Review

Jürgen Lassak*, Franziska Koller, Ralph Krafczyk and Wolfram Volkwein

Exceptionally versatile – arginine in bacterial post-translational protein modifications

https://doi.org/10.1515/hsz-2019-0182
Received March 5, 2019; accepted June 1, 2019; previously published online June 11, 2019

Abstract: Post-translational modifications (PTM) are the evolutionary solution to challenge and extend the boundaries of genetically predetermined proteomic diversity. As PTMs are highly dynamic, they also hold an enormous regulatory potential. It is therefore not surprising that out of the 20 proteinogenic amino acids, 15 can be post-translationally modified. Even the relatively inert guanidino group of arginine is subject to a multitude of mostly enzyme mediated chemical changes. The resulting alterations can have a major influence on protein function. In this review, we will discuss how bacteria control their cellular processes and develop pathogenicity based on post-translational protein-arginine modifications.

Keywords: advanced glycation end products; arginine acetylation; arginine ADP-ribosylation; arginine glycosylation; arginine hydroxylation; arginine methylation; arginine phosphorylation.

Introduction

The complexity of proteomes is so tremendous that it exceeds the predictions based on genome data by up to three orders of magnitude (Walsh et al., 2005). The major contribution to proteome diversification comes from post-translational modifications (PTMs). As the name says, PTMs are changes that take place subsequent to protein biosynthesis either directly at the nascent chain or at the fully translated and even folded protein. On the one hand, a highly reactive chemical compound can react spontaneously with an amino acid side chain in a rather random manner. On the other hand, post-translational proteome diversification is more specifically achieved by an arsenal of specialized enzymes. In higher eukaryotes, up to 5% of the total genome are dedicated to such modifiers (Walsh et al., 2005). Although bacteria are often considered as simple organisms with very basal cellular regulation, their proteome is also subject to substantial directed post-translational changes (Grangeasse et al., 2015). Notably, prokaryotic modification enzymes are not limited to those that act on endogenous proteins but also encompass effectors, which are secreted to act within a host. Most often, these modification enzymes alter or inhibit a protein function and thus work as pathogenicity factors. Accordingly, PTMs are utilized as a means to promote survival and proliferation of the bacterium inside a host. As those prokaryotic modification systems can be very unique they might be attractive targets for biomedical application (Lassak et al., 2015) or the development of novel molecular tools (Nadal et al., 2018).

In general, PTMs can be classified into two major categories (Walsh et al., 2005): the first comprises the alteration of the primary protein sequence by non-ribosomal formation of a new peptide bond or conversely its cleavage by limited proteolysis. The second category, which is the main focus of this review, means the alteration of an amino acid side chain preexisting in the polypeptide. Most commonly this is achieved by the addition of a chemical group, which itself can be highly diverse. Accordingly, the corresponding catalog of PTMs is large and includes but is not limited to methylation, acylation, phosphorylation as well as glycosylation. Out of the 20 canonical amino acids 15 can serve as acceptor substrates (Walsh et al., 2005) and if N-terminal acylation is included in this calculation, then there is no exception at all (Karve et al., 2011). Of these, arginine has a distinguishing feature – the guanidino group, which has the highest pKa value of all side chains (pKa = 13.8 in aqueous solution at 25°C) (Fitch et al., 2015). The resulting positive charge promotes the location of arginine on the hydrophilic protein surface and enables the binding...
of negatively charged molecules such as DNA (Schlundt et al., 2017). In combination with its side chain length and flexibility, the guanidino group allows for perfect hydrogen-bonding geometries with not only nucleobases and phosphate residues but also other amino acids (Fuhrmann et al., 2015a,b). Thus, it rapidly becomes clear that arginine is a physiologically relevant target for post-translational modification (Figure 1). It is therefore not surprising, that this amino acid is frequently targeted by diverse PTMs. However, the inert guanidine group impedes a nucleophilic attack onto a donor substrate, which is the prerequisite for its post-translational modification.

In the following, we will discuss the strategies to overcome difficulties in activating the guanidino group to allow for modification and to exploit the full regulatory potential inherent in changing the chemical properties of arginine. In this regard, this review article puts particular emphasis on the unique mechanisms that microorganisms utilize to glycosylate, to ADP-ribosylate and to phosphorylate this residue.

**Arginine glycosylation – sweet switches to control protein function**

Protein-glycosylation is the attachment of a sugar moiety onto an amino acid side chain within a polypeptide. The modification reaction is catalyzed by a class of enzymes known as glycosyltransferases (GTs) and can be found in all domains of life (Schwarz and Aebi, 2011). Most commonly, the hydroxyl oxygen of serine or threonine (O-glycosylation) or the amide nitrogen of asparagine (N-glycosylation) serve as acceptor for type of modification (Lairson et al., 2008; Lafite and Daniellou, 2012). N-glycosylation represents the most common protein glycosylation and is implicated in numerous cellular processes (Helenius and Aebi, 2004; Abouelhadid et al., 2019). For a long time, it was thought that the general pathway in the endoplasmic reticulum of eukaryotes is the only biosynthetic route to generate N-linked proteoglycans (Nothaft and Szymanski, 2010). In 1999 however, the identification of the oligosaccharyltransferase (OST) undecaprenyl-diphosphooligosaccharide-protein glycotransferase B (PglB) in the ε-proteobacterium *Campylobacter jejuni* demonstrated that certain bacterial species possess a homologous system (Szymanski et al., 1999). PglB is a periplasmic, membrane integrated glycosyltransferase that adopts a so-called GT-C fold that is characterized by an α-helical transmembrane domain and an alternating α/β periplasmic domain (Kikuchi et al., 2003; Liu and Mushegian, 2003; Lizak et al., 2011, 2013). It is structurally similar to the STT3 subunit of the eukaryotic OST complex (Lizak et al., 2011; Wild et al., 2018) and mediates the en bloc transfer of a preassembled heptasaccharide from a lipid-linked oligosaccharide (LLO) onto more than 60 different extra-cytoplasmic protein targets (Feldman et al., 2005; Naegeli and Aebi, 2015) (Figure 2A). This broad
acceptor spectrum results from the target sequon D/E-
X1-N-X2-S/T (where X represents any amino acid except
proline; Marshall, 1972) which constitutes an extension of
the more general eukaryotic OST sequon N-X-S/T (Kowarik
et al., 2006; Schwarz et al., 2011; Schäffer and Messner,
2017). An alternative route for the biosynthesis of aspar-
agine-linked protein glycosylation was identified in the
γ-proteobacterium Haemophilus influenzae (Grass et al.,
2003, 2010). In contrast to the general pathway for N-gl-
cyosylation, the cytosolic glycosyltransferase HMW1C (for
high molecular weight protein 1C) modifies its sole protein
target, the adhesin HMW1, using activated nucleotide
sugars as donor substrates (Figure 2B). This mode of
action represents an unusual mechanism for N-linked
glycosylation that has so far only been observed in bac-
teria. HMW1C is capable of mono- or sequentially digly-
cyosylating HMW1 at 31 distinct sites, 30 of which exhibit
the eukaryote-like NX(S/T) sequon. The modified glycoprotein
is subsequently translocated into the periplasm by the Sec (white) secretory pathway. From there it is excreted and concomitantly tethered
to the outer membrane by HMW1B (white). (C) Modification and inactivation of the FAS associated death domain protein (FADD, gray)
TNFR1 associated death domain proteins (TRADD, gray) and the receptor-interacting serine/threonine-protein kinase 1 (RIPK1, gray)
by the effector glycosyltransferases (yellow) NleB1/2 from enteropathogenic Escherichia coli (EPEC) and SseK1/2/3 from Salmonella.
The effector glycosyltransferases are injected into the host cell via Type III secretion (white). In the host, the effector transfers a single
N-acetylgalactosamin moiety from UDP-N-acetylgalactosamin onto an arginine. This residue is localized within a highly conserved tryptophan-
arginine (WR) motif on the death domain of the target protein. (D) Modification and activation of the specialized translation elongation
donor (TDP-β-L-rhamnose to modify EF-P (gray) at a highly conserved arginine (R) in an unstructured acceptor loop region.
Recognition and binding of the acceptor substrate is both sequence- and structure dependent.

Figure 2: Bacterial pathways for asparagine and arginine linked N-glycosylation.
(A) Block transfer of the heptasaccharide [GalNAc,GlcGalNAc,diNAcBac] from undecaprenyl pyrophosphate (black zigzag line and orange
sphere) to the target sequon (D/E)X1NX2(S/T) in C. jejuni and Campylobacter lari. The lipid-linked oligosaccharide is translocated across
the cytoplasmic membrane by PglK and subsequently used as donor substrate by the membrane bound OST PglB (yellow) that mediates
the transfer to the acceptor asparagine. (B) Sequential transfer of monosaccharide moieties in H. influenzae. The high molecular weight adhesin
glycosyltransferase 1C (HMW1C, yellow) utilizes the nucleotide sugar donors UDP-glucose and UDP-galactose to modify its target protein
HMW1 (gray) at 31 distinct positions. Thirty of these modification sites exhibit the eukaryote-like NX(S/T) sequon. The modified glycoprotein
is subsequently translocated into the periplasm by the Sec (white) secretory pathway. From there it is excreted and concomitantly tethered
to the outer membrane by HMW1B (white). (C) Modification and inactivation of the FAS associated death domain protein (FADD, gray)
TNFR1 associated death domain proteins (TRADD, gray) and the receptor-interacting serine/threonine-protein kinase 1 (RIPK1, gray)
by the effector glycosyltransferases (yellow) NleB1/2 from enteropathogenic Escherichia coli (EPEC) and SseK1/2/3 from Salmonella.
The effector glycosyltransferases are injected into the host cell via Type III secretion (white). In the host, the effector transfers a single
N-acetylgalactosamin moiety from UDP-N-acetylgalactosamin onto an arginine. This residue is localized within a highly conserved tryptophan-
arginine (WR) motif on the death domain of the target protein. (D) Modification and activation of the specialized translation elongation
donor (TDP-β-L-rhamnose to modify EF-P (gray) at a highly conserved arginine (R) in an unstructured acceptor loop region.
Recognition and binding of the acceptor substrate is both sequence- and structure dependent.
Thus, the protein products of both bacterial systems PglB and HMW1C are exclusively outward facing with implications for cell-cell interaction, pathogenicity and virulence. Moreover, the carbohydrate is exclusively coupled to the carboxamide group of asparagine. A first indication for a broader acceptor spectrum of protein N-glycosylation was the discovery of a putative arginine auto-glucosylation of sweetcorn amylogenin (Singh et al., 1995). Similarly, the UDP-arabinopyranose mutase of rice might be self-glucosylated at arginine (Konishi et al., 2010). However, in both cases, the data stem from mass spectrometry analyses with no further experimental corroboration. Accordingly, the catalytic mechanisms for side chain activation and transfer remain obscure.

