Prokaryotic Ubiquitin-Like Protein and Its Ligase/Deligase Enyzmes

Review Article

Author(s):
Delley, Cyrille L.; Müller, Andreas U.; Ziemski, Michal; Weber-Ban, Eilika

Prokaryotic ubiquitin-like protein and its ligase/deligase enyzmes

Cyrille L. Delley[a], Andreas Müller[a], Michal Ziemski[a] and Eilika Weber-Ban[a*]

[a]ETH Zurich, Institute of Molecular Biology & Biophysics, Otto-Stern-Weg 5, 8093 Zurich, Switzerland

Correspondence to Eilika Weber-Ban: eilika@mol.biol.ethz.ch; Phone: +41446333678

Abstract

Prokaryotic ubiquitin-like protein (Pup) and the modification enzymes involved in attaching Pup to or removing it from target proteins present a fascinating example of convergent evolution with respect to eukaryotic ubiquitination. Like ubiquitin (Ub), Pup is a small protein that can be covalently attached to lysine side chains of cellular proteins, and like Ub it can serve to recruit tagged proteins for proteasomal degradation. However, unlike Ub, Pup is conformationally highly dynamic, exhibits a different linkage connectivity to its target lysines and its ligase belongs to a different class of enzymes than the E1/E2/E3 cascade of ubiquitination. A specific feature of actinobacteria (aside from sporadic cases in a few other lineages), pupylation appears to have evolved to provide an advantage to the bacteria under certain environmental stresses rather than act as a constitutive modification. For *Mycobacterium tuberculosis*, pupylation and the recruitment of pupylated substrates to the proteasome supports persistence inside host macrophages during pathogenesis, rendering the Pup-proteasome system an attractive drug target.

In this review, we consider the dynamic nature of Pup in relation to its function, discuss the reaction mechanism of ligation to substrates as well as cleavage from pupylated

substrates and put it in perspective of the evolutionary history of this post-translational modification.

Introduction

Post-translational modifications (PTMs) of proteins change the properties of their targets to diversify and regulate their interactions with other cellular components [1]. Many diverse consequences result from these modifications, amongst them also the targeted degradation of the modified protein, thereby changing its cellular half-life. PTMs known to be involved in targeting proteins for degradation include for example phosphorylation [2, 3] or acetylation [4].

A particularly intriguing group amongst PTMs are the small protein modifiers. Post-translational modifications of proteins by covalent attachment of other, usually small, proteins exist in all domains of life [5]. The first protein modifier, Ubiquitin (Ub), was discovered in eukaryotes, and has since become the paradigm for this class of post-translational modifiers [6-8]. Compared to PTMs with small chemical moieties (e.g. phosphorylation, acetylation) or with small molecules (e.g. glycosylation, palmitoylation), use of a protein tag provides additional unique features. Due to their larger size, polypeptide modifiers can feature higher functional density, as different portions of the polypeptide can take over different functions, they can engage in interactions with larger footprints, can provide large recognition handles and, perhaps most importantly, they provide the possibility for modification of the modifier, since they can become themselves targets of other modifications [9-11]. Small protein modifiers are attached to their protein targets through isopeptide bonds, formally by attack of a lysine side chain amino-group from the target protein on an activated carboxylate group of the modifier protein.

Prokaryotic ubiquitin-like protein (Pup) was the first such small protein modifier identified in prokaryotes, demonstrating that the functional principle of tagging a protein

with another protein is not restricted to eukaryotes (Figure 1) [12, 13]. Modification of target proteins with Pup, a process referred to as pupylation, has evolved in the phylum of actinobacteria independently of eukaryotic ubiquitination. In mycobacteria and many other actinobacteria, pupylation is tightly connected to proteasomal degradation [14, 15], while many other members of the phylum do not possess proteasomal subunit genes. The general arrangement of the gene locus is conserved (Figure 2), featuring the *pup* gene at its center, while the genes for the modification enzymes and the proteasome components are found upstream or downstream of the *pup* gene. In addition to actinobacteria, the Pup locus was shown to be sporadically present in a few members of nitrospira, verrucomicrobia, planctomycetes, the delta subdivision of proteobacteria and armatimonadetes (Figure 2) [16].

The Pup-proteasome system (PPS) is functionally related to the eukaryotic Ubiquitin-proteasome system (UPS), but the number of the involved players is smaller [5, 17, 18]. A single ligase, PafA (proteasome accessory factor A), is responsible for modification of all target proteins with Pup, and reversibility of the modification is ensured by a single depupylase enzyme Dop (deamidase of Pup) [13, 19-22]. In mycobacteria and other proteasome-harboring actinobacteria, pupylated substrates can be recruited to a specific proteasome complex comprising the 20S core and an ATPase ring complex of the AAA family [14, 15], called Mpa (mycobacterial proteasome ATPase) in mycobacteria [23] and ARC (ATPase forming ring-shaped complexes) in other actinobacteria [24].

Despite the obvious homology between the bacterial and the eukaryotic proteasome complexes, ubiquitination and pupylation have evolved independently. In this review, we highlight the distinct features of the Pup-proteasome system, focusing on the

structural and biophysical properties of the modifier Pup itself as well as on the two opposing enzymes of this reversible post-translational modification, the Pup ligase PafA and the depupylase Dop.

Flexibility as a functional principle in prokaryotic ubiquitin-like protein Pup

Prokaryotic ubiquitin-like protein Pup was named for its functional analogies to eukaryotic ubiquitin [12, 13]. Indeed, the common principle of covalently attaching a small protein via its C-terminal residue to a lysine side chain of a target protein (Figure 1), is another striking example of how evolution tends to arrive at similar solutions. In this case, both systems exploit the ability of the lysine side chain to act as a nucleophile in its deprotonated state, whereby the obvious target of the nucleophilic attack on a protein is a carboxylate. In case of ubiquitin, this is the C-terminal α-carboxylate, while in Pup it is the γ-carboxylate of the C-terminal glutamate residue (Figure 1) [25]. Modification via a terminal residue allows for easy accessibility into the active sites of the modification enzymes and sufficient flexibility for attachment to a range of different protein target surfaces.

In mycobacteria, the bacterial genus where pupylation was discovered, the functional analogy is also present in the role of Pup as a recruitment tag to proteasomal degradation of the pupylated substrate [12, 13]. Many other actinobacteria, for example corynebacteria and bifidobacteria, feature the pupylation modification pathway in absence of a proteasome (Figure 2), suggesting functions for pupylation that are independent of degradation. This in turn allows drawing parallels to the proteasome-independent roles of ubiquitin or other ubiquitin-like small protein modifiers.

