# Modulation of the bacterial cell wall by N-acetylmuramoyl-L-alanine amidases

**Dissertation**

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Dipl. Biochem. Felix Michael Büttner

aus Kirchheim unter Teck

Tübingen

2016

*Always Look on the Bright Side of Life*

Eric Idle, Monty Python

# ABSTRACT

The bacterial cell wall is a highly dynamic structure that undergoes constant change in order to fulfill its various tasks, which range from physical protection against exterior stress and maintaining homoeostasis to immune evasion. A major component of the bacterial cell wall is the peptidoglycan network (PGN). Built up by a carbohydrate backbone of repeating units of N-acetylglucosamine and N-acetylmuramic acid linked to a peptide stem containing non-proteinogenic amino acids, PGN is a net-like structure that harbors various proteins and anchors further components of the cell wall. Depending on the composition of the peptide stems and the type of cross-linkage between the peptide stems, PGN can be a very dense network or a rather loose mesh.

N-acetylmuramoyl-L-alanine amidases cleave the amide bond between the carbohydrate backbone and the peptide stem and represent a class of PGN-modulating enzymes that ensure its plasticity. My work focused on this class of enzymes in order to better understand the mechanisms that underlie PGN cleavage, and thus its plasticity, by using biochemical and cell biological tools in combination with X-ray crystallography.

The bifunctional major autolysin AtlA of *Staphylococcus aureus* contains a glucosaminidase and an amidase, which are post-translationally processed and separated. Deletion of AtlA leads to cell clusters with irregular division patterns, indication a crucial role in cell division. I solved the atomic structure of the catalytic domain of the amidase, AmiA-cat, in complex with its substrate component muramoyltetrapeptide. Close investigation of the molecular interactions between enzyme and substrate, along with the analyses of the apo-structure and enzymatic activity assays, elucidated the likely reaction mechanism as well as substrate specificity. Since the intact substrate, including the scissile bond, is present in the complex structure, it moreover serves as a starting point for therapeutics against methicillin-resistant *Staphylococcus aureus*. Further studies with AmiA-cat in this regard involve a fragment-based screening approach using X-ray crystallography and the production and evaluation of therapeutic antibodies against AmiA-cat as possible active agents.

AmiC2 of the filamentous cyanobacterium *Nostoc punctiforme* fulfills a unique task in order to enable communication of neighboring cells within a filament. In contrast to cell-splitting amidases, AmiC2 drills holes into the septal disk that separates neighboring cells, thus generating a nanopore array used for nutrient exchange and communication. My cooperation partner located AmiC2 in the maturating septum and I solved the structure of the catalytic domain of this enzyme, AmiC2-cat. In comparison with the homologous enzyme AmiC $^{E.\ coli}$, a regulatory α-helix is missing, and AmiC2-cat exhibits high activity, which can be abolished by mutation of a catalytic glutamate. Ongoing research is focused on the mechanism that governs activity and specificity of this unusual amidase. In particular, I study the separate and / or cooperative influence of the additional domains, AMIN-A, AMIN-B, and the proline-rich linker of the AmiC2 holo-enzyme on catalysis and specificity. Furthermore, in cooperation, I am working on elucidating the exact chemical composition of *Nostoc* PGN, perhaps even differences between nascent, septal, and mature PGN. The results will be essential to generate complex structures, and elucidating potential PGN differences will provide insights into specificity.

## ZUSAMMENFASSUNG

Die Bakterienzellwand ist eine hochdynamische Struktur, die einem ständigen Wandel unterliegt, um ihre verschiedenen Aufgaben zu erfüllen. Diese reichen von physischem Schutz gegen äußere Belastungen über die Aufrechterhaltung der Zellhomöostase bis zur Immunevasion. Ein Hauptbestandteil der bakteriellen Zellwand ist das Peptidoglycan (PGN). Es ist aus einem Kohlenhydratgerüst mit abwechselnden Einheiten von N-Acetylglucosamin und N-Acetylmuraminsäure sowie einem Peptidstamm, der auch nicht-proteinogene Aminosäuren beinhaltet, aufgebaut. PGN ist eine netzartige Struktur, die außerdem verschiedene Proteine und weitere Komponenten der Zellwand verankert. Je nach Zusammensetzung des Peptidstammes selbst und der Art der Vernetzung zwischen den Peptidstämmen kann das PGN ein sehr dichtes Netz oder ein eher lockeres Geflecht sein.

N-Acetylmuramoyl-L-Alanin-Amidasen spalten die Amidbindung zwischen dem Kohlenhydratgerüst und dem Peptidstamm und stellen eine Klasse von PGN-modulierenden Enzymen dar, die seine Plastizität sicherstellen. Die vorliegende Arbeit konzentriert sich auf diese Enzymklasse und soll zum Verständnis der zugrunde liegenden Mechanismen jener enzymatischen Spaltung beitragen, die für die Plastizität von PGN verantwortlich ist. Dieser Fragestellung wurde mit Hilfe biochemischer und zellbiologischer Methoden sowie der Röntgenstrukturanalyse nachgegangen.

Das bi-funktionelle *Major Autolysin AtlA* von *Staphylococcus aureus* besteht aus einer Glucosaminidase und einer Amidase, welche posttranslational voneinander getrennt werden. Das gezielte Abschalten (*Knockout,* Nullmutante) von AtlA führt zu Zellclustern mit unregelmäßigem Teilungsmuster, was eine entscheidende Rolle bei der Zellteilung nahelegt. Ich habe die atomare Struktur der katalytischen Domäne der Amidase, AmiA-cat, im Komplex mit ihrem Substratbestandteil Muramoyltetrapeptid gelöst. Sowohl die genaue Untersuchung der molekularen Wechselwirkungen zwischen Enzym und Substrat sowie die Analyse der apo-Struktur als auch enzymatische Aktivitätstests haben Anhaltspunkte für den wahrscheinlichen Reaktionsmechanismus sowie die Substratspezifität des Enzyms geliefert. Da das intakte Substrat einschließlich der zu spaltenden Bindung in der Komplexstruktur vorhanden ist, dient sie ferner als ein guter Startpunkt für Therapeutika gegen den Methicillin-resistenten *Staphylococcus aureus*. Weitere Studien mit AmiA-cat in diese Richtung beinhalten neben

einem fragmentbasierten Screeningansatz unter Zuhilfenahme von Röntgenkristallographie auch die Produktion und Tests von therapeutischen Antikörpern gegen AmiA-cat als mögliche Wirkstoffe.

AmiC2 des filamentösen Cyanobakteriums *Nostoc punctiforme* führt eine einzigartige Reaktion aus, um die Kommunikation von benachbarten Zellen innerhalb eines Filaments zu ermöglichen. Im Gegensatz zu zellspaltenden Amidasen bohrt AmiC2 Löcher in das Septum, welches Nachbarzellen voneinander trennt. Dadurch entsteht ein *Nanopore Array*, das für Nährstoffaustausch und Kommunikation verwendet wird. Meine Kooperationspartner haben AmiC2 im ausreifenden Septum lokalisiert, und ich habe die Struktur der katalytischen Domäne dieses Enzyms gelöst (AmiC2-cat). Interessanterweise fehlt eine regulatorische α-Helix, wie man sie in dem homologen Enzym AmiC $^{E.\ coli}$ findet. AmiC2-cat ist katalytisch sehr aktiv, was durch die Mutation eines katalytischen Glutamats aber aufgehoben werden kann. Weitergehende Forschung zielt auf die Aufklärung des Mechanismus ab, der die Aktivität und Spezifität dieser ungewöhnlichen Amidase regelt. Insbesondere wird momentan der getrennte und / oder kooperative Einfluss der zusätzlichen Domänen des AmiC2-Holoenzyms - AMIN-A, AMIN-B sowie Prolin-reicher Linker - auf die Katalyse und Spezifität von AmiC2-cat erforscht. Außerdem wird die genaue chemische Zusammensetzung des PGN von *Nostoc* untersucht, um den physiologischen Liganden von AmiC2 für eine Komplexstruktur zu identifizieren. Weiterhin könnten eventuelle Unterschiede zwischen jungem, septalem und reifem PGN eine Rolle bei der enzymatischen Spezifität spielen.

# TABLE OF CONTENTS

# ABBREVIATIONS

| | |
|---|---|
| AmiA | an amidase of *Staphylococcus aureus*, part of AtlA |
| AmiA-cat | catalytic domain of AmiA |
| AmiC | an amidase of *Escherichia coli* |
| AmiC2 | an amidase of *Nostoc punctiforme* |
| AmiC2-cat | catalytic domain AmiC2 |
| AmiD | an amidase of *Escherichia coli* |
| AmiE | an amidase of *Staphylococcus epidermidis* |
| AMIN A | amino-terminal domain A of AmiC2 |
| AMIN AB | amino-terminal domains A and B of AmiC2 |
| AMIN B | amino-terminal domain B of AmiC2 |
| AmpD | an amidase of *Citrobacter freundii* |
| AmpDH2 | an amidase of *Pseudomonas aeruginosa* |
| AmpDH3 | an amidase of *Pseudomonas aeruginosa* |
| AtlA | bifunctional major autolysin of *Staphylococcus aureus* |
| AtlE | bifunctional major autolysin of *Staphylococcus epidermidis* |
| AtlL | bifunctional major autolysin of *Staphylococcus lugdunensis* |
| CA-MRSA | community-associated MRSA |
| CDC | Centers for Disease Control and Prevention |
| CNS | coagulase negative staphylococci |
| CPS | coagulase positive staphylococci |
| D- / L-Ala | D- / L-alanine |
| D- / L-Ser | D- / L-serine |
| D-iGln | D-isoglutamine |
| D-iGlu | D-isoglutamic acid |
| D-iGlx | D-isoglutamine or glutamic acid |
| D-Lac | D-lactate |
| DRA | dye release assay |
| EC | enzyme commission number |
| EPS | exopolysaccharide |
| GlcNAc | N-acetylglucosamine |
| HA-MRSA | hospital-acquired MRSA |
| IM | inner membrane |
| LA-MRSA | livestock-associated MRSA |
| LPS | lipopolysaccharides |
| LTA | lipoteichoic acid |
| LytA | an amidase of *Streptococcus pneumoniae* |
| *m*DAP | *meso*-diaminopimelic acid |
| MES | 2-(*N*-morpholino)ethanesulfonic acid |

| | |
|---|---|
| MGE | mobile genetic elements |
| MRSA | methicillin-resistant *Staphylococcus aureus* |
| MtetP | muramoyl tetrapeptide (MurNAc-L-Ala-D-iGln-L-Lys(NHAc)-D-Ala-NH$_2$) |
| MurNAc | N-acetylmuramic acid |
| NAGase | glucosaminidase of *Staphylococcus aureus*, part of AtlA |
| OM | outer membrane |
| PBP | penicillin binding protein |
| PDB-ID | deposition code for structures in the Protein Data Bank |
| PGN | peptidoglycan network |
| PGRP | peptidoglycan recognition protein |
| ProLinker | proline-rich linker domain of AmiC2 |
| ProLinker-cat | construct of AmiC2 comprising ProLinker and AmiC2-cat |
| PVL | Panton–Valentine leukocidin |
| rmsd | root-mean-square deviation |
| ROS | reactive oxygen species |
| Rv3717 | an amidase of *Mycobacterium tuberculosis* |
| SAXS | small-angle X-ray scattering |
| SCC | staphylococcal cassette chromosome |
| T6SS | bacterial Type Six Secretion System |
| UPLC / MS | ultra-performance liquid chromatography with mass spectrometry |
| VISA | vancomycin-intermediate sensitive *Staphylococcus aureus* |
| VRSA | vancomycin-resistant *Staphylococcus aureus* |
| VSSA | vancomycin-sensitive *Staphylococcus aureus* |
| wt | wild-type |
| WTA | wall teichoic acid |
| wwPDB | worldwide Protein Data Bank |

Abbreviations used in chapter 6 in order of appearance:

| | |
|---|---|
| JNK3 | c-Jun N-terminal kinase 3 |
| MAPK | mitogen-activated protein kinase |
| QM | quantum mechanical |
| FP | Fluorescence Polarization |
| ITC | Isothermal Titration Calorimetry |

Abbreviations used in the Appendix are listed in the individual publications.

For clarity, individual amidases may be labelled with the organism of origin in the main text.

The common abbreviations of units of measurement as well as proteinogenic amino acids are used.

# CONTRIBUTIONS TO PUBLICATIONS

## Journal articles

### Structure-Function Analysis of *Staphylococcus aureus* Amidase Reveals the Determinants of Peptidoglycan Recognition and Cleavage

I generated active site mutants for this work and purified wild-type as well as mutant enzyme for activity measurements and crystallization trials. I crystallized AmiA-cat, prepared a ligand-soaked crystal, and solved its complex structure. Structure refinement and analysis were performed by me. Also, I carried out structural and sequence-based comparisons to further elucidate the enzymatic function. Figures 1-5, 7-8, and table 1 were prepared by me, and I wrote the bulk of the manuscript, together with Friedrich Götz and Thilo Stehle.

### Enabling cell-cell communication via nanopore formation: structure, function, and localization of the unique cell wall amidase AmiC2 of *Nostoc punctiforme*

I purified AmiC2-cat for activity assays and crystallization. I solved, refined, as well as analyzed its structure. Furthermore, active site mutants were designed and generated by me. I also purified them and checked for proper folding with analytical size exclusion chromatography and circular dichroism spectroscopy. Structural and sequence-based comparisons were performed by me. I prepared all figures and tables, except for data of figures 1, 2, and 5b. I wrote the bulk of the manuscript, together with Katharina Faulhaber, Karl Forchhammer, Iris Maldener, and Thilo Stehle.

**X-ray crystallography and its impact on understanding bacterial cell wall remodeling processes**

For this review, I prepared figures 2 and 3 and wrote the section "*A case study of the impact of X-ray crystallography: The cell wall remodeling enzyme class of N-acetylmuramoyl-L-alanine amidases*". Together with Michaela Renner-Schneck and Thilo Stehle, I developed and proofread the manuscript.

**Targeting the Gatekeeper MET146 of C-Jun N-terminal Kinase3 (JNK3) Induces a Bivalent Halogen / Chalcogen Bond**

I obtained crystallographic data of a JNK3 complex crystal produced by Andreas Lange. The structure was solved, refined, and interpreted by me, Andreas Lange, Thilo Stehle, and Frank M. Boeckler. I wrote part of the section "4. Crystal Structure of 7d in Complex with JNK3 (4X21) and its Comparison to 2P33".

## Protein structures

PDB-ID 4KNK   –   AmiA-cat: I crystallized the protein and solved the structure.

PDB-ID 4KNL   –   AmiA-cat in complex with its ligand MtetP: I crystallized the protein, soaked the crystal, and solved the complex structure.

PDB-ID 4X21   –   JNK3 in complex with an iodine ligand: I measured the crystallographic data and was involved in solving the complex structure.

PDB-ID 5EMI   –   AmiC2-cat: I crystallized the protein and solved the structure.

# 1. INTRODUCTION

## 1.1. Bacteria in our environment

The vast majority of biomass on earth is made up by microorganisms and, in fact, for three billion years they represented the only living beings. A major representative, and one of the three domains of life, are bacteria [1]. Microorganisms were first observed in 1665 by Robert Hooke, who saw mold through his microscope and, shortly after, Antoni van Leeuwenhoek discovered bacteria in 1676 (published in 1684) [1]. In the nineteenth century, three important individuals laid the basis for what we call bacteriology and microbiology. To list only highlights, Ferdinand Cohn defined bacteriology and discovered bacterial endospores, Louis Pasteur invented the sterilization process ("*pasteurization*") and developed attenuated vaccines, and Robert Koch established methods for culturing pure microbial cultures and linked microorganisms to infectious diseases [1]. Therefore, most people associate bacteria with disease and see them as pathogens, however, this is only part of the truth.

When cyanobacteria evolved into oxygenic phototrophs 3.3 billion years ago, they started oxygenizing Earth and laid a basis for "higher" life [1, 2]. Respiration of oxygen provides more energy, and the formation of the ozone layer protects land-life from UV radiation [2-4]. Furthermore, according to the endosymbiotic theory, all eukaryotic cells harbor organelles of bacterial origin [5]. Bacteria are also true symbiotic partners for multicellular organisms including humans [6]. Giving one example, the microbiome in our gut helps with digestion, provides vitamins or other essential compounds, and competes with pathogenic microorganisms [1, 6]. Imbalances in the microbiome can be linked to a variety of health problems, including obesity and diabetes (e.g. reviewed in [7, 8]).

One could list countless other examples to highlight the biological importance of bacteria for the environment (e.g. decomposition processes) but also for their economic use (e.g. food and beverage industries), which is even increasing due to modern microbiological methods and techniques [1].

It is thus not only important to find out how we can contain pathogenic bacteria, but we need to seek out how bacteria are able to perform the variety of sometimes incredible tasks, too. This is reflected in this thesis with respect to the close investigation of enzymes from the pathogen *S. aureus* as well as the filamentous cyanobacterium *N. punctiforme*.

## 1.2. Staphylococci

### 1.2.1. The genus *Staphylococcus*

The genus *Staphylococcus* was first identified by Sir Alexander Ogston in 1881 [9] and comprises more than 40 species and subspecies. He found two forms of "micrococci" in pus from abscesses: the chain-forming *Streptococcus* (Greek: στρεπτός, streptos, meaning easily bent or twisted, like a chain and κόκκος, kókkos, for "granule") and a group of cocci that look like grapes or "the roe of fish" [9] (Figure 1). They were later termed *Staphylococcus* (Greek: σταφυλή, staphylē, for "grape" and again κόκκος, kókkos, meaning "granule"). Staphylococci are Gram-positive bacteria that colonize the skin or mucous membranes of the upper respiratory tract [10, 11]. Most members of this facultatively anaerobic genus produce catalase but are oxidase-negative. Staphylococci withstand harsh environments such as drought and high salt concentrations [10, 11].
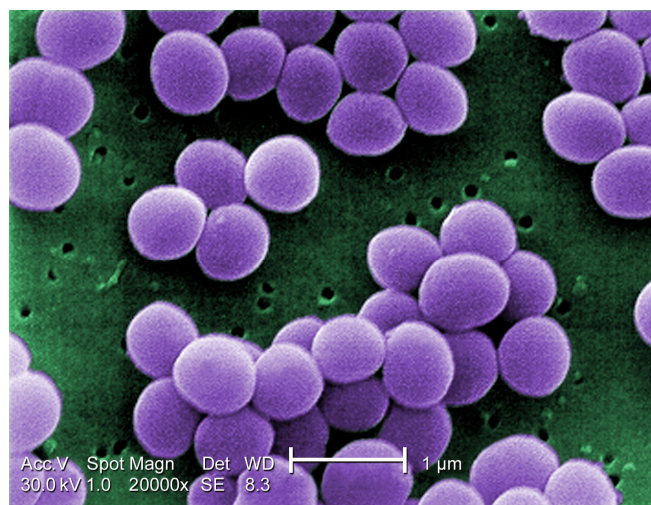


Figure 1 - *Staphylococcus aureus* in a colored electron microscope image with typical grape-like arrangement. The picture is free of any copyright restrictions and was taken from the "Public Health Image Library" (PHIL). Photo Credit: Janice Haney Carr; Content Providers(s): CDC/ Matthew J. Arduino, DRPH.

The two most abundant human pathogens of the Staphylococcus genus are *S. aureus* and *S. epidermidis*. Both species produce an extracellular polymeric substance, called biofilm, for protection against the host immune system [12] and antibiotics. The expression of coagulase distinguishes coagulase positive staphylococci (CPS) such as *S. aureus* and coagulase negative staphylococci (CNS) such as *S. epidermidis*, which are considered less pathogenic. Coagulase processes fibrinogen to fibrin, thereby inducing blood clotting [13] and thus preventing the host immune system from efficient counteraction [10].

### 1.2.2. *Staphylococcus aureus*

*Staphylococcus aureus* produces the carotenoid pigments *staphyloxanthin* that gave it its name (Latin: aureus, for "golden"). Staphyloxanthin has antioxidant capabilities, thereby enabling *S. aureus* to cope with oxidative stress from reactive oxygen species (ROS) and making it more resistant to neutrophil killing [14].

Although harmless to healthy individuals, *S. aureus* can cause severe infections in immunocompromised individuals. Apart from the immune-evading function, the biofilm also contributes to persistent infections, which ultimately lead to complications [15, 16]. The multilayered polysaccharide matrix [17] also contributes to pathogenicity. Still, the primary function is the protection from the host's immune system [18].

*S. aureus* colonizes roughly one third of the population asymptomatically as a commensal bacterium [19]. However, *S. aureus* is responsible for a broad variety of diseases that can be life-threatening. Mild skin infections such as abscesses or lesions producing pus, as well as food poisoning, endocarditis, meningitis, pneumonia, septicemia, and toxic shock syndrome can be caused by *S. aureus* [10, 20, 21]. The bacteria secrete several enzymes and toxins, e.g. coagulase, but also hyaluronidase, fibrinolysin, lipase, ribonuclease, deoxyribonuclease, Protein A, Panton–Valentine leukocidin (PVL), or toxic shock syndrome toxin I [10]. These enzymes facilitate destruction of host tissue and promote the spread of infection, while Protein A binds to immunoglobulins and helps to evade opsonization and phagocytosis. PVL introduces pores into leukocytes and thereby lyses these immune cells. The toxic shock syndrome toxin I acts as a superantigen, recruits a large number of T-cells and, in so doing, causes a major inflammatory response with a fatality rate of 70% [10, 20].

The adaptive and pathogenic variety of *S. aureus* is achieved via mobile genetic elements (MGE) that make up 15-20 % of its genome and carry resistance-genes, host-adaption mechanisms, or encode for toxins [22]. MGEs include staphylococcal cassette chromosomes (SCC), plasmids, transposons, or *S. aureus* pathogenicity islands, and are mostly transferred between *S. aureus* cells via phage-mediated generalized transduction. Conjugation or transformation are less frequent but also possible and moreover allow gene acquisition from other genera [22].

## 1.2.3. Antibiotic resistance in Staphylococci

Most antibiotics are natural products taken directly or derived from bacteria (or fungi), which use them to defend their ecological niche against adversaries. When Alexander Fleming discovered penicillin in 1929, mankind, too, started to utilize those compounds to treat bacterial infections [23]. However, most antibiotics require a resistance mechanism that protects the producing organism form it. Due to horizontal gene transfer, the first resistances against the "miracle drug" emerged even before mass production of penicillin started [24, 25].

Drug resistance mechanisms of bacteria are [26] (Figure 2):

- drug inactivation or modification
- alteration of the target site
- reduced drug accumulation via lower intake or higher efflux
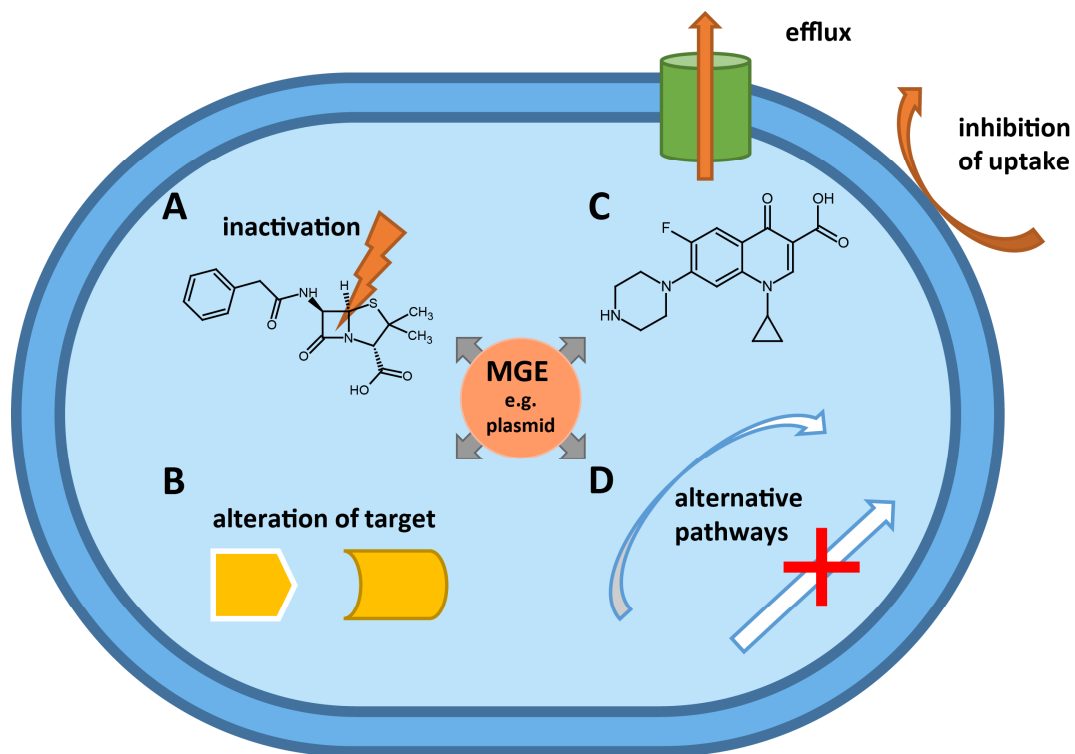- alteration of metabolic pathways



Figure 2 - Mechanisms of resistance against antibiotics used by bacteria: (A) drug inactivation or modification; (B) alteration of target site; (C) reduced drug accumulation via lower intake or higher efflux; (D) alteration of metabolic pathways. Resistances often arise by horizontal gene transfer via mobile gene elements (MGE), e.g. by the acquisition of plasmids.

The increased usage of penicillin further promoted the acquisition of resistance genes due to selection pressure. Although chemically modified and sterically enhanced β-lactam antibiotics were introduced, the bacteria responded, and first strains of methicillin-resistant *Staphylococcus aureus* (MRSA) were discovered in the early 1960s [27-29].

MRSA strains are resistant to all β-lactam antibiotics because of a special penicillin binding protein (PBP) that allows transpeptidation in the presence of e.g. methicillin [30]. This is achieved by the integration of a SCCmec gene cluster (staphylococcal cassette chromosome mec) into their genome, carrying the *mecA* gene [10]. Eight different SCCmecs are known so far [31]. Kos et al. found that *"CC5 strains were identified among early methicillin-resistant isolates in the 1960s (…) and were shown to have acquired staphylococcal cassette chromosome (SCCmec) at least 23 separate times"* [32]. The constant and high selection pressure in healthcare facilities from a broad variety of antibiotics forces MRSA strains to adapt and evolve and, consequently, to acquire new antibiotic resistances against several classes of antibiotics (e.g. gyrase or bacterial RNA polymerase inhibitors and glycopeptide based antibiotics) [32-34], posing a serious threat to human health [31]. This is reflected in rising numbers of multi-resistant strains, e.g. the vancomycin sensitive, intermediate-sensitive, and resistant *S. aureus* (VSSA, VISA, VRSA, respectively).

In the US, over 80,000 cases of MRSA infections lead to 11,000 casualties per year as indicated by the Centers for Disease Control and Prevention (CDC) [35]. According to the *Robert Koch Institut*, reported case numbers in Germany decreased in the period of 2012 to 2014 from 4,498 to 3,841 cases [36]. However, with an actual prevalence of around 10,000 annual MRSA infections, the numbers are still concerning [37]. Higher hygienic standards may play a role in decreasing numbers of hospital-acquired MRSA (HA-MRSA) [38] as intensive screening of patients and staff as well as consequent treatment have kept the MRSA infection rate in the Netherlands under 3% [39]. Nevertheless, community-associated MRSA (CA-MRSA) and livestock-associated MRSA (LA-MRSA) infections have been increasing and spreading significantly over the last years [27, 31, 36, 40]. Furthermore, LA-MRSA gain resistances against more antibiotic drug classes that are used in animal-fattening farms and also acquire further virulence factors [31]. Numerous novel combinations of new resistances with so far non-human virulence factors arise and threaten our well-being.

## 1.3.   Cyanobacteria of the genus *Nostoc*

### 1.3.1.   *Nostoc punctiforme* ATCC 29133

Cyanobacteria set the basis for higher life on earth by introducing oxygen into the atmosphere through oxygenic photosynthesis [2-4]. The taxonomy of cyanobacteria is controversial. One method of classification is the division into five groups based on different characteristics and properties that allow the adaption to a broad variety of habitats [41]. The group of *Nostocales* comprises filamentous bacteria that have the ability to fix nitrogen (Figure 3). Moreover, *Nostocales* are true multicellular organisms and able to differentiate, which is usually thought to be an exclusively eukaryotic capability [42] (Figure 3). A common outer membrane surrounds a filament of cells and they subsequently share cell wall as well as periplasm [43-48].
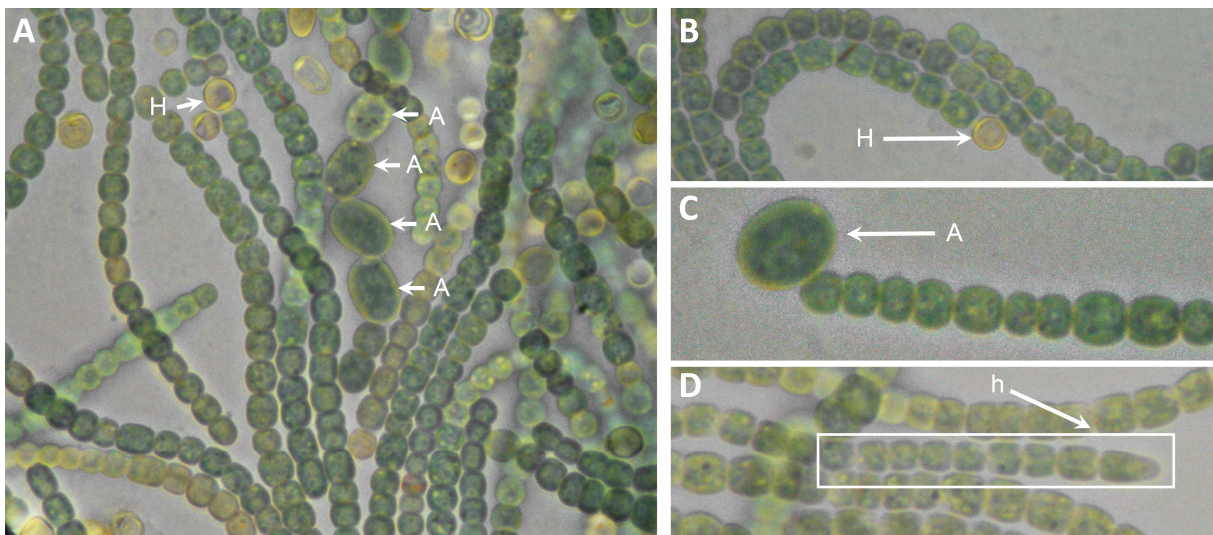


Figure 3 - The filamentous cyanobacterium *Nostoc punctiforme*. (A) A population of *Nostoc punctiforme* with vegetative cells in several filaments. The larger akinetes are indicated by an arrow and a white "A"; nitrogen fixing heterocysts are indicated by a "H", respectively, and can also be identified by less green chlorophyll a as they are not able to perform oxygenic photosynthesis. (B) One heterocyst positioned at the end of a filament of vegetative cells. (C) An akinete is shown at the left edge of a filament. (D) Hormogonia, highlighted by a white box and indicated with a "h" are smaller and shaped differently compared to vegetative cells, especially the terminal cell of a filament.
*The microscopic images were provided by Katharina Faulhaber.*

### 1.3.2.   Multicellularity and differentiation

Specialization and division of labor can significantly increase the productivity of a population. In case of most organisms, this translates to proliferation and adaptation to the environment. The large genome of *Nostoc punctiforme* [49] is beneficial in this endeavor.

Vegetative cells of the filament are able to differentiate into akinetes and hormogonia as well as heterocysts (reviewed in [50, 51], Figure 3 and Figure 4). Akinetes are spore-like cells that persist even under extreme conditions and germinate into a new filament in a better environment [52-54] (Figure 3 A+C). Hormogonia are short gliding filaments that are formed upon stress signals and able to engage plants in order to start a symbiosis [51, 54-56] (Figure 3 D).
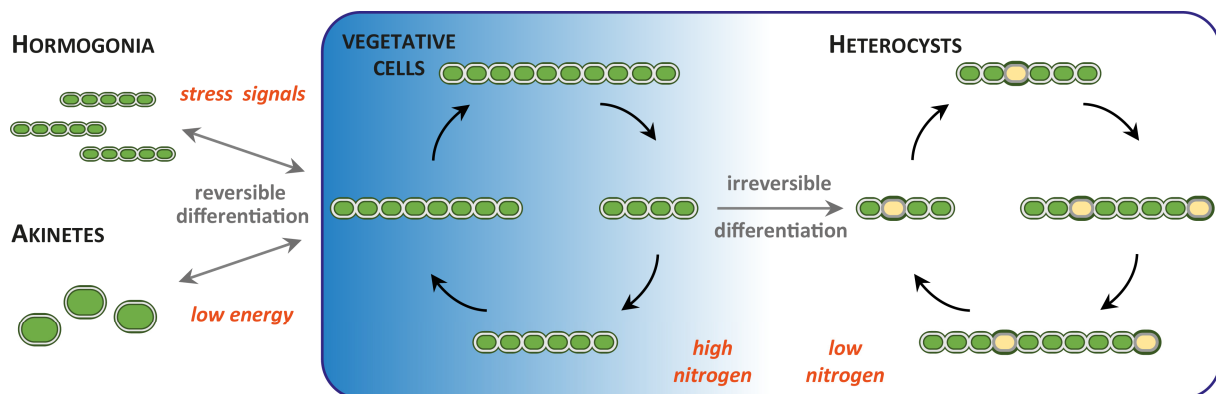


Figure 4 – Overview of proliferation and differentiation of *Nostoc punctiforme*. Vegetative cells proliferate (black arrows) and extend the filaments under beneficial conditions. Vegetative cells differentiate (grey arrows) to hormogonia in order to start symbioses with plants, while akinetes are formed under energy deprivation and can give rise to new filaments once better conditions arise. Finally, the differentiation into heterocysts upon low levels of fixed nitrogen is irreversible. The heterocysts provide fixed nitrogen for vegetative cells so that they can proliferate. Based on [50, 54].

At low levels of fixed nitrogen, *Nostoc* filaments form heterocysts at termini or along the filament in a semi-regular pattern (Figure 3 A+B and Figure 4). Differentiation into heterocysts is irreversible and linked to severe changes in order to perform nitrogen fixation [54]. As oxygen is harmful to nitrogenase and inhibits $N_2$-fixation, heterocysts form a microoxic environment by formation of a special cell envelope on top of the cell wall to prevent $O_2$ diffusion into the cell. They are also no longer able to perform oxygenic photosynthesis because they lack photosystem II [41]. As a result, heterocysts need carbon sources for the production of ATP and, equally importantly, as electron donors to reduce nitrogen [41]. The vegetative cells and heterocysts are therefore interdependent, as the heterocysts provide reduced nitrogen but need carbon sources and vice versa [41, 52].

## 1.3.3.   Cellular communication

The exchange of differentiation- and other signals as well as nutrients between cells can either take place by the slow diffusion process through the periplasm [47], or via a directed cell-cell

connection. Such connections were termed *microplasmodesmata*, *septosomes*, *channels*, or *septal junctions* [48, 57-61] and probably traverse the periplasmic space [62] (Figure 5 B+C). To do so, the channels also need to cross the septal PGN. Indeed, in the so-called nanopore array the septal disk is perforated by approximately 150 pores with a diameter of 20 nm each [44, 58, 63] (Figure 5 A). It was shown that these nanopores are essential for cell-cell communication and that N-acetylmuramoyl-L-alanine amidases are involved in their formation [43, 44, 58, 64]. Mutational studies revealed a pleiotropic phenotype, where irregular septal planes lead to distorted filaments. Furthermore, the transfer of fluorescent dye is no longer observed, the specialization of cells is abolished, and diazotrophic growth is therefore not possible [43] (Figure 5 D).
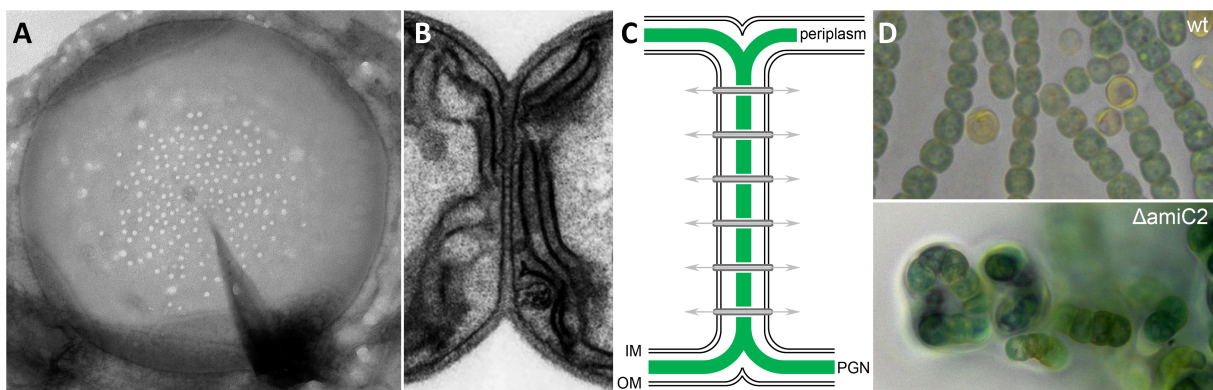


Figure 5 - Nanopore array, septal junctions, wt and ΔamiC2 *N. punctiforme*. (A) View onto the septal disk reveals the nanopore array with ~ 150 pores of ~ 20 nm diameter. (B) Side-view on the septum with septal junctions. (C) Schematic representation of nanopores and septal junctions traversing the septa. (D) ΔamiC2 *N. punctiforme* showing an impaired phenotype compared to wt. Figure is based on [43, 44].
*Pictures were provided by Iris Maldener (A+B, credit goes to Josef Lehner; Lehner, Maldener, and Forchhammer, unpublished results) and by Katharina Faulhaber (D).*

## 1.4.  Bacterial cell wall and peptidoglycan

### 1.4.1.  The cell wall and its function

Bacteria are surrounded by the cell wall, a protective coat that prevents the cells from bursting due to their high osmotic inner pressure. It also defines shape and rigidity of the cells and shields them from outside influences [65]. Bacterial cell walls are divided into two classes based on their response to Gram staining with crystal violet and safranin [66]. Gram-positive cell walls consist of one lipid bilayer (bilayers are approximately 4-5 nm thick) and a thick PGN of 20-80 nm with up to 80 sublayers. Teichoic acids and associated proteins are also found in the Gram-positive cell wall. Wall teichoic acids (WTA) are linked to the PGN, while lipoteichoic acids (LTA) are anchored in the plasma membrane [65] (Figure 6).
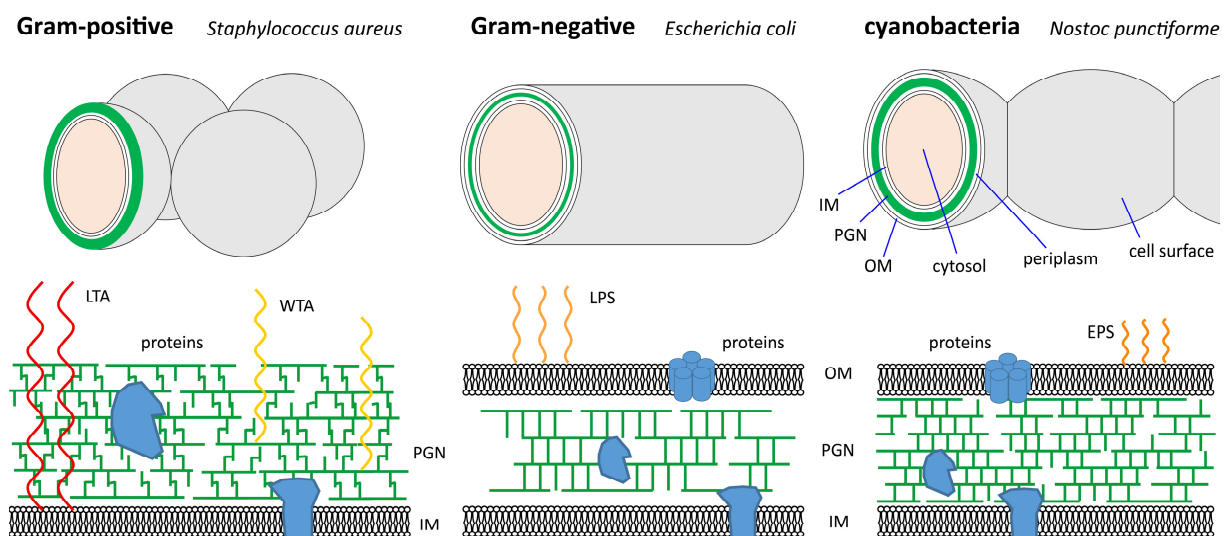


Figure 6 - Scheme of bacterial cell walls of different classes with respective representatives. The Gram-positive cell wall is exemplarily shown for *S. aureus*, *E. coli* represents Gram-negative bacteria and *N. punctiforme* covers the special case of cyanobacteria. Schematic cells display differences in cellular shape and organisation, in thicknesses of PGN (in green), and in the presence or absence of an outer membrane (OM). A close-up on the cell wall shows the individual PGN pattern and additional components of respective cell walls like lipo- and wall teichoic acids (LTA + WTA), membrane and periplasmic proteins, lipopolysaccharides (LPS), and exopolysaccharides (EPS). IM - inner membrane.

Gram-negative bacteria have an inner and an outer membrane with a thin PGN-layer of 2.5-7 nm separating the two membranes [67]. Again, associated proteins are located in the cytoplasmatic membrane and the 15 nm wide periplasmic space [65] (Figure 6). The outer membrane additionally features porins, lipoproteins, as well as lipopolysaccharides (LPS). LPS consist of the anchor Lipid A, the core saccharide, and diverse O-polysaccharides [65].

The cyanobacterial cell wall is considered Gram-negative because of its outer membrane, however, the PGN layer is significantly thicker than usual for Gram-negative bacteria [48, 68-71] (Figure 6). Furthermore, some filamentous cyanobacteria share a common sacculus along the whole filament [43-48].

## 1.4.2. Peptidoglycan

The bacterial peptidoglycan (PGN) has the general composition of a carbohydrate backbone and a peptide stem (Figure 7 and Figure 8). The carbohydrate backbone consists of repeating units of β 1→4 linked N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc). The peptide stem is generally composed of L-Ala, D-iGlu, L-Lys or *mDAP*, and D-Ala-D-Ala, where the amino acids in position three (L-Lys or *m*DAP) and in position four (D-Ala) cross-link the peptide stems (3 → 4) either directly or via a peptide bridge [65, 72-74] (Figure 7 and Figure 8).
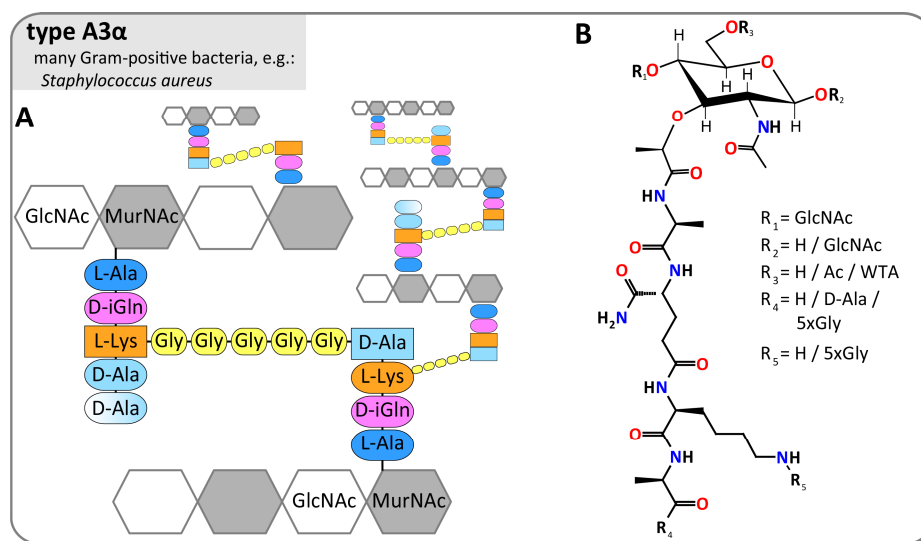


Figure 7 - Schematic representation of the bacterial PGN of type A3α. (A) A3α-PGN (Gram-positive) is highly crosslinked via a pentapeptide linker and usually contains L-Lys and D-iGln. A second D-Ala is sporadically present. (B) The molecular structure of staphylococcal muramoyltetrapeptide (MtetP, MurNAc-L-Ala-D-iGln-L-Lys-D-Ala) with different substituents depending on the degree of cross-linkage and modification. Only the most common linkages are shown.

Schleifer and Kandler summarized the knowledge on PGN in their review in 1972 [74], which still is our basic notion today. They report, however, that the actual PGN composition is more complex. Individual bacteria implement a variety of alterations, modifications, and substitutions in order to adapt their cell wall (Table 1). This is also reflected in the annotation of PGN-subtypes like "A3α" for *S. aureus* or "A1γ" for e.g. *E. coli*, *B. subtilis,* and *N. punctiforme* [74, 75] (Figure 7 and Figure 8). Most Gram-negative bacteria use iGlu in position two and

*mDAP* in position three (Figure 8), whereas Gram-positive bacteria amidate iGlu to iGln and incorporate L-Lys instead of *m*DAP (Figure 7). Another alteration is the exchange of the terminal D-Ala with D-Ser or D-lactate (D-Lac), which correlates with certain antibiotic resistances. An overview on amino acid usage is given in Table 1 [72]. Furthermore, Gram-negative PGN is usually directly cross-linked, while a pentaglycine bridge, which also may contain other amino acids, connects Gram-positive peptide stems [74, 76]. Other variables include the occasional absence of unlinked, terminal D-Ala residues, different amidation patterns of carboxyl groups (e.g. from *m*DAP), or deviations from the 3 → 4 cross-linkage pattern [74].
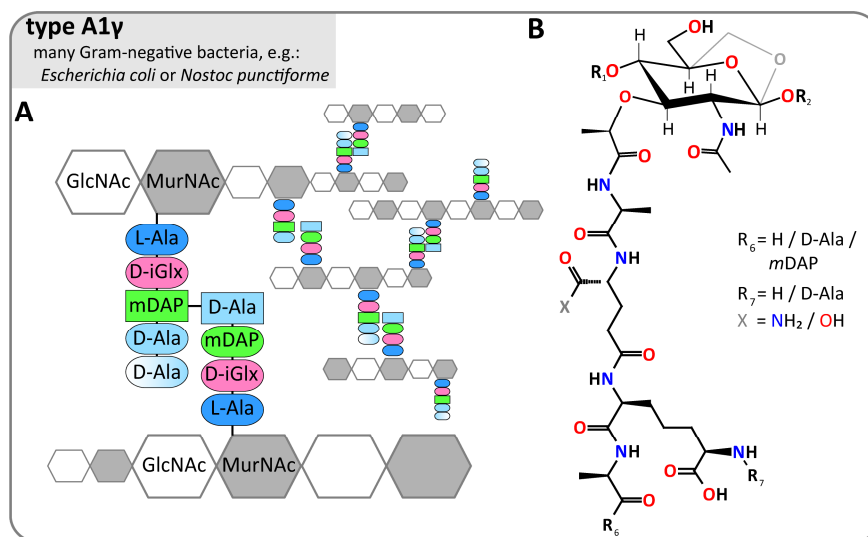


Figure 8 - Schematic representation of the bacterial PGN of type A1γ. (A) A1γ-PGN (Gram-negative) is often less dense and utilizes *m*DAP. D-iGlu may be amidated to D-iGln as indicated by D-iGl<u>x</u> and peptide stems are directly cross-linked. (B) MtetP of A1γ-PGN (MurNAc-L-Ala-D-iGlx-*m*DAP-D-Ala). Terminal MurNAc may present in the non-reducing 1,6-anhydro form (grey). Otherwise, as Figure 7.

The carbohydrate backbone is also highly diverse and varies in length from 3 to over 100 subunits. However, the common ranges are 10 - 40 carbohydrate rings in Gram-negative and 3 - 10 in Gram-positive bacteria [67, 77]. The reducing end of a carbohydrate chain is typically a MurNAc that is present in the 1,6-anhydro form in most organisms, featuring a second ring structure [67] (Figure 8 B). Nevertheless, *S. aureus*' PGN has reducing MurNAc, not the 1,6-anhydro species [77]. Furthermore, WTA can be linked to MurNAc, and the common MurNAc O6-acetylation mediates tolerance towards lysozyme [78].

| POSITION | AMINO ACID | EXAMPLARY SPECIES |
|:---:|---|---|
| **1** | L-Ala | most *species* |
| | Gly | *Mycobacterium leprae, Brevibacterium imperiale* |
| | L-Ser | *Butyribacterium rettgeri* |
| **2** | D-isoglutamate | most Gram-negative species |
| | D-isoglutamine | most Gram-positive species*, Mycobacteria* |
| | threo-3-hydroxyglutamate | *Microbacterium lacticum* |
| **3** | *m*DAP | most Gram-negative species*, Bacilli, Mycobacteria* |
| | L-Lys | most Gram-positive species |
| | L-ornithine | *Spirochetes, Thermus thermophilus* |
| | L-Lys/L-ornithine | *Bifidobacterium globosum* |
| | L-Lys/D-Lys | *Thermotoga maritima* |
| | LL-DAP | *Streptomyces albus, Propionibacterium petersonii* |
| | *meso*-lanthionine | *Fusobacterium nucleatum* |
| | L-2,4-diaminobutyrate | *Corynebacterium aquaticum* |
| | L-homoserine | *Corynebacterium poinsettiae* |
| | L-Ala | *Erysipelothrix rhusiopathiae* |
| | L-Glu | *Arthrobacter J. 39* |
| | amidated *m*DAP | *Bacillus subtilis* |
| | 2,6-diamino-3-hydroxypimelate | *Ampuraliella regularis* |
| | L-5-hydroxylysine | *Streptococcus pyogenesz* |
| | Nγ-acetyl-L-2,4-diaminobutyrate | *Corynebacterium insidiosum* |
| **4** | D-Ala | all bacteria |
| **5** | D-Ala | most bacteria |
| | D-Ser | *Enterococcus gallinarum* |
| | D-Lac | *Lactobacillus casei,* Enterococci with acquired resistance to vancomycin |

Table 1 - Variations of amino acid usage in bacterial PGN; adapted from [72].

## 1.5.  Cell wall modulation by N-acetylmuramoyl-L-alanine amidases

### 1.5.1.  Different classes of cell wall modulating enzymes

In order to fulfill its functions, the cell wall and thus PGN also need to be highly dynamic and undergo constant adjustments or remodeling to ensure growth and survival. This plasticity is achieved by the continuous synthesis, degradation, and recycling of components by a broad range of enzymes. Figure 9 presents the cleavage sites of such lytic enzymes that may target almost every bond of PGN and were in the focus of several reviews [79-81]. The different enzyme classes, their tasks, and some respective representative examples are given below:

- Endopeptidases cleave peptide bonds at different locations of the peptide stem or the interpeptide bridge. They may have diverse stereospecificities towards D‑D, D‑L, or L‑D peptide bonds, e.g. lysostaphin [82].
- Carboxypeptidases cleave peptide bonds to release terminal amino acids, e.g. PBP5 [83].
- Muramidases cleave the β-1→4 glycosidic bond between MurNAc and GlcNAc.
    - Lysozymes result in reducing MurNAc, e.g. Lysozyme [84].
    - lytic transglycosylases produce 1,6-anhydro MurNAc [85], e.g. MltB [86].
- Glucosaminidases cleave the β-1→4 bond of GlcNAc and MurNAc, e.g. NAGase [77, 87].
- N-acetylmuramoyl-L-alanine amidases cleave the amide bond between N-acetylmuramic acid and L-alanine and are discussed in more detail in the following section.
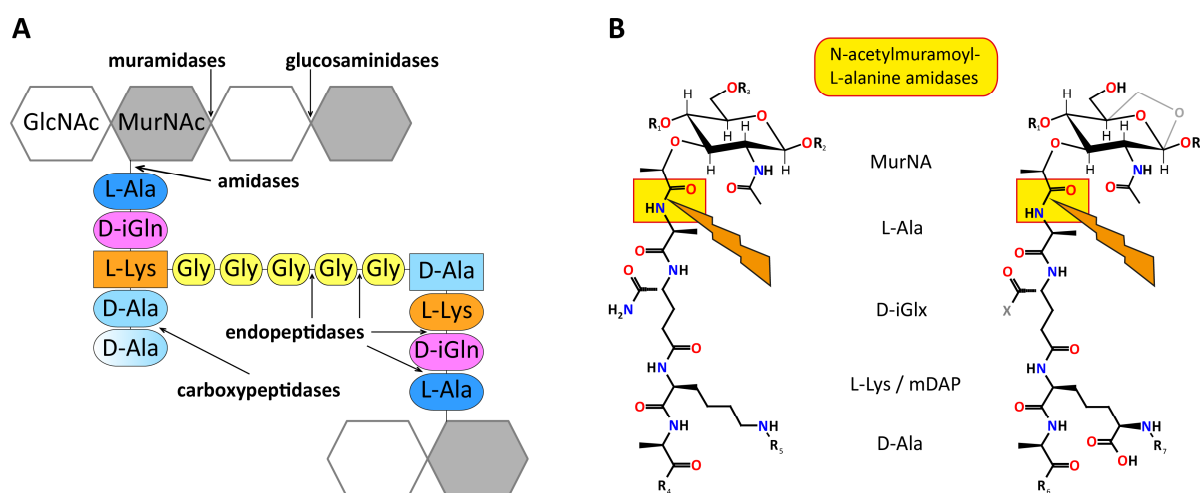


Figure 9 – (A) Several hydrolytic enzymes process PGN in order to keep it dynamic. Arrows indicate the cleavage sites of the different enzyme classes. (B) The cleavage site of amidases shown in more detail for PGN A3α and A1γ, respectively. Residues "R" as in Figures 7 and 8.