Two independent research groups simultaneously published the first unambiguous data of enzyme-mediated arginine glycosylation in 2013 (Li et al., 2013; Pearson et al., 2013). The authors of these studies demonstrated that the entero pathogenic Escherichia Coli (EPEC) protein NleB1 is a so-called effector glucosyltransferase. After injection into the host cell via type III secretion, the translocated enzyme disrupts death receptor signaling by adding a single N-acetylglucosamine (GlcNAc) moiety to conserved arginine residues in the death domains (DD) of the tumor necrosis factor receptor 1 (TNFR1) and its downstream adaptors (Figure 2C). Specifically, glycosylation of Arg235 of the TNFR1-associated death domain protein (TRADD) and Arg117 of the FAS-associated death domain protein (FADD) were shown to inhibit DD interactions (Li et al., 2013; Pearson et al., 2013).

Fold recognition models of NleB1 suggested that the protein is a GT-A type glucosyltransferase (Gao et al., 2013). The GT-A fold is characterized by a central open twisted β-sheet that is formed by two closely associated alternating β/α/β domains (Coutinho et al., 2003). This notion was further supported by the finding that effector function was abolished upon substitution of a conserved D-X-D motif (Li et al., 2013; Pearson et al., 2013). This motif is a conserved and functionally important feature of GT-A enzymes (Lairson et al., 2008). The dyad of two negatively charged residues is responsible for coordinating a divalent cation. This in turn neutralizes the developing negative charge on the phosphoryl group upon transfer of a carbohydrate from an activated nucleotide sugar donor (Breton et al., 1998; Breton and Imberty, 1999; Hu and Walker, 2002). Extensive mutational analysis of NleB1 further identified the conserved residues Tyr219 and Glu253 to be catalytically relevant (Wong Fok Lung et al., 2016). Crystallographic analyses of the NleB1 paralog NleB2 (Park et al., 2018) and the orthologous proteins in Salmonella enterica SseK1, SseK2 and SseK3 (Esposito et al., 2018) subsequently confirmed that the glycosyltransferases indeed exhibit a GT-A fold and contain a conserved D-X-D motif responsible for binding divalent cations (Esposito et al., 2018; Park et al., 2018). These structures also revealed a conserved tyrosine (corresponding to Tyr219 in NleB1) being involved in coordinating the uracil moiety of the nucleotide sugar donor substrate (Esposito et al., 2018). In silico docking of arginine to the SseK2 structure suggests close contacts between the acceptor guanidinium and a glutamate (corresponding to Glu253 in NleB1). Together with a neighboring asparagine and a proximal histidine this residue forms the highly conserved HEN motif that appears to be important for catalysis. While its functional importance is experimentally validated, the exact mechanistic role remains to be determined (Park et al., 2018). Interestingly, NleB-like arginine glycosyltransferases modify distinct sets of target DD (El Qaidi et al., 2017) although all of them rely on the presence of a conserved tryptophan-arginine pair in the protein acceptor (Park et al., 2018). Molecular dynamics analyses detected significant differences in the flexibility of the helix-loop-helix domains of SseK1 and SseK2, indicating that dynamics within this domain might be important for target selectivity (Park et al., 2018).

Donor substrate hydrolysis in the absence of the acceptor substrate (Esposito et al., 2018) as well as nuclear magnetic resonance (NMR) analysis of arginine N-acetylgalactosaminylated glycerdehyde 3-phosphate dehydrogenase (Park et al., 2018) suggest that the glycosyl transfer reaction results in retention of the anomeric configuration. Based on this observation a S_{2,1}-like double displacement mechanism was initially suggested (Esposito et al., 2018). As this mechanism depends on the presence of a suitable catalytic base that could not be conclusively identified, a S_{2,2}-like mechanism was proposed instead (Park et al., 2018).

The previously discussed findings about the stereochemical outcome and thus the catalytic mechanism by which NleB mediates the glycosyl transfer reaction was recently challenged (Ding et al., 2019). In their study, Ding et al. reported the crystal structures of the N-acetylgalactosaminylated TRADD and RIPK1 death domains (Ding et al., 2019). Supported by NMR analyses a β-anomeric configuration of the sugar linkage on the Nε atom of the modified arginine guanidinium was found. These data contrasts previous findings (Esposito et al., 2018; Park et al., 2018) as it demands for inverting glycosyl transfer (Ding et al., 2019). The reaction is likely to occur via a S_{2,2}-like direct displacement mechanism. As the side chain of Glu253 forms bidentate interactions with the acceptor arginine this amino acid was proposed to act as base catalyst that deprotonates and thereby activates the guanidinium. This suggestion is in line with the previously described critical
role of Glu253 for NleB function (Esposito et al., 2018; Park et al., 2018). However, catalytic deprotonation of the basic guanidinium by a glutamate side chain appears to be challenging under physiological conditions. Accordingly, the authors suggest that the unique architecture of the catalytic site furthers the deprotonation of the guanidinium. Here, the side chains of His182, His281, Tyr273, Tyr284 and Trp329 of NleB closely surround the acceptor amino acid, thereby possibly lowering the otherwise very high pH of arginine and facilitating the formation of an oxocarbonium-like transition state that initiates the $S_n2$-like direct displacement reaction.

Structural analysis of the NleB/FADD-DD complex further identified prerequisites for death-domain selectivity of the glycosyltransferase (Ding et al., 2019). Enzyme-substrate binding is mediated by two main binding interfaces. The first one is represented by the αB/αC face of FADD-DD that directly contacts alpha helix α9 of NleB. Here, the negatively charged amino acids Asp123, Asp127 and Glu130 of the DD form a strong hydrogen-bond network with the positively charged side chains of Lys292, Lys289 and Lys277 of NleB. Another hydrogen-bond network with inverted distribution of charge is formed between Arg113, Arg140 and Trp112 of the FADD-DD and Asp285 and Asp279 of NleB. The second binding interface is established by the C-terminal part of FADD-DD and the α3/α4 insertion of NleB (Ding et al., 2019). Primary sequence analyses of NleB sensitive and NleB resistant death domain proteins revealed that charged residues within the binding interface of these proteins were either conserved or substituted by side chains of opposite charge, respectively (Ding et al., 2019). After substituting Lys369 of the NleB resistant death domain protein DR5 by Asp and thereby achieving an inversion of the charge on the interface, Ding et al. were able to glycosylate DR5 in an NleB-dependent manner. Likewise, an NleB variant Lys289Asp/Lys292Gln was capable of modifying the wildtype DR5-DD (Ding et al., 2019). Taken together, these results demonstrate the importance of charge distribution within the binding interface of death domain targeting glycosyltransferases and their corresponding protein targets and specifically highlight the functional relevance of Lys289 and Lys292 of NleB in acceptor binding (Ding et al., 2019).

A second case of arginine glycosylation was reported to take place on the bacterial translation elongation factor P (EF-P) (Lassak et al., 2015). EF-P is needed to overcome ribosome stalling during the synthesis of consecutive prolines (Peil et al., 2013; Ude et al., 2013; Starosta et al., 2014c; Lassak et al., 2016). The factor binds to the stalled ribosome and stimulates peptidyl transferase activity, thus allowing translation to continue. Activation requires post-translational modification of a conserved positively charged residue to position and orient the P-site tRNA in a way that is optimal for proline-proline peptide bond formation (Lassak et al., 2015; Huter et al., 2017). In about 25% of all bacteria, including *E. coli*, EF-P is modified by β-lysylation and hydroxylation of a conserved lysine (Bailly and de Crecy-Lagard, 2010; Navarre et al., 2010; Šumida et al., 2010). Analogously a position equivalent arginine in the EF-Ps of another 10% of prokaryotes is glycosylated to ultimately fulfill the same function (Lassak et al., 2015). Specifically, mass-spectrometry and mutational analyses revealed that a rhamnose moiety is covalently linked to this amino acid side chain. A protein with (at that time) a domain of unknown function DUF2331 was identified as the modification system. It was therefore renamed into EF-P arginine 32 rhamnosyltransferase essential for post-translational activation (EarP) (Lassak et al., 2015; Rajkovic et al., 2015; Yanagisawa et al., 2016). *In vitro* experiments with purified protein variants and biochemically synthesized TDP-β-L-rhamnose (TDP-Rha) unambiguously demonstrated that EarP is a glycosyltransferase that mono-rhamnosylates its protein acceptor EF-P (Lassak et al., 2015) (Figure 2D).

Bio synthesis of this modification is markedly different compared to previously described mechanisms of N-glycosylation (Kawai et al., 2011; Lizak et al., 2011; Lassak et al., 2015) although the mode of transfer is somewhat reminiscent of the HMW1C pathway, as it also involves cytoplasmic, direct transfer of a single sugar moiety onto the acceptor from an activated nucleotide sugar donor (Grass et al., 2010; Nothaft and Szymanski, 2010; Kawai et al., 2011). However, rhamnosylated and thereby activated EF-P, is not exported but instead fulfills its molecular function within the bacterial cytosol (Lassak et al., 2015) (Figure 3D). Thus, the elongation factor is to-date the only known example of a bacterial cytoplasmic glycoprotein (Figure 2).