From a structural point of view however, Pup is a misnomer, as the two modifiers Pup and Ub fall to diametrically opposite ends of the conformational and the associated energy landscape spectrum. Ubiquitin is a stably folded protein that populates one dominant conformational state of lowest energy and shows, with the exception of the C-terminal four residues, very low flexibility (Figure 3) [26]. Its highly compact fold consists of a twisted, 5-stranded β-sheet, across which an α-helix packs diagonally, making it appear as if the β-sheet grasps the helix by cradling it in its concave surface. Pup on the other hand was shown to be an intrinsically disordered protein (IDP) without a stable fold [27-29]. Its sequence features the characteristic enrichment in charged residues (Figure 4B), a contributing factor to preventing formation of a compact core for most IDPs. The ^{15}N-HSQC (heteronuclear single quantum correlation) NMR spectrum of Pup, which contains a peak for each backbone NH, shows low dispersion under physiological conditions [27-29], which is another characteristic for a disordered protein. IDPs are known to exist as dynamic conformational ensembles with each conformation representing only a shallow minimum on a rather flat energy landscape [30, 31]. The presence of NOE (nuclear Overhauser effect) NMR signals characteristic for α-helical structure in a segment from residue 52-58 (of 64 total) indicates, that the Pup conformation ensemble occasionally samples an α-helical state in this region [29].

As will be further discussed below, Pup undergoes disorder-to-order transitions upon interaction with its dedicated binding partners, the Pup ligase PafA [32], the depupylase Dop [33] and the proteasomal ATPase Mpa [34]. It is not uncommon for an IDP to already in the unbound state sample conformational elements of the structure it will adopt during binding [31, 35]. However, the structured conformations that Pup

samples are poorly populated and transient, so that the disordered nature prevents structural characterization by the usual methods [27-29]. Free Pup also remains mainly disordered under the crowded conditions inside the bacterial cytosol, as has been demonstrated by in-cell NMR experiments upon overexpression of labeled Pup in *E. coli* [36]. This does not preclude the possibility that Pup could adopt more structure under very specific conditions. A recent study showed that Pup can shift its conformational ensemble further toward helical conformation in presence of SDS micelles and even more significantly toward helical conformation at low pH (pH3 – pH1) [37]. Whether this might play a role under certain conditions in vivo remains to be seen. It should be noted in this context, that it has been suggested Pup might specifically be attached to membrane proteins in certain sporadic members outside actinobacteria, since the Pup ligases in these organisms feature predicted trans-membrane helices [16].

However, to a protein with its main function as mediator of recognition between interaction partners, intrinsic disorder coupled to disorder-order transitions during the interaction process can bring advantages [38]. It is a very efficient way to allow small proteins to undergo interactions with different binding partners, a phenomenon that has been described as "high functional density". Furthermore, through the process of disorder-to-order transition, IDPs are able to form rather large and therefore specific binding interfaces [31, 38, 39]. It is now generally accepted that binding of IDPs to a folded target often involves a mixture of conformational selection before and induced folding after/during binding, allowing the IDP to "anneal" to its target and adopt a perfect fit into extended binding interfaces [39-41]. At the same time, the interplay between loss of conformational flexibility (entropy) and favorable interaction with the partner protein allows

for a large range in affinities depending on the conformations adopted during binding [40]. This is particularly important in recognition scenarios, where high affinity is not desired, as we see in the next paragraph for the case of Pup recognition by the proteasome complex.

Pupylation-dependent route to proteasomal degradation – a conformational journey for Pup

In all mycobacteria and many other actinobacteria, pupylation serves as a recruitment tool for proteasomal degradation [12, 13]. The Pup-proteasome pathway for degradation exists in these bacteria in addition to the typical bacterial chaperone-protease degradation systems like the Clp proteases, Lon or FtsH, and the bacteria do not rely on it for survival under normal conditions [42]. Rather, the adoption of a proteasome by horizontal gene transfer and the evolution of a recruitment pathway around it, afforded actinobacteria an advantage under specific conditions and for specific purposes. The precise roles played by pupylation and Pup-driven degradation in these organisms, we are only beginning to understand. Most progress in this direction has been made in *Mycobacterium tuberculosis* (Mtb) and its non-pathogenic, close relative *Mycobacterium smegmatis* (Msm). Mtb persistence inside host macrophages is supported by a functional PPS locus [43, 44], which can at least in part be traced back to the prevention of nitric oxide (NO) toxicity for the pathogen by proteasomal degradation of an enzyme responsible for production of cytokinins [45]. Although the exact mechanism is not fully understood, the breakdown products of cytokinins together with the NO produced by the macrophage lead to strongly bacteriotoxic effects. This is a striking example where an advantage for survival can be connected to degradation of an individual proteasomal

substrate. A growth advantage due to more global proteasomal degradation for the purpose of amino acid recycling was proposed for Msm under nitrogen starvation, since disruption of pupylation significantly impairs survival in medium lacking nitrogen sources [46]. A recent proteomic study suggests in addition a more specific role for proteasomal degradation during nitrogen starvation by adjusting levels of enzymes involved in assimilation of nitrogen, since key players of nitrogen assimilation showed significant changes in abundance when comparing an MsmΔ*pup* strain with the parent strain [47]. In *Streptomyces coelicolor*, disruption of pupylation was recently shown to cause a sporulation defect [48].

Irrespective of the roles played by individual degradation substrates or groups of substrates, it is well-supported that the substrate clientele consisting of pupylated proteins (the so-called pupylome) for the proteasome complex under certain growth conditions encompasses a hundred or more cellular proteins differing in structure, assembly state and abundance [48-51]. Their unifying recognition property is the covalent modification of one or more of their lysine side chains with Pup, which mediates their interaction with a ring-shaped, hexameric proteasomal ATPase complex referred to as Mpa in mycobacteria or ARC in other actinobacteria. The bacterial proteasome core particle (20S proteasome, CP) has been shown to interact with two alternative, ring-shaped regulator complexes, that are responsible for substrate recruitment, namely the ATPase ring Mpa [52, 53] and the non-ATPase ring Bpa (bacterial proteasome activator) [54, 55]. Only the Mpa-CP complex recognizes pupylated substrates, while the Bpa-CP complex recruits substrates in a Pup-independent manner based on their conformational state [56].