## 1.5.2. N-acetylmuramoyl-L-alanine amidases

Zinc-dependent N-acetylmuramoyl-L-alanine amidases cleave the amide bond between the lactyl moiety of N-acetylmuramic acid and L-alanine (Figure 10, EC: 3.5.1.28). The enzymes comprise the catalytic domain ("cat") and so-called repeat ("R") or amidase N-terminal ("AMIN") domains, which are connected by flexible linker domains that may consist of over a hundred amino acids (see chapters 1.5.4 & 1.5.5). The R or AMIN domains mediate targeting and anchoring of the amidase to its cellular site of action [88, 89]. They furthermore can be located N-terminally (e.g. [90]) as well as C-terminally (e.g. [89]) to the catalytic domains and have differing copy numbers [91, 92].
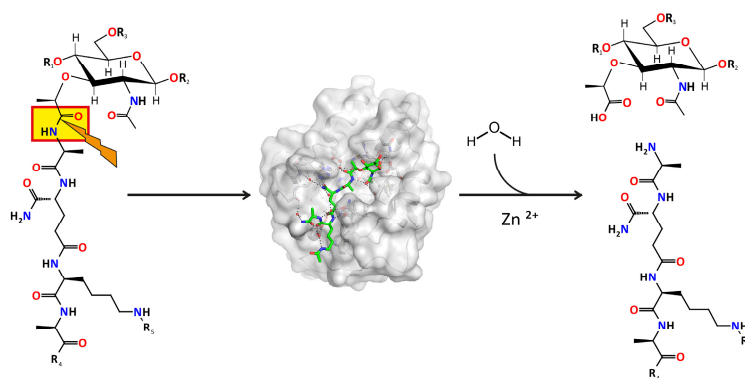


Figure 10 – The general function of N-acetylmuramoyl-L-alanine amidases (EC: 3.5.1.28) is to cleave the amide bond between the lactyl moiety of MurNAc and L-Ala. First, the PGN-substrate binds to the amidase and is subsequently hydrolyzed via a zinc-dependent mechanism. Residues "R" as in Figures 7 and 8.

The variability of PGN in different bacteria is reflected by changes in the specificity of amidases processing these networks. Furthermore, one bacterium can have several amidases that fulfill different tasks [80]. Most amidases are transported either to the periplasm (Gram-negative species) or the extracellular space (Gram-positive species) where they cleave the amide bonds of PGN, some of which in their folded state (e.g. AmiA [E. coli], AmiC [E. coli] [93]) and others unfolded (e.g. AmiB [E. coli], AmiC2 [N. punctiforme] [43, 93]). In most cases, the bacterial N-acetylmuramoyl-L-alanine amidases facilitate cell separation (e.g. AmiE [S. epidermidis], AmiA [S. aureus] [94, 95]). They moreover play a role in PGN fragment recycling [96] in the periplasm (e.g. Rv3717 [M. tuberculosis] [97]) as well as in the cytosol after import of PGN-fragments (e.g. AmpD [C. freundii] [98]). Last but not least, it was shown that they are also important for intercellular communication and nutrient exchange in filamentous bacteria (AmiC2 [N. punctiforme] [43]).

As part of the bacterial defense, amidases can also be secreted to digest the cell walls of adversaries (Type 6 Secretion System, "T6SS" [99]) and bacteriophages utilize these enzymes during the infection process to perforate the cell walls of their hosts more efficiently [100].

### 1.5.3. Amidase families and relatives

The Pfam (Protein <u>fam</u>ilies) - webpage is a database that contains more than 16,000 entries and groups related enzymes into families [91]. Five of those are N-acetylmuramoyl-L-alanine amidases (Table 2). The Amidase_2 and Amidase_3 families are rather well investigated and listed as the two major representatives. Both hydrolases are zinc-dependent and the catalytic domains adopt globular, mixed αβ-folds. So far, little is known about the remaining three members of the clan (Amidase_5, Amidase_6, and Amidase_02c), and no structures of members of these families have been deposited with the wwPDB to date (Table 2).

| FAMILY | PFAM-ID | PHYLOGENY | NAME | *ORGANISM* | *GRAM* |
|--------|---------|-----------|------|------------|--------|
| **AMIDASE_2** | PF01510 | | AmiA | *Staphylococcus aureus* | *Gram-positive* |
| | | | AmiE | *Staphylococcus epidermidis* | *Gram-positive* |
| | | | PlyL | *Bacillus anthracis* | *Gram-positive* |
| | | | XlyA | *Bacillus subtilis* | *Gram-positive* |
| | | | LytA | *Streptococcus pneumaniae* | *Gram-positive* |
| | | | AmiD | *Escherichia coli* | *Gram-negative* |
| | | | AmpDh3 | *Pseudomonas aeruginosa* | *Gram-negative* |
| | | | AmpDh2 | *Pseudomonas aeruginosa* | *Gram-negative* |
| | | | AmpD | *Citrobacter freundii* | *Gram-negative* |
| **AMIDASE_3** | PF01520 | | AmiC2 | *Nostoc punctiforme* | *Gram-negative* |
| | | | CwlV | *Paenibacillus polymyxa* | *Gram-positive* |
| | | | AmiC | *Escherichia coli* | *Gram-negative* |
| | | | AmiB | *Bartonella henselae* | *Gram-negative* |
| | | | Rv3717 | *Mycobacterium tuberculosis* | *Gram-positive* |
| **AMIDASE_5** | PF05382 | - - - | - - - | - - - | - - - |
| **AMIDASE_6** | PF12671 | - - - | - - - | - - - | - - - |
| **AMIDASE_02C** | PF12123 | - - - | - - - | - - - | - - - |

Table 2 - Pfam Amidase families with ID and representatives for which a structure is available. Gram characteristics and phylogenetic relation, based on the sequences of the respective structures, are also given.

The phylogenetic relationships among the amidases correlate with the Pfam families and also reveal parallels within a family that can be connected to PGN type or function [101] (Table 2). Although there is little sequence homology among different amidases, the structurally defined amidases share similar folds, thus demonstrating that the same fold can be achieved with a low level of amino acid conservation. Accordingly, conservation is restricted to few catalytic and structural residues [101].

Eukaryotic peptidoglycan recognition proteins (PGRPs) are part of the innate immune system and able to recognize bacterial PGN (e.g. reviewed in [102-104]). In some cases, PGRPs exhibit catalytic activity [105] and degrade bacterial PGN, possibly leading to cell lysis. PGRPs comprise at least one peptidoglycan-binding domain that is structurally homologous to the catalytic domain of bacterial enzymes of the Amidase_2 family, as seen for the seven known human PGRP structures [106-110].

### 1.5.4.   Major Autolysin of *Staphylococcus aureus* AtlA

AtlA, AtlE, and AtlL of *S. aureus*, *S. epidermidis* and *S. lugdunensis*, respectively, are well-studied representatives of the major autolysin Atl [87, 95, 111, 112], which is highly conserved among Staphylococci, especially the amidase domain [113]. AtlA is a pre-pro-protein that consists of a signal peptide for translocation to the extracellular space and a pro-peptide of so far unknown function, as well as the PGN-processing enzymes AmiA (N-acetylmuramoyl-L-alanine amidase) and NAGase (endo-$\beta$-N-acetylglucosaminidase) [87, 95, 111, 112] (Figure 11). The two catalytic domains ("cat") of AmiA and NAGase are each linked to targeting repeats (R1-R3), where AmiA-cat has R1R2 located C-terminally and NAGase-cat is N-terminally preceded by R3. Posttranslational processing, as indicated by orange arrows, generates the active enzymes (Figure 11).
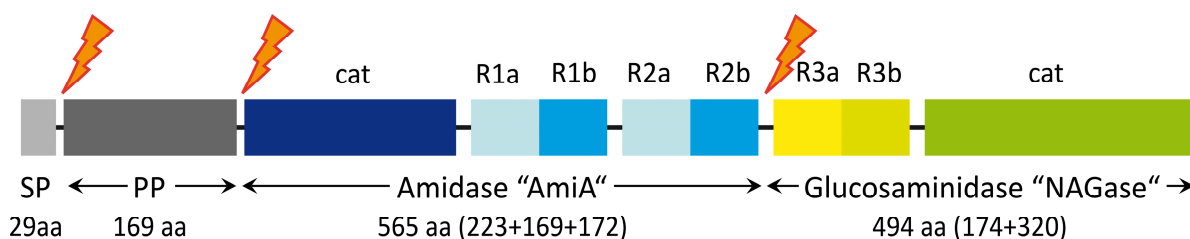


Figure 11 - Domain arrangement of the major autolysin AtlA of *Staphylococcus aureus* with posttranslational processing sites indicated by orange arrows. SP, signal peptide; PP, pro-peptide; Amidase "AmiA", EC: 3.5.1.28; Glucosaminidase "NAGase", EC: 3.2.1.96; cat, catalytic domain; R, repeat domain.

AmiA and NAGase localize at the septal region [114] and process staphylococcal PGN during cell growth and division [95]. Deletion of AtlA leads to impaired cell division and clustering of cells [95, 111], making it and its functional enzymes interesting therapeutic targets.

AmiE of *S. epidermidis* is highly homologous to AmiA (81% sequence identity). In its tandem repeats, R1a is highly similar to R2a, as is R1b to R2b (Figure 11). The R domains avoid WTA but bind LTA and target AmiE to the septum [89]. The flexible linker likely allows the catalytic domain to move and execute its hydrolytic function at several positions [89]. AmiE-cat was used to identify the minimal substrate of the Atl amidases as a muramoyltripeptide [115]. Several residues in and near the active site are conserved and critical for function, including the zinc-binding residues and a catalytic glutamate [94]. The structure of AmiE-cat reveals an Amidase_2 fold and was used for *in silico* docking studies to gain insights on the PGN binding mode [94]. Nevertheless, experimental biological and structural information on amidase-PGN interaction in Gram-positive bacteria and on amidase specificity was missing.

## 1.5.5. Unique cell wall amidase AmiC2 of *Nostoc punctiforme* ATCC 29133

The N-acetylmuramoyl-L-alanine amidase AmiC2 of *N. punctiforme* belongs to the Amidase_3 family. It differs from the homologous AmiC [E. coli] in domain composition and function [43, 44, 64]. AmiC2 carries a Sec-signal sequence and is secreted to the periplasm in an unfolded state. Moreover, it has two AMIN domains instead of the single AMIN domain found in the enzymes of unicellular bacteria such as *E. coli* [90, 116]. A stretch of approximately 160 residues contains many proline residues and connects the AMIN and catalytic domains (Figure 12). This proline-rich linker ("ProLinker") may be ordered to some degree, however, its secondary structure and its function are so far unknown (unpublished data). The catalytic domain AmiC2-cat shows low sequence similarity to related enzymes, however, catalytic and structural residues are conserved. *In vivo* and *in vitro* samples of AmiC2 revealed a possible autocatalytic cleavage product of 55 kDa that likely lacks AMIN-A (Figure 12, [44, 117]).
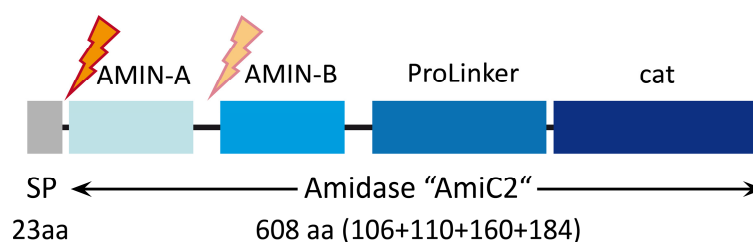


Figure 12 - Domain arrangement of AmiC2 from *Nostoc punctiforme* ATCC 29133 with the cleavage site of the signal peptide indicated by an orange arrow and a possible autoregulatory processing site given by a semi-transparent arrow.

The AMIN domains are thought to govern AmiC2's localization [88], which concurs with the septal localization of AmiC [E. coli] [118, 119]. AmiC2 is targeted to the septal disk that separates neighboring cells of the filament and is responsible for the formation of a nanopore array [44]. Lehner et al. demonstrated the importance of AmiC2 for cellular communication by generating mutant strains of *N. punctiforme* that lack this enzyme and no longer formed proper filaments, exhibited irregular septa, and lost the ability of differentiation [43, 44]. Furthermore, intercellular calcein exchange is abolished in the absence of AmiC2 [43].

AmiC2 lacks additional sequence features as found in previously characterized Amidase_3 enzymes AmiC [E. coli] and its ortholog [90, 120] as well as Rv3717 [M. tuberculosis] [97, 121] that fold into elements which regulate enzymatic activity. The control of AmiC2 activity is thus elusive. Moreover, it is unknown how the unique function of nanopore formation is accomplished, as AmiC2-cat fused to AMIN [E. coli] exhibits cell-splitting function in *E. coli* [44].

## 2. OBJECTIVES

At the beginning of the AmiA project, Zoll et al. already established a crystal structure and a docking model of a bound substrate fragment for the homologous AmiE from *S. epidermidis* [94, 115]. Several crucial residues and a minimal substrate motif were also defined. Those findings were to be validated for AmiA of the highly pathogenic *S. aureus*. In order to determine the substrate specificity, the mode of recognition, and the catalytic mechanism of AmiA-cat, well resolved crystal structures of the enzyme needed to be determined in the absence and presence of its substrate component muramoyltetrapeptide (MtetP, MurNAc-L-Ala-D-iGln-L-Lys-NHAc-D-Ala-NH$_2$). Furthermore, functional assays were to be performed to determine activity and specificity of AmiA-cat and a mutated variant with different substrates.

AmiC2 of *N. punctiforme* is an amidase with a unique function. Lehner et al. established that this enzyme is essential for formation of a nanopore array in the septal disk and important for intercellular exchange [43, 44]. An inhibitory α-helix that covers the active site in a homologous enzyme is missing, and regulation is therefore likely entirely different. The localization of native AmiC2 to the septa of a filament was to be validated using immunofluorescence. Furthermore, my focus was to solve and examine the crystal structure of the catalytic domain of AmiC2 (AmiC2-cat). Additional insights into function and specificity should be gained from functional analyses, including dye release assays, and the mutation of zinc-binding residues as well as the potentially catalytic glutamate 578.

N-acetylmuramoyl-L-alanine amidases exhibit a low level of amino acid conservation, generally only twenty to thirty percent, although they perform the same reaction and serve similar functions. In the course of this thesis it thus became tempting to speculate that sequence diversity of amidases is linked to differential recognition of substrates, diverse interactions with additional domains, and subsequently variable functions. A comparison of all amidase structures available in the worldwide Protein Data Bank (wwPDB) needed to be conducted to reveal such relations.

# 3. RESULTS AND DISCUSSION

This chapter summarizes the published results on bacterial N-acetylmuramoyl-L-alanine amidases that are presented as full texts in the appendix under *Publications* [101, 117, 122]. For details and references, please consult that section.

## 3.1. Structure-function analysis of *Staphylococcus aureus* amidase reveals the determinants of peptidoglycan recognition and cleavage

*Felix Michael Büttner*, *Sebastian Zoll, Mulugeta Nega, Friedrich Götz, and Thilo Stehle.*

The major autolysin AtlA of *Staphylococcus aureus* features two enzymes, a glucosaminidase (NAGase) and the N-acetylmuramoyl-L-alanine amidase (AmiA). The deletion of AtlA leads to severely impaired cells that cluster. Although such cells exhibit no lethal phenotype, they are not able to proliferate. Inhibition of the autolysin may therefore be a promising lead for active compounds against methicillin-resistant *S. aureus* (MRSA).

To establish knowledge about peptidoglycan (PGN) engagement and binding, the structure of the catalytic domain of AmiA (AmiA-cat) was solved to the atomic resolution of 1.12 Å. AmiA-cat features a mixed αβ-fold with a central, six-stranded β-sheet that is surrounded by seven helices. The active site is located in a well-defined binding cleft that is made up of two α-helices and loops. A catalytic zinc ion is coordinated by two histidine and one aspartate side chains.

In order to obtain data on the molecular interactions upon substrate binding, we sought to solve a protein-ligand structure as well. AmiA-cat is highly active even in crystallized form. Prior to soaking, crystals needed to be inactivated by EDTA-treatment over three days to remove the zinc ion from the active site. Muramoyltetrapeptide (MtetP) was chosen as this modified ligand containing an amidated D-Ala and acetylated L-Lys approximates the physiologic ligand best. MtetP was soaked over another sixty hours into EDTA-treated AmiA-cat crystals. Binding of the ligand to AmiA-cat was successful, and a complex structure

at 1.55 Å resolution with excellent statistics was obtained. The enzyme did not rearrange upon substrate binding, as indicated by only 0.2 Å Cα-rmsd between apo structure and complex structure. It has to be stated that only two out of four copies in the asymmetric unit exhibit ligand density because of smaller solvent channels and less accessible binding sites.

A simulated annealing omit map provides bias-reduced electron density for MtetP along the entire peptide stem, while the MurNAc moiety exhibits less well-defined density due to higher flexibility, which is also mirrored in elevated B-factors.

The complex structure reveals a dense interaction network of the four amino acids of the stem peptide with AmiA-cat, although some hydrogen bonds are water-mediated. While L-Ala only makes one hydrophobic interaction, D-iGln forms the majority of hydrogen bonds and contributes strongly towards binding and orientation of the ligand. D-Ala and L-Lys form one further direct and four water-mediated hydrogen bonds. Furthermore, hydrophobic interactions by side chains with ligand methyl groups as well as those of W310 with L-Lys promote proper ligand binding. In contrast, the carbohydrate moiety of MtetP is exclusively bonded via the $CH_3$ of the N-acetylation near F293 and hydrogen bonds of MurNAc's O1 with AmiA-cat. Interestingly, the pyranose opened and closed again during soaking, leaving the anomeric carbon in axial (α-) instead of equatorial (β-) configuration.

Substrate binding induces a tilt of the H382 side chain as the imidazole moiety engages the lactyl-oxygen of MtetP via a hydrogen bond. The active site is formed by H265, H370, and D384, which coordinate the zinc ion, as well as a catalytic E324. A structurally conserved water molecule, "w10", is present in both apo and ligand structures (slightly shifted) and in good position to attack the scissile bond. This catalytic water makes it possible to infer a plausible mechanism in which E324 and D266 bind the hydrogens and the zinc ion contacts one free electron pair of w10. Well oriented, the remaining lone pair of w10 attacks the lactyl carbon, introducing a tetrahedral intermediate. H382 stabilizes this tetrahedral intermediate by providing a proton for the resulting oxanion. Re-formation of a carbonyl leads to the release of cleaved carbohydrate backbone and peptide stem as leaving groups.

The entire mechanism of PGN engagement is, however, more complex. To enhance our understanding, we also examined surface properties of AmiA-cat and the conservation of residues among related amidases.

Within the MurNAc binding site, interacting residues identified in our complex structure are strictly conserved, as are further residues that define the border of the pocket. Binding of the PGN carbohydrate backbone is thus likely similar in all amidases. Mapping the conserved residues onto the molecular surface and examination of the electrostatic surface potential gave further insights where neighboring N-acetylglucosamines may bind. Based on NMR structures of PGN fragments, we were able to model two adjacent GlcNAc moieties. Similarly, we examined the properties of Ami t at the lower end of the binding site to infer hints on binding of the pentaglycine bridge. Nonetheless, no clear indication points to one certain area although a biological assay shows that the enzyme discriminates PGN based on the presence of the pentaglycine bridge. AmiA-cat digests over 95% of mutanolysin pre-treated *S. aureus* PGN, as analyzed by reversed-phase high-performance liquid chromatography and mass spectroscopy, while similarly treated *B. subtilis* PGN on the other hand was not digested. The substitution of L-Lys with *m*DAP and the direct cross-linkage of *B. subtilis* PGN is likely not accepted by AmiA-cat, whereas iGlu instead of iGln should be tolerated [115].

The comparison with the enzyme-substrate complex of AmiD [E. coli] [123], at that time the only other published complex structure with uncleaved substrate, reveals structural differences and the substitution of peptide stem binding residues. In line with this, *E. coli* and *B. subtilis* share the same type of PGN (A1γ), which AmiA-cat cannot cleave. The substrate in the AmiD [E. coli] structure also contains 1,6-anhydro MurNAc, which is not present in *S. aureus* PGN. Therefore, this complex structure is not suitable to infer insights into *S. aureus* PGN engagement and thus highlights the importance of our findings. Moreover, the AmiA-cat complex improved previous docking studies from the homologous AmiE-cat and presents new interactions that set the basis for proper structure-based inhibitor design.

A potential inhibitor of AmiA-cat may feature an uncleavable amide mimic, for example a molecule containing a tetrahedral dihydroxyethylene sub-structure. The carbohydrate moiety is important for binding [115] but could be exchanged with a similar cyclic compound that is easier to handle, improves pharmacokinetics, and is financially more feasible.

Taken together, we provide a highly resolved complex structure of AmiA-cat with its ligand that significantly broadens our knowledge on how Staphylococci process their PGN. These findings build the basis for development of therapeutics against this important human pathogen.

## 3.2. Enabling cell-cell communication via nanopore formation: structure, function, and localization of the unique cell wall amidase AmiC2 of *Nostoc punctiforme*

***Felix Michael Büttner***, *Katharina Faulhaber, Karl Forchhammer, Iris Maldener, Thilo Stehle*

*FEBS J, (2016), 283: 1336–1350. doi:10.1111/febs.13673*

*Nostoc punctiforme* is a cyanobacterium that organizes in filaments and is able to differentiate into specialized cells. Its amidase AmiC2 exhibits the novel function of creating a nanopore array. In contrast to cell splitting amidases that merely cleave the scissile amide bond for proliferation, AmiC2 drills exactly defined holes into the septal disk between neighboring cells. These nanopores are involved in intercellular communication and nutrient exchange.

Sequence analyses revealed that AmiC2 is a member of the Amidase_3 family. In contrast to the homologous, cell-splitting amidase from *E. coli*, AmiC2 possesses two AMIN domains that are possibly involved in regulation and targeting to young septa. It also contains a long proline-rich linker domain of unknown function, and, furthermore, AmiC2 is secreted to the periplasm, not translocated as AmiC *E. coli*. Nevertheless, in previous experiments AmiC2-cat was fused to the *E. coli* AMIN domain and fulfilled cell splitting function in an *E. coli* knockout strain [44]. The only other structurally and functionally characterized member of the Amidase_3 family is Rv3717 *M. tuberculosis*, a recycling amidase with broad-spectrum activity.

AmiC2 was detected in young septa of the filament by immunofluorescence. A specific antibody without cross-reactivity recognizes the *Nostoc* sequence within AMIN-AB. It was shown that AmiC2 follows the invaginating cell wall and finally occupies the whole septal disk, while only little AmiC2 was present in old septa. It is noteworthy that a truncated and possibly autoproteolytic fragment of AmiC2 (55 kDa in size, lacking the AMIN-A domain; not shown) was also detected.

The catalytic domain (AmiC2-cat) exhibits higher activity than lysozyme in a dye release assay (DRA), which is based on the release of dye from stained *E. coli* PGN upon enzymatic breakdown. The AMIN domains have no catalytic effect. The strong hydrolytic activity of

AmiC2-cat may be linked to the thick PGN of cyanobacteria compared to the rather thin PGN of Gram-negative bacteria.

To gain further knowledge on the amidase, we solved the structure of AmiC2-cat at 1.12 Å resolution with excellent statistics. The protein comprises a mixed αβ-fold of a central, six-stranded β-sheet, which is surrounded by six α-helices. The binding groove is defined by helices α2 and α4 as well as loops. However, this groove is wide and shallow, and AmiC2-cat may therefore accommodate the substrate in various conformations. Zinc coordination is facilitated by H447, H515, and E462, while two additional residues, D451 and S513, help to properly orient the histidine side chains. We also found a buffer molecule and the N-terminus of AmiC2-cat's crystallographic symmetry mate in the binding groove. Both may have implications for substrate binding and catalysis. The high catalytic activity concurs with a positive surface potential of AmiC2 around the active site, which also implies a negatively charged substrate.

A tight spatio-temporal regulation is imperative for the precise function of AmiC2. Nonetheless, it is unknown to date how the catalytic activity is controlled. AmiC2 lacks an inhibitory α-helix that is found in homologs and that is allosterically removed when amidase activity is needed [90]. As a product of autoproteolysis, the 55 kDa species of AmiC2 may be part of an autoregulation mechanism. However, the uniform size and shape of the nanopores may need a rather complex machinery. It was shown that certain PGN-binding proteins influence the nanopore formation in *Anabaena* [58, 124], but this has not yet been investigated for *Nostoc*.

The sequence conservation of AmiC2-cat compared to homologous enzymes was investigated and also mapped onto AmiC2-cat's surface. All residues involved in proper zinc binding as well as the catalytic E578 are strictly conserved, making a common enzymatic mechanism within this family likely. On the other hand, further conservation was observed only for structurally important residues and not for additional surface-exposed patches that relate to a defined binding site. Comparison with the closely related Amidase_2 family, where a definite and rather narrow binding cleft allows credible docking of the substrate (AmiE [S. epidermidis] [94], LytA [S. pneumoniae] [125]), also gave no indications on where and how AmiC2-cat might exactly bind to PGN in order to cleave it.

The mutation of conserved surface residues revealed a structural dependency of AmiC2 on the zinc ion. Once AmiC2 was no longer able to bind zinc due to a missing coordination site, it aggregated. Accordingly, zinc-removal from AmiC2-cat in solution by EDTA also lead to precipitation. This structural dependency on zinc is not consistent within the amidase families. Mutants of E578, however, were all stable and properly folded. The exchange of glutamate with alanine, aspartate, asparagine, or glutamine shows that both charge and correct distance are important for catalysis as none of the mutants exhibited residual activity.

A crystallographic occurrence gives reason to hypothesize on the mechanism of AmiC2: The N-terminus of the symmetry mate reaches into the active site of AmiC2-cat and participates in zinc binding. The position of the terminal $NH_2$ group can also be occupied by a water molecule. As length and charge of E578 are particularly critical, the residue likely acts analogously to E324 of AmiA-cat [S. aureus]. In cooperation with a backbone carbonyl, and of course the zinc ion, E578 correctly positions the hydrogens and free electron pairs of the proposed water molecule. The distances and angles of these interactions fit well so that the water is ready for an attack on the scissile bond. A possible proton relay, as suggested by Mayer et al. [126], is nearby and may assist in catalysis. The MES molecule from the crystallization buffer bound in a pocket close to the zinc ion and might occupy the carbohydrate binding site.

We present the high-resolution structure of the catalytic domain of AmiC2 and show the localization to young septa. Furthermore, we highlight the importance of E578 and proper zinc binding. Although it was not possible to gain direct knowledge on how AmiC2 fulfills its task, we provide an excellent basis for future studies to elucidate the binding site, a catalytic mechanism, as well as regulation and specificity of AmiC2.

## 3.3. X-ray crystallography and its impact on understanding bacterial cell wall remodeling processes

X-ray crystallography turned one hundred years old in 2014. To celebrate this occasion, several prestigious journals including *Nature* and *Science* published special issues featuring this key method in structural biology. This review also relates to that anniversary and highlights the importance of X-ray crystallography for microbiology even today using the example of bacterial cell wall remodeling by N-acetylmuramoyl-L-alanine amidases. The study also aims to give non-structural biologists a set of guidelines to critically evaluate protein structures and to work with them.

The most abundant function of amidases is cell splitting in the septal region, which is important for proliferation. However, there are additional amidases, for example involved in PGN fragment recycling, cellular communication, bacterial defense against competitors, or bacteriophage infections. Amidases furthermore have a close structural relationship to eukaryotic peptidoglycan recognition proteins (PGRPs). Nevertheless, we focused on the families Amidase_2 and Amidase_3, as only protein structures from these families are deposited in the wwPDB.

Molecular or atomic protein structures give insights into the function of enzymes and concern mostly basic research. Yet, one can obtain clues for therapeutic agents from the analysis of amidase structures of pathogenic bacteria such as *S. aureus*, *S. epidermidis*, *P. aeruginosa*, *E. coli*, *M. tuberculosis*, or *N. meningitidis*.

Most amidases possess AMIN or repeat domains that lie N- or C-terminally to the catalytic domains and often target and anchor the enzymes to / at their designated place of function. Due to limitations of space we focused on catalytic domains. Although the sequence conservation among amidases is generally low (around 30 %), they share a globular, mixed αβ-fold featuring a central β-sheet with surrounding α-helices that define a core fold.

Nevertheless, distinct differences between families exist, and the families are characterized in separate chapters.

The Amidase_2 family featured 23 structures belonging to nine different enzymes in mid-2014. All share a six-stranded central β-sheet, while helices α4 and α7 stabilize the fold on the rear side of this sheet. Helices α2 (right), α5, and α6 (left) along with several loops are on the front side of the β-sheet and define the binding grove as well as the active site. Two histidines and one aspartate (cysteine for phage-derived amidases) coordinate the catalytically crucial zinc ion. Further crucial residues that are conserved include Thr267 for carbohydrate binding, Trp310 for hydrophobic and CH–π interactions with L-Lys or *m*DAP of the substrate, and the structurally important Asn317 that also engages the substrate (AmiA $^{S\ .aureus}$ numbering). His382 (or rarely a lysine) stabilizes the tetrahedral reaction intermediate and, of course, the catalytic Glu324. Natural presence or the introduction of a positive surface charge increases lytic activity. Although the reaction mechanism is likely the same throughout the Amidase_2 family, the regulation of catalytic activity seems not to be conserved. Shape and size of the binding sites differ among species, accounting for different specificities. Phylogeny and characteristics of the organism furthermore suggest additional classification of the Amidase_2 family into an AmiA-type and an AmiD-type.

Supplementary features of the AmiD-type are a β-loop insertion that is involved in carbohydrate binding and (except for AmpD $^{C.\ freundii}$) an additional N-terminal extension for multimerization as well as a C-terminal domain that is possibly involved in PGN recognition. AmpD $^{C.\ freundii}$ is a cytosolic enzyme, which has a recycling function with a specificity for PGN-monomers, and it is the only amidase for which structural rearrangements are reported [98].

Nine structures of five different proteins of the Amidase_3 family had been solved when the review was written. Again, a six-stranded central β-sheet is stabilized by helices α1, α3 and α5 on the rear side. The binding site is predominantly defined by loops, not helices. Furthermore, it is wider and more shallow than for the Amidase_2 family. Conserved residues next to the active site facilitate correct orientation of two histidine side chains that bind the zinc ion together with a glutamate. Glu373 (AmiC $^{E.\ coli}$ numbering) is the catalytic residue, the carbonyl oxygen of Ala266 may aide in the water-activation, analogously to Asp266 in Amidase_2 enzymes, and Asp/Asn267 may interact with

26

MurNAc. An unusual kink in helix α5 of all enzymes is caused by side chain–backbone hydrogen bonds of Ser/Thr/Asn at that position.

Some members of the Amidase_3 family have additional residues inserted into the catalytic domain. A lid over the active site initiates PGN-fragment recycling by Rv3717 [M. tuberculosis], possibly with an autoregulatory function, and an inhibitory α-helix blocks the active site of AmiC [E. coli] and only moves out of it upon interaction with LytM activators that are associated with septal ring factors.

Complex structures of substrate components bound to enzymes provide most mechanistic insights, however, obtaining them is challenging. Due to their diverse structures, it is difficult to acquire PGN-derived substrates in sufficient amounts and purity. They are not commercially available but have to be custom-synthesized. Furthermore, most complexes need non-cleavable substrate analogs or inactive / inactivated enzymes as they often exhibit residual activity even once crystallized.

The complex structure of MtetP bound to AmiA-cat [S .aureus] was depicted exemplarily. Differences compared to other complexes link amino acid substitutions in the enzymes to altered specificities. The exchange of H370 and T380 to lysine and arginine, respectively, shifts substrate specificity from D-iGln in the peptide stem to D-iGlu. Likewise, deletion of several residues shrinks the carbohydrate pocket and changes specificity towards terminal PGN-fragments with 1,6-anhydro MurNAc. Additional arginine residues at the end of the binding groove may stabilize substrates that carry free carboxylates. The complexes of AmpDH2 [P. aeruginosa] and AmpDH3 [P. aeruginosa] shed light on PGN engagement by amidases in a greater aspect as they bound larger fragments of PGN that are composed of several carbohydrate rings and peptide stems.

The review summarizes the structural knowledge on N-acetylmuramoyl-L-alanine amidase as of mid-2014 and highlights common as well as distinct features within and between amidase families.

# 4. CONCLUSIONS AND OUTLOOK

## 4.1. Remaining challenges in the field

### 4.1.1. Peptidoglycan

We still know too little about the bacterial PGN to fully understand the underlying mechanisms that govern targeting and specificity for different amidases of different species. Knowledge on the PGN of individual species in their three-dimensional form, their composition, their modifications, and their linkages as well as the degree of polymerization is up to date limited for most species. Even when experimental data on those properties is available, it is unknown whether there are distinct PGN-types for different stages of maturation or different places in the cell wall, e.g. the septum [67]. As de Pedro and Cava highlight in their outlook, it is imperative to broaden our knowledge on peptidoglycan, e.g. with the new techniques at our disposal by the recent "-omics" boom [67]. That knowledge will help to design more precise experiments on activity and specificity. In combination with chemical synthesis of various compounds of then identified substrates, structural biology methods can provide more detailed and dependable insights into what most likely happens during PGN-modulation and the different subsequent cellular processes. Falling in line with this, just recently, a Japanese group developed a "*Peptidoglycan Microarray as A Novel Tool to Explore Protein-Ligand Recognition*" [127].

### 4.1.2. N-acetylmuramoyl-L-alanine amidases

Following up, more complex structures of amidases bound to their substrates need to be determined to diversify the database of known interactions. Usage of more sophisticated substrate fragments that contain several carbohydrate moieties and peptide stems is an additional long-term goal. This way, one can better determine common features of all amidases or particular (sub-) families and distinct residues or patches that govern specificity and activity. That will likely happen first with catalytic domains. However, we also need to gain more knowledge on the other amidase domains and putative linker regions. As shown for AmiE [S. epidermidis], for example, the repeat domains determine the locus of catalytic action. More structures of repeat and AMIN domains need to be determined in an integrated approach with functional data. That is ideally accomplished by usage of full-length proteins instead of

independent domains, in order to find out more about the interplay of different domains regarding targeting and regulation. Since full-length amidases consist of several domains that are in some cases connected by flexible linkers, X-ray crystallography may come to its limits and needs to be combined with other techniques such as small angle X-ray scattering (SAXS), as was successfully done for AmiE *S. epidermidis* [89] and LytA *S. pneumoniae* [128] earlier. Despite the recent advancements of cryo electron microscopy, this technique is not yet feasible to study full-length amidases, due to its limitation to ensembles greater 200 kDa [129]. While it is possible to "bulk-up" smaller proteins with antibodies or Fab-fragments, this is probably not applicable for these multi domain enzymes with their long, flexible linkers. Negative stain electron microscopy, on the other hand, is an easy to handle method that yields immediate low-resolution results that could complement SAXS measurements and give rough information on domain interplay [129].

### 4.1.3. Interplay of proteins

Furthermore, it is advisable to also investigate other proteins that are directly or mediately involved in amidase activity. Such proteins may act as activators or inhibitors, and they could be involved in, or control localization. There might also be adapter proteins that govern co-localization directly with the biosynthesis machinery, or certain membrane or cell wall features. Finally, one needs to determine whether a specific amidase acts independently, or if it is involved in a degradation complex.

## 4.2. Specific challenges for the investigated amidases

### 4.2.1. Staphylococcal autolysin AtlA

The major autolysin AtlA of *S. aureus* and its close relative AtlE of *S. epidermidis* have undergone intensive research. The amidases of AtlA / AtlE are well characterized by structures of the repeat domains and the catalytic domains, including a substrate-bound structure, many functional data, and SAXS-derived information on how the full-length enzyme behaves in solution. Still, several points remain unanswered.

Although we determined key interactions of AmiA-cat with the PGN peptide stem, we lack data on how larger PGN fragments are engaged. Therefore, additional complex structures with more complex ligands that contain several carbohydrates and more than one peptide stem could verify or correct our proposed carbohydrate binding pocket. A complex structure with a substrate made of two peptide stems that are cross-linked by the physiologic pentaglycine bridge would be even more interesting. Enzymatic evidence by the digest of *S. aureus* and *B. subtilis* PGN showed that this feature is critical for activity, yet it was not possible to infer strong hints on interactions of AmiA-cat with this bridge or adjacent peptide stems.

Since AtlA is transcribed and translated as a large pre-pro-Atl polypeptide, which is secreted and further processed into three separate proteins, it is intriguing to ask whether those three proteins also functionally interact. While the function of the first component, the pro-peptide, is elusive to date, the third and last protein - called NAGase - also cleaves PGN. The characterization of the N-acetylglucosaminidase, which processes the carbohydrate backbone, is another piece in the puzzle of AtlA action. The structure and function of NAGase can be best judged by additional information based on a complex structure with several carbohydrate moieties. Such a structure may also show what a minimal substrate would look like. The investigation whether there is a cross talk between AmiA and NAGase and possibly even a co-localization with cooperative effects would finally bring our understanding of AtlA to a new level that could be used as a rationale for lead structure design to attack the important human pathogen *Staphylococcus aureus*.

**Ongoing research**

With reference to possible compounds that may be utilized to counteract *S. aureus* or more precisely and importantly, the highly resistant MRSA, a diploma student in our laboratory, Melanie Oelker, investigated the potential of fragment based screening [130]. In this process, a defined library of compounds is applied to a large-scale crystallization setup. The resulting crystals are measured and the subsequent structures screened for ligand binding. Some binders to the active site and / or the binding grove of AmiA-cat were identified using this approach. Those compounds now need to be tested in our established activity assay. A second round of screening with a broader library could increase the number of candidates. Future steps would include the development of a larger compound that contains the interacting fragment and improves pharmaceutically relevant characteristics. We expect it to be more likely to identify an inhibitor by this technique as most fragments bound in the active site. However, instead of merely stalling cell growth by the use of an inhibitor, an activator of AmiA is more promising to battle MRSA since such a compound could induce cell lysis.

Another strategy in treating MRSA infections is being tested in cooperation with Dr. Beatrix Förster from the Gene Center of the Ludwig-Maximilians-University in Munich. Antibodies against the extracellular AmiA-cat, purified by us, are raised and evaluated for therapeutic use.

### 4.2.2. AmiC2 of *Nostoc punctiforme* ATCC 29133

Unlike AmiA, the unique amidase AmiC2 of *Nostoc punctiforme* is not a therapeutically interesting target. The elucidation of its novel function that enables cell-cell communication is nonetheless important and fascinating. We are therefore following up on several leads that include many of the above-mentioned aspects.

**Ongoing research**

The proper targeting of AmiC2 is essential for drilling nanopores into the septum. Sebastian Grün investigated the AMIN domains under my supervision in the course of his Bachelor thesis [131]. Three constructs, AMIN-A, AMIN-B, and AMIN-AB, were investigated. Protein purification resulted in unstable AMIN-A, and low yields of highly soluble AMIN-B. AMIN-AB on the other hand gave good yields of highly pure protein. However, the linker connecting AMIN-A and AMIN-B was labile, resulting in a mixture of the two domains that could not be separated. Crystallization of the three constructs was unsuccessful.

In order to identify the binding grove of AmiC2-cat, I am currently pursuing a complex structure. A first prerequisite is the inactivation of the enzyme to avoid crystal damage by flexible and charged reaction products, as seen for AmiA-cat. Although AmiC2-cat did not tolerate zinc removal in solution either by addition of the chelator EDTA or the mutation of the zinc binders, the zinc ion can be removed from crystallized AmiC2-cat without destroying the crystal. In fact, the crystals suffer strongly by visual examination but retain their full diffraction potential to 1 Å resolution. A second approach is the crystallization of the E578 mutants. Two mutants were tested and both, E578A as well as E578Q, yielded well diffracting crystals after cross-seeding where microscopic wt-crystal fragments are used to initiate the formation of mutant crystals. Resolution of the mutant crystals was better than 1.5 Å.

With the goal to use a ligand most similar to the actual substrate, several custom suppliers were asked to synthesize a respective compound consisting of MurNAc-L-Ala-D-iGlx-*m*DAP-D-Ala. Unfortunately, it was not possible to synthesize the compound in sufficient amounts and / or at a feasible price. The *meso*-diaminopimelic acid with its two stereo centers appears to be an obstacle for efficient synthetic chemistry. We thus performed soaking experiments

as well as co-crystallization of wt and mutant AmiC2-cat with commercially available muramoyldipeptide and pentapeptide in numerous experiments. Nevertheless, no complex formation could be observed even though the solvent channels are of sufficient size and the active site is solvent accessible. The presence of *m*DAP in the substrate may therefore be crucial for tight binding in a crystal complex or for any binding at all.

Only recently, Wang et al. reported the successful synthesis of PGN-fragments containing *m*DAP [132]. This may be a "silver lining" for future attempts to obtain complex structures of enzymes that process PGN of the A1γ-type.

We are curious about the exact composition of *Nostoc's* PGN and collaborate with Ute Bertsche and Christoph Mayer to investigate this by characterization of PGN fragments using ultra-performance liquid chromatography coupled with mass spectroscopy (UPLC / MS). If possible, a chemical distinction of the local PGN composition in *Nostoc's* septal disk can help in defining the actual substrate of AmiC2 and may explain the specificity if substantial differences arise.

To tie on that, we wonder how the functionality of AmiC2 is regulated. In *Nostoc*, it forms nanopores, however, *in vitro* and heterologously expressed in *E. coli*, AmiC2 exhibits unrestricted and cell splitting activity, respectively. The full length AmiC2 is unstable and a 55 kDa fragment is more prominent *in vivo* as well as in overexpressed settings. This 55 kDa fragment is lacking the AMIN-A domain. It is possible that AmiC2 is not continuously active, but first localized and at some point activated by an autocatalytic reaction.

In order to evaluate the cross talk of AmiC2's domains, the 55 kDa fragment, AmiC2-cat, the AMIN domains, the ProLinker, and a ProLinker-cat construct are currently tested for activity in all possible combinations using the dye release assay.

## Concluding remark

Despite those open questions, we learned a lot about the fascinating class of bacterial N-acetylmuramoyl-L-alanine amidases. As presented in this thesis, we now have a better understanding of this enzyme class through a comparative approach and especially broadened the knowledge on two very different representatives, AmiA of *Staphylococcus aureus* and AmiC2 of *Nostoc punctiforme* ATCC 29133.

# 5. REFERENCES

1. **Madigan, M.T., Martinko, J.M., Bender, K.S., Buckley, D.H., and Stahl, D.A.,** *Brock Biology of Microorganisms: Chapter 1: Microorganisms and Microbiology. 14th Edition ed.* **2014.** <u>ISBN:</u> **9780321897398**.

2. **Madigan, M.T., Martinko, J.M., Bender, K.S., Buckley, D.H., and Stahl, D.A.,** *Brock Biology of Microorganisms: Chapter 12.2: Photosynthesis. 14th Edition ed.* **2014.** <u>ISBN:</u> **9780321897398**.

3. **Holland, H.D.,** *The oxygenation of the atmosphere and oceans. Philos Trans R Soc Lond B Biol Sci,* **2006.** *361(1470): p. 903-15* <u>DOI:</u> **10.1098/rstb.2006.1838.**

4. **Tomitani, A., Knoll, A.H., Cavanaugh, C.M., and Ohno, T.,** *The evolutionary diversification of cyanobacteria: molecular-phylogenetic and paleontological perspectives. Proc Natl Acad Sci U S A,* **2006.** *103(14): p. 5442-7* <u>DOI:</u> **10.1073/pnas.0600999103.**

5. **Madigan, M.T., Martinko, J.M., Bender, K.S., Buckley, D.H., and Stahl, D.A.,** *Brock Biology of Microorganisms: Chapter 2.21: Mitochondria, Hydrogenosomes and Chloroplasts. 14th Edition ed.* **2014.** <u>ISBN:</u> **9780321897398**.

6. **Madigan, M.T., Martinko, J.M., Bender, K.S., Buckley, D.H., and Stahl, D.A.,** *Brock Biology of Microorganisms: Chapter 23 - I: Normal Human-Microbiol Interactions. 14th Edition ed.* **2014.** <u>ISBN:</u> **9780321897398**.

7. **Bull, M.J. and Plummer, N.T.,** *Part 1: The Human Gut Microbiome in Health and Disease. Integr Med (Encinitas),* **2014.** *13(6): p. 17-22* <u>DOI:</u> **10.1053/j.gastro.2014.03.032.**

8. **Bull, M.J. and Plummer, N.T.,** *Part 2: Treatments for Chronic Gastrointestinal Disease and Gut Dysbiosis. Integr Med (Encinitas),* **2015.** *14(1): p. 25-33* <u>PMCID:</u> **PMC4566455.**

9. **Ogston, A.,** *Report upon Micro-Organisms in Surgical Diseases. Br Med J,* **1881.** *1(1054): p. 369.b2-375* <u>DOI:</u> **10.1136/bmj.1.1054.369.**

10. **Madigan, M.T., Martinko, J.M., Bender, K.S., Buckley, D.H., and Stahl, D.A.,** *Brock Biology of Microorganisms: Chapter 29.9: Staphylococcus aureus infections. 14th Edition ed.* **2014.** <u>ISBN:</u> **9780321897398**.

11. **Madigan, M.T., Martinko, J.M., Bender, K.S., Buckley, D.H., and Stahl, D.A.,** *Brock Biology of Microorganisms: Chapter 15.7: Firmicutes. 14th Edition ed.* **2014.** <u>ISBN:</u> **9780321897398**.

12. **O'Gara, J.P. and Humphreys, H.,** *Staphylococcus epidermidis biofilms: importance and implications. J Med Microbiol,* **2001.** *50(7): p. 582-7* <u>DOI:</u> **10.1099/0022-1317-50-7-582.**

13. **Chapman, G.H., Berens, C., and Stiles, M.H.,** *The Coagulation of Plasma by Staphylococci. J Bacteriol,* **1941.** *41(4): p. 431-40* <u>PMCID:</u> **PMC374708.**

14. **Clauditz, A., Resch, A., Wieland, K.P., Peschel, A., and Götz, F.,** *Staphyloxanthin Plays a Role in the Fitness of Staphylococcus aureus and Its Ability To Cope with Oxidative Stress. Infect Immun,* **2006.** *74(8): p. 4950-3* <u>DOI:</u> **10.1128/IAI.00204-06.**

15. **Götz, F.,** *Staphylococcus and biofilms. Mol Microbiol,* **2002.** *43(6): p. 1367-78* <u>DOI:</u> **10.1046/j.1365-2958.2002.02827.x.**

16. **Hall-Stoodley, L., Costerton, J.W., and Stoodley, P.,** *Bacterial biofilms: from the natural environment to infectious diseases. Nat Rev Microbiol,* **2004.** *2(2): p. 95-108* <u>DOI:</u> **10.1038/nrmicro821.**

17. **Heilmann, C., Schweitzer, O., Gerke, C., Vanittanakom, N., Mack, D., and Götz, F.,** *Molecular basis of intercellular adhesion in the biofilm-forming Staphylococcus epidermidis. Mol Microbiol,* **1996.** *20(5): p. 1083-91* <u>DOI:</u> **10.1111/j.1365-2958.1996.tb02548.x.**

18. **Archer, N.K., Mazaitis, M.J., Costerton, J.W., Leid, J.G., Powers, M.E., and Shirtliff, M.E.,** *Staphylococcus aureus biofilms: properties, regulation, and roles in human disease. Virulence,* **2011.** *2(5): p. 445-59* <u>DOI:</u> **10.4161/viru.2.5.17724.**

19. **Wertheim, H.F., Melles, D.C., Vos, M.C., van Leeuwen, W., van Belkum, A., Verbrugh, H.A., and Nouwen, J.L.,** *The role of nasal carriage in Staphylococcus aureus infections. Lancet Infect Dis,* **2005**. *5(12): p. 751-62* DOI: **10.1016/S1473-3099(05)70295-4.**

20. **Lowy, F.D.,** *Staphylococcus aureus infections. N Engl J Med,* **1998**. *339(8): p. 520-32* DOI: **10.1056/NEJM199808203390806.**

21. **Madigan, M.T., Martinko, J.M., Bender, K.S., Buckley, D.H., and Stahl, D.A.,** *Brock Biology of Microorganisms: Chapter 31.8: Staphylococcal Food Poisening. 14th Edition ed.* **2014**. ISBN: **9780321897398.**

22. **Lindsay, J.A.,** *Staphylococcus aureus genomics and the impact of horizontal gene transfer. Int J Med Microbiol,* **2014**. *304(2): p. 103-9* DOI: **10.1016/j.ijmm.2013.11.010.**

23. **Fleming, A.,** *On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. Bull World Health Organ,* **2001**. *79(8): p. 780-90* PMCID: **PMC2566493.**

24. **Abraham, E.P. and Chain, E.,** *An enzyme from bacteria able to destroy penicillin. 1940. Rev Infect Dis,* **1988**. *10(4): p. 677-8* DOI: **10.1038/146837a0.**

25. **Spink, W.W. and V., F.,** *Penicillin-resistant staphylococci: mechanisms involved in the development of resistance. Journal of Clinical Investigation,* **1947**. *26(3): p. 379-393* DOI: **10.1172/JCI101820.**

26. **Hawkey, P.M.,** *The origins and molecular basis of antibiotic resistance. BMJ,* **1998**. *317(7159): p. 657-60* DOI: **10.1136/bmj.317.7159.657.**

27. **Moellering, R.C., Jr.,** *MRSA: the first half century. J Antimicrob Chemother,* **2012**. *67(1): p. 4-11* DOI: **10.1093/jac/dkr437.**

28. **Jevons, M.P., Coe, A.W., and Parker, M.T.,** *Methicillin resistance in staphylococci. Lancet,* **1963**. *1(7287): p. 904-7* DOI: **10.1016/S0140-6736(63)91687-8.**

29. **Jevons, M.,** *Celbenin-resistant staphylococci. Br Med J,* **1961***(124–5)* PMCID: **PMC1952888.**

30. **Ubukata, K., Nonoguchi, R., Matsuhashi, M., and Konno, M.,** *Expression and inducibility in Staphylococcus aureus of the mecA gene, which encodes a methicillin-resistant S. aureus-specific penicillin-binding protein. J Bacteriol,* **1989**. *171(5): p. 2882-5* PMCID: **PMC209980.**

31. **Robert-Koch-Institut**. *Auftreten und Verbreitung von MRSA in Deutschland 2010. Epidemiologisches Bulletin,* **2011**. *Accessed on* **2016-02-20**. *Available from:* http://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2011/Ausgaben/26_11.pdf?__blob=publicationFile.

32. **Kos, V.N., Desjardins, C.A., Griggs, A., Cerqueira, G., Van Tonder, A., Holden, M.T., Godfrey, P., Palmer, K.L., Bodi, K., Mongodin, E.F., Wortman, J., Feldgarden, M., Lawley, T., Gill, S.R., Haas, B.J., Birren, B., and Gilmore, M.S.,** *Comparative Genomics of Vancomycin-Resistant Staphylococcus aureus Strains and Their Positions within the Clade Most Commonly Associated with Methicillin-Resistant S. aureus Hospital-Acquired Infection in the United States. MBio,* **2012**. *3(3)* DOI: **10.1128/mBio.00112-12.**

33. **Mwangi, M.M., Wu, S.W., Zhou, Y., Sieradzki, K., de Lencastre, H., Richardson, P., Bruce, D., Rubin, E., Myers, E., Siggia, E.D., and Tomasz, A.,** *Tracking the in vivo evolution of multidrug resistance in Staphylococcus aureus by whole-genome sequencing. Proc Natl Acad Sci U S A,* **2007**. *104(22): p. 9451-6* DOI: **10.1073/pnas.0609839104.**

34. **McDougal, L.K., Steward, C.D., Killgore, G.E., Chaitram, J.M., McAllister, S.K., and Tenover, F.C.,** *Pulsed-field gel electrophoresis typing of oxacillin-resistant Staphylococcus aureus isolates from the United States: establishing a national database. J Clin Microbiol,* **2003**. *41(11): p. 5113-20* DOI: **10.1128/JCM.41.11.5113-5120.2003.**

35. **CDC:Centers_for_Disease_Control_and_Prevention**. *Antibiotic Resistance Threats in the United States, 2013.* **2013**. *Accessed on* **2016-02-20**. *Available from:* http://www.cdc.gov/drugresistance/threat-report-2013/.