Employing NMR spectroscopy Li and Krafczyk et al. showed that the transfer reaction results in the formation of α-rhamnosylarginine, thus making EarP an inverting glycosyltransferase (Li et al., 2016). This finding triggered the synthesis of mono-α-rhamnosylarginine containing peptides that were used to raise polyclonal antibodies capable of detecting arginine-rhamnosylated proteins (Li et al., 2016; Krafczyk et al., 2017). Using alternative strategies for the anomer-specific glycopeptide synthesis, Wang and colleagues subsequently confirmed the anomeric configuration of rhamnose on EF-P (Wang et al., 2017). This experiment was based on chemically synthesized peptides that correspond to the *Pseudomonas aeruginosa* EF-P acceptor loop carrying either mono-α- or β-rhamnosylarginine: by
comparing nano-UHPLC elution profiles, the α-anomeric configuration and therefore the inverting mode of transfer was confirmed (Wang et al., 2017). In a crystallographic and biochemical study on EarP from *Pseudomonas putida* (EarP\textsubscript{Ppu}) Krafczyk et al. were able to show that EarP consists of two opposing Rossmann-like domains and thus adopts a GT-B fold (Coutinho et al., 2003; Krafczyk et al., 2017). The binding site for the donor substrate TDP-Rha was identified in the protein C-domain. The mode of binding is similar to those found in other GT-B glycosyltransferases and mediated by several conserved amino acids (Hu et al., 2003; Kawai et al., 2011; Krafczyk et al., 2017). NMR titration experiments and subsequent C-terminal truncation of the protein acceptor EF-P revealed that the KOW-like EF-P-N-domain is sufficient for rhamnosylation by EarP. Structural comparison of the binding site architecture of EarP and the inverting asparagine-OST PglB (Lizak et al., 2011, 2013) together with bioinformatic analyses suggested that two invariant aspartate residues (Asp13 and Asp17) are important for catalysis. This idea was further supported by mutant analyses revealing that substitution of any of these residues by alanine results in enzyme inactivation (Krafczyk et al., 2017). In 2018, Sengoku et al. reported the crystal structure of EarP from *Neisseria meningitidis* (EarP\textsubscript{Nm}) in complex with its protein acceptor EF-P\textsubscript{Nm} (Sengoku et al., 2018). These data demonstrated that the EarP N-domain is dedicated with binding of the protein acceptor and implies that acceptor binding is mediated by both structure and sequence elements. This idea is supported by the fact that EarP is capable of rhamnosylating a variant of the non-cognate EF-P orthologue from *E. coli* in which lysine was substituted by arginine (Volkwein et al., 2019). Sengoku et al. further corroborated the functional importance of the negatively charged dyad (Asp16 and Asp20 in EarP\textsubscript{Nm} corresponding to Asp13 and Asp17 in EarP\textsubscript{Ppu}) by showing its direct interaction with the acceptor guanidinium via salt bridge formation (Sengoku et al., 2018). Structural alignment of the EarP\textsubscript{Nm}-EF-P\textsubscript{Nm} complex with the EarP\textsubscript{Ppu}-TDP-β-L-rhamnose complex further revealed that the donor substrate is initially bound in a conformation that is unfavorable for glycosyl transfer. Thus, EF-P binding might induce a structural rearrangement of EarP loop1 that in turn forces the donor substrate into an active conformation and thereby allows the inverting transfer of the rhamnose moiety to occur (Sengoku et al., 2018). Recently a crystallographic analysis of EarP from *P. aeruginosa* in its apo form as well as in binary complex with either TDP or TDP-Rha and in ternary complex with TDP and EF-P from *P. aeruginosa* was published (He et al., 2019). This extensive structural study confirmed both the functional importance of the aspartate dyad (Krafczyk et al., 2017; Sengoku 2018).
et al., 2018) and the necessity for rearrangement of the TDP-Rha conformation as a prerequisite for the glycosyl transfer reaction (Sengoku et al., 2018).

Inverting glycosyl transfer reactions are generally considered to occur via a $S_{N}2$-like reaction mechanism (Lairson et al., 2008). The O-linked glycosylation of serine and threonine residues is initiated by a catalytic base that increases the nucleophilicity of the acceptor oxygen. This reaction facilitates the nucleophilic attack onto the anomeric center. Due to its extensive hydrogen-bond interaction with the acceptor guanidinium, Asp20 of EarP$_{Nm}$ was suggested as the general base for activation of the EF-P arginine (Sengoku et al., 2018).

An alternative mode of activation for inverting N-linked glycosylation, the twisted amide mechanism was proposed for PglB (Lizak et al., 2013). Similar to the activation of arginine by EarP, PglB depends on the presence of two negatively charged residues in close proximity of the accepting amino acid (Lizak et al., 2011, 2013; Krafczyk et al., 2017; Sengoku et al., 2018). As the conjugation of the nitrogen $\pi$ electron represents a major problem in the nucleophilicity of both asparagine and arginine, a similar mechanism might be employed by inverting arginine glycosyltransferases. Interestingly, activation of arginine during methyl transfer by the protein arginine methyltransferase 1 (PRMT1) is also achieved by two negatively charged sidechains (Zhang et al., 2013). Here, acid/base activation was initially suggested as well (Fuhrmann et al., 2015a). However, it was later shown that the hydrogen atom on the modified amino group is lost directly after the transfer reaction, thereby dismissing this mode of activation (Rust et al., 2011). Instead, these interactions abolish the planarity of the guanidinium and localize the positive charge to one of the $\omega$-nitrogen atoms. This leaves a lone electron pair on the other $\omega$-nitrogen atom, which is thereby primed for the nucleophilic attack (Zhang et al., 2013). The orientation of the catalytically important amino acids relative to the acceptor arginine in PRMT1 is highly similar to that in EarP (Figure 4), suggesting a common mode of substrate activation. Specifically, localization and hydrogen bond formation of Asp16, Asp20 and Tyr288 in EarP of *N. meningitidis* are essentially mimicked by Glu144, Glu153 and Tyr35 in PRMT1 (Zhang et al., 2013; Fuhrmann et al., 2015a,b; Sengoku et al., 2018). While this does not generally rule out activation by an acid-base mechanism, the similar architecture of the corresponding binding pockets indicates, that both ways of arginine activation are plausible. Future studies will unveil which mechanism is responsible for the activation of arginine by inverting glycosyltransferases.

The extensive structural and biochemical studies on EarP have turned the rhamnosyltransferase from a curiosity into one of the most thoroughly characterized GT-B glycosyltransferases (Figure 3A–C) (Krafczyk et al., 2017; Sengoku et al., 2018; He et al., 2019). Together with the arginine glycosyltransferase NleB1 and its orthologs, this enzyme represents several new concepts in N-linked glycosylation. Most prominent is the observation that N-glycosylation is not limited to asparagine but can also occur on arginine residues (Pearson et al., 2013; Lassak et al., 2015). Second, the activating (Lassak et al., 2015; Krafczyk et al., 2017) and deactivating (Gao et al., 2013; Pearson et al., 2013) capabilities of arginine N-glycosylation expand the functional potential of this modification. Lastly, the occurrence of arginine glycoproteins in the cytosol might be indicative that this cellular compartment is an especially unappreciated place for this kind of modification. In summary, the occurrence of the

![Figure 4: Architecture of the PglB, EarP and PRMT1 catalytic sites.](image-url)

Amino acids involved in forming the active site are indicated by single letter code specifiers. Hydrogen bonds are indicated by dashed lines. Hydrogen bonding donors of the acceptor substrates are highlighted by a red background. Hydrogen bonding acceptors of the modification enzymes are highlighted by a green background. (A) PglB of *C. jejuni* as described in (Lizak et al., 2013). The position of the divalent metal ion ($M^{2+}$) is indicated by a gray sphere. The amido group of the acceptor asparagine is highlighted by a gray backdrop. (B) EarP of *N. meningitidis* as described in Sengoku et al. (2018). The guanidinium of the acceptor arginine is highlighted by a gray backdrop. (C) PRMT1 of *Rattus norvegicus* as described in Zhang and Cheng (2003). The guanidinium of the acceptor arginine is highlighted by a gray backdrop.
two very distinct arginine targeting glycosyltransferases both of which challenge common principles in glycobiology prompt to the future discovery of a multitude of new glycoproteins (Krafczyk et al., 2017; Sengoku et al., 2018).

Pedal to the methyl – increasing evidence for protein arginine methylation in bacteria

Protein methylation is an abundant post-translational modification that was initially discovered on a flagellar protein of the Gram-negative enterobacterium Salmonella typhimurium (Ambler and Rees, 1959). The most commonly identified target sites for protein methyl transferases (PMT) are the amino groups of lysine [protein lysin (K) methyl transferase, PKMT] and the terminal \( \omega \)-nitrogen atoms of arginine [protein arginine (R) methyl transferase, PRMT] (Clarke, 2013; Alban et al., 2014). Both PKMTs and PRMTs use S-adenosylmethionine as a methyl group donor (Kim and Paik, 1965; Bauerle et al., 2015). Depending on the corresponding modification enzyme, methylation of arginine results in the formation of monomethylarginine, symmetric dimethylarginine or asymmetric dimethylarginine (Fuhrmann et al., 2015a; Murn and Shi, 2017) (Figure 5). Accordingly, PRMTs are classified with respect to the preferential outcome of the reaction. In eukaryotes, arginine methylation is involved in a variety of cellular processes such as the regulation of chromatin structure and transcription (Bird and Wolffe, 1999). Other functions include the alteration of both protein-protein interactions and protein activity (Raposo and Piller, 2018). In HeLa cells up to 50% of all methylations are found on arginine sidechains according to a mass spectrometric analysis. This is likely due to the important functional role that arginine methylation plays during regulation of transcriptional and post-transcriptional processes in eukaryotes (Zhang et al., 2018). Arginine methylation is especially prevalent within the nucleus. Around 2% of rat liver nuclei proteins are methylated (Boffa et al., 1977). The special role of arginine methylation on histones has been extensively reviewed in 2015 (Fuhrmann et al., 2015a,b).

In prokaryotes, protein methylation is involved in diverse cellular functions, including chemotaxis (Levit and Stock, 2002). The reversible methylation of sensory proteins allows the bacterial response to attractants and repellents. However, the necessary target amino acid sidechains do not include arginine and until recently, there was no clear overall evidence on the presence of PRMTs and arginine methylation in bacteria (Levit and Stock, 2002; Shahul Hameed et al., 2018; Zhang et al., 2018).

In 2018, a mass spectrometry-based proteomic analysis by Zhang and colleagues found methylation of Arg60 on the translocation and assembly module subunit TamA (Zhang et al., 2018). While the authors of this study did not identify a corresponding PRMT, indications on the existence of such an enzyme in bacteria was provided by others (Shahul Hameed et al., 2018). The Mitochondrial Dysfunction protein A (MidA), a PRMT from Dictyostelium discoideum shows structural similarities to the putative protein Q6N1P6 (PDB: 1ZKD) of Rhodopseudomonas palustris and two other hypotheticals (PDB: 4F3N, 4G67) (Baugh et al., 2013) from Burkholderia thailandensis, the latter of which was annotated as likely methyltransferase. Hameed et al. therefore concluded that the so far uncharacterized proteobacterial proteins might be PRMTs (Shahul Hameed et al., 2018). Collectively these findings provide increasing evidence for protein arginine methylation in bacteria. Future endeavors might therefore identify further arginine methylation sites in prokaryotes and reveal the full

Figure 5: Chemical diversity of arginine methylation. Hydrogen-bonding donor sites of arginine, monomethylated arginine (MMA), asymmetrical dimethylated arginine (ADMA) and symmetric dimethylated arginine (SDMA) are highlighted by a red background. Methyl groups are highlighted by a gray background.
regulatory and functional potential of this post-translational modification.

**Murder and suicide with bacterial ADP-ribosylation**

ADP-ribosylation is a reversible post-translational modification in which one or more ADP-ribosyl moieties (mono-/poly-ADP-ribosylation) are added to a target protein (Ueda and Hayaishi, 1985). The reaction is catalyzed by ADP-RibosylTransferases (ARTs). ARTs transfer the ADP-ribose moiety to an acceptor residue, which can be glutamate, aspartate, lysine and arginine. The ADP-ribose is bound to negatively charged amino acids via ester bonds (glutamate, aspartate) or to positively charged amino acids via ketamine bonds (lysine, arginine) (Adamietz and Hilz, 1976; Moss et al., 1983; Altmeyer et al., 2009; Hottiger, 2011). The best studied mono-ADP-ribosylations are those that are catalyzed by bacterial toxins. Most of these enzymes modify a specific arginine residue in one single protein target (Corda and Di Girolamo, 2003; Cohen and Chang, 2018) thus promoting bacterial pathogenesis (Deng and Barbieri, 2008; Berthold et al., 2009; Simon et al., 2014) (Table 1).

One example is arginine (Arg177) of human globular actin (G-actin), which is recognized by several bacterial toxins. ADP-ribosylated G-actin binds to the barbed ends of filamentous actin (F-actin), where it acts as a capping protein and prevents the addition of further unmodified G-actin molecules. Thus, actin polymerization is inhibited and the cytoskeleton is destroyed (Aktories et al., 1986; Wegner and Aktories, 1988; Margarit et al., 2006). ADP-ribosylation of G-actin is catalyzed by *Bacillus cereus* VIP2, *Clostridium botulinum* C2, *Clostridium perfringens* iota toxin, *Clostridium difficile* CDT, *Clostridium spiroforme* CST and *S. enterica* SpVB (Aktories et al., 1986; Vandekerckhove et al., 1987, 1988; Popoff and Boquet, 1988; Simpson et al., 1989; Perelle et al., 1997; Han et al., 1999; Otto et al., 2000; Hochmann et al., 2006). It is assumed that targeting of the actin cytoskeleton promotes bacterial invasion and intercellular spread (Lesnick et al., 2001; Guiney and Lesnick, 2005). All of these toxins have a binary structure, consisting of two non-linked proteins; one component with ART activity and another, which mediates binding and uptake of the ART component (Barth and Aktories, 2011).

In *P. aeruginosa* exotoxin S (ExoS) was identified as mono-ART and is an example of a promiscuous ART. It shows less substrate specificity than other bacterial toxins and ribosylates the filament protein vimentin as well as Ras and several Ras related proteins (Rab3, Rab4, Ras, Rap1A and Rap2) (Coburn et al., 1989, Coburn and Gill, 1991). Ras itself is modified at two target sites (Arg41 and Arg128) (Ganesan et al., 1998). Moreover, ExoS is able to ribosylate a third arginine residue (Arg135) when Arg41 and Arg128 are mutated to lysine (Ganesan et al., 1999). Nevertheless, ExoS modifies all of its substrates at arginine residues (Coburn et al., 1989).

ExoS is injected into the cytosol of eukaryotic cells via type III secretion system of *P. aeruginosa*. The enzyme is a bifunctional cytotoxin possessing an N-terminal RhoGAP domain and a C-terminal ADP-ribosyltransferase domain (Barbieri and Sun, 2004). The inactivation of Rho GTPases leads to reorganization of the actin cytoskeleton, while the ADP-ribosylation strongly affects Ras and Rap signaling (Pederson et al., 1999; Riese et al., 2001). As ExoS shares several properties with vertebrate ARTs (substrate promiscuity, ADP-ribosylation at multiple sites, similarities of the catalytic domain), it was postulated that ExoS represents an evolutionary link between bacterial and human ARTs.

<table>
<thead>
<tr>
<th>Bacterial strain</th>
<th>Enzyme</th>
<th>Target(s) (residue)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Bacillus cereus</em></td>
<td>VIP2</td>
<td>G-actin (Arg177) (Han et al., 1999)</td>
</tr>
<tr>
<td><em>Bacillus sphaericus</em></td>
<td>MTX</td>
<td>EF-Tu (Schirmer et al., 2002)</td>
</tr>
<tr>
<td><em>Clostridium botulinum</em></td>
<td>C2</td>
<td>G-actin (Arg177) (Aktories et al., 1986; Vandekerckhove et al., 1988)</td>
</tr>
<tr>
<td><em>Clostridium perfringens</em></td>
<td>Iota toxin</td>
<td>G-actin (Arg177) (Vandekerckhove et al., 1987)</td>
</tr>
<tr>
<td><em>Clostridium difficile</em></td>
<td>CDT</td>
<td>G-actin (Arg177) (Perelle et al., 1997)</td>
</tr>
<tr>
<td><em>Clostridium spiroforme</em></td>
<td>CST</td>
<td>G-actin (Arg177) (Popoff and Boquet, 1988; Simpson et al., 1989)</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td>LT1/LT2</td>
<td>Gα&lt;sub&gt;3&lt;/sub&gt;, Gα&lt;sub&gt;t&lt;/sub&gt; (Arg187) (Moss and Richardson, 1978)</td>
</tr>
<tr>
<td><em>Pseudomonas aeruginosa</em></td>
<td>ExoS</td>
<td>Ras family (Arg41/Arg128) (Coburn et al., 1989; Coburn and Gill, 1991), vimentin (Coburn et al., 1989)</td>
</tr>
<tr>
<td><em>Pseudomonas aeruginosa</em></td>
<td>ExoT</td>
<td>Crk (Arg20) (Sun and Barbieri, 2003; Deng et al., 2005)</td>
</tr>
<tr>
<td><em>Salmonella enterica</em></td>
<td>SpVB</td>
<td>Actin (Arg177) (Otto et al., 2000; Hochmann et al., 2006)</td>
</tr>
<tr>
<td><em>Vibrio cholerae</em></td>
<td>CT</td>
<td>Gα&lt;sub&gt;3&lt;/sub&gt;, Gα&lt;sub&gt;t&lt;/sub&gt; (Arg187) (Moss and Vaughan, 1977; Spangler, 1992)</td>
</tr>
<tr>
<td><em>Rhodospirillum rubrum</em></td>
<td>DraT</td>
<td>Dinitrogen reductase (Arg101) (Pope et al., 1985)</td>
</tr>
</tbody>
</table>
vertebrate mono-ARTs (Ganesan et al., 1999; Koch-Nolte et al., 2008).

The cholera toxin (CT) of *Vibrio cholerae* consists of a catalytic A-subunit (green), which is non-covalently associated with five B-subunits (gray). The B subunits bind to the ganglioside GM1 (gray) at the cell surface of intestinal epithelial cells, enabling the entry of the A subunit, which possesses ART activity. The substrate of ART is the protein Gs (gray), which binds GTP and hydrolyzes to GDP through its intrinsic GTPase activity. ADP-ribosylation inhibits GTPase activity of Gs and the enzyme remains in the active GTP-bound state. Active Gs stimulates the enzyme adenylate cyclase (AC, gray) leading to increased intracellular levels of cyclic AMP (cAMP, gray). The activation of the cAMP-regulated chloride channel (CFTR) leads to a loss of chloride in the intestinal lumen. Further ions and water are released from intestinal epithelial cells by osmosis, causing diarrhea.

**Figure 6:** Action of cholera toxin.

The cholera toxin (CT) of *V. cholerae* consists of a catalytic A-subunit (green), which is non-covalently associated with five B-subunits (gray). The B subunits bind to the ganglioside GM1 (gray) at the cell surface of intestinal epithelial cells, enabling the entry of the A subunit, which possesses ART activity. The substrate of ART is the protein Gs (gray), which binds GTP and hydrolyzes to GDP through its intrinsic GTPase activity. ADP-ribosylation inhibits GTPase activity of Gs and the enzyme remains in the active GTP-bound state. Active Gs stimulates the enzyme adenylate cyclase (AC, gray) leading to increased intracellular levels of cyclic AMP (cAMP, gray). The activation of the cAMP-regulated chloride channel (CFTR) leads to a loss of chloride in the intestinal lumen. Further ions and water are released from intestinal epithelial cells by osmosis, causing diarrhea.

The heat-labile enterotoxin (LT) of *E. coli* is acting by a mechanism that is similar to that of CT (Moss and Richardson, 1978). Moreover, both enzymes are homologs and accordingly also share a common structural architecture (Dallas and Falkow, 1980). Interestingly, besides Gs it was reported that CT and LT also mono-ADP-ribosylate the human defensin proteins HNP-1 and HBd1 

**in vitro** (Castagnini et al., 2012). Defensins are arginine-rich enzymes playing an important role in the defense against bacteria, fungi and viruses (Lehrer et al., 1993). The modification of these antimicrobial components by bacterial ADP-ribosylating toxins may provide a so far unappreciated mechanism to facilitate bacterial colonization (Castagnini et al., 2012).

As their bacterial counterparts vertebrate ARTs can – among other amino acids – recognize arginine as an acceptor substrate. Canonically, the eukaryotic enzymes are larger and more complex and act exclusively endogenously. Vertebrate ARTs show a promiscuous activity; they recognize multiple arginine residues on the same protein target, thereby regulating different cellular pathways. Most of them are anchored in the outer leaflet of the plasma membrane (Okazaki and Moss, 1999). ART1, ART5 and ART2.1/2.2 are arginine-specific mono-ARTs in humans and mice. All of them possess the same active site motif, which is also found in the prokaryotic arginine-specific ARTs (Glowacki et al., 2002).