The subunit structure of Mpa/ARC resembles the one found in the ATPase subunits (Rpts) of the eukaryotic 19S regulatory particle: a helical N-terminal region, two (instead of only one in the Rpts) OB (oligonucleotide binding) domains each forming a β-barrel and a C-terminal AAA module [57]. The six AAA modules arrange into the main ring body with a central pore, into which flexible, ATPase-driven loops responsible for substrate unfolding and translocation protrude; on top of the ATPase ring the two OB barrels each arrange into hexameric rings of their own, creating a stiff tubular conduit for entering substrates [53]. The N-terminal helices emerge from the upper OB ring and two neighboring helices form a coiled-coil each. In contrast to the Rpts, where the coiled-coil domains are involved in holding together the components of the regulatory particle [58-60], it was shown that in Mpa/ARC the three coiled-coil domains present the site of substrate capture [29]. Their flexible, extended conformation is perfectly shaped for easy accessibility and sampling of the surrounding solution, "fishing" for interaction partners. By the same token, IDPs have been proposed to have a greater capture radius than globular proteins, a property that might further enhance the capture of Pup by Mpa [61]. Inspection of the pattern and conservation of residues in Pup sequences from various actinobacteria reveals a distinctly bipartite character of the Pup primary sequence: the C-terminal half (residues 31-64 in Mtb Pup) shows a high level of sequence conservation paired with a characteristic pattern of alternating hydrophobic and hydrophilic residues, indicative of helical structure with a propensity for formation of coiled-coils, while the N-terminal region shows poor conservation and features mostly disorder-promoting residues (Figure 4) [29]. This dual flavor reflects the functional roles encoded in the two halves of Pup. The C-terminal region shows high functional density as a binding and

interaction template, since it contains the covalent attachment site, the interaction region with the ligase (see next paragraph) and the binding region for docking transiently to the proteasomal ATPase (Figure 4).

Based on the coiled-coil propensity encoded in the Pup sequence and the identification of the coiled-coil domains of Mpa as the docking site for Pup, an intriguing disorder-to-helix transition model was suggested for the recruitment of Pup to the Mpa-CP complex, where upon binding to one of the Mpa coiled-coil domains, Pup adopts a helical fold in the region featuring the alternating hydrophobic/hydrophilic sequence motif, forming a shared coiled-coil with Mpa [29]. This model was later experimentally confirmed by a crystal structure of the complex between a Pup fragment (residues 21-64) and an Mpa fragment (residues 46-96) corresponding to the coiled-coil region, as well as a low resolution structure of Pup with a truncated Mpa ring complex corresponding to the double OB domains bearing the three N-terminal coiled-coils [34]. The structures beautifully demonstrate that Pup forms a long helix (residues 21-51) upon binding and that it interacts with the lower (C-terminal) half of the Mpa coiled-coil domain in an antiparallel manner to form the predicted shared coiled-coil (Figure 4A). Although three coiled-coils are available on Mpa, Pup binds with a 1:1 stoichiometry to the Mpa hexamer [29]. As Pup binds at the base of the coiled-coil, right at the rim of the pore formed by the OB-domains, this leads to a space constraint that likely forces this stoichiometry. Consequently, the other two coiled-coils are free to form additional interactions with the substrate portion of a pupylated protein, which might explain the observation that at least in vitro, pupylated proteins outcompete free Pup for recruitment to the Mpa-CP complex [52].

The poorly conserved N-terminal half of Pup exhibits less propensity to adopt transient secondary structure elements and might remain disordered under most conditions, even when Pup is in complex with interaction partners, further underlining the bipartite nature of Pup. For the recruitment to proteasomal degradation the N-terminal region of Pup plays an important role, although it has been demonstrated that it does not participate in the interaction with the Mpa coiled-coil domain [29]. Proteasomal substrates that are modified with a truncated version of Pup can bind to the Mpa-CP complex, but cannot be degraded [52]. An analogy can be drawn here to degradation of ubiquitinated proteins by the 26S proteasome, where polyubiquitinated substrates require a so-called "initiation site", a loosely structured conformational element on the substrate protein itself, from which threading into the ATPase central pore can be initiated [62, 63]. In case of Pup, the N-terminal unstructured region of the degradation tag itself fulfills this role [52, 64]. Access of degradation substrates to the active sites inside the bacterial core particle is gained by translocation of the substrate through the Mpa pore, a movement that requires the ATPase-driven movement of pore loops contacting the translocating substrate starting from the Pup initiation site (Figure 5). The antiparallel arrangement of the shared coiled-coil points the disordered Pup N-terminal region directly into the pore entrance. Once translocation is initiated, directional transport into the proteasome core is supported by the ATPase-driven loop movements and eventually the pulling force will lead to the disruption of the shared coiled-coil interaction and a helix-to-disorder transition in this stretch of Pup. In the context of this transient interaction, it is important that the strength of Pup binding to the coiled-coil domain, despite the specificity, is not too high,

keeping a fine balance between the loss of conformational flexibility (entropy) and favorable interaction.

One aspect of Pup-dependent proteasomal degradation that remains elusive is the rather weak association of Mpa with the 20S core in vitro [53], which necessitated a truncation of the 20S proteasome gate in order to allow in vitro analysis of Pup-driven degradation [53, 65]. It is unknown, how efficient degradation is established in vivo by the full-length proteasome. Post-translational modification, specifically phosphorylation has been suggested as one possibility [66]. Alternatively, it is also possible that an additional, as yet unidentified proteinaceous factor is involved. With the exact state or composition of the Mpa-CP degradation complex unknown, this also leaves open the question, whether Pup is degraded along with the substrate, as is the case with open-gate proteasome in vitro, or whether there might be an as yet undiscovered mechanism of Pup-recycling.

Structural analysis of Pup and its binding to the proteasome had so far only been investigated for Pup alone or even fragments of Pup, with the implied assumption that Pup remains disordered when covalently attached to proteasomal substrate proteins. However, considering the existence of many shallow energy minima for the possible conformations of an IDP and the usually low energy barriers between them, proximity to the surface of a large substrate protein might tip the balance in favor of a more ordered state. Should Pup adopt structure, this would have the potential to change the binding preference for one binding partner over another. Recruitment to the Mpa-CP could potentially be favored by formation of a long helix, while a different conformation at the surface of the substrate could make the Pup-binding region unavailable for interaction

with Mpa. A recent NMR study resolved this open question by generating isotopically labeled Pup and ligating it to unlabeled substrate in vitro [67]. Chemical shift comparison between free Pup and the Pup covalently ligated to proteasomal substrates revealed that Pup adopts an intrinsically disordered state in the conjugate similar to the one observed for the free, unbound form. It can therefore be excluded that a preformed structure either enhances or hinders the interaction with Mpa or other binding partners, but that rather, at least in the majority of proteasomal substrates, Pup remains conformationally naïve and thereby competent for binding to different interactors requiring different disorder-to-order transitions of Pup. On its way to proteasomal degradation, Pup therefore undergoes a conformational journey with a range of disorder-to-order transitions [67]. As the next paragraph will show, its interaction with the modification enzymes requires yet a different conformation.