36. **Robert-Koch-Institut**. *Eigenschaften, Häufigkeit und Verbreitung von MRSA in Deutschland – Update 2013/2014: Daten zur Häufigkeit von MRSA in Deutschland. Epidemiologisches Bulletin,* **2015.** *Accessed on* **2016-02-20**. *Available from:* http://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2015/Ausgaben/31_15.pdf?__blob=publicationFile.

37. **Herrmann, M.,** *[Screening for MRSA: important and sensible]. Dtsch Med Wochenschr,* **2015**. *140(19): p. 1447-8* DOI: **10.1055/s-0041-105587.**

38. **Kallen, A.J., Mu, Y., Bulens, S., Reingold, A., Petit, S., Gershman, K., Ray, S.M., Harrison, L.H., Lynfield, R., Dumyati, G., Townes, J.M., Schaffner, W., Patel, P.R., and Fridkin, S.K.,** *Health care-associated invasive MRSA infections, 2005-2008. JAMA,* **2010**. *304(6): p. 641-8* DOI: **10.1001/jama.2010.1115.**

39. **Vos, M.C., Ott, A., and Verbrugh, H.A.,** *Successful search-and-destroy policy for methicillin-resistant Staphylococcus aureus in The Netherlands. J Clin Microbiol,* **2005**. *43(4): p. 2034; author reply 2034-5* DOI: **10.1128/JCM.43.4.2034-2035.2005.**

40. **David, M.Z. and Daum, R.S.,** *Community-associated methicillin-resistant Staphylococcus aureus: epidemiology and clinical consequences of an emerging epidemic. Clin Microbiol Rev,* **2010**. *23(3): p. 616-87* DOI: **10.1128/CMR.00081-09.**

41. **Madigan, M.T., Martinko, J.M., Bender, K.S., Buckley, D.H., and Stahl, D.A.,** *Brock Biology of Microorganisms: Chapter 14.3: Cyanobacteria. 14th Edition ed.* **2014**. ISBN: **9780321897398**.

42. **Kaiser, D.,** *Building a multicellular organism. Annu Rev Genet,* **2001**. *35: p. 103-23* DOI: **10.1146/annurev.genet.35.102401.090145.**

43. **Lehner, J., Zhang, Y., Berendt, S., Rasse, T.M., Forchhammer, K., and Maldener, I.,** *The morphogene AmiC2 is pivotal for multicellular development in the cyanobacterium Nostoc punctiforme. Mol Microbiol,* **2011**. *79(6): p. 1655-69* DOI: **10.1111/j.1365-2958.2011.07554.x.**

44. **Lehner, J., Berendt, S., Dorsam, B., Perez, R., Forchhammer, K., and Maldener, I.,** *Prokaryotic multicellularity: a nanopore array for bacterial cell communication. FASEB J,* **2013**. *27(6): p. 2293-300* DOI: **10.1096/fj.12-225854.**

45. **Lang, N.J. and Fay, P.,** *The heterocysts of blue-green algae II. Details of ultrastructure. Proc. Roy. Soc. Lond. B,* **1971**. *178: p. 193-203* DOI: **10.1098/rspb.1971.0061**

46. **Wildon, D.C. and Mercer, F.V.,** *The ultrastructure of the heterocyst and akinete of the blue-green algae. Arch Mikrobiol,* **1963**. *47: p. 19-31* DOI: **10.1007/BF00408286.**

47. **Flores, E., Herrero, A., Wolk, C.P., and Maldener, I.,** *Is the periplasm continuous in filamentous multicellular cyanobacteria? Trends Microbiol,* **2006**. *14(10): p. 439-43* DOI: **10.1016/j.tim.2006.08.007.**

48. **Wilk, L., Strauss, M., Rudolf, M., Nicolaisen, K., Flores, E., Kuhlbrandt, W., and Schleiff, E.,** *Outer membrane continuity and septosome formation between vegetative cells in the filaments of Anabaena sp. PCC 7120. Cell Microbiol,* **2011**. *13(11): p. 1744-54* DOI: **10.1111/j.1462-5822.2011.01655.x.**

49. **Meeks, J.C., Elhai, J., Thiel, T., Potts, M., Larimer, F., Lamerdin, J., Predki, P., and Atlas, R.,** *An overview of the genome of Nostoc punctiforme, a multicellular, symbiotic cyanobacterium. Photosynth Res,* **2001**. *70(1): p. 85-106* DOI: **10.1023/A:1013840025518.**

50. **Flores, E. and Herrero, A.,** *Compartmentalized function through cell differentiation in filamentous cyanobacteria. Nat Rev Microbiol,* **2010**. *8(1): p. 39-50* DOI: **10.1038/nrmicro2242.**

51. **Maldener, I., Summers, M.L., and Sukenik, A.,** *Cellular differentiation in filamentous cyanobacteria, in The Cell Biology of Cyanobacteria., E. Flores and A. Herrero, Editors.* **2014**, *Norfolk, UK: Caister Academic Press. 263-291.* ISBN: **978-1-908230-38-6**.

52. **Adams, D.G. and Duggan, P.S.,** *Heterocyst and akinete differentiation in cyanobacteria. New Phytol,* **1999**. *144: p. 3-33* DOI: **10.1046/j.1469-8137.1999.00505.x**

53. **Argueta, C. and Summers, M.L.,** *Characterization of a model system for the study of Nostoc punctiforme akinetes. Arch Microbiol,* **2005**. *183(5): p. 338-46* DOI: **10.1007/s00203-005-0778-5.**

54. **Meeks, J.C., Campbell, E.L., Summers, M.L., and Wong, F.C.,** *Cellular differentiation in the cyanobacterium Nostoc punctiforme. Arch Microbiol,* **2002**. *178(6): p. 395-403* DOI: **10.1007/s00203-002-0476-5.**

55. **Meeks, J.C. and Elhai, J.**, *Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. Microbiol Mol Biol Rev,* **2002***. 66(1): p. 94-121; table of contents* DOI: **10.1128/MMBR.66.1.94-121.2002.**

56. **Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., and Stanier, R.Y.**, *Generic assignments, strain stories and properties of pure cultures of cyanobacteria. J Gen Microbiol,* **1979***. 111: p. 1-61* DOI: **10.1099/00221287-111-1-1.**

57. **Giddings, T.H., Jr., and Staehelin, L.A.**, *Plasma membrane architecture of Anabaena cylindrica: occurrence of microplasmodesmata and changes associated with heterocyst development and the cell cycle. Eur J Cell Biol* **1978***. 16: p. 235-249***.

58. **Nürnberg, D.J., Mariscal, V., Bornikoel, J., Nieves-Morion, M., Krauss, N., Herrero, A., Maldener, I., Flores, E., and Mullineaux, C.W.**, *Intercellular diffusion of a fluorescent sucrose analog via the septal junctions in a filamentous cyanobacterium. MBio,* **2015***. 6(2): p. e02109* DOI: **10.1128/mBio.02109-14.**

59. **Omairi-Nasser, A., Mariscal, V., Austin, J.R., 2nd, and Haselkorn, R.**, *Requirement of Fra proteins for communication channels between cells in the filamentous nitrogen-fixing cyanobacterium Anabaena sp. PCC 7120. Proc Natl Acad Sci U S A,* **2015***. 112(32): p. E4458-64* DOI: **10.1073/pnas.1512232112.**

60. **Omairi-Nasser, A., Haselkorn, R., and Austin, J., 2nd**, *Visualization of channels connecting cells in filamentous nitrogen-fixing cyanobacteria. FASEB J,* **2014***. 28(7): p. 3016-22* DOI: **10.1096/fj.14-252007.**

61. **Flores, E., Herrero, A., Forchhammer, K., and Maldener, I.**, *Septal Junctions in Filamentous Heterocyst-Forming Cyanobacteria. Trends Microbiol,* **2016***. 24(2): p. 79-82* DOI: **10.1016/j.tim.2015.11.011.**

62. **Mariscal, V.**, *Cell-cell joining proteins in heterocyst-forming cyanobacteria, in The cell biology of cyanobacteria, E. Flores and A. Herrero, Editors.* **2014***, Norfolk, UK: Caister Academic Press. 239-304.* ISBN: **9781908230386**.

63. **Maldener, I. and Forchhammer, K.**, *Requirement of Cell Wall Remodelling for Cell-Cell Communication and Cell Differentiation in Filamentous Cyanobacteria of the Order Nostocales, in Biological Nitrogen Fixation, F.J. de Bruijn, Editor.* **2015***: Wiley-Blackwell. 873-878.* ISBN: **978-1-118-63704-3**.

64. **Berendt, S., Lehner, J., Zhang, Y.V., Rasse, T.M., Forchhammer, K., and Maldener, I.**, *Cell wall amidase AmiC1 is required for cellular communication and heterocyst development in the cyanobacterium Anabaena PCC 7120 but not for filament integrity. J Bacteriol,* **2012***. 194(19): p. 5218-27* DOI: **10.1128/JB.00912-12.**

65. **Madigan, M.T., Martinko, J.M., Bender, K.S., Buckley, D.H., and Stahl, D.A.**, *Brock Biology of Microorganisms: Chapter 2.10: Peptidoglycan and Chapter 2.11 LPS: the outer membrane. 14th Edition ed.* **2014**. ISBN: **9780321897398**.

66. **Gram, H.C.**, *Über die isolierte Färbung der Schizomyceten in Schnitt- und Trockenpräparaten. Fortschritte der Medizin,* **1884***. 2: p. 185-89* DOI: **citeulike-article-id:1574067.**

67. **de Pedro, M.A. and Cava, F.**, *Structural constraints and dynamics of bacterial cell wall architecture. Front Microbiol,* **2015***. 6: p. 449* DOI: **10.3389/fmicb.2015.00449.**

68. **Hoiczyk, E. and Baumeister, W.**, *Envelope structure of four gliding filamentous cyanobacteria. J Bacteriol,* **1995***. 177(9): p. 2387-95* PMCID: **PMC176896.**

69. **Hoiczyk, E. and Hansel, A.**, *Cyanobacterial cell walls: news from an unusual prokaryotic envelope. Journal of Bacteriology,* **2000***. 182(5): p. 1191* DOI: **10.1128/JB.182.5.1191-1199.2000**

70. **Hahn, A. and Schleiff, E.**, *The cell envelope, in The cell biology of cyanobacteria, E. Flores and A. Herrero, Editors.* **2014***, Norfolk, UK: Caister Academic Press. 29-88.* ISBN: **9781908230386**.

71. **Yao, X., Jericho, M., Pink, D., and Beveridge, T.**, *Thickness and elasticity of gram-negative murein sacculi measured by atomic force microscopy. J Bacteriol,* **1999***. 181(22): p. 6865-75* PMCID: **PMC94159.**

72. **Vollmer, W., Blanot, D., and de Pedro, M.A.**, *Peptidoglycan structure and architecture. FEMS Microbiol Rev,* **2008***. 32(2): p. 149-67* DOI: **10.1111/j.1574-6976.2007.00094.x.**

73. **Vollmer, W.**, *Structural variation in the glycan strands of bacterial peptidoglycan. FEMS Microbiol Rev,* **2008***. 32(2): p. 287-306* DOI: **10.1111/j.1574-6976.2007.00088.x.**

74. **Schleifer, K.H. and Kandler, O.**, *Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriological reviews,* **1972**. *36(4): p. 407-77* <u>PMCID:</u> **PMC408328.**

75. **Xu, Q., Sudek, S., McMullan, D., Miller, M.D., Geierstanger, B., Jones, D.H., Krishna, S.S., Spraggon, G., Bursalay, B., Abdubek, P., Acosta, C., Ambing, E., Astakhova, T., Axelrod, H.L., Carlton, D., Caruthers, J., Chiu, H.J., Clayton, T., Deller, M.C., Duan, L., Elias, Y., Elsliger, M.A., Feuerhelm, J., Grzechnik, S.K., Hale, J., Han, G.W., Haugen, J., Jaroszewski, L., Jin, K.K., Klock, H.E., Knuth, M.W., Kozbial, P., Kumar, A., Marciano, D., Morse, A.T., Nigoghossian, E., Okach, L., Oommachen, S., Paulsen, J., Reyes, R., Rife, C.L., Trout, C.V., van den Bedem, H., Weekes, D., White, A., Wolf, G., Zubieta, C., Hodgson, K.O., Wooley, J., Deacon, A.M., Godzik, A., Lesley, S.A., and Wilson, I.A.**, *Structural basis of murein peptide specificity of a gamma-D-glutamyl-l-diamino acid endopeptidase. Structure,* **2009**. *17(2): p. 303-13* <u>DOI:</u> **10.1016/j.str.2008.12.008.**

76. **de Jonge, B.L., Chang, Y.S., Gage, D., and Tomasz, A.**, *Peptidoglycan composition of a highly methicillin-resistant Staphylococcus aureus strain. The role of penicillin binding protein 2A. J Biol Chem,* **1992**. *267(16): p. 11248-54* <u>PMCID:</u> **PMID: 1597460**

77. **Boneca, I.G., Huang, Z.H., Gage, D.A., and Tomasz, A.**, *Characterization of Staphylococcus aureus cell wall glycan strands, evidence for a new beta-N-acetylglucosaminidase activity. J Biol Chem,* **2000**. *275(14): p. 9910-8* <u>DOI:</u> **10.1074/jbc.275.14.9910.**

78. **Clarke, A.J. and Dupont, C.**, *O-acetylated peptidoglycan: its occurrence, pathobiological significance, and biosynthesis. Can J Microbiol,* **1992**. *38(2): p. 85-91* <u>PMCID:</u> **PMID: 1521192.**

79. **Vollmer, W., Joris, B., Charlier, P., and Foster, S.**, *Bacterial peptidoglycan (murein) hydrolases. FEMS Microbiol Rev,* **2008**. *32(2): p. 259-86* <u>DOI:</u> **10.1111/j.1574-6976.2007.00099.x.**

80. **van Heijenoort, J.**, *Peptidoglycan hydrolases of Escherichia coli. Microbiol Mol Biol Rev,* **2011**. *75(4): p. 636-63* <u>DOI:</u> **10.1128/MMBR.00022-11.**

81. **Firczuk, M. and Bochtler, M.**, *Folds and activities of peptidoglycan amidases. FEMS Microbiol Rev,* **2007**. *31(6): p. 676-91* <u>DOI:</u> **10.1111/j.1574-6976.2007.00084.x.**

82. **Schindler, C.A. and Schuhardt, V.T.**, *Lysostaphin: A New Bacteriolytic Agent for the Staphylococcus. Proc Natl Acad Sci U S A,* **1964**. *51: p. 414-21* <u>PMCID:</u> **PMC300087.**

83. **Nicholas, R.A., Krings, S., Tomberg, J., Nicola, G., and Davies, C.**, *Crystal structure of wild-type penicillin-binding protein 5 from Escherichia coli: implications for deacylation of the acyl-enzyme complex. J Biol Chem,* **2003**. *278(52): p. 52826-33* <u>DOI:</u> **10.1074/jbc.M310177200.**

84. **Laschtschenko, P.**, *Über die keimtötende und entwicklungshemmende Wirkung von Hühnereiweiß. Zeitschrift für Hygiene und Infektionskrankheiten,* **1909**. *64(1): p. 419-427* <u>DOI:</u> **10.1007/bf02216170.**

85. **Holtje, J.V., Mirelman, D., Sharon, N., and Schwarz, U.**, *Novel type of murein transglycosylase in Escherichia coli. J Bacteriol,* **1975**. *124(3): p. 1067-76* <u>PMCID:</u> **PMC236007.**

86. **Reid, C.W., Blackburn, N.T., Legaree, B.A., Auzanneau, F.I., and Clarke, A.J.**, *Inhibition of membrane-bound lytic transglycosylase B by NAG-thiazoline. FEBS Lett,* **2004**. *574(1-3): p. 73-9* <u>DOI:</u> **10.1016/j.febslet.2004.08.006.**

87. **Oshida, T., Sugai, M., Komatsuzawa, H., Hong, Y.M., Suginaka, H., and Tomasz, A.**, *A Staphylococcus aureus autolysin that has an N-acetylmuramoyl-L-alanine amidase domain and an endo-beta-N-acetylglucosaminidase domain: cloning, sequence analysis, and characterization. Proc Natl Acad Sci U S A,* **1995**. *92(1): p. 285-9* <u>PMCID:</u> **PMC42863.**

88. **de Souza, R.F., Anantharaman, V., de Souza, S.J., Aravind, L., and Gueiros-Filho, F.J.**, *AMIN domains have a predicted role in localization of diverse periplasmic protein complexes. Bioinformatics,* **2008**. *24(21): p. 2423-6* <u>DOI:</u> **10.1093/bioinformatics/btn449.**

89. **Zoll, S., Schlag, M., Shkumatov, A.V., Rautenberg, M., Svergun, D.I., Götz, F., and Stehle, T.**, *Ligand-binding properties and conformational dynamics of autolysin repeat domains in staphylococcal cell wall recognition. J Bacteriol,* **2012**. *194(15): p. 3789-802* <u>DOI:</u> **10.1128/JB.00331-12.**

90. **Rocaboy, M., Herman, R., Sauvage, E., Remaut, H., Moonens, K., Terrak, M., Charlier, P., and Kerff, F.**, *The crystal structure of the cell division amidase AmiC reveals the fold of the AMIN domain, a new peptidoglycan binding domain. Mol Microbiol,* **2013**. *90(2): p. 267-77* <u>DOI:</u> **10.1111/mmi.12361.**

91. *Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., and Bateman, A.*, The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res, **2016**. 44(D1): p. D279-85 <u>DOI</u>: **10.1093/nar/gkv1344.**

92. *Pfam_Database*. Pfam: Family: Amidase_2 (PF01510): Domain organisation. **2016.** Accessed on **2016-02-26**. Available from: <u>http://pfam.xfam.org/family/PF01510#tabview=tab1</u>.

93. *Bernhardt, T.G. and Boer, P.A.J.d.*, The Escherichia coli amidase AmiC is a periplasmic septal ring component exported via the twin-arginine transport pathway. Molecular Microbiology, **2003**. 48(5): p. 1171-1182 <u>DOI</u>: **10.1046/j.1365-2958.2003.03511.x.**

94. *Zoll, S., Pätzold, B., Schlag, M., Götz, F., Kalbacher, H., and Stehle, T.*, Structural basis of cell wall cleavage by a staphylococcal autolysin. PLoS Pathog, **2010**. 6(3): p. e1000807 <u>DOI</u>: **10.1371/journal.ppat.1000807.**

95. *Heilmann, C., Hussain, M., Peters, G., and Götz, F.*, Evidence for autolysin-mediated primary attachment of Staphylococcus epidermidis to a polystyrene surface. Mol Microbiol, **1997**. 24(5): p. 1013-24 <u>DOI</u>: **10.1046/j.1365-2958.1997.4101774.x.**

96. *Park, J.T. and Uehara, T.*, How bacteria consume their own exoskeletons (turnover and recycling of cell wall peptidoglycan). Microbiol Mol Biol Rev, **2008**. 72(2): p. 211-27, table of contents <u>DOI</u>: **10.1128/MMBR.00027-07.**

97. *Prigozhin, D.M., Mavrici, D., Huizar, J.P., Vansell, H.J., and Alber, T.*, Structural and biochemical analyses of Mycobacterium tuberculosis N-acetylmuramyl-L-alanine amidase Rv3717 point to a role in peptidoglycan fragment recycling. J Biol Chem, **2013**. 288(44): p. 31549-55 <u>DOI</u>: **10.1074/jbc.M113.510792.**

98. *Carrasco-Lopez, C., Rojas-Altuve, A., Zhang, W., Hesek, D., Lee, M., Barbe, S., Andre, I., Ferrer, P., Silva-Martin, N., Castro, G.R., Martinez-Ripoll, M., Mobashery, S., and Hermoso, J.A.*, Crystal structures of bacterial peptidoglycan amidase AmpD and an unprecedented activation mechanism. J Biol Chem, **2011**. 286(36): p. 31714-22 <u>DOI</u>: **10.1074/jbc.M111.264366.**

99. *Chou, S., Bui, N.K., Russell, A.B., Lexa, K.W., Gardiner, T.E., LeRoux, M., Vollmer, W., and Mougous, J.D.*, Structure of a peptidoglycan amidase effector targeted to Gram-negative bacteria by the type VI secretion system. Cell Rep, **2012**. 1(6): p. 656-64 <u>DOI</u>: **10.1016/j.celrep.2012.05.016.**

100. *Low, L.Y., Yang, C., Perego, M., Osterman, A., and Liddington, R.C.*, Structure and lytic activity of a Bacillus anthracis prophage endolysin. J Biol Chem, **2005**. 280(42): p. 35433-9 <u>DOI</u>: **10.1074/jbc.M502723200.**

101. *Büttner, F.M., Renner-Schneck, M., and Stehle, T.*, X-ray crystallography and its impact on understanding bacterial cell wall remodeling processes. Int J Med Microbiol, **2015**. 305(2): p. 209-16 <u>DOI</u>: **10.1016/j.ijmm.2014.12.018.**

102. *Dziarski, R. and Gupta, D.*, The peptidoglycan recognition proteins (PGRPs). Genome Biol, **2006**. 7(8): p. 232 <u>DOI</u>: **10.1186/gb-2006-7-8-232.**

103. *Royet, J. and Dziarski, R.*, Peptidoglycan recognition proteins: pleiotropic sensors and effectors of antimicrobial defences. Nat Rev Microbiol, **2007**. 5(4): p. 264-77 <u>DOI</u>: **10.1038/nrmicro1620.**

104. *Dziarski, R. and Gupta, D.*, Review: Mammalian peptidoglycan recognition proteins (PGRPs) in innate immunity. Innate Immun, **2010**. 16(3): p. 168-74 <u>DOI</u>: **10.1177/1753425910366059.**

105. *Wang, Z.M., Li, X., Cocklin, R.R., Wang, M., Fukase, K., Inamura, S., Kusumoto, S., Gupta, D., and Dziarski, R.*, Human peptidoglycan recognition protein-L is an N-acetylmuramoyl-L-alanine amidase. J Biol Chem, **2003**. 278(49): p. 49044-52 <u>DOI</u>: **10.1074/jbc.M307758200.**

106. *Cho, S., Wang, Q., Swaminathan, C.P., Hesek, D., Lee, M., Boons, G.J., Mobashery, S., and Mariuzza, R.A.*, Structural insights into the bactericidal mechanism of human peptidoglycan recognition proteins. Proc Natl Acad Sci U S A, **2007**. 104(21): p. 8761-6 <u>DOI</u>: **10.1073/pnas.0701453104.**

107. *Guan, R., Brown, P.H., Swaminathan, C.P., Roychowdhury, A., Boons, G.J., and Mariuzza, R.A.*, Crystal structure of human peptidoglycan recognition protein I alpha bound to a muramyl pentapeptide from Gram-positive bacteria. Protein Sci, **2006**. 15(5): p. 1199-206 <u>DOI</u>: **10.1110/ps.062077606.**

108. **Guan, R., Wang, Q., Sundberg, E.J., and Mariuzza, R.A.,** *Crystal structure of human peptidoglycan recognition protein S (PGRP-S) at 1.70 A resolution. J Mol Biol,* **2005**. *347(4): p. 683-91* DOI: **10.1016/j.jmb.2005.01.070.**

109. **Guan, R., Roychowdhury, A., Ember, B., Kumar, S., Boons, G.J., and Mariuzza, R.A.,** *Structural basis for peptidoglycan binding by peptidoglycan recognition proteins. Proc Natl Acad Sci U S A,* **2004**. *101(49): p. 17168-73* DOI: **10.1073/pnas.0407856101.**

110. **Guan, R., Malchiodi, E.L., Wang, Q., Schuck, P., and Mariuzza, R.A.,** *Crystal structure of the C-terminal peptidoglycan-binding domain of human peptidoglycan recognition protein Ialpha. J Biol Chem,* **2004**. *279(30): p. 31873-82* DOI: **10.1074/jbc.M404920200.**

111. **Biswas, R., Voggu, L., Simon, U.K., Hentschel, P., Thumm, G., and Götz, F.,** *Activity of the major staphylococcal autolysin Atl. FEMS microbiology letters,* **2006**. *259(2): p. 260-8* DOI: **10.1111/j.1574-6968.2006.00281.x.**

112. **Bourgeois, I., Camiade, E., Biswas, R., Courtin, P., Gibert, L., Götz, F., Chapot-Chartier, M.P., Pons, J.L., and Pestel-Caron, M.,** *Characterization of AtlL, a bifunctional autolysin of Staphylococcus lugdunensis with N-acetylglucosaminidase and N-acetylmuramoyl-l-alanine amidase activities. FEMS Microbiol Lett,* **2009**. *290(1): p. 105-13* DOI: **10.1111/j.1574-6968.2008.01414.x.**

113. **Albrecht, T., Raue, S., Rosenstein, R., Nieselt, K., and Götz, F.,** *Phylogeny of the staphylococcal major autolysin and its use in genus and species typing. J Bacteriol,* **2012**. *194(10): p. 2630-6* DOI: **10.1128/JB.06609-11.**

114. **Schlag, M., Biswas, R., Krismer, B., Kohler, T., Zoll, S., Yu, W., Schwarz, H., Peschel, A., and Götz, F.,** *Role of staphylococcal wall teichoic acid in targeting the major autolysin Atl. Mol Microbiol,* **2010**. *75(4): p. 864-73* DOI: **10.1111/j.1365-2958.2009.07007.x.**

115. **Lützner, N., Pätzold, B., Zoll, S., Stehle, T., and Kalbacher, H.,** *Development of a novel fluorescent substrate for Autolysin E, a bacterial type II amidase. Biochem Biophys Res Commun,* **2009**. *380(3): p. 554-8* DOI: **10.1016/j.bbrc.2009.01.140.**

116. **Pfam_Database.** *Pfam: Family: Amidase_3 (PF01520): Domain organisation.* **2016.** *Accessed on* **2016-02-26**. *Available from: http://pfam.xfam.org/family/PF01520#tabview=tab1.*

117. **Büttner, F.M., Faulhaber, K., Forchhammer, K., Maldener, I., and Stehle, T.,** *Enabling cell-cell communication via nanopore formation: structure, function, and localization of the unique cell wall amidase AmiC2 of Nostoc punctiforme. FEBS J,* **2016**. *283(7): p. 1336–1350* DOI: **10.1111/febs.13673.**

118. **Uehara, T., Dinh, T., and Bernhardt, T.G.,** *LytM-domain factors are required for daughter cell separation and rapid ampicillininduced lysis in Escherichia coli. J Bacteriol* **2009**. *191: p. 5094-5107* DOI: **10.1128/JB.00505-09.**

119. **Uehara, T., Parzych, K.R., Dinh, T., and Bernhardt, T.G.,** *Daughter cell separation is controlled by cytokinetic ring-activated cell wall hydrolysis. EMBO J,* **2010**. *29(8): p. 1412-22* DOI: **10.1038/emboj.2010.36.**

120. **Yang, D.C., Tan, K., Joachimiak, A., and Bernhardt, T.G.,** *A conformational switch controls cell wall-remodelling enzymes required for bacterial cell division. Mol Microbiol,* **2012**. *85(4): p. 768-81* DOI: **10.1111/j.1365-2958.2012.08138.x.**

121. **Kumar, A., Kumar, S., Kumar, D., Mishra, A., Dewangan, R.P., Shrivastava, P., Ramachandran, S., and Taneja, B.,** *The structure of Rv3717 reveals a novel amidase from Mycobacterium tuberculosis. Acta Crystallogr D Biol Crystallogr,* **2013**. *69(Pt 12): p. 2543-54* DOI: **10.1107/S0907444913026371.**

122. **Büttner, F.M., Zoll, S., Nega, M., Götz, F., and Stehle, T.,** *Structure-function analysis of Staphylococcus aureus amidase reveals the determinants of peptidoglycan recognition and cleavage. J Biol Chem,* **2014**. *289(16): p. 11083-94* DOI: **10.1074/jbc.M114.557306.**

123. **Kerff, F., Petrella, S., Mercier, F., Sauvage, E., Herman, R., Pennartz, A., Zervosen, A., Luxen, A., Frere, J.M., Joris, B., and Charlier, P.,** *Specific structural features of the N-acetylmuramoyl-L-alanine amidase AmiD from Escherichia coli and mechanistic implications for enzymes of this family. J Mol Biol,* **2010**. *397(1): p. 249-59* DOI: **10.1016/j.jmb.2009.12.038.**

124. *Rudolf, M., Tetik, N., Ramos-Leon, F., Flinner, N., Ngo, G., Stevanovic, M., Burnat, M., Pernil, R., Flores, E., and Schleiff, E.,* The Peptidoglycan-Binding Protein SjcF1 Influences Septal Junction Function and Channel Formation in the Filamentous Cyanobacterium Anabaena. MBio, **2015**. 6(4): p. e00376 <u>DOI:</u> **10.1128/mBio.00376-15.**

125. *Li, Q., Cheng, W., Morlot, C., Bai, X.H., Jiang, Y.L., Wang, W., Roper, D.I., Vernet, T., Dong, Y.H., Chen, Y., and Zhou, C.Z.,* Full-length structure of the major autolysin LytA. Acta Crystallogr D Biol Crystallogr, **2015**. 71(Pt 6): p. 1373-81 <u>DOI:</u> **10.1107/S1399004715007403.**

126. *Mayer, M.J., Garefalaki, V., Spoerl, R., Narbad, A., and Meijers, R.,* Structure-based modification of a Clostridium difficile-targeting endolysin affects activity and host range. J Bacteriol, **2011**. 193(19): p. 5477-86 <u>DOI:</u> **10.1128/JB.00439-11.**

127. *Wang, N., Hirata, A., Nokihara, K., Fukase, K., and Fujimoto, Y.,* Peptidoglycan Microarray as A Novel Tool to Explore Protein-Ligand Recognition. Biopolymers, **2016** <u>DOI:</u> **10.1002/bip.22807.**

128. *Mellroth, P., Sandalova, T., Kikhney, A., Vilaplana, F., Hesek, D., Lee, M., Mobashery, S., Normark, S., Svergun, D., Henriques-Normark, B., and Achour, A.,* Structural and Functional Insights into Peptidoglycan Access for the Lytic Amidase LytA of Streptococcus pneumoniae. Mbio, **2014**. 5(1) <u>DOI:</u> **10.1128/mBio.01120-13.**

129. Personal Communication, **John Johnson Jr., Ph.D.,** from the The Scripps Research Institute (TSRI) in La Jolla, California, USA, on **2016-02-18** to F.M. Büttner.

130. *Oelker, M.,* Fragment-based screening for drug development or active site mapping in case of three proteins, in Faculty of Science - Interfaculty Institute of Biochemistry. 2015, Eberhard Karls Universität.

131. *Grün, S.,* Reinigung und Charakterisierung der Domäne AMIN-AB von AmiC2 aus Nostoc punctiforme, in Faculty of Science - Interfaculty Institute of Biochemistry. 2014, Eberhard Karls Universität.

132. *Wang, Q., Matsuo, Y., Pradipta, A.R., Inohara, N., Fujimoto, Y., and Fukase, K.,* Synthesis of characteristic Mycobacterium peptidoglycan (PGN) fragments utilizing with chemoenzymatic preparation of meso-diaminopimelic acid (DAP), and their modulation of innate immune responses. Org Biomol Chem, **2016**. 14(3): p. 1013-23 <u>DOI:</u> **10.1039/c5ob02145f.**

133. *Wang, W., Ji, B., and Zhang, Y.,* Chalcogen bond: a sister noncovalent bond to halogen bond. J Phys Chem A, **2009**. 113(28): p. 8132-5 <u>DOI:</u> **10.1021/jp904128b.**

134. *Bleiholder, C., Werz, D.B., Koppel, H., and Gleiter, R.,* Theoretical investigations on chalcogen-chalcogen interactions: what makes these nonbonded interactions bonding? J Am Chem Soc, **2006**. 128(8): p. 2666-74 <u>DOI:</u> **10.1021/ja056827g.**

135. *Politzer, P., Murray, J.S., and Clark, T.,* Halogen bonding and other sigma-hole interactions: a perspective. Phys Chem Chem Phys, **2013**. 15(27): p. 11178-89 <u>DOI:</u> **10.1039/c3cp00054k.**

136. *Murray, J.S., Lane, P., Clark, T., and Politzer, P.,* Sigma-hole bonding: molecules containing group VI atoms. J Mol Model, **2007**. 13(10): p. 1033-8 <u>DOI:</u> **10.1007/s00894-007-0225-4.**

137. *Alam, M., Beevers, R.E., Ceska, T., Davenport, R.J., Dickson, K.M., Fortunato, M., Gowers, L., Haughan, A.F., James, L.A., Jones, M.W., Kinsella, N., Lowe, C., Meissner, J.W., Nicolas, A.L., Perry, B.G., Phillips, D.J., Pitt, W.R., Platt, A., Ratcliffe, A.J., Sharpe, A., and Tait, L.J.,* Synthesis and SAR of aminopyrimidines as novel c-Jun N-terminal kinase (JNK) inhibitors. Bioorg Med Chem Lett, **2007**. 17(12): p. 3463-7 <u>DOI:</u> **10.1016/j.bmcl.2007.03.078.**

# 6. SUPPLEMENT: HALOGEN BONDING

## 6.1. Targeting the gatekeeper MET146 of c-Jun N-terminal kinase 3 (JNK3) induces a bivalent halogen / chalcogen bond

*Andreas Lange, Marcel Günther, **Felix Michael Büttner**, Markus O. Zimmermann, Johannes Heidrich, Susanne Hennig, Stefan Zahn, Christoph Schall, Adrian Sievers-Engler, Francesco Ansideri, Pierre Koch, Michael Laemmerhofer, Thilo Stehle, Stefan A. Laufer, and Frank M. Boeckler, 2015, 137, (46), 14640-14652, DOI: 10.1021/jacs.5b07090 © 2015 American Chemical Society.*

The c-Jun N-terminal kinase 3 (JNK3) is a mitogen-activated protein kinase (MAPK) and therefore an interesting drug target. The "gatekeeper" residue blocks the hydrophobic pocket of kinases and is most commonly a methionine. Threonines, leucines, and phenylalanines are also found in that position. So far, only classical approaches were taken to target the gatekeeper of JNK3 (Met146) by hydrophobic interactions that utilize the flexibility of methionine. Molecular design using computational, synthetic, and biophysical techniques can point to new strategies that involve the sulfur atom of methionine and exemplify the potential and applicability of such. Halogen bonds are one way to increase selectivity and affinity towards a sulfur-containing residue, and the utilization of this interaction may help to find new potent compounds.

Halogen bonds exploit the anisotropic electron distribution of members of the halogen group, which leads to patches of positive electrostatic potential ("σ-holes"). Those σ-holes can be addressed by non-bonding valence electrons ("lone pairs") of binding partners.
Similarly to halogen bonds, chalcogen bonds harness the anisotropic electron distribution of heavier members of the oxygen group ("chalcogens") that lead to σ-holes, which again can be addressed by lone pairs of binding partners [133-136].

We investigated halogen bonding in quantum mechanical (QM) studies and generated an existing chlorinated pyrimidine derivative [137] and analogs of it. The chlorine was substituted with hydrogen, bromine, and iodine. In addition to the *in silico* studies, biophysical

characterization by a Fluorescence Polarization assay (FP) and Isothermal Titration Calorimetry (ITC) of the interaction between the different compounds and JNK3 were used to gain experimental evidence for the positive effects of halogen bonding. The mutation of the gatekeeper methionine to alanine, leucine, or threonine emphasizes its importance for compound affinity. The effects of the mutations also give hints for selectivity towards JNK3 and show that other effects are able to compensate for the loss of the halogen bond.

However, some results were somewhat contradictory because usage of bromine and iodine did not result in a higher affinity as the QM studies had suggested. A complex structure of JNK3 with the iodine compound was generated and solved at 1.95 Å resolution to elucidate the reasons for this discrepancy. The iodine compound exhibits well defined electron density in both copies in the asymmetric unit and binds to JNK3 in the same way as the chlorine compound that was previously reported (PDB-ID 2P33, [137]). Hydrogen bonds with the hinge region and a lysine as well as the halogen bond with the gatekeeper methionine tightly bind the ligand.

Careful inspection of the complex structure furthermore unveiled the rotameric rearrangement of another methionine (Met115) near the binding site, which was not observed in the chlorine complex. The sulfur atoms of both methionines are found in good distance and angles for a chalcogen bond.

The combination of the interactions with the hinge region and the mixed halogen-chalcogen bond result in a somewhat restricted arrangement of the ligand in the JNK3 binding site. As the distances between the halogen and the sulfur of the gatekeeper consequently cannot vary too much, the full potential of the larger σ-holes for bromine and iodine is not exploitable. This restriction explains the inconsistency between biophysical measurements and initial QM calculations.

As only 39 % of human kinases have a methionine as gatekeeper and merely 9.3 % exhibit the second methionine in proximity for chalcogen bonding, our findings contribute to the development of kinase inhibitors with increased selectivity.

## 6.2. Outlook

Although it is not trivial, in order to gain favorable contributions from halogen bonding one needs to have precise distances and angles between donor and acceptor. Very subtle changes of only 20‑30 pm or minor angle shifts by a few degrees can change the interaction from attractive to repulsive.

The exchange of halogen atoms should not only be a way to increase affinity, it can also be utilized to optimize selectivity of a compound towards its target. Following up on this point: a carefully engineered compound may trigger the formation of an additional chalcogen bond as seen for the complex in the PDB entry 4X21. Other kinds of interactions, ranging from hydrophobic to ionic, can also arise depending on the individual environments and improve a compound's qualities.

While important, the adaptive behavior of a binding site is hard to predict and different in each case. Still, the combination of screening halogen compound libraries with QM-based algorithms is a beneficial strategy to exploit halogen bonds in compound optimization.

# Appendix

## List of figures and tables

# Curriculum vitae

| | |
|---|---|
| Name | Felix Michael Büttner |
| Date of birth | |
| Place of birth | |
| Nationality | |

| | |
|---|---|
| 10-2012 – 06-2016 | **Doctoral studies**, research, and thesis; Prof. Thilo Stehle, Tübingen |
| 10-2011 – 06-2012 | **Diploma thesis**; Prof. Thilo Stehle, Tübingen. "Structural analysis of bacterial amidase AmiA from *Staphylococcus aureus* and its substrate, the peptidoglycan component N-acetylmuramoyl-tetrapeptide" |
| 10-2006 – 06-2012 | **Biochemistry studies** at Eberhard Karls Universität Tübingen |
| <u>final exams:</u> | Biochemistry, Pharmaceutical Chemistry, and Pharmaceutical Biology |
| <u>internships:</u> | RNA trafficking; Prof. Jansen, Tübingen |
| | Synthesis of Kinase inhibitors; Prof. Laufer, Tübingen |
| | Herpes Virus inhibitors; Prof. Coen & Prof. Hogle, Harvard, Boston |
| 2006 | **Abitur**, at Hohenzollern Gymnasium, Sigmaringen |

Conferences:

| | |
|---|---|
| 2011 | Retreat of the CRC-TRR34, *Münster* |
| 2012 | Pathophysiology of Staphylococci in the post-genomic era, *Banz*, **poster** |
| 2013 | Rhine-Knee Regional Meeting on Biocrystallography, *Schluchsee*, **talk** |
| 2014 | Retreat of the CRC766, *Freudenstadt*, **talk** and **poster** |
| 2014 | Evaluation of the CRC-TRR34, *Greifswald*, **poster** |
| 2014 | Pathogen-Host Interactions at Cellular Barriers, *Münster*, **poster** |
| 2014 | Rhine-Knee Regional Meeting on Biocrystallography, *St. Otile* |
| 2015 | Evaluation of the CRC766, *Tübingen*, **poster** |
| 2015 | PhD-retreat of the CRC-TRR34, *Würzburg*, **talk** |
| 2015 | Rhine-Knee Regional Meeting on Biocrystallography, *Engelberg* |
| 2016 | CCP4 study weekend, *Nottingham* |
| 2016 | Assembly and Function of Bacterial Surfaces from the Molecular to Intercellular Scales *Gordon Research Seminar and Conference, Mt. Snow*, **poster** |
| 2016 | PhD-retreat of the CRC-TRR34, *Münster*, **talk** |

## Acknowledgments

## Publications

**Felix Michael Büttner**, Sebastian Zoll, Mulugeta Nega, Friedrich Götz, and Thilo Stehle. *Structure-Function Analysis of Staphylococcus aureus Amidase Reveals the Determinants of Peptidoglycan Recognition and Cleavage*. J. Biol. Chem. 2014 289: 11083-11094. First Published on March 5, 2014, doi:10.1074/jbc.M114.557306 © American Society for Biochemistry and Molecular Biology.

**Felix Michael Büttner**, Katharina Faulhaber, Karl Forchhammer, Iris Maldener and Thilo Stehle (2016), Enabling cell–cell communication via nanopore formation: structure, function and localization of the unique cell wall amidase AmiC2 of *Nostoc punctiforme*. FEBS J, 283: 1336–1350. doi:10.1111/febs.13673.
License number: 3800611430780

Reprinted from: International Journal of Medical Microbiology, *The Bacterial Cell Envelope: Structure, Function, and Infection Interface*, Volume 305, Issue 2, February 2015, **Felix Michael Büttner**, Michaela Renner-Schneck, Thilo Stehle, *X-ray crystallography and its impact on understanding bacterial cell wall remodeling processes,* Pages 209-216, © 2015, with permission from Elsevier.
License number: 3751270540797

Reprinted with permission from: *Targeting the Gatekeeper MET146 of C-Jun N-Terminal Kinase 3 Induces a Bivalent Halogen/Chalcogen Bond,* Andreas Lange, Marcel Günther, **Felix Michael Büttner**, Markus O. Zimmermann, Johannes Heidrich, Susanne Hennig, Stefan Zahn, Christoph Schall, Adrian Sievers-Engler, Francesco Ansideri, Pierre Koch, Michael Laemmerhofer, Thilo Stehle, Stefan A. Laufer, and Frank M. Boeckler, Journal of the American Chemical Society, 2015, 137, (46), 14640-14652, DOI: 10.1021/jacs.5b07090, © 2015, American Chemical Society.

# Structure-Function Analysis of *Staphylococcus aureus* Amidase Reveals the Determinants of Peptidoglycan Recognition and Cleavage*

Felix Michael Büttner[‡1], Sebastian Zoll[‡1], Mulugeta Nega[§], Friedrich Götz[§], and Thilo Stehle[‡¶2]

From the ‡Interfaculty Institute of Biochemistry, University of Tübingen, Hoppe-Seyler-Strasse 4, 72076 Tübingen, Germany, the §Microbial Genetics, Interfaculty Institute of Microbiology and Infection Medicine, University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany, and the ¶Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee 37232

**Background:** Autolysins ensure the plasticity of bacterial cell walls, and deletion leads to impaired cell clusters.
**Results:** High resolution structures of *Staphylococcus aureus* amidase AmiA shed light on peptidoglycan binding and cleavage.
**Conclusion:** AmiA distinguishes peptidoglycan mostly by the peptide, and cleavage is facilitated by a zinc-activated water molecule.
**Significance:** These structures will inform strategies to develop new therapeutics against MRSA.

The bifunctional major autolysin AtlA of *Staphylococcus aureus* cleaves the bacterium's peptidoglycan network (PGN) at two distinct sites during cell division. Deletion of the enzyme results in large cell clusters with disordered division patterns, indicating that AtlA could be a promising target for the development of new antibiotics. One of the two functions of AtlA is performed by the *N*-acetylmuramyl-L-alanine amidase AmiA, which cleaves the bond between the carbohydrate and the peptide moieties of PGN. To establish the structural requirements of PGN recognition and the enzymatic mechanism of cleavage, we solved the crystal structure of the catalytic domain of AmiA (AmiA-cat) in complex with a peptidoglycan-derived ligand at 1.55 Å resolution. The peptide stem is clearly visible in the structure, forming extensive contacts with protein residues by docking into an elongated groove. Less well defined electron density and the analysis of surface features indicate likely positions of the carbohydrate backbone and the pentaglycine bridge. Substrate specificity analysis supports the importance of the pentaglycine bridge for fitting into the binding cleft of AmiA-cat. PGN of *S. aureus* with L-lysine tethered with D-alanine via a pentaglycine bridge is completely hydrolyzed, whereas PGN of *Bacillus subtilis* with *meso*-diaminopimelic acid directly tethered with D-alanine is not hydrolyzed. An active site mutant, H370A, of AmiA-cat was completely inactive, providing further support for the proposed catalytic mechanism of AmiA. The structure reported here is not only the first of any bacterial amidase in which both the PGN component and the water molecule that carries out the nucleophilic attack on the carbonyl carbon of the scissile bond are present; it is also the first peptidoglycan amidase complex structure of an important human pathogen.

Gram-positive, spherical staphylococci arrange in clusters and colonize the skin or mucous membranes, causing severe infections in infants, the elderly, transplantation patients, and people suffering from immunocompromising diseases, such as AIDS. Staphylococci are among the main causes of hospital-acquired infections (1). *Staphylococcus aureus* and *Staphylococcus epidermidis* are the most abundant human pathogens of the *Staphylococcus* genus. Both species form a biofilm that protects them from the human immune system (2) and antibiotics as well as contributing to persistent infections (3). In the case of *S. epidermidis*, this multilayered polysaccharide matrix (4) is responsible for infections of patients with implants, such as intravasal catheters, prostheses, or pacemakers (2, 5). This may require implant replacement and cause severe complications for the affected patients (6). The biofilm of *S. aureus* primarily serves as protection but also contributes to its pathogenicity (7). *S. aureus* is responsible for a large number of life-threatening infections that can result in diseases, such as endocarditis, meningitis, pneumonia, septicaemia, and toxic shock syndrome (8).

Resistance against *S. aureus* is on the rise, posing a serious threat to human health. There is therefore an urgent need for the development of new antibiotics to control emerging methicillin-resistant and vancomycin-resistant *S. aureus* strains (MRSA[3] and VRSA, respectively). Worldwide numbers are not available, but with about 132,000 cases in Germany per year (9), hospital-acquired MRSA currently accounts for ~20% of all staphylococcus infections (10), whereas in the early 1990s, the MRSA fraction was only ~1% (11). In high risk areas, such as

[3] The abbreviations used are: MRSA and VRSA, methicillin-resistant and vancomycin-resistant *S. aureus*, respectively; AtlA, bifunctional major autolysin of *S. aureus*; AmiA, amidase of *S. aureus*; AmiE, amidase of *S. epidermidis*; AmiA-cat and AmiE-cat, catalytic domain of AmiA and AmiE, respectively; AmiD, amidase of *E. coli*; NAGase, glucosaminidase of *S. aureus*; PGN, peptidoglycan network; MtetP, muramyl tetrapeptide (MurNAc-L-Ala-D-iGln-L-Lys(NHAc)-D-Ala-NH₂); MTP, muramyl tripeptide (L-Ala-D-iGlu-L-Lys); anhydro-MTP, 1,6-anhydro-MTP; MurNAc, *N*-acetylmuramic acid; D-iGln, D-isoglutamine; D-iGlu, D-isoglutamic acid; Wat, water; *meso*-DAP, *meso*-diaminopimelic acid; NAc, *N*-acetyl.

intensive care units, the MRSA infection rate increases up to 37% (12), causing 5,000 deaths and leading to additional costs of ~380 million € (9) per year in Germany alone. In the United States, annual *S. aureus* infections have reached 475,000, 275,000 of which are MRSA-related, with $1–10 billion in extra expenses for the health care system and ~11,000 to ~19,000 deaths (13, 14). Targeting staphylococcal enzymes, critical for survival and growth of the bacterium, represents an attractive strategy for the development of new antibiotics.

Several hydrolytic enzymes ensure the plasticity of the staphylococcal cell wall by processing the complex PGN network. One of these, the major autolysin AtlA, is composed of two enzymes with hydrolytic activity (an amidase (AmiA) and a glucosaminidase (NAGase)) that cleave PGN at different locations (15). In the precursor AtlA protein, the two catalytic functions (cat) are each linked to targeting repeats (R1–R3) and also connected to a propeptide and a signal peptide (Fig. 1A). Posttranslational processing generates active AmiA and NAGase, which both localize at the septal region (16), where they process staphylococcal PGN during cell growth and division. *S. aureus* AtlA deletion mutants show a severely impaired phenotype that is unable to proliferate, forming large cell clusters instead (17). These findings demonstrate the essential function of AtlA in the *S. aureus* life cycle and also highlight a therapeutic potential for specific inhibition of AtlA.

The staphylococcal Atl domain organization is highly conserved in all of the species, with the amidase being the most conserved domain. It has been shown that the Atl-based phylogenetic tree correlates well with the corresponding 16 S rRNA- and core genome-based tree and represents a useful tool for staphylococcal genus and species typing (18).

The catalytic domain of AmiA (referred to here as AmiA-cat) is a zinc-dependent amidase that cleaves the amide bond between the peptide stem and carbohydrate backbone of PGN (16, 19, 20). The previously proposed mechanism for hydrolysis of the lactyl-alanine bond was based on *in silico* docking studies of the homologous catalytic domain AmiE from *S. epidermidis* (19). Further data from structures of a homologous protein originate from *Escherichia coli* (21). Consequently, detailed structural information on amidase-PGN interaction in Gram-positive bacteria is limited to date.

To determine the specificity of recognition and the mechanism of catalysis of AmiA-cat, we determined crystal structures of the enzyme in the absence (Fig. 1B) and presence (Fig. 2A) of the muramyltetrapeptide MurNAc-L-Ala-D-iGln-L-Lys-NHAc-D-Ala-NH$_2$ (MtetP), a ligand that includes the previously characterized minimal ligand for the *S. epidermidis* amidase (muramyltripeptide) (19). Both structures were solved to high resolution, and they unambiguously establish the specificity of interaction as well as the reaction mechanism used by this essential cell wall enzyme. Our results form an excellent basis for the design of new antibiotic lead structures.

## EXPERIMENTAL PROCEDURES

*Molecular Biology*—The cDNA coding for AmiA-cat (residues 199–421) was cloned into a pGEX-4-3T vector for expression. The expressed protein contains an N-terminal GST tag fused to AmiA-cat with a six-amino acid thrombin-cleav-

able linker. Active site mutants were created using site-directed mutagenesis as described in the QuikChange® protocol (22).

*Protein Expression and Purification*—Proteins were expressed in *E. coli* BL21 (DE3). After induction, cultures were incubated for 72 h at 20 °C. Harvested cells were then resuspended in buffer (150 mM NaCl, 50 mM Tris, pH 8.0) supplemented with PMSF and Roche Applied Science Complete protease inhibitor mix. Filtered cell lysate was loaded onto a 5-ml GSTrap FF column (GE Healthcare). 100 units of thrombin were added for on-column overnight cleavage at 20 °C and release of the fusion protein. Size exclusion chromatography removed the remaining small impurities and aggregates from the protein. Purity was confirmed by SDS-PAGE and MALDI-MS.

*Protein Crystallization*—AmiA-cat crystals belong to space group C2 and contain two protomers in the asymmetric unit, giving rise to a solvent content of 41.2%. Crystals were grown using the hanging drop vapor diffusion method at 20 °C. 1 μl of protein solution (11 mg/ml) was mixed with 1 μl of a well solution containing 0.1 M MES/imidazole buffer at pH 6.5 and a mix of sodium formate, ammonium acetate, sodium citrate, racemic sodium/potassium tartrate, and sodium oxalate at concentrations of 0.02 M each as well as 12.5% (w/v) PEG 1000, 12.5% (w/v) PEG 3350, and 12.5% (w/v) 2-methyl-2,4-pentanediol (Molecular Dimensions). Microseeding improved crystal quantity and quality. In order to obtain catalytically inactive enzyme, the AmiA-cat crystals were first incubated for 72 h in well solution supplemented with 20 mM EDTA to remove the active site zinc ion. For complex formation, crystals were next soaked for 60 h in well solution containing 20 mM EDTA and 20 mM MtetP. Crystals yielding a complex belong to space group P2$_1$ with four protomers in the asymmetric unit and similar solvent content. All crystals could be directly flash-frozen in liquid nitrogen because the well solutions contained sufficient cryoprotectant.

*X-ray Diffraction*—All data were collected at 100 K on PILATUS detectors using synchrotron radiation at beamlines X06DA and X06SA of the Swiss Light Source in Villigen, Switzerland.

*Structure Determination*—Indexing, integrating, and scaling were done with the XDS software package (23). Molecular replacement for all AmiA-cat structures was performed with PHASER (24, 25). Initial phases for AmiA-cat were determined with an AmiE search model (Protein Data Bank accession code 3LAT, 81% identity). The refined AmiA-cat structure was then used to solve the ligand structure by molecular replacement. Model building, refinement, and validation were performed with Coot (26, 27), the CCP4 suite (28–31), Phenix (32), and the MolProbity Web page (33). The coordinate and parameter files for MtetP were obtained from the PRODRG2 server (34). Simulated annealing (Phenix) was performed to remove model bias for the ligand structure, especially in the active site. The ligand was then added manually and modeled according to difference or simulated annealing omit density maps in Coot. The final structures contain almost all of the 223 residues, with the exception of 13–15 poorly ordered N-terminal and 1–4 C-terminal amino acids. Figures were generated with PyMOL (35), and electrostatic potentials were calculated with PBD2PQR and APBS 2.1 (36), implemented in PyMOL.

*Purification of Peptidoglycan*—PGN was isolated from *S. aureus* SA113 (37) or *Bacillus subtilis* ATCC 6051 using the method of de Jonge *et al.* (38) with some modifications. Briefly, cells were grown to $A_{578\,nm}$ of 0.6 and harvested by centrifugation at $3,000 \times g$ for 30 min, boiled with 5% SDS for 30 min, and broken with glass beads. Insoluble PGN was harvested by centrifugation at $30,000 \times g$ for 30 min and washed several times with lukewarm water to remove SDS. Broken cell walls were suspended in 100 mM Tris-HCl, pH 7.2, treated with 10 $\mu$g/ml DNase (Sigma) and 50 $\mu$g/ml RNase (Sigma) for 2 h, and subsequently treated with 100 $\mu$g/ml trypsin for 16 h at 37 °C. To remove wall teichoic acid, the PGN preparations were incubated with 48% hydrofluoric acid for 24 h at 4 °C while stirring. PGN was harvested by centrifugation at $30,000 \times g$ for 30 min and washed several times with water until complete removal of hydrofluoric acid. The final PGN product was lyophilized.

*Preparation and HPLC/MS Analysis of PGN*—Purified PGN (5 mg) was resuspended in 1 ml of 25 mM sodium phosphate buffer (pH 6.8) and digested with mutanolysin for 16 h at 37 °C. The enzyme reaction was stopped by boiling the sample for 5 min at 95 °C, and insoluble contaminants were removed by centrifugation. 50 $\mu$l of AmiA-cat, AmiA-H370A, or AmiA gel filtration buffer as a negative control were added to 100 $\mu$l of mutanolysin-digested PGN and incubated overnight at 37 °C while stirring. HPLC separation of digestion products was carried out on a reversed-phase column (Poroshell 120 EC-C18 $4.6 \times 150$ mm, 2.7 $\mu$m; Agilent Technologies, Waldbronn, Germany) fitted with a poroshell EC-C18, 4.6-mm guard column using an Agilent 1200 system operating ChemStation software.

HPLC was performed using a linear 150-min gradient from 100% HPLC-buffer A (100 mM sodium phosphate, pH 2.2, and 5% methanol) to 100% HPLC-buffer B (100 mM sodium phosphate, pH 2.8, and 30% methanol) with a column temperature of 52 °C. Detection was made at 205 nm. Samples were prepared by mixing 100 $\mu$l of the product with an equal volume of 0.5 M sodium borate buffer (pH 8.0) containing freshly dissolved sodium borohydride (10 mg/ml) and reduced for 30 min at room temperature. The reaction was stopped, and excess borohydride was deactivated by lowering the pH to <3 using 20% phosphoric acid. 100 $\mu$l of the prepared sample was injected.

LC/MS analysis of these samples was performed with an Agilent HPLC-electrospray ionization-MS system (LC/MSD Ultra Trap System XCT 6330), using a gradient of A ($H_2O$ with 0.1% formic acid) and B (0.06% formic acid in acetonitrile) as follows: 0–10% B:A over 25 min, 10% B to 27 min, and 100% B to 30 min at 0.4 ml min$^{-1}$ at 40 °C on a Nucleosil 100 C18 3 $\mu$m column ($100 \times 2$-mm inner diameter) with a precolumn ($10 \times 2$-mm inner diameter, Dr. Maisch, Ammerbuch, Germany). Detection of $m/z$ values consistent with AmiA-cat digestion products was conducted using Agilent Data Analysis for 6300 Series Ion Trap LC/MS 6.1 version 3.4 software (Bruker-Daltonik GmbH).

## RESULTS

*Overall Structures of Unliganded and MtetP-bound AmiA-cat*—The structures of unliganded and liganded AmiA-cat were determined at high resolution (1.12 and 1.55 Å, respectively, Table 1), allowing us to analyze their salient features with confidence. The proteins adopt a globular, mixed $\alpha/\beta$ fold that is

highly similar to that of the previously reported unliganded AmiE-cat structure (19). Essential features of this fold are seven $\alpha$-helices that surround a central six-stranded $\beta$-sheet (Fig. 1*B*). The rear of the $\beta$-sheet is shielded by two long $\alpha$-helices, whereas its front is solvent-accessible and forms the bottom of a recessed groove. In the case of unliganded AmiA-cat, one end of this groove accommodates a zinc ion that is required for catalysis. Residues His-265, His-370, and Asp-384 directly coordinate the zinc ion (Fig. 1*B*), and nearby residues Glu-324 and His-382 are positioned to participate in catalysis. To produce a complex with MtetP, the zinc ion was first removed by treatment of the AmiA-cat crystals with EDTA, followed by extensive soaking with MtetP (see "Experimental Procedures"). Ligand binding was confirmed by a simulated annealing omit density map that displays unbiased electron density for MtetP over the entire tetrapeptide and as far as the *N*-acetylmuramic acid (MurNAc) moiety of the ligand. The tetrapeptide portion of MtetP is well defined by electron density (Fig. 2*B*), showing that this part of the ligand binds in an extended conformation to the recessed groove. The lactate moiety of MurNAc linking the peptide to the carbohydrate as well as the cyclic MurNAc moiety are less well ordered, suggesting higher mobility due to fewer defined interactions. These differences in mobility are reflected in the *B*-factor distribution of the ligand (Fig. 2*B*). The MtetP-bound AmiA-cat crystals contain four molecules in their asymmetric unit, only two of which contain fully occupied ligand binding sites. The flexible, cyclic MurNAc moiety presented sufficient electron density for model building in both copies. The binding sites of the remaining two molecules are less accessible due to smaller solvent channels. As a result, they contain electron density features that suggest only partially bound ligand.

With the exception of the active site region, unliganded and liganded AmiA-cat exhibit no structural differences. The two structures can be superimposed with a root mean square deviation of 0.2 Å (DaliLite pairwise (39)) onto each other, demonstrating that ligand binding does not induce larger structural changes.

*Interactions of AmiA-cat with MtetP*—Inspection of the structure of AmiA-cat in complex with MtetP reveals a sophisticated interaction network that significantly extends knowledge derived from earlier *in silico* docking (19). The tetrapeptide backbone of MtetP is anchored into the binding groove along its entire length via several direct or water-mediated hydrogen bonds to protein residues (Fig. 3). In addition, all four amino acid side chains of MtetP (L-Ala, D-iGln, L-Lys-NHAc, and D-Ala-NH$_2$) are also involved in individual contacts with the protein, accounting for specificity. These interactions are detailed in Fig. 3 and summarized below.

The methyl side chain of the first amino acid of MtetP, L-Ala, inserts into a small hydrophobic pocket formed by residues Ala-288 and Val-290 (Fig. 3, *A* and *C*). The methyl group of L-Ala is also only 4.4 Å away from the C$\alpha$ carbon of the conserved Gly-311. Lack of a side chain at position 311 contributes to specificity because even a medium sized side chains would lead to clashes with the L-Ala group.

The short amide side chain of D-isoglutamine (D-iGln), which is a key determinant of ligand binding (20), projects into a shal-

**TABLE 1**

**Data collection and refinement statistics for AmiA-cat (Protein Data Bank code 4KNK) and liganded AmiA-cat (Protein Data Bank code 4KNL)**

Values for the highest resolution bin are given in parentheses.

| Parameters | AmiA-cat | Liganded AmiA-cat |
|---|---|---|
| Beamline | X06DA (PX III) | X06SA (PX I) |
| Space group | C2 | P2$_1$ |
| Cell dimensions (Å) | $a$ = 96.9, $b$ = 81.7, $c$ = 68.7 | $a$ = 67.7, $b$ = 82.8, $c$ = 77.6 |
| | $\alpha = \gamma = 90°$. $\beta = 128.8°$ | $\alpha = \gamma = 90°$. $\beta = 95.9°$ |
| Wavelength (Å) | 1.00000 | 1.00605 |
| Detector | Pilatus 2M | Pilatus 6M |
| Resolution (Å) | 50-1.12 (1.15-1.12) | 50.0-1.55 (1.59-1.55) |
| Measured reflections | 885,445 (14,326) | 442,448 (27,121) |
| Unique reflections | 146,370 (6,937) | 122,742 (8,927) |
| $R_{meas}$ (%) | 4.7% (60.3%) | 6.5% (82.3%) |
| Completeness (%) | 92.6% (59.7%) | 99.5% (98.1%) |
| Redundancy | 6.0 (2.0) | 3.6 (3.0) |
| $I/\sigma$ ($I$) | 21.1 (1.8) | 13.9 (1.7) |
| CC1/2 (%) | 100 (71.7) | 99.9 (69.7) |
| Resolution (Å) | 40.9-1.12 | 48.2-1.55 |
| $R_{work}/R_{free}$ | 12.09/13.89 | 17.12/19.81 |
| No. of atoms | 3,894 | 7,321 |
| Protein | 3,318 | 6,531 |
| Water | 535 | 630 |
| MtetP | | 102 |
| Zn$^{2+}$ | 2 | |
| Others | 39 | 58 |
| *B*-Factors (Å$^2$) | | |
| Wilson | 13.7 | 24.5 |
| Mean | 16.9 (7.9–41.8) | 25.9 (11.9–69.9) |
| Protein | 15.1 (7.9–38.8 ) | 25.3 (11.9–69.9) |
| Water | 27.9 (9.0–41.8) | 33.1 (14.8–55.9) |
| MtetP | | 36.7 (18.0–61.5) |
| Zn$^{2+}$ | 12.5 (12.2–12.8) | |
| Others | 27.8 (11.0–38.7) | 37.0 (18.8–53.0) |
| Root mean square deviations | | |
| Bond length (Å) | 0.009 | 0.012 |
| Bond angles (degrees) | 1.389 | 1.380 |
| Ramachandran plot | | |
| Most favorable (%) | 97.4 | 97.6 |
| Allowed (%) | 2.6 | 2.4 |
| Disallowed (%) | 0.0 | 0.0 |

low pocket and forms direct hydrogen bonds with the hydroxyl group of Thr-380 and N$\epsilon$ of His-370, as well as two water-mediated hydrogen bonds to Asp-381 (Fig. 3, *B* and *C*). The side chain of His-382 seals the pocket on one side, whereas Thr-380 is located at the other end.

The side chain of the third amino acid, L-Lys, is acetylated (L-Lys-NHAc) in MtetP to approximate the physiologic state of this residue in PGN, where it is linked to additional glycine residues (see Fig. 7*A*). Specificity of recognition derives primarily from hydrophobic interactions between the aliphatic side chain of L-Lys-NHAc and the side chain of Trp-310. The entire lysine side chain lies parallel to the large indole ring of Trp-310, with distances below 4 Å (Fig. 3*B*). In addition, a water molecule (Wat-7) bridges the carbonyl oxygen of D-Ala and N$\epsilon$ of L-Lys-NHAc, thus helping to fix the orientation of the lysine side chain. Similarly, the peptide oxygen of Lys-NHAc forms a water-mediated (Wat-1) hydrogen bond with Asn-317.

The fourth residue, D-Ala-NH$_2$, is engaged in backbone hydrogen bond formation, but its side chain is at least partially exposed to solvent. The terminal amide group of D-Ala-NH$_2$ participates in a network of water-mediated hydrogen bonds that help to fix its orientation with respect to the protein and the preceding L-Lys-NHAc side chain (Fig. 3, *B* and *C*). These interactions would still be possible in the context of either a pentaglycine bridge or a free carboxyl terminus.

The carbohydrate moiety of MtetP engages in fewer interactions with AmiA-cat than the peptide moiety, resulting in its higher flexibility. The *N*-acetyl oxygen of MurNAc forms an intramolecular hydrogen bond with the L-Ala nitrogen and via Wat-4 water-mediated contacts to AmiA-cat. The NAc-methyl group inserts in the hydrophobic portion (Ala-288, Val-290, and Phe-293) of the otherwise hydrophilic binding pocket (Fig. 3, *A* and *C*). Only Thr-267, Asn-269, and Glu-277 directly bond with MurNAc. Interestingly, the anomeric carbon is in $\alpha$-configuration instead of $\beta$-configuration.

*Active Site*—The bound zinc ion in the unliganded structure marks the center of the active site of the enzyme (Fig. 1*B*). The ion is coordinated by the side chains of His-265, His-370, and Asp-384, with a water (Wat-10) completing its almost perfect tetrahedral coordination sphere. Although the zinc had to be removed to prepare the complex with MtetP, its position can be reliably inferred from superposition. Inspection of the unliganded and liganded structures shows two water molecules (Wat-9 and Wat-10) in each case, at most shifted by 1 Å. Wat-9 is located next to the position of the zinc ion, as well as Wat-10. Coordination of Wat-9 involves His-370, Asp-384, and Wat-10 (Fig. 3*A*). A fourth hydrogen bond is formed between Wat-9 and the carbonyl group of the scissile amide bond between L-Ala and the lactate of MurNAc in the ligand (Fig. 3, *A* and *C*). The high mobility or partial occupancy of Wat-9, indicated by a high *B*-factor, make a vital role unlikely. However, it may support stabilization of intermediate states during catalysis. The scissile bond is positioned directly adjacent to Wat-10, suggesting that the reaction mechanism proceeds by a Wat-10-medi-
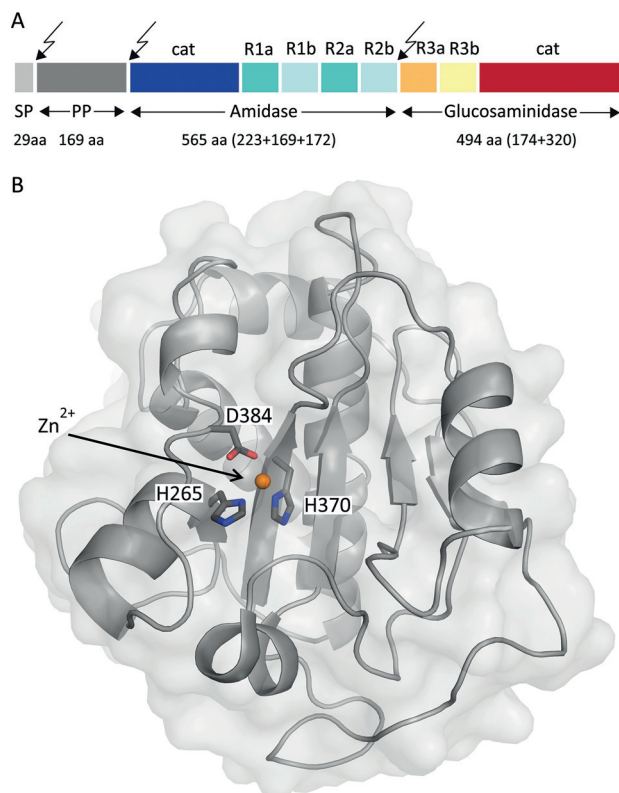
FIGURE 1. **Prepro-AtlA holoenzyme and structure of AmiA-cat.** *A*, domain arrangement of AtlA with sites of post-translational cleavage indicated by *arrows*. *SP*, signal peptide; *PP*, propeptide; *cat*, catalytic domain; *R*, repeat domain. This figure was adapted and modified from Ref. 19. *B*, mixed $\alpha/\beta$ fold of unliganded AmiA-cat in a schematic representation with a *transparent surface*. The zinc ion (*orange*) and its coordinating residues (*dark gray*) are highlighted.

ated nucleophilic attack. Two residues, Asp-266 and Glu-324, lie next to Wat-10 and probably serve to enhance its nucleophilicity, thus favoring an attack on the lactyl amide bond. The carbonyl atom of the scissile amide bond of MtetP is hydrogen-bonded to His-382, which helps to orient the lactic acid moiety in the active site. We note that the His-382 side chain has different orientations in the unliganded and liganded AmiA-cat structures.

*Putative Reaction Mechanism*—The architecture of the active site and the observed interactions between protein and substrate are very much consistent with a reaction mechanism in which Wat-10 attacks the scissile bond. This water is hydrogen-bonded to the Asp-266 carbonyl and Glu-324 carboxyl groups, which would lead to both hydrogens of Wat-10 facing toward these residues and the free electron pairs of the Wat-10 oxygen facing toward zinc and the bond connecting the lactyl and peptide moieties (Fig. 4*A*). The likely role of the zinc ion is to polarize the oxygen of Wat-10, rendering it more reactive. The remaining free electron pair of the Wat-10 oxygen would then be able to perform a nucleophilic attack on the scissile bond of MtetP. His-382 could provide a proton for stabilization of the resulting oxyanion, whereas Glu-324 could accept a hydrogen from Wat-10 (Fig. 4*B*). Stabilization of the tetrahedral intermediate would involve the zinc ion and N$\delta$ of His-382, which could each interact with one of the resulting hydroxyl groups as well as the side chains of Asp-266 and Glu-324. In the next step, the

tetrahedral intermediate again forms a carbonyl group but with the peptide moiety as the leaving group rather than the previously attacking water molecule, leading to the separation of the carbohydrate and peptide moieties of PGN (Fig. 4*C*). At the same time, the N terminus of the new peptide is poised to accept a hydrogen from Glu-324, whereas His-382 can accept a hydrogen atom from the tetrahedral intermediate. Finally, the cleavage products are released from the active site (Fig. 4*D*).

*Implications for PGN Binding and Cleavage*—The MtetP compound is a substrate for AmiA-cat, which cleaves large PGN structures in its physiologic setting. It is possible and indeed likely that additional contacts between PGN components and AmiA-cat exist and that the interactions between PGN and AmiA-cat are somewhat more complex than depicted here. Nevertheless, analysis of surface properties provides at least some clues as to how AmiA-cat would engage components of PGN that extend beyond the muramyltetrapeptide (*i.e.* the MurNAc-GlcNAc glycan polymer and the pentaglycine bridge) (Fig. 5*A*).

A number of residues largely conserved among bacterial amidases (Fig. 5, *B* and *C*) define the spacious carbohydrate binding pocket. Thr-267 and Glu-277 together with Met-281 and Phe-293 form the bottom of the pocket, whereas Tyr-280 on one side and Ala-268, Asn-269, and Ser-273 on the other side enclose the carbohydrate and form the lateral edges. Analysis of the complex shows that MurNAc does not engage in many specific contacts on its own. Based on chemical and geometric restraints, we therefore modeled a plausible conformation of GlcNAc-MurNAc-GlcNAc in the carbohydrate pocket (Fig. 5*A*), which is concurrent with a previously proposed three-dimensional structure of PGN (40) and nicely follows the carbohydrate groove. Steric constraints lead to a minor shift of MurNAc out of the binding pocket in the presence of $\beta$-1,4-linked GlcNAc molecules. Hydrogen bonds formed with O1 of MurNAc fall away in this model, whereas hydrophobic interactions and hydrogen bonds of the NAc moiety remain. Nevertheless, only the oxygen atoms O6 of preceding and succeeding GlcNAc may each form additional hydrogen bonds (preceding GlcNAc with His-382 or Wat-5 and succeeding GlcNAc with Gly-276, Glu-277, and Wat-66; data not shown). Physiologically, weaker interactions with MurNAc and the glycan strand make sense because the enzyme achieves its specificity through engagement of the tetrapeptide stem and must dissociate from the PGN after cleavage has occurred.

Investigation of the electrostatic surface of AmiA-cat gives two plausible orientations in terms of uncharged contact area for the pentaglycine bridge linked to L-Lys in PGN (Fig. 5*A*), although C$\epsilon$ and N$\epsilon$ would not superimpose with the complex. In both cases, glycines would be able to loosely interact with conserved (Fig. 5*B*) and uncharged AmiA-cat surface residues. Previous data showed that the presence or absence of the glycine bridge has little effect on catalysis (19, 20), indicating that interactions of the glycines with AmiA-cat contribute little binding energy. Nonetheless, the pentaglycine bridge is most likely important for specificity.

*Substrate Specificity*—In order to relate the structural data to functional experiments, we investigated wild-type AmiA-cat and an active site mutant AmiA-H370A for its ability to digest
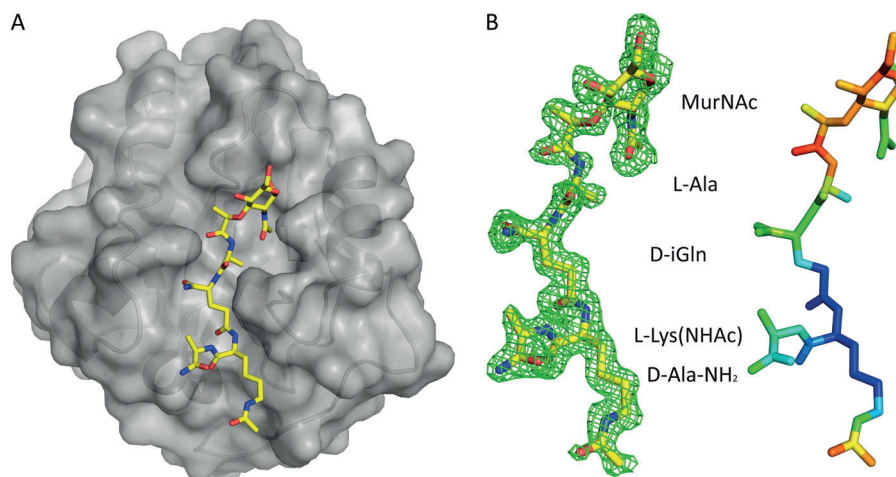
**FIGURE 2. AmiA-cat in complex with MtetP.** *A*, AmiA-cat (*semitransparent gray surface*) bound MtetP (*yellow sticks*) in the active site. *B*, the *close-up* on MtetP surrounded by omit density illustrates that the ligand is well defined. Still, the *B*-factor distribution of MtetP displays the elevated flexibility of its MurNAc moiety. The difference omit map is shown at a σ level of 2.0, and the *color scale* for *B*-factors ranges from *blue* (20 Å²) to *red* (50 Å²).
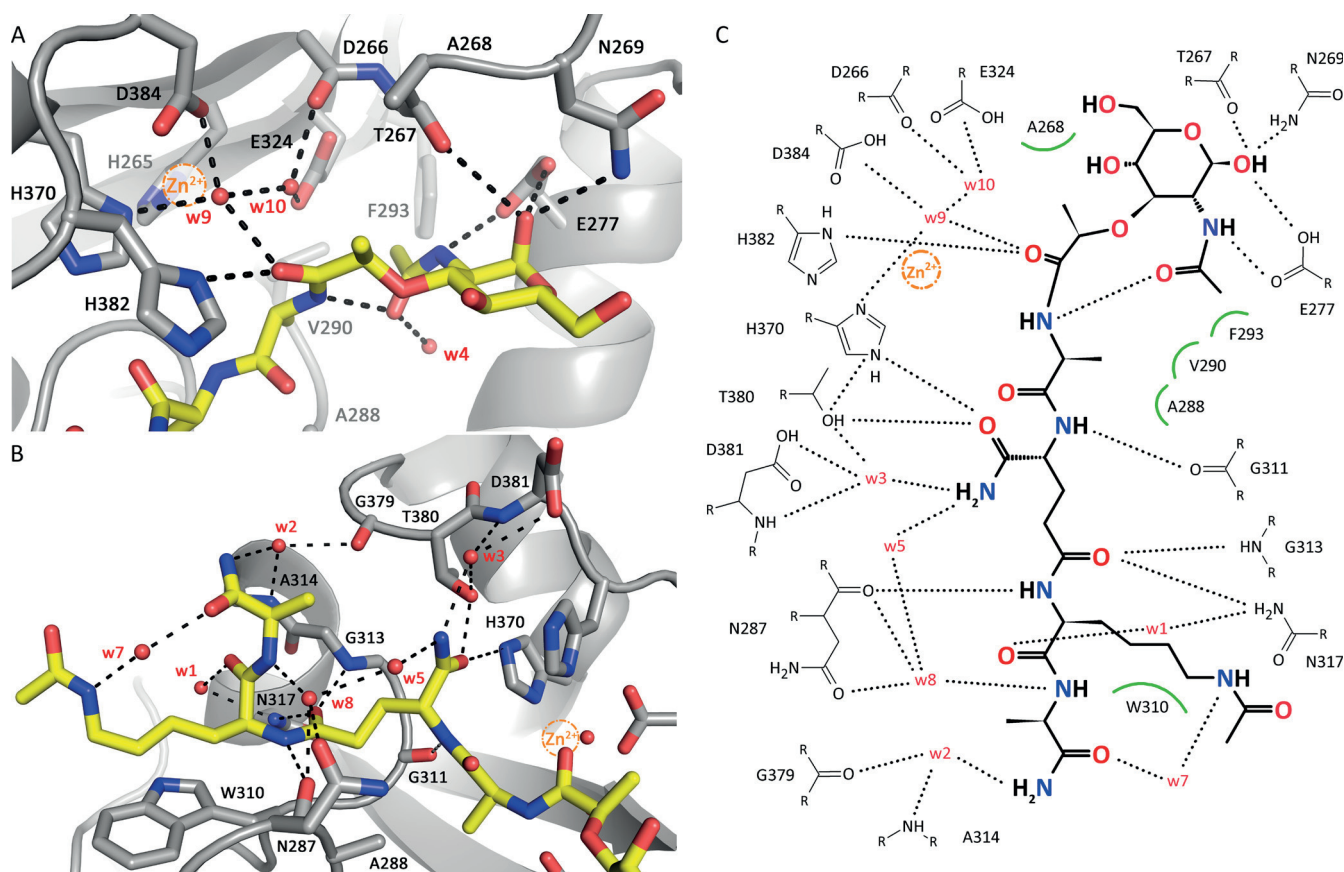


**FIGURE 3. Interactions between AmiA-cat and MtetP.** *A*, interactions of AmiA-cat with the MurNAc moiety and ʟ-Ala of MtetP at the active site. The zinc ion from the native structure (*orange sphere*) is superimposed on the AmiA-cat complex (*gray*) with MtetP (*yellow*). *B*, *close-up* of the interactions of the peptide moiety of MtetP with AmiA. *C*, ChemSketch (57) plot of interactions between MtetP and AmiA-cat. Van-der-Waals interactions are depicted as *green arcs*, and hydrogen bonds are shown as *black dashed lines*. Coordination of Wat-9 involves His-370, Asp-384, and Wat-10 and a hydrogen bond with the carbonyl oxygen of the scissile amide bond in the ligand, which itself is positioned by interaction with His-382. Wat-10 lies next to Asp-266, Glu-324, Wat-9, and the carbonyl carbon of the scissile bond. MurNAc forms an intramolecular and four further hydrogen bonds with Glu-277, Thr-267, and Asn-269. Hydrophobic interactions of the methyl groups involve Ala-268 and Phe-293, respectively. ʟ-Ala inserts into a small hydrophobic pocket formed by residues Ala-288 and Val-290. The amide side chain of ᴅ-iGln forms direct hydrogen bonds with Thr-380 and His-370 as well as two water-mediated hydrogen bonds to Asp-381. The carbonyl oxygen bonds with Gly-313 and Asn-317. ʟ-Lys is stabilized by interactions with Asn-287 as well as Asn-317, whereas the acetylated side chain engages in hydrophobic interactions with Trp-310. A water (Wat-7 (*w7*)) bridges Nε of ʟ-Lys-NHAc and the carbonyl oxygen of ᴅ-Ala, which engages additional water-mediated interactions with AmiA-cat residues Asn-287, Ala-314, and Gly-379.

PGN from *S. aureus* as well as *B. subtilis*. Both PGN structures were first predigested with mutanolysin, which cleaves the glycosidic bond between GlcNAc and MurNAc, to facilitate the

digestion. AmiA-cat completely hydrolyzed *S. aureus* PGN, whereas the AmiA-H370A mutant was inactive (Fig. 6, *A–C*). Analysis by HPLC shows that the main product generated by
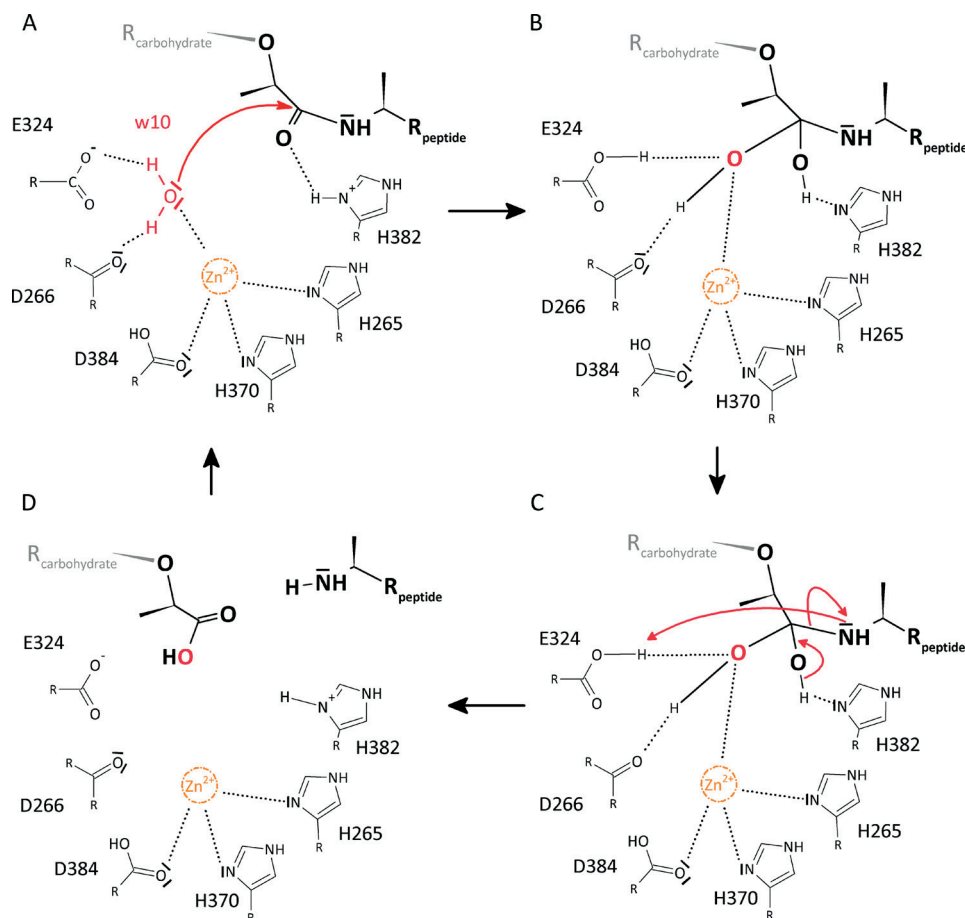
FIGURE 4. **Proposed reaction mechanism of AmiA.** *A*, Wat-10 (*w10*) is hydrogen-bonded to Asp-266 and Glu-324, and its free electron pairs face toward the scissile bond. $Zn^{2+}$ is complexed by His-265, His-370, and Asp-384 and probably renders Wat-10 more reactive, enabling a nucleophilic attack. *B*, tetrahedral intermediate is stabilized by hydrogen bonds of the resulting hydroxyl groups with Nδ of His-382 as well as Asp-266, Glu-324, and zinc, respectively. *C*, reformation of a carbonyl group with the peptide moiety as leaving group. His-382 can accept a hydrogen atom from the tetrahedral intermediate, whereas the peptide is poised to accept a hydrogen from Glu-324. *D*, product release.

AmiA-cat is the disaccharide GlcNAc-MurNAc, with a mass of $m/z$ 496 (Fig. 6, *B* and *F*). Interestingly, the intensity of the GlcNAc-MurNAc peak (Fig. 6*B*) represents the sum of the PGN oligomers seen in Fig. 6*A*, indicating that >95% of the PGN substrate was digested.

In contrast to *S. aureus*, *B. subtilis* PGN could not be hydrolyzed by AmiA-cat (Fig. 6, *D* and *E*). The major differences in PGN structure of *S. aureus* and *B. subtilis* are illustrated in Fig. 7. *B. subtilis* incorporates D-isoglutamic acid into the peptide linking the glycan chains (Fig. 7*B*), whereas *S. aureus* converts this amino acid to D-iGln (Fig. 7*A*). It is, however, unlikely that this difference is solely responsible for the inability of AmiA-cat to cleave *B. subtilis* PGN because the homologous AmiE enzyme still hydrolyzed synthetic substrates composed of MurNAc-L-Ala-D-iGlu-L-Lys, albeit with lower efficiency (20). We consider it more likely that two other alterations in the PGN structures are responsible for the observed difference in activity. The *B. subtilis* PGN carries a *meso*-diaminopimelic acid (*meso*-DAP), which has an amidated ε-carboxylate and a free α-carboxylate and also features a directly cross-linked peptide stem (Fig. 7*B*), without the pentaglycine bridge found in the *S. aureus* PGN. These two major differences probably prevent a proper fit of the *B. subtilis* PGN structure into the binding cleft of AmiA-cat.

## DISCUSSION

We have determined high resolution structures of the catalytic region of the *S. aureus* amidase AmiA in its unliganded form and in complex with a compound that includes MurNAc and the tetrapeptide L-Ala-D-iGln-L-Lys-D-Ala. To crystallize the wild-type enzyme in complex with ligand while avoiding cleavage, it proved critical to first remove the catalytic zinc ion through extensive incubation with EDTA and then soak crystals with ligand, whose high concentrations reflect the environment in PGN. Analysis of the two structures provides insights into the parameters that govern specificity as well as the catalytic mechanism.

AmiA-cat folds into a compact structure with a long, exposed ligand binding groove that appears ideally suited to access its specific cleavage sites within the dense, highly cross-linked PGN structure. In addition to the catalytically active AmiA-cat domain, the mature enzyme also contains four repeat domains (Fig. 1*A*). Analysis of the highly homologous *S. epidermidis* amidase has shown that these repeats probably anchor the protein to lipoteichoic acid protruding from the staphylococcal cell wall (41). Importantly, the repeats are flexibly linked to the catalytically active domain, allowing cleavage to proceed efficiently (41). The position of the bound MtetP ligand sheds light
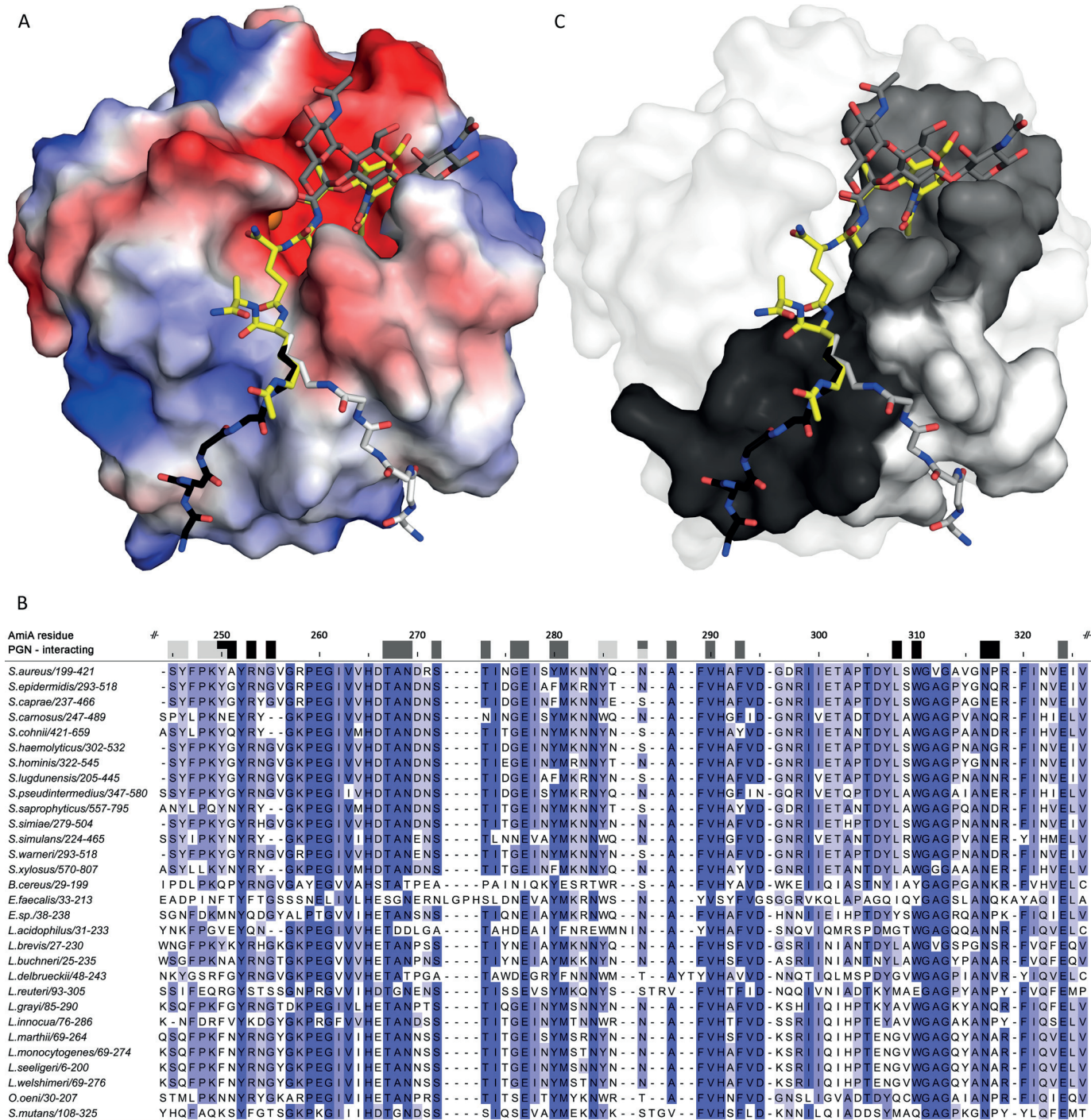
**FIGURE 5. AmiA-cat binding to PGN.** *A*, electrostatic surface of AmiA-cat with MtetP (*yellow*), zinc (*orange*), and modeled PGN components (*dark gray*, *black*, and *light gray*). Uncharged (*white surface area*), positively charged (*blue areas*), and negatively charged residues (*red surface*) are shown. The spacious hydrophilic pocket harboring the zinc ion and active site also accommodates MurNAc. Adjacent GlcNAc rings (*dark gray sticks*) shift MurNAc slightly when modeled as a polymer. The lower peptide moiety of MtetP binds in the mostly uncharged region of the binding cleft. Two conformations for the pentaglycine bridge (*black* and *light gray sticks*, respectively) linked to L-Lys of MtetP were modeled according to uncharged surface area and possible hydrogen bonds. *B*, a multisequence alignment of bacterial amidases with identical residues *colored* by conservation from *light* to *dark blue*. Residues forming the carbohydrate binding pocket, marked by *dark gray boxes*, are highly conserved among all compared amidases. Surface-exposed amino acids near the two pentaglycine bridges are marked in *black* and *light gray* for the respective conformation models. Conservation, especially among staphylococci is high for both. Alignment was calculated using Clustal Omega (58), and the output was created using Jalview (59). *C*, conserved residues mapped on the AmiA-cat surface according to the *color scheme* used in *A* and *B*.

on the interaction of amidases with more complex, branched PGN fragments. Thus, the structure of liganded AmiA-cat allows us to visualize how the enzyme acts in a physiologic setting. In addition to the groove that accommodates the tetra-peptide of MtetP, AmiA-cat contains surface features that probably allow for binding of MurNAc-GlcNAc polymers at one end and a pentaglycine bridge at the other end of the groove. The peptide composition of PGN in staphylococci is
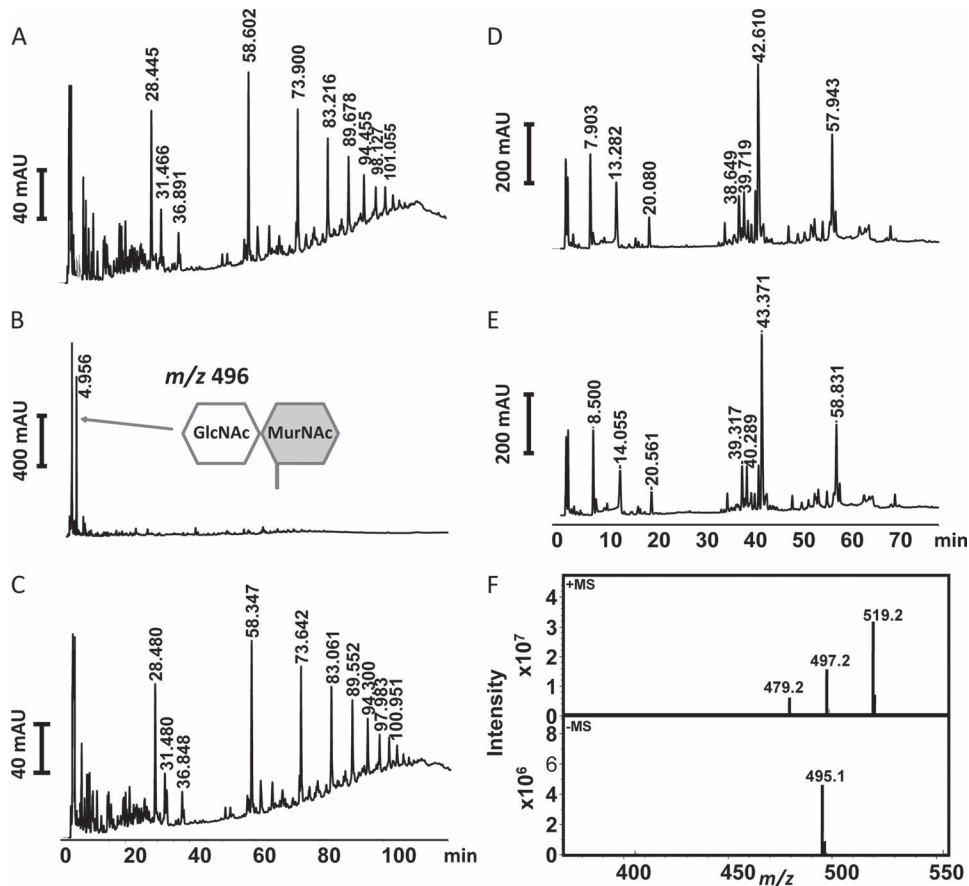
FIGURE 6. **RP-HPLC profile of mutanolysin-digested PGN fragments after incubation with enzyme.** *A, S. aureus* PGN fragments incubated with elution buffer of AmiA-cat as control. Three peaks around 28 min contain monomer species of one peptide stem with carbohydrates, peaks at 58 min comprise species with two peptide stems, peaks at 73 min cover fragments with three peptide stems, and continuing accordingly. *B, B. subtilis* PGN fragments incubated with AmiA-cat result in completely digested polymers with GlcNAc-MurNAc fragments remaining. *C, S. aureus* PGN fragments incubated with active site mutant AmiA-H370A show no catalytic activity. *D, B. subtilis* PGN fragments incubated with elution buffer of AmiA-cat result in a pattern similar to *A* but with specific retention times for *B. subtilis*. *E, B. subtilis* PGN fragments incubated with AmiA-cat exhibit no activity of the staphylococcal amidase for *B. subtilis* PGN. *F,* mass analysis of the major product shown in *B*, which corresponds to the disaccharide GlcNAc-MurNAc (*m/z* 496).
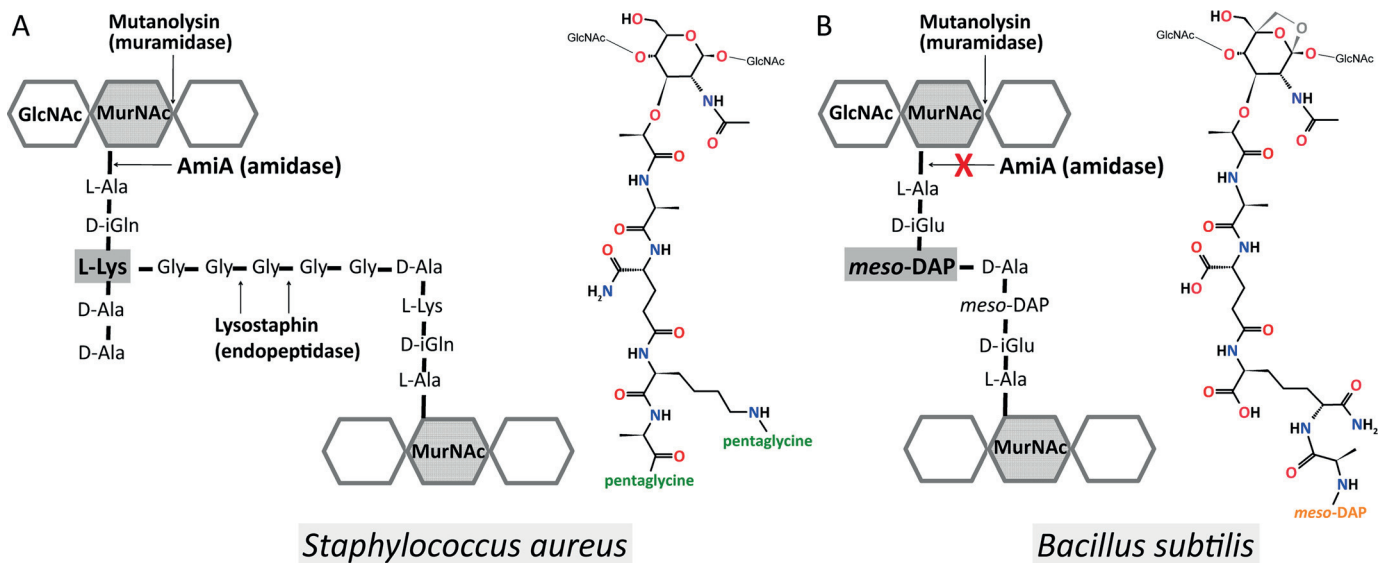


FIGURE 7. **Schematic and molecular illustration of the PGN structures.** *A,* in *S. aureus*, the PGN subunit is composed of D-isoglutamine and L-lysine and is cross-linked via a pentaglycine bridge (43). It comprises C-terminally either D-Ala-D-Ala or D-Ala-pentaglycine. *B, B. subtilis* PGN from vegetative cells (53, 55), however, contains D-isoglutamic acid and amidated *meso*-DAP and lacks a D-alanine at the *meso*-DAP α-carboxyl group, and it is directly cross-linked with D-alanine of the next subunit. In addition, *B. subtilis* PGN has 1,6-anhydro-MurNAc, shown in *gray*, at its terminus instead of a reducing MurNAc.

well conserved, typically using a pentaglycine bridge that connects D-Ala of one peptide stem with the Nε from L-Lys of another stem (42–44). Variations in *S. aureus* and among subspecies exist (43, 44) but are rare. Also, they concentrate on the variation of the pentaglycine bridge in terms of length and composition. The number of glycines may vary from four to six, or one glycine may be substituted by a serine or an alanine (43, 44). Although structures of several bacterial amidases have been determined (19, 45–52), the amidase AmiD from *E. coli* is the only catalytically active amidase for which structural data of an uncleaved ligand-enzyme complex have been available prior to this work (21).

The PGN of *E. coli* and *B. subtilis* is highly similar (42) and belongs to the type of variation A1γ (43). It differs in amino acid composition (D-iGln and L-Lys are substituted with D-iGlu and *meso*-DAP, respectively) as well as direct cross-linkage of amino acids 3 and 4 (no interpeptide bridge) from the PGN of *S. aureus* (42), which has a PGN of the A3α type of variation (43). Additionally, *E. coli* and *B. subtilis* PGN has anhydro-MurNAc at the terminus of the carbohydrate backbone (42, 53), whereas PGN of *S. aureus* always presents reducing MurNAc (54). Further differences within the *E. coli* and the *B. subtilis* A1γ type of variation include deviations in the presence or absence of an unlinked, terminal D-Ala and amidation of *meso*-DAP carboxyl groups (53, 55).

The inability to digest *B. subtilis* PGN presumably is a cumulative effect because D-iGlu instead of D-iGln is tolerated (20), and a *meso*-DAP, which is amidated at its second carboxyl group, itself is flexible enough to be accommodated at the lysine binding site of AmiA-cat. In addition, Asn-317 could partially compensate for a negative charge resulting from the absent, unlinked, terminal D-Ala of *B. subtilis* PGN. However, the direct cross-linkage is most likely to cause a clash with AmiA-cat due to a high rigidity and bulky side chain compared with a pentaglycine bridge. Since each of these differences is likely tolerated on its own, we suggest that the combined differences account for substrate specificity.

Although the *E. coli* enzyme AmiD and AmiA both belong to the amidase 2 family and are zinc-dependent, the two enzymes share a sequence identity of only 21%, have different substrates, and exhibit large structural differences (Fig. 8), including differences in the active site shown by a 2.0-Å root mean square deviation of α-carbon atoms (Dali server (56)). For example, the proton donor in AmiD is a lysine instead of a histidine residue (His-382 in AmiA). His-370 and Thr-380 that stabilize D-iGln in AmiA are replaced by a histidine and an arginine, which are better suited to bond with D-iGlu from A1γ-PGN variants. The free carboxylate of *meso*-DAP or D-Ala could be stabilized by another arginine of AmiD. Interestingly, this is not the case in the two available complex structures of AmiD (cocrystallized with tripeptide L-Ala-D-iGlu-L-Lys (MTP) and soaked with anhydro-MurNAc-L-Ala-D-iGlu-L-Lys (anhydro-MTP)) due to crystal contacts. Also, lysine was used as the third amino acid of the peptide stem instead of *meso*-DAP. Because anhydro-MurNAc does not occur in *S. aureus* PGN and the carbohydrate binding pocket of AmiD is smaller than in AmiA, one can draw only limited conclusions for binding and catalysis of staphylococcal PGN from these structures. Moreover, the
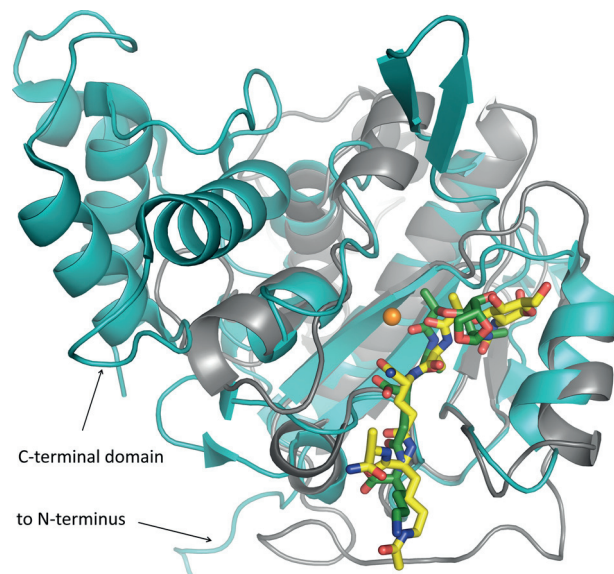


FIGURE 8. **Comparison of the AmiA-cat complex (*gray schematic*) with AmiD from *E. coli* (*cyan schematic*; Protein Data Bank code 3D2Y).** Superimposition of the two bacterial amidases reveals a similar fold solely around the binding and active site. AmiD deviates by a 2.0-Å root mean square deviation (DaliLite pairwise) from the AmiA-cat main chain, amino acids in the active and binding sites differ, and AmiD contains additional motifs at its N and C termini. Ligand positioning of anhydro-MTP (*green sticks*) to AmiD is comparable with the AmiA-cat complex with MtetP (*yellow sticks*). However, anhydro-MTP has an overall shift in relation to MtetP, and interactions of enzyme with ligand are unlike. Additionally, the MurNAc moiety, including the scissile bond, lies in the direct vicinity of the zinc binding residues and is in the anhydro form, which does not occur in staphylococci. $Zn^{2+}$ from the unliganded AmiA-cat structure in *orange* was superimposed.

AmiD complex structure of anhydro-MTP lacks space for both the zinc ion and a water molecule in the active site, which could be activated by the zinc ion and attack the scissile bond. Thus, the structure of AmiA-cat from *S. aureus* presented here provides new information both about how the ligand is contacted and also about how catalysis is performed.

The presented results also significantly extend prior knowledge obtained from *in silico* molecular docking of a tripeptide into the highly homologous amidase structure of *S. epidermidis*, AmiE-cat (19). Although the overall model and the orientation of the ligand agree with our structure, the ligand used for docking AmiE-cat lacked the C-terminal D-Ala and the acetyl group at the L-Lys side chain. The latter resulted in a non-physiologic positive charge due to the protonated Nε of the lysine. Although some of the predicted hydrogen bonds and hydrophobic interactions could be validated by our complex structure, the majority of interactions identified in this work were not projected. Furthermore, the docking model did not take the nucleophilic water in the active site into account, causing the substrate to locate closer to the zinc ion. Additionally, a non-physiological zinc ion caused the backbone to shift at residue 63 (Ala-268 in AmiA-cat), making the carbohydrate binding pocket smaller.

Inspection of the liganded AmiA-cat structure directly suggests strategies for inhibition. A competitive inhibitor of AmiA should contain the tetrapeptide stem of MtetP because all four residues make strong interactions with the enzyme. D-Ala, however, contributes least and could therefore be omitted or replaced. The main contact maintained by L-Lys is the hydro-

phobic interaction with Trp-310. L-Ala and especially D-iGln are conserved in staphylococcal PGN and should be part of a lead structure. Changes in these residues are not tolerated by the enzyme (19) and would therefore not lead to a competitive compound. A reasonable approach would be the introduction of an noncleavable, peptide-mimicking group at the position of the L-Ala-lactyl peptide bond. Suitable bioisosteres are hydroxyethylene or dihydroxyethylene that imitate the tetragonal transition state. Modification of the carbohydrate (*e.g.* at C1 to form a salt bridge with Glu-277) would also increase affinity for AmiA-cat. On the other hand, it may be advantageous to exchange the sugar moiety and end the lead structure with a non-cleavable lactate analog linked to a cyclic compound because carbohydrate synthesis is expensive.

In conclusion, we have solved the complex structure of AmiA-cat with its ligand, the PGN component MtetP, at high resolution. Inspection of the interactions observed in the complex leads to a plausible model for catalysis and allows us to predict the binding locations of the adjacent GlcNAc molecules and the pentaglycine bridge. We show that the crystallized enzyme is active and that it possesses a narrow specificity that enables it to digest PGN structures from *S. aureus* but not from *B. subtilis*. The complex structure offers plausible explanations for these differences in catalytic activity. Our work moreover provides new data on the mechanism of the crucial amidase reaction that may help in the development of therapeutics against MRSA.

**REFERENCES**

1. Baron, S. (1996) in *Medical Microbiology* (Baron, S., ed) 4th Ed., Chapter 12, University of Texas, Galveston, TX
2. O'Gara, J. P., and Humphreys, H. (2001) *Staphylococcus epidermidis* biofilms: importance and implications. *J. Med. Microbiol.* **50,** 582–587
3. Götz, F. (2002) Staphylococcus and biofilms. *Mol. Microbiol.* **43,** 1367–1378
4. Heilmann, C., Schweitzer, O., Gerke, C., Vanittanakom, N., Mack, D., and Götz, F. (1996) Molecular basis of intercellular adhesion in the biofilm-forming *Staphylococcus epidermidis*. *Mol. Microbiol.* **20,** 1083–1091
5. Götz, F. G. P. (2000) in *Infections Associated with Indwelling Medical Devices* (Waldvogel, F. A., and Bisno, A. L., eds) pp. 55–88, American Society for Microbiology Press, Washington, D. C.
6. Hall-Stoodley, L., Costerton, J. W., and Stoodley, P. (2004) Bacterial biofilms: from the natural environment to infectious diseases. *Nat. Rev. Microbiol.* **2,** 95–108
7. Archer, N. K., Mazaitis, M. J., Costerton, J. W., Leid, J. G., Powers, M. E., and Shirtliff, M. E. (2011) *Staphylococcus aureus* biofilms: properties, regulation, and roles in human disease. *Virulence* **2,** 445–459
8. Lowy, F. D. (1998) *Staphylococcus aureus* infections. *N. Engl. J. Med.* **339,** 520–532
9. Köck, R., Becker, K., Cookson, B., van Gemert-Pijnen, J. E., Harbarth, S., Kluytmans, J., Mielke, M., Peters, G., Skov, R. L., Struelens, M. J., Tacconelli, E., Navarro Torné, A., Witte, W., and Friedrich, A. W. (2010) Methicillin-resistant *Staphylococcus aureus* (MRSA): burden of disease and control challenges in Europe. *Euro Surveill.* **15,** 19688
10. European Antimicrobial Resistance Surveillance System (2009) *EARSS Annual Report 2008*, pp. 55–58, European Centre for Disease Prevention and Control, Bilthoven, The Netherlands
11. Kresken, M., Hafner, D., Schmitz, F.-J., and Wichelhaus, T. A. (2009) Resistenzsituation bei klinisch wichtigen Infektionserregern gegenüber Antibiotika in Deutschland und im mitteleuropäischen Raum. Bericht über die Ergebnisse einer multizentrischen Studie der Arbeitsgemeinschaft Empfindlichkeitsprüfungen & Resistenz der Paul-Ehrlich-Gesellschaft für Chemotherapie e.V. aus dem Jahre 2007, p. 9, Antiinfectives Intelligence, Rheinbach, Germany
12. Kohlenberg, A., Schwab, F., Geffers, C., Behnke, M., Rüden, H., and Gastmeier, P. (2008) Time-trends for Gram-negative and multidrug-resistant Gram-positive bacteria associated with nosocomial infections in German intensive care units between 2000 and 2005. *Clin. Microbiol. Infect.* **14,** 93–96
13. Klein, E., Smith, D. L., and Laxminarayan, R. (2007) Hospitalizations and deaths caused by methicillin-resistant *Staphylococcus aureus*, United States, 1999–2005. *Emerg. Infect. Dis.* **13,** 1840–1846
14. Zeller, J. L., Burke, A. E., and Glass, R. M. (2007) JAMA patient page. MRSA infections. *JAMA* **298,** 1826
15. Biswas, R., Voggu, L., Simon, U. K., Hentschel, P., Thumm, G., and Götz, F. (2006) Activity of the major staphylococcal autolysin Atl. *FEMS Microbiol. Lett.* **259,** 260–268
16. Schlag, M., Biswas, R., Krismer, B., Kohler, T., Zoll, S., Yu, W., Schwarz, H., Peschel, A., and Götz, F. (2010) Role of staphylococcal wall teichoic acid in targeting the major autolysin Atl. *Mol. Microbiol.* **75,** 864–873
17. Heilmann, C., Hussain, M., Peters, G., and Götz, F. (1997) Evidence for autolysin-mediated primary attachment of *Staphylococcus epidermidis* to a polystyrene surface. *Mol. Microbiol.* **24,** 1013–1024
18. Albrecht, T., Raue, S., Rosenstein, R., Nieselt, K., and Götz, F. (2012) Phylogeny of the staphylococcal major autolysin and its use in genus and species typing. *J. Bacteriol.* **194,** 2630–2636
19. Zoll, S., Pätzold, B., Schlag, M., Götz, F., Kalbacher, H., and Stehle, T. (2010) Structural basis of cell wall cleavage by a staphylococcal autolysin. *PLoS Pathog.* **6,** e1000807
20. Lützner, N., Pätzold, B., Zoll, S., Stehle, T., and Kalbacher, H. (2009) Development of a novel fluorescent substrate for Autolysin E, a bacterial type II amidase. *Biochem. Biophys. Res. Commun.* **380,** 554–558
21. Kerff, F., Petrella, S., Mercier, F., Sauvage, E., Herman, R., Pennartz, A., Zervosen, A., Luxen, A., Frère, J. M., Joris, B., and Charlier, P. (2010) Specific structural features of the *N*-acetylmuramoyl-L-alanine amidase AmiD from *Escherichia coli* and mechanistic implications for enzymes of this family. *J. Mol. Biol.* **397,** 249–259
22. Stratagene (2006) QuikChange® Site-Directed Mutagenesis Kit: Instruction Manual, Stratagene, La Jolla, CA
23. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132
24. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674
25. McCoy, A. J. (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* **63,** 32–41
26. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
27. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501
28. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 235–242
29. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Biol. Crystallogr.* **59,** 1131–1137
30. Matthews, B. W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* **33,** 491–497
31. Kantardjieff, K. A., and Rupp, B. (2003) Matthews coefficient probabilities: improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.* **12,** 1865–1871
32. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., Mc-

Coy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221

33. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 12–21

34. Schüttelkopf, A. W., and van Aalten, D. M. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 1355–1363

35. DeLano, W. L. (2012) *The PyMOL Molecular Graphics System*, version 1.5.0.4, Schrödinger, LLC, New York

36. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 10037–10041

37. Iordanescu, S., and Surdeanu, M. (1976) Two restriction and modification systems in *Staphylococcus aureus* NCTC8325. *J. Gen. Microbiol.* **96,** 277–281

38. de Jonge, B. L., Chang, Y. S., Gage, D., and Tomasz, A. (1992) Peptidoglycan composition in heterogeneous Tn551 mutants of a methicillin-resistant *Staphylococcus aureus* strain. *J. Biol. Chem.* **267,** 11255–11259

39. Hasegawa, H., and Holm, L. (2009) Advances and pitfalls of protein structural alignment. *Curr. Opin. Struct. Biol.* **19,** 341–348

40. Meroueh, S. O., Bencze, K. Z., Hesek, D., Lee, M., Fisher, J. F., Stemmler, T. L., and Mobashery, S. (2006) Three-dimensional structure of the bacterial cell wall peptidoglycan. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 4404–4409

41. Zoll, S., Schlag, M., Shkumatov, A. V., Rautenberg, M., Svergun, D. I., Götz, F., and Stehle, T. (2012) Ligand-binding properties and conformational dynamics of autolysin repeat domains in staphylococcal cell wall recognition. *J. Bacteriol.* **194,** 3789–3802

42. Vollmer, W., Joris, B., Charlier, P., and Foster, S. (2008) Bacterial peptidoglycan (murein) hydrolases. *FEMS Microbiol. Rev.* **32,** 259–286

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

44. de Jonge, B. L., Chang, Y. S., Gage, D., and Tomasz, A. (1992) Peptidoglycan composition of a highly methicillin-resistant *Staphylococcus aureus* strain. The role of penicillin binding protein 2A. *J. Biol. Chem.* **267,** 11248–11254

45. Low, L. Y., Yang, C., Perego, M., Osterman, A., and Liddington, R. (2011) Role of net charge on catalytic domain and influence of cell wall binding domain on bactericidal activity, specificity, and host range of phage lysins. *J. Biol. Chem.* **286,** 34391–34403

46. Yang, D. C., Tan, K., Joachimiak, A., and Bernhardt, T. G. (2012) A conformational switch controls cell wall-remodelling enzymes required for bacterial cell division. *Mol. Microbiol.* **85,** 768–781

47. Cheng, X., Zhang, X., Pflugrath, J. W., and Studier, F. W. (1994) The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **91,** 4034–4038

48. Low, L. Y., Yang, C., Perego, M., Osterman, A., and Liddington, R. C. (2005) Structure and lytic activity of a *Bacillus anthracis* prophage endolysin. *J. Biol. Chem.* **280,** 35433–35439

49. Liepinsh, E., Généreux, C., Dehareng, D., Joris, B., and Otting, G. (2003) NMR structure of *Citrobacter freundii* AmpD, comparison with bacteriophage T7 lysozyme and homology with PGRP domains. *J. Mol. Biol.* **327,** 833–842

50. Martínez-Caballero, S., Lee, M., Artola-Recolons, C., Carrasco-López, C., Hesek, D., Spink, E., Lastochkin, E., Zhang, W., Hellman, L. M., Boggess, B., Mobashery, S., and Hermoso, J. A. (2013) Reaction products and the x-ray structure of AmpDh2, a virulence determinant of *Pseudomonas aeruginosa. J. Am. Chem. Soc.* **135,** 10318–10321

51. Carrasco-López, C., Rojas-Altuve, A., Zhang, W., Hesek, D., Lee, M., Barbe, S., André, I., Ferrer, P., Silva-Martin, N., Castro, G. R., Martínez-Ripoll, M., Mobashery, S., and Hermoso, J. A. (2011) Crystal structures of bacterial peptidoglycan amidase AmpD and an unprecedented activation mechanism. *J. Biol. Chem.* **286,** 31714–31722

52. Prigozhin, D. M., Mavrici, D., Huizar, J. P., Vansell, H. J., and Alber, T. (2013) Structural and biochemical analyses of *Mycobacterium tuberculosis N*-acetylmuramyl-L-alanine amidase Rv3717 point to a role in peptidoglycan fragment recycling. *J. Biol. Chem.* **288,** 31549–31555

53. Warth, A. D., and Strominger, J. L. (1971) Structure of the peptidoglycan from vegetative cell walls of *Bacillus subtilis. Biochemistry* **10,** 4349–4358

54. Boneca, I. G., Huang, Z. H., Gage, D. A., and Tomasz, A. (2000) Characterization of *Staphylococcus aureus* cell wall glycan strands, evidence for a new β-*N*-acetylglucosaminidase activity. *J. Biol. Chem.* **275,** 9910–9918

55. Atrih, A., Bacher, G., Allmaier, G., Williamson, M. P., and Foster, S. J. (1999) Analysis of peptidoglycan structure from vegetative cells of *Bacillus subtilis* 168 and role of PBP 5 in peptidoglycan maturation. *J. Bacteriol.* **181,** 3956–3966

56. Holm, L., and Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38,** W545–W549

57. Advanced Chemistry Development, Inc. (2012) *ACD/ChemSketch Freeware*, version 10.00, Advanced Chemistry Development, Inc., Toronto, Canada

58. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7,** 539

59. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009) Jalview Version 2: a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25,** 1189–1191

ASBMB

# Enabling cell–cell communication via nanopore formation: structure, function and localization of the unique cell wall amidase AmiC2 of *Nostoc punctiforme*

Felix M. Büttner[1],*, Katharina Faulhaber[2],*, Karl Forchhammer[2], Iris Maldener[2] and Thilo Stehle[1,3]

1 Interfaculty Institute of Biochemistry, University of Tübingen, Germany
2 Interfaculty Institute for Microbiology and Infection Medicine, Department of Organismic Interactions, University of Tübingen, Germany
3 Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, TN, USA

To orchestrate a complex life style in changing environments, the filamentous cyanobacterium *Nostoc punctiforme* facilitates communication between neighboring cells through septal junction complexes. This is achieved by nanopores that perforate the peptidoglycan (PGN) layer and traverse the cell septa. The *N*-acetylmuramoyl-ʟ-alanine amidase AmiC2 (Npun_F1846; EC 3.5.1.28) in *N. punctiforme* generates arrays of such nanopores in the septal PGN, in contrast to homologous amidases that mediate daughter cell separation after cell division in unicellular bacteria. Nanopore formation is therefore a novel property of AmiC homologs. Immunofluorescence shows that native AmiC2 localizes to the maturing septum. The high-resolution crystal structure (1.12 Å) of its catalytic domain (AmiC2-cat) differs significantly from known structures of cell splitting and PGN recycling amidases. A wide and shallow binding cavity allows easy access of the substrate to the active site, which harbors an essential zinc ion. AmiC2-cat exhibits strong hydrolytic activity *in vitro*. A single point mutation of a conserved glutamate near the zinc ion results in total loss of activity, whereas zinc removal leads to instability of AmiC2-cat. An inhibitory α-helix, as found in the *Escherichia coli* AmiC$^{E.\ coli}$ structure, is absent. Taken together, our data provide insight into the cell-biological, biochemical and structural properties of an unusual cell wall lytic enzyme that generates nanopores for cell–cell communication in multicellular cyanobacteria. The novel structural features of the catalytic domain and the unique biological function of AmiC2 hint at mechanisms of action and regulation that are distinct from other amidases.

**Abbreviations**

AmiC2-cat, catalytic domain of AmiC2; AMIN, amidase N-terminal domains of AmiC2; DRA, dye release assay; MES, (2-(*N*-morpholino) ethanesulfonic acid); PGN, peptidoglycan; RBB-PGN, peptidoglycan coupled to remazol brilliant blue; RT, room temperature.

# Introduction

Cyanobacteria are widely distributed in the illuminated biosphere. As a result of numerous adaptation mechanisms, including an exceptional variety of morphologies among bacteria, cyanobacteria flourish in many different habitats. Through the invention of oxygenic photosynthesis, they laid the ground for the evolution of higher organisms [1,2]. Although multicellularity refers commonly to eukaryotes, the filamentous cyanobacteria of the order *Nostocales* are true multicellular organisms [3]. Their cells are embedded in a cell wall, a giant macromolecule that spans the entire filament and is surrounded by the outer membrane [4–7]. The periplasm is consequently continuous along the filament [8,9]. Different cells within a filament perform different tasks for a common goal: adaptation and proliferation. With one of the largest genomes among bacteria [10], *Nostoc punctiforme* ATCC 29133 (furthermore: *N. punctiforme*) can adapt to many environmental changes. The filament differentiates into akinetes (spore-like cells) to survive harsh environmental conditions. Under more favorable conditions, akinetes give rise to new filaments [11,12]. In response to nitrogen starvation, heterocysts are formed in a semi-regular pattern along the filament. Heterocysts perform nitrogen fixation and therefore provide the filament with suitable nitrogen sources. In exchange, heterocysts receive carbon sources from surrounding vegetative cells [11]. Hormogonia are small motile filaments that infect mosses and other plants to start a symbiosis [13–15].

Exchange of signal molecules and nutrients within a filament is pivotal for this complex lifestyle. Two routes within a filament are currently discussed. Molecules can be released to the periplasm and diffuse along the filament to be taken up by other cells downstream of a filament [8]. A faster way to exchange information is via cell–cell connections, referred to as microplasmodesmata, septosomes, channels or septal junctions [9,16–19]. Septal protein complexes may span the periplasm and connect neighboring cells as channels [20]. To cross the septal cell wall, a nanopore array consisting of approximately 150 pores of 20 nm in diameter is present in the septal peptidoglycan (PGN) [4,17,21]. These nanopores are essential for cell–cell communication [4,5,17,22]. Mutational analysis showed that *N*-acetylmuramoyl-L-alanine amidases (EC 3.5.1.28) are involved in nanopore formation in *N. punctiforme* (Npun_F1846 or AmiC2) and *Anabaena* sp. PCC 7120 (Alr0092 or AmiC1$^{ANA}$). The Δ*amiC2* mutant strain of *N. punctiforme* shows a pleiotropic phenotype. Aberrantly placed septal planes

lead to distorted filaments that can no longer differentiate into specialized cells. Therefore, diazotrophic growth is not possible, and movement of fluorescent marker molecules between cells is abolished [5].

The cyanobacterial amidases are homologs of amidases that mediate daughter cell separation in unicellular bacteria [23]. Initially, the function of the cyanobacterial amidases was elusive. However, when it became clear that these enzymes have a pore-forming function, it became intriguing to define the structural features that allow the enzyme to adapt to its new task. In line with their unusual task, AmiC amidases of heterocyst-forming filamentous cyanobacteria differ in architecture (Fig. 1A) compared, for example, to the homologous *E. coli* enzyme AmiC$^{E.\ coli}$ [4,22]. In filamentous cyanobacteria, AmiC amidases comprise two AMIN (amidase N-terminal) domains [24] instead of the single AMIN domain found in the enzymes of unicellular bacteria. Furthermore, they carry a long, proline-rich linker domain of approximately 160 residues that separates the AMIN and catalytic domains (Fig. 1A). The AMIN domain was shown to be responsible for septal localization of AmiC$^{E.\ coli}$ in *E. coli* [25,26]. Moreover, there are differences in transport because the *N. punctiforme* AmiC2 enzyme is secreted to the periplasm in an unfolded state via a Sec-signal sequence, whereas AmiC$^{E.\ coli}$ is transported to the periplasm in its folded state via the Tat translocon.

AmiC2 belongs to the Amidase_3 protein family (PF01520) [27], whose structure–function relationship remains under-investigated. Five Amidase_3 enzymes have been structurally determined previously [28–31 and Protein Data Bank (PDB) codes: 1JWQ, 3CZX]; however, only AmiC$^{E.\ coli}$ and its close ortholog AmiB from *Bartonella henselae*, as well as Rv3717 of *Mycobacterium tuberculosis*, were also functionally characterized. The Rv3717 amidase lacks AMIN domains and is considered to play a role in PGN recycling [28,29], whereas AmiC$^{E.\ coli}$ possesses one AMIN domain, carries an inhibitory α-helix obstructing the active site and thus requires activation of associated proteins to facilitate daughter-cell separation [30,31]. Such a helix is not present in the *N. punctiforme* AmiC2, and regulation of activity is therefore elusive.

To define the mechanism of AmiC2 action, we undertook a structure–function analysis of this unusual enzyme. We located native AmiC2 to the septa of a filament using immunofluorescence. We furthermore solved the crystal structure of the catalytic domain of AmiC2 (AmiC2-cat) at high resolution, and found that key features of the enzyme differ significantly from
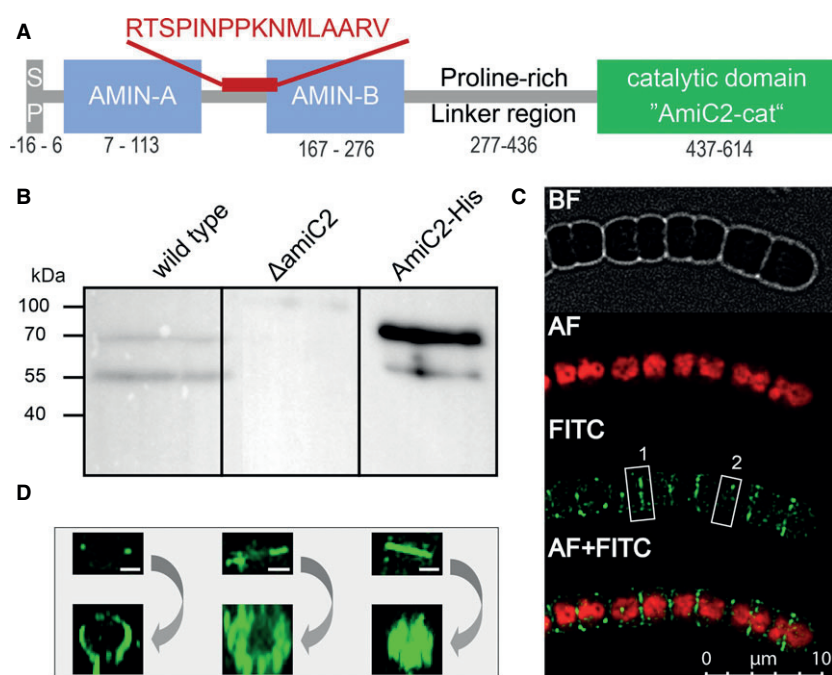
**Fig. 1.** Domain structure and localization of AmiC2 in *N. punctiforme* (A) Domain architecture of AmiC2, which is composed of a sec-signal peptide (SP), the AMIN-A and -B domains, followed by a proline-rich linker region and the catalytic domain. The peptide sequence used for antibody production is shown above the scheme, and residues of respective domains are given below. (B) Immunodetection of AmiC2 in crude cell extracts of the wild-type *N. punctiforme* and the ΔamiC2 mutant (20 µg total protein) and purified AmiC2-6xHis protein (20 ng) probed with the peptide antibody. (C) Immunofluorescence images of a *N. punctiforme* filament using the peptide antibody. AmiC2 localized preferentially to young septa (e.g. frame 1) and was less abundant in older septa (e.g. frame 2). BF, bright field; AF, autofluorescence; FITC, fluorescein isothiocyanate. (D) Three-dimensional projection of AmiC2 localization within septa of different of constriction stages. Top: side view on the septa (as in C) with respective scale bars (= 1 µm). Bottom: top view on the septal plane.

those of cell separation amidases in unicellular bacteria. Interestingly, an inhibitory α-helix that covers the active site in the *E. coli* homolog is missing in AmiC2, implying an entirely different regulation mechanism of the *Nostoc* amidase. Finally, we showed that AmiC2-cat is highly active in a dye release assay (DRA), whereas mutation of zinc binders or otherwise removal of zinc leads to precipitation. Mutation of the potential catalytic glutamate 578 (Glu578) causes total loss of function.

## Results

### Localization of AmiC2 in filaments of *N. punctiforme*

In previous studies, use of AmiC2 (Npun_F1846) with a fused GFP-tag and under the export signal of a cloned N-terminal TAT signal peptide provided preliminary information on the septal localization of this cell wall lytic enzyme in filaments of *N. punctiforme* [4,5]. Because native AmiC2 is translocated via the Sec system and not the Tat-translocation system, we rein-

vestigated the localization in the wild-type filaments using immunofluorescence. A polyclonal peptide antibody against a region solely present in AmiC2 (Fig. 1A) was used to detect native AmiC2 protein *in situ*. In western blots of *N. punctiforme* wild-type and Δ*amiC2* crude cell extracts, this antibody recognized the 69-kDa full length protein and a 55-kDa truncated version of AmiC2 in the wild-type extract only (Fig. 1B). The unique peptide region is not present in the very similar AmiC1 protein (Npun_F1845); hence, the antibody did not show cross-reactivity with AmiC1 in the Δ*amiC2* crude extracts. The observation that a 55-kDa fragment is also detected in fractions of the purified protein suggests that this fragment is produced by auto-proteolysis [4]. The highly specific antibody was used for the subsequent *in situ* studies. As seen in the immunofluorescence micrograph, AmiC2 primarily localized to young septa and was less abundant in old septa and lateral walls (Fig. 1C, frames 1 and 2, respectively). Septum maturation is a slow process that results in old septa, which show advanced constriction visible in phase contrast images (Fig. 1C, frame 2) [4]. Dense Z-stacks (Δ0.1 µm) and three-

dimensional projection enabled localization of AmiC2 within septa of different stages of septum formation (Fig. 1D). In an early state of septum formation, AmiC2 localizes as a circular structure, resembling the Z-ring at the site of septum formation. AmiC2 moves inwards with the invaginating cell wall and finally occupies the entire discoid surface of the young septum. Because of the resolution limit of light microscopy, it is not possible to further resolve the fine-structured localization of AmiC2 on the discoid surface. In old septa and in the lateral cell walls, only traces of AmiC2 could be detected. This localization is in line with a function of AmiC2 in drilling nanopores in the septal PGN during septum formation.

### AmiC2 catalytic activity

From our previous study using full-length AmiC2 protein, we knew that this cyanobacterial amidase shows activity against *E. coli* crude cell extracts in zymogram analysis [4]. To characterize the enzymatic activity of the AmiC2 protein in more detail, AmiC2-cat, which lacks the N-terminal part of AmiC2 including the AMIN domains and the linker domain (Fig. 1A), was overexpressed as a GST-fusion protein in *E. coli* and purified. Using a DRA, the lytic activity of the purified catalytic domain was analyzed against purified peptidoglycan from *E. coli* MG1655 coupled to remazol brilliant blue (RBB-PGN) (Fig. 2A). Remarkably, the activity of AmiC2-cat was very high in comparison to lysozyme. A concentration of 0.08 μM AmiC2-cat

resulted in saturation of the reaction after only 30 min of incubation. Heat-inactivated AmiC2-cat protein did not show any activity, confirming the enzymatic origin of the dye release (Fig. 2A). To investigate a possible lytic activity of the N-terminal parts of AmiC2, we investigated a possible activity of the purified AMIN domains (A + B) (Fig. 2B). No dye release, and thus no activity, was observed for this part of AmiC2.

### Overall structure and active site architecture of AmiC2-cat

To advance an understanding of the enzyme, we solved the structure of AmiC2-cat at 1.12 Å resolution. Excellent data and refinement statistics (Table 1) provide a highly accurate basis for the analysis of the overall structure and the description of its features. AmiC2-cat adopts a mixed αβ-fold that is characteristic for Amidase_3 enzymes (Fig. 3) [32]. This fold comprises a central, six-stranded β-sheet that is surrounded by six α-helices. Helices α1, α3, α5 and α6 shield the rear of the β-sheet, whereas the remaining two helices, α2 and α4, cover part of its front.



**Fig. 2.** Catalytic activity of AmiC2 domains (A) Supernatants of a DRA. Blue color indicates hydrolytic activity on RBB-PGN. Heat-inactivated catalytic domain or buffer alone did not cause dye release. (B) DRA of lysozyme, AmiC2-cat and AMIN domains (AB). All samples were incubated for 30 min each, and all proteins were used at equimolar concentration of 0.08 μM. Mean ± SD are based on three technical replica (*n* = 3).

**Table 1.** Data and refinement statistics for AmiC2-cat. Values for highest resolution bin are given in parentheses. Coordinates are available through PDB code: 5EMI.

|  | 5EMI |
|---|---|
| Data statistics | |
| Beamline | X06DA (PXIII) |
| Space group | $P2_12_12_1$ |
| Cell dimensions (Å) | 35.0 62.8 65.3 |
| (°) | 90  90  90 |
| Wavelength (Å) | 1.000 |
| Detector | PILATUS 2M |
| Resolution (Å) | 50–1.12 (1.15–1.12) |
| Measured reflections | 626 001 (18 460) |
| Unique reflections | 52 924 (2806) |
| CC 1/2 | 100 (88.4) |
| Completeness (%) | 94.3 (69.3) |
| $I/\sigma(I)$ | 31.54 (3.65) |
| Refinement statistics | |
| Resolution (Å) | 1.12 |
| $R_{work}/R_{free}$ (%) | 13.4/14.5 |
| B-factors (Å$^2$) | |
| Wilson | 13.5 |
| Mean | 12.7 |
| rmsd | |
| Bond lengths (Å) | 0.009 |
| Bond angles (°) | 1.426 |
| Ramachandran plot | |
| Most favorable (%) | 98.4 |
| Allowed (%) | 1.6 |
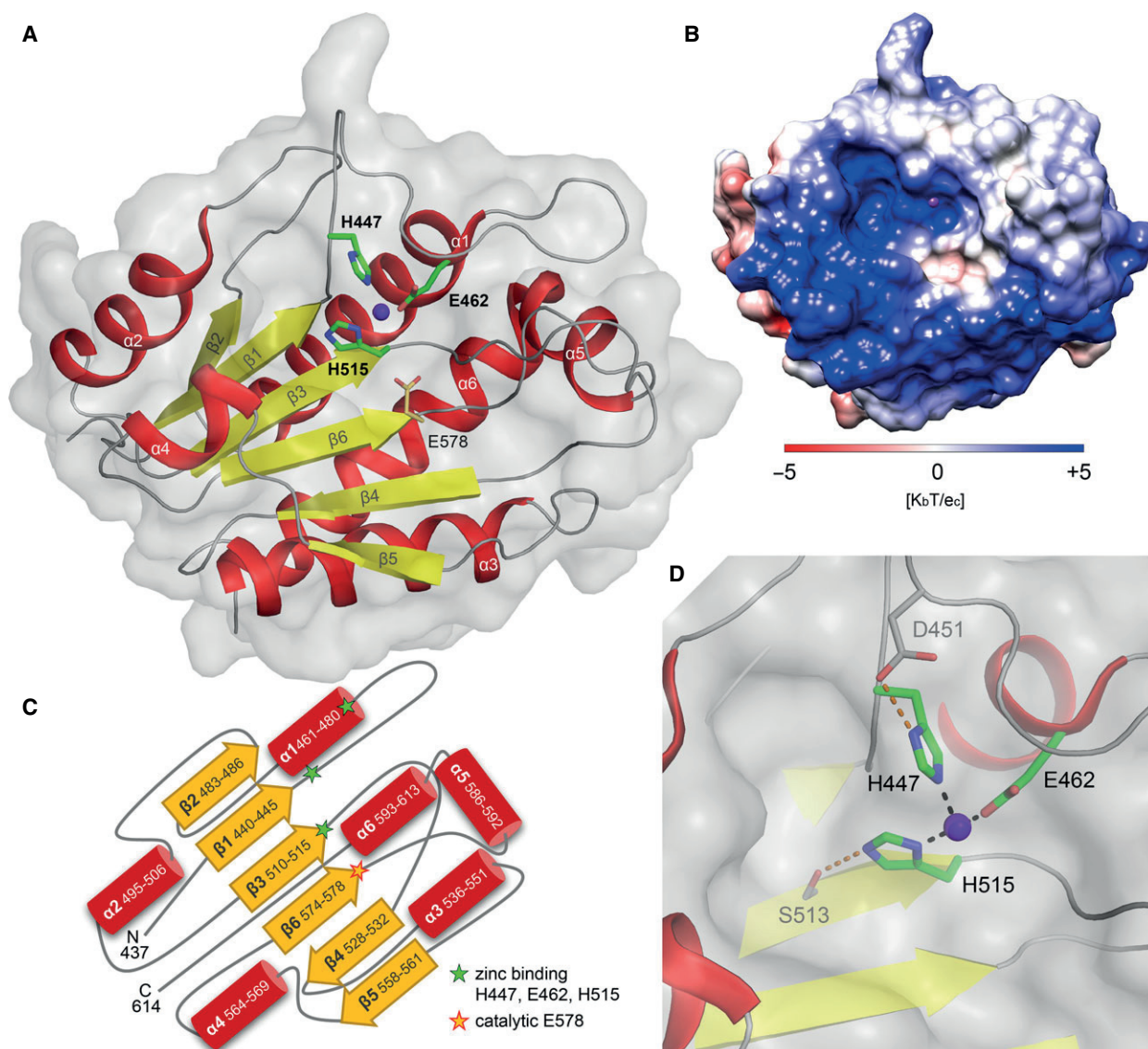
**Fig. 3.** Structure and topology of AmiC2-cat (A) The structure of AmiC2-cat in cartoon representation with a semi-transparent surface reveals a mixed αβ-fold in an arrangement typical for the Amidase_3 family. A central, six-stranded β-sheet builds the core of AmiC2-cat, which is stabilized by six surrounding α-helices and loops. Residues H447, E462 and H515 (green sticks) coordinate a zinc ion (purple sphere) and mark the center of the active site. The neighboring E578 (orange–yellow sticks) plays a pivotal role upon PGN cleavage. The α-helices are shown in red, β-strands in yellow and loops in grey. (B) Electrostatic surface representation of AmiC2-cat showing a strong electropositive potential around the active site. A patch of electronegative potential at the termini of AmiC2-cat is seen at the left edge of the enzyme. (C) The topology diagram depicts the fold of AmiC2-cat. Green stars highlight the positions of the zinc-binding residues H447 in a loop, E462 at the start of helix α1 and H515 at the end of strand β3. Furthermore, the catalytic residue E578 is located at the end of strand β6 (orange–yellow star). In accordance with (A), α-helices are shown as red cylinders and β-sheets as yellow arrows. (D) The active site forms a shallow cavity with a pocket to the left of the zinc ion. The zinc-binding residues H447, E462 and H515 have distances of 2.1 Å each towards the zinc (highlighted by black dashes). Side chains of D451 and S513 (grey sticks) help orient the respective zinc-ligating side chains of H447 (2.8 Å) and H515 (3.2 Å) by hydrogen bonds (shown as orange dashed lines).

Together with loop regions, helices α2 and α4 also define a wide and shallow cavity that harbors the active site and a bound zinc ion. The zinc ion is coordinated by two histidines (His447, His515) and a glutamate (Glu462) at the upper end of a wide and

shallow cavity (Fig. 3A,B). The conserved GHGG motif of residues Gly446, His447, Gly448 and Gly449 after strand β1 permits two close turns and the tight folding at the top of the active site by the absence of side chains. This delicate arrangement and the side

chain of Asp451, lying near His447, facilitate proper orientation of the imidazole residue by a hydrogen bond (2.8 Å), whereas Ser513 anchors the side chain of His515 in a similar manner (3.2 Å) (Fig. 3D). These features, as well as the hydrogen bonds, enable accurate zinc binding by AmiC2-cat. Furthermore, a 2-(*N*-morpholino)ethanesulfonic acid (MES) molecule from the crystallization buffer and the N-terminus of a symmetry copy bound in the active site; they represent crystallographic artifacts (not shown).

## Surface features and conservation

The structurally known and related enzyme from a unicellular bacterium, AmiC$^{E.\ coli}$, differs from AmiC2-cat in terms of the shape of the active site region (Fig. 4A). AmiC$^{E.\ coli}$ has a narrower groove that could accommodate the substrate, whereas AmiC2-cat has a much more open and wider binding region. Moreover, the active site of AmiC$^{E.\ coli}$ is covered with an inhibitory α-helix that must change position for substrate to be processed. Because of these differences, few conclusions for substrate binding of AmiC2-cat can be drawn from a comparison with the AmiC$^{E.\ coli}$ structure despite their ability to cleave the same substrate (Fig. 2). There is no enzyme–substrate complex structure available for AmiC$^{E.\ coli}$, although the structure of the catalytic domain of the *Staphylococcus aureus* amidase (AmiA-cat$^{S.\ aureus}$) was recently solved bound to its substrate [33] (Fig. 4A). AmiA-
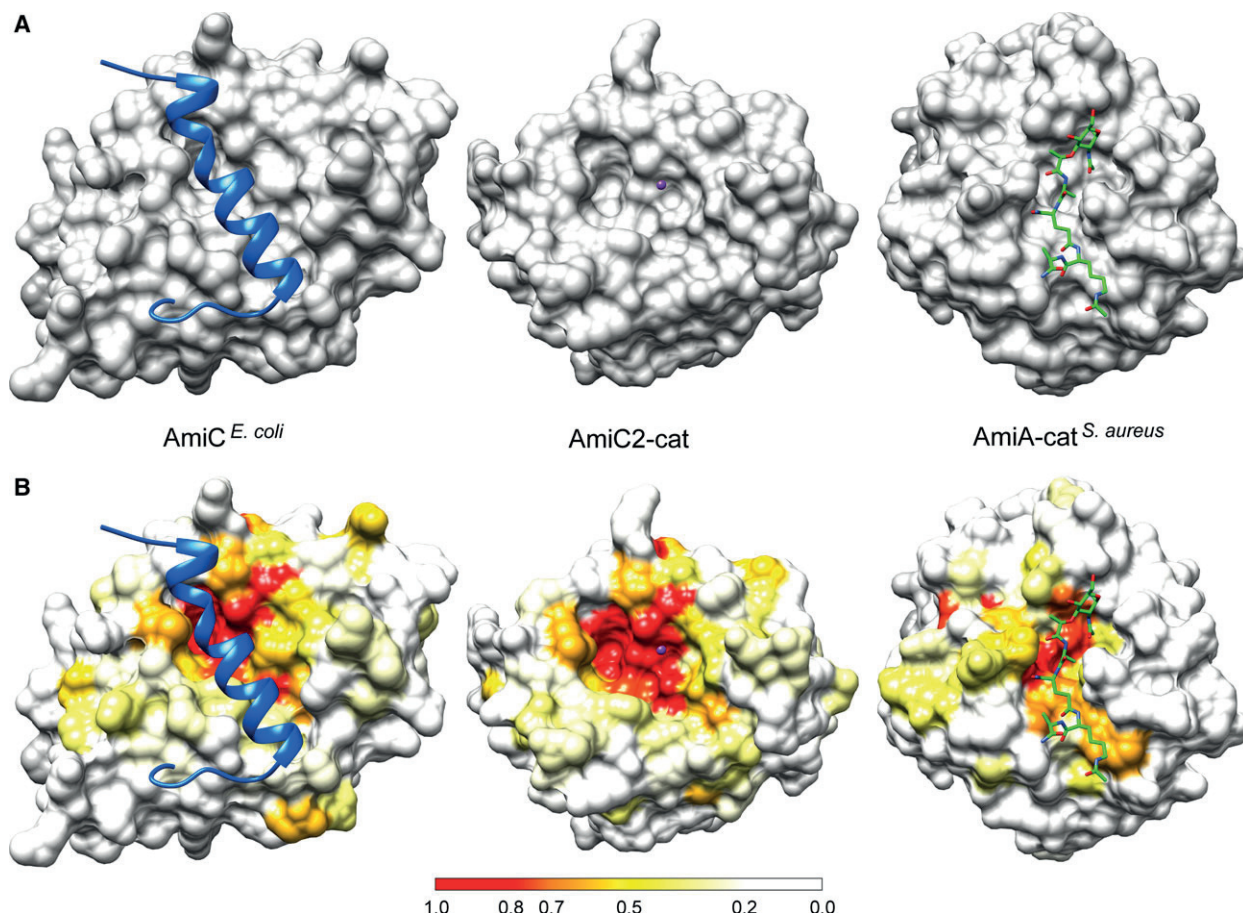


**Fig. 4.** Surface conservation of AmiC2-cat (middle), AmiC$^{E.\ coli}$ (left) and AmiA-cat$^{S.\ aureus}$ (right). (A) Surface representations of the three selected amidases reveal no recessed binding cleft for AmiC2-cat (middle) and AmiC$^{E.\ coli}$ (left). The inhibitory α-helix of AmiC$^{E.\ coli}$ also gives no clear indication on substrate binding. By contrast, the *S. aureus* amidase AmiA-cat$^{S.\ aureus}$ shows a clear Y-shaped binding grove where its ligand is bound. (B) AmiC2-cat, a member of the Amidase_3 family, only displays conservation in the zinc-binding region of the protein, whereas the surface of AmiA-cat$^{S.\ aureus}$ (an Amidase_2 family member) is conserved along the entire binding groove of the bound substrate muramoyltetrapeptide. This conservation enables substrate modelling options for Amidase_2 enzymes but none for an Amidase_3. The level of conservation is depicted by a color bar. Seed sequences from the PFAM database, ClustalOmega and Chimera were used to generate images in (B).

cat$^{S. aureus}$ also features a narrow, groove-like binding region that engages the linear substrate through numerous contacts. Comparison of the binding surfaces shows that AmiC2-cat does not have such a narrow groove, indicating that the enzyme could conceivably accommodate several different substrate orientations, perhaps in the form of PGN monomers, as well as highly cross-linked PGN (Fig. 4A).

To gain insight into regions that might be involved in ligand binding, we mapped conserved residues onto the surface of the AmiC2-cat structure (Fig. 4B). A bioinformatic analysis of seed sequences from the PFAM database was chosen over BLAST hits to exclude bias from the alignment and conservation analysis. The zinc-coordinating residues His447, Glu462 and His515, as well as Glu578 and the active site residues highlighted in Fig. 3, are strongly conserved among members of the Amidase_3 family, indicating a generally conserved mechanism of cleavage.

However, a surprisingly small number of surface-exposed amino acids are conserved beyond the active site (Fig. 4B) and most of these are conserved because of structural requirements. For example, the conserved Pro573 enables a tight turn just before strand β3 with the zinc-coordinating Glu578 at its end and, similarly, the conserved Gly481 allows the two amino acid turn between helix α1 and strand β2. The analysis of conservation does not allow us to delineate a putative substrate binding groove for AmiC2-cat (Fig. 4B). This is in contrast to representatives of the Amidase_2 family, where such analyses revealed putative substrate binding regions in AmiA-cat$^{S. aureus}$ homologs [34,35].

## Mutational analysis

Based on the AmiC2-cat structure and sequence conservation, we designed several single point mutations in the active site. The three zinc-binding residues His447, Glu462 and His515, as well as Asp451, were separately mutated to alanine (H447A, E462A, H515A and D451A, respectively), and the four mutant proteins were expressed and purified similarly to the wild-type protein. However, all four mutants were unstable and aggregated during size exclusion chromatography (data not shown) and they could therefore not be analyzed for residual activity (Fig. 5B). Because the four point mutations likely affected zinc binding, we conclude that the zinc ion is important for protein stability and the integrity of the overall fold. In support of this, removal of zinc from wt AmiC2-cat using the chelator EDTA in solution also led to precipitation of the protein (data not shown).

According to a previously proposed catalytic mechanism for the related AmiA-cat$^{S. aureus}$ amidase, a conserved glutamate accepts a proton from a nucleophilic water next to the zinc, enabling it to attack the PGN and facilitate cleavage [33]. To analyze whether this conserved glutamate, Glu578 in AmiC2-cat, is also relevant for the catalytic activity of AmiC2-cat, single point mutants of Glu578 to alanine, glutamine, aspartate and asparagine (E578A, E578Q, E578D and E578N, respectively) were generated. These mutations were designed to evaluate the potential importance of charge and side chain length for enzymatic activity. The resulting proteins could easily be purified after
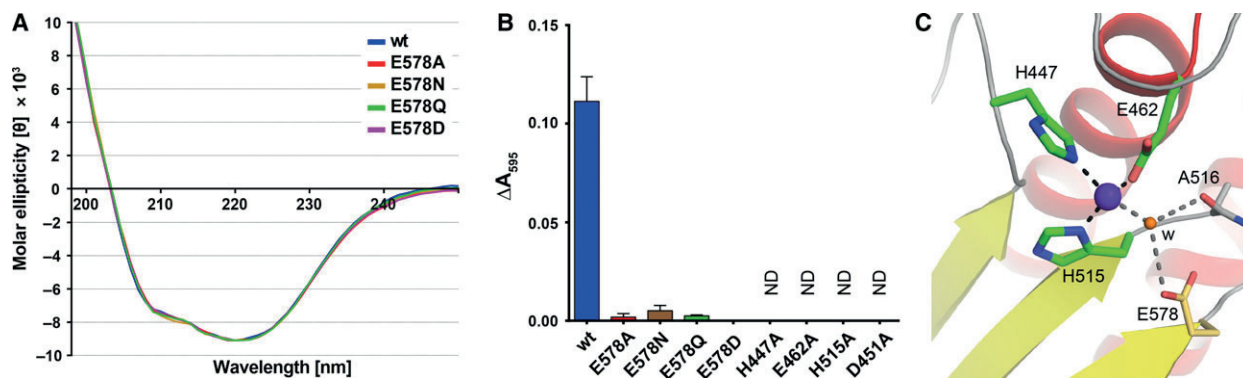


**Fig. 5.** AmiC2-cat mutant stability and activity. (A) CD spectroscopy with stable mutants revealed unaltered folding compared to wild-type AmiC2-cat. (B) DRA of AmiC2-cat and E578 mutants shows total loss of activity for all substitutions. The activities of other mutants could not be determined as a result of protein instability and aggregation. ND, not determined. Mean ± SD are based on three technical replica ($n = 3$). (C) The active site is depicted similar to Fig. 3D with backbone of A516 (grey sticks) and the E578 side chain (orange–yellow sticks). The position of the N-terminal amino group of a crystallographic symmetry-related molecule that bound in the active site is substituted with a water molecule 'w' in orange. This water molecule is in perfect distance to the zinc ion (2.2 Å, grey dashes) and the assisting residues E578 and A516 (3.0 Å each, grey dashes). Thus, the water is well positioned for activation and performing a nucleophilic attack on the scissile bond of the substrate.

adaptation of the purification protocol (see Experimental procedures). Analytical size exclusion chromatography (data not shown) and CD spectroscopy (Fig. 5A) confirmed stability, as well as proper folding, without any significant difference from the wild-type AmiC2-cat. The Glu578 mutants were investigated using the DRA. None of the Glu578-variants exhibited dye release, suggesting that this residue is indispensable for catalysis (Fig. 5B). Thus, the Glu578 side chain likely activates the water molecule located next to the bound zinc ion.

## Discussion

With the formation of nanopore arrays, the cell wall amidase AmiC2 fulfills a pivotal function in filamentous *N. punctiforme* [4]. Such arrays consist of approximately 150 nanopores that are located in the center of the septal PGN, and allow for the formation of cell–cell communication structures. A *Nostoc* mutant lacking AmiC2 forms aberrant septa during cell division, cannot differentiate specialized cell types and cannot establish cell–cell communication via septal junction proteins [5]. In general, cell wall lytic enzymes must be under stringent control to avoid cell lysis. In case of the highly specific AmiC2 amidase, however, the enzyme has to be spatiotemporally regulated. The formation of the nanopore array requires precise activation of AmiC2 on distinct spots on the septal disc at a certain time during septum maturation. In *E. coli*, LytM domain proteins such as NlpD and EnvC fulfill the regulatory task by displacing an obstructing α-helix from the active site of AmiC$^{E.\ coli}$ during cell division by an unknown mechanism [30,31,36]. We performed *in situ* localization studies, as well as biochemical and structural analyses of AmiC2, aiming to define the substrate specificity of the enzyme and the mechanism of regulation of its lytic activity.

*In situ* immunofluorescence confirmed the septal localization of AmiC2 and allowed us to correlate the distribution of AmiC2 with the state of septum formation. The cellular localization of the enzyme correlates well with its proposed function in nanopore formation in the septal cell walls [4]. In old septa, the formation of new nanopores may not be needed anymore, and hence the amidase protein disappears. Nevertheless, the time course of amidase activation is still unclear and could occur in two different ways: the AmiC2 protein is either active from the beginning of septum formation or AmiC2 is present from the beginning (as shown here) but activated after the septum has completely formed, and then disappears completely. In *E. coli*, the AMIN$^{E.\ coli}$ domain, which binds to the PGN, has

been shown to be necessary for the correct localization of amidases to the septum during cell division [24,31]. Because the AMIN domains of AmiC2 from *N. punctiforme* do not contribute to lytic activity, it is probable that they help to mediate septal localization.

The DRA demonstrates that the purified AmiC2-cat domain has a significantly elevated activity compared with lysozyme. The substrate used in all DRA was Gram-negative *E. coli* PGN. This compound is similar to the cyanobacterial PGN, which is considered to be Gram-negative because of the presence of an outer membrane. Nevertheless, the PGN of most cyanobacterial strains is thicker than in typical Gram-negative bacteria. Its thickness is estimated at approximately 12 nm versus approximately 6 nm in *E. coli* [9,37–41]. The higher activity of AmiC2 may therefore reflect the need for efficient processing of the thicker, more complex PGN in cyanobacteria.

It is not known whether the structure and composition of septal PGN of filamentous cyanobacteria differs from lateral murein or PGN of unicellular organisms. Hence, the septal PGN could contain structural features that are involved in activation of AmiC2 and its special function in drilling nanopores. The N-terminal domains of AmiC2, such as the AMIN domains and/or the proline-rich linker, may be involved in regulation. Limited proteolysis could be another possibility for governing AmiC2 activity. Recently, Singh *et al.* [42] showed that a murein endopeptidase is regulated by a two-protein system comprising NlpI, an outer membrane lipoprotein, and Prc, a periplasmic protease. Because shorter fragments of AmiC2 appear in western blot analysis with crude cell extracts, as well as with purified AmiC2, the possibility of regulation by (auto-) proteolysis should be considered and deserves further investigation.

A complex machinery of catalytic regulation may be required to drill pores of uniform size and shape with a precise spatial distribution across the septum surface. In the closely-related strain *Anabaena* sp. PCC 7120, the protein SjcF1 was found to influence the size of the nanopores, and a *sjcF1* mutant was hampered in cellular exchange of molecules [43]. In addition, SjcF1 was shown to interact with the putative channel forming FraC and SepJ proteins, which also have influence on the formation of the nanopore array [17]. However, protein interaction with PGN processing proteins such as AmiC1 or AmiC2 has not yet been investigated.

The high-resolution crystal structure of AmiC2-cat reveals a typical Amidase_3 family fold [27,32]. Although five Amidase_3 structures have previously been determined, only AmiC$^{E.\ coli}$ [30,31] and Rv3717 of *M. tuberculosi*s [28,29] were also functionally char-

acterized. The structure of AmiC$^{E.\ coli}$ [31] differs mostly by the insertion of a regulatory α-helix and the elongation of helix α3 (Fig. 6A). The inserted α-helix blocks the active site and prevents catalysis. Inhibition of AmiC$^{E.\ coli}$ even occurs, if a deletion mutant is supplemented with the missing obstructing α-helix [31]. The interaction with LytM domains, which are part of the divisome machinery that forms a septal ring below the inner membrane, induces a displacement of the regulatory α-helix in a possibly ATP-dependent manner [30,31]. Rv3717 of *Mycobacterium tuberculosis* has broad-spectrum activity [28] and carries an insertion that forms a $3_{10}$-helix and an additional β-hairpin that is stabilized by a disulfide bridge (Fig. 6B) [28,29]. The insertion acts as a lid for the binding site and generates a tunnel that can accommodate substrate but limits the activity to smaller PGN fragments [29]. Furthermore, product fragments and a docked Mur-NAc are located in this tunnel. Both studies also assume a regulatory function of the insertion because Rv3717 lacks a cell wall binding domain. Along with functional data, the structure of AmiC2-cat is therefore the first to characterize a subclass of the Amidase_3 family that features no regulatory insertions and is thus distinct in regulation.

The zinc ion bound to AmiC2-cat likely serves an essential function for enzyme stability as zinc removal by EDTA or mutation of zinc-coordinating residues led to protein aggregation. Although some amidases can tolerate zinc removal, such as AmiE [44], Ply-PSA [45] and Rv3717 [29], others exhibit similar sensitivity, such as Rv3717 [28] and AmiA (F. M. Büttner, personal communication). Furthermore, the electrostatic surface potential reveals strong, mostly electropositive features in the active site (Fig. 3B). Thus, the PNG substrate of AmiC2-cat likely carries an overall negative charge. PGN of cyanobacteria belongs to type A1γ and comprises building blocks with free carboxyl groups such as meso-diaminopimelic acid, as well as unlinked D-Ala and feasibly non-amidated D-iGlu [38,46], thereby possibly supporting substrate specificity.

Mutational analysis of Glu578 demonstrates the relevance of this residue for catalysis, and provides evidence for a reaction mechanism that is based on the activation of a water molecule next to the zinc ion, consistent with reaction mechanisms described for other zinc-dependent amidases. The mutants show that even a small change in side chain length (E578D) or a change in charge (E578Q) leads to a large reduction in
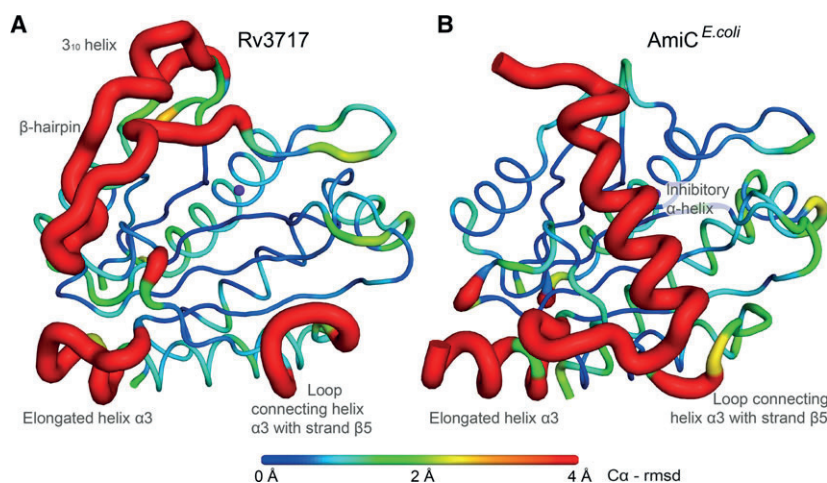


**Fig. 6.** Structural comparison of Amidase_3 enzymes with AmiC2-cat by Cα-rmsd color plot. (A) Rv3717 (PDB code: 4M6G, Cα rmsd of 1.5 Å; 29% sequence identity) deviates from AmiC2-cat by an insertion of a $3_{10}$-helix and a β-hairpin that acts as a lid on the active site (shown in red). The resulting tunnel reduces substrate variability and indicates PGN recycling activity of Rv3717. Helix α3 is extended, and the loop connecting helix α3 with strand β5 further constricts the lower end of the binding site. The inner core, however, aligns very well with AmiC2-cat, as indicated by a low Cα-rmsd (blue). (B) AmiC$^{E.\ coli}$ (PDB code: 4BIN, Cα rmsd of 1.5 Å; 34% sequence identity) exhibits two major differences compared to AmiC2-cat, both being involved in regulation. An inhibitory α-helix with long connecting loops that is inserted between strand β4 and helix α3 blocks the active site, thereby inactivating the enzyme. In addition, helix α3 is significantly extended, allowing the movement of the inhibitory α-helix in and out of the active site upon interaction with LytM domains. AmiC2-cat and the respective enzymes were aligned, and rmsd values for positions of every Cα atom were calculated. The rmsd values were used to color the respective enzymes from no deviation in blue to 4 Å or more deviation in red. Three N- and one C-terminal residues of Rv3717 (4M6G), as well as the AMIN domain along with 13 N-terminal residues of AmiC$^{E.\ coli}$ (4BIN), were omitted from this representation for clarity.

activity, suggesting that proper placement of the E578 side chain is essential for the reaction mechanism. In the AmiC2-cat crystals, N-terminal residues of a symmetry-related molecule reach into the active site and participate in zinc binding, completing the almost perfect octahedral coordination. A water molecule is likely to coordinate the zinc at this position in a more physiologic setting. Contacts with the backbone carbonyl of Ala516 and the Glu578 carboxylate (3.0 Å each) and the zinc ion with 2.2 Å (Fig. 5C) would enable this water to initiate a nucleophilic attack on the scissile amide bond of PGN between the peptide stem and the carbohydrate backbone. Although Glu578 could temporarily accept a proton from this water molecule, the glutamine amide group could not be protonated at physiologic pH. The nearby hydrophilic surface-exposed residues Asn517, Glu529 and Arg560 are partly conserved, and they might participate in substrate binding or form part of a proton relay system for Glu578, as suggested by Mayer *et al.* [47] for a bacteriophage amidase.

AmiC2-cat has surprisingly few conserved surface residues, which is perhaps in keeping with its unusual function in nanopore formation. By providing a large groove that can host large PGN molecules, as present in cyanobacteria, the open and shallow binding surface of AmiC2, along with the lack of conservation, suggests that its substrate may be accommodated in somewhat different orientations. This would also be in line with the observed high enzymatic activity against *E. coli* PGN *in vitro* and *in vivo* because fast substrate processing requires easy access of the substrate to the active site, and quick dissociation of the products. Moreover, an electropositive potential was shown to be critical for (and even enhanced) amidase activity in the Amidase_2 enzymes PlyA and XlyA [48]. Likewise, the electropositive potential of AmiC2-cat could contribute to its high activity, perhaps by facilitating substrate binding and docking into the active site through long-range electrostatic forces. It is tempting to speculate that MurNAc could fit into the pocket to the left of the zinc ion where a MES molecule of the crystallization buffer bound. This would position the scissile bond near a zinc-activated water molecule that would engage the nucleophilic attack. The remaining portion of the PGN chain substrate would then lie across the shallow binding region.

Taken together, the data reported in the present study shed light on the pivotal role of the *N*-acetylmuramoyl-L-alanine amidase AmiC2 in context of cellular communication in *N. punctiforme* filaments. Our results provide an excellent basis for further studies

involving the molecular mechanism, as well as substrate and site specificity.

# Experimental procedures

## Strains and growth conditions

*E. coli* strains were grown in LB medium [49]. Strain MG1655 grew without the addition of antibiotics, and transformed clones of *E. coli* BL21 were selected on 100 µg·mL$^{-1}$ ampicillin. *N. punctiforme* and its Δ*amiC2* mutant strain were grown in medium of BG11 supplemented with NO3$^-$ [15]. The mutant was grown in presence of 50 µg·mL$^{-1}$ neomycin. Both cyanobacterial strains were cultivated under horizontal shaking (80 r.p.m.) and permanent light [5].

## Immunodetection

A 50 mL culture of *N. punctiforme* filaments was harvested (20 000 *g* and 4 °C) and the filaments resuspended in 1 mL 50 mM Hepes. Cells were broken by sonication, and the protein concentration determined according to Bradford [50]. The AmiC2-6xHis protein was purified as described previously [4]. After separation by SDS/PAGE, the proteins were blotted on a poly(vinylidene difluoride) membrane [51], which was blocked in blocking solution (1× PBS 0.1% Tween 20, +5% skim milk) for 1 h at room temperature (RT). Peptide antibodies were used at a dilution of 1 : 100 in blocking solution and incubated overnight at 4 °C. After three washes, the membrane was incubated with a horseradish peroxidase-conjugated secondary antibody (dilution 1 : 10 000; Sigma-Aldrich, Munich, Germany) for 2 h at RT. After four rounds of washing, the LumiLight mixture (Roche, Basel, Switzerland) was used and the luminescence captured by the Gel Logic 1500 imaging system (Kodak, Rochester, NY, USA).

## Immunofluorescence localization

A polyclonal peptide antibody was raised in rabbits against the synthetic peptide RTSPINPPKNGMLAARV (Fig. 1A) by Dr J. Pineda Antikörper-Service (Berlin, Germany). *N. punctiforme* filaments from exponentially growing cultures were harvested and washed with 2–5 mL of PBS. The cells were fixed by 2 mL of HistoChoice Tissue Fixative (Sigma-Aldrich) for 10 min at RT, followed by incubation at 4 °C for 30 min. After three washes with PBS, the cells were resuspended in 1 mL of ice-cold 70% EtOH, again washed, and resuspended in 2 mL of GTE buffer (50 mM glucose, 20 mM Tris-HCl, pH 7.5, 10 mM EDTA). Next, 2 mg·mL$^{-1}$ lysozyme (Sigma Aldrich) was added and left at RT for 5 min. Then, 200 µL of the cell suspension were transferred to Polylysine® slides (Thermo

Scientific, Waltham, MA, USA) and dried. After rehydration with 200 µL of PBS for 5 min at RT, the slides were blocked by 200 µL of 2% BSA in PBS for 20 min and then incubated with primary antibody (dilution 1 : 5 in PBS containing 2% BSA) at 4 °C overnight in a wet chamber. After eight washes with PBS, the FITC coupled secondary antibody (Sigma Aldrich; dilution 1 : 200 in PBS/BSA) was incubated for 2 h at room temperature in a dark wet chamber and washed eight times with PBS. One drop of Mounting Medium H1200 (Vector Laboratories, Inc., Burlingame, CA, USA) was added to each slide and then overlaid with cover glasses (24 × 60 mm) and sealed with transparent nail polish. The slides were stored at 4 °C in the dark.

To capture fluorescence and bright field images, a DM5500B microscope (Leica Microsystems, Wetzlar, Germany) was used with a ×100/1.3 numerical aperture oil objective lens. Image acquisition was carried out using a DFC360FX monochrome camera (Leica Microsystems). A BP470 40 nm excitation filter and a BP525 50 nm emission filter were used for detection. The exposure time was adjusted according to green fluorescence background in samples without primary antibody incubation. Autofluorescence of cyanobacterial cells was monitored with a filter cube of BP535 50 nm excitation filters and a BP610 75 nm emission filter. Exposure times ranged from 6 ms (bright field) to 80 ms (fluorescein isothiocyanate) and 170 ms (autofluorescence). Images were taken as Z-stacks with 0.1 µm intervals. Z-stacks were used to perform three-dimensional deconvolution using the built-in function of the LEICA ASF software (Leica Microsystems). Images were recolored by the LEICA ASF software based on the filter used, and intensity levels were adjusted using ADOBE PHOTOSHOP CS3 (Adobe Systems Inc., San Jose, CA, USA).

## PGN isolation from *E. coli* MG1655

The PGN of *E. coli* was purified according to the protocol of Glauner [51], with some modifications. Bacteria grown overnight at 37 °C in LB broth were cooled to 4 °C and harvested (4000 $g$ for 30 min at 4 °C). Cells were resuspended in ice-cold distilled water (0.2 g·mL$^{-1}$) and added dropwise to an equal volume of 8% boiling SDS. Boiling was continued for 30 min under vigorous stirring. After cooling to RT, the SDS insoluble fraction was collected (100 000 $g$ for 45 min at RT). The pellet was freed from SDS by repeated washing and/or dialysis against water. The SDS content was controlled by methylene blue assay [52]. Subsequently, the pellet was resuspended in H$_2$O. Then, 100 µg of α-amylase (from *Bacillus subtilis*; Fluka BioChemika, Buchs, Switzerland) per mL were added and incubated at 37 °C for 2 h. To inhibit the activity of contaminated lysozyme in the amylase preparation, 0.32 M imi-

dazole was added. Spatula tip amounts of DNAse and RNAse (Genaxxon, Ulm, Germany) were also added. During 2 h of incubation (37 °C), pronase (from *Streptomyces griseus*; Merck, Darmstadt, Germany) was pretreated for 90 min at 60 °C followed by the addition of 200 µg·mL$^{-1}$ pronase to remove lipoproteins covalently bound to the PGN. Again, the sample was added to an equal volume of 8% SDS and boiled for at least 15 min. SDS was removed afterwards.

## DRA

Coupling of PGN to RBB and DRA was performed according to Rocaboy *et al.* [31] with some modifications: 500 µL of RemazolBrilliantBlue (1% w/v H$_2$O), 500 µL of H$_2$O and 100 µL (20 g of Na$_2$SO$_4$, 2.32 g of Na$_3$PO$_4$ × 12 H$_2$O in 10 mL of H$_2$O) were added to 1 mL of purified PGN and incubated at 50 °C for 30 min. Afterwards, the PGN was washed with 50 mM sodium phosphate buffer (pH 7) until the supernatant was clear. For measuring hydrolytic activity, 50 µL of RBB-PGN were added to 0.16 µM enzyme in 50 mM TrisHCl, 150 mM NaCl buffer, pH 7.8 (final enzyme concentration: 0.08 µM) and immediately incubated at 28 °C by horizontal shaking with 300 r.p.m. Heat-inactivated samples were denatured at 95 °C for 20 min prior to addition to RBB-PGN. The reaction was stopped at defined time points by adding 100 µL of 20% (w/v) TCA. The samples were centrifuged (1700 $g$ for 10 min at RT) and 150 µL of the supernatant was added to 150 µL of H$_2$O in micro-cuvettes. The absorbance was measured at 595 nm. All activities were normalized to a control containing solely RBB-PGN and buffer. Graphs were prepared using PRISM, version 6.07 (GraphPad Software, La Jolla, CA, USA).

## Cloning

Plasmids and primers used for cloning are listed in Tables S1 and S2. Whole cells of *N. punctiforme* were used as template to amplify the AmiC2-cat fragment in PCR using primers No 993 and 994, introducing restriction sites (*Bam*HI-*Eco*RI) for cloning, and high fidelity Q5-polymerase (NEB, Ipswich, MA, USA). The amplified fragment was ligated into a pGEX-4T-3 vector, resulting in an N-terminal fusion of GST-tag to the AmiC2-cat domain. The sequence encoding the AMIN domains was ordered as an optimized sequence (MWG Eurofins, Ebersberg, Germany) and delivered in the pEX-K standard cloning plasmid. PCR was performed with primers No 997 and 998 introducing restriction sites *Bam*HI and *Eco*RI and the resulting fragment ligated in pGEX-4T-3. All plasmids were transformed in *E. coli* BL21(DE3). The integrity of all PCR-based products was verified by sequencing.

## Mutagenesis

Site-directed mutagenesis was used to generate point mutations of zinc-coordinating residues (H447A, E462A and H515A), as well as active site residues (D451A, E578A, E578Q, E578D, E578N). Oligonucleotides introducing the mutations are given in Table S2. The reactions were performed as described in the QuikChange protocol (Agilent Technologies, Santa Clara, CA, USA).

## Protein expression and purification

Wild-type AmiC2-cat carrying an N-terminal GST-tag was expressed in *E. coli* BL21 (DE3) cells in auto-induction medium (ZYM-5052) [53] over 72 h at 21 °C. Harvested cells were resuspended in buffer (150 mM NaCl, 50 mM Tris, pH 8), supplemented with 1 mM PMSF and 1% Tween 20, and lysed with a FrenchPress (EmulsiFlexC3; Avestin, Ottawa, ON, Canada). The cell lysate was filtered and incubated with 2 mM ATP + 10 mM $MgSO_4$ for 15 min, when rotating at 37 °C. The supernatant was applied to a GSTrap FF column (GE Healthcare, Little Chalfont, UK). After thorough washing, a buffered salt gradient reaching 1 M NaCl removed impurities. Thirty units of thrombin per milligram fusion protein were added for on-column cleavage at 4 °C over 3 days and released AmiC2-cat. The remaining impurities and aggregate were removed by size exclusion chromatography with a 215 nm detection system (Superdex75 on an ÄKTA FPLC; GE Healthcare).

AmiC2-cat mutants were expressed and initially processed as described for the wild-type enzyme. In a modification of the above protocol, the fusion protein was eluted using buffer supplemented with 15 mM reduced glutathione after affinity chromatography and then size exclusion chromatography was performed. In the case of H447A, E462A and H515A, as well as D451A, the proteins all ran in the aggregate peak. Fusion proteins of E578A/Q/D/N, however, were stable and could be cleaved in solution over 60 h. GST and thrombin were removed by GSTrap FF and Benzamidine FF columns (GE Healthcare) before final gel filtration.

AmiC2_AMIN-AB was cloned similarly to AmiC2-cat, with an N-terminal GST-tag, expressed using *E. coli* BL21 (DE3) cells in auto-induction medium over 3 days, and purified using the same protocol.

Purity of proteins was confirmed by SDS/PAGE and analytical size exclusion chromatography, their identity was verified by MALDI and proper folding was analyzed using circular dichroism.

## Circular dichroism spectroscopy (CD)

Wild-type AmiC2-cat and AmiC2-cat mutants stored in gel filtration buffer were concentrated to approximately 3.0–3.5 mg·mL$^{-1}$ and subsequently diluted with ddH$_2$O to abolish interference by chloride ions. The final concentration of the proteins was approximately 0.2 mg·mL$^{-1}$ in diluted buffers of approximately 9 mM NaCl and 3 mM Tris. All samples were transferred into a 1 mm quartz cuvette and measured using a J-720 spectrometer (Jasco, Gross-Umstadt, Germany). Spectra were recorded at room temperature with an accumulation of 10 in the range 250–199 nm. The spectra manager (Jasco) was used to subtract the buffer baseline, smoothen the spectra and calculate molar ellipticity values. Because of imprecise measurements of the final protein concentrations, the molar ellipticity values of all AmiC2-cat mutants needed to be scaled to wild-type AmiC2-cat. The individual scaling factors were determined by the ratio of molar ellipticity of wild-type at 220 nm over the molar ellipticity of mutant at 220 nm and applied to all values of the respective mutant. CD spectra were prepared using PRISM.

## Crystallization

AmiC2-cat was concentrated to 3.7 mg·mL$^{-1}$, and 1 μL of protein was mixed with 1 μL of well solution containing 120 mM alcohols (mixture of 1,6-hexanediol, 1-butanol, 1,2-propanediol, 2-propanol, 1,4-butanediol and 1,3-propanediol), 100 mM MES/imidazole (pH 6.5) and 37.5% poly(ethylene glycol)1000 / poly(ethylene glycol)3350 / 2-methyl-2,4-pentanediol using the hanging drop vapor diffusion method. Orthorhombic crystals grew over 2 weeks at 20 °C and were directly flash frozen in liquid nitrogen.

## Data collection and structure determination

Data were collected at 100 K using a PILATUS 2M detector at beamline X06DA of the Swiss Light Source (Villigen, Switzerland). Indexing, integrating and scaling were performed using POINTLESS [54] implemented in the CCP4 suite [55] and XDS [56]. The structure of Cwlv, an amidase from *Bacillus* (Paenibacillus) Polymyxa var. *colistinus* (PDB code: 1JWQ), with 38% sequence identity was modified using CHAINSAW [57], also implemented in CCP4, to serve as a search model for molecular replacement by PHASER [58]. Iterative model building, refinement and validation were performed with COOT [59] and PHENIX [60]. All 178 residues and the two-residue linker were built in the final model. Images were generated using PYMOL [61] and CHIMERA [62], exploiting bioinformatic information from NCBI (http://www.ncbi.nlm.nih.gov), CLUSTAL OMEGA [63] and PFAM [27].

experiments and the Swiss Light Source, Villigen, Switzerland, for beam time and support. The *Deutsche Forschungsgemeinschaft* (DFG), 'CRC766 – The bacterial cell envelope: Structure, function and infection interface' funded this work.

## Author contributions

FMB, TS, KFa, IM and KF planned the experiments. FMB and KFa performed the experiments. FMB, TS, KFa, IM and KF analyzed data. FMB, TS, KFa, IM and KF wrote the paper.

## References

1 Holland HD (2006) The oxygenation of the atmosphere and oceans. *Philos Trans R Soc Lond B Biol Sci* **361**, 903–915.

2 Tomitani A, Knoll AH, Cavanaugh CM & Ohno T (2006) The evolutionary diversification of cyanobacteria: molecular-phylogenetic and paleontological perspectives. *Proc Natl Acad Sci USA* **103**, 5442–5447.

3 Kaiser D (2001) Building a multicellular organism. *Annu Rev Genet* **35**, 103–123.

4 Lehner J, Berendt S, Dorsam B, Perez R, Forchhammer K & Maldener I (2013) Prokaryotic multicellularity: a nanopore array for bacterial cell communication. *FASEB J* **27**, 2293–2300.

5 Lehner J, Zhang Y, Berendt S, Rasse TM, Forchhammer K & Maldener I (2011) The morphogene AmiC2 is pivotal for multicellular development in the cyanobacterium Nostoc punctiforme. *Mol Microbiol* **79**, 1655–1669.

6 Lang NJ & Fay P (1971) The heterocysts of blue-green algae II. Details of ultrastructure. *Proc R Soc Lond B Biol Sci* **178**, 193–203.

7 Wildon DC & Mercer FV (1963) The ultrastructure of the heterocyst and akinete of the blue-green algae. *Arch Mikrobiol* **47**, 19–31.

8 Flores E, Herrero A, Wolk CP & Maldener I (2006) Is the periplasm continuous in filamentous multicellular cyanobacteria? *Trends Microbiol* **14**, 439–443.

9 Wilk L, Strauss M, Rudolf M, Nicolaisen K, Flores E, Kuhlbrandt W & Schleiff E (2011) Outer membrane continuity and septosome formation between vegetative cells in the filaments of *Anabaena* sp. PCC 7120. *Cell Microbiol* **13**, 1744–1754.

10 Meeks JC, Elhai J, Thiel T, Potts M, Larimer F, Lamerdin J, Predki P & Atlas R (2001) An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium. *Photosynth Res* **70**, 85–106.

11 Adams DG & Duggan PS (1999) Heterocyst and akinete differentiation in cyanobacteria. *New Phytol* **144**, 3–33.

12 Argueta C & Summers ML (2005) Characterization of a model system for the study of *Nostoc punctiforme* akinetes. *Arch Microbiol* **183**, 338–346.

13 Maldener I, Summers ML & Sukenik A (2014) Cellular differentiation in filamentous cyanobacteria. In The Cell Biology of Cyanobacteria (Flores E & Herrero A, eds), pp. 263–291. Caister Academic Press, Norfolk, UK.

14 Meeks JC & Elhai J (2002) Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. *Microbiol Mol Biol Rev* **66**, 94–121.

15 Rippka R, Deruelles J, Waterbury JB, Herdman M & Stanier RY (1979) Generic assignments, strain stories and properties of pure cultures of cyanobacteria. *J Gen Microbiol* **111**, 1–61.

16 Giddings TH Jr & Staehelin LA (1978) Plasma membrane architecture of *Anabaena cylindrica*: occurrence of microplasmodesmata and changes associated with heterocyst development and the cell cycle. *Eur J Cell Biol* **16**, 235–298.

17 Nürnberg DJ, Mariscal V, Bornikoel J, Nieves-Morion M, Krauss N, Herrero A, Maldener I, Flores E & Mullineaux CW (2015) Intercellular diffusion of a fluorescent sucrose analog via the septal junctions in a filamentous cyanobacterium. *MBio* **6**, e02109.

18 Omairi-Nasser A, Mariscal V, Austin JR II & Haselkorn R (2015) Requirement of Fra proteins for communication channels between cells in the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. PCC 7120. *Proc Natl Acad Sci USA* **112**, E4458–E4464.

19 Omairi-Nasser A, Haselkorn R & Austin J II (2014) Visualization of channels connecting cells in filamentous nitrogen-fixing cyanobacteria. *FASEB J* **28**, 3016–3022.

20 Mariscal V (2014) Cell-cell joining proteins in heterocyst-forming cyanobacteria. In The Cell Biology of Cyanobacteria (Flores E & Herrero A, eds), pp. 239–304. Caister Academic Press, Norfolk, UK.

21 Maldener I & Forchhammer K (2015) Requirement of cell wall remodelling for cell-cell communication and cell differentiation in filamentous cyanobacteria of the order Nostocales. In Biological Nitrogen Fixation (de Bruijn FJ, ed.), pp. 873–878. Wiley-Blackwell, Hoboken, New Jersey.

22 Berendt S, Lehner J, Zhang YV, Rasse TM, Forchhammer K & Maldener I (2012) Cell wall amidase AmiC1 is required for cellular communication and heterocyst development in the cyanobacterium *Anabaena* PCC 7120 but not for filament integrity. *J Bacteriol* **194**, 5218–5227.

23 Priyadarshini R, de Pedro MA & Young KD (2007) Role of peptidoglycan amidases in the development and morphology of the division septum in *Escherichia coli*. *J Bacteriol* **189**, 5334–5347.

24 de Souza RF, Anantharaman V, de Souza SJ, Aravind L & Gueiros-Filho FJ (2008) AMIN domains have a predicted role in localization of diverse periplasmic protein complexes. *Bioinformatics* **24**, 2423–2426.

25 Uehara T, Parzych KR, Dinh T & Bernhardt TG (2010) Daughter cell separation is controlled by cytokinetic ring-activated cell wall hydrolysis. *EMBO J* **29**, 1412–1422.

26 Uehara T, Dinh T & Bernhardt TG (2009) LytM-domain factors are required for daughter cell separation and rapid ampicillininduced lysis in *Escherichia coli*. *J Bacteriol* **191**, 5094–5107.

27 Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J *et al.* (2014) Pfam: the protein families database. *Nucleic Acids Res* **42**, D222–D230.

28 Kumar A, Kumar S, Kumar D, Mishra A, Dewangan RP, Shrivastava P, Ramachandran S & Taneja B (2013) The structure of Rv3717 reveals a novel amidase from *Mycobacterium tuberculosis*. *Acta Crystallogr D* **69**, 2543–2554.

29 Prigozhin DM, Mavrici D, Huizar JP, Vansell HJ & Alber T (2013) Structural and biochemical analyses of *Mycobacterium tuberculosis* N-acetylmuramyl-L-alanine amidase Rv3717 point to a role in peptidoglycan fragment recycling. *J Biol Chem* **288**, 31549–31555.

30 Yang DC, Tan K, Joachimiak A & Bernhardt TG (2012) A conformational switch controls cell wall-remodelling enzymes required for bacterial cell division. *Mol Microbiol* **85**, 768–781.

31 Rocaboy M, Herman R, Sauvage E, Remaut H, Moonens K, Terrak M, Charlier P & Kerff F (2013) The crystal structure of the cell division amidase AmiC reveals the fold of the AMIN domain, a new peptidoglycan binding domain. *Mol Microbiol* **90**, 267–277.

32 Büttner FM, Renner-Schneck M & Stehle T (2015) X-ray crystallography and its impact on understanding bacterial cell wall remodeling processes. *Int J Med Microbiol* **305**, 209–216.

33 Büttner FM, Zoll S, Nega M, Götz F & Stehle T (2014) Structure-function analysis of *Staphylococcus aureus* amidase reveals the determinants of peptidoglycan recognition and cleavage. *J Biol Chem* **289**, 11083–11094.

34 Zoll S, Patzold B, Schlag M, Götz F, Kalbacher H & Stehle T (2010) Structural basis of cell wall cleavage by a staphylococcal autolysin. *PLoS Pathog* **6**, e1000807.

35 Li Q, Cheng W, Morlot C, Bai XH, Jiang YL, Wang W, Roper DI, Vernet T, Dong YH, Chen Y *et al.* (2015) Full-length structure of the major autolysin LytA. *Acta Crystallogr D* **71**, 1373–1381.

36 Bernhardt TG & Vollmer W (2013) Mechanisms of bacterial morphogenesis and their subversion by phages. *Curr Opin Microbiol* **16**, 728–730.

37 Hoiczyk E & Baumeister W (1995) Envelope structure of four gliding filamentous cyanobacteria. *J Bacteriol* **177**, 2387–2395.

38 Jürgens UJ, Drews G & Weckesser J (1983) Primary structure of the peptidoglycan from the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6714. *J Bacteriol* **154**, 471–478.

39 Hahn A & Schleiff E (2014) The cell envelope. In The Cell Biology of Cyanobacteria (Flores E & Herrero A, eds), pp. 29–88. Caister Academic Press, Norfolk, UK.

40 Matias VR, Al-Amoudi A, Dubochet J & Beveridge TJ (2003) Cryo-transmission electron microscopy of frozen-hydrated sections of *Escherichia coli* and *Pseudomonas aeruginosa*. *J Bacteriol* **185**, 6112–6118.

41 Yao X, Jericho M, Pink D & Beveridge T (1999) Thickness and elasticity of gram-negative murein sacculi measured by atomic force microscopy. *J Bacteriol* **181**, 6865–6875.

42 Singh SK, Parveen S, SaiSree L & Reddy M (2015) Regulated proteolysis of a cross-link-specific peptidoglycan hydrolase contributes to bacterial morphogenesis. *Proc Natl Acad Sci USA* **112**, 10956–10961.

43 Rudolf M, Tetik N, Ramos-Leon F, Flinner N, Ngo G, Stevanovic M, Burnat M, Pernil R, Flores E & Schleiff E (2015) The peptidoglycan-binding protein SjcF1 influences septal junction function and channel formation in the filamentous Cyanobacterium *Anabaena*. *MBio* **6**, e00376.

44 Lutzner N, Patzold B, Zoll S, Stehle T & Kalbacher H (2009) Development of a novel fluorescent substrate for Autolysin E, a bacterial type II amidase. *Biochem Biophys Res Commun* **380**, 554–558.

45 Korndorfer IP, Danzer J, Schmelcher M, Zimmer M, Skerra A & Loessner MJ (2006) The crystal structure of the bacteriophage PSA endolysin reveals a unique fold responsible for specific recognition of *Listeria* cell walls. *J Mol Biol* **364**, 678–689.

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

47 Mayer MJ, Garefalaki V, Spoerl R, Narbad A & Meijers R (2011) Structure-based modification of a *Clostridium difficile*-targeting endolysin affects activity and host range. *J Bacteriol* **193**, 5477–5486.

48 Low LY, Yang C, Perego M, Osterman A & Liddington R (2011) Role of net charge on catalytic domain and influence of cell wall binding domain on bactericidal activity, specificity, and host range of phage lysins. *J Biol Chem* **286**, 34391–34403.

49 Sambrook J, Fritsch EF & Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York.

50 Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.

51 Glauner B, Holtje JV & Schwarz U (1988) The composition of the murein of *Escherichia coli*. *J Biol Chem* **263**, 10088–10095.

52 Hayashi K (1975) A rapid determination of sodium dodecyl sulfate with methylene blue. *Anal Biochem* **67**, 503–506.

53 Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207–234.

54 Evans PR (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D* **67**, 282–292.

55 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D* **67**, 235–242.

56 Kabsch W (2010) Xds. *Acta Crystallogr D* **66**, 125–132.

57 Stein N (2008) CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *J Appl Crystallogr* **41**, 641–643.

58 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.

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

60 Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* **66**, 213–221.

61 Schrödinger LLC. The PyMOL Molecular Graphics System, Version 1.7.2. https://www.pymol.org/citing

62 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004) UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612.

63 McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP & Lopez R (2013) Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res* **41**, W597–W600.

## Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:
**Table S1.** Plasmids.
**Table S2.** Primers.

Mini-Review

# X-ray crystallography and its impact on understanding bacterial cell wall remodeling processes

Felix Michael Büttner[a], Michaela Renner-Schneck[a], Thilo Stehle[a,b,*]

[a] Interfaculty Institute of Biochemistry, University of Tübingen, Hoppe-Seyler-Straße 4, 72076 Tübingen, Germany
[b] Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, TN, USA

## ABSTRACT

The molecular structure of matter defines its properties and function. This is especially true for biological macromolecules such as proteins, which participate in virtually all biochemical processes. A three dimensional structural model of a protein is thus essential for the detailed understanding of its physiological function and the characterization of essential properties such as ligand binding and reaction mechanism. X-ray crystallography is a well-established technique that has been used for many years, but it is still by far the most widely used method for structure determination. A particular strength of this technique is the elucidation of atomic details of molecular interactions, thus providing an invaluable tool for a multitude of scientific projects ranging from the structural classification of macromolecules over the validation of enzymatic mechanisms or the understanding of host–pathogen interactions to structure-guided drug design. In the first part of this review, we describe essential methodological and practical aspects of X-ray crystallography. We provide some pointers that should allow researchers without a background in structural biology to assess the overall quality and reliability of a crystal structure. To highlight its potential, we then survey the impact X-ray crystallography has had on advancing an understanding of a class of enzymes that modify the bacterial cell wall. A substantial number of different bacterial amidase structures have been solved, mostly by X-ray crystallography. Comparison of these structures highlights conserved as well as divergent features. In combination with functional analyses, structural information on these enzymes has therefore proven to be a valuable template not only for understanding their mechanism of catalysis, but also for targeted interference with substrate binding.

© 2015 Elsevier GmbH. All rights reserved.

## X-ray crystallography: principles, possibilities and challenges

### Overview

The determination of the crystal structure of diamonds by William Henry Bragg and his son William Lawrence Bragg in 1913 can be viewed as the birth of X-ray crystallography. Thus, this method has only recently turned one hundred years old, and can now look back on a century in which it has often enriched and sometimes revolutionized scientific research. To mark the occasion, the year 2014 has been announced as the "International Year of Crystallography" (iycr2014, 2014) and the journals *Nature* and *Science* have dedicated special issues to review the historical milestones of X-ray crystallography, its achievements, developments, and future prospects (NATURE, 2014; SCIENCE, 2014). The increasing level of automation in the crystallographic pipeline over these last 100 years has led to a tremendous and constantly growing number of structures deposited in the PDB today, the PDB features over 100.000 structures. In combination with the rapid technical evolution of X-ray crystallographic techniques in areas such as membrane protein structure determination or room-temperature structure determination at synchrotrons, as well as the possibilities offered by free electron laser X-ray (XFEL) sources (Garman, 2014), this promises that macromolecular crystallography will continue to have considerable impact in a broad range of scientific research fields in the future.

X-ray crystallography is not a direct imaging technique that focuses visible light scattered from objects through refractive lenses to create a magnified image of the object. Rather, it exploits the fact that X-rays with wavelengths between 0.05 and 5.0 nm are scattered by the electron shells of atoms and thus provide the

possibility to obtain structural information of molecules at near-atomic resolution. However, due to the refractive index of X-rays in different materials, which is essentially equal and close to unity, it is not possible to obtain direct atomic resolution images of a single protein molecule (or other macromolecule) through simple focusing of scattered X-rays (Rupp, 2010; Sumner, 2014). Nevertheless, diffraction images can be obtained from protein crystals in X-ray diffraction experiments. These images carry information about the content of the crystal's unit cell (the protein of interest), but this information is encoded in intensity distributions of reflections in "reciprocal space" and thus not easily accessible. With the help of Fourier transformations, this information can be "translated" back into molecular "real" space, giving rise to an image of the crystallized molecule. The Fourier transformation is a straightforward mathematical operation that requires two terms as Fourier coefficients: (i) the structure factor amplitudes, which can be obtained from the recorded diffraction spot intensities, and (ii) the relative phase angle corresponding to each observed diffraction spot (Rupp, 2010). Since these phase angles are not directly accessible by experimental methods, they must be obtained in additional so-called phasing approaches, which can involve further experiments or molecular replacement calculations with the help of the phases from related known structures. This is generally known as the "phase problem" in crystallography, and it is one reason X-ray crystallographic structure determinations can remain challenging even today. Once initial phases are determined, a first electron density map can be calculated, and this map provides the basis for molecular model building and structural refinement (Fig. 1).

*Experimental approach and challenges*

An X-ray crystallographic structure determination requires one or typically several crystals of the molecule of interest. In 1937 James Sumner demonstrated that proteins can be crystallized and must therefore have a regular, ordered structure (Sumner, 1937). Despite efforts to standardize and automate this process, growing crystals remains a major bottleneck in crystallography. Crystal formation requires sufficient amounts of highly pure protein, which is usually obtained using recombinant expression systems in bacteria, yeast, insect cells, or mammalian cell lines. Using these approaches, even challenging proteins carrying post-translational modifications or membrane proteins can often be expressed.

Although the parameters governing the process of protein crystallization are now better understood, it is not yet possible to predict the conditions under which a particular protein will crystallize (Garman, 2014), and thus crystallization conditions have to be screened empirically. Moreover, since diffraction power can vary tremendously from crystal to crystal, several additional rounds of fine-screening and crystal optimization are often necessary to obtain diffraction-quality crystals. Even with the use of robotic platforms that can routinely dispense low-volume drops (as low as 50 nL protein + 50 nL of precipitant solution), this screening typically still requires milligram amounts of pure, homogeneous protein solutions that are not always easy to obtain.

Wavelength (energy) and brilliance (flux) of the X-ray beam itself are also important factors influencing diffraction power and data quality. In some cases, such as in experimental phasing strategies with the help of anomalous scatterers, it is even necessary to record several data sets (see below) at different wavelengths. The evolution of storage ring sources to the currently available third-generation synchrotron sources with tunable wavelengths in conjunction with fast and accurate X-ray detectors has greatly facilitated the performance and efficiency of X-ray crystallography in the last decades. Nowadays, even weakly diffracting or smaller crystals can be used for structure determination (Garman, 2014).
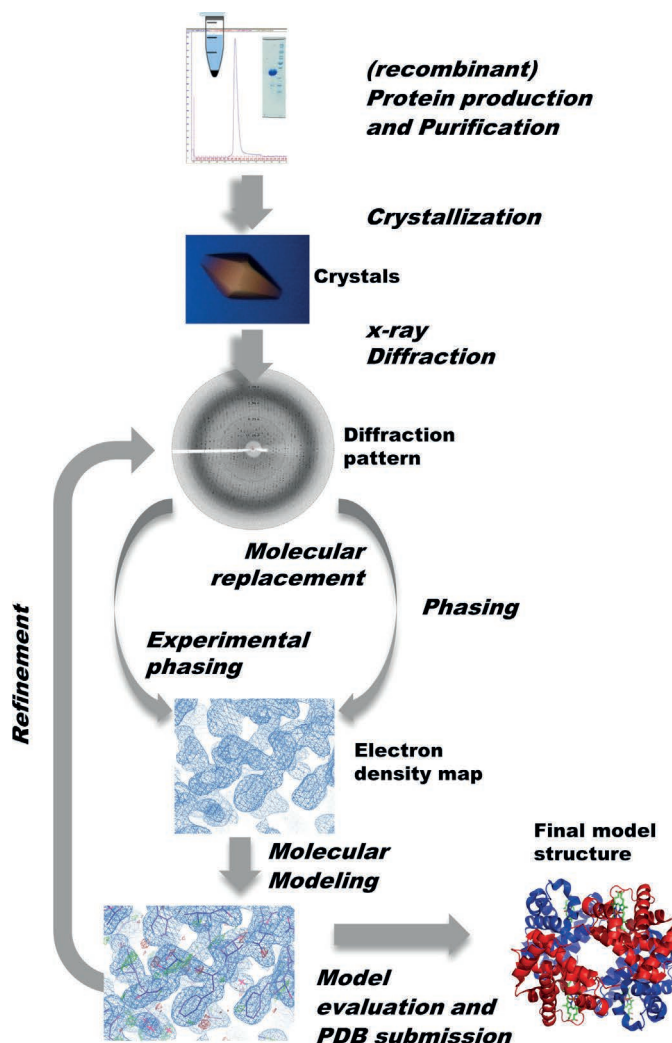


**Fig. 1.** X-ray crystallographic workflow. Schematic diagram showing the workflow for macromolecular structure determination by X-ray crystallography.

Once X-ray data from a crystal are available, the intensities of the reflections need to be extracted from the data images and processed further. Here, all spots recorded on the diffraction images are indexed according to the crystal's space group, and their intensities are subsequently integrated. The processed data (the "data set") then forms the basis for phase determination. A number of streamlined program packages are available nowadays that can overcome many difficulties in data interpretation and phasing with limited user adjustment. In some cases, crystallographic software packages are even capable of solving structures without human intervention. However, since data processing and phasing have a major impact on the resulting structural information, while leaving room for dramatic misinterpretations at several stages, it is still essential to assess their outputs for biochemical plausibility. Moreover, the automated approaches typically fail when challenging macromolecules, such as large, poorly diffracting complexes or complicated crystal packing arrangements are analyzed. Therefore, the input of human intelligence and experience is still essential.

*Evaluating crystal structures*

A typical scientist does perhaps not need to understand the intricate details and challenges of molecular structure determination, but he or she must be able to critically evaluate a structure

that has been published. With very few exceptions, X-ray crystallographic structures are deposited in the Protein Data Bank (PDB) (www.pdb.org). The PDB employs a set of quality controls which ensure that the deposited structures conform to standard geometrical parameters, such as bond distances, bond angles, or protein backbone torsion angles. Bond distances and bond angles are commonly reported using root mean square (rms)-deviations that describe how far (in Å or degrees, respectively – see below) bonds or angles deviate from ideal values. A trustworthy structure might have rms deviations of <0.015 Å and <1.5° for its bonds and angles, respectively. Usually protein geometry is also evaluated in a Ramachandran diagram, which is a highly useful indicator of the stereochemical quality of the backbone conformation of a structural model (Ramachandran, 1963). In such diagrams, the conformation of the backbone of a protein is plotted with two separate torsion angle values for each amino acid, and as very few values are allowed for the backbone of protein residues, the diagram can serve to highlight suspect amino acids. Ideally, a trustworthy structural model should have favorable torsion angle values for over 90% of its residues, and no residues in disallowed regions.

Important parameters assessing how well a model fits the experimental data are the so-called reliability values (R-values: $R_{work}$ and $R_{free}$). $R_{work}$ numerically quantifies the overall agreement between the observed diffraction amplitudes and "virtual" amplitudes calculated using the structural model. A model that agrees well with its diffraction data will give rise to calculated amplitudes that are very similar to the observed ones, and this is reflected in a low $R_{work}$ value. $R_{work}$ is reported as a percentage value, with a good value ranging from 15% to 25%. The typically also reported $R_{free}$ is a cross-validation R-value. Its careful monitoring helps to prevent overparameterization of the model during model building and refinement. The $R_{free}$ value is calculated in the same way as the $R_{work}$, but only from a small subset of the experimental data that is separated from the working data set and not used during structural refinement. The $R_{free}$ should closely track the value for $R_{work}$. A small difference between them ($\leq$5%) is acceptable, while a difference >10% might well indicate a problem with the structure determination or the quality of the model. B-factors (or temperature factors) provide important information on the mobility of the crystallized protein or its components (such as ligands or solvent). Here, a high value (e.g. 80 Å$^2$) might indicate a problem, particularly if ligand atoms have significantly elevated B-factors compared to surrounding amino acids. In such a case, the ligand may not be bound well in the binding pocket, and might have high mobility and/or low occupancy.

Last but not least, the resolution to which reflections could be measured on the diffraction images gives an impression on how distinct and reliable atomic details can be deduced from the electron density map, and thus how significant conclusions on molecular interactions are. Good resolutions in X-ray crystallography range from 1.5 to 2.5 Å (1 Å = 0.1 nm), but higher resolutions of 1.0 Å and beyond have also been achieved. Most structures deposited in the PDB have a resolution between 1.5 and 2.5 Å (www.pdb.org), but there are also quite a few structures at low resolution >3 Å. While these structures also can provide valuable information about overall fold and domain organization, the interpretation of details such as hydrogen bond distances or side chain conformations must be done with caution.

Apart from interpreting the statistics for a published structural model, a good strategy to evaluate its reliability is to inspect the electron density map, which can also be downloaded from the PDB for many structures. This allows for an easy assessment of how well a structural model agrees with a corresponding electron density map. Structures that contain bound ligands, such as cofactors, ions, or substrates, can also be evaluated in this manner, and inspecting electron density maps can often provide clues about how well a ligand is bound to a binding site, and how reliable the assigned contacts are.

## A case study of the impact of X-ray crystallography: the cell wall remodeling enzyme class of N-acetylmuramoyl-ʟ-alanine amidases

### Introduction

The impact of X-ray crystallography on our detailed understanding of biological systems can of course be illustrated with many different examples. We have chosen to examine the contribution this method has made to the characterization of a class of bacterial cell wall remodeling enzymes. We focus on structurally and functionally similar domains of a range of N-acetylmuramoyl-ʟ-alanine amidases. These enzymes cleave the amide bond between the carbohydrate and peptide moieties of the bacterial peptidoglycan (PGN), a major component of the bacterial cell wall. Composed of the alternating carbohydrates N-acetylmuramic acid and N-acetyl glucosamine as well as a peptide stem, the PGN forms the backbone of all bacterial cell walls. While it serves as a barrier that protects bacteria, the cell wall is necessarily also a very dynamic polymer that undergoes constant change, or remodeling. Synthesis, degradation, and recycling of components by a broad variety of enzymes are responsible for this plasticity. Interestingly, N-acetylmuramoyl-ʟ-alanine amidases exhibit rather low amino acid conservation although they perform generally similar functions, and it is tempting to speculate that the sequence diversity is linked to differential recognition of substrates and different interaction mechanisms with additional domains. Thirty-three amidase structures of fifteen different bacterial N-acetylmuramoyl-ʟ-alanine amidases are currently known, and the majority of these have been determined by X-ray crystallography (www.pdb.org). Given this substantial data base of structures, a structural comparison is warranted to better understand the function, targeting, and regulation of bacterial cell wall amidases and to unravel generally conserved structural features, as well as distinct properties.

### Bacterial amidases

Bacterial cell walls are divided into two major classes: Gram-positive and Gram-negative. Gram-positive cell walls consist of one lipid bilayer and a thick PGN layer, as well as associated proteins and teichoic acids. Gram-negative bacteria have an inner and an outer lipid bilayer, associated proteins, lipopolysaccharides, and a thin PGN-layer separating the two membranes. In addition, the amino acid composition differs between the two types of cell walls (reviewed by Schleifer and Kandler (1972)). The PGN of Gram-positive bacteria generally contains ʟ-lysine and a pentaglycine bridge between two connected peptide stems, while Gram-negative bacteria incorporate *meso*-diaminopimelic acid (*meso*-DAP) instead of lysine and link the peptide stems directly with each other. Furthermore, different species modify their PGN in manifold ways. A prominent representative of Gram-positive bacteria, *S. aureus*, has a PGN of subtype A3α, meaning its PGN comprises ʟ-Ala-ᴅ-iGlu-ʟ-Lys-ᴅ-Ala-ᴅ-Ala and a pentaglycine linker (de Jonge et al., 1992; Schleifer and Kandler, 1972). Moreover, *S. aureus* does not form anhydro-MurNAc (Boneca et al., 2000), but amidates isoglutamate to form isoglutamine. Further variations such as different linker lengths or substitution of glycines with alanine or serine residues may also occur (de Jonge et al., 1992; Schleifer and Kandler, 1972). On the other hand, the PGN of bacteria such as *B. subtilis* and *E. coli* belong to the A1γ type, indicating a direct linkage

and use of *meso*-DAP (Schleifer and Kandler, 1972). Modifications in these PGN include anhydro-MurNAc at the carbohydrate terminus, the occasional absence of unlinked, terminal D-Ala residues, and different amidation patterns of *meso*-DAP carboxyl groups, whereas the backbone isoglutamate is not amidated (Atrih et al., 1999; Schleifer and Kandler, 1972; Warth and Strominger, 1971).

It is not surprising that the variability of PGN in different species is reflected by differences in the specificity of amidases that process these networks. In addition to cleaving PGN at defined locations, amidases have a large number of additional requirements. They are needed for mother-daughter cell separation during cytokinesis (AmiA (Heilmann et al., 1997)), they play a crucial role in recycling of PGN components (AmpD (Carrasco-Lopez et al., 2011), Rv3717 (Prigozhin et al., 2013)), and they are even required for intercellular communication and nutrient exchange in filamentous bacteria (AmiC2, Lehner et al., 2011). Amidases furthermore play a role in bacterial defense, as some types are secreted to digest the cell walls of adversaries (T6SS, (Chou et al., 2012)) and bacteriophages utilize these enzymes to infect their hosts more efficiently. The eukaryotic peptidoglycan recognition proteins (PGRPs) are structurally closely related to bacterial amidases. They share the discussed Amidase_2 fold (see below) and recognize similar structural motifs. While all PGRPs trigger an immune response against bacteria upon infection, some are even able to cleave PGN (Wang et al., 2003).

Structural biology techniques have contributed significantly to a better understanding of the function, targeting, and regulation of bacterial amidases. The analysis of crystal structures, particularly of complexes with ligands or ligand components, allows the identification of key residues of the enzymes, providing a direct link to function in catalysis or ligand binding. Insights into the binding mode and the catalytic mechanism obviously facilitate the development of substrate-based ligand mimetics that can serve as inhibitors. Although the current research focus still lies mostly on acquisition of basic knowledge on bacterial mechanisms, therapeutic interests are a major long-term objective in the structure/function analyses of amidases. Enzymes of potentially pathogenic organisms such as *S. aureus*, *S. epidermidis*, *P. aeruginosa*, *E. coli*, *Mycobacterium tuberculosis*, and *Neisseria meningitides* have been investigated, with one of the goals being the development of new antibiotics.

*Phylogeny and composition of amidases*

Amidases are divided into families, and the Pfam database (Bateman et al., 2002) lists Amidase_2 and Amidase_3 as the two major families. Both are $Zn^{2+}$-dependent hydrolases in which the catalytic domains adopt a globular, mixed αβ-fold. Three other amidase families, Amidase_5, Amidase_6, and Amidase_02C, are listed, however, knowledge on them is very limited and no structures of members of these families have been solved to date. We will therefore review structural features of the Amidase_2 and Amidase_3 families.

Most amidase domains are embedded in larger protein chains that contain additional domains, often called repeat or AMIN domains. These domains typically play a role in targeting and anchoring the amidase to the desired environment (Zoll et al., 2012). The repeats may also act as regulatory domains, which can be connected by a variable and often flexible linker to the catalytically active amidase. Repeat domains can lie N-terminally (Zoll et al., 2012), as well as C-terminally (Rocaboy et al., 2013) of the catalytic domains, and they are present in different numbers in different amidases. Due to space limitations, we will focus here on a comparison of catalytic amidase domains (Fig. 2).

A growing number of amidase structures are available in the PDB, the majority of which were published within the last five years (Fig. 2A). However, these structures are still from relatively few organisms, and so far a comprehensive view of amidase structures is lacking. The sequence homology is generally low among the currently solved structures, and conservation is limited to a small number of residues (Fig. 2B and C). Surprisingly, however, the amidases nevertheless share similar folds, demonstrating that even a low level of amino acid conservation can still lead to the same fold. Sequence analysis of all fourteen enzymes solved by X-ray crystallography reveals a phylogenic tree with three major branches (Fig. 2A, Rv3717 being listed twice). The upper two belong to the Amidase_2 family, whereas the lower branch represents structures of the Amidase_3 family. Interestingly, the two branches within the Amidase_2 family correlate with the Gram-characteristics of the source organisms. The three branches will subsequently be referred to as AmiA-type, AmiD-type, and AmiC-type (see Fig. 2A). Moreover, the two protein families both have a globular, mixed αβ-fold, but present distinct differences that render structural comparison amongst families less meaningful. The assessment of all currently published bacterial amidase structures is therefore divided into two chapters that discuss the Amidase_2 and Amidase_3 family members separately.

*The Amidase_2 family*

Nine Amidase_2 enzymes have been crystallized, with 23 total structures in native, mutant, or liganded states. The secondary structure features are depicted in Fig. 2B, and they are almost completely conserved among the Amidase_2 enzymes. Superposition of all structures (Fig. 3A) shows that a six-stranded, mostly parallel, β-sheet forms the central structural element. The back of this sheet faces two of the α-helices (α4, α7), stabilizing the overall fold, while its front is largely solvent accessible and forms the bottom of the substrate binding groove. This narrow groove is built by the flanking helix α2 on the right side and two additional α-helices (α5, α6) on the left hand side (Fig. 2B and 3A). The shape and size of the groove varies substantially among the family members, in accord with the somewhat different PGN structures found in bacteria, which results in different ligand structures bound by the enzymes. The groove is also bordered by a loop and harbors the three zinc-binding as well as catalytically and structurally stabilizing residues. The catalytic zinc ion is typically coordinated by histidines 265 and 370 (numbering and described interactions according to AmiA (Büttner et al., 2014) for whole Amidase_2 section) and aspartic acid 384 (Fig. 2B). In two enzymes, PlyL and XlyA, which are annotated as "bacterial" enzymes but have a bacteriophage background, this aspartate is replaced with a cysteine.

Although the Amidase_2 sequences are highly divergent apart from the zinc-ligating residues, some additional amino acids are highly conserved in all structures, and these conservations provide clues about residues required for function. Threonine 267 is likely a part of the carbohydrate binding site and forms contacts with MurNAc, while tryptophan 310 interacts with L-lysine or *meso*-DAP of PGN. Asparagine 317 stabilizes the tertiary structure of the area and also interacts with the backbone carbonyl oxygen atoms of D-Ala and L-Lys/*meso*-DAP of PGN. Glutamate 324 is the catalytic residue that acts as proton acceptor during catalysis, enabling a water molecule to attack the scissile amide bond of PGN. Histidine 382 or a lysine stabilize the tetrahedral intermediate state during catalysis, while proline 385 enables the correct positioning of the preceding, zinc binding aspartate 384 (or cysteine in PlyL and XlyA) with unusual angles. The remaining well conserved and largely hydrophobic residues play a structural role as they fix secondary structures and stabilize the amidase fold.

AmiD, AmpDh2, and AmpDh3 have an additional β-loop that is inserted between strand β5 and helix α4. Martinez-Caballero and colleagues propose that this loop plays an important role in carbohydrate binding, and they could indeed show for AmpDh2
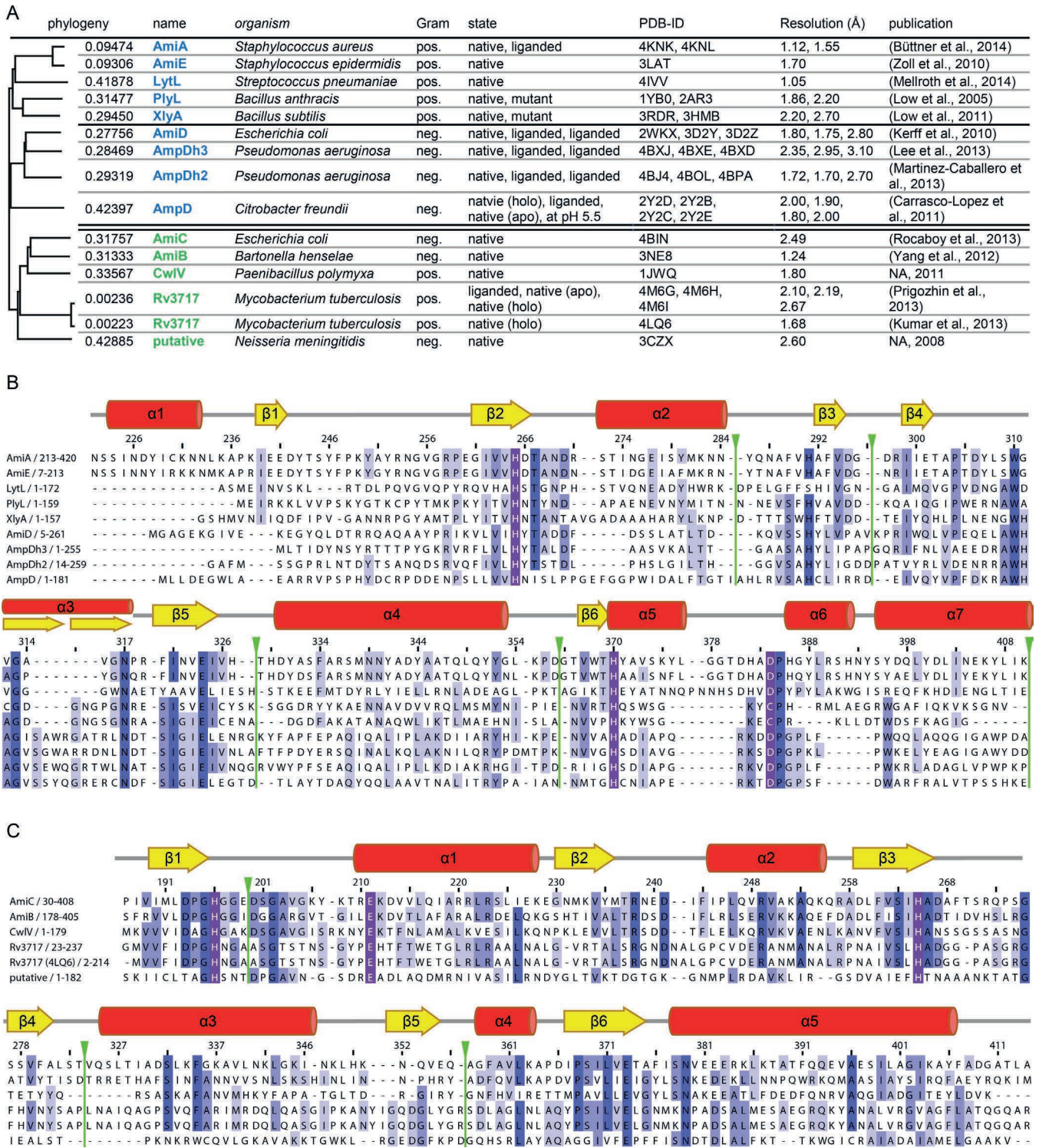
**Fig. 2.** Multivalent analysis of N-acetylmuramoyl-ʟ-alanine amidases. (A) N-acetylmuramoyl-ʟ-alanine amidases structures solved to date, with a phylogram obtained from the ClustalW2 webserver (Larkin et al., 2007) using the sequences deposited with the PDB files. The respective organism, Gram-stain, enzyme state, PDB-code, resolution, and original publication are listed in each case. The upper part of the panel presents the Amidase_2 family (in blue), subdivided into AmiA and AmiD type, the lower part shows the Amidase_3 family (in green). (B) Multisequence alignment of Amidase_2 sequences listed in panel A. Secondary structure features are represented with yellow arrows for β-strands, red ribbons for α-helices, and gray lines for loops. Identical residues are indicated with a blue background, with the color increasing in intensity for higher conservation; zinc-coordinating residues are highlighted with purple boxes; green lines mark omission of unconserved residues. (C) Multisequence alignment of the Amidase_3 sequences listed in panel A, using the same color code as in panel B.

that the loop shifts upon binding of a tetrasaccharide (Martinez-Caballero et al., 2013). The additional C-terminal domain (Fig. 3A), which contains four α-helices, may participate in PGN-binding or enhance substrate specificity and binding of larger PGN-fragments

(Kerff et al., 2010). The N-terminal extension plays a role in the dimerization of these enzymes. AmiD and AmpDh2 are thought to be dimers, while AmpDh3 is a tetramer (Lee et al., 2013; Martinez-Caballero et al., 2013). In addition, the N-terminus of one copy
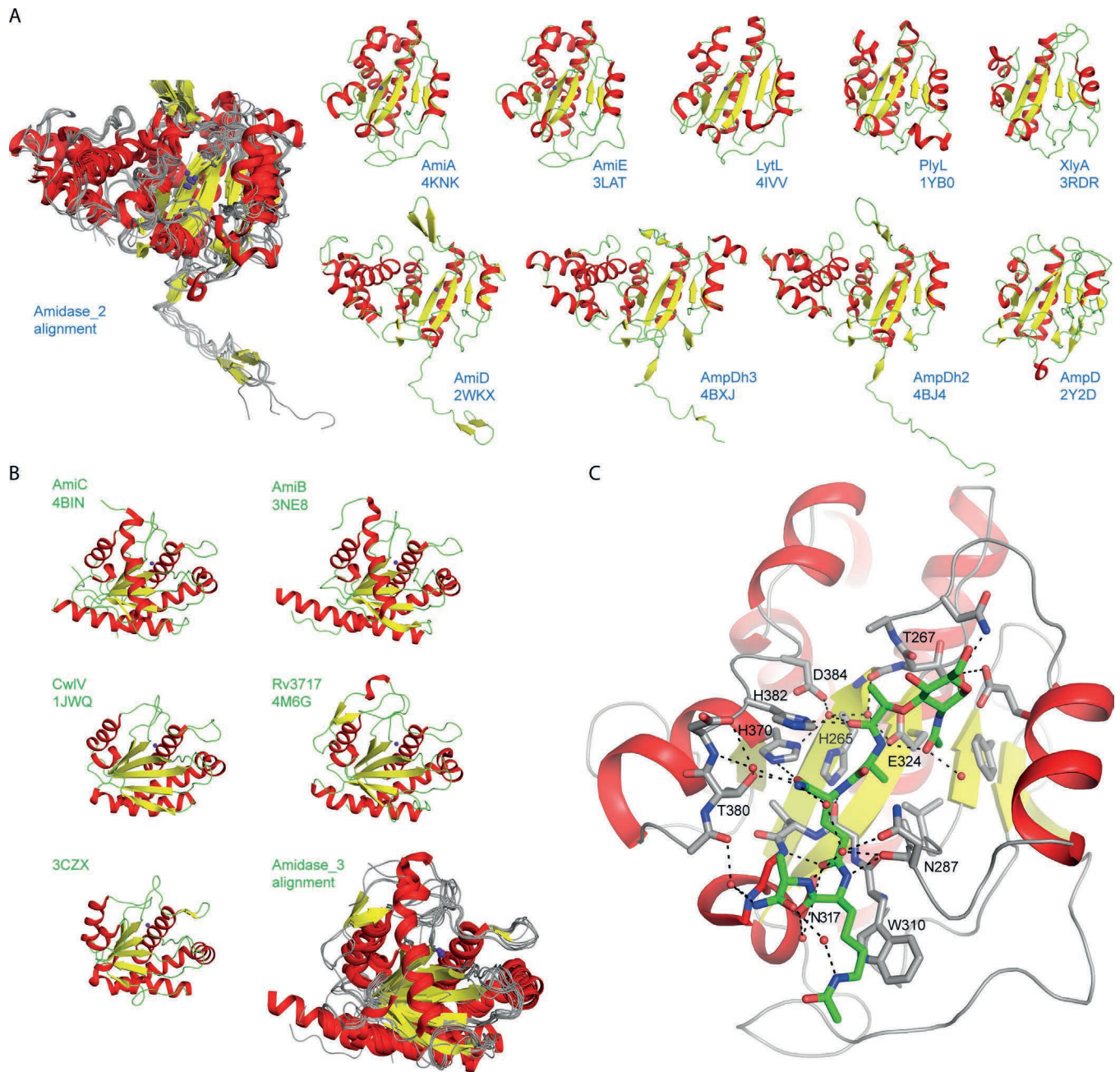
**Fig. 3.** Structures of N-acetylmuramoyl-ʟ-alanine amidases. (A) Collage of nine known native Amidase_2 enzyme structures placed next to an overlay of all Amidase_2 structures. Secondary structure features are represented with yellow arrows for β-strands, red ribbons for α-helices and gray (overlay) or green (single structures) lines for loops. Names and pdb codes are listed next to each structure. (B) Collage of the five known native Amidase_3 enzyme structures next to an overlay of all Amidase_3 structures, with the color code introduced in panel (A). (C) Structure of the AmiA amidase domain, a member of the Amidase_2 family, bound to its ligand muramyltetrapeptide (MtetP). Coloring and orientation of AmiA are as in (A), MtetP is shown in green (carbons), red (oxygens) and blue (nitrogens). Hydrogen bonds are depicted as black dashes.

completes the peptide stem binding site of the other. Models with larger PGN-fragments hint at an explanation for how the AmiD-type amidases can bind and cleave polymeric PGN. AmpD is the only cytosolic enzyme of the AmiD-type, and it becomes involved in PGN-recycling once PGN subunits have been cleaved by a lytic transglycosylase and transported into the cell by AmpG. It therefore has a preference for PGN-monomers with anhydro-MurNAc and lacks the additional features discussed for AmiD, AmpDh2, and AmpDh3 (Carrasco-Lopez et al., 2011). Especially interesting is the comparison of the X-ray structure of AmpD in an active "open" state with its NMR-structure in an inactive "closed" state. Major rearrangements, as much as 17 Å shifts of loops at the N-terminus

and after strands and helices β2, β5, and α5 along with breaking several salt bridges converts the enzyme between states. In order to elucidate the consequences of these rearrangements, the authors performed molecular dynamics calculations, and these hint at an allosteric triggering of the rearrangement (Carrasco-Lopez et al., 2011). As no such rearrangements have been reported for other Amidase_2 members, the described activation mechanism may only be required for a cytosolic amidase.

AmiE and LytA have been extensively studied, with focus on the substrate specificity and the role of anchoring cell wall binding domains, including small angle X-ray scattering (SAXS) to determine the structures of these multi-domain enzymes in solution

(Mellroth et al., 2014; Zoll et al., 2010, 2012). Although AmiE is excreted into the periplasm and LytA resides in the cytosol, similar patterns of substrate recognition were observed. However, in contrast to AmiE, LytA forms dimers, and the dimer is significantly more active compared to the monomer (Mellroth et al., 2014). Investigation of the closely related AmiA from *S. aureus* led to the identification of a dense interaction network between enzyme and substrate, and subsequently the specificity as well as a plausible reaction mechanism (Büttner et al., 2014). Since the participating residues are highly conserved, this mechanism likely is used by all Amidase_2 enzymes (Büttner et al., 2014; Paget et al., 1999).

The phage-derived enzymes XlyA and PlyA were instigated with a special focus on lytic activity. XlyA and a mutant with altered surface charge were tested, and a positive charge increased activity. In addition to that, the cell wall binding domain seems not to significantly interact with the catalytic domain (Low et al., 2011). In contrast, PlyA is likely to be regulated by the C-terminal regulator and cell wall binding domain, which also may play a role for specificity of the target (Low et al., 2005).

*The Amidase_3 family*

The Amidase_3 family is less well characterized, and only five different enzymes with a total of eight structures are deposited in the PDB (Fig. 2A). This smaller data base compared to the Amidase_2 family may be the reason for a seemingly better sequence conservation and less additional features to the common fold of the AmiC-type amidases (Fig. 2C). The overall fold is again determined by a central six-stranded and mostly parallel β-sheet (Fig. 3B). The rear site interacts with the three stretched α-helices one, three, and five, forming the hydrophobic core. The front of the β-sheet is again solvent accessible and forms the bottom of the binding cleft, however, in contrast to the Amidase_2 family the binding groove is bordered by loops, not helices. Furthermore, the putative active site with the coordinated zinc ion and binding region for peptide stem as well as carbohydrate backbone is much wider than for Amidase_2 enzymes, such as AmiA or AmiD.

The zinc-binding residues histidine 196, glutamate 211, and histidine 265 (numbering according to AmiC (Rocaboy et al., 2013)) are located next to conserved proline, glycine and serine residues, and these latter residues probably are required for correct positioning and orientation of the zinc ion (Fig. 2C). The carbonyl oxygen of alanine 266 likely stabilizes an intermediate state during catalysis, analogous to aspartate 266 of AmiA (Büttner et al., 2014). Asparagine or aspartate 267 are well positioned to engage contacts with MurNAc, again analogous to threonine 267 of AmiA. Glutamate 373 is the catalytically active residue, whose carboxyl group is within 5 Å of the zinc ion and well positioned to act as co-activator of the water molecule conducting the nucleophilic attack on the scissile amide bond. Threonine 388 forms hydrogen bonds with pre- and succeeding backbone amide and carbonyl atoms, enabling an unusual kink of helix α5 toward the rear of the central β-sheet that occurs in the absence of known helix breakers. The threonine is not strictly conserved, but other hydrophilic residues such as serine, asparagine, or aspartate can act similarly. Again, the remaining well conserved and largely hydrophobic residues play a structural role as they anchor secondary structures and stabilize the fold.

Rv3717 of *Mycobacterium tuberculosis* possesses no cell wall binding domain but an insertion of twenty amino acids that forms a short $3_{10}$-helix and a disulfide-bonded β-hairpin (Fig. 3B) (Kumar et al., 2013; Prigozhin et al., 2013). This insertion is close to the active site and may play a role in autoregulation, as Rv3717 has no potential regulating cell-wall binding domain (Kumar et al., 2013). The disulfide bond is not needed for activity, however, zinc and glutamate 200 (glutamate 373 in AmiC) are required for cleavage. Limited proteolysis of peptidoglycan and complex structure

analysis suggest that the insertion may regulate specificity and that Rv3717 is involved in PGN-fragment recycling (Prigozhin et al., 2013). CwlV from *Paenibacillus polymyxa* (Ishikawa et al., 1999) and the putative amidase from *Neisseria meningitides* (Tettelin et al., 2000) are not yet well investigated, but the published structures show the Amidase_3 core fold without additional features, suggesting unrestricted amidase activity. AmiC of *E. coli* (Rocaboy et al., 2013) and the orthologous AmiB from *Bartonella henselae* (Yang et al., 2012) have an insertion of over forty amino acids after β-sheet β4. A proposed self-regulatory α-helix lies within the binding groove as catalytic-domain-internal repressor (Fig. 3B), contacting the catalytic zinc ion by a glutamate or glutamine and blocking the active site via predominantly hydrophobic interactions. A flexible loop, not visible in the crystal structures, connects the insert to helix α3, which is significantly longer in AmiB and AmiC. LytM activators associated with septal ring factors allow a conformational shift of the inhibiting α-helix out of the binding site and therefore trigger amidase activity (Rocaboy et al., 2013; Yang et al., 2012).

*Structures of complexes – a challenging task to obtain liganded enzyme structures*

Obtaining an enzyme-substrate complex is usually a challenging task, and even more so for amidases. As PGN is a variable and heavily modifiable biopolymer and the affinity of amidases for peptidoglycan is comparatively low due to its catalytic environment with high local substrate concentration, the particular substrates usually are not commercially available and need to be produced by customized synthesis in sufficient quantities. To avoid merely fragment-bound structures, leading to incomplete data on enzyme-substrate interactions, it is further necessary to inactivate the amidase by removal of zinc or mutation, or synthesis of a non-cleavable substrate analog.

As an example, Fig. 3C shows the enzyme-substrate complex of AmiA (4KNL, (Büttner et al., 2014)). A dense network of direct and water-mediated interactions orients the peptide stem of PGN within the narrow, elongated AmiA binding groove. As described above, conserved residues coordinate the zinc ion and key areas of the ligand. Comparison with the complex structures of AmiD (3D2Y and 3D2Z, substrates (Kerff et al., 2010)), AmpD (2Y2B, reaction product (Carrasco-Lopez et al., 2011)), AmpDh2 (4BOL, 4BPA, pentapeptide and tetrasaccharide (Martinez-Caballero et al., 2013)), and AmpDh3 (4BXD, 4BXE, pentapeptide and pentapeptide-tetrasaccharide (Lee et al., 2013)) reveal that minor substitutions are able to modulate substrate specificity. Histidine 370 and threonine 380 for example stabilize the PGN-isoglutamine, and histidine 382 the tetrahedral reaction intermediate in AmiA, whereas a lysine and arginines stabilize the PGN-isoglutamate and the tetrahedral intermediate in AmpD. The positive charge of arginine is better suited to stabilize the negative charge of glutamate, thus adapting specificity. The incorporation of arginines at the end of the binding groove enables binding of a substrate that contains free carboxylate(s) at terminal D-Ala and/or *meso*-DAP, as seen for AmpD, AmpDh2 and AmpDh3. This effect is not seen in AmiD due to crystal contacts preventing the stabilization. The deletion of few amino acids in AmiD after the first zinc-binding histidine shortens a loop and reduces the size of the carbohydrate binding pocket, leading to a preference for anhydro-MurNAc.

The complex structures of AmpDh2 and AmpDh3 with pentapeptide and tetrasaccharide ligands in combination with a cross-linked PGN, based on a solution NMR-structure of a larger PGN fragment, provide the basis for a plausible model of how these enzymes engage PGN. The only complex structure of an Amidase_3 is of Rv3717 (4M6G, reaction product (Prigozhin et al., 2013)). The position of the bound reaction products in a pocket, created by the

additional β-hairpin motif, supports the assumption of a recycling function of Rv3717.

*Conclusions and outlook*

As we have shown here, structural analyses of a significant number of amidases have been performed. Some of these enzymes have little sequence homology and were not expected to share the same fold prior to the determination of the structures. In this context, X-ray crystallography has helped to uncover close structural relationships among these enzymes. The amidase fold is even conserved in human proteins such as the PGRPs that serve a very different function. It is clear that, taken together, the different structural analyses are able to provide general insights into the function of these enzymes that each structure, standing alone, does not. Moreover, in most cases there are also interesting differences in amidase structures, and these offer clues about the binding modes of the substrates, which also differ among species. Finally, some amidases have been modified with loops or extra domains to gain specific additional functions in cell wall anchoring and regulation. It is important to keep in mind that the amidases discussed here are embedded into larger protein chains, and we still lack an understanding of how the amidases interact with neighboring domains. Deciphering the interplay among these domains, and what this means for regulation and targeting, remains one of the major challenges for this field of research. Even so, the knowledge available today on the ligand binding properties and reaction mechanisms of so many bacterial amidases is poised to lead for the development of a new class of antibacterial drugs as it is clear that amidases, such as the AmiA from *S. aureus*, are required for proper cell division of bacteria.

## References

Atrih, A., Bacher, G., Allmaier, G., Williamson, M.P., Foster, S.J., 1999. Analysis of peptidoglycan structure from vegetative cells of *Bacillus subtilis* 168 and role of PBP 5 in peptidoglycan maturation. J. Bacteriol. 181, 3956–3966.

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S.R., Griffiths-Jones, S., Howe, K.L., Marshall, M., Sonnhammer, E.L., 2002. The Pfam protein families database. Nucleic Acids Res. 30, 276–280.

Boneca, I.G., Huang, Z.H., Gage, D.A., Tomasz, A., 2000. Characterization of *Staphylococcus aureus* cell wall glycan strands, evidence for a new beta-N-acetylglucosaminidase activity. J. Biol. Chem. 275, 9910–9918.

Büttner, F.M., Zoll, S., Nega, M., Gotz, F., Stehle, T., 2014. Structure-function analysis of *Staphylococcus aureus* amidase reveals the determinants of peptidoglycan recognition and cleavage. J. Biol. Chem. 289, 11083–11094.

Carrasco-Lopez, C., Rojas-Altuve, A., Zhang, W., Hesek, D., Lee, M., Barbe, S., Andre, I., Ferrer, P., Silva-Martin, N., Castro, G.R., Martinez-Ripoll, M., Mobashery, S., Hermoso, J.A., 2011. Crystal structures of bacterial peptidoglycan amidase AmpD and an unprecedented activation mechanism. J. Biol. Chem. 286, 31714–31722.

Chou, S., Bui, N.K., Russell, A.B., Lexa, K.W., Gardiner, T.E., LeRoux, M., Vollmer, W., Mougous, J.D., 2012. Structure of a peptidoglycan amidase effector targeted to Gram-negative bacteria by the type VI secretion system. Cell Rep. 1, 656–664.

de Jonge, B.L., Chang, Y.S., Gage, D., Tomasz, A., 1992. Peptidoglycan composition of a highly methicillin-resistant *Staphylococcus aureus* strain. The role of penicillin binding protein 2A. J. Biol. Chem. 267, 11248–11254.

Garman, E.F., 2014. Developments in X-ray crystallographic structure determination of biological macromolecules. Science 343, 1102–1108.

Heilmann, C., Hussain, M., Peters, G., Gotz, F., 1997. Evidence for autolysin-mediated primary attachment of *Staphylococcus epidermidis* to a polystyrene surface. Mol. Microbiol. 24, 1013–1024.

Ishikawa, S., Kawahara, S., Sekiguchi, J., 1999. Cloning and expression of two autolysin genes, cwlU and cwlV, which are tandemly arranged on the chromosome of Bacillus polymyxa var. colistinus. Mol. Gen. Genet. 262, 738–748.

iycr2014, 2014. International Year of Crystallography 2014.

Kerff, F., Petrella, S., Mercier, F., Sauvage, E., Herman, R., Pennartz, A., Zervosen, A., Luxen, A., Frere, J.M., Joris, B., Charlier, P., 2010. Specific structural features of the N-acetylmuramoyl-L-alanine amidase AmiD from *Escherichia coli* and mechanistic implications for enzymes of this family. J. Mol. Biol. 397, 249–259.

Kumar, A., Kumar, S., Kumar, D., Mishra, A., Dewangan, R.P., Shrivastava, P., Ramachandran, S., Taneja, B., 2013. The structure of Rv3717 reveals a novel amidase from *Mycobacterium tuberculosis*. Acta Crystallogr. D: Biol. Crystallogr. 69, 2543–2554.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.

Lee, M., Artola-Recolons, C., Carrasco-Lopez, C., Martinez-Caballero, S., Hesek, D., Spink, E., Lastochkin, E., Zhang, W., Hellman, L.M., Boggess, B., Hermoso, J.A., Mobashery, S., 2013. Cell-wall remodeling by the zinc-protease AmpDh3 from *Pseudomonas aeruginosa*. J. Am. Chem. Soc. 135, 12604–12607.

Lehner, J., Zhang, Y., Berendt, S., Rasse, T.M., Forchhammer, K., Maldener, I., 2011. The morphogene AmiC2 is pivotal for multicellular development in the cyanobacterium *Nostoc punctiforme*. Mol. Microbiol. 79, 1655–1669.

Low, L.Y., Yang, C., Perego, M., Osterman, A., Liddington, R., 2011. Role of net charge on catalytic domain and influence of cell wall binding domain on bactericidal activity, specificity, and host range of phage lysins. J. Biol. Chem. 286, 34391–34403.

Low, L.Y., Yang, C., Perego, M., Osterman, A., Liddington, R.C., 2005. Structure and lytic activity of a *Bacillus anthracis* prophage endolysin. J. Biol. Chem. 280, 35433–35439.

Martinez-Caballero, S., Lee, M., Artola-Recolons, C., Carrasco-Lopez, C., Hesek, D., Spink, E., Lastochkin, E., Zhang, W., Hellman, L.M., Boggess, B., Mobashery, S., Hermoso, J.A., 2013. Reaction products and the X-ray structure of AmpDh2, a virulence determinant of *Pseudomonas aeruginosa*. J. Am. Chem. Soc. 135, 10318–10321.

Mellroth, P., Sandalova, T., Kikhney, A., Vilaplana, F., Hesek, D., Lee, M., Mobashery, S., Normark, S., Svergun, D., Henriques-Normark, B., Achour, A., 2014. Structural and functional insights into peptidoglycan access for the lytic amidase LytA of *Streptococcus pneumoniae*. Mbio 5.

NATURE s.i., 2014. Crystallography turns 100. Nature 505, 585–716.

Paget, M.S., Chamberlin, L., Atrih, A., Foster, S.J., Buttner, M.J., 1999. Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). J. Bacteriol. 181, 204–211.

Prigozhin, D.M., Mavrici, D., Huizar, J.P., Vansell, H.J., Alber, T., 2013. Structural and biochemical analyses of *Mycobacterium tuberculosis* N-acetylmuramyl-L-alanine amidase Rv3717 point to a role in peptidoglycan fragment recycling. J. Biol. Chem. 288, 31549–31555.

Ramachandran, R.a.S., 1963. Stereochemistry of polypeptide chain configurations. J. Mol. Biol. 7, 95–99.

Rocaboy, M., Herman, R., Sauvage, E., Remaut, H., Moonens, K., Terrak, M., Charlier, P., Kerff, F., 2013. The crystal structure of the cell division amidase AmiC reveals the fold of the AMIN domain, a new peptidoglycan binding domain. Mol. Microbiol. 90, 267–277.

Rupp, B., 2010. Biomolecular Crystallography. Garland Science Taylor & Frncis Group, LLC.

Schleifer, K.H., Kandler, O., 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol. Rev. 36, 407–477.

SCIENCE s.i., 2014. Crystallography at 100. Science 343, 1049–1168.

Sumner, J.B., 1937. The story of urease. J. Chem. Educ. 14, 225.

Sumner, T., 2014. Dazzling-history. Science, 343.

Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Cittone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H., Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Masignani, V., Pizza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappuoli, R., Venter, J.C., 2000. Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58. Science 287, 1809–1815.

Wang, Z.M., Li, X., Cocklin, R.R., Wang, M., Fukase, K., Inamura, S., Kusumoto, S., Gupta, D., Dziarski, R., 2003. Human peptidoglycan recognition protein-L is an N-acetylmuramoyl-L-alanine amidase. J. Biol. Chem. 278, 49044–49052.

Warth, A.D., Strominger, J.L., 1971. Structure of the peptidoglycan from vegetative cell walls of *Bacillus subtilis*. Biochemistry 10, 4349–4358.

www.pdb.org, PDB, RCSB.

Yang, D.C., Tan, K., Joachimiak, A., Bernhardt, T.G., 2012. A conformational switch controls cell wall-remodelling enzymes required for bacterial cell division. Mol. Microbiol. 85, 768–781.

Zoll, S., Patzold, B., Schlag, M., Götz, F., Kalbacher, H., Stehle, T., 2010. Structural basis of cell wall cleavage by a *Staphylococcal autolysin*. PLoS Pathog. 6, e1000807.

Zoll, S., Schlag, M., Shkumatov, A.V., Rautenberg, M., Svergun, D.I., Gotz, F., Stehle, T., 2012. Ligand-binding properties and conformational dynamics of autolysin repeat domains in staphylococcal cell wall recognition. J. Bacteriol. 194, 3789–3802.

# Targeting the Gatekeeper MET146 of C-Jun N-Terminal Kinase 3 Induces a Bivalent Halogen/Chalcogen Bond

Andreas Lange,[†,‡,§] Marcel Günther,[∥,§] Felix Michael Büttner,[#] Markus O. Zimmermann,[†,‡] Johannes Heidrich,[†,‡] Susanne Hennig,[†] Stefan Zahn,[¶] Christoph Schall,[#] Adrian Sievers-Engler,[⊥] Francesco Ansideri,[∥] Pierre Koch,[∥] Michael Laemmerhofer,[⊥] Thilo Stehle,[#,∇] Stefan A. Laufer,[∥] and Frank M. Boeckler*,[†,‡]

[†]Molecular Design and Pharmaceutical Biophysics, Institute of Pharmaceutical Sciences, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

[‡]Center for Bioinformatics Tübingen (ZBIT), Eberhard Karls Universität Tübingen, Sand 1, 72076 Tübingen, Germany

[∥]Pharmaceutical and Medicinal Chemistry, Institute of Pharmaceutical Sciences, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

[#]Interfaculty Institute of Biochemistry, Eberhard Karls Universität Tübingen, Hoppe-Seyler-Str. 4, 72076 Tübingen, Germany

[¶]Institute of Physical Chemistry, Justus-Liebig-Universität Gießen, Heinrich-Buff-Ring 17, 35392 Gießen, Germany

[⊥]Pharmaceutical (Bio)Analysis, Institute of Pharmaceutical Sciences, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

[∇]Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, United States

**S** *Supporting Information*

**ABSTRACT:** We target the gatekeeper MET146 of c-Jun N-terminal kinase 3 (JNK3) to exemplify the applicability of X···S halogen bonds in molecular design using computational, synthetic, structural and biophysical techniques. In a designed series of aminopyrimidine-based inhibitors, we unexpectedly encounter a plateau of affinity. Compared to their QM-calculated interaction energies, particularly bromine and iodine fail to reach the full potential according to the size of their $\sigma$-hole. Instead, mutation of the gatekeeper residue into leucine, alanine, or threonine reveals that the heavier halides can significantly influence selectivity in the human kinome. Thus, we demonstrate that, although the choice of halogen may not always increase affinity, it can still be relevant for inducing selectivity. Determining the crystal structure of the iodine derivative in complex with JNK3 (4X21) reveals an unusual bivalent halogen/chalcogen bond donated by the ligand and the back-pocket residue MET115. Incipient repulsion from the too short halogen bond increases the flexibility of $C_\varepsilon$ of MET146, whereas the rest of the residue fails to adapt being fixed by the chalcogen bond. This effect can be useful to induce selectivity, as the necessary combination of methionine residues only occurs in 9.3% of human kinases, while methionine is the predominant gatekeeper (39%).
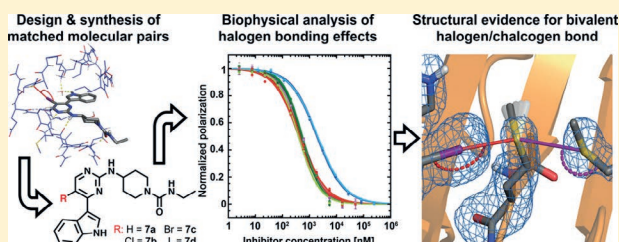
## INTRODUCTION

In the past decade halogen bonding has increasingly gained attention in life sciences and drug discovery.[1−8] This attractive interaction can typically be described as a R−X···D−R′ contact, where X represents chlorine, bromine, and iodine (as electron pair acceptor) and D can be any kind of Lewis base.[9] For a broader recognition of halogen bonding in molecular design, we have recently studied halogen bonding contacts with different interaction partners in protein binding sites.[4−7] From these studies it is evident that halogen bonds can provide innovative ideas, how to access interesting parts of the biological space (e.g., therapeutic targets) from hitherto widely neglected parts of the chemical space (chemotypes featuring halogen bonds as essential key recognition motifs in the binding mode).

On the basis of the lack of directional interactions toward methionine (MET) in classical interactions, this amino acid has been particularly difficult to target, except for using hydrophobic contacts. Only recently, Meanwell and colleagues have highlighted the role of noncovalent sulfur interactions in drug design.[10] Recently, we have studied the targetability of MET by halogen bonding systematically on a quantum chemical level of theory.[4] One of our goals was to provide new insights for utilizing halogen···sulfur contacts in molecular design. Here, we

14640

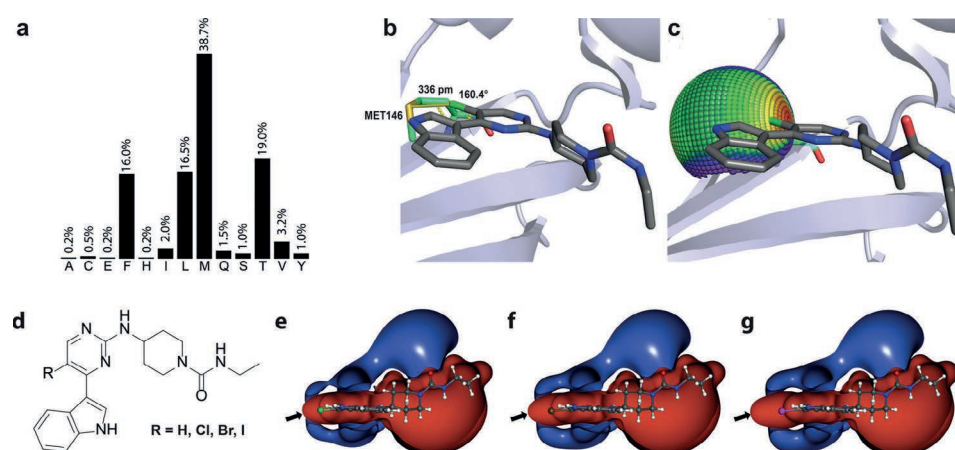**Figure 1.** (a) Percentage distribution of gate keeper residues in human kinases based on kinase sequence alignments[16] obtained from http://kinase. com/human/kinome/phylogeny.html. (b) 4-{[5-Chloro-4-(1H-indol-3-yl)pyrimidin-2-yl]amino}-N-ethylpiperidine-1-carboxamide interacting with the gatekeeper MET146 of JNK3 (PDB: 2P33).[15] The chlorine···sulfur distance ($d_{X···S}$ = 336 pm) and the $\sigma$-hole angle ($\alpha_{C-X···S}$ = 160.4°) are depicted explicitly. (c) Data from a spherical interaction scan on the MP2/TZVPP level of theory using iodobenzene as a model ligand and dimethylsulfide as a model for methionine[4,17] were plotted onto MET146, indicating that the substitution of chlorine by iodine should be favorable (color gradient from red: very favorable to purple: unfavorable). Spherical interaction scans were prepared previously[4] by systematically rotating the ligand system around the methionine sulfur as the center of the sphere, while maintaining the orientation of the model ligand, i.e., the same $\sigma$-hole angle $\alpha_{C-X···S}$. The exact spherical orientation of any ligand interacting through a halogen bond with methionine can be described by the spherical coordinates of the iodine atom (distance $d_{S···X}$, azimuth and elevation angle with respect to a certain plane of symmetry of dimethylsulfide) and the $\sigma$-hole angle $\alpha_{C-X···S}$. The figures in panel (b) and (c) were prepared using PyMOL.[18] (d) Structures of the designed matched molecular pair series comprising the hydrogen (**7a**), chlorine (**7b**), bromine (**7c**), and iodine derivative (**7d**). (e–g) ESP isosurfaces (MP2/TZVPP) of compounds **7b–d**. Negative ESP isosurfaces contoured at an energy of −0.009 au are colored in dark blue, while positive ESP isosurfaces contoured at 0.009 au are colored in red. Toward the heavier halides, the $\sigma$-hole (positive potential) on the halogen in elongation of the R−X bond (indicated by a black arrow) increases. All pictures were prepared with MOLCAD.[19,20]

target the gatekeeper MET146 of c-Jun N-terminal kinase 3 (JNK3) to exemplify the applicability of X···S halogen bonding in molecular design using computational, structural and biophysical techniques. Figure 1a shows the percentage distribution of gatekeeper residues in all human kinases. The term "gatekeeper" is commonly used to describe the residue in a kinase that blocks the hydrophobic pocket and, therefore, is a key feature of the recognition elements in the ATP-binding site of the kinase.[11] From Figure 1a, it is obvious that methionine (with ~39%) is by far the most common gatekeeper in human kinases. It is followed by threonine (with ~19%), leucine (~17%), and phenylalanine (~16%). All other amino acids together are only found in a small fraction (~10%) as gatekeeper residues. JNK3 is a member of the mitogen-activated protein kinases (MAPK) family and a popular target for drug discovery.[12,13] Because of the lack of specific ways to target the gatekeeper methionine residue (MET146) of JNK3, traditional drug discovery approaches have focused on utilizing the side chain flexibility of this residue to shift it into a position granting access to the hydrophobic pocket.[14] However, an exhaustive PDB survey revealed that MET146 can be targeted by halogen bonding (Figure 1b,c). Actually, Alam et al. had reported a crystal structure of the complex between 4-{[5-chloro-4-(1H-indol-3-yl)pyrimidin-2-yl]amino}-N-ethylpiperi-dine-1-carboxamide (**7b**) and JNK3 (PDB-ID: 2P33, further referred to as 2P33), in which a chlorine is interacting with MET146.[15]

In this study, we generated the halogen/hydrogen analogues to complete the series of matched molecular pairs (Figure 1d). We used quantum chemical methods to assess the difference in $V_{max}$, the maximal positive electrostatic potential representing the size and strength of the $\sigma$-hole. We plotted the electrostatic

potential isosurfaces of the halogenated analogues (Figure 1e–g) and optimized all four ligands in the binding pocket using TPSS-D2/SV(P). We synthesized all four ligands, establishing a new route to obtain the bromine and iodine derivatives. We expressed and purified JNK3 to characterize the differences in ligand affinities using biophysical techniques such as differential scanning fluorimetry (DSF), fluorescence polarization (FP), and isothermal titration calorimetry (ITC). Since JNK3 and p38$\alpha$ MAP kinase (p38$\alpha$) are notoriously similar in structure, we also assessed the ligand selectivity against p38$\alpha$. To elucidate the contribution of the halogen bond, we prepared the gatekeeper mutants JNK3-M146A, JNK3-M146L, and JNK3-M146T and investigated mutant-induced changes in the binding affinity compared to JNK3 wildtype (JNK3-wt) by FP competition experiment. In order to rationalize our findings, we have determined the crystal structure of JNK3-wt in complex with the iodine derivative (**7d**).

## ■ RESULTS

**1. Halogen Bond Tuning ($V_{max}$) and Rationalization of the Designed MMPs by Quantum Chemistry.** The chlorine···sulfur distance $d_{X···S}$ = 336 pm and the $\sigma$-hole angle $\alpha_{C-X···S}$ = 160.4° suggest that the halogen bond to the gatekeeper MET146 in 2P33 is favorable (Figure 1b). Derived from previous[4] and current theoretical work,[17] we also rationalized that the spherical geometry is beneficial (Figure 1c). The spherical geometry can be represented by the spherical coordinates (distance, azimuth, elevation) of iodine with respect to the methionine sulfur as the center of the sphere. Therefore, we generated in this study the matched molecular pairs complementing the crystallized ligand (**7b**) with the
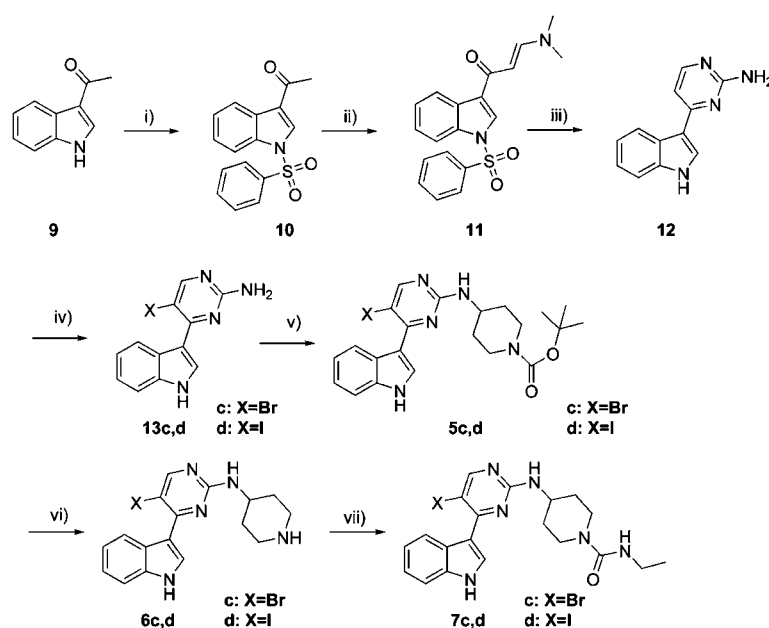
**Figure 2.** Synthesis of compounds **7c** and **7d**. Reagents and conditions: (i) MeOH, KOH, acetone, benzenesulfonyl chloride, rt, 92%; (ii) dimethylformamide-dimethylacetale, reflux, 24 h, 89%; (iii) MeOH, guanidine carbonate, sodium methoxide, reflux, 48 h, 61%; (iv) DMF, N-halosuccinimide, rt, overnight, 59–65%; (v) DCM, N-Boc-piperidin-4-on, sodium triacetoxy borohydride, TFA, rt, 70–80%; (vi) DCM, TFA, rt, 2 h; (vii) DCM, triethylamine, ethyl isocyanate, overnight, 40–59% (yield for step (vi) + (vii)).

analogues containing bromine (**7c**), iodine (**7d**) or hydrogen (**7a**) in position five of the pyrimidine ring (Figure 1d).

The strength of a halogen bond is tuned by the scaffold.[21−23] To elucidate this tuning effect, we calculated the electrostatic potential isosurfaces and the $V_{max}$ value (most positive electrostatic potential on the $\sigma$-hole when mapping the ESP onto the isodensity surface) for each halogenated compound (MP2/SV(P), Figure 1e−g). **7b** appears to be hardly tuned, showing a $V_{max}$ value of 0.123 au (chlorobenzene: 0.117 au, $\Delta V_{max}$ = 0.006 au). **7c** and **7d**, however, possess increasingly tuned $\sigma$-holes, as expected. **7c** has a $V_{max}$ value of 0.168 au (bromobenzene: 0.150 au, $\Delta V_{max}$ = 0.018 au) and **7d** has a value of 0.219 au (iodobenzene: 0.182 au, $\Delta V_{max}$ = 0.037 au). With the $V_{max}$ value being almost twice as large for iodine as for chlorine, a significantly stronger halogen bond should result from the proposed halogen exchange.

We further investigated, whether these tuning differences translate into changes of the binding mode. Using calculations on a TPSS-D2/SV(P) level of theory and a carefully selected binding site model to rationalize the designed MMPs. A detailed report about the protocol and results is provided in the Supporting Information. Essentially, the hydrogen bond network to the hinge binding region appeared to be intact in all compounds. The halogen bond geometry of chlorine is preserved perfectly. Despite their size, the iodine and bromine atoms in **7c** and **7d** are located 10 and 14 pm closer to the sulfur atom of MET146, respectively. This can reflect that the stronger $\sigma$-hole tuning permits a closer contact to the sulfur, but could also indicate an onset of an increasingly repulsive contact.

**2. Synthesis.** Compounds **7a** and **7b** were synthesized according to a route (Scheme S1) published by Alam et al.[15] with some modifications. Further details can be found in the Supporting Information. On the basis of a modified Bredereck synthesis,[24] we utilized a linear route including a pyrimidine

cyclization for the synthesis of compounds **7c** and **7d** as depicted in Figure 2.

To prevent N-methylation during step ii, we first protected the nitrogen of commercially available 3-acetylindole as a benzenesulfonamide using already described conditions.[25] Further treatment of **10** with dimethylformamid-dimethylacetale under neat and refluxing conditions afforded **11** in high yields. The following cyclization conditions utilizing guanidine carbonate and sodium methoxide in refluxing MeOH resulted in both, the formation of the 2-aminopyrimidyl product **12** including the cleavage of the benzenesulfonamide group in the same step in acceptable yields. According to Rossignol et al.[26] 2-amino-4-indolylpyrimidines can be regioselectively halogenated using N-halosuccinimides in polar solvents. It is noteworthy that we isolated **5c** and **5d** in moderate yields without the occurrence of regioisomeric halogenated side products under the depicted conditions in contrast to our attempts in halogenating **3a** or **4a** (Scheme S1, Supporting Information), suggesting a crucial role of the primary 2-amino group in selective electrophilic halogenation. The 2-aminopyrimidines **13c** and **13d** can be smoothly converted into the secondary amines **14c** and **14d** via reductive amination in DCM using TFA as acid and sodium triacetoxy borohydrid as reducing agent at room temperature without cleavage of the Boc protective group or reductive dehalogenation.[27] Again, we used 20% TFA in DCM to quantitatively remove the Boc group from the 4-aminopiperidino side chain and successfully introduced the urea structure via electrophilic addition to ethyl isocyanate under basic conditions in DCM at room temperature to yield **7c** and **7d** in moderate overall yields.

**3. Biophysical Characterization.** *3.1. Fluorescence Polarization (FP) Assay.* We used a fluorescence polarization assay as the primary biophysical method to determine $K_I$ values for **7a−d**. When suitable fluorescent reporter molecules are available, the FP assay is an elegant, rapid, and dependable
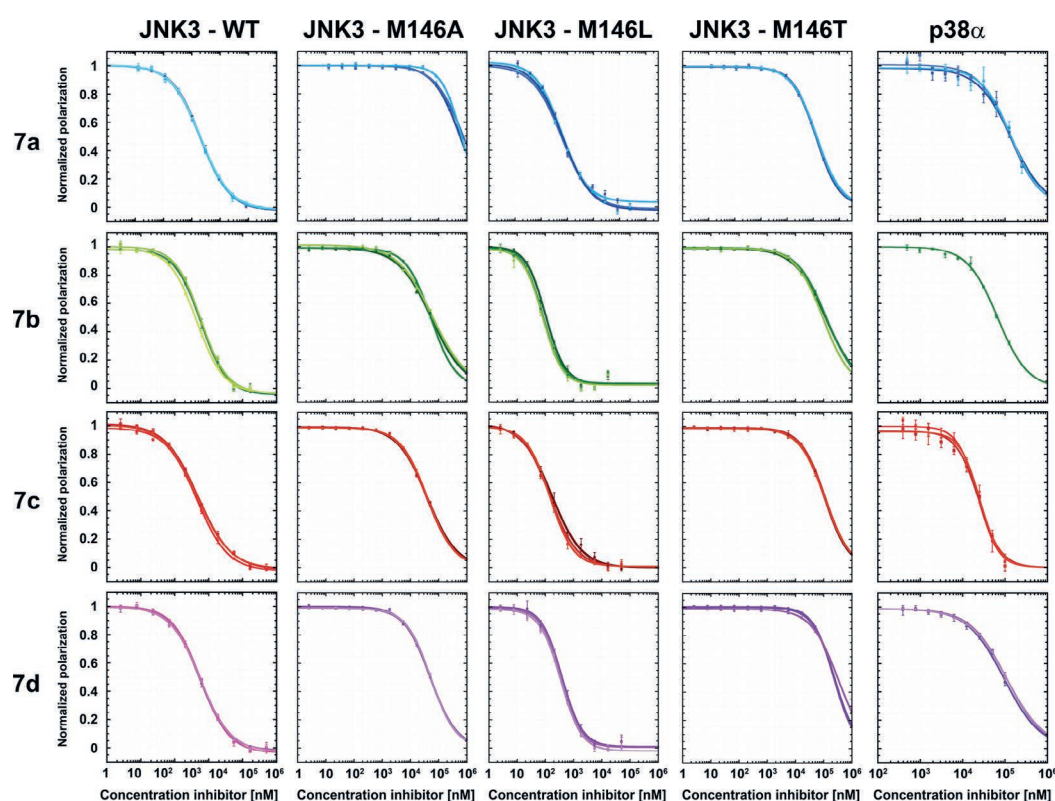
**Figure 3.** Fluorescence polarization-based competition binding experiments using compounds **7a**–**7d** and JNK3-wt, the JNK3 mutants M146A, M146L, and M146T, or p38$\alpha$. The polarization signal (in mP units) was read using a BMG CLARIOstar reader. The values were normalized to a range from 100% to 0% polarization. Eleven concentrations of the respective inhibitor were measured in quadruplicate and each competition binding experiment was repeated three times, except for **7b** with p38$\alpha$, where limited compound availability allowed only one experiment. Because of the onset of ligand precipitation, concentrations >100 $\mu$M were neglected and removed from the plot. The four parameter logistic nonlinear regression model was used for all curve fits. For JNK3-M146A and JNK3-M146T the asymptote for full displacement was fixed to 0% normalized polarization. Plots were prepared using KaleidaGraph.[33]

**Table 1. Results of the Normalized FP-Assay for JNK3 and p38$\alpha$**

| | JNK3 | | | | p38$\alpha$ | JNK3-wt preference | | | |
| | wt | M146A | M146L | M146T | | | | | |
| compound | $K_D$ [nM] | $K_D$ [nM] | $K_D$ [nM] | $K_D$ [nM] | $K_D$ [nM] | $K_D$ (M146A)/ $K_D$ (wt) | $K_D$ (M146L)/ $K_D$ (wt) | $K_D$ (M146T)/ $K_D$ (wt) | $K_D$ (p38$\alpha$)/ $K_D$ (JNK3-wt) |
|---|---|---|---|---|---|---|---|---|---|
| PIT0105016[a] | 4.0 ± 0.2 | 2.3 ± 0.2 | 31 ± 3.5 | 3.0 ± 0.2 | n.d. | 0.58 | 7.8 | 0.75 | – |
| RN4646[b] | n.d. | n.d. | n.d. | n.d. | 7.3 ± 1.3 | – | – | – | – |
| **7a**[c] | 270 ± 7.9 | 56000 ± 6000[e] | 270 ± 15 | 14000 ± 470[e] | 27000 ± 1100 | 210 | 1.0 | 52 | 100 |
| **7b**[c] | 68 ± 9.1 | 4800 ± 230[e] | 31 ± 4.6 | 12000 ± 1400[e] | 14000[d] | 71 | 0.46 | 180 | 210 |
| **7c**[c] | 58 ± 2.3 | 3300 ± 100[e] | 61 ± 5.5 | 12000 ± 280[e] | 4800 ± 220 | 57 | 1.1 | 210 | 83 |
| **7d**[c] | 72 ± 2.5 | 4500 ± 96[e] | 160 ± 5.6 | 28000 ± 4400[e] | 21000 ± 1300 | 63 | 2.2 | 390 | 290 |

[a]PIT0105016 is a novel fluorescein-labeled, pyridinylimidazole-based high-affinity probe for JNK3.[31] [b]RN4646 is a fluorescein-labeled analogue of SB203580.[30] $K_D$ values of both reporter ligands were determined by three independent direct titrations with 12 concentrations of protein measured in quadruplicate. [c]The $K_D$ values have been determined by three individual competition binding assays with 11 compound concentrations measured in quadruplicate. $K_D$ values were derived using the $K_I$ calculator of Shaomeng Wang and co-workers (http://sw16.im.med.umich.edu/software/calc_ki/).[32] [d]Because of the limited availability of **7b**, p38$\alpha$ affinity was only determined as a single measurement of 11 concentrations of **7b** in quadruplicate. [e]Because of low affinities and limited compound solubility of **7a**–**d**, no complete displacement of the reporter ligand could be observed. The normalization was, therefore, done by using the average FP values of the reporter ligand in buffer as the expected minimum of the curve. From the other experiments we can deduce that the given $K_D$ values are the lower limit of the possible values obtainable from curve fitting.

method to assess a broad range of affinities (over several orders of magnitude).[28,29] On the basis of a previously reported fluorescein-labeled analogue of SB203580[30] (RN4646) as a high-affinity fluorescent probe for p38$\alpha$ ($K_D$ of 7.3 nM) and a novel fluorescein-labeled, pyridinylimidazole-based high-affinity

probe for JNK3 (PIT0105016, $K_D$ of 4.0 nM),[31] we were able to determine both, affinities and selectivities, with good precision. Figure 3 shows the curve fits of three individual competition binding experiments per compound for **7a**–**d**. Every individual curve is based on 11 concentrations, each

measured in quadruplicate. Mean values and standard deviations for the $K_D$ values were obtained by three individual direct titrations of the reporter compounds (PIT0105016 and RN4646) and the $K_I$ values of **7a–d** were obtained by three individual competition experiments (Table 1).

For JNK3 (wt = wildtype), well-defined, reproducible sigmoidal curves were obtained, showing only small standard deviations for each concentration measured in quadruplicate. **7c** exhibits the highest affinity with a $K_I$ of 58 nM, followed by **7b** with 68 nM, and **7d** with 72 nM. For the unsubstituted ligand **7a**, a $K_I$ value of only 270 nM is obtained. According to ANOVA applying the Bonferroni post hoc test, the difference between **7a** and all other ligands is highly significant ($p < 0.0001$); however, the results for **7b–d** are not significantly different. Consistent with our QM-based analysis of the proposed interactions and previous DSF experiments (see Supporting Information: Figure S3 and Table S3), there is little difference between **7c** and **7d**. In addition, all halogenated ligands exhibit better binding to JNK3 than **7a**. However, it is surprising that **7b** seems to have the same affinity for JNK3 as **7c,d**.

We tested **7a–d** on p38α, as well, to evaluate similarities and differences in JNK3 selectivity. JNK3 and p38α are both MAPK kinases, showing a sequence identity of 51% and substantial three-dimensional similarity, particularly in the binding site.[14] Still, some crucial residues in the binding site are well-known for inducing selectivity. For example, the gatekeeper in JNK3 is a methionine (MET146), whereas in the p38α the gatekeeper is a threonine (THR106). Another important difference between these two proteins is located in the hinge region, where ASP150 in JNK3 corresponds to GLY110 in p38α, which facilitates a flipped conformation in the hinge that can be used to obtain highly selective ligands for p38α.[34] According to DSF experiments with p38α, we expected to find only weak binding (see Supporting Information). The results are given in Figure 3 and summarized in Table 1. On the basis of the decrease of affinity for p38α, we had to increase the ligand concentration up to 200 μM. We encountered that limited solubility of all compounds under the given buffer conditions at concentration higher than 100 μM leads to an onset of precipitation. Consequently, these concentrations were removed from the normalized plots. To still provide a reasonable curve fitting to the remaining data, we normalized the fluorescence polarization using the polarization of the labeled compound alone in buffer (∼70 to 50 mP) as the lowest possible value. **7a** is the weakest binder (27 μM), followed by **7d** (21 μM) and **7b** (14 μM). Interestingly, **7c** shows the highest affinity for p38α with a $K_I$ value of 4.8 μM. According to ANOVA applying the Bonferroni post hoc test, all measured $K_I$ values are different with high statistical significance ($p < 0.005$). In the case of **7b** only one independent experiment could be conducted based on substance limitations. Thus, assessment of statistical significance is invalid in this case.

To evaluate the contribution of the σ-hole interaction of the halogens with the gatekeeper residue M146 to the observed structure-affinity relationships, we exchanged methionine into alanine, leucine, and threonine by side-directed mutagenesis. Expression and purification of all mutants was done as for JNK3-wt. Confirmation of the correct mutation was obtained by plasmid sequencing (GATC) and determination of the molecular mass of the mutant protein by ESI-MS. JNK3-M146A was used as control for the loss of the direct interaction (σ-hole bond or other) with methionine. The affinity of the halogenated ligands **7b–d** is reduced by a factor of 57-fold to 71-fold (Table 1), while the unsubstituted ligand **7a** even suffers a 210-fold loss in affinity. Because of the limited solubility of all compounds and the increased affinity of the fluorescence reporter PIT0105016 for JNK3-M146A, the highest concentration of the compounds displaced PIT0105016 by only up to 50%. Therefore, the fluorescence polarization signal of PIT0105016 in buffer was used as the target value for full displacement of PIT0105016 and the asymptote for full displacement was fixed to 0% normalized polarization when performing the curve fit using a four parameter logistic nonlinear regression model (Figure 3). In the context of all other FP experiments, it can be deduced that the $K_D$ values obtained for JNK3-M146A are at the lower limit of the values expectable without the given intrinsic compound property limitations. The overall size of the affinity loss is very well in line with the importance of the gatekeeper interactions, whether they are based on σ-hole bonding or not.

With 16.5%, leucine is one of the most frequent gatekeeper residues found in the human kinome. It has a rather similar size compared to methionine, however, lacking electron pair donor functions, it cannot engage in σ-hole bonds. Thus, JNK3-M146L was prepared as an interaction-specific control, indicating possible changes in the selectivity profile in human kinases. As expectable, **7a** did not show any affinity difference for JNK3-M146L versus JNK3-wt. However, there is a clear differentiation for the halogenated ligands: Bromine (**7c**) shows no significant difference ($K_D = 61$ nM), while chlorine (**7b**) benefits with statistical significance (2.2-fold; $p < 0.01$; $K_D = 31$ nM) from the exchange into leucine. In contrast, iodine binds with significantly less affinity to JNK3-M146L ($K_D = 160$ nM), resulting in a 2.2-fold preference for JNK3-wt ($p < 0.001$). A plausible explanation of this observation is that several dispersive CH-X interactions onto the negative electrostatic potential of the halogen can compensate for the loss of the halogen bond. For the smaller chlorine, these can even overcompensate the effect of the halogen bond, while bromine is indifferent and the potential repulsion of the larger iodine leads to a reduced affinity. This data highlights that, although there is a plateau of affinity in this ligand series, the type of halogen (Cl, Br, I) actually has an important influence on the selectivity profile.

The JNK3-M146T mutant was prepared as link between JNK3 and p38α. In addition, threonine is the second most frequent gatekeeper residue (19%) in human kinases. Thus, this control provides important information about the extended selectivity profile. The affinities obtained for JNK3-M146T closely resemble the data obtained for p38α, as we expected. In contrast to the M146A mutant, **7a** shows a smaller decrease in affinity (52-fold, $K_D = 14$ μM), whereas **7b–d** prefers JNK3-wt by a factor of 180-fold (**7b**) to 390-fold (**7d**). This strong affinity loss for all halogenated compounds implies that no alternative halogen bond can be formed onto the threonine oxygen. We also conclude that the interaction toward the gatekeeper residue threonine is the most influential parameter for p38α/JNK3 selectivity. Similar to the M146A mutant, the fluorescence reporter PIT0105016 could not be fully displaced by increasing concentrations of **7a–d**. Nevertheless, we can infer from other experiments that the given $K_D$ values (obtained by setting the asymptote for full displacement of PIT0105016 to the fluorescence polarization of PIT0105016 in buffer) represent the lower limit of the possible $K_D$ values.
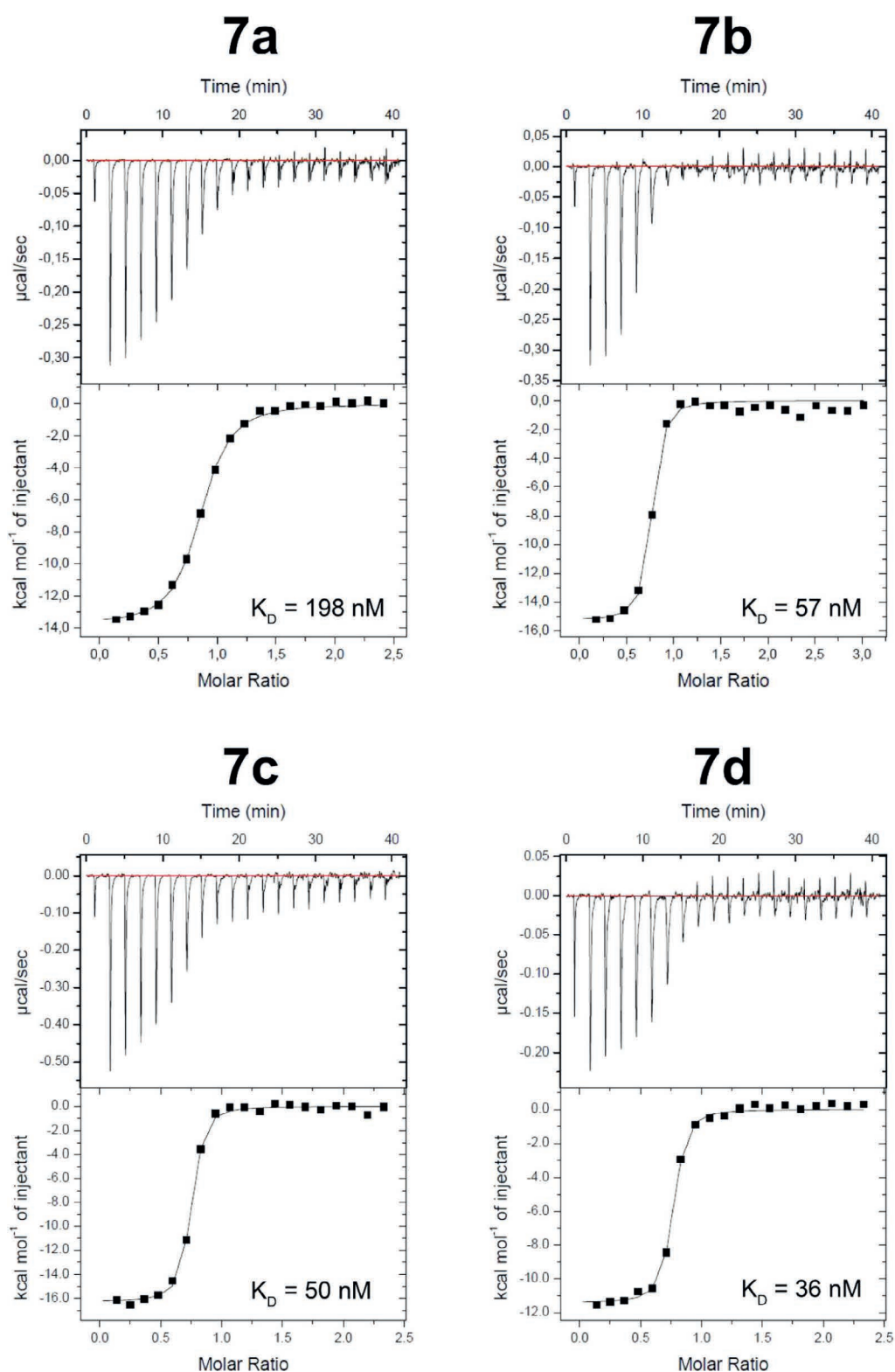
**Figure 4.** Results of the reverse ITC experiment. Protein concentration was 115 $\mu$M and compound concentration was 10 $\mu$M (for **7a**, **7c**, and **7d**) and 8 $\mu$M (for **7b**).

The FP-assay clearly confirms moderate to good JNK3 selectivity in the investigated compounds, similar to the DSF data. Using the point mutations M146A, M146L and M146T, we have shown that (a) the gatekeeper contact is important for

maintaining good kinase affinity, (b) threonine as a gatekeeper is not able to accept a halogen bond with its hydroxyl oxygen atom, (c) threonine is also not able to engage in any other alternative, attractive interaction (neither dispersive/hydro-

**Table 2. Results of the ITC Measurements of the Compounds with JNK3**

|  | $N$ | $K$ [M$^{-1}$] | $K_D$ [nM] | $c$-value | $\Delta H$ [kcal·mol$^{-1}$] | $\Delta S$ [cal·mol$^{-1}$K$^{-1}$] |
|---|---|---|---|---|---|---|
| **7a** | 1.20 | $5.06 \times 10^7$ | $198 \pm 28$ | 60 | $-10.7 \pm 0.5$ | $-5.12 \pm 2.04$ |
| **7b** | 1.39 | $1.75 \times 10^7$ | $57 \pm 5$ | 180 | $-11.3 \pm 0.5$ | $-4.95 \pm 1.50$ |
| **7c** | 1.48 | $2.00 \times 10^7$ | $50 \pm 7$ | 300 | $-10.3 \pm 1.1$ | $-1.17 \pm 3.58$ |
| **7d** | 1.36 | $2.80 \times 10^7$ | $36 \pm 7$ | 380 | $-8.3 \pm 0.7$ | $6.13 \pm 2.10$ |

phobic, nor polar) to compensate for the loss of interaction, (d) not only halogen bonds, but also dispersive/hydrophobic interactions can preserve ligand affinity, however, (e) leucine differentiates between the halogens, preferring the small chlorine featuring a less pronounced $\sigma$-hole over the large iodine with a strongly pronounced $\sigma$-hole. As a consequence, halogen bonds by different halogens can affect not only affinity, but also selectivity, making them interesting tools for kinase drug discovery.

*3.2. Isothermal Titration Calorimetry (ITC).* Using isothermal titration calorimetry, we aimed for an independent technique with a complementary biophysical readout in order to confirm our findings in the FP-assay and to resolve the binding constants into their thermodynamic parameters. We optimized the titration protocol to obtain well-defined binding isotherms with suitable c-values for **7a**−**d** in the range of 60 to 380. Under these conditions both, $K_D$ and $\Delta H$, can be derived from the binding isotherm with good reliability.[29] To achieve this, we had to apply a reverse ITC setup, based on the lack of solubility of the halogenated compounds. The sample cell was filled with a solution of **7a**−**d** in buffer containing 5% DMSO (at a concentration of 8−10 $\mu$M). A JNK3 protein solution in buffer containing 5% DMSO (at a concentration of 115 $\mu$M in the syringe) was titrated into the sample cell. Raw data and the corresponding binding isotherms of a characteristic titration experiment for each ligand are presented in Figure 4. Statistics for the relevant thermodynamic parameters from four independent titrations per compound are summarized in Table 2.

The $K_D$ values obtained by ITC are in good agreement with the affinities measured in the FP-assay. Small differences between the ITC and FP results might arise from the intrinsic characteristics of both biophysical techniques/experimental protocols (direct binding vs competition experiment, equilibration times, titration vs simultaneous measurement of all concentrations, solubility-based errors, etc.), as well as from differences in the buffer conditions (2 mM TCEP for ITC experiments vs 10 mM $\beta$-mercaptoethanol for FP experiments).

While **7a** binds clearly weaker (198 nM) than any of the halogenated compounds, there seems to be a slight trend toward better affinities for the heavier halogens (57 nM for **7b**, 50 nM for **7c**, and 36 nM for **7d**). However, by applying one-way ANOVA with Bonferroni post hoc testing, it becomes evident that there is no statistical significance in the different $K_D$ values found for **7b**−**d**. Only **7a** is different to all other compounds with high statistical significance ($p < 0.0001$). Thus, we conclude that the insignificant difference found with both methods for **7b**−**d** clearly show that for this series of matched molecular pairs, no gain of affinity can be induced by replacing chlorine by the heavier halides. Interestingly, there is also no apparent loss of affinity.

Decomposition of the free energy of binding into enthalpy ($\Delta H$) and entropy ($-T \cdot \Delta S$) contributions, however, draws attention to some important differences between **7b**, **7c**, and **7d**. The enthalpy decreases from −11.3 kJ/mol for the chlorine

derivative **7b** to −10.3 kJ/mol for the bromine derivative **7c**, and −8.3 kJ/mol for the iodine derivative **7d**. On the basis of the slightly higher standard deviation for bromine, ANOVA with Bonferroni post hoc analysis clearly shows that only the difference between **7d** and the others is significant ($p < 0.015$). Considering that iodine could in principle form a much stronger halogen bond than bromine or chlorine, this result is at least surprising. To elucidate the reason for this behavior, we solved the crystal structure of **7d** in complex with JNK3.

From Table 2, it is also evident that for the heavier halides the entropic contribution to the free energy of binding becomes more favorable, compensating almost exactly the loss of enthalpy. Thus, although $\Delta G$ is hardly affected by the type of halogen forming the halogen bond to MET146, the thermodynamic parameters are clearly deviating between chlorine, bromine, and iodine. Similar to $\Delta H$, ANOVA with Bonferroni post hoc testing showed a significant difference in $\Delta S$ only between iodine (**7d**) and the other compounds ($p < 0.01$), which again might be attributed to the high standard deviation of **7c**. A summary plot of $\Delta H$, $-T \cdot \Delta S$, and $\Delta G$ highlighting the enthalpy−entropy compensation in **7b**−**d**, can be found in the Supporting Information (Figure S4).

We likewise tried to perform ITC experiments with **7a-d** and p38$\alpha$. However, because of the limited solubility of the compounds and their weak binding to p38$\alpha$, neither direct nor indirect titration experiments yielded satisfying c-values and interpretable binding isotherms.

**4. Crystal Structure of 7d in Complex with JNK3 (4X21) and its Comparison to 2P33.** In order to find possible explanations for the significantly different thermodynamics of **7d**, we have solved the structure of **7d** soaked into a JNK3 crystal at a resolution of 1.95 Å with a final $R_{work}/R_{free}$ of 19.86/22.30 (PDB code 4X21; for further refinement statistics see Table 3). During model building, it was not possible to trace the peptide chain between ALA211 and VAL225, as well as ILE375 and ASP391 due to lack of electron density. Poorly defined electron densities for these comparatively flexible loops were also encountered in previous crystal structures of JNK3.[14,35,36]

The crystals contain two structurally very similar copies of ligand-bound JNK3 in their asymmetric unit. Because of the higher overall B-factors in chain B, we will first describe the binding situation of chain A. Although we provide comparisons and measurements on the pm scale, it should be emphasized that the possible coordinate error of the crystal structure (Table 3) certainly limits the interpretability of the very small differences sometimes observed. Therefore, we have additionally performed some accompanying calculations to elaborate on the meaning of the observed trends. Figure 5a shows that **7d** binds tightly to the hinge region and forms a halogen bond to MET146. The hydrogen bond distance to the hinge region is 205 pm for N···NH and 206 pm for NH···OC. This slightly exceeds the predicted values by 20 or 14 pm, respectively (see Table S1). In contrast, the CH···OC weak hydrogen bond involving the GLU147 backbone was found to be 241 pm long,

**Table 3. Crystallographic Data and Refinement Statistics for 4X21**

| Data Collection | |
| --- | --- |
| Beamline | HZB-BESSYII BL14.1 |
| Space group | $P2_12_12$ |
| Cell dimensions (Å) | $a = 155.94$ $b = 109.84$ $c = 43.91$ |
| | $\alpha = \beta = \gamma = 90°$ |
| Wavelength (Å) | 0.918409 |
| Resolution (Å)$^a$ | 50.0−1.95 (2.00−1.95) |
| Measured reflections | 741250 (55236) |
| Unique reflections | 56051 (4113) |
| CC1/2$^a$ | 99.9 (67.9) |
| Completeness (%)$^a$ | 100 (100) |
| Redundancy$^a$ | 13.2 (13.4) |
| $I/\sigma(I)^a$ | 14.35 (2.42) |
| Refinement | |
| Resolution (Å) | 46.89−1.95 |
| $R_{work}/R_{free}$ $^b$ | 19.86/22.30 |
| Number of atoms | 5595 |
| Protein | 5195 |
| Water | 344 |
| **7d** | 56 |
| B-factors (Å$^2$)$^{c,d}$ | |
| Protein | 35.7 (13.9−79.4) |
| Water | 38.6 (16.3−61.9) |
| **7d** | 36.3 (24.7−49.1) |
| r.m.s.d. | |
| Bond length (Å) | 0.014 |
| Bond angles (deg) | 1.241 |
| Ramachandran Plot | |
| most favorable (%) | 98.5 |
| allowed (%) | 1.5 |

$^a$Values in parentheses are for the highest-resolution shell. $^b R_{free}$ represents a 5% subset of the total reflections, that are excluded from refinement. $^c$The overall mean B-factor is found to be in good agreement with the Wilson B-factor. $^d$The overall coordinate error was calculated as approximately 0.2 Å.

as predicted by the QM model. The iodine···sulfur distance $d_{I\cdots S}$ in chain A is with 315 pm rather short, featuring an excellent σ-hole angle of 170.3°. Thus, our prediction (319 pm/165°) closely matches the experimentally found parameters. The ligand in chain B binds in an almost identical mode to JNK3 (see Figure S5). The hydrogen bonds to the hinge in chain B have a length of 204 pm (N···NH) and 215 pm (NH···OC), deviating by only 1 and 9 pm from chain A. The distance of the weak hydrogen bond in chain B ($d_{CH\cdots OC}$ = 234 pm) also closely matches chain A. The halogen bond has a distance of only 305 pm and a σ-hole angle of 168.4°. The deviation of the σ-hole angle is only marginal and the shortening of $d_{I\cdots S}$ by 10 pm appears small. Still, this very short contact can be actually detrimental for the energy of the system. Indeed, as demonstrated in Figure 5d, the attractiveness of the interaction is further decreased at a distance of 305 pm. However, this very short length can also be explained by significant radiation damage to the iodine of **7d** in chain B. In addition, there is also a small change in the geometry of the residues lining the binding site, caused by a crystallization interface in chain A and another in chain B. We provide a detailed account of these structural changes and their effect on the ligand binding in the Supporting Information (Figure S7). It should be noted that in chain B, the electron density of the *N*-ethylcarboxamide moiety

is more clearly defined, confirming that the 180° flipped orientation of the ethyl group in comparison to the structure of **7b** in 2P33 is likely favored.

Figure 5b shows chain A (orange) superposed onto the crystal structure of 2P33 (cyan). Both share an almost identical binding mode. The hydrogen bonds to the hinge in chain A of 4X21 only deviate by 2 pm (N···NH) and 9 pm (NH···OC) from 2P33. The distance of the weak hydrogen bond (CH···OC) is likewise altered by 3 pm only. The difference in the hinge binding geometry, thus, appears negligible. The halogen bond of the iodine in 4X21 is 21 pm shorter than for chlorine and the σ-hole angle improves by ∼10° to an almost linear arrangement. While the small deviation of the σ-hole angle is almost optimal, the significantly shorter $d_{I\cdots S}$ could imply an improved, tighter binding, but could also represent a repulsive contribution. Similar to **7b**, **7d** binds to LYS93 via an interstitial water molecule. The hydrogen bond network around LYS93 in 4X21 appears to be slightly extended by another water molecule, defined with good electron density and a low B-factor.

In summary, the comparison of both crystal structures demonstrates that **7d** and **7b** are both tightly bound by their hydrogen bonding network involving the hinge region (GLU147, MET149) and the bridged interaction with LYS93. From the QM-based prediction we inferred that the indole-substituted 2-aminopyrimidine scaffold will not shift significantly to accommodate the larger iodine, because the stronger σ-hole of iodine leads to a decreased halogen bond length of 319 pm, in comparison to 333 pm for chlorine. Despite a reasonably good consistency between our QM-model and the crystal structure (see Table S1), we deemed it interesting that the halogen bond length for chlorine (**7b**) was predicted to be slightly shorter, while the length for iodine (**7d**) was predicted to be slightly longer than in the crystal structure.

**5. Bivalent, Mixed Halogen and Chalcogen Bonding and Torsional Flexibility of the MET146 $\chi_3$ Dihedral Angle $\delta_{C\beta-C\gamma-S-C\epsilon}$.** In order to analyze the relative strength of the experimentally found halogen bond with a distance of 315 pm, we employed QM calculations on a TPSS-D2/TZVPP level of theory including the entire ligand **7d** together with the capped methionine (cMET146 = 2-acetamido-*N*-methyl-4-(methylthio)butanamide). We also used this model system to determine another crucial parameter: During refinement of the crystal structure 4X21, the electron density of the $C_\epsilon$ of MET146 is not as well-defined as the rest of the residue. The electron density map for the neighbor atom sulfur is spherically elongated, making two alternative, preferential orientations of the terminal $C_\epsilon$ most plausible. It should be noted that the lack of electron density for this atom clearly indicates a higher flexibility, which might be connected to the stronger entropy contribution to binding of **7d**, which we observed in ITC experiments.

Even more importantly, the orientation of $C_\epsilon$ is of utmost importance for the strength of the halogen bond.[4] The highest electron densities (of the "lone pairs") are found almost perpendicular to the plane defined by $C_\gamma$, S, and $C_\epsilon$, located above and below the sulfur. In plane, the sulfur exhibits σ-holes of itself, efficiently preventing a halogen bond to be formed in this orientation. Thus, the MET146 $\chi_3$ dihedral angle $\delta_{C\beta-C\gamma-S-C\epsilon}$ has an essential influence on the energy evaluation of the system. From the experimental electron density, we concluded that the most likely dihedral angle is $\delta(C_\beta-C_\gamma-S-C_\epsilon) = -95°$.
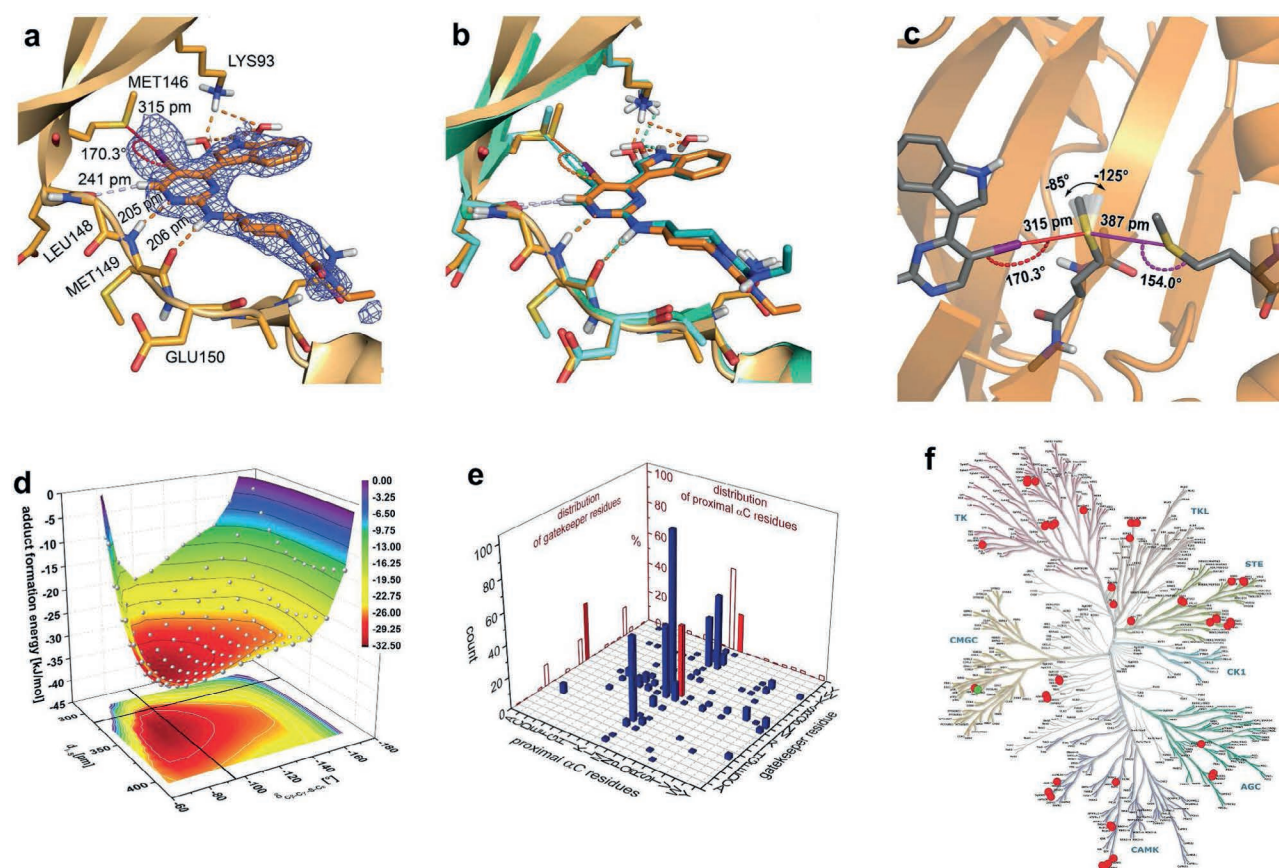
**Figure 5.** Crystal structure of the 5-iodopyrimidine derivative (**7d**) in complex with JNK3 and computational analysis of the iodine···sulfur halogen bond. (a) The binding site of chain A of 4X21 is presented with emphasis on the tight hydrogen bond network toward the hinge region and via two bridging water molecules toward LYS93. The protein is shown as a bright orange cartoon representation, with selected residues in the cavity depicted as stick models. The ligand is shown as an orange stick model, and hydrogen bonds are highlighted with broken, orange lines. The weak hydrogen bond (C—H···O=C) involving GLU147 is indicated by light blue, broken lines. The halogen bond targeting the gatekeeper MET146 is shown as a red line, while the $\sigma$-hole angle is depicted as a curved, dotted, red line. Characteristic distances and the $\sigma$-hole angle are denoted explicitly (see also Table S1). An unbiased simulated-annealing omit electron-density map ($F_o - F_c$) for the ligand is shown at a contour level of 3.0 $\sigma$. Only density within a maximal distance of 220 pm of any atom of the ligand is displayed for clarity reasons. The $N$-ethyl-substituent of the piperidinyl-1-carboxamide moiety of 7d protruding from the cavity was not clearly resolved in the crystal structure. (b) Overlay of chain A of 4X21 (bright orange) and 2P33 (bright cyan). The respective ligands **7d** of 4X21 and **7b** of 2P33 are colored in orange and cyan. The polar interactions (hydrogen, weak hydrogen and halogen bonds) are depicted similar to panel (a). For better differentiation between the structures, the color code of the ligands is likewise used for the interactions. Alignment and superposition has been done in MOE.[37] Pictures have been prepared and rendered in PyMOL.[18] (c) MET146 is engaged in a bivalent halogen bond (red) and chalcogen bond (purple) donated by **7d** and MET115, respectively. Distances ($d_{I···S}$ and $d_{S···S}$) and $\sigma$-hole angles ($\alpha_{C-I···S}$ and $\alpha_{C\gamma-S···S}$) characterizing the quality of the interactions are depicted in the model system (shown as sticks). This consists of **7d** and the two capped methionine residues (cMET = 2-acetamido-$N$-methyl-4-(methylthio)butanamide) in position 146 and 115. The flexibility of the $\chi_3$ dihedral angle ($\delta_{C\beta-C\gamma-S-C\epsilon}$) is represented by partially transparent replicas of the $C_\epsilon$-S bond of MET146. (d) The depicted 3D plot shows the results (white dots) of the systematic variation of $\delta_{C\beta-C\gamma-S-C\epsilon}$ in increments of 10° and $d_{I···S}$ between 295 and 405 pm in steps of 10 pm. Adduct formation energies (in kJ/mol) were calculated on a TPSS-D2/TZVPP level of theory by subtracting the DFT-energy of the three educts (**7d**, cMET146, and cMET115) from the DFT-energy of the complex. Fitting a surface to the data, which was color coded with a gradient representing the adduct formation energies, and projection of the energy profile onto the $xy$-plane was done in Origin9.1.[38] The point of intersection between the two black lines plotted in the projected map indicates the most reasonable experimental geometry. (e) 3D bar chart depicting the number of human kinases for each possible combination of gatekeeper residue and their proximal residue in the $\alpha$C-helix. Residues were identified and statistics derived based on kinase sequence alignments[16] obtained from http://kinase.com/human/kinome/phylogeny.html. Red 2D bar charts projected onto the side and back of the 3D plot provide the relative overall distribution of residues for both positions in 486 human kinases. The red 3D bar highlights 45 different kinases (including JNK3) featuring methionine in both positions. (f) Distribution of these 45 kinases highlighted in the dendrogram of human protein kinases according to Manning et al.[16] by red dots. For clarity, JNK3 is represented by a green dot.

Closer inspection of the proximal amino acids in the back pocket behind MET146, reveals an unexpected, but highly interesting contact featured consistently in both chains of 4X21, however, not in 2P33. MET115, located in helix $\alpha$C, adopts a different rotameric state in 4X21, forming a chalcogen bond with the sulfur electron density of MET146 from the opposite

direction of the halogen bond (Figure 5c). Similar to halogens, the heavier chalcogens such as sulfur show an anisotropic electron distribution, leading to $\sigma$-holes (positive electrostatic potentials) in elongation of the $C_{\gamma/\epsilon}$-S bond vectors.[39–43] The distance $d_{S···S}$ between MET115 and MET146 corresponds with 387 pm (or 382 pm in chain B) quite well to the minima

reported in the literature for a similar model system (403 pm).[39] An almost linear arrangement ($\alpha_{C\gamma-S\cdots S} = 180°$) indicates a good alignment of the $\sigma$-hole of the donor methionine with the sulfur of the acceptor methionine. We find a deviation of only $\Delta\alpha_{C\gamma-S\cdots S} = 26.0°$ for chain A and $20.3°$ for chain B. Additionally, a very good overlap of the donor $\sigma$-hole of MET115 with the highest electron density (lone pairs) above the $C_\gamma-S-C_\varepsilon$ plane of MET146 is observed for $\chi_3$ dihedral angle $\delta_{C\beta-C\gamma-S-C\varepsilon}$ close to $-95°$. Hence, the chalcogen bond donated from MET115 and the halogen bond donated from **7d** are engaged in a bivalent interaction with MET146. As shown in Figure 5c, we have therefore extended our model system by the capped MET115 (cMET115) to elucidate, how this additional interaction restricts the conformational freedom of MET146.

We rotated $\delta_{C\beta-C\gamma-S-C\varepsilon}$ in steps of $10°$ between $-65°$ and $-175°$. All dihedral angles beyond this span clearly led to significant clashes with the ligand or MET115. We simultaneously altered the distance $d_{I\cdots S}$ from 295 to 405 pm in increments of 10 pm, giving a total of 144 single points to be evaluated. The results can be seen in Figure 5d. According to the most reasonable geometry derived from the electron density, the interaction energy amounts to approximately $-28.6$ kJ/mol. It should be noted that this interaction energy is only the simple adduct formation energy of the complex, not including desolvation effects or entropy contributions. Thus, the numbers cannot be interpreted as absolute, but only relative values. If the dihedral angle is changed by more than $20°$ toward the iodine atom ($\delta_{C\beta-C\gamma-S-C\varepsilon} = -75°$), the mismatch of the sulfur electron density and the iodine $\sigma$-hole will be dramatically increased and steric clashes start to arise from the too close $C_\varepsilon\cdots S$ contact, leading to an almost complete loss of the attractive interaction. On the basis of steric repulsion with MET115, dihedral angles of $\delta_{C\beta-C\gamma-S-C\varepsilon} < -165°$ are not possible. The interaction energy in the range $-85° > \delta_{C\beta-C\gamma-S-C\varepsilon} > -115°$ is quite similar at the given experimental distance of $d_{I\cdots S} = 315$ pm (black line in Figure 5d). Moreover, at a larger distance, the energy surface becomes even more shallow, differing between $-75°$ and $-145°$ by only up to 6 kJ/mol. The chalcogen bond provides between a $\delta_{C\beta-C\gamma-S-C\varepsilon}$ of $-65°$ and $-145°$ an almost constant stabilization of the complex ($-9.6$ to $-10.8$ kJ/mol on a TPSS-D2/TZVPP-level of theory), quite independent of the dihedral angle.

The strength of the interaction could be further improved by a larger distance of the iodine atom ($\sim$345 pm instead of 315 pm). It is obvious from Figure 5d that the strength of the halogen bond at 315 pm will be improved or maintained up to a distance of more than 400 pm at these dihedral angles. Thus, we conclude that the experimentally determined geometry represents a still favorable, attractive interaction, however, fails to utilize the potential of the stronger $\sigma$-hole of iodine. This is due to the binding mode being clearly dominated by the tight hydrogen-bonding network to the hinge and LYS93. In addition, the too close iodine atom of **7d** increases the flexibility of $C_\varepsilon$ of MET146, however, the rest of the residue cannot adapt, because it accepts the chalcogen bond from MET115.

We further evaluated the relative conformational strain of the MET146 side chain, when it adopts different dihedral angles. We calculated the relative energy of our cMET146 model system, when modifying $\delta_{C\beta-C\gamma-S-C\varepsilon}$. The identified torsional minimum at approximately $-75°$ causes an almost repulsive interaction. The most plausible dihedral of $-95°$ causes a

slightly increased conformational strain of $+4.0$ kJ/mol, which appears to be less relevant in comparison to the gain of interaction energy (in Figure 5d). Still, combining the trend observed for the conformational strain of MET146 with the trend in the interaction energies, it appears plausible that a certain fluctuation of the $\chi_3$ dihedral angle of MET146 of $\pm15°$ with respect to the best angle might be possible. Interestingly, in comparison to 4X21, the crystal structure of **7b** (2P33) appears to have a more well-defined electron density for $C_\varepsilon$, probably also correlating with the larger halogen bond distance and the smaller size of chlorine. Beside desolvation effects, this observation is one possible explanation for changes in enthalpy and entropy between these ligands, while the overall $\Delta G$ is only insignificantly different. We have calculated solvation effects using COSMO-RS (see Supporting Information, Figure S2, and Table S2) for the ligands **7a**−**d** in complex with the capped gatekeeper cMET146. The results show a good correlation between the COSMO-RS corrected and the gas phase energies ($R^2 = 0.96$). The desolvation energy ranges between $-18.0$ kJ/mol (4X21) and $-19.8$ kJ/mol (2P33).

## ■ CONCLUSION

We performed a matched molecular pair analysis of a series of aminopyrimidine-based JNK3 inhibitors, designed to investigate halogen bonding effects with the sulfur of MET146, the gatekeeper residue of this kinase. With a significant portion of the human kinome sharing a methionine residue as the gatekeeper, this case study is an interesting starting point for a more systematic investigation of X$\cdots$S halogen bonding effects on affinity and selectivity tuning in kinase drug discovery. There are important lessons learned from this case study, which may help to guide future molecular design projects.

Despite the usual assumption that halogen bonding strength increases from chlorine to bromine to iodine, in this ligand series we encounter a plateau of affinity, reached already for chlorine (**7b**). Neither bromine (**7c**), nor iodine (**7d**) were found to provide any statistically significant advance in affinity in fluorescence polarization-based measurements or isothermal titration calorimetry. Thus, these two halogens clearly fail to reach the full potential according to the size of their $\sigma$-hole, which we have demonstrated by calculating their $V_{max}$ value and by visualizing their electrostatic potentials. It should be emphasized that the lack of a significant improvement does not imply that they do not have greater potential for gaining affinity; however, it means that in this system other factors did not permit **7c** or **7d** to make use of this potential. On the basis of the crystal structure of **7d** in complex with JNK3 (4X21), its comparison to the original structure of **7b** in complex with JNK3 (2P33), and exploiting a systematic QM-based evaluation of parameters crucial for binding, we are able to identify these factors and to rationalize, when halogen exchange toward bromine and iodine might be a rewarding strategy and when it may not provide any advantage. Such cases, where chlorine already is the optimum can certainly be beneficial, particularly when being concerned about attrition caused by adverse effects that could be related to these heavier halogens.

From Figure 5, there is clear evidence that **7d** is held so tightly in place by the hydrogen bonding network involving the hinge (MET149, GLU147) and LYS93 that a significant displacement of the scaffold is not expectable. Despite being often considered rather flexible, methionine seems not to be able to adapt to the strongly fixed ligand by increasing the halogen bond distance from 315 pm (in the crystal structure)

to roughly 335 to 365 pm, where the optimum is expected according to our model calculations (shown in Figure 5d). We have shown additionally that the conformational strain of such adaptive processes should not be underestimated and can be a real liability for ligand binding. Interestingly, in our JNK3 complex with **7d**, MET146 is bivalently engaged by a halogen bond from the ligand and a chalcogen bond from MET115. This dual interaction strongly restricts the flexibility of MET146, except for the $\chi_3$ dihedral angle. The slightly too close iodine atom in **7d** enhances the flexibility of $C_\varepsilon$ and causes variations in $\chi_3$ within the boundaries of the bivalent halogen/chalcogen bond, counterbalanced by the increasing conformational strain of MET146. In the context of the significant shifts between enthalpy ($\Delta\Delta H(\mathbf{7d{-}7b}) = +3.0$ kcal·mol$^{-1}$) and entropy ($\Delta\Delta S(\mathbf{7d{-}7b}) = +11.1$ cal·mol$^{-1}$·K$^{-1}$), this change in flexibility might be one relevant contribution. Of course, changes in desolvation can likewise contribute to the observed enthalpy−entropy compensation between **7b** and **7d**.[44]

As we have shown for JNK3 versus p38$\alpha$, halogen bonding to the gatekeeper residue methionine can modify the selectivity profile of kinase inhibitors. This effect can, of course, depend not only on the difference between the gatekeeper residues (e.g., MET vs THR) and the discriminative way both can be targeted by halogen bonding, but it also may depend on small to subtle geometric variances in the binding site featuring the same gatekeeper. As shown in this study, the difference between an attractive and a repulsive interaction can be as small as 20−30 pm. In addition, small changes in the spherical orientation of the residue (e.g., the $\chi_3$ dihedral angle in methionine) or deviations from a linear arrangement ($\sigma$-hole angle of 180°) by 10−30° can be sufficient to cause a substantial shift in selectivity. We prepared the gatekeeper mutants M146A, M146L, and M146T of JNK3 to scrutinize the role of the halogen bond in molecular recognition of **7b−d**. The strong loss of affinity for M146A suggests that an interaction with the gatekeeper residue is key to achieving high affinity for all ligands. The high similarity of the ligand binding data between M146T and p38$\alpha$ implies that the exchange of the gatekeeper into threonine is the major determinant of p38$\alpha$/JNK3-selectivity. Apparently, none of the interaction patterns (halogen bond, hydrogen bond, dispersive/hydrophobic interaction) which threonine can form is able to compensate for the loss of the halogen bond toward methionine. In contrast, leucine as a gatekeeper of similar size can compensate, based on dispersive/hydrophobic interactions, for the loss of the methionine interaction. It shows an interesting differentiation between the halogens: the small chlorine is clearly preferred over the larger bromine and even more over iodine. Thus, preference by leucine increases with decreasing $\sigma$-hole size and decreasing volume of the halogen. Hence, although failing to modify affinity toward JNK3-wt, the choice of halogen can have an important impact on kinase selectivity.

A fascinating paradigm observed in this study is the stabilization of MET146 by the proximal MET115. The quite favorable chalcogen bond geometry restricts the motility of MET146, preventing a more efficient adaption to the too close iodine atom. Thus, this combination of chalcogen bonding partners in this position may give rise to some affinity cliffs in the structure−activity relationship (SAR) of ligand series. As highlighted in Figure 5e, when the gatekeeper is a methionine residue, the probability of a methionine as its proximal neighbor in the $\alpha$C-helix of the back pocket is only 24.6%. The overall

chance of two methionine residues occurring in these two positions, is only 9.3% with respect to the human kinome. Leucine has a much higher prevalence in this position of the $\alpha$C-helix, however, it only can interact by forming hydrophobic contacts. Interestingly, analysis of the distribution of the 45 kinases featuring both methionine residues reveals an unequal distribution among the kinase subfamilies. While JNK3 and JNK1, both share this feature, neither JNK2, nor any other kinase in the CMGC subtree (see Figure 5f) show this combination of methionine residues. It seems to occur in small clusters of two to four closely related kinases, however, is spread rather equally across the different subfamilies, except for the Casein Kinase 1 family. Still, close proximity between two methionine residues in these two positions will not by itself guarantee formation of a chalcogen bond, as demonstrated in 2P33. However, the described paradigm may provide new opportunities for selectivity profiling in the human kinome. Clearly, more test cases and a more exhaustive comparison is needed to highlight general trends.

In a nutshell, what did we learn from this model system? Molecular design of halogen bonds in binding sites starting from molecules of high complexity[45] that comprise a substantial number of highly preoptimized interactions is certainly not trivial. Best results can be expected when the carbon···sulfur distance between the aromatic scaffold and methionine is approximately 540−560 pm and the $\sigma$-hole angle only deviates by up to 20° from linearity. It is dangerous to rely on adaptive behavior of the binding site, even with a side chain such as methionine that is usually perceived to be more flexible than most other residues. An alternative, much better strategy for such a situation is to select one of the most efficient fragments from a library such as HEFLibs (halogen-enriched fragments),[46,47] which shows an optimized halogen bond geometry with this gatekeeper and then try to grow or merge the fragment to introduce an optimized network of other interactions. Still, molecular design is a most valuable method for lead optimization. As a consequence, QM-based scoring functions for docking approaches, accounting precisely for the geometric restraints of halogen bonding[17,48] can be valuable tools complementary to the direct use of quantum chemistry or integrated QM/SQM/MM-approaches.[22,49,50]

## ■ MATERIALS AND METHODS

See Supporting Information.

## ■ ASSOCIATED CONTENT

### Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b07090.

> Detailed information about halogen bond tuning ($V_{max}$), design of matched molecular pairs by quantum chemistry, synthesis procedures, biophysical characterization (DSF). Materials and methods for: QM calculations, solvation effects, organic chemistry, biochemistry, DSF, ITC, FP assay, statistical data analysis, crystallization of JNK3, data collection and structure determination of 4X21. Schemes for synthesis of **7a** and **7b** (S1), loss of regioselectivity during Suzuki cross coupling (S2), and structure of RN4646 (S3). Tables with detailed measurements of computed versus experimental complexes (S1), COSMO-RS solvation-corrected interaction energies (S2), and DSF measure-

ments (S3). Figures showing computed structures of MMPs in complex with JNK3 (S1), model system for calculation of solvation effects (S2), DSF curves (S3), summary of ITC results decomposed into $\Delta G$, $\Delta H$ and $-T\Delta S$ (S4), superposition of chain A and B of 4X21 (S5), unbiased simulated-annealing omit ($F_o - F_c$) electron-density map for **7d** and MET146 in 4X21 (S6), and crystal contacts in 4X21 (S7). Detailed synthesis and characterization of compounds **3−7** and additional references. (PDF)

## ■ AUTHOR INFORMATION

**Corresponding Author**
*frank.boeckler@uni-tuebingen.de

**Author Contributions**
§A.L. and M.G. contributed equally.

**Notes**
The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

## ■ REFERENCES

(1) Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16789.

(2) Politzer, P.; Lane, P.; Concha, M. C.; Ma, Y. G.; Murray, J. S. *J. Mol. Model.* **2007**, *13*, 305.

(3) Hardegger, L. A.; Kuhn, B.; Spinnler, B.; Anselm, L.; Ecabert, R.; Stihle, M.; Gsell, B.; Thoma, R.; Diez, J.; Benz, J.; Plancher, J.-M.; Hartmann, G.; Banner, D. W.; Haap, W.; Diederich, F. *Angew. Chem., Int. Ed.* **2011**, *50*, 314.

(4) Wilcken, R.; Zimmermann, M. O.; Lange, A.; Zahn, S.; Kirchner, B.; Boeckler, F. M. *J. Chem. Theory Comput.* **2011**, *7*, 2307.

(5) Wilcken, R.; Zimmermann, M.; Lange, A.; Zahn, S.; Boeckler, F. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 935.

(6) Wilcken, R.; Zimmermann, M. O.; Lange, A.; Joerger, A. C.; Boeckler, F. M. *J. Med. Chem.* **2013**, *56*, 1363.

(7) Lange, A.; Zimmermann, M. O.; Wilcken, R.; Zahn, S.; Boeckler, F. M. *J. Chem. Inf. Model.* **2013**, *53*, 3178.

(8) Politzer, P.; Murray, J. S.; Clark, T. *Phys. Chem. Chem. Phys.* **2013**, *15*, 11178.

(9) Desiraju, G. R.; Ho, P. S.; Kloo, L.; Legon, A. C.; Marquardt, R.; Metrangolo, P.; Politzer, P.; Resnati, G.; Rissanen, K. *Pure Appl. Chem.* **2013**, *85*, 1711.

(10) Beno, B. R.; Yeung, K.-S.; Bartberger, M. D.; Pennington, L. D.; Meanwell, N. A. *J. Med. Chem.* **2015**, *58*, 4383.

(11) Cohen, M. S.; Zhang, C.; Shokat, K. M.; Taunton, J. *Science* **2005**, *308*, 1318.

(12) Gehringer, M.; Muth, F.; Koch, P.; Laufer, S. A. *Expert Opin. Ther. Pat.* **2015**, *25*, 849.

(13) Koch, P.; Gehringer, M.; Laufer, S. A. *J. Med. Chem.* **2015**, *58*, 72.

(14) Scapin, G.; Patel, S. B.; Lisnock, J.; Becker, J. W.; LoGrasso, P. V. *Chem. Biol.* **2003**, *10*, 705.

(15) Alam, M.; Beevers, R. E.; Ceska, T.; Davenport, R. J.; Dickson, K. M.; Fortunato, M.; Gowers, L.; Haughan, A. F.; James, L. A.; Jones, M. W.; Kinsella, N.; Lowe, C.; Meissner, J. W. G.; Nicolas, A.-L.; Perry, B. G.; Phillips, D. J.; Pitt, W. R.; Platt, A.; Ratcliffe, A. J.; Sharpe, A.; Tait, L. J. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 3463.

(16) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. *Science* **2002**, *298*, 1912.

(17) Ruff, M.; Zimmermann, M. O.; Lange, A.; Boeckler, F. M., unpublished data, 2015.

(18) Delano, W. L. *The PyMOL Molecular Graphics System*, 1.2r.1; DeLano Scientific LLC: Palo Alto, CA, 2009.

(19) Brickmann, J.; Exner, T. E.; Gimmler, J.; Lautenschläger, P.; Heiden, W.; Moeckel, G.; Zahn, D. *MOLCAD II*, V1.4; MOLCAD GmbH: Darmstadt, Germany; 2010, http://www.molcad.de.

(20) Brickmann, J.; Exner, T. E.; Keil, M.; Marhofer, R. J. *J. Mol. Model.* **2000**, *6*, 328.

(21) Lange, A.; Zimmermann, M. O.; Heidrich, J.; Exner, T. E.; Boeckler, F. M., unpublished data, 2015.

(22) Fanfrlík, J.; Kolář, M.; Kamlar, M.; Hurný, D.; Ruiz, F. X.; Cousido-Siah, A.; Mitschler, A.; Řezáč, J.; Munusamy, E.; Lepšík, M.; Matějíček, P.; Veselý, J.; Podjarny, A.; Hobza, P. *ACS Chem. Biol.* **2013**, *8*, 2484.

(23) Riley, K. E.; Murray, J. S.; Fanfrlik, J.; Rezac, J.; Sola, R. J.; Concha, M. C.; Ramos, F. M.; Politzer, P. *J. Mol. Model.* **2011**, *17*, 3309.

(24) Fresneda, P. M.; Delgado, S.; Francesch, A.; Manzanares, I.; Cuevas, C.; Molina, P. *J. Med. Chem.* **2006**, *49*, 1217.

(25) Ottoni, O.; Cruz, R.; Alves, R. *Tetrahedron* **1998**, *54*, 13915.

(26) Rossignol, E.; Youssef, A.; Moreau, P.; Prudhomme, M.; Anizon, F. *Tetrahedron* **2007**, *63*, 10169.

(27) Caldarelli, M.; Angiolini, M.; Colombo, R.; Disingrini, T.; Nuvoloni, S.; Posteri, H.; Salsa, M.; Silvagni, M., Pyrazolo-quinazolines. Eur. Pat. Appl. 2303891, April 6, 2011.

(28) Lea, W. A.; Simeonov, A. *Expert Opin. Drug Discovery* **2011**, *6*, 17.

(29) Wätzig, H.; Oltmann-Norden, I.; Steinicke, F.; Alhazmi, H. A.; Nachbar, M.; El-Hady, D. A. A.; Albishri, H. M.; Baumann, K.; Exner, T. E.; Boeckler, F. M.; El Deeb, S. *J. Comput.-Aided Mol. Des.* **2015**, *29*, 1.

(30) Munoz, L.; Selig, R.; Yeung, Y. T.; Peifer, C.; Hauser, D.; Laufer, S. *Anal. Biochem.* **2010**, *401*, 125.

(31) Ansideri, F.; Lange, A.; Elgokha, A. A.; Boeckler, F. M.; Koch, P., unpublished data, 2015.

(32) Nikolovska-Coleska, Z.; Wang, R.; Fang, X.; Pan, H.; Tomita, Y.; Li, P.; Roller, P. P.; Krajewski, K.; Saito, N. G.; Stuckey, J. A.; Wang, S. *Anal. Biochem.* **2004**, *332*, 261.

(33) *KaleidaGraph*, 4.02; Synergy Software: Reading, PA, 1986−2013.

(34) Koeberle, S. C.; Romir, J.; Fischer, S.; Koeberle, A.; Schattel, V.; Albrecht, W.; Grütter, C.; Werz, O.; Rauh, D.; Stehle, T.; Laufer, S. A. *Nat. Chem. Biol.* **2012**, *8*, 141.

(35) Xie, X.; Gu, Y.; Fox, T.; Coll, J. T.; Fleming, M. A.; Markland, W.; Caron, P. R.; Wilson, K. P.; Su, M. S. S. *Structure* **1998**, *6*, 983.

(36) Swahn, B.-M.; Huerta, F.; Kallin, E.; Malmström, J.; Weigelt, T.; Viklund, J.; Womack, P.; Xue, Y.; Öhberg, L. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 5095.

(37) *Molecular Operating Environment (MOE)*, 2013.08; Chemical Computing Group Inc.: Montreal, QC, 2015.

(38) *Origin* (OriginLab: Northampton, MA).

(39) Bleiholder, C.; Werz, D. B.; Koppel, H.; Gleiter, R. *J. Am. Chem. Soc.* **2006**, *128*, 2666.

(40) Murray, J. S.; Lane, P.; Clark, T.; Politzer, P. *J. Mol. Model.* **2007**, *13*, 1033.

(41) Murray, J. S.; Lane, P.; Politzer, P. *Int. J. Quantum Chem.* **2008**, *108*, 2770.

(42) Wang, W.; Ji, B.; Zhang, Y. *J. Phys. Chem. A* **2009**, *113*, 8132.

(43) Politzer, P.; Murray, J. S.; Clark, T. *Phys. Chem. Chem. Phys.* **2013**, *15*, 11178.

(44) Ferenczy, G. G.; Keserű, G. M. *J. Chem. Inf. Model.* **2010**, *50*, 1536.

(45) Hann, M. M.; Leach, A. R.; Harper, G. *J. Chem. Inf. Model.* **2001**, *41*, 856.

(46) Zimmermann, M. O.; Lange, A.; Wilcken, R.; Cieslik, M. B.; Exner, T. E.; Joerger, A. C.; Koch, P.; Boeckler, F. M. *Future Med. Chem.* **2014**, *6*, 617.

(47) Wilcken, R.; Liu, X.; Zimmermann, M. O.; Rutherford, T. J.; Fersht, A. R.; Joerger, A. C.; Boeckler, F. M. *J. Am. Chem. Soc.* **2012**, *134*, 6810.

(48) Zimmermann, M. O.; Lange, A.; Boeckler, F. M. *J. Chem. Inf. Model.* **2015**, *55*, 687.

(49) Dobeš, P.; Řezáč, J.; Fanfrlík, J.; Otyepka, M.; Hobza, P. *J. Phys. Chem. B* **2011**, *115*, 8581.

(50) Brahmkshatriya, P. S.; Dobes, P.; Fanfrlik, J.; Rezac, J.; Paruch, K.; Bronowska, A.; Lepsik, M.; Hobza, P. *Curr. Comput.-Aided Drug Des.* **2013**, *9*, 118.