In addition to the vertebrate and bacterial, three viral ARTs (Alt, ModA and ModB) have been described, all of which mono-ribosylate specific arginine residues. These three enzymes are encoded by bacteriophage T4 and help to gain control over the host cell (Corda and Di Girolamo, 2003; Palazzo et al., 2017). They ribosylate host RNA polymerase at Arg265 and S1 ribosomal protein, leading to an increase in viral transcription (Rohrer et al., 1975; Goff, 1984; Tiemann et al., 1999; Depping et al., 2005).

ARTs transfer an ADP-ribose moiety from beta-nicotinamide adenine dinucleotide (β-NAD⁺) to the guanidino group of a peptidyl arginine residue thereby releasing nicotinamide. ARTs bind the β-NAD⁺ in a conformation enabling the nucleasephilic attack of the arginine guanidinium on the β-N-glycosidic bond between nicotinamide and the C1'-atom of the ribose-group. Thereby the configuration is inverted (Margarit et al., 2006; Tsuge et al., 2008). ADP-ribosylation presumably occurs in an S₁ reaction; however, other mechanisms like S₂, are also conceivable (Bazan and Koch-Nolte, 1997; Tsuge et al., 2003, 2008; Tsurumura et al., 2013). In this regard, the *Clostridium perfringens* iota toxin is investigated in molecular detail. Here, an S₁ strain-alleviation mechanism is proposed to occur with two intermediates; an oxocarbenium ion and a cationic intermediate (Tsuge et al., 2008; Tsurumura et al., 2013).
Bacterial ARTs can be divided into two classes: ARTC (cholera toxin-like) and ARTD (diphtheria toxin-like). This classification is based on the three amino acids that are essentially involved in the binding of NAD\(^+\); arginine-serine-glutamate (R-S-E) triad motif for ARTC and histidine-tyrosine-glutamate (H-Y-E) triad motif for ARTD (Hottiger et al., 2010). Arginine-specific ARTs belong to the ARTC subclass and exhibit an extended triad motif: R-S-EXE (Laing et al., 2011). The R-S arginine forms electrostatic interactions with the diphosphate backbone of the NAD\(^+\) donor, while the R-S serine builds hydrogen bonds with nicotinamide ribose. EXE is located within the so-called ADP-ribosyltransferase turn-turn loop (ARTT) and is important for acceptor substrate recognition (Han et al., 2001; Koch-Nolte et al., 2001; Sun et al., 2004). The first glutamate (R-S-E) is suspected to be essential for arginine specificity (Koch-Nolte et al., 2008). It is presumed that the second glutamate (R-S-EXE) has a catalytic role and stabilizes the oxocarbenium intermediate (Barth et al., 1998; Wilde et al., 2002). Notably, all members of the ARTC family only attach one ADP-ribose (Luscher et al., 2018) and consequently all bacterial arginine-specific ARTs are also mono-ARTs.

ADP-ribosylation of arginine alters its chemical properties significantly. The attachment of a double negatively charged ADP-ribose to the positively guanidinium inverts the net charge of the side chain. In addition, there is an increase in molecular weight; an ADP-ribose moiety (540 Da) is about 3 times as heavy as arginine itself (174 Da). These changes lead to altered protein function and protein-protein interactions. In the case of ADP-ribosylated G-actin the steric hindrance of the bulky ADP-ribose prevents

![Diagram](https://example.com/diagram.png)

**Figure 7**: Reversible ADP-ribosylation of Fe protein.

In the presence of ammonium, the Dinitrogenase Reductase ADP-ribosylTransferase (DraT, orange) attaches an ADP-ribose moiety from β-NAD\(^+\) to the Arg101 residue of one subunit of the homodimeric Fe protein (green), whose activity is thereby switched off. The Dinitrogenase Reductase-Activating Glycohydrolase (DraG, orange) removes the ADP-ribose as soon as the ammonium is exhausted and thereby restores the nitrogenase activity.
polymerization into F-actin (Wille et al., 1992; Lesnick et al., 2001; Guiney and Lesnick, 2005; Margarit et al., 2006).

As stated in the beginning ADP-ribosylation is a dynamic PTM and thus can be removed by ADP-ribosylarginine hydrolases (ARHs), thereby releasing ADP-ribose (Ludden, 1994). So far, only a few bacteria have been described to possess genes for potential ARTs as well as ARHs. The enzymatic function of ART and ARH is particularly well studied in *Rhodospirillium rubrum* (Figure 7).

The toxin Dinitrogenase Reductase ADP-ribosyltransferase (DraT) is known to regulate nitrogen fixation by ADP-ribosylation of the dinitrogenase reductase (Fe protein) (Pope et al., 1985). The Fe protein is one of two components of nitrogenase and therefore essential in nitrogen fixation. When fixed nitrogen is present in the environment, the nitrogenase is inactivated to save energy (Huergo et al., 2012). Inactivation occurs by attaching an ADP-ribose to the homodimeric Fe protein (Arg101). This arginine residue is located close to the (4Fe-4S) cluster, which is the docking site for the other component of nitrogenase, the dinitrogenase (MoFe) (Georgiadis et al., 1992). Modification prevents the association of Fe protein and MoFe and interrupts electron transfer (Schindelin et al., 1997; Moure et al., 2015). The dinitrogenase reductase-activating glycohydrolase (DraG) acts as antitoxin that hydrolyzes the ADP-ribosylarginine. These two enzymes form an ADP-ribosylation circuit and thus regulate the central enzyme of nitrogen fixation (Pope et al., 1986; Moure et al., 2015).

Genome analyses showed that DraT homologues occur exclusively in diazotrophic bacteria. In contrast, DraG homologues were found in bacteria as well as in eukaryotes and archaea (Moure et al., 2015). Three DraG homologues (ARH1-3) have been identified in humans. Interestingly, ARH1 exhibits the same enzymatic activity as DraG and is able to remove ADP-ribose from arginine residues (Moss et al., 1988, 1992). Experiments with ARH1-deficient mice showed that their cells were more sensitive to the cytotoxic effects of cholera toxins. This suggests that ARH1 can also reverse the ADP-ribosylation of bacterial ARTs (Kato et al., 2007; Watanabe et al., 2018).

**Sentenced to death – protein arginine phosphorylation as degradation signal in Gram-positive bacteria**

The reversible protein phosphorylation is one of the most widespread post-translational modifications (Khoury et al., 2011) and occurs in a myriad of proteins thus affecting nearly every cellular aspect (Lassak et al., 2010; Binnenkade et al., 2011; Fuhrmann et al., 2015a; Lucero et al., 2019; Song and Luo, 2019). Protein kinases and corresponding protein phosphatases catalyze the synthesis and removal of this modification. Although other types of protein phosphorylations (histidine, aspartate) are reported (Lassak et al., 2013; Jung et al., 2018; Potel et al., 2018) the best-characterized form is the attachment of a phosphoryl group to the side chain hydroxyl group of either serine, threonine or tyrosine (Matthews, 1995; Besant et al., 2009). The resulting phosphomonoesters are accessible to methods involving acidic treatments in standard biochemical, proteomics and antibody production protocols (Ciesla et al., 2011; Suskiewicz and Clausen, 2016). In comparison, protein N-phosphorylation is far less characterized due to the acid labile character of the phosphoramidate formed when the phosphoryl group is attached to the side chain nitrogen of, histidine, lysine or arginine (Klumpp and Kriegstein, 2002; Attwood et al., 2007; Kowalewska et al., 2010; Kee and Muir, 2012; Engholm-Keller and Larsen, 2013; Fuhrmann et al., 2015a). Despite the challenges arising from the acid labile character of the phosphoramidates, there have been sporadic reports on activities of protein arginine kinases in eukaryotes and viruses (Smith et al., 1976; Sikorska and Whitfield, 1982; Wilson and Consigli, 1985; Levy-Favatier et al., 1987; Wakim and Aswad, 1994; Matthews, 1995; Wakim et al., 1995; Besant et al., 2009). However, none of these studies could identify a responsible enzyme, which led to controversial discussions about the existence of arginine phosphorylation in proteins. The controversy was ended by Fuhrmann and co-workers who investigated the bacterial CtsR/McsB heat shock response system in *Bacillus subtilis* and found that arginine residues within the DNA binding domain of the class three stress repressor (CtsR) are specifically phosphorylated (Fuhrmann et al., 2009). This phosphorylation inactivates the protein by introducing negative charges in the DNA binding domain thus interfering with protein-DNA complex formation. This in turn leads to an increased expression of the heat shock genes *clpC* and *clpP* (Derre et al., 1999), which constitute the ClpCP protease complex that mediates the degradation of misfolded proteins. In the same work (Fuhrmann et al., 2009), McsB – which was previously misannotated as a protein tyrosine kinase – was identified as the corresponding protein arginine kinase. This enzyme shows close homology to the catalytic domain of phosphen (guanidino) kinases catalyzing the transfer of the γ-phosphoryl group from ATP to small
molecules bearing a guanidino group as also found in arginine (Ellington, 2001). At the same time McsB lacks the N-terminal domain of phosphagen kinases responsible for their substrate specificity (Kruger et al., 2001; Kirstein and Turgay, 2005; Fuhrmann et al., 2009, 2015a). Conversely, McsB harbors a structurally distinct C-terminal domain. This domain is presumably responsible for the shift of the substrate specificity from free to peptidyl-arginine (Fuhrmann et al., 2009, 2015a). The catalytic mechanism by which McsB phosphorylates the guanidino group is assumed to be similar to that of L-arginine kinases (Figure 8, right). The latter apply a single bimolecular nucleophilic substitution (S_{N2}) reaction, in which the \( \gamma \)-phosphoryl group from ATP is directly transferred to the guanidino group of peptidyl-arginine (Kirstein and Turgay, 2005; Fuhrmann et al., 2009, 2015a). In this regard the correct orientation of the triphosphate is most likely achieved by magnesium (Fuhrmann et al., 2015a). In addition to the function as a CtsR kinase McsB also serves as an adapter protein for ClpC and as such stimulates its ATPase activity thereby promoting the degradation of CstR (Kirstein et al., 2007; Elsholz et al., 2010, 2011). Subsequent studies in *B. subtilis* (Elsholz et al., 2012; Schmidt et al., 2014; Trentini et al., 2016) and *Staphylococcus aureus* (Junker et al., 2018) revealed that McsB not only phosphorylates CstR but in fact hundreds of proteins. These findings demonstrate that McsB mediated arginine phosphorylation has a more general role in stress response and other cellular processes. To regulate protein arginine phosphorylation levels according to the physiological conditions McsB activity is counteracted by the protein arginine phosphatase YwlE (Kirstein and Turgay, 2005; Kirstein et al., 2008; Hahn et al., 2009; Elsholz et al., 2010, 2011, 2012; Fuhrmann et al., 2013). YwlE is a member of the low-molecular-weight protein-tyrosine phosphatases (LMW-PTP) family and catalyzes the hydrolysis of phosphoarginine residues in two consecutive nucleophilic substitutions (S_{N2}) (Fuhrmann et al., 2015a, 2016) (Figure 8, left). In the first reaction, the highly nucleophilic cysteine residue C7 of YwlE attacks the incoming phosphorus atom of the guanidino group, thereby generating a phosphoryl enzyme intermediate. Then, an incoming water molecule is deprotonated by the aspartic acid residue D118, generating a nucleophilic hydroxyl anion which attacks the phosphoryl enzyme intermediate in the second reaction. This attack leads to the release of a phosphate ion on the one hand, and regenerates C7 on the other hand (Fuhrmann et al., 2015a, 2016). Interestingly, C7 is not only essential for catalysis but also provides the cell with a sensor for oxidative stress based on the formation of an intramolecular disulfide bridge between the cysteine residues C7 and the structurally neighboring C14. In this state YwlE is inactive in turn leading to an increase in McsB-mediated peptidyl arginine-phosphorylation and thus activation of the stress response system due to increased levels of phosphorylated CtsR. When encountering reducing conditions the disulfide bridge is resolved and YwlE gets activated (Fuhrmann et al., 2016; Suskiewicz and Clausen, 2016).

Another breakthrough in the field came in 2016 when Trentini, Suskiewicz and colleagues showed that in *B. subtilis* McsB-mediated arginine phosphorylation can also acts as signal for the ClpCP degradation system that allows for recognition of heat damaged misfolded proteins (Trentini et al., 2016; Tripathi and Gottesman, 2016) (Figure 9). In their study, the authors used catalytically inactive variants of the ClpP protease (Flynn et al., 2003; Feng et al., 2013; Trentini et al., 2016) expressed them under heat shock conditions and co-purified trapped proteins. By using this technique in combination with mass spectrometry Trentini et al. determined that about 25% of those were arginine phosphorylated (Trentini et al., 2016). To demonstrate the function of this modification as a degradation tag, the researchers further reconstituted the ClpCP/McsB system in vitro. Thereby, they showed that the ClpCP protease complex recognizes phosphoarginine modified \( \beta \)-casein [an intrinsically unfolded protein (Halwer, 1954)] directly, whereas degradation of the unmodified form requires the simultaneous presence of one of two adaptor proteins, MecA (Schlothauer et al., 2003) or McsB (Kirstein and Turgay, 2005). Further, Trentini et al. demonstrated that only phosphorylated \( \beta \)-casein was able to bind to the N-terminal domain (NTD) of ClpC and thus allowed adaptor free assembly of the ClpCP complex (Trentini et al., 2016). In vitro experiments using \( \beta \)-casein pre-incubated with McsB additionally confirmed that McsB mediated arginine phosphorylation is sufficient to tag \( \beta \)-casein for degradation by the ClpCP complex (Trentini et al., 2016). Ultimately, the co-crystal structure of the NTD of ClpC in complex with phosphoarginine revealed two almost identical binding sites for phosphoarginine in the NTD, which partly overlap with the binding sites for the MecA adapter protein. Taken together, these data provide convincing evidence that arginine phosphorylation is a degradation signal for ClpCP (Trentini et al., 2016). However, it remains puzzling why arginine-phosphorylated CtsR is not recognized by the ClpCP complex in the absence of adaptor proteins (Kirstein et al., 2007; Elsholz et al., 2017). Thus, further research is required to fully understand the role of McsB during heat stress. Along the
Figure 8: Peptidyl-arginine phosphorylation and dephosphorylation.

(Right) Suggested mechanism for the phosphorylation of peptidyl arginine by McsB. The mechanism is exemplified on the mechanism for L-arginine kinases. There, arginine is clamped between two carboxylate groups leading to a polarization of the guanidino group (blue) which is then attacked in a $S_2$ reaction by the electrophilic $\gamma$-phosphoryl group from ATP (red) (1), resulting in the transfer of the $\gamma$-phosphoryl group from ATP to the guanidino group (2). In this mechanism it is assumed that a magnesium ion coordinates the three phosphoryl groups of ATP and thus ensures the correct orientation of the $\gamma$-phosphoryl group for the $S_2$ reaction. (Left) Peptidyl-arginine phosphorylation can be reversed by YwlE mediated hydrolysis of the phosphoramidate (N-P) bond. This mechanism is composed of two consecutive $S_2$ reaction steps. In the first reaction step, the phosphorous atom is attacked by the activated nucleophilic residue of cysteine C7 of YwlE (3), resulting in a thio phosphate intermediate (4). Afterward an incoming water molecule is deprotonated leading to the second $S_2$ attack of the hydroxyl anion on the thio phosphate intermediate (5) which leads to the cleavage of the intermediate, release of the phosphate ion and regeneration of the active nucleophilic C7 residue for further dephosphorylation reactions (6). Figure adopted from Fuhrmann et al. (2015a, 2016).
same lines one might argue that more protein arginine modifications might exist that do not serve as degradation signals. The discovery of those demands for efficient detection methods (Fuhrmann et al., 2009, 2015a; Schmidt et al., 2013). In this regard, the chemical synthesis of phosphoarginine peptides and the generation of acid stable analogs (Fuhrmann et al., 2016; Ouyang et al., 2016) allowed for the generation of modification specific antibodies (Fuhrmann et al., 2013, 2015b; Hauser et al., 2017). These already have and will contribute further to a comprehensive understanding of the peptidyl-arginine phosphorylation in the future.

The beginning of the end – arginine as N-terminal acceptor or donor substrate determines the fate of proteins

More than 50 years ago, a group of researchers discovered that in cell-free extracts amino acids can be incorporated into proteins N-terminally and independent of ribosomes (Kaji et al., 1963a,b; Kaji, 1968). This phenomenon is domain-spanning and does not reconstitute an alternative translational process. It is instead a post-translational modification (Kaji et al., 1965a,b; Kaji and Kaji, 2011; Saha and Kashina, 2011). In eukaryotic cells, on the one hand, the catalysis involves the formation of a peptide bond with the $\alpha$-carboxyl group of arginine as the donor substrate and is accordingly termed arginylation (Soffer and Horinshi, 1969). In bacteria, on the other hand, leucine and phenylalanine (Leibowitz and Soffer, 1969), less frequently also methionine (Scarpulla et al., 1976) and tryptophan (Kaji et al., 1965a,b) are incorporated (Figure 10A). In both cases and similar to translation charged tRNAs serve as activated precursors. By contrast however, peptide bond formation is not mediated by rRNA, but is a protein-based chemical reaction catalyzed by specialized modification enzymes. At the molecular level the reaction is most likely inverse to the peptide hydrolysis mediated by serine proteases and involves an electron relay system that includes two negatively charged residues (Figure 10B) (Watanabe et al., 2007). Whereas the $\alpha$-amino group of arginine, lysine and methionine serve as the main acceptors for the prokaryotic leucyl/phenylalanyl-tRNA protein transferase (LFTRs), the eukaryotic LFTR counterpart, ATE1 (for arginine transfer enzyme 1), preferentially recognizes the one of aspartate and glutamate instead (Kaji and Rao, 1976). Glutamate is also a substrate for another unusual form of arginylation of the human regulatory peptide neurotensin (Eriste et al., 2005), where the $\gamma$-carboxyl group is the acceptor to form the peptide bond. This observation suggests that either further unknown enzymes are involved or that the ATE1-mediated modification is more promiscuous than originally thought (Saha and Kashina, 2011). The latter is plausible as the bacterial LFTR has a relatively large donor and acceptor substrate spectrum. How exactly substrate recognition works has been elucidated crystallographically with E. coli LFTR (Suto et al., 2006; Watanabe et al., 2007). The enzyme has a compact structure consisting of two domains. In particular, the C-terminal domain is catalytically active. Its folding pattern matches the superfamily of GCN5-like N-acetyltransferases (GNAT), specifically the non-ribosomal FemABX-related peptidyltransferases (Hegde and Shrader, 2001; Benson et al., 2002). Via a hydrophobic C-shaped pocket, LFTR recognizes tRNAs charged with nonpolar residues that lack branched $\beta$-carbons (Suto et al., 2006; Watanabe et al., 2007). The cognition of the tRNA also involves the N-terminal domain of LFTR which recognizes particularly the 3’-aminoacyladenosine. In addition to this, neighboring nucleotide sequence elements and the tRNA-D stem also play a role (Fung et al., 2014). On the acceptor side, specifically the positive charges of the arginine side chain are accommodated in a negatively charged pocket of LFTR. In a second pocket,
the immediately following amino acids are predominantly recognized via backbone interactions (Watanabe et al., 2007). However, this alone does not explain how, for example, the binding of incompatible methionine can serve as an acceptor. It is therefore possible that the immediately following amino acid residue compensates for sterically unfavorable effects caused by a negatively charged environment (Dougan et al., 2010). Here, a better understanding of the binding properties of the second pocket could provide new insights into the specificity of LFTR or ATE1 and its homologs. In this context, two aminoacyltransferases with an unusual substrate spectrum are of particular interest (Graciet et al., 2006). One of the two enzymes is the ATE1 ortholog of the malaria pathogen *Plasmodium falciparum*, which arginylates an LFTR identical sequon. Conversely, the bacterial protein transferase Bpt from *Vibrio vulnificus* preferentially links leucine to glutamate and aspartate residues. This diversity of acceptor peptides not only raises questions about the general diversity of LFTR/ATE1 recognition sequences, but also implies species-specific protein target spectra. This is fascinating regarding the function of these enzymes in protein degradation and the application of the so-called N-end rule (Bachmair et al., 1986; Dougan et al., 2010; Fung et al., 2016). This states that the stability of a polypeptide starting with tryptophan, tyrosine, phenylalanine or leucine is recognized by ClpS and recruited to the ClpAP protease complex, to be degraded.
protease complex (Figure 10C) (Wang et al., 2007; Kirstein et al., 2009; Ninnis et al., 2009). Directed intracellular proteolysis is a fundamental process that is conserved in all domains of life. While in the case of the previously discussed arginine phosphorylation, the linear degradation signal is hidden in the protein structure and is only exposed by unfolding stress, the N-terminus is often accessible without such a trigger. This fundamental difference already makes it possible to determine the protein turnover rate in the coding sequence by the preselection of the N-terminal amino acid. For the exposure of a primary destabilizing residue – in *E. coli* these are tryptophan, tyrosine, phenylalanine and leucine (Bachmair et al., 1986) – several routes are discussed. It is well known that protein synthesis in bacteria is generally initiated with formyl methionine, which is accordingly not a primary degradation signal, nor is it recognized as a substrate by LFTR. Thus, a prior post-translational modification is required in order to mark a protein for degradation. One possibility is deformylation catalyzed by peptide deformylase (PDF). The resulting methionine can then be directly labeled with a destabilizing N-terminal residue with the help of LFTR. As an alternative to PDF, an endopeptidase-mediated protein cleavage could also directly produce such a residue, or LFTR acceptor amino acids such as arginine or lysine could be N-terminally liberated for a destabilizing modification. Such a scenario was simulated by the fusion of the Small Ubiquitin-like MOdifier (SUMO) (Hay, 2005) with model proteins (Wang et al., 2007; Starosta et al., 2014b). Through the simultaneous expression of the corresponding SUMO protease, an N-terminus with destabilizing residues can be generated, which in turn opens up its use as a molecular tool for targeted degradation (Starosta et al., 2014b; Sekar et al., 2016). In addition, the LFTR catalyzed peptide bond formation offers an attractive possibility to load the N-terminus with non-natural amino acids (Hamamoto et al., 2011; Kawaguchi et al., 2013) and thus make it usable for synthetic biology. Finally, the integration of tRNA into the proteolysis process leads to an interesting regulatory scenario resulting from the interplay with the translation machinery: under normal conditions, there is competition between LFTR and the ribosome for the donor substrate, which is mainly decided in favor of the latter. However, under translational stress conditions (Starosta et al., 2014a), charged tRNAs may accumulate, the LFTR substrate limitation is removed and more N-degrons are formed (Fung et al., 2016). In particular, the so-called stringent response could play an important physiological role here. Stringent response is an adaptive mechanism triggered by a variety of stressors such as amino acid starvation (Stent and Brenner, 1961; Hauryliuk et al., 2015). GTP cleavage produces the alarmones guanosine pentao- and guanosine tetraphosphate, which in turn inhibit translation (Pesavento and Hengge, 2009). As a result of the LFTR-initiated increased proteolysis activity, more amino acids are now liberated, which in turn eliminate the corresponding deficiency for protein synthesis. The consequent restoration of ribosome activity finally completes the circle (Fung et al., 2016).

### Post-translational toothache – citrullination as bacterial weapon to cause chronic periodontitis

Protein citrullination means the post-translational conversion of arginine into citrulline (Fearon, 1939; Rogers, 1962). The replacement of the imino group by a carbonyl moiety is also referred to as a deimination reaction. Accordingly, the enzymes which accomplish this post-translational conversion are called peptidyl arginine deiminases (PAD) (Rogers and Taylor, 1977; Vossenaar et al., 2003). The PAD catalyzed hydrolysis of the guanidino group has a tremendous influence on the possible hydrogen bonds and electrostatic interactions: the citrulline side chain is no longer positively charged but neutral with having now two acceptors – but only three donor sites for hydrogen bonds. Especially in eukaryotes this conversion plays an important role, for example, in the development of rigid structures like skin, hair and myelin sheaths (Mangat et al., 2010). Accordingly, the modification is important both under physiological normal conditions and during processes of epidermal differentiation (Mechin et al., 2005; Nachat et al., 2005), the maturation of hair follicles (Mechin et al., 2005) and the insulation of nerve fibers. Aberrant citrullination can contribute to skin diseases such as psoriasis and neurological disorders such as multiple sclerosis and Alzheimer’s disease (György et al., 2006). The manifold physiological significance of this post-translational modification in humans is contrasted by one prokaryotic example. The pathogen of chronic periodontitis *Porphyromonas gingivalis*, a representative of Bacteroidetes (Naito et al., 2008), is to date the only bacterium with known functional peptidylarginine deiminase (PPAD) (Figure 11) (McGraw et al., 1999). In the host, PPAD is secreted and inactivates anaphylatoxin C5a – a component of the complement system – by citrullination of a critical C-terminal arginine (Bielecka et al., 2014). In addition, the ammonium released in this
process might promote survival in the host (Mangat et al., 2010). Although the eukaryotic and prokaryotic enzymes catalyze the same reaction, the respective PADs differ significantly regarding their primary structure, the dependence on cofactors as well as the substrate spectrum (Arita et al., 2004; Goulas et al., 2015; Slade et al., 2015; Montgomery et al., 2016). Nevertheless, all PADs belong to the same protein superfamily, the penteins (Shirai et al., 2001; Linsky and Fast, 2010). This relatively heterogeneous group shares a common β/α propeller folding pattern. PADs are further characterized by a catalytic cysteine which establishes a covalent bond with the donor guanidino group and thereby forms a thiouronium intermediate (Linsky and Fast, 2010). While human PADs are activated by Ca²⁺ (Arita et al., 2004), PPAD works independently of the cofactor. Moreover, the bacterial enzyme recognizes terminal arginine as acceptor amino acid (Goulas et al., 2015), while the eukaryotic counterparts act as endodeiminases (Hagiwara et al., 2002). Ultimately, PPAD can also modify free arginine (McGraw et al., 1999). In this context, it is worth mentioning that PPAD is structurally closer to agmatine deiminases (AgDI). Agmatine itself is produced by the decarboxylation of arginine which is deaminated by AgDI to N-carbamoyl putrescine accompanied by the release of ammonium (Shek et al., 2017). Unlike eukaryotic PADs, AgDIs can be easily identified in various bacteria. It is therefore possible that the N-terminal domain of PPAD is evolutionarily derived from an AgDI ancestor (Goulas et al., 2015). The fusion with an approximately 75-amino acid-long C-terminal domain for maturation and translocation through the outer membrane (Sato et al., 2013) finally led to the emergence of this extraordinary virulence factor. Future studies will show whether peptidyl deimination by PPAD remains a prokaryotic single case or whether post-translational citrullination is more common in bacteria than assumed today.

**Friend or foe – advanced glycation end products to poison or feed bacteria**

Glycation or non-enzymatic glycosylation refers to the reaction of amino compounds with reducing sugars (Ulrich and Cerami, 2001). The phenomenon was first described in 1912 by Louis-Camille Maillard as being responsible for the aroma, taste and the appearance of thermally processed food (Maillard, 1912a,b). Proteinogenic glycations normally occur at the amino termini of proteins, the ε-amino groups of lysine residues (Hellwig and Henle, 2014; Hellwig, 2019), but also the guanidino groups of arginine (Zhu and Yaylayan, 2017). Initially, Schiff bases are formed, which spontaneously but reversibly rearrange into more stable Amadori or Heyns products (McGraw et al., 1999). In this context, it is worth mentioning that PPAD is structurally closer to agmatine deaminases (AgDI). Agmatine itself is produced by the decarboxylation of arginine which is deaminated by AgDI to N-carbamoyl putrescine accompanied by the release of ammonium (Shek et al., 2017). Unlike eukaryotic PADs, AgDIs can be easily identified in various bacteria. It is therefore possible that the N-terminal domain of PPAD is evolutionarily derived from an AgDI ancestor (Goulas et al., 2015). The fusion with an approximately 75-amino acid-long C-terminal domain for maturation and translocation through the outer membrane (Sato et al., 2013) finally led to the emergence of this extraordinary virulence factor. Future studies will show whether peptidyl deimination by PPAD remains a prokaryotic single case or whether post-translational citrullination is more common in bacteria than assumed today.

**Friend or foe – advanced glycation end products to poison or feed bacteria**

Glycation or non-enzymatic glycosylation refers to the reaction of amino compounds with reducing sugars (Ulrich and Cerami, 2001). The phenomenon was first described in 1912 by Louis-Camille Maillard as being responsible for the aroma, taste and the appearance of thermally processed food (Maillard, 1912a,b). Proteinogenic glycations normally occur at the amino termini of proteins, the ε-amino groups of lysine residues (Hellwig and Henle, 2014; Hellwig, 2019), but also the guanidino groups of arginine (Zhu and Yaylayan, 2017). Initially, Schiff bases are formed, which spontaneously but reversibly rearrange into more stable Amadori or Heyns products (Heyns and Noack, 1962). A consequence of further parallel and/or sequential reactions (such as oxidation, fragmentation, condensation, reduction, dehydration, isomerization, cyclization, etc.) leads to the formation of irreversible advanced glycation endproducts (AGE) (Hellwig and Henle, 2014; Henning and Glomb, 2016). In addition, reactive α-dicarbonyl compounds such as glyoxal or methylglyoxal can be released from the early glycation products but also from free aldoses and ketoses. These can in turn react with lysine and arginine side chains to form AGEs (Ulrich and Cerami, 2001; Grillo and Colombatto, 2008). A large number of compounds can be derived from arginine conjugates (Figure 12):
incubation with glyoxal leads to the formation of 1-(4-amino-4-carboxybutyl)2-imino-5-oxo-imidazolidine (Glarg) (Schwarzenbolz et al., 1997), which hydrolyzes under physiological conditions to the acid-labile carboxymethylarginine (CMA) (Glomb and Lang, 2001). A reaction with methylglyoxal results mainly in N\textsubscript{δ}-(5-methyl-4-oxo-5-hydroimidazolinone-2-yl)-L-ornithine (MG-H1) (Henle et al., 1994) and to a lesser extent to 2-amino-5-(2-amino-5-hydro-5-methyl-4-imidazolon-1-yl) pentanone (MG-H2) and 2-amino-5-(2-amino-4-hydro-4-methyl-5-imidazolon-1-yl) pentanone (MG-H3) (Ahmed et al., 2002). At room temperature and under basic conditions, MG-H3 is hydrolyzed quantitatively to carboxyethyl-L-arginine (CEA) within days (Gruber and Hofmann, 2005). Furthermore, the imidazolinone products can react with another molecule of methylglyoxal to form the compounds N\textsubscript{δ}-(5-hydroxy-4,6-dimethylpyrimidin-2-yl)-L-ornithine (argpyrimidine) (Shipanova et al., 1997) or alternatively N\textsubscript{δ}-(4-carboxy-4,6-dimethyl-5,6-dihydroxy-1,4,5,6-tetrahydropyrimidin-2-yl)-L-ornithine (tetrahydroargpyrimidine) (Oya et al., 1999). As arginine residues are often of catalytic importance (Lehoux and Mitra, 2000; Lee et al., 2001; Keenholtz et al., 2013), their glycation may have an effect on protein function. The discovery that the Maillard reaction also takes place under physiological conditions in humans, or that its products are ingested via the diet led to the intensive investigation of its pathophysiological significance. AGEs serve as markers for diabetes mellitus (Ahmed and Thornalley, 2007), atherosclerosis (Price and Knight, 2007), uremia (Thornalley and Rabbani, 2009) and aging processes (Ahmed et al., 2003). Maillard products consumed through the diet are discussed as 'glycotoxins' and thus as a nutritional risk, but more and more their positive effects are also recognized (Sebekova and Brouder Sebekova, 2019). After incorporation in the body, AGEs are absorbed relatively quickly and can be detected in the blood within a few hours (Koschinsky et al., 1997). The interaction of such exogenous AGEs with specific receptor proteins leads to an increased inflammatory response and oxidative tissue stress (Goh and Cooper, 2008). On the other hand, a diet enriched with AGEs can reduce glycemic effects (Chuyen et al., 2005). Given their clinical significance, the majority of studies on non-enzymatic glycosylation were conducted primarily on higher eukaryotes. In lower eukaryotes such as baker’s yeast, only the glycation of aspartyl-tRNA synthetase (Colas and Boulanger, 1983) or phosphoglycoisomerase is well documented (Zahner et al., 1989). The literature on protein glycation in bacteria is similarly scarce. This may be due to the belief that the time required to form protein glycation exceeds the lifetime of bacterial proteins. In fact, however, there is evidence that AGEs are formed even under normal growth
conditions (Mironova et al., 2001). AGEs were found both in recombinantly produced human interferon γ and in E. coli total protein extract. There is a direct correlation between the amount of AGEs produced and the volume of the culture, the cultivation time and, in particular, the composition of the medium (Dimitrova et al., 2004; Kram and Finkel, 2014, 2015). Under conditions that lead to the enrichment of glycolytic phosphorylated intermediates and the resulting phosphate limitation, methylglyoxal is formed from dihydroxyacetone phosphate after dephosphorylation by methylglyoxal synthase (Green and Lewis, 1968; Totemeyer et al., 1998). Alternatively, methylglyoxal can also arise from the oxidation of aminoacetone, an intermediate of threonine catabolism (Mathys et al., 2002; Kim et al., 2004). In addition to intrinsic methylglyoxal production, there are also external sources to which bacteria are exposed. The consumption of food containing methylglyoxal (Griffith and Hammond, 1989; Majtan et al., 2012) could promote glycation, especially in intestinal bacteria. While the detoxification of methylglyoxal takes place directly and predominantly through the glyoxalase system (Reiger et al., 2015), the glycation of proteins in E. coli is removed by the two paralogs YhbO and Yajl. (Wilson et al., 2005; Abdallah et al., 2016). These act as deglycascases repairing damage of methylglyoxal and glyoxal glycated cysteine, arginine and lysine residues (Richarme et al., 2015). While there is so far no work on the positive effects of intrinsically occurring glycations, at least the external presence of Maillard products is reported to stimulate the growth of health-promoting bacteria in the intestinal flora (Morales et al., 2012; Richarme et al., 2015). In this context it is worth mentioning that both the model organisms E. coli (Wiaume et al., 2002; Hellwig et al., 2019) and B. subtilis (Wiaume et al., 2004) as well as pathogens such as S. enterica (Ali et al., 2014) can use Amadori sugars as the sole carbon source. Future studies will show to what extent endogenous glycation in general and arginine glycation in particular affect bacterial physiology. If one speculates at this point, it would be plausible that such modifications could influence especially bacterial persistence.

To be continued with protein arginine ac(etyl)ylation and hydroxylation

The landscape of arginine modifications is further growing and encompasses two additional chemical changes namely acylation and hydroxylation.

Acetylation

Protein acylation refers to the attachment of an acyl group such as acetyl (Weinert et al., 2013; Kuhn et al., 2014; Volkwein et al., 2017), propionyl (Sun et al., 2016), succinyl (Weinert et al., 2013) or malonyl (Qian et al., 2016). These PTMs can be a result of a spontaneous reaction with high energy phosphodonsors (Wagner et al., 2017) but are also facilitated by enzymes (Christensen et al., 2018). In this regard, acetyltransferases catalyze the transfer of an acetyl group from an acetyl coenzyme A (acetyl CoA) donor (Christensen et al., 2019). Primarily the acyl moiety is accepted by the ε-amino group of lysine or alternatively an N-terminal α-amino group of various amino acids including arginine (Ouidir et al., 2015; Jedlicka et al., 2018). In contrast to eukaryotes, in which the majority of α-amino groups are acetylated (~80% in humans), reports of N-terminal acetylations in prokaryotes are rather scarce (Polevoda and Sherman, 2003; Jones and O’Connor, 2011; Ouidir et al., 2015). However, it remains to be seen whether this type of PTM is in fact such a rare event or whether the few references are due to a lack of systematic proteomic approaches. Interestingly, the majority of prokaryotic proteins with acetylated N-terminus that have been found so far are associated with translation. In E. coli for example, the three N-terminal acetyltransferases RimL, RimJ and RimL specifically modify the ribosomal proteins S18, S5 and L12 with Ala-Arg, Ala-His and Ser-Ile N-termini (Yoshikawa et al., 1987; Tanaka et al., 1989). At least in the case of L12 the modification levels seem to be correlated with the growth phase (Ramagopal and Subramanian, 1974). Physiologically this acetylation increases the intramolecular interaction in the so-called ribosomal stalk complex and thus makes the bacterial cell more resistant to stress (Gordiyenko et al., 2008). Nevertheless, more research is required to unveil the physiological relevance of bacterial α-N-terminal protein acylation in general and of arginine in particular. It is worth mentioning here that there are indications coming from mass spectrometry analyses that the guanidino group of peptide bound arginine is also subject to acetylation (Jedlicka et al., 2018). However, these data so far lack biochemical proof.

Hydroxylation

Hydroxylation means the chemical reaction in which a hydroxyl group (-OH) is added to an organic compound. In proteins, mainly three amino acids serve as acceptors: most commonly the C3- or C4-position of proline, the C5-position of lysine and the C3-position of asparagine.

C5-position of lysine and the C3-position of asparagine...

Funding: Deutsche Forschungsgemeinschaft, Funder Id: http://dx.doi.org/10.13039/501100001659, Grant Numbers: DFG Research Training Group, GRK2062/1, LA 3658/11.

References


reductase activating glycohydrolase (activating enzyme) from *Rhodospirillum rubrum*. J. Biol. Chem. 261, 10104–10111.


Bionotes

Jürgen Lassak
Center for Integrated Protein Science Munich (CIPSM), Department of Biology I, Microbiology, Ludwig-Maximilians-Universität München, Grosshaderner Strasse 2-4, D-82152 Planegg, Germany, juergen.lassak@lmu.de.
https://orcid.org/0000-0003-3936-3389

Jürgen Lassak studied Biology at the Eberhard Karls Universität Tübingen with a main focus in microbiology. From 2005 on he worked as a scientist for 2 years at the Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute (HKI) in Jena on antibiotic resistance mechanisms in Gram-positive bacteria. He then moved to Marburg to gain his PhD at the Max Planck Institute for Terrestrial Microbiology on prokaryotic signal transduction mechanisms. Since 2010 he has been affiliated with the Department of Microbiology at the Ludwig-Maximilians-Universität München: after a 3-year period as postdoctoral fellow he first became a senior researcher and since 2015 is an independent research group leader. In 2018 he did his habilitation in Microbiology. His fascination has always been with bacteria and the molecular mechanisms by which microorganisms convert external or internal stimuli into an adequate biochemical response. In this regard he is especially interested in translational stress and post-translational modifications.

Franziska Koller
Center for Integrated Protein Science Munich (CIPSM), Department of Biology I, Microbiology, Ludwig-Maximilians-Universität München, Grosshaderner Strasse 2-4, D-82152 Planegg, Germany
https://orcid.org/0000-0003-2003-7129

Franziska Koller studied Biology at the Universität Regensburg and the Technische Universität München, Germany and graduated in 2017 (MSc). In 2018 she started her PhD focusing on protein rhamnosylation under the supervision of Jürgen Lassak at the Ludwig-Maximilians-Universität München, Germany.

Ralph Krafczyk
Center for Integrated Protein Science Munich (CIPSM), Department of Biology I, Microbiology, Ludwig-Maximilians-Universität München, Grosshaderner Strasse 2-4, D-82152 Planegg, Germany
https://orcid.org/0000-0001-6171-6086

Ralph Krafczyk received his Master of Science degree in Biology from the Ludwig-Maximilians-Universität München in 2015. He proceeded to perform his doctoral studies in Molecular Microbiology under the supervision of Professor Kirsten Jung and PD Jürgen Lassak at LMU München, Germany. During this time, he focused on the structural and biochemical characterization of arginine glycosylation in Pseudomonas putida. He is currently a post-doctoral fellow in the Department of Microbiology at LMU, working on the synthetic engineering of protein glycosyltransferases.

Wolfram Volkwein
Center for Integrated Protein Science Munich (CIPSM), Department of Biology I, Microbiology, Ludwig-Maximilians-Universität München, Grosshaderner Strasse 2-4, D-82152 Planegg, Germany
https://orcid.org/0000-0003-7129

Wolfram Volkwein studied Biology at Ludwig-Maximilians-Universität München, Germany where he obtained his MSc degree in 2014. In the same year he started his PhD thesis under the supervision of Kirsten Jung and Jürgen Lassak at the Ludwig-Maximilians-Universität München, Germany, working on the synthetic activation of elongation factor P. Since 2019 he has been employed at the Bavarian Health and Food Safety Authority, Germany.