The pupylation-depupylation enzymes – two birds of a feather

Despite the striking analogy between pupylation and ubiquitination in covalently attaching the C-terminal residue of a small modifier protein to lysine side chains in the target protein via an isopeptide bond [12, 13], enzymes related to the E1/E2/E3 modification cascade are not present in mycobacteria and other actinobacteria. Instead, in actinobacteria a pair of enzymes unrelated to the ubiquitination enzymes was identified by pull-down experiments with Pup-decorated beads and shown to act in the pupylation pathway [21]. The two proteins, PafA and Dop, are encoded in the Pup-proteasome gene locus (Figure 2), and pupylated proteins are undetectable in mycobacterial knockout strains of either gene [13, 20]. Strikingly, PafA and Dop are close structural homologs (Figure 6), consisting of two tightly interacting domains, a small C-terminal domain and a

14

large N-terminal domain that shares its fold with members of the γ-carboxylate-amine ligase superfamily [16, 33]. Despite this close homology, however, it was shown that they play quite different roles in the pupylation pathway [19, 20, 22, 68]. The enzyme PafA acts as the single existing ligase attaching Pup covalently to target proteins [13, 21], whereas its twin Dop catalyzes the opposing activity by cleaving the isopeptide bond between Pup and the target protein (Figure 7) [19, 68]. Notably, in mycobacteria and many other actinobacteria, depupylase Dop also plays an important role in the pupylation branch of the modification cycle. The reason is that in those members of actinobacteria Pup is encoded in a ligation-incompetent form bearing a glutamine residue at its C-terminus (PupQ) rather than a glutamate (PupE). The same way Dop can cleave the C-N bond between Pup and the target lysine, it can catalyze the removal of ammonia from glutamine, thereby deamidating PupQ to the coupling-competent PupE [21].

Like glutamine synthetase (GS) and glutamate cystein ligase (GCL), both γ-carboxylate-amine ligase family members, the homologous N-terminal domain of PafA and Dop features an antiparallel twisted β-sheet, carrying the active site on its concave surface (Figure 6) [33]. One end of this β-sheet cradle is closed off by loops from both the C- and N-terminal domains and contains the conserved nucleotide binding site characteristic for this family of enzymes. The opposite end is more open and leads into a highly conserved, 40-50 Å long groove, characteristic only for Dop and PafA members. A structure of the complex between PafA and Pup showed that Pup binds into this highly conserved groove, placing its C-terminal glutamate residue into the active site at the equivalent position of the glutamate in the GCL family of enzymes (Figure 6, upper left) [32, 69]. Upon binding to PafA, Pup undergoes a disorder-to-order transition in the C-

terminal, conserved half of its primary sequence, resulting in the formation of two helices connected by only a short linker (H1: S38-L47; H2: A51-Y58) (Figure 4 and 6). The two helices fit closely into the binding groove on PafA, following it perfectly and burying an impressive surface area of more than 1500 Å^2 in the interaction interface. The order-to-disorder transition likely involves some element of conformational selection, given that Pup transiently samples small stretches of helical conformation (Figure 3), followed by annealing along the groove to adopt the final conformation. This mode of binding allows for a particularly snug fit into the interaction groove, providing the interaction with the necessary affinity. This is important, since Pup is not available at high millimolar concentrations like the glutamate serving as substrate in the homologous GS and GCL.

The helix further away from the active site (H1) provides the main portion of the thermodynamic driving force for the Pup-PafA interaction, while the second helix employs stacking interactions between its conserved F54 and Y58 aromatic residues to three aromatic residues on PafA (Y114, F116 and Y127) thereby anchoring the Pup C-terminal helix and the C-terminal glutamate in the correct position for the reaction in the active site to take place [32]. The surface residues in the equivalent binding groove on Dop are highly conserved, so that it is expected that the mode of interaction and the structure induced in Pup are very similar [33]. Nevertheless, some differences remain and, in particular, the conformations of flexible loops in the Dop structure could have some influence on the accessibility of the binding groove or the potential locking in of Pup. Specifically an extended loop present only in Dop members but not in PafA members between α-helix 1 and β-strand 2, referred to as the "Dop loop", is unresolved in the Dop crystal structure [33]. Given the high degree of conservation of residues in this loop, it is

expected to play an important role that remains to be elucidated. One intriguing possibility is that, like Pup, it undergoes a disorder-to-order transition upon Pup binding. A structure of Dop with Pup bound could provide interesting novel insights on that possibility.

Reaction mechanism of pupylation and depupylation

The evolutionary relationship of the Pup ligase to the γ-carboxylate-amine ligase superfamily, encompassing members like GS and GCL [16], and the fact that PafA turns over ATP to ADP stoichiometrically with pupylation events [21], provided clues about the potential reaction mechanism. It had been suggested for members of the carboxylate-amine ligase family of enzymes, that a phosphate mixed anhydride intermediate is formed [70-72], although the intermediate could never be isolated.

For the Pup ligase PafA, using [γ-^{32}P]-ATP it could be shown that upon ATP cleavage the γ-phosphate is transferred to Pup, thereby activating it for nucleophilic attack by the lysine side chain of the target protein [73]. Unprecedented for this enzyme family, the generated phospho-Pup intermediate could in this case be detected by mass spectrometry, and it was shown that ADP and the phosphorylated Pup-intermediate remains poised in the active site awaiting the arrival of the nucleophilic substrate. The incoming lysine side chain is likely activated for nucleophilic attack through abstraction of a proton by a highly conserved aspartate residue on PafA (D64). The pupylation rate is limited by binding of the nucleophilic substrate under low substrate concentration, but when the target protein is present at saturating concentrations, the pupylation rate is limited by the rate of activation [73].

PafA is responsible for the modification of several hundred target proteins, collectively referred to as the pupylome [48-51, 74], but nevertheless shows some degree of specificity, since not all exposed lysines get modified. Furthermore, the specificity constant (k_{cat}/K_m) for a bona fide substrate is three orders of magnitude higher than the one observed for free lysine as a model substrate [73]. A recent study suggests that the flexible loop N-terminal of strand 7 at one edge of the β-sheet cradle is involved in making contacts to pupylation substrate proteins, since mutations in this region (196-216) lead to diminished activity in protein pupylation but not in Pup amidation using ammonia as a model substrate [75]. In the crystal structure, the loop runs roughly parallel to β-strand 7, partially projecting above the active site cradle [33]. It thereby forms part of the surface around the active site that needs to be approached by the protein carrying the nucleophilic lysine in order for the lysine to reach into the active site.

Although many of the Pup orthologs feature in their primary sequence additional lysine residues, formation of Pup-chains on the substrate, as is observed in case of ubiquitin, does not appear to play a significant role in vivo [48-51, 74]. An in vitro study showed for two known pupylation substrates (IdeR and PanB) that poly-pupylation, when it occurs, proceeds via sequential addition of Pup to a growing Pup chain on the substrate rather than by attaching an existing poly-Pup [76]. The study further concludes that in case the affinity for mono-pupylated substrate is similar to unpupylated target as for IdeR, polypupylation occurs along with monopupylation, whereas for much higher affinity of unpupylated target, as is the case for PanB, monopupylation is preferred. Of course, considering that for polypupylation the Pup unit on the mono-pupylated substrate must bind to the active site cradle rather than the target protein portion, a competition for

binding between the different portions of the pupylated target (Pup or target protein portion) is expected to occur also in absence of unpupylated target protein.

Interestingly, the Pup ligase PafA itself is a pupylation target [46, 49, 51], and it was shown that in an Msm proteasome knockout, PafA is stabilized [46], indicating that PafA levels can be regulated by pupylation. Furthermore, a PafA variant (K320R) lacking the main modified lysine residue, when expressed in Msm, is more stable than the wild type PafA [77].

The formation of an isopeptide bond linkage between the side chain carboxylate of Pup's C-terminal glutamate and the lysine side chain of the target by PafA requires the activation of the carboxylate through phosphorylation, and therefore results in stoichiometric turnover of ATP to ADP with ligation events [21, 73]. In contrast, it was shown early on that depupylase/deamidase Dop, featuring a homologous nucleotide binding site, does not turn over ATP for each catalytic cleavage event [21, 68]. Nevertheless, non-hydrolyzable ATP analogs do not support catalysis. Co-crystallization experiments with Dop and ATP resulted in a structure that clearly featured ADP and inorganic phosphate in the active site, and in vitro depupylation and deamidation assays demonstrated that ADP and Pi, but not ADP alone, can support catalysis (Figure 6, left) [78]. In fact, depupylation time courses upon addition of ATP exhibit a lag phase that is absent when ADP and Pi are supplied directly, indicating that ATP must be turned over to ADP and Pi before catalysis can occur. Together, the evidence points to a mechanistic role of inorganic phosphate for catalysis [78]. In light of the evolutionary relationship to the carboxylate-amine ligase family, it is plausible that this role is formation of a transient phospho-Pup intermediate (Figure 6, left). Such a mechanism would involve the

nucleophilic attack of an inorganic phosphate oxygen on the side chain carbonyl carbon of the isopeptide bond to form the same phospho-Pup intermediate that occurs in the forward reaction of the Pup ligase PafA. However, other than in the active site of PafA, the mixed anhydride intermediate is immediately attacked by water, releasing Pup from the active site and regenerating the inorganic phosphate active site species. Aspartate residue D94, the homologous residue to D64 in the ligase PafA, where it was suggested to act as catalytic base to accept a proton from the attacking lysine, might play an equivalent role in Dop by activating water for the attack on the phospho-Pup intermediate [33]. Another study suggested that this aspartate residue might directly attack the γ-carboxylate of Pup, resulting in a covalent intermediate that has not been observed in any other enzyme of this family [79]. The mechanism outlined in Figure 6 (right side) on the other hand takes into account the high degree of structural homology observed between PafA and Dop and the carboxylate-amine ligase family of the glutamine synthetase type, since it resembles the conversion of glutamine to glutamate by GS in presence of ADP and arsenate [78, 80].

Little is known about the substrate specificity of depupylation, although it appears to be similarly broad as the ligase activity of PafA, since depupylation of the purified "pupylome" leads to a general decrease in the ladder of pupylated proteins [68]. As Dop acts on the conjugated substrate, accessibility of the covalent linkage plays a role. In vitro experiments have shown that the presence of Mpa alongside Dop can enhance depupylation of a bona fide pupylated substrate tenfold and that this enhancement is dependent on the ATPase-driven unfolding of the pupylated substrate, presumably making the isopeptide bond more accessible [68]. It was suggested that this might be one

reason why the proteasomal ATPase gene (*arc*) is maintained even in those actinobacteria that do not harbor the proteasome core subunits. In addition to rendering the pupylated substrate more accessible to the depupylase, the unfolded substrate emerging from the ARC complex could also become a degradation substrate of other cellular proteases. Furthermore, ARC can function in disassembly of larger complexes, if some of their components are pupylated. Such an example was recently shown for *Corynebacterium glutamicum*, where pupylation of the iron storage protein ferritin leads to recruitment to the unfoldase ARC causing disassembly of the ferritin complex and thereby release of iron during iron starvation [81].

The dual role of the depupylase Dop in some of the actinobacteria is a curious feature. As mentioned earlier, in all mycobacteria, Pup is encoded with a C-terminal glutamine, necessitating deamidation of Pup's C-terminal residue before conjugation to a substrate lysine is possible (Figure 7). Other actinobacteria encode the already ligation-competent form of Pup featuring a glutamate at the C-terminus. The reason could be that deamidation is a fine-tuning feature of the pupylation pathway in certain organisms, guaranteeing that pupylation does not occur unless the opposing player, the depupylase is present and active. Notwithstanding the role of deamidation in some actinobacteria, the interplay of ligation and deligation in all actinobacteria is sure to play an important role in determining the fate of pupylation substrates.

Closing Remarks

With high-resolution structural information available now for most components of the Pup-proteasome system, and the careful mechanistic analysis of the

pupylation/depupylation enzymes, an important goal will be the translation of this information into the development of novel antituberculosis drugs that target the problem of persistence. A challenge that will require the concerted efforts of synthetic chemistry, medicinal chemical biology and in vivo screening.

Several mechanistic questions remain open, particularly in the context of the cellular environment. For example, it is an unresolved question, whether in vivo Pup is degraded along with the pupylated substrate as it was shown to be the case in vitro or whether additional factors or regulatory circuits might exist that would allow for Pup to be recycled from the pupylated substrate. In vitro studies of bacterial proteasomal degradation have so far relied on only a handful of well-characterized model substrates and on a truncated variant of the proteasome core that exhibits increased affinity for the proteasomal ATPase ring.

A very exciting, yet rather uncharted territory is the potential interplay of pupylation and other post-translational modifications. This will most likely require state-of-the art mass spectrometric analysis in combination with in vivo studies to unravel any existing modification crosstalk.

It is also evident from the multitude of pupylation substrates that have been identified and from the conditions already characterized, where the pupylation locus imparts a survival advantage, that the roles of pupylation for the bacteria are likely to be manifold as well. Valuable lessons will be learned from the study of pupylation and the conditions under which it makes a difference for survival and fitness in other important actinobacterial model organisms. With every uncovered new facet, a more complete picture of this fascinating modification pathway emerges.

Figure legends

Figure 1. **Bacterial pupylation and eukaryotic ubiquitination are post-translational modification pathways that covalently attach a small protein modifier to a target protein.** A cartoon representation of the overall reaction of pupylation (a) and ubiquitination (b) is shown. In both pathways the lysine side chain in a cellular target protein is covalently attached to the C-terminal residue of a small marker protein via an isopeptide bond. A) The ligation-competent form of intrinsically disordered prokaryotic ubiquitin-like protein Pup carries a C-terminal glutamate residue, featuring a C-terminal α-carboxylate as well as a side chain carboxylate. The isopeptide bond to the lysine residue of the target protein occurs via the side chain carboxylate. B) The stably folded ubiquitin carries a C-terminal glycine residue and the isopeptide bond is made through the C-terminal α-carboxylate.

Figure 2. **Gene arrangement in the pupylation gene locus.** The pupylation gene locus exists in all actinobacteria and also occurs sporadically in a few other lineages (verrucomicrobia, nitrospira, planctomycetes, the delta subdivision of proteobacteria and armatimonadetes). The arrangement of genes belonging to the pupylation gene locus is conserved, although some members have lost the proteasomal subunit genes and in some of the sporadic members only the *pup* gene along with the Pup ligase gene *pafA*

23

remain. The alignment of the gene locus is centered on the *pup* gene (red), which is organized in an operon with the proteasomal subunit genes (beige) (where they are present). The gene encoding the depupylase, *dop* (green), is usually found directly upstream of the *pup* gene. The proteasomal ATPase Mpa is encoded in a separate operon (orange) located even a bit further upstream of the *dop* gene. The Pup ligase gene *pafA* (blue) on the other hand is located downstream of the *pup* operon, sometimes separated from it by several genes unrelated to the pupylation gene locus. The intergenic distances are not presented to scale. Organism abbreviations are as follows: Mtb, *Mycobacterium tuberculosis*; Msm, *Mycobacterium smegmatis*; Scoe, *Streptomyces coelicolor*; Rery, *Rhodococcus erythropolis*; Cglu, *Corynebacterium glutamicum*; Bado, *Bifidobacterium adolescentis*; Mlut, *Micrococcus luteus*; Cfla, *Chthoniobacter flavus*; Lept, *Leptospirillum* sp.; Rbal, *Rhodopirellula baltica*; Ppac, *Plesiocystis pacifica*; Ccal, *Chthonomonas calidirosea*.

Figure 3. **Intrinsically disordered prokaryotic ubiquitin-like protein Pup is structurally "unlike" the stably folded ubiquitin.** Ubiquitin (X-ray structure, on the left) exhibits a stable, compact fold referred to as the "β-grasp fold", in which an α-helix packs diagonally against the concave surface of a strongly twisted β-sheet (PDB code: 1UBQ [26]). In contrast, Pup (on the right) is intrinsically disordered and exists as a dynamic conformational ensemble, occasionally and transiently sampling conformational elements of the structures it adopts during binding to its interaction partners. The depiction of the overlaid random coil Pup structures is not based on experimentally determined structures, but was generated by random assignment of peptide bond phi and psi angles using

24

Ramachandran-favored constraints for the purpose of visualization. Short secondary structure parts were copied from the available co-crystal structures of Pup with binding partners (PDB codes: 4BJR, 3M9D).

Figure 4. **The C-terminal region of Pup is well conserved and exhibits high functional density as a binding and interaction template.** A) The primary sequence of Mtb Pup is shown colored by residue conservation from the alignment shown in B. Regions involved in the interaction with Mpa or PafA are indicated by black bars above or below the primary sequence, respectively. Also shown next to the regions is the structure adopted by Pup when bound to Mpa (above) or PafA (below). B) Alignment of Pups from different actinobacteria colored according to conservation (shaded from white for no conservation to red for high conservation). Organism abbreviations are as follows: Mtb, *Mycobacterium tuberculosis*; Msm, *Mycobacterium smegmatis*; Cglu, *Corynebacterium glutamicum*; Rery, *Rhodococcus erythropolis*; Aaur, *Arthrobacter aurescens*; Scoe, *Streptomyces coelicolor*; Bado, *Bifidobacterium adolescentis*; Tfus, *Thermobifida fusca*; Rsal, *Renibacterium salmoninarum*; Stro, *Salinispora tropica*; Krhi, *Kocuria rhizophila*; Mlut, *Micrococcus luteus*; Acel, *Acidothermus cellulolyticus*.

Figure 5. **Pup serves as both recognition and threading element in proteasomal degradation of pupylated substrates.** The degradation substrate (grey) covalently modified with Pup (red) is recruited to the Mpa-proteasome complex (orange/beige) by docking to the Mpa coiled-coil domain, which triggers Pup to undergo a disorder-to-order transition forming an extended helix (Pup residues 21-51) that associated into a shared

three-stranded coil with the Mpa N-terminal coiled-coil domains. The disordered N-terminal region of Pup points into the Mpa central pore, where it is engaged by the ATPase-driven pore loops (depicted as dark grey cog wheels) for unfolding and directional translocation into the proteasome core for degradation.

Figure 6. **The pair of enzymes constituting the pupylation/depupylation cycle are close structural homologs with the same evolutionary origin and shared mechanistic features.** Pup ligase PafA (blue, on the left; PDB code: 4BJR [32]) and depupylase Dop (green, on the right; PDB code: 5LRT [78]) are depicted as cartoon models on a background of their surface outlines. The C-terminal domain (CTD) of each enzyme is bracketed by a half-circle. The active site aspartate is shown in yellow stick representation (D64N for PafA, D94 for Dop). Unresolved loops in PafA are shown as blue dashed lines. The extended, unresolved loop in Dop ("Dop loop") is shown as a black dashed line. PafA-bound Pup is colored in red with its C-terminal glutamate shown in stick representation. Helices H1 and H2 of Pup which form upon binding to PafA are indicated. Underneath each enzyme, its reaction mechanism is shown on a blue (for PafA) or green (for Dop) background. For simplicity, the tetrahedral intermediates that are transiently formed before formation of phospho-Pup and upon nucleophilic attack on phospho-Pup by either the lysine side chain or water are not shown in the reaction schemes.

Figure 7. **The Pup-proteasome system in mycobacteria.** In mycobacteria, Pup is encoded in a pupylation-incompetent form, featuring a glutamine at its C-terminus. Deamidase/depupylase Dop catalyzes the conversion of the C-terminal glutamine into a

glutamate, thereby rendering Pup ligation competent. Pup can then be covalently attached to the lysine side chain of a substrate protein (grey) to form the Pup-substrate conjugate. The modified substrate is then either recognized by the Mpa-proteasome complex (orange/beige), unfolded and degraded, or the modification is reversed by the depupylase Dop.

References

[1] Walsh CT. Posttranslational modifications of proteins expanding nature's inventory. Englewood, Colo.: Roberts and Company Publishers; 2006.

[2] Trentini DB, Suskiewicz MJ, Heuck A, Kurzbauer R, Deszcz L, Mechtler K, et al. Arginine phosphorylation marks proteins for degradation by a Clp protease. Nature. 2016;539:48-53.

[3] Barik S, Sureka K, Mukherjee P, Basu J, Kundu M. RseA, the SigE specific anti-sigma factor of Mycobacterium tuberculosis, is inactivated by phosphorylation-dependent ClpC1P2 proteolysis. Molecular microbiology. 2010;75:592-606.

[4] Hwang CS, Shemorry A, Varshavsky A. N-terminal acetylation of cellular proteins creates specific degradation signals. Science. 2010;327:973-7.

[5] Burroughs AM, Iyer LM, Aravind L. The natural history of ubiquitin and ubiquitin-related domains. Front Biosci (Landmark Ed). 2012;17:1433-60.

[6] Finley D, Ulrich HD, Sommer T, Kaiser P. The ubiquitin-proteasome system of Saccharomyces cerevisiae. Genetics. 2012;192:319-60.

[7] Kleiger G, Mayor T. Perilous journey: a tour of the ubiquitin-proteasome system. Trends Cell Biol. 2014;24:352-9.

[8] Hershko A, Ciechanover A. The ubiquitin system. Annual review of biochemistry. 1998;67:425-79.

[9] Herhaus L, Dikic I. Expanding the ubiquitin code through post-translational modification. EMBO reports. 2015;16:1071-83.

[10] Komander D. The emerging complexity of protein ubiquitination. Biochem Soc Trans. 2009;37:937-53.

[11] Ravid T, Hochstrasser M. Diversity of degradation signals in the ubiquitin-proteasome system. Nature reviews Molecular cell biology. 2008;9:679-90.

[12] Burns KE, Liu WT, Boshoff HI, Dorrestein PC, Barry CE, 3rd. Proteasomal protein degradation in Mycobacteria is dependent upon a prokaryotic ubiquitin-like protein. The Journal of biological chemistry. 2009;284:3069-75.

[13] Pearce MJ, Mintseris J, Ferreyra J, Gygi SP, Darwin KH. Ubiquitin-like protein involved in the proteasome pathway of Mycobacterium tuberculosis. Science. 2008;322:1104-7.

[14] Burns KE, Darwin KH. Pupylation: proteasomal targeting by a protein modifier in bacteria. Methods in molecular biology. 2012;832:151-60.

[15] Striebel F, Imkamp F, Ozcelik D, Weber-Ban E. Pupylation as a signal for proteasomal degradation in bacteria. Biochimica et biophysica acta. 2014;1843:103-13.

[16] Iyer LM, Burroughs AM, Aravind L. Unraveling the biochemistry and provenance of pupylation: a prokaryotic analog of ubiquitination. Biology direct. 2008;3:45.

[17] Burns KE, Darwin KH. Pupylation versus ubiquitylation: tagging for proteasome-dependent degradation. Cellular microbiology. 2010;12:424-31.

[18] Barandun J, Delley CL, Weber-Ban E. The pupylation pathway and its role in mycobacteria. BMC biology. 2012;10:95.

[19] Burns KE, Cerda-Maira FA, Wang T, Li H, Bishai WR, Darwin KH. "Depupylation" of prokaryotic ubiquitin-like protein from mycobacterial proteasome substrates. Molecular cell. 2010;39:821-7.

[20] Imkamp F, Rosenberger T, Striebel F, Keller PM, Amstutz B, Sander P, et al. Deletion of dop in Mycobacterium smegmatis abolishes pupylation of protein substrates in vivo. Molecular microbiology. 2010;75:744-54.

[21] Striebel F, Imkamp F, Sutter M, Steiner M, Mamedov A, Weber-Ban E. Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to substrates by distinct but homologous enzymes. Nature structural & molecular biology. 2009;16:647-51.

[22] Cerda-Maira FA, Pearce MJ, Fuortes M, Bishai WR, Hubbard SR, Darwin KH. Molecular analysis of the prokaryotic ubiquitin-like protein (Pup) conjugation pathway in Mycobacterium tuberculosis. Molecular microbiology. 2010;77:1123-35.

[23] Darwin KH, Lin G, Chen Z, Li H, Nathan CF. Characterization of a Mycobacterium tuberculosis proteasomal ATPase homologue. Molecular microbiology. 2005;55:561-71.

[24] Wolf S, Nagy I, Lupas A, Pfeifer G, Cejka Z, Muller SA, et al. Characterization of ARC, a divergent member of the AAA ATPase family from Rhodococcus erythropolis. Journal of molecular biology. 1998;277:13-25.

[25] Sutter M, Damberger FF, Imkamp F, Allain FH, Weber-Ban E. Prokaryotic ubiquitin-like protein (Pup) is coupled to substrates via the side chain of its C-terminal glutamate. Journal of the American Chemical Society. 2010;132:5610-2.

[26] Vijay-Kumar S, Bugg CE, Cook WJ. Structure of ubiquitin refined at 1.8 A resolution. Journal of molecular biology. 1987;194:531-44.

[27] Chen X, Solomon WC, Kang Y, Cerda-Maira F, Darwin KH, Walters KJ. Prokaryotic ubiquitin-like protein pup is intrinsically disordered. Journal of molecular biology. 2009;392:208-17.

[28] Liao S, Shang Q, Zhang X, Zhang J, Xu C, Tu X. Pup, a prokaryotic ubiquitin-like protein, is an intrinsically disordered protein. The Biochemical journal. 2009;422:207-15.

[29] Sutter M, Striebel F, Damberger FF, Allain FH, Weber-Ban E. A distinct structural region of the prokaryotic ubiquitin-like protein (Pup) is recognized by the N-terminal domain of the proteasomal ATPase Mpa. FEBS letters. 2009;583:3151-7.

[30] Tompa P. Intrinsically disordered proteins: a 10-year recap. Trends in biochemical sciences. 2012;37:509-16.

[31] Uversky VN, Dunker AK. Multiparametric analysis of intrinsically disordered proteins: looking at intrinsic disorder through compound eyes. Analytical chemistry. 2012;84:2096-104.

[32] Barandun J, Delley CL, Ban N, Weber-Ban E. Crystal structure of the complex between prokaryotic ubiquitin-like protein and its ligase PafA. Journal of the American Chemical Society. 2013;135:6794-7.

[33] Ozcelik D, Barandun J, Schmitz N, Sutter M, Guth E, Damberger FF, et al. Structures of Pup ligase PafA and depupylase Dop from the prokaryotic ubiquitin-like modification pathway. Nature communications. 2012;3:1014.

[34] Wang T, Darwin KH, Li H. Binding-induced folding of prokaryotic ubiquitin-like protein on the Mycobacterium proteasomal ATPase targets substrates for degradation. Nature structural & molecular biology. 2010;17:1352-7.

[35] Fuxreiter M, Simon I, Friedrich P, Tompa P. Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. Journal of molecular biology. 2004;338:1015-26.

[36] Maldonado AY, Burz DS, Reverdatto S, Shekhtman A. Fate of pup inside the Mycobacterium proteasome studied by in-cell NMR. PloS one. 2013;8:e74576.

[37] Ye K, Tu X, Zhang X, Shang Q, Liao S, Yu J, et al. Induced Folding Under Membrane Mimetic and Acidic Conditions Implies Undiscovered Biological Roles of Prokaryotic Ubiquitin-Like Protein Pup. Protein and peptide letters. 2016;23:756-62.

[38] Tompa P, Schad E, Tantos A, Kalmar L. Intrinsically disordered proteins: emerging interaction specialists. Current opinion in structural biology. 2015;35:49-59.

[39] Oldfield CJ, Dunker AK. Intrinsically disordered proteins and intrinsically disordered protein regions. Annual review of biochemistry. 2014;83:553-84.

[40] Dogan J, Gianni S, Jemth P. The binding mechanisms of intrinsically disordered proteins. Physical chemistry chemical physics : PCCP. 2014;16:6323-31.

[41] Espinoza-Fonseca LM. Reconciling binding mechanisms of intrinsically disordered proteins. Biochemical and biophysical research communications. 2009;382:479-82.

[42] Laederach J, Leodolter J, Warweg J, Weber-Ban E. Chaperone-Preoteases of Mycobacteria. The Molecular Chaperones Interaction Network in Protein Folding and Degradation: Springer; 2014. p. 419-44.

[43] Darwin KH, Ehrt S, Gutierrez-Ramos JC, Weich N, Nathan CF. The proteasome of Mycobacterium tuberculosis is required for resistance to nitric oxide. Science. 2003;302:1963-6.

[44] Gandotra S, Schnappinger D, Monteleone M, Hillen W, Ehrt S. In vivo gene silencing identifies the Mycobacterium tuberculosis proteasome as essential for the bacteria to persist in mice. Nature medicine. 2007;13:1515-20.

[45] Samanovic MI, Tu S, Novak O, Iyer LM, McAllister FE, Aravind L, et al. Proteasomal control of cytokinin synthesis protects Mycobacterium tuberculosis against nitric oxide. Molecular cell. 2015;57:984-94.

[46] Elharar Y, Roth Z, Hermelin I, Moon A, Peretz G, Shenkerman Y, et al. Survival of mycobacteria depends on proteasome-mediated amino acid recycling under nutrient limitation. The EMBO journal. 2014;33:1802-14.

[47] Fascellaro G, Petrera A, Lai ZW, Nanni P, Grossmann J, Burger S, et al. Comprehensive Proteomic Analysis of Nitrogen-Starved Mycobacterium smegmatis Deltapup Reveals the Impact of Pupylation on Nitrogen Stress Response. Journal of proteome research. 2016;15:2812-25.

[48] Compton CL, Fernandopulle MS, Nagari RT, Sello JK. Genetic and Proteomic Analyses of Pupylation in Streptomyces coelicolor. Journal of bacteriology. 2015;197:2747-53.

[49] Festa RA, McAllister F, Pearce MJ, Mintseris J, Burns KE, Gygi SP, et al. Prokaryotic ubiquitin-like protein (Pup) proteome of Mycobacterium tuberculosis [corrected]. PloS one. 2010;5:e8589.

[50] Poulsen C, Akhter Y, Jeon AH, Schmitt-Ulms G, Meyer HE, Stefanski A, et al. Proteome-wide identification of mycobacterial pupylation targets. Molecular systems biology. 2010;6:386.

[51] Watrous J, Burns K, Liu WT, Patel A, Hook V, Bafna V, et al. Expansion of the mycobacterial "PUPylome". Molecular bioSystems. 2010;6:376-85.

[52] Striebel F, Hunkeler M, Summer H, Weber-Ban E. The mycobacterial Mpa-proteasome unfolds and degrades pupylated substrates by engaging Pup's N-terminus. The EMBO journal. 2010;29:1262-71.

[53] Wang T, Li H, Lin G, Tang C, Li D, Nathan C, et al. Structural insights on the Mycobacterium tuberculosis proteasomal ATPase Mpa. Structure. 2009;17:1377-85.

[54] Delley CL, Laederach J, Ziemski M, Bolten M, Boehringer D, Weber-Ban E. Bacterial proteasome activator Bpa (Rv3780) is a novel ring-shaped interactor of the mycobacterial proteasome. PloS one. 2014;9:e114348.

[55] Jastrab JB, Wang T, Murphy JP, Bai L, Hu K, Merkx R, et al. An adenosine triphosphate-independent proteasome activator contributes to the virulence of Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences of the United States of America. 2015;112:E1763-72.

[56] Imkamp F, Ziemski M, Weber-Ban E. Pupylation-dependent and -independent proteasomal degradation in mycobacteria. Biomolecular concepts. 2015;6:285-301.

[57] Djuranovic S, Hartmann MD, Habeck M, Ursinus A, Zwickl P, Martin J, et al. Structure and activity of the N-terminal substrate recognition domains in proteasomal ATPases. Molecular cell. 2009;34:580-90.

[58] Beck F, Unverdorben P, Bohn S, Schweitzer A, Pfeifer G, Sakata E, et al. Near-atomic resolution structural model of the yeast 26S proteasome. Proceedings of the National Academy of Sciences of the United States of America. 2012;109:14870-5.

[59] Lander GC, Estrin E, Matyskiela ME, Bashore C, Nogales E, Martin A. Complete subunit architecture of the proteasome regulatory particle. Nature. 2012;482:186-91.

[60] Matyskiela ME, Lander GC, Martin A. Conformational switching of the 26S proteasome enables substrate degradation. Nature structural & molecular biology. 2013;20:781-8.

[61] Shoemaker BA, Portman JJ, Wolynes PG. Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proceedings of the National Academy of Sciences of the United States of America. 2000;97:8868-73.

[62] Inobe T, Fishbain S, Prakash S, Matouschek A. Defining the geometry of the two-component proteasome degron. Nat Chem Biol. 2011;7:161-7.

[63] Prakash S, Tian L, Ratliff KS, Lehotzky RE, Matouschek A. An unstructured initiation site is required for efficient proteasome-mediated degradation. Nature structural & molecular biology. 2004;11:830-7.

[64] Burns KE, Pearce MJ, Darwin KH. Prokaryotic ubiquitin-like protein provides a two-part degron to Mycobacterium proteasome substrates. Journal of bacteriology. 2010;192:2933-5.

[65] Lin G, Hu G, Tsu C, Kunes YZ, Li H, Dick L, et al. Mycobacterium tuberculosis prcBA genes encode a gated proteasome with broad oligopeptide specificity. Molecular microbiology. 2006;59:1405-16.

[66] Anandan T, Han J, Baun H, Nyayapