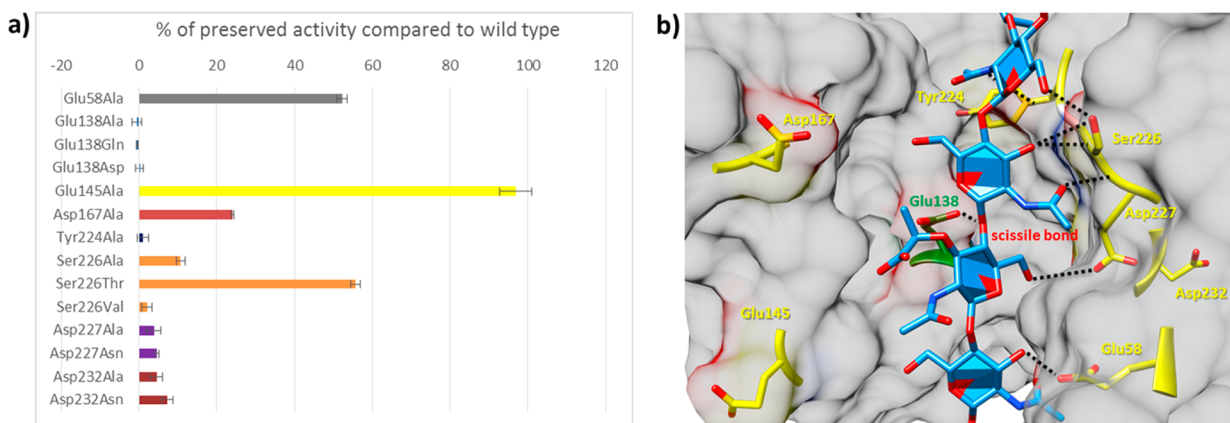


Figure 2. (a) Percentage of catalytic activity in AtIE mutants in comparison with the wild-type enzyme. (b) Mutated residues Glu58, Glu145, Asp167, Tyr224, Ser226, Asp227, and Asp232 (depicted in yellow) and E138 (depicted in green) mapped on AtIE.

are lacking because of experimental limitations (mostly the availability of appropriate substrates), leaving room for hypotheses. It has been suggested that more than one mechanism remains possible.⁸

Members of the GH73 family are structurally similar to those of the glycoside hydrolases from the GH19, GH22, GH23, GH24, and GH46 families of enzymes.⁹ The core regions of these enzymes are comprised of five to six helices, whereas the left and right lobe regions above the core region vary in size and structure.¹⁰ The highly conserved residues and regions in the families of enzymes usually indicate catalytic sites. Mutational studies on GH73 family members show that the only conserved catalytic residue is Glu138 (using AtIE numbering), whereas the second catalytic carboxylic group remains elusive.^{8,11–17} On the basis of predictions of a large conformational change that are, however, unsupported by structural analysis, it was suggested that the variability of the left lobe region, with either the presence or absence of a second acidic amino acid, is a determinant of the catalytic mechanisms within the GH73 family.⁸ The recently determined crystal structure of the AtIE, an N-acetylglucosaminidase from *Staphylococcus aureus* Mu50 and its complexes with fragments of peptidoglycan (PDB

identifiers: 4PI7, 4PI8, and 4PI9) suggest the substrate binding geometry (Figure 1).¹⁰

To address the catalytic mechanisms of the N-acetylglucosaminidases of the GH73 family, we conducted site-directed mutagenesis and applied the minimalist modeling approach using the available structural data.¹⁰ In the search of a possible catalytic mechanism for AtIE, we used molecular dynamics (MD) simulations and a replica path quantum mechanics/molecular mechanics (QM/MM) method^{18,19} that defines reaction pathways by simultaneous optimization of a series of geometries of the reacting system along the reaction pathway,^{20–22} unlike the more frequently used restrained coordinate driving (RCD) methods.

RESULTS

Site-Directed Mutagenesis. To assess the roles of residues that are potentially involved in substrate binding and/or catalysis in AtIE, we mutated all acidic amino acid residues within a 15 Å radius from the catalytic residue Glu138, as well as from Tyr and Ser residues from the highly conserved YASD region located on the right-hand side of the active site cleft near Glu138 (Figure 2).

To assess the effects of mutations on the folding and stability of the enzyme, near and far circular dichroism (CD) spectra were measured (data not shown). They confirmed the native folds of all mutants. In addition, the crystal structure of the Glu138Ala mutant (4PI8) closely matched the structure of the native enzyme.¹⁰ To corroborate these data, we additionally tested the stability of the mutants by differential scanning fluorimetry (DSF). The Glu138Ala and Glu232Ala mutants were very unstable, as they did not produce melting curves with a linear beginning. The Glu138Asp, Asp167Ala, and Asp227Ala mutants were significantly less stable than the wild-type enzyme (WT) (Figure S1), while Glu145Ala, Tyr224A, and Asp232Asn were only slightly less stable than the WT.

As expected, mutations of Glu138 completely abolished the enzymatic activity (Figure 2). The Tyr224Ala mutation in the highly conserved Tyr-Ala-Ser-Asp region within the loop in the right lobe also yielded a completely inactive enzyme. The Phe and Trp mutations of Tyr at the equivalent position in AcmA and in Atl(WM) preserved enzyme function, indicating that the aromatic ring, and not the hydroxyl group, is crucial for enzymatic activity.^{12,16} Hence, Tyr224 seems to facilitate substrate accommodation and is not involved in catalysis. The Ser226 and Asp227 positions are not absolutely conserved. Their mutants indicated reduced, but not abolished, enzyme activity. This suggests that the region is involved in substrate binding and not in catalysis, assuming that these enzymes use the same catalytic mechanisms. Asp232 is too far from the active site cleft, and a putative conformational change that would move it closer to the catalytic region seems very unlikely; hence, it appears to be important for stabilizing the R-lobe structure, including the Tyr-Ala-Ser-Asp region, and is not directly involved in substrate binding.

Mutants of Glu58 and Asp167 showed partial activity. The lower stability of the Asp167Ala mutant revealed by CD and DSF is likely responsible for the lower enzymatic activity, whereas the Glu145Ala mutant exhibited nearly wild-type activity.

MD Simulations of the Enzyme–Substrate Complex.

To obtain structural insight into the dynamics of substrate binding, AtlE, in complex with the substrate fragment (NAG-NAM)₃, was prepared and subjected to an MD equilibration run followed by a 100 ns MD simulation. During the MD run, all of the protein atom RMSD values were 1.8 ± 0.4 Å, while the substrate fragment RMSD value was 1.9 ± 0.3 Å (Figure S2a,b). These values indicate that the structure of the complex was stable and well equilibrated.

To further substantiate the results of site-directed mutagenesis, the MD trajectory of the complex was inspected. First, the candidates for the second catalytic carboxylate, in the vicinity of 15 Å from the substrate scissile glycosidic oxygen atom, were analyzed. From the observed distances between the scissile glycosidic oxygen atom and the carboxylic side chain of Glu58 (9.3 ± 0.8 Å), Asp167 (14.0 ± 3.2 Å), Asp227 (6.4 ± 0.4 Å), and Asp232 (15.0 ± 1.3 Å), we concluded that there was no additional Glu or Asp residue, other than Glu138, that was directly involved in catalysis (Figure S2c–f). Throughout the MD simulation, the β -hairpin region placed above the active site was stable, and the carboxyl side chains of Asp227 and Asp232 pointed away from the scissile bond. Moreover, the Asp227 side chain formed hydrogen bonds with the substrate's NAM(+1) N-acetyl group and Lys233, Arg229, and Trp230, which restrained its position and excluded its active role in catalysis.

A secondary analysis of the MD trajectory suggested the residues listed in Figure 3 as the key residues involved in

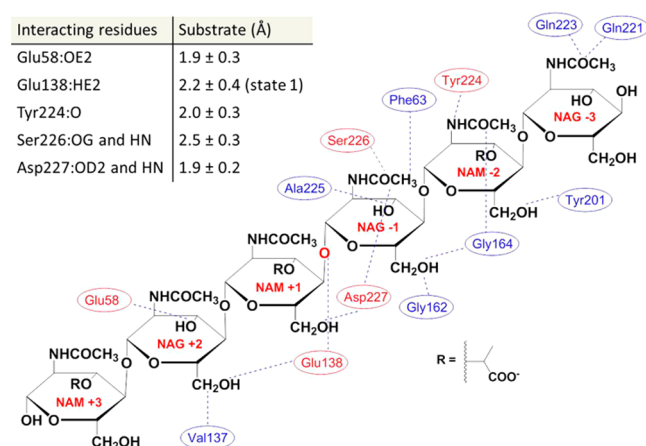


Figure 3. Schematic presentation of the interactions between the substrate (NAG-NAM)₃ and the AtlE. The Glu58, Glu138, Tyr224, Ser226, and Asp227 residues shown in red formed hydrogen bonds with the substrate fragment for most of the simulation time, whereas the Val137, Gly162, Gly164, Tyr201, Gln221, Gln223, and Ala225 residues shown in blue formed hydrogen bonds with the (NAG-NAM)₃ substrate for approximately half of the MD simulation time. The table to the left of the diagram shows the calculated average interacting distances between the selected AtlE binding site residues and the substrate.

substrate binding. The Glu58, Glu138, Tyr224, Ser226, and Asp227 residues shown in red formed at least one hydrogen bond with the main chain and/or side chain of the (NAG-NAM)₃ substrate fragment in the majority of the MD time frames, whereas the Val137, Gly162, Gly164, Tyr201, Gln221, Gln223, and Ala225 residues shown in blue formed hydrogen bonds with the substrate in half of the MD time frames.

Interestingly, in the MD trajectory, two different positions of the catalytic Glu138 (HE2 atom) with respect to the scissile glycosidic bond oxygen atom were observed: from 0 to 14 ns and from 60 to 100 ns of the MD simulation, the average distance between them was 2.2 Å, whereas from 14 to 60 ns of the MD simulation, their average distance was 4 Å (Figure 4). In the 14–60 ns interval, we noticed the presence of structural frames, with a water molecule positioned between Glu138 and the glycosidic oxygen atom, prompting an investigation into the role of this water molecule in the catalysis (Figure 4a, right).

Calculation of the Proposed Reaction Mechanisms.

The combined analysis of site-directed mutagenesis data and MD simulations suggested two possible mechanisms of the hydrolysis of the glycosidic bond with respect to proton transfer to the scissile glycoside bond oxygen atom: path A, where the Glu138 proton is transferred directly, and path B, where proton transfer is assisted by a water molecule.

To evaluate which of the two reaction pathways more likely corresponds to the mechanism of the AtlE-catalyzed reaction, we selected a pair of corresponding MD frames (Figure 4) and analyzed them both using the replica path method.¹⁸ A description of the QM region is available in Figure S3, and the replica path method is described in the Experimental Section. The details of the modeled pathways are shown in Scheme 1.

The first step in both possible mechanisms is the transfer of a proton to the glycosidic oxygen atom. In the second step, the glycosidic bond breaks, and either a nucleophilic Glu138 anion

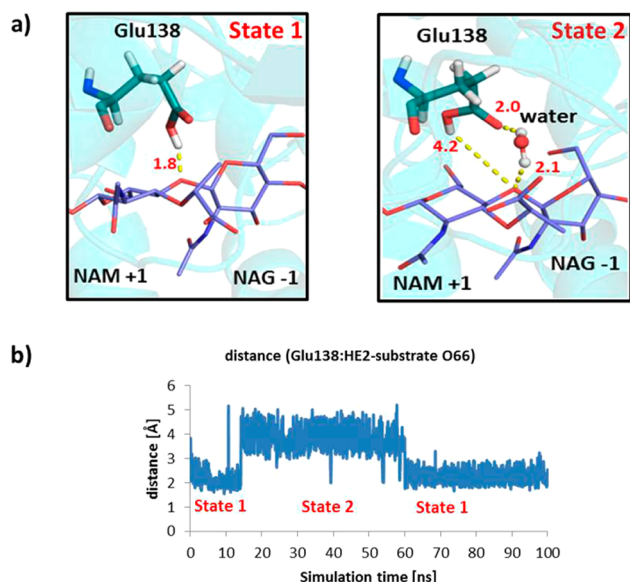


Figure 4. (a) Selected frames from the MD simulations for the QM/MM RPATH simulations. State 1, at approximately 76 ns, corresponds to the direct involvement of Glu138, and state 2, at 52 ns, corresponds to water-mediated proton transfer. (b) Distance between Glu138:HE2 and the scissile glycoside bond oxygen atom in the 100 ns MD simulation of the AtLE–substrate complex corresponding to states 1 and 2.

or a nucleophilic hydroxyl anion is formed via path A or path B, respectively. Next, in path A, Glu138 attacks the water molecule to form the hydroxyl anion, which then attacks the anomeric carbon positioned on the NAG-1 and forms the products. Similarly, in path B, the hydroxyl anion attacks the anomeric carbon and forms the products (Scheme 1).

Optimization of the Reactant and Product Structures. The frames selected from the MD simulation are minimized to obtain the starting conformations for the QM/MM study. In addition to the identified key residues (Figure 3), two additional residues, His222 and Lys233, were observed to form stable hydrogen bonds with the AtLE substrate after QM/MM minimization (Figure S4). Substrate structures for paths A and B (Scheme 1) are very similar, as their RMSD is just 0.1 Å after the structural protein alignment. Both are stabilized by almost the same residues as described before, the only difference being the substrate hydrogen bond with Glu138 that is unique to path A. All the aforementioned AtLE binding site residues and their corresponding distances to the (NAG-NAM)₃ substrate are comparable to the distances observed in the MD simulations (see Figure 3) and the experimentally observed interactions of the ligand NAG-NAM in the AtLE complex crystal structure.¹⁰ These results suggest that the presence of a water molecule does not significantly influence the overall position of the (NAG-NAM)₃ substrate in the AtLE binding site.

The product structures were generated from the substrate structures from paths A and B by breaking the scissile bond between the glycosidic oxygen and anomeric carbon atoms in (NAG-NAM)₃ using the CHARMM restrain distance methodology.²³ The product structures consisted of two molecules: the first one comprised NAG-3, NAM-2, and NAG-1, and the second one comprised NAM+1, NAG+2, and NAM+3 trisaccharides (see Figure 4b). In the substrate structures of both paths A and B, the first part of the products forms

hydrogen bonds with Gly162, Gly164, Gln223, Ser226, and Asp227, whereas the second part of the products forms hydrogen bonds with Glu58, Glu138, Asp227, and Lys233 (Figure S4). Overall, the interaction pattern does not change considerably in the product structures in comparison with the QM/MM reactant structures, indicating its stability in both the substrate and product configurations.

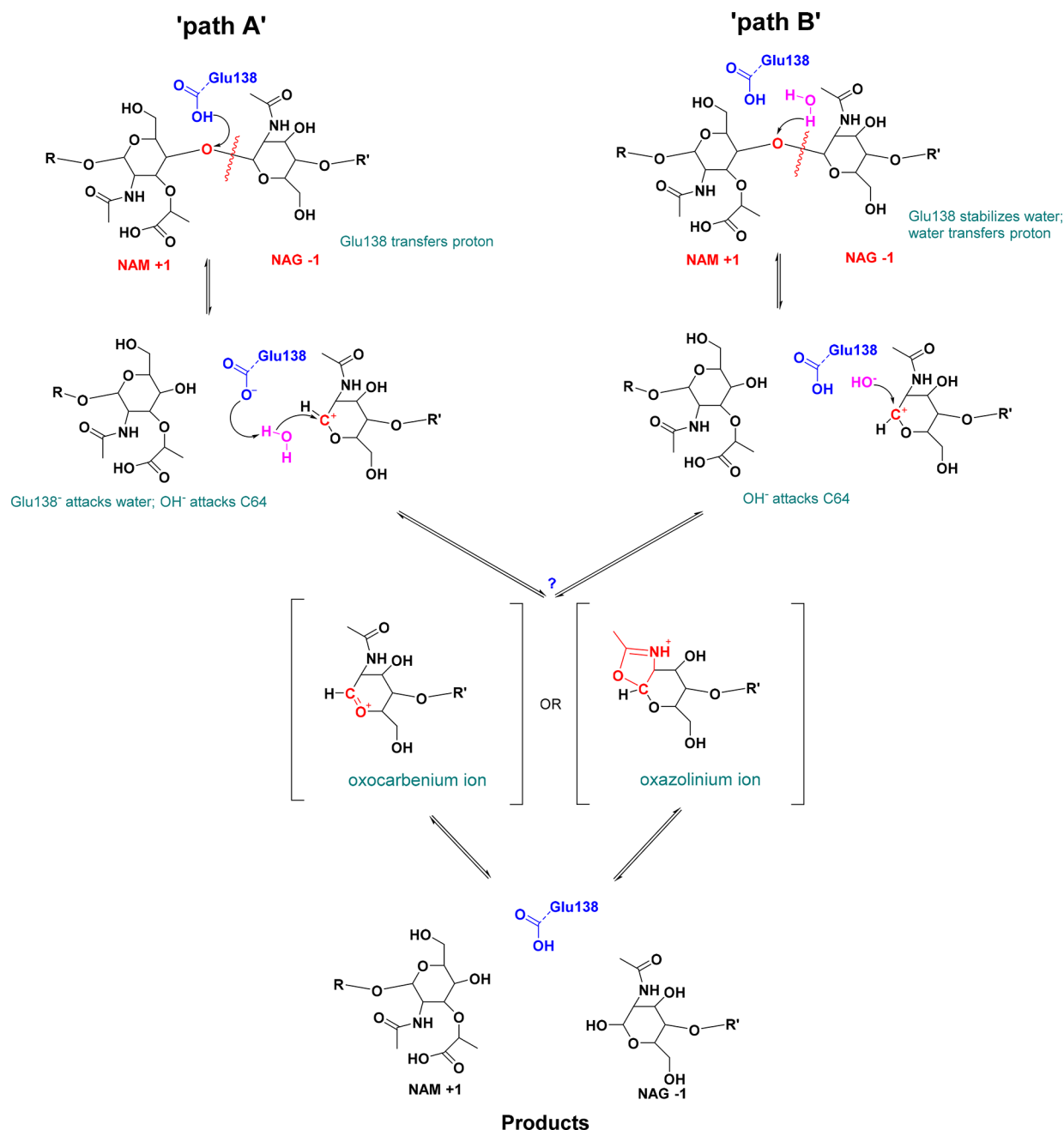
Reaction Pathways A and B. For both path A and path B, 14 replicas (Figures 5a and 6a, respectively) were generated by linearly interpolating the QM/MM optimized coordinates between the QM/MM optimized corresponding pairs of AtLE structures containing reactants and products to generate the initial structures on the reaction coordinate. The replica path method was used to QM/MM optimize the geometries of the reaction pathways. After the reaction paths were optimized, the pathways were visualized (see reaction animations of the calculated paths A and B in the Supporting information), and the crucial distances between atoms of the interacting partners along the replica pathway were determined (Figures 5b and 6b).

We identified four distances for path A, which effectively monitor the reaction mechanism of the AtLE catalysis (Figure 5b).

From a mechanistic viewpoint, R1 represents the attack of the water hydroxyl anion to the anomeric carbon atom of NAG-1, R2 corresponds to the proton transfer from the water to Glu138, R3 corresponds to the proton transfer from Glu138 to the glycosidic bond, and R4 corresponds to the breaking of the glycosidic bond. In the starting reactant structure replica 1, the following values were obtained: R1 = 5.0 Å; R2 = 2.0 Å; R3 = 1.8 Å; R4 = 1.4 Å.

In the first eight replicas of path A, the R1 distance decreased considerably, from 5.0 to 4.0 Å, corresponding geometrically to the water molecule repositioning itself to its reactive position in the AtLE enzyme. In these replicas, the NAG-1 substrate unit has a chair ⁴C₁ hexopyranose conformation. In replica 9, the R1 distance between the water and the anomeric carbon atom of NAG-1 dropped further to 3.6 Å, probably optimally placing the reactants before the initiation of the reaction. We also observed the conformational distortion of the ⁴C₁ hexopyranose NAG-1 into the half-chair/envelope conformation, approximately corresponding to a state between half-chair ⁴H₃ and envelope E₃ conformations based on the Cremer-Pople parameters.

In replica 10, proton transfer from the protonated Glu138 to the glycosidic oxygen occurred (R3, 1.0 Å) along with breakage of the glycosidic bond (R4 = 2.1 Å). In addition, the properly oriented oxygen of the water molecule was positioned near the positively charged anomeric carbon of the NAG-1 substrate (R1 = 2.8 Å), and Glu138 formed a H bond with the proton from the water molecule, causing the gradual formation of the reacting nucleophilic hydroxyl anion (R2 = 1.7 Å). The geometry of the NAG-1 hexopyranose moiety here closely resembled the oxocarbenium ion like TS reported previously.²⁴ In replica 10, the conformation of the NAG-1 hexopyranose state still resembled the previous half-chair ⁴H₃ and envelope E₃ conformations with the NAG-1 conformation displaying boat conformation properties. The reacting species was stabilized by H bonding with Phe63, Ala225, Ser226, and Asp227 in the catalytic site (Figure S5a). The rest of the substrate molecule was stabilized by Glu58, Glu138, Tyr201, Gln221, His222, Gln223, Tyr224, and Lys233; this was not significantly different from the reactant structure interaction pattern (Figure S4). The

Scheme 1. Reaction Schemes of the Proposed Reaction Mechanisms of the Hydrolysis of the Glycosidic Bond in AtIE^a

^aPath A represents a direct proton transfer from the Glu138 to the scissile glycosidic bond, while path B represents a reaction via a water molecule. In the second step of both reactions, the glycosidic bond breaks, while in the third step, the hydroxyl anion attacks the anomeric carbon to form products in the final step. The possible oxocarbenium ion and oxazolinium ion intermediates/transition states leading to the products are depicted in brackets.

structure of replica 10 thus represents an approximation of the transition state (TS) at the highest point on the reaction energy diagram with a relative energy of 45 kcal/mol.

In replica 11, product formation, represented by a covalent bond between the anomeric carbon and the oxygen on the water molecule, is observed, following the nucleophilic hydroxyl ion attack as the R1 distance between them drops to 1.5 Å. The configuration at the newly formed anomeric center of the NAG-1 substrate is identical with the stereochemistry in the substrate structure caused by the position of the reacting hydroxyl nucleophile, indicating the retaining mechanism. The

charged oxygen of Glu138 accepts the remaining proton from the water molecule (R2 = 1 Å). In replica 11, a complete separation of the NAG-1 and NAM+1 units of the initial substrate (NAG-NAM)₃ is observed as the R4 distance between them is increased to 2.6 Å.

In replicas 12–16, only smaller conformational changes that lead to the hydrolyzed products occur. The R4 distance between the hydroxyl oxygen and anomeric carbon atom drops further to its final value of 1.4 Å, and the distance between the products (R4) increases to 3.3 Å. From replica 10, the NAG-1 conformation is transformed from the formed boat con-

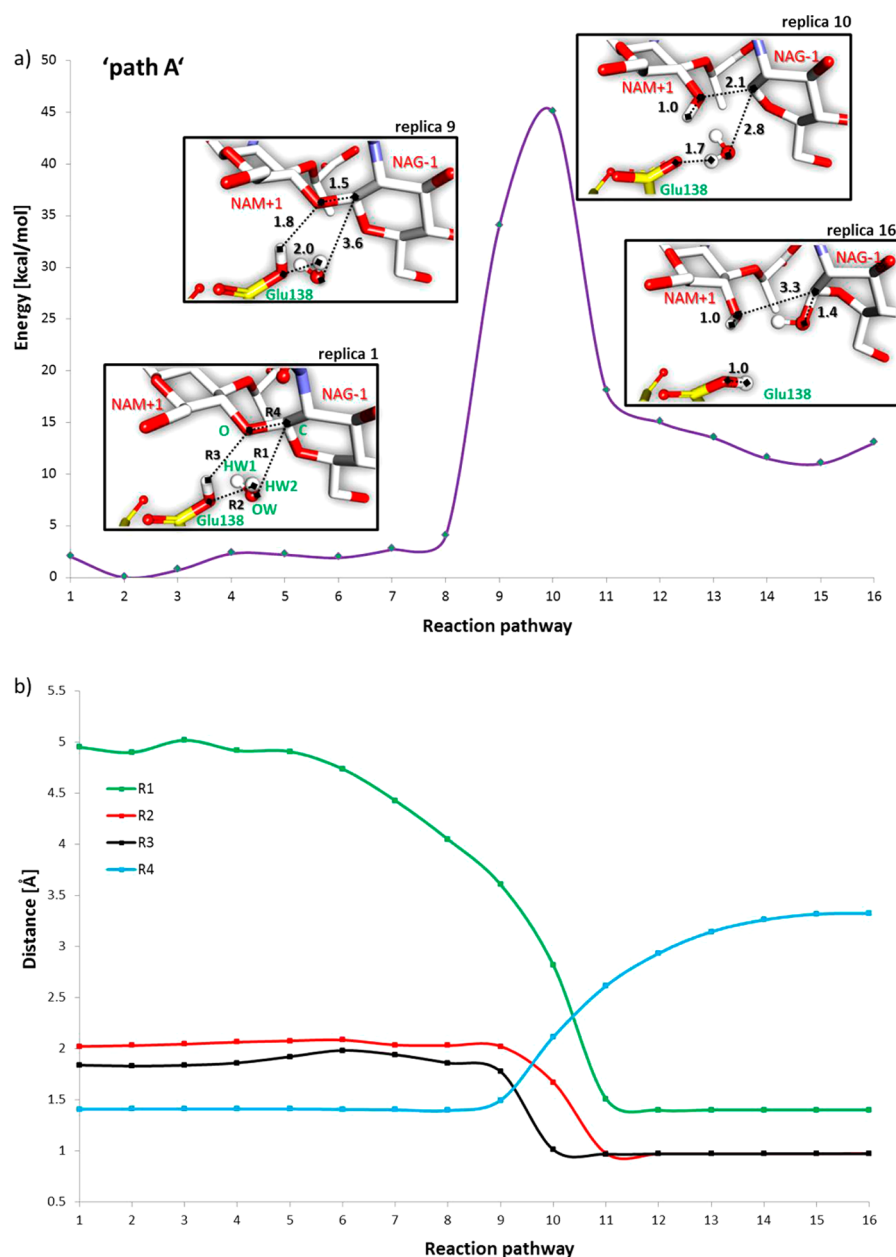


Figure 5. (a) Optimized energy reaction profile of reaction path A determined by the replica path method with the distances that monitor this reaction mechanism catalyzed by AtLE: R1 is the distance between the water oxygen OW and the anomeric carbon of NAG-1, R2 is the distance between the water hydrogen HW2 and the Glu138-OH, R3 is the distance between the Glu138-OH and the glycosidic oxygen, and R4 is the distance between the glycosidic oxygen and the anomeric carbon. (b) Values of the selected distances along the reaction pathway: R1, green line; R2, red line; R3, black line; R4, blue line.

formation to a skew boat, suggesting a final development of the NAG-1 moiety back to a chair conformation. Other hexopyranose units of the (NAG-NAM)₃ substrate preserve their initial conformation throughout the calculated reaction pathway. Overall, in path A, the oxocarbenium ion like TS (replica 10) appears to be the most energetically demanding reaction state, where proton transfer is completed and the glycosidic bond is cleaved just before the final attack of the hydroxyl anion on the anomeric carbon.

We used an equivalent approach to describe path B with the four representative distances R5, R6, R7, and R8 (Figure 6).

Mechanistically, R5 describes the proton transfer event from the water to the glycosidic oxygen, R6 describes the distance between the transferred water proton and the water oxygen, R7

represents the attack of the water-derived hydroxyl anion on the anomeric carbon atom of NAG-1, and finally, R8 describes the cleavage of the glycosidic bond. The starting reactant structure, replica 1, has R5 = 1.8 Å, R6 = 1 Å, R7 = 3.5 Å, and R8 = 1.4 Å.

Similar to path A, in the first eight replicas of the optimized path B, active repositioning of the water molecule to its reactive location is observed. The water molecule forms two hydrogen bonds, one with Glu138 and the other with the glycosidic oxygen. In the first eight replica frames, the NAG-1 substrate unit is in the chair ⁴C₁ hexopyranose conformation. In replica 9, the water-reacting proton is brought closer to the glycosidic oxygen (R5 = 1.6 Å). Concurrently, this movement also drives the molecule toward the anomeric carbon atom of NAG-1 (R7

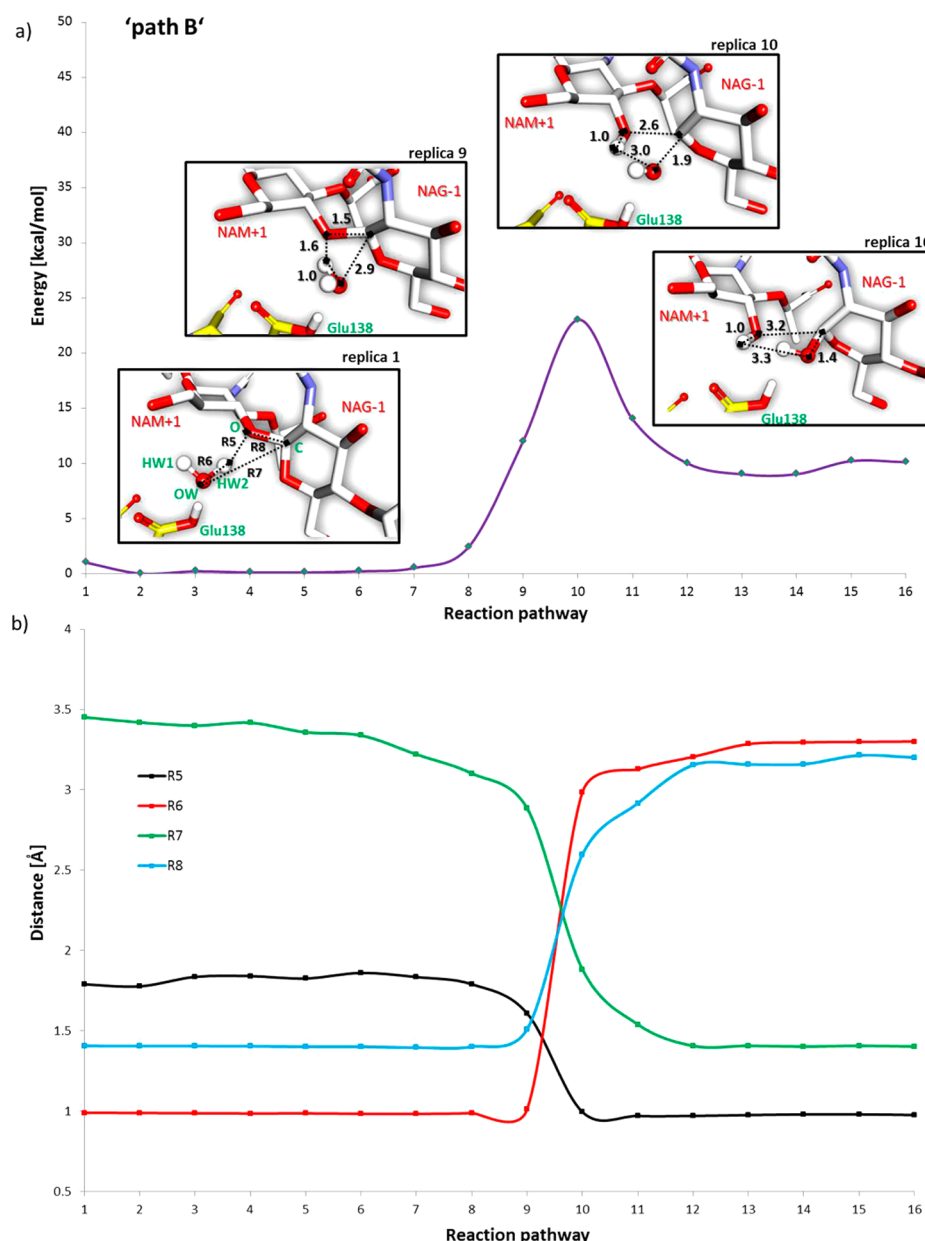


Figure 6. (a) Optimized energy reaction profile of reaction path B, determined by the replica path method. The representative distances important for the description of this AtIE catalytic mechanism are as follows: R5, the distance between the water proton, HW2, and the glycosidic oxygen; R6, the distance between the water oxygen, OW, and the glycosidic covalent bond between the glycosidic oxygen atom and the anomeric carbon of the NAG-1 unit. (b) Values for the R5 (green line), R6 (red line), R7 (black line), and R8 (blue line) distances along the reaction pathway.

= 2.9 Å). Additionally, in this replica, the $^4C_1$ hexopyranose conformation starts to slightly distort to a state comparable to that in path A, between half-chair $^2H_3$ and envelope E_3 conformations.

In replica 10, we observed a proton transfer of the water molecule proton to the glycosidic oxygen of NAG-1 ($R_5 = 1$ Å) and the formation of the reactive hydroxyl ion ($R_6 = 3$ Å) being in a favorable position for the nucleophilic attack on the anomeric carbon ($R_7 = 1.9$ Å) (Figure 6). This event is accompanied by cleavage of the glycosidic bond ($R_8 = 3$ Å). Similarly, as in path A, the geometry of the NAG-1 hexopyranose moiety closely resembles the oxocarbenium ion like TS. The NAG-1 hexopyranose conformation stays in an in-between half-chain/envelope conformation similar to that in

the previous replica stabilized by H-bonding with Phe63, Ala225, Ser226, and Asp227, while the remaining substrate units interact with Glu58, Glu138, Gly164, Tyr201, His222, Gln223, Tyr224, and Lys233, which are comparable to the observations in path A and not significantly different from the situation in the QM/MM-optimized reactant structure (Figure S5b). Replica 10 in path B represents an approximation of the TS with the highest relative energy (23 kcal/mol).

Replica 11 closely resembles the final product configuration when the covalent bond between the anomeric carbon and hydroxyl oxygen is formed ($R_8 = 1.5$ Å). The stereochemistry relative to NAG-1 is retained. In replicas 12–16, minor repositioning of the final products is observed. In these replicas, the NAG-1 hexopyranose half-chain/envelope conformation

undergoes a comparable conformational change through a boat conformation to a skew boat in replica 16, further implying conversion to the chair conformation of the formed product.

It should be noted that we unsuccessfully attempted to obtain stable oxocarbenium ion and oxazolinium ion intermediates and an oxazolinium ion like TS, the latter ruling out a substrate-assisted mechanism (see the [Supporting Information and Figure S6](#)). Hence, our results suggest that formation of the stable intermediates is not possible for AtlE, unlike the case for some other studied systems.^{25,26}

DISCUSSION

For the retention and inversion acid/base mechanisms, the average distances between the proton donor and general base are approximately 5.5 and 10 Å, respectively.²⁷ The importance of Glu224, Glu223, Glu223, Glu156, and Glu65 as putative secondary catalytic residues in substrate degradation was indicated by the structural and mutational studies of FlgJ from *Sphingomonas* sp. A1,¹¹ *Salmonella enterica*,¹⁴ *Salmonella typhimurium*,²⁸ Auto from *Listeria monocytogenes*,¹³ and TM0633 from *Termotoga maritima*.⁸ However, for the enzymes of the GH73 family examined in this study, the distances between the carboxylic groups are all greater than 10 Å. Hence, to position the putative second catalytic residue imbedded in the hairpin region above the active site during catalysis close enough to the substrate, the hairpin region would have to drastically change its conformation. This points to the fact that these proposals lack the support of structural evidence. In addition, our MD simulation of the NAG-NAM polymer bound to the active site of the crystal structure of AtlE indicated that the modeled structure remained stable during the simulation. There is neither unfolding of the lobe structure nor significant flexibility of the substrate model. Moreover, in the AtlE structure, the L-lobe is structured and appears to adopt a stable form that is not a hairpin, and the two potentially similar Glu residues (Glu179 and Glu181) are positioned away from the active site cleft. Furthermore, the equivalence in the position of the proposed secondary Glu residue involved in the catalysis is not clearly exposed by a conserved position in the sequence alignment throughout the GH73 family.⁸

The next argument for the absence of the secondary catalytic residue is provided by the comparison of *N*-acetylglucosaminidase substrate binding to the c- and g-type lysozyme mechanisms.^{6,7} The secondary catalytic carboxylate group in lysozyme approaches the C1 atom from the same side of the carbohydrate ring as the primary catalytic residue. In contrast, in our recent work,¹⁰ we showed that the *N*-acetylglucosaminidases bind their substrates from the other side as that of muramidases (lysozymes). This suggests that the second acidic residue is to be found on the right side of the active site cleft and not on the left, as lysozymes and glucosaminidases display mirror selectivity. On the right-hand side of the active site cleft of AtlE, there is a highly conserved residue, Asp227, positioned at a similar distance of 8.5 Å to the catalytic glutamate as the Asp101 in the Atlantic code lysozyme structure.⁷ However, this residue did not attract much attention as the candidate for the second catalytic residue in spite of its high conservation. The Asp227Gln/Ala mutants showed that Asp227 is a sensitive residue affecting the rate of hydrolysis; however, the Asp227Asn mutant retained low activity, and in an MD simulation, a solvent molecule is not inserted between the Asp227 and the substrate, analogous to the g-type lysozyme.⁷ Therefore, we concluded that the role of Asp227 is to position

the substrate using a hydrogen bond with the NAM(+1) residue hydroxyl group. Similarly, the mutants of the equivalent position in *Streptococcus pneumoniae* LytB²⁹ and *Staphylococcus warneri*¹⁶ yielded an almost fully active protein.¹⁶ In addition, this Asp is not strictly conserved throughout the entire GH73 family.⁸

Altogether, this conundrum can be explained by the absence of the second catalytic carboxylic residue and the fact that the spread of the important positions of acidic residues reflects the differences in specificity of various enzymes. Hence, the water-assisted reaction mechanism we describe here is a plausible explanation for the catalytic mechanism of the GH73 family of enzymes.

The reactant structures were based on two clusters/states derived from the 100 ns MD simulation using the crystal structure of AtlE. This MD simulation allowed sufficient sampling of the conformational space of the protein–substrate complex to ensure stability of the complex.²⁴ During the MD simulation also the modeled structure and the substrate remained stable and there was neither any observable unfolding of the lobe structure nor significant flexibility of the substrate model (see RMSs graphs in [Figure S2](#)). Therefore, no additional less frequent clustered structures observed could provide valuable alternative starting structures for AtlE studied mechanisms. The analysis showed that the water molecule can be present between Glu138 and the AtlE substrate which exposed water-assisted and substrate-assisted catalytic mechanisms as two possible alternatives.

In search of a possible catalytic mechanism of AtlE, we used the replica path method to model paths A and B. Unlike popularly used restrained coordinate driving (RCD) methods, the replica path method removes the need for the predetermination of the reaction coordinate. Thus, it circumvents the biased procedure of manually defining reaction steps and TS-like structures and independently provides the sequence of reaction steps.¹⁸ Site-directed mutagenesis studies coupled with MD simulations identified two possible mechanisms of proton transfer to the scissile glycoside bond oxygen atom, via Glu138 or via a water molecule. Both mechanisms do not require the secondary catalytic carboxylic group, and both retain their configuration at the anomeric centers. The Glu138 catalytic residue is in a syn orientation relative to the endocyclic C-5 to O-5 bond. Both calculated pathways proceed via the oxocarbenium ion like TS, and no stable intermediates are observed. The difference between the paths is the role of the catalytic Glu138 in the mechanism. In path A, Glu138 directly transfers the proton to the glycosidic oxygen, which leads to cleavage of the glycoside bond followed by a nucleophilic attack of the polarized water molecule on the C1 atom of the NAG(−1) residue and, finally, the regeneration of the proton to Glu138 by the remaining water proton. In path B, the proton transfer to the glycosidic oxygen is facilitated by a water molecule, which leads to the cleavage of the glycoside bond followed by nucleophilic attack of the C1 atom of the NAG(−1) residue by the formed hydroxyl ion. GH reaction mechanisms that use alternative nucleophiles have been described previously: for example, 1,2- α -fucosidase from GH95, where two Asn residues withdraw a proton from a nucleophilic water molecule, while two acidic residues enhance the nucleophilicity of water.³⁰

Our performed point mutation studies showed that mutating the AtlE Glu138 residue to Gln138 completely abolished the AtlE enzymatic activity. In visualization of path B (see the

movie of path B in the Supporting Information) the water molecule oxygen forms a hydrogen bond with the proton of the Glu138 oxygen, thus positioning it properly for the reaction to commence. This implies that such an orientation of both interacting species is crucial. The Glu proton is more acidic in comparison to Gln amide protons. Thus, it can be assumed that the Glu residue can form a stronger H bond with the reacting water molecule favorably contributing to a suitable interaction with it as well as properly polarizing it. The Glu residue can also more adequately position the water molecule to enter the reaction.

The calculated energy change (ΔE) suggests an endothermic reaction, as the ΔH value for both path A and path B was slightly above +10 kcal/mol (Figures 5a and 6a), in accordance with similar systems.²⁵ The lower energy of the transition state of 23 kcal/mol for path B in comparison with 45 kcal/mol for path A indicates that this pathway has a reasonably lower potential energy barrier and appears to be more favorable, in comparison to similar systems.²⁵ The energy difference observed between the path A and path B oxocarbenium ion like TS structures could be attributed to a more energetically favorable proton transfer to the glycosidic oxygen via a water molecule and the nucleophilic attack of a nearly ideally positioned water hydroxyl ion on the anomeric carbon. Energetic favorability of path B could further be rationalized by the fact that at its transition state the hydroxyl anion has already been formed and stabilizes the nascent positive charge at the anomeric carbon. It should, however, be noted that minimal energy pathways (MEP) like that produced by the replica path do not incorporate entropic effects. Nevertheless, these METs can be proficiently used as starting points to further approximate the full free energy of barrier crossings as well as entropy contributions by using a variety of methods such as, for example, harmonic analysis approaches.³¹

Furthermore, we unsuccessfully attempted to obtain a stable oxocarbenium ion and oxazolinium ion intermediates, ruling out substrate-assisted mechanisms (see the Supporting Information). Our results suggest that formation of the stable intermediates is not possible in our system, unlike in some other studied systems.^{25,26} The reaction in path A first involves a conformational change of the hexopyranose NAG-1 unit from the chair ${}^4C_1$ to an intermediate half-chair ${}^4H_3$ /envelope E_3 conformation of the transition state. Further reaction steps reveal a boat conformation that ends in a skew-boat state close to the chair conformation. In path B, the reaction commenced from a ${}^4C_1$ NAG-1 hexopyranose state that proceeds to a midway half-chair ${}^2H_3$ /envelope E_3 NAG-1 conformation with a final development of the NAG-1 ring conformational changes similar to those described for path A. Such conformation behavior of hexopyranose is in accordance with observations reported from similar GH enzymes.²⁴ The understanding of the reaction geometries, especially those close to the transition state structures for each GH, GH family, or type of substrate, is essential for the rational development of TS mimics, and thus, understanding the AtIE mechanism could be a step toward developing active agents that target this enzyme.³²

In this study we considered five different reaction mechanisms to obtain insight into the AtIE enzymatic reaction. In addition, unusual and rare GH mechanisms³³ usually require cofactors, exogenous bases, or alternative nucleophiles and are therefore not realistic explanations for the catalytic mechanism of the GH73 family, including AtIE. There is no experimental evidence that GH73 family enzymes require a cofactor or

exogenous base.¹⁰ Our analysis of the active site, as well as sequence conservation and other point-mutation studies of this family,⁸ gives no indication of the possible alternative (nonacidic) amino acid involved in the catalysis. Next, the retaining and inverting mechanisms utilizing two catalytic residues were excluded by the mutational studies, structural data, and conservation analysis reported in the literature.⁸ The remaining substrate-assisted and the potential water-assisted mechanism were simulated along with the reaction pathway to probe for potential stability of the oxocarbenium ion like transition state.

CONCLUSION

Site-directed mutagenesis and MD simulations suggest two possible mechanisms of proton transfer to the scissile glycosidic bond oxygen atom, directly from the catalytic Glu138 or via a water molecule. Both mechanisms retain the configuration at the anomeric centers, and neither of them requires the involvement of the secondary catalytic carboxylic group. The lower energy of the water-mediated pathway with the oxocarbenium ion like transition state obtained by QM/MM calculations suggests that hydrolysis of the glycosidic bond by AtIE proceeds via a water molecule assisted mechanism. We expect that the insights presented here will assist in setting up the experimental verification of the suggested mechanism and aid in drug discovery projects targeting the GH73 family of N-acetylglucosaminidases.

EXPERIMENTAL SECTION

Cloning and Production of AtIE Protein and Its Single-Point Mutants. Wild-type protein was produced as described previously.¹⁰ Mutants were constructed using the primers given in Table S2. Each strand was multiplied using KOD Hot Start Polymerase (EMD Millipore), and matrix DNA was degraded using DpnI (NEB). The resulting plasmids were transformed into *E. coli* DH5 α , multiplied, and isolated. The mutated genes were sequenced to confirm successful mutant construction. Mutant proteins were produced by induction with IPTG at an OD₆₀₀ = 0.6 and then incubated for 18 h at 18 °C and 250 rpm. Purification procedures were the same as for wild-type protein.

Differential Scanning Fluorimetry Experiments. DSF was conducted as described,³⁴ using a final concentration of 0.25 mg/mL of AtIE, 50 mM of MES, pH 5.5, and 5x SYPRO Orange Protein Gel Stain (Life Technologies).

Peptidoglycan Preparation and AtIE Activity Assay. *B. subtilis* peptidoglycan was purified as described previously³⁵ with minor adjustments; a sonicator was used to break up the cells. Peptidoglycan was dyed with Remazol Brilliant Blue (Sigma), and the activity assay was conducted as described previously³⁶ in 50 mM of MES, pH 5.5, at 20 °C and 250 rpm with 5 μ M of enzymes. Aliquots were taken out at 5 min intervals, and the reaction was stopped with a half volume of 96% ethanol. A₅₉₅ was plotted against time, and the initial velocity rate was determined as the slope of the curve. The experiments were repeated three times. Percentages of activities were expressed as the initial velocity rate of the mutant protein divided by the initial velocity rate of the wild-type enzyme.

Molecular Dynamics (MD) of the AtIE Remodeled Complex with (NAG-NAM)₃. MD calculations of AtIE from the 4PIA crystal structure in complex with the remodeled substrate were performed using the CHARMM molecular modeling suite.³⁷ The remodeled substrate positioned in the

complex was derived from a previous study of ours.¹⁰ The remodeled substrate, consisting of six monosaccharide units with the alternating NAG-NAM sequence in the active site cleft of AtIE, was constructed using alignment with several structural complexes of glucosaminidase enzymes and G-type lysozyme (PDB IDs: 4PI7, 4PI9, 3GXR, and 148L). His79 and Lys80 from the 4PIA crystal structure were not seen on the experimental electron density map. The Discovery Studio (BIOVIA) module of automatic protein preparation was used to construct missing residues, which were not positioned in the vicinity of the substrate binding groove.

The protonation states of all ionizable amino acid residues were assigned on the basis of the pK_a estimated via PROPKA version 3.0 software³⁸ at a pH of 5.5. Aspartate and glutamate residues were deprotonated, with the exception of the catalytic Glu138, which was protonated with the PROPKA pK_a value of 6.36, consistent with the proposed glycosylation mechanism of similar enzymes^{25,39} and the position of chlorine ions in the vicinity of the catalytic Glu138 in the crystal structure of AtIE. Histidine residues, none of which were located near the active site, were modeled in their neutral states; their tautomeric state was assigned on the basis of hydrogen bonding using WHATIF.⁴⁰

The CHARMM-GUI environment⁴¹ was used for protein manipulation and construction of the solvated protein–substrate complexes. The CHARMM parameter and topology files (version 27)^{42,43} were utilized to specify the force field parameters of the amino acid residues comprising the protein. The CHARMM General Force Field (CGenFF)⁴⁴ was used to describe the atom types and partial charges of the substrate in the Discovery Studio module simulations. The determined partial charges and assigned atom types for the substrate are given in Table S1 in the Supporting Information.

The protein and its (NAM-NAG)₃ unit substrate were immersed into a sphere TIP3⁴⁵ of water molecules with a truncated octahedral shape and an edge distance of 10 Å. A single potassium ion was added to make the system electroneutral. Ion placement was done using the Monte Carlo method. The periodic boundary conditions were applied on the basis of the shape and size of the system. Grid information for the Particle-Mesh Ewald Fast Fourier Transform was generated automatically. The final system prepared for the MD simulation was composed of 37246 atoms. To remove bad steric contacts, short steps of energy minimization were performed. The system was then minimized for 1000 steps by the steepest descent method followed by 1000 steps of the modified Adopted Basis Newton–Raphson method and an MD equilibration run of 350 ps. The production MD trajectory was generated using a leapfrog integration scheme and 2 fs simulation step using the SHAKE algorithm. A 100 ns MD simulation production run was performed. Conformations were sampled every 100th step, resulting in 25000 conformations for subsequent analysis. Visualization and analysis of the geometry parameters of the production MD trajectory were performed using the VMD program.⁴⁶

Definition of the QM Region for the QM/MM Simulations. The QM region for the QM/MM simulations was selected by taking into account the atoms present in the theoretical catalytic mechanisms of the glycoside hydrolase family members^{24,47} and the data from site-directed mutagenesis studies¹⁰ and consisted of a total of 43 atoms (Figure S3). To begin with, the Glu138 residue side chain was included in the QM region, as it is the only assumed catalytic amino acid

directly involved in the catalytic process. Next, the identified potential catalytic water molecule, essential for attacking the anomeric carbon of NAG-1 in the water-assisted mechanism, was included. Furthermore, the scissile glycoside bond between the NAM+1 and NAG-1 substrates was also included as part of the QM region, since it represents the place where the catalytic reaction occurs. Finally, the whole NAG-1 unit of the substrate, including the anomeric carbon and acetamido side group essential for forming the oxocarbenium/oxazolinium ion like transition states/intermediates for both reaction pathways, directed proton transfer via Glu138, and water-mediated catalysis were included in the QM region.

Reactants, Products, and Oxazolinium Ion Generation for QM/MM Simulations. QM/MM calculations⁴⁸ were performed using the CHARMM biomolecular simulation program.³⁶ The ab initio and density functional theory (DFT) calculations were performed with General Atomic and Molecular Electronic Structure System (GAMESS) software⁴⁹ interfaced with CHARMM. The QM region was treated with the B3LYP/6-31G(d) level of theory. Molecular mechanics calculations were performed with a dielectric constant of 1 ($\epsilon = 1$) using a classical force-field shift method and a cutoff distance of 12 Å.

Selected frames from the MD simulation were initially minimized using 10000 steps of an adopted-basis Newton–Raphson scheme to form the reactants for the QM/MM RPATH simulation. Next, the products were restrained from the structure of the reactants by making a bond between the Glu138-OH proton and the glycosidic oxygen, breaking the glycosidic bond between the glycosidic oxygen and the anomeric carbon atoms, and creating a bond between the OH[−] (water molecule) and the anomeric carbon.

For the generation of the desired oxazolinium ion, the starting points were the product structures. Subsequently, the oxazolinium ion was restrained from the first product structure (units NAG-3, NAM-2, and NAG-1) on the NAG-1 unit by closing the ring via the formation of the bond between the acetylic O88 oxygen and the anomeric carbon atom and breaking the bond between the anomeric carbon and the glycosidic oxygen atoms.

For the products and the oxazolinium ion generation CHARMM restrain distance (RESD) methodology, $k_r = 100.0$ kcal/(mol Å²) was applied, and no weighting factor was used.²³ RESD methodology was applied in 100 QM/MM minimization steps. Subsequently, the restraints were removed, and the system was minimized by another 10000 steps using the Adopted-Basis Newton–Raphson scheme. Schematic structures of the reactants and products are presented in Figure 1b and Figure S4, and the schematic structure of the oxazolinium ion is presented in Scheme 1 and Figure S6.

Calculation of the Reaction A and B Pathways Using the Replica Path Method. The replica path (RPATH) method¹⁹ is an extension of the self-penalty walk method by Elber and co-workers.^{50,51} The 14 initial replicas for each of the two investigated reaction pathways, path A and path B, were generated by linearly interpolating the coordinates between all the atoms of the QM/MM geometrically optimized starting structure, A or B, and the desired corresponding product, A or B. The RPATH method uses a combined minimization involving the sum of the configurational energies coupled with the two RPATH penalty energy terms. One RPATH penalty term restrains the distances between adjacent replica points, ensuring that the generated pathway is smooth and

evenly spaced. The second RPATH term restrains the angle between an adjacent and the next adjacent pathway. In this way, the RPATH method ensures that the optimized replicas represent the reaction pathway.^{18,19}

QM/MM RPATH geometry optimizations of the initially generated path A and path B were performed using the Adopted Basis Newton–Raphson method.⁵² Parameters included in the QM/MM replica penalty terms were set to the following values: $K_{\text{rms}} = 2000.0 \text{ kcal}/(\text{mol } \text{\AA}^2)$, $K_{\text{angle}} = 100.0 \text{ kcal}/(\text{mol } \text{\AA})$, and $\text{COSMAX} = 0.95 \text{ rad}$. Only the QM region was weighted in the replica path RMS calculation to avoid overestimating the contributions of the atoms not significantly involved in the reaction.⁵² Reaction pathways were minimized for 10000 steps or until the total pathway root-mean-square gradient was less than $0.01 \text{ kcal}/(\text{mol } \text{\AA})$ and the change in total pathway energy was less than 1.0 kcal/mol for at least 30 consecutive steps. Finally, the energy profiles were calculated for both the studied reaction mechanisms. Calculations were carried out on 128 CPU units (8 CPU/replica) within the cluster at the National Institute of Chemistry in Ljubljana.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscatal.8b01064.

Supporting scheme, figures, and tables as described in the text and a description of attempts to obtain stable intermediates and the oxazolinium ion like TS (PDF)

Reaction animation – Path A (AVI)

Reaction animation – Path B (AVI)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail for J.B.: jure.borisek@ki.si.

*E-mail for A.P.: andrej.perdih@ki.si.

*E-mail for M.N.: marjana.novic@ki.si.

Author Contributions

[†]J.B. and S.P. contributed equally to the paper.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge financial support from the Slovenian Research Agency (Research core funding nos. P1-0017, P1-0012, and P1-0048). We also acknowledge the Supercomputing Center at the National Institute of Chemistry of Slovenia for computational resources.

■ REFERENCES

- (1) Rabin, N.; Zheng, Y.; Opoku-Temeng, C.; Du, Y. X.; Bonsu, E.; Sintim, H. O. Biofilm formation mechanisms and targets for developing antibiofilm agents. *Future Med. Chem.* **2015**, *7*, 647.
- (2) Zoll, S.; Patzold, B.; Schlag, M.; Gotz, F.; Kalbacher, H.; Stehle, T. Structural basis of cell wall cleavage by a staphylococcal autolysin. *PLoS Pathog.* **2010**, *6*, e1000807.
- (3) Koshland, D. E. Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev.* **1953**, *28*, 416.
- (4) Vocadlo, D. J.; Davies, G. J. Mechanistic insights into glycosidase chemistry. *Curr. Opin. Chem. Biol.* **2008**, *12*, 539.
- (5) Whitworth, G. E.; Macauley, M. S.; Stubbs, K. A.; Dennis, R. J.; Taylor, E. J.; Davies, G. J.; Greig, I. R.; Vocadlo, D. J. Analysis of PUGNAc and NAG-thiazoline as transition state analogues for human

O-GlcNAcase: mechanistic and structural insights into inhibitor selectivity and transition state poise. *J. Am. Chem. Soc.* **2007**, *129*, 635–644.

(6) Vocadlo, D. J.; Davies, G. J.; Laine, R.; Withers, S. G. Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature* **2001**, *412*, 835.

(7) Helland, R.; Larsen, R. L.; Finstad, S.; Kyomuhendo, P.; Larsen, A. N. Crystal structures of g-type lysozyme from Atlantic cod shed new light on substrate binding and the catalytic mechanism. *Cell. Mol. Life Sci.* **2009**, *66*, 2585.

(8) Lipski, A.; Herve, M.; Lombard, V.; Nurizzo, D.; Mengin-Lecreulx, D.; Bourne, Y.; Vincent, F. Structural and biochemical characterization of the β -N-acetylglucosaminidase from *Thermotoga maritima*: toward rationalization of mechanistic knowledge in the GH73 family. *Glycobiology* **2015**, *25*, 319.

(9) Wohlkonig, A.; Huet, J.; Looze, Y.; Wintjens, R. Structural relationships in the lysozyme superfamily: significant evidence for glycoside hydrolase signature motifs. *PLoS One* **2010**, *5*, e15388.

(10) Mihelic, M.; Vlahovick-Kahlina, K.; Renko, M.; Mesnage, S.; Dobersek, A.; Taler-Vercic, A.; Jakas, A.; Turk, D. The mechanism behind the selection of two different cleavage sites in NAG-NAM polymers. *IUCrJ* **2017**, *4*, 185.

(11) Maruyama, Y.; Ochiai, A.; Itoh, T.; Mikami, B.; Hashimoto, W.; Murata, K. Mutational studies of the peptidoglycan hydrolase FlgJ of *Sphingomonas* sp. strain A1. *J. Basic Microbiol.* **2010**, *50*, 311.

(12) Inagaki, N.; Iguchi, A.; Yokoyama, T.; Yokoi, K.; Ono, Y.; Yamakawa, A.; Taketo, A.; Kodaira, K. Molecular properties of the glucosaminidase AcmA from *Lactococcus lactis* MG1363: Mutational and biochemical analyses. *Gene* **2009**, *447*, 61.

(13) Bubltz, M.; Polle, L.; Holland, C.; Heinz, D. W.; Nimtz, M.; Schubert, W. D. Structural basis for autoinhibition and activation of Auto, a virulence-associated peptidoglycan hydrolase of *Listeria monocytogenes*. *Mol. Microbiol.* **2009**, *71*, 1509.

(14) Herlihey, F. A.; Moynihan, P. J.; Clarke, A. J. The Essential Protein for Bacterial Flagella Formation FlgJ Functions as a beta-N-Acetylglucosaminidase. *J. Biol. Chem.* **2014**, *289*, 31029.

(15) Tamai, E.; Sekiya, H.; Goda, E.; Makihata, N.; Maki, J.; Yoshida, H.; Kamitori, S. Structural and biochemical characterization of the *Clostridium perfringens* autolysin catalytic domain. *FEBS Lett.* **2017**, *591*, 231.

(16) Yokoi, K. J.; Sugahara, K.; Iguchi, A.; Nishitani, G.; Ikeda, M.; Shimada, T.; Inagaki, N.; Yamakawa, A.; Taketo, A.; Kodaira, K. I. Molecular properties of the putative autolysin AtIWM encoded by *Staphylococcus warneri* M: Mutational and biochemical analyses of the amidase and glucosaminidase domains. *Gene* **2008**, *416*, 66.

(17) Rico-Lastres, P.; Diez-Martinez, R.; Iglesias-Bexiga, M.; Bustamante, N.; Aldridge, C.; Heseck, D.; Lee, M.; Mobashery, S.; Gray, J.; Vollmer, W.; Garcia, P.; Menendez, M. Substrate recognition and catalysis by LytB, a pneumococcal peptidoglycan hydrolase involved in virulence. *Sci. Rep.* **2015**, *5*, 16198.

(18) Woodcock, H. L.; Hodoscek, M.; Sherwood, P.; Lee, Y. S.; Schaefer, H. F.; Brooks, B. R. Exploring the quantum mechanical/molecular mechanical replica path method: a pathway optimization of the chorismate to prephenate Claisen rearrangement catalyzed by chorismate mutase. *Theor. Chem. Acc.* **2003**, *109*, 140.

(19) Woodcock, H. L.; Hodoscek, M.; Gilbert, A. T. B.; Gill, P. M. W.; Schaefer, H. F.; Brooks, B. R. Interfacing Q-chem and CHARMM to perform QM/MM reaction path calculations. *J. Comput. Chem.* **2007**, *28*, 1485.

(20) Sosic, I.; Gobec, M.; Brus, B.; Knez, D.; Zivec, M.; Konc, J.; Lesnik, S.; Ogrizek, M.; Obreza, A.; Zigon, D.; Janezic, D.; Mlinaric-Rascan, I.; Gobec, S. Nonpeptidic Selective Inhibitors of the Chymotrypsin-Like (Si) Subunit of the Immunoproteasome. *Angew. Chem., Int. Ed.* **2016**, *55*, 5745.

(21) Ogrizek, M.; Konc, J.; Bren, U.; Hodoscek, M.; Janezic, D. Role of magnesium ions in the reaction mechanism at the interface between Tm1631 protein and its DNA ligand. *Chem. Cent. J.* **2016**, *10*, 41.

(22) Perdih, A.; Hodoscek, M.; Solmajer, T. MurD ligase from *E. coli*: Tetrahedral intermediate formation study by hybrid quantum

mechanical/molecular mechanical replica path method. *Proteins: Struct., Funct., Genet.* **2009**, 74, 744.

(23) Monard, G.; Prat-Resina, X.; Gonzalez-Lafont, A.; Lluch, J. M. Determination of enzymatic reaction pathways using QM/MM methods. *Int. J. Quantum Chem.* **2003**, 93, 229.

(24) Ardevol, A.; Rovira, C. Reaction Mechanisms in Carbohydrate-Active Enzymes: Glycoside Hydrolases and Glycosyltransferases. Insights from ab Initio Quantum Mechanics/Molecular Mechanics Dynamic Simulations. *J. Am. Chem. Soc.* **2015**, 137, 7528.

(25) Jitonnorn, J.; Limb, M. A. L.; Mulholland, A. J. QM/MM Free-Energy Simulations of Reaction in *Serratia marcescens* Chitinase B Reveal the Protonation State of Asp142 and the Critical Role of Tyr214. *J. Phys. Chem. B* **2014**, 118, 4771.

(26) Jitonnorn, J.; Sattayanon, C.; Kungwan, N.; Hannongbua, S. A DFT study of the unusual substrate-assisted mechanism of *Serratia marcescens* Chitinase B reveals the role of solvent and mutational effect on catalysis. *J. Mol. Graphics Modell.* **2015**, 56, 53.

(27) Davies, G.; Henrissat, B. Structures and Mechanisms of Glycosyl Hydrolases. *Structure* **1995**, 3, 853.

(28) Zaloba, P.; Bailey-Elkin, B. A.; Derksen, M.; Mark, B. L. Structural and Biochemical Insights into the Peptidoglycan Hydrolase Domain of FlgJ from *Salmonella typhimurium*. *PLoS One* **2016**, 11, e0149204.

(29) Rico-Lastres, P.; Diez-Martinez, R.; Iglesias-Bexiga, M.; Bustamante, N.; Aldridge, C.; Heseck, D.; Lee, M.; Mobashery, S.; Gray, J.; Vollmer, W.; Garcia, P.; Menendez, M. Substrate recognition and catalysis by LytB, a pneumococcal peptidoglycan hydrolase involved in virulence. *Sci. Rep.* **2015**, 5, 16198.

(30) Nagae, M.; Tsuchiya, A.; Katayama, T.; Yamamoto, K.; Wakatsuki, S.; Kato, R. Structural basis of the catalytic reaction mechanism of novel 1,2- α -L-fucosidase from *Bifidobacterium bifidum*. *J. Biol. Chem.* **2007**, 282, 18497.

(31) Tao, P.; Hodošek, M.; Larkin, J. D.; Shao, Y.; Brooks, B. R. Comparison of Three Chain-of-States Methods: Nudged Elastic Band and Replica Path with Restraints or Constraints. *J. Chem. Theory Comput.* **2012**, 8, 5035.

(32) Gloster, T. M.; Davies, G. J. Glycosidase inhibition: assessing mimicry of the transition state. *Org. Biomol. Chem.* **2010**, 8, 305.

(33) Jongkees, S. A.; Withers, S. G. Unusual enzymatic glycoside cleavage mechanisms. *Acc. Chem. Res.* **2014**, 47, 226–35.

(34) Niesen, F. H.; Berglund, H.; Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2007**, 2, 2212.

(35) Wheeler, R.; Turner, R. D.; Bailey, R. G.; Salama, B.; Mesnage, S.; Mohamad, S. A. S.; Hayhurst, E. J.; Horsburgh, M.; Hobbs, J. K.; Foster, S. J. Bacterial Cell Enlargement Requires Control of Cell Wall Stiffness Mediated by Peptidoglycan Hydrolases. *mBio* **2015**, 6, e00660-15.

(36) Zhou, R. Q.; Chen, S. G.; Recsei, P. A Dye Release Assay for Determination of Lysostaphin Activity. *Anal. Biochem.* **1988**, 171, 141.

(37) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caffisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodosscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, 30, 1545.

(38) Olsson, M. H. M.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK(a) Predictions. *J. Chem. Theory Comput.* **2011**, 7, 525.

(39) Bowman, A. L.; Grant, I. M.; Mulholland, A. J. QM/MM simulations predict a covalent intermediate in the hen egg white lysozyme reaction with its natural substrate. *Chem. Commun.* **2008**, 37, 4425.

(40) Vriend, G. What If - a Molecular Modeling and Drug Design Program. *J. Mol. Graphics* **1990**, 8, 52.

(41) Kumar, R.; Iyer, V. G.; Im, W. CHARMM-GUI: A graphical user interface for the CHARMM users. *Abstr. Pap. Am. Chem. Soc.* **2007**, 233, 273.

(42) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, 102, 3586.

(43) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **2004**, 25, 1400.

(44) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell, A. D. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2009**, 31, 671.

(45) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, 79, 926.

(46) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, 14, 33.

(47) Vuong, T. V.; Wilson, D. B. Glycoside Hydrolases: Catalytic Base/Nucleophile Diversity. *Biotechnol. Bioeng.* **2010**, 107, 195.

(48) Eurenium, K. P.; Chatfield, D. C.; Brooks, B. R.; Hodosscek, M. Enzyme mechanisms with hybrid quantum and molecular mechanical potentials. I. Theoretical considerations. *Int. J. Quantum Chem.* **1996**, 60, 1189.

(49) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General Atomic and Molecular Electronic-Structure System. *J. Comput. Chem.* **1993**, 14, 1347.

(50) Czereminski, R.; Elber, R. Reaction-Path Study of Conformational Transitions in Flexible Systems - Applications to Peptides. *J. Chem. Phys.* **1990**, 92, 5580.

(51) Elber, R.; Karplus, M. A Method for Determining Reaction Paths in Large Molecules - Application to Myoglobin. *Chem. Phys. Lett.* **1987**, 139, 375.

(52) Woodcock, H. L.; Hodosscek, M.; Brooks, B. R. Exploring SCC-DFTB paths for mapping QM/MM reaction mechanisms. *J. Phys. Chem. A* **2007**, 111, 5720.

RESEARCH PAPER



Discovery of (phenylureido)piperidinyl benzamides as prospective inhibitors of bacterial autolysin E from *Staphylococcus aureus*

Jure Borišek^a, Sara Pintar^{b,c}, Mitja Ogrizek^a, Simona Golič Grdadolnik^a, Vesna Hodnik^d, Dušan Turk^{b,e}, Andrej Perdih^a and Marjana Novič^a

^aNational Institute of Chemistry, Ljubljana, Slovenia; ^bDepartment of Biochemistry, Molecular and Structural Biology, Jozef Stefan Institute, Ljubljana, Slovenia; ^cJozef Stefan International Postgraduate School, Ljubljana, Slovenia; ^dBiotechnical Faculty, Infrastructural Center for Surface Plasmon Resonance, Ljubljana, Slovenia; ^eCentre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins, Ljubljana, Slovenia

ABSTRACT

Autolysin E (AtlE) is a cell wall degrading enzyme that catalyzes the hydrolysis of the β -1,4-glycosidic bond between the *N*-acetylglucosamine and *N*-acetylmuramic acid units of the bacterial peptidoglycan. Using our recently determined crystal structure of AtlE from *Staphylococcus aureus* and a combination of pharmacophore modeling, similarity search, and molecular docking, a series of (Phenylureido)piperidinyl benzamides were identified as potential binders and surface plasmon resonance (SPR) and saturation-transfer difference (STD) NMR experiments revealed that discovered compounds bind to AtlE in a lower micromolar range. (phenylureido)piperidinyl benzamides are the first reported non-substrate-like compounds that interact with this enzyme and enable further study of the interaction of small molecules with bacterial AtlE as potential inhibitors of this target.

ARTICLE HISTORY

Received 4 June 2018
Revised 22 June 2018
Accepted 22 June 2018

KEYWORDS

Autolysin E; glycoside hydrolase; virtual screening; SPR; STD NMR

Introduction

In the last decade resistance to antibiotics and other antimicrobial compounds became a global problem with increasing importance¹. One of the promising strategies to confront aforementioned challenge is to investigate new antibiotic targets aiming at discovery of new classes of antibiotics². In particular, remodeling and hydrolysis of cell wall remains largely unexploited territory as opposed to widely explored area of targeting peptidoglycan synthesis pathways³. Autolysin E (AtlE) from *Staphylococcus aureus* is one of the peptidoglycan degrading enzymes⁴ and its inhibition could pave the way to potential novel class of antimicrobial agents.

Staphylococcus aureus is a widespread gram-positive pathogen of humans and animals. In most cases, it lives as a commensal organism on the skin and mucous membranes and poses no threat to healthy individuals. However, when it enters into the body through breaches in skin or membranes, it may pose a serious risk especially for immunocompromised individuals. It can cope with hostile conditions encountered in the bloodstream of the living host, a scarce supply of certain nutrients, attacks of the immune system and anti-infective measures undertaken in the clinical field. List of infections it causes includes bacteremia, infective endocarditis, impetigo, surgical site infections, cutaneous abscesses, purulent cellulitis, osteomyelitis, septic arthritis, prosthetic device infections, and toxic shock syndrome^{5,6}. Another feature that makes it even more difficult to treat is its ability to form biofilm. Biofilm is a community of microorganism that is attached to the surface and plays a significant role in persistence of

bacterial infections⁷. Bacteria within biofilms are several orders of magnitude more resilient to antibiotics, compared with planktonic bacteria⁸. The huge repertoire of different virulence factors and additional supportive gene products that increase its capability to survive within the living host make *S. aureus* one of the most threatening microorganisms causing hospital and community-acquired infections⁹.

Wide-spread use of antibiotics in recent decades has resulted in emergence of antibiotic and multiple antibiotic resistant strains, such as methicillin (MRSA) and vancomycin resistant *S. aureus*. Before 1980s MRSA was considered a healthcare-associated disease, however there has been a dramatic increase in the number of community-based infections due to it¹⁰. For most of the known antibiotics, strains resistant to each, have been isolated¹¹. The threat of the spread of even more resistant *S. aureus* strains urges the development of new antibiotics targeting this organism.

The genome of *S. aureus* strain Mu50 codes for five *N*-acetylglucosaminidases, one of them being AtlE, which belongs to the glycoside hydrolases (GH) 73 family. During growth and division bacteria have to synthesise and degrade peptidoglycan. Cells with inactivated *N*-acetylglucosaminidases have morphological defects as a consequence of impaired ability to increase in size after division and adopt correct mature shape^{12,13}. Further it has been shown that some autolysins have a role in concealing bacteria from the receptor proteins that detect peptidoglycan and thus enable the pathogen to evade the immune system of the host¹⁴. Therefore, enzymes involved in the cell wall degradation can be considered as novel, valuable extracellular drug/vaccine targets. It may seem at a first glance contradictory that blocking the enzyme

CONTACT Marjana Novič ✉ marjana.novic@ki.si National Institute of Chemistry, Hajdrihova 19, SI-1001, Ljubljana, Slovenia.

 Supplemental data for this article can be accessed [here](#).

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

responsible for the degradation of the bacterial cell wall would have antimicrobial effect. However, as described above, the bacteria needs to degrade peptidoglycan in certain stages of growth and division, thus the inhibition of such enzymes would be beneficial. Interestingly, also an alternative molecular engineering approach was developed to enhance the bacteriolytic activity of autolysins and thus make them promising antimicrobial agents¹⁵.

AtIE from *S. aureus* is an enzyme from the GH73 family¹⁶. In our previous studies, the crystal structure of AtIE and its structures with fragments of its substrate have been determined and thereby the binding groove for the substrate has been experimentally identified⁴. AtIE has a binding site that can accommodate three NAG-NAM units of its natural peptidoglycan substrate and cleaves the β -1,4-glycosidic bond between the *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid (NAM). Furthermore, active site residue Glu138 was confirmed as crucial for the activity of the enzyme (Figure 1)^{4,17}.

In order to discover AtIE inhibitors, several computational structure-based drug design tools were employed. Binding affinities of the most promising virtual hits obtained by iterative processes of pharmacophore modeling, similarity search and molecular docking were determined experimentally using surface plasmon resonance (SPR) and saturation-transfer difference (STD) NMR experiments.

Materials and methods

Investigation of the AtIE binding site and pharmacophore-based virtual screening experiments

In order to determine the starting point for the virtual screening, experimentally remodeled AtIE substrate was derived from previous study⁴. As remodeled AtIE substrate contains 3 (NAG-NAM) units, we focused on the central NAG-NAM unit being bound near the catalytic Glu138 residue. The remodeled substrate structure of the NAG-NAM unit in the AtIE binding site (PDB ID: 4PIA) was used to explore interactions using LigandScout software¹⁸ and to derive structure-based pharmacophore for the NAG-NAM ligand (see Figure 1). The pharmacophore model was used to screen approximately 5 million commercially available compounds, all of which were previously converted into multifunctional format (25 conformers for each compound in the database) using the LigandScout screening module¹⁹. The conformers of the molecules in the screening library were generated using the idbgen module available in LigandScout, coupled with the OMEGA software²⁰. The default high-throughput settings were used for the library generation: maximum number of output conformers per molecule = 25; RMS threshold to duplicate conformers = 0.8 Å; maximum number of all generated conformers per molecule = 30,000; and maximum

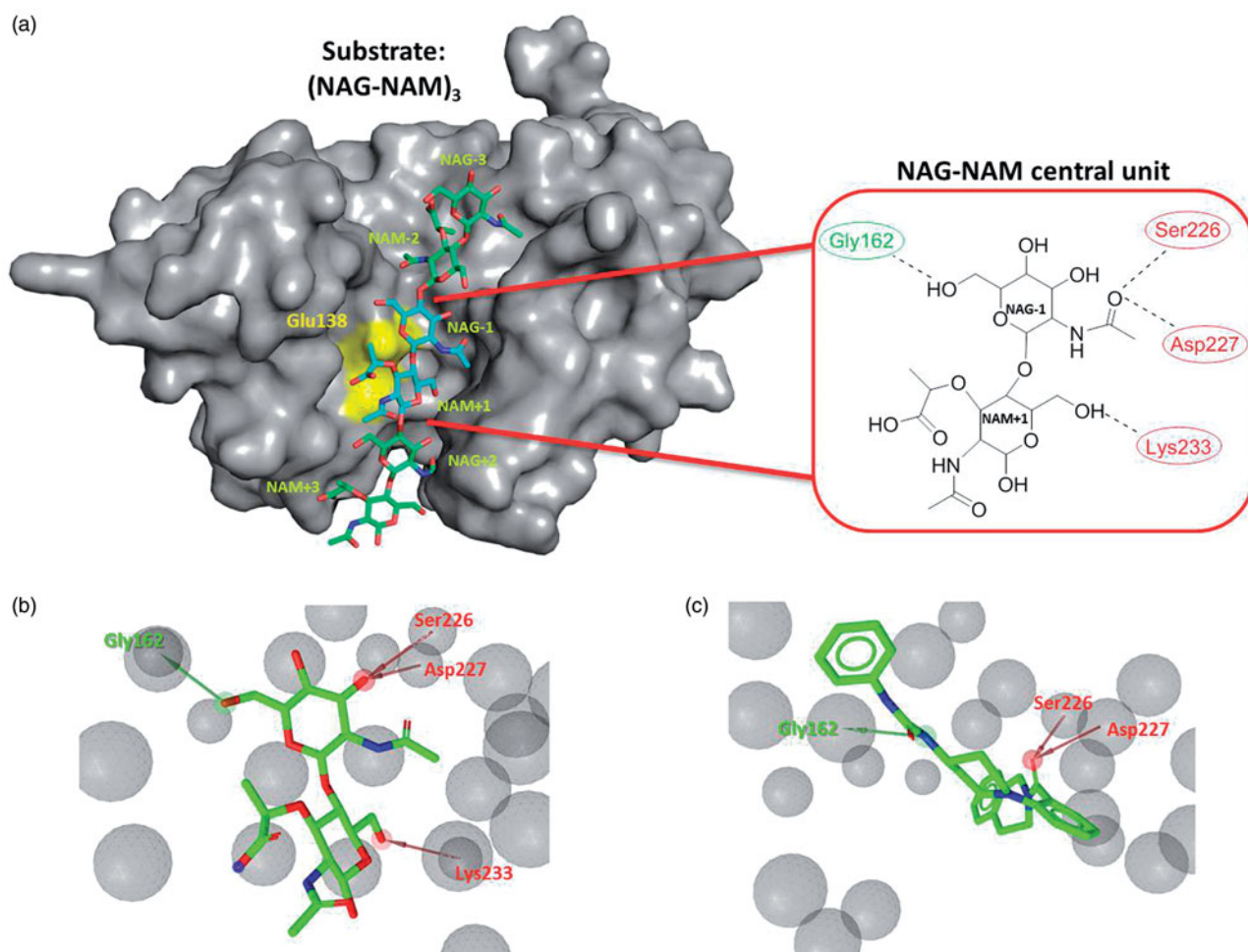


Figure 1. (a) Remodeled substrate in AtIE (PDB: 4PIA) and schematic presentation of the selected NAG-NAM central unit in AtIE for further virtual screening campaign; yellow region represents catalytic amino acid Glu138. Dotted lines represent hydrogen bonds between the central substrate unit and the protein active site residues. Green and red residues represent hydrogen bond acceptors and donors, respectively. (b) LigandScout structure-based pharmacophore for the modeled central NAG-NAM ligand within the AtIE binding site and (c) micromolar hit 3 obtained from the first screening campaign aligned to the pharmacophore model. Green arrows represent hydrogen bond donors, red arrows hydrogen bond acceptors and gray spheres exclusion volume spheres. Pharmacophore model comprised one hydrogen bond donor (interaction with Gly162) and three hydrogen bond acceptors, describing possible interactions with Ser226, Asp227 and Lys233.

number of intermediate conformers per molecule = 4000. Pharmacophore Fit scoring function was used to score the matching of the hit molecules to the pharmacophore model. The LigandScout screening procedure retrieved approximately 10,000 hit compounds that were subsequently docked in the AtIE binding site.

Molecular docking calculations

Molecular docking studies were performed using GOLD Suite v5.1 docking package²¹. Binding site was defined around remodeled substrate coordinates of the NAG-NAM ligand in radius of 12 Å. Validation of the docking settings²² was performed by redocking of the NAG-NAM ligand of the remodeled substrate into the AtIE binding site. The observed RMSD of obtained docked poses versus the original NAG-NAM position were within the accepted limits and CHEMPLP scoring function gave the best results.

Next, each ligand was docked into the binding site by the scoring function CHEMPLP and applying the following parameters of the GOLD genetic algorithm: population size = 100, selection pressure = 1.1, no. of Operations = 100,000, no. of islands = 5, Niche Size = 2, migrate = 10, mutate = 95, crossover = 95.42.

Surface plasmon resonance (SPR) experiments

SPR experiments were performed using a Biacore T100 (GE Healthcare) equipped with a Series S sensor chip CM5 (GE Healthcare). PBS buffer (67 mM Na₂HPO₄, 12.5 mM KH₂PO₄, 70 mM NaCl, 0.05% P20, pH 7.4) was used for immobilisation of the protein. AtIE was produced as previously described¹ and attached to the surface of the chip by amine coupling. The surface of flow cells 1 and 2 was initially activated with a 10-min pulse of a 1:1 mixture of 0.4 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride and 0.1 M N-hydroxysuccinimide, according to manufacturer recommendations. AtIE was diluted into 10 mM sodium acetate buffer, pH 5.8, to a final concentration of 100 µg/mL, and injected twice for 180 s over the second flow cell. Both flow cells were blocked with a 7-min pulse of 1.0 M ethanolamine. The final immobilisation level was about 4100 RU. Screening was run at 25 °C using PBS (67 mM Na₂HPO₄, 12.5 mM KH₂PO₄ pH 7.4, 70 mM NaCl, 0.05% P20) supplemented with 1% or 5% DMSO (Merck) as the running buffer. Selected compounds from the virtual screening studies were tested at two different concentrations: 40 and 400 µM. Each compound was injected for 1 min at a flow-rate of 30 µL/min, and the dissociation was monitored for 30 s. Regeneration was not needed, although 30 s of buffer-flow was used to stabilise the surface after each injection. We compared the responses obtained by the injections of the buffer and the two different sample concentrations, and continued the study with the compounds that showed binding. Selected compounds were tested at eight different concentrations (depending on their solubility) in three parallel titrations. KD values were determined by fitting of the data to a 1:1 steady state binding model.

Nuclear magnetic resonance (NMR) experiments

The high-resolution NMR spectra were recorded on a Varian DirectDrive 800 MHz spectrometer at 25 °C. All data were collected using pulse sequences and phase-cycling routines provided in Varian libraries of pulse programs. The cryogenic triple-resonance NMR probe was used. NMR samples were prepared in a buffer containing 20 mM Tris-D11, 10% glycerol-D8, and 100 mM NaCl in

D₂O, pD 7.2. DSS (0.1 mM) was used as an internal standard. All the spectra were recorded at a protein/ligand ratio of 1:100, the protein concentration was 1 µM and the ligand concentration was 0.1 mM. The STD experiments²³ were performed with a 8802.8 Hz spectral width, 16384 data points, a saturation time of 2 s, a relaxation delay of 1.3 s, and 3800 scans. Selective saturation was achieved by a train of 50 ms long Gauss-shaped pulses separated by 1 ms delay. Water was suppressed via excitation sculpting. The on-resonance selective saturation of AtIE was applied at -0.9 ppm, which is well separated from the methyl group of compounds at 1.1 ppm. The off-resonance irradiation was applied at 30 ppm for the reference spectrum. Subtraction of the on and off-resonance spectra was performed internally via phase cycling. Spectra were zero-filled twice and apodised by an exponential line-broadening function of 1 Hz.

Results and discussion

Design phase I: virtual screening

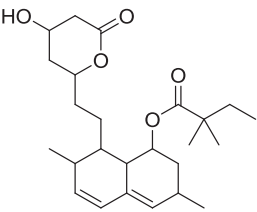
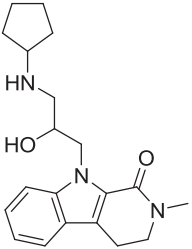
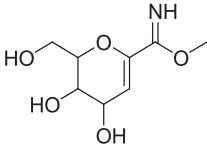
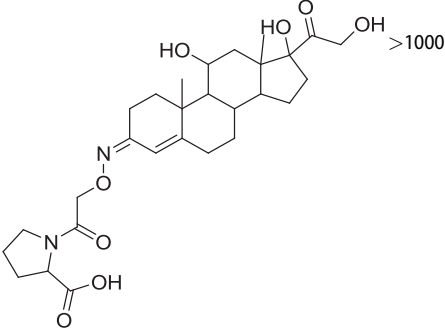
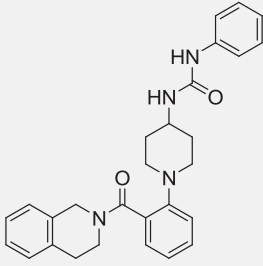
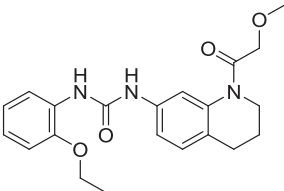
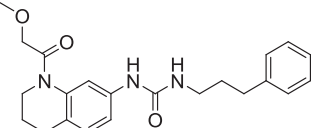
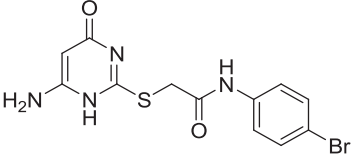
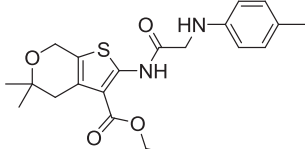
To identify AtIE inhibitors, our design starting point were the available structural information of the enzyme AtIE (PDB ID: 4PIA) and its remodeled substrate from our previous study⁴. Briefly, the substrate model in the active site cleft of AtIE was constructed by merging information from substrate fragments of superimposed complexes of AtIE and G-type lysozyme (PDB IDs: 4PI7⁴, 4PI9⁴, 3GXR²⁴ and 148L²⁵). The AtIE substrate model contained 3 NAG-NAM units. The central unit NAG-1-NAM +1 representing the cleavage site was positioned next to the catalytic Glu138⁴. This disaccharide unit was selected to derive structure based pharmacophore for the initiation of the virtual screening campaign (Figure 1(a)) and remodeled⁴ using ligands NAG and MUB from the complexes 3GXR and 148L, respectively (Figure 1(b)). It contained the interaction pattern with the AtIE based on hydrogen bonds with Gly162, Ser226, Asp227 and Lys233.

To identify molecules that retained molecular recognition pattern of the selected NAG-NAM unit, several *in silico* techniques were employed as powerful drug design tools. First, a structure-based pharmacophore model mimicking the interactions of the selected central NAG-NAM unit with the AtIE binding site was generated using LigandScout software¹⁸. Pharmacophores consisted of the following features: hydrogen bond acceptor (describing the interaction with the Gly162) and three hydrogen bond donors (observed with Ser226, Ser227 and Lys233) (Figure 1). Exclusion volume spheres were also derived, mimicking the sterical boundary conditions of the AtIE-binding site. Interaction between scissile glycosidic substrate bond and Glu138 was not identified by LigandScout as potential H-bond pharmacophore feature due to not optimal geometry. This is in accordance with our previous MD simulations of the AtIE bound with (NAG-NAM)₃ substrate which have suggested a complex role of Glu138 residue in molecular recognition and catalysis¹⁷.

In order to broaden the chemical space of the resulted virtual hits, we introduced an additional criterion that three out of initial four pharmacophore features had to be satisfied for a compound to be considered as a hit. Subsequently, this pharmacophore model was used in a large scale virtual screening campaign, using available library of approximately 5 million commercially available compounds¹⁹. This procedure resulted in approximately 10,000 virtual hits from different structural classes that were able to comply with the requested pharmacophore features.

Hits obtained were subsequently docked into the AtIE-binding site using the GOLD molecular docking tool (see Experimental

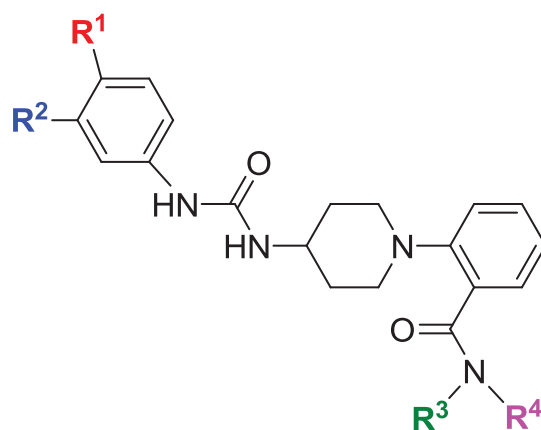
Table 1. Tested compounds with binding affinities to AtIE.

Compound	Structure	K_d (μ M)	Compound	Structure	K_d (μ M)
1		>1000	6		>1000
2		>1000	7		>1000
3		177	8		>1000
4		>1000	9		>1000
5		>1000			

Most promising hit **3** marked in gray.

section for GOLD parameter settings) to explore the proposed orientations of these molecules. A successful validation of the docking model was made by redocking of the NAG-NAM molecule in the substrate binding groove. The investigated binding site was defined as a 12 Å spherical cavity around ligand NAG-NAM coordinates. The docking procedure was performed by applying the

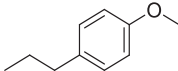
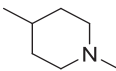
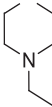
ChemPLP scoring function²⁶, and top 200 ranked docking solutions by scoring function were visually inspected for the fitness and orientation in the AtIE binding site. We were aware that scoring fitness function approach beared certain inherent limitations such as accurate ranking of docking solutions for an investigated ligand and adequate description of entropic contributions²⁷. Since

Table 2. Compounds from the structural class of (phenylureido)piperidinyI benzamides tested for AtIE binding affinity in SPR experiments.

Compd	R ¹	R ²	R ³	R ⁴	K _d (μM)
10	-H	-Cl			1.9
11	-Me	-H		-H	2.4
12	-H	-Cl		-H	2.5
13	-Me	-H		-H	6.1
14	-Me	-H		-H	11.3
15	-Me	-H		-H	11.7
16	-Me	-H			18.2
17	-H	-H		-H	21.1
18	-Me	-H		-Et	38.8
3	-H	-H			177

(continued)

Table 2. Continued.

Compd	R ¹	R ²	R ³	R ⁴	K _d (μM)
19	-Me	-H		-H	>1000
20	-Me	-H		-Me	>1000
21	-H	-H			>1000
22	-H	-H		-Et	>1000
23	-Me	-H		-H	>1000
24	-H	-H		-H	>1000
25	-Me	-H			>1000

no ligands are known for this target, we decided to utilise this option as selection criteria. There are cases in the literature that support the potential of this approach²⁸. Finally, 41 compounds from different structural classes were selected from the pool of approximately 200 hits by applying previously derived criteria of fitness, orientation and scoring function, and their binding affinities to AtIE was determined using SPR (see [Supplementary Table 1](#)). In the first design phase, several compounds from the diverse structural classes exhibited binding affinity in millimolar range ($K_d > 1000 \mu\text{M}$) [Table 1](#).

We were especially pleased to observe that compound **3** from the structural class of (phenylureido)piperidinyl benzamides was found to exhibit promising binding affinity to AtIE in the micromolar range (see [Figure 1\(c\)](#) and [Figure S1c](#) for SPR sensorgram of **3**) making it to the best of our knowledge the first identified reported compound that binds this target.

To gain an insight into the molecular recognition of **3** by AtIE molecular docking was applied. A binding mode suggested hydrogen bonds with AtIE residues Gly162, Tyr224, Ser226 and Asp227,

hydrophobic interactions with Phe63 and Tyr201, and electrostatic cation- π interaction with Lys233 ([Figure S1](#)).

Design phase II: screening and evaluation of a (phenylureido)piperidinyl benzamides focused library

The first phase of our VS campaign, described in Design Phase I, yielded a promising micromolar AtIE ligand, to the best of our knowledge the first ever described small molecule binder for this target. In order to provide initial SAR a focused library of all commercially available, (phenylureido)piperidinyl benzamides was extracted from the eMolecules library.²⁹ It consisted of 16 molecules (**10–25**), which were tested for their binding affinity to AtIE by SPR. None of these molecules were present in our initial library of screened compounds. The results of the SPR binding affinity assay are presented in [Table 2](#); SPR sensorgrams for active compounds are shown in [Figure S2](#).

Additionally, SPR experiments revealed fast and tight binding to the AtIE at low micromolar concentrations. All commercially available compounds possessing AtIE binding affinity were

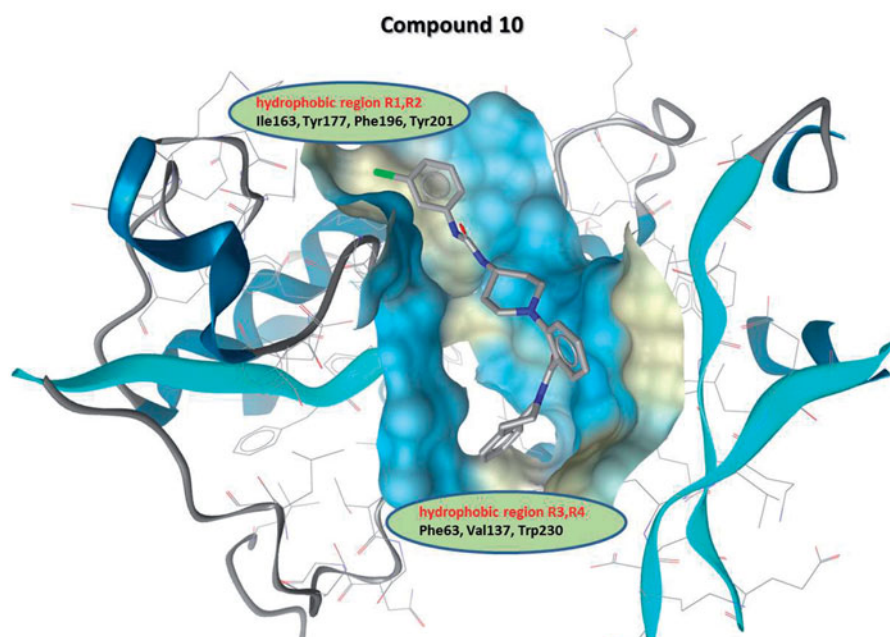


Figure 2. Hydrophilic and hydrophobic regions of AtIE binding surface for modeled compound **10** are depicted in blue and yellow, respectively. Hydrophobic region R1, R2 and R3, R4 represent the area where substituents of (phenylureido)piperidinyl benzamides R1 and R2, and R3, R4 bind, respectively.

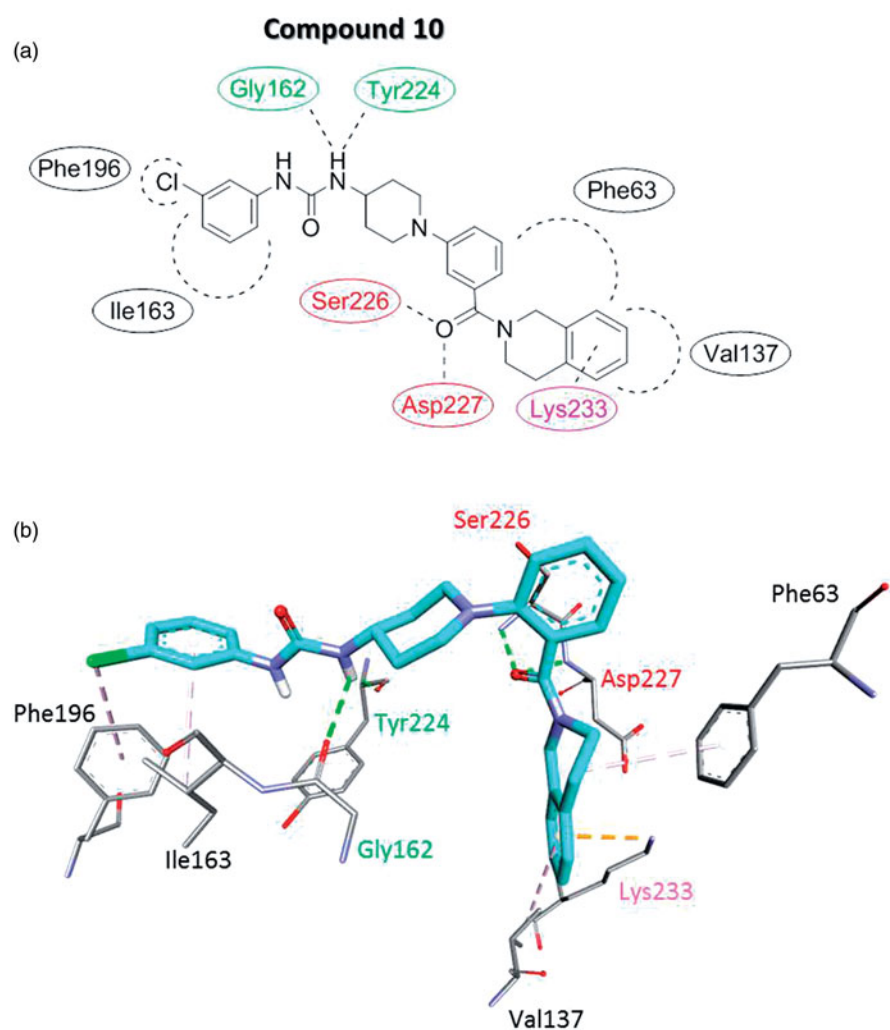


Figure 3. Proposed interaction binding pattern between the active (phenylureido)piperidinyl benzamides compound **10** and the AtIE binding site based on molecular docking calculations, depicted in (a) 2D and (b) 3D. It comprises two hydrogen bonds acceptors (Gly162 and Tyr224), two hydrogen bond donors (Ser226 and Asp227), several hydrophobic interactions with Phe63, Val137, Ile163, and Phe196 and Lys233 electrostatic cation-Pi interaction. Green residue represents hydrogen bond acceptor, red residues hydrogen bond donors, black residues form hydrophobic interactions and magenta residue forms electrostatic cation-Pi interaction.

characterised for their identity by HR-MS techniques (see [Supplementary Information](#)).

Our further SPR measurements of the focused library yielded additional nine compounds with the binding affinity to AtIE in the low micromolar range from the same structural class of (phenylureido)piperidinyl benzamides ([Table 2](#)). This enabled rough estimation of SAR properties required for binding to this target.

Based on the binding affinities and on the molecular docking calculations of compounds **3**, **10** and **16**, we concluded that the chlorine substituent at the R2 position is the most important for binding affinity as the R1 methyl containing **16** has 10 times lower affinity ($K_d = 18.2 \mu\text{M}$) and R1, R2 non-substituted **3** ($K_d = 177 \mu\text{M}$) has 100 times lower affinity than **10** ($K_d = 1.9 \mu\text{M}$), respectively. We could hypothesise that R1 and R2 substituents bind into the hydrophobic region of AtIE-binding site comprising Ile163, Tyr177, Phe196 and Tyr201 residues ([Figures 2 and 3](#)).

Compounds with cyclic lipophilic/aromatic fragments on the R3 and/or the R4 display a higher affinity ($K_d = \text{low } \mu\text{M}$ range) than compounds **22**, **23**, **24** ($K_d > 1000 \mu\text{M}$) with alkyl fragments on the R3 and/or R4 positions. The models suggested that the substituents on the R3 and/or R4 positions bind into the elongated hydrophobic region, comprised by Phe63, Val137, and Trp230 residues and also might form additional electrostatic interactions with Glu138 and Lys233. Most of the active (phenylureido)piperidinyl benzamides compounds form hydrogen bonds with Gly162, Gly164, Tyr224, Ser226 and Asp227.

Binding model of the ligands in AtIE binding site based on the molecular docking calculations was similar for all the (phenylureido)piperidinyl benzamides compounds. It is schematically presented for the model compound **10** ([Figure 3](#)).

The AtIE substrate binding site is quite extensive as it can accommodate the substrate with six aminosugar NAG-NAM units. Thus alternative binding modes seemed also possible for the relatively small molecule as **10**³⁰. Nevertheless, the interaction binding pattern of compound **10** was well defined. It comprised two hydrogen bonds acceptors (Gly162 and Tyr224), two hydrogen bond donors (Ser226 and Asp227), and several hydrophobic interactions with Phe63, Val137, Ile163, and Phe196 and Lys233 electrostatic cation- π interaction. Compared to the initial structure-based pharmacophore for the NAG-NAM (see [Figure 1](#)), **10** retained the binding pattern of one hydrogen bond donor (Gly162) and two hydrogen bond acceptors (Ser226 and Asp227).

We unsuccessfully attempted to perform enzymatic assays due to poor compounds solubility. In order for this assays to be successful the concentration of the inhibitor needs to be above its solubility in used assay settings. However, (phenylureido)piperidinyl benzamide are the first known series of nonsubstrate-like small molecules with confirmed binding to AtIE and are as such a good starting point for subsequent optimisation studies. To confirm the SPR results that revealed binding affinity of selected compounds in micro-molar concentration range, we have applied the STD NMR method, which provides well recognised alternative binding experiments³¹.

STD NMR binding measurements

Additionally, we tested the binding nature of selected binders with STD NMR spectroscopy. For the STD NMR experiments, we selected the most soluble compounds among the active ones, i.e. two (phenylureido)piperidinyl benzamides, **13** and **14** ([Table 2](#)). Due to overall low solubility of the compounds from this class their binding to AtIE was probed at experimental conditions for qualitative STD detection along with long selective saturation time

of AtIE. We obtained weak STD signals in the aromatic region of compounds ([Figure S3](#)), which indicates that the compounds **13** and **14** bind to AtIE.

In addition to the poor solubility of the compounds (up to $100 \mu\text{M}$), their low binding affinities ($K_d = 10 \mu\text{M}$) near the limit of detection of the method²³ may provide explanation for weak STD signals. Furthermore, due to the presence of methyl groups in ligands, we could not irradiate signals of protein in positive ppm range to improve transfer of saturation along protein.

Conclusions

In conclusion, using recently determined crystal structures of AtIE from *S. aureus* coupled with pharmacophores and molecular docking, a small series of (phenylureido)piperidinyl benzamides was discovered. Virtual screening and subsequent SPR and STD NMR binding experiments revealed that compounds interact with the AtIE in the lower micromolar range. (phenylureido)piperidinyl benzamides thus represent the first reported non-substrate like compounds that interact with this bacterial target with shown direct binding data obtained by two independent biophysical experiments. Thus, these compounds offer promise to pave the way to the first hit compounds of this cell wall degrading enzyme and enable its validation as drug target.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The authors acknowledge the financial support from the Slovenian Research Agency [research core funding No. P1-0017, P1-0012, P1-0048 and J1-8145].

References

1. Rossolini GM, Arena F, Pecile P, et al. Update on the antibiotic resistance crisis. *Curr Opin Pharmacol* 2014;18:56–60.
2. Brown ED, Wright GD. Antibacterial drug discovery in the resistance era. *Nature* 2016;529:336–343.
3. Kohanski MA, Dwyer DJ, Collins JJ. How antibiotics kill bacteria: from targets to networks. *Nat Rev Microbiol* 2010;8:423–435.
4. Mihelic M, Vlahovick-Kahlina K, Renko M, et al. The mechanism behind the selection of two different cleavage sites in nag-nam polymers. *IUCRJ* 2017;4:185–198.
5. Kulhankova K, King J, Salgado-Pabon W. Staphylococcal toxic shock syndrome: superantigen-mediated enhancement of endotoxin shock and adaptive immune suppression. *Immunol Res* 2014;59:182–187.
6. Tong SYC, Davis JS, Eichenberger E, et al. *Staphylococcus aureus* infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin Microbiol Rev* 2015;28:603–661.
7. Mancl KA, Kirsner RS, Ajdic D. Wound biofilms: lessons learned from oral biofilms. *Wound Repair Regen* 2013;21:352–362.
8. Rabin N, Zheng Y, Opoku-Temeng C, et al. Biofilm formation mechanisms and targets for developing antibiofilm agents (vol 7, pg 493, 2015). *Future Med Chem* 2015;7:493–1362.

9. Rosenstein R, Gotz F. What distinguishes highly pathogenic staphylococci from medium- and non-pathogenic? *Betw Pathogen Commensal* 2013;358:33–89.
10. Knox J, Uhlemann A-C, Lowy FD. *Staphylococcus aureus* infections: transmission within households and the community. *Trends Microbiol* 2015;23:437–444.
11. Xia JF, Gao JJ, Kokudo N, et al. Methicillin-resistant *Staphylococcus aureus* antibiotic resistance and virulence. *Biosci Trends* 2013;7:113–121.
12. Wheeler R, Turner RD, Bailey RG, et al. Bacterial cell enlargement requires control of cell wall stiffness mediated by peptidoglycan hydrolases. *Mbio* 2015;6:e00660–15.
13. Chan YGY, Frankel MB, Missiakas D, Schneewind O. SagB glucosaminidase is a determinant of *Staphylococcus aureus* glycan chain length, antibiotic susceptibility, and protein secretion. *J Bacteriol* 2016;198:1123–1136.
14. Atilano ML, Pereira PM, Vaz F, et al. Bacterial autolysins trim cell surface peptidoglycan to prevent detection by the drosophila innate immune system. *Elife* 2014;3:e02277.
15. Osipovitch DC, Therrien S, Griswold KE. Discovery of novel *S. aureus* autolysins and molecular engineering to enhance bacteriolytic activity. *Appl Microbiol Biotechnol* 2015;99:6315–6126.
16. Lombard V, Ramulu HG, Drula E, et al. The carbohydrate-active enzymes database (cazy) in 2013. *Nucleic Acids Res* 2014;42:D490–D495.
17. Borišek J, Pintar S, Ogrizek M, et al. water-assisted catalytic mechanism in glycoside hydrolases demonstrated on the *Staphylococcus aureus* autolysin e. *ACS Catal* 2018;8:4334–4345.
18. Wolber G, Langer T. Ligandscout: 3-d pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J Chem Informat Model* 2005;45:160–169.
19. Skedelj V, Perdih A, Brvar M, et al. Discovery of the first inhibitors of bacterial enzyme d-aspartate ligase from *Enterococcus faecium* (asl(fm)). *Eur J Med Chem* 2013;67:208–220.
20. Hawkins PCD, Skillman AG, Warren GL, et al. Conformer generation with omega: algorithm and validation using high quality structures from the protein databank and cambridge structural database. *J Chem Informat Model* 2010;50:572–584.
21. Verdonk ML, Cole JC, Hartshorn MJ, et al. Improved protein-ligand docking using gold. *Proteins* 2003;52:609–623.
22. Kirchmair J, Markt P, Distinto S, et al. Evaluation of the performance of 3d virtual screening protocols: Rmsd comparisons, enrichment assessments, and decoy selection - what can we learn from earlier mistakes?. *J Comput Aided Mol Design* 2008;22:213–228.
23. Mayer M, Meyer B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew Chem Int Ed Engl* 1999;38:1784–1788.
24. Helland R, Larsen RL, Finstad S, et al. Crystal structures of g-type lysozyme from atlantic cod shed new light on substrate binding and the catalytic mechanism. *Cell Mol Life Sci* 2009;66:2585–2598.
25. Kuroki R, Weaver LH, Matthews BW. A covalent enzyme-substrate intermediate with saccharide distortion in a mutant t4 lysozyme. *Science* 1993;262:2030–2033.
26. Korb O, Stutzle T, Exner TE. Empirical scoring functions for advanced protein-ligand docking with plants. *J Chem Inform Model* 2009;49:84–96.
27. Liebeschuetz JW, Cole JC, Korb O. Pose prediction and virtual screening performance of gold scoring functions in a standardized test. *J Comput Aided Mol Des* 2012;26:737–748.
28. Brus B, Kosak U, Turk S, et al. Discovery, biological evaluation, and crystal structure of a novel nanomolar selective butyrylcholinesterase inhibitor. *J Med Chem* 2014;57:8167–8179.
29. Database emolecules. <http://www.emolecules.com/>.
30. Sousa SF, Fernandes PA, Ramos MJ. Protein-ligand docking: current status and future challenges. *Proteins Structure Funct Bioinform* 2006;65:15–26.
31. Viegas A, Manso J, Nobrega FL, et al. Saturation-transfer difference (std) NMR: a simple and fast method for ligand screening and characterization of protein binding. *J Chem Educ* 2011;88:990–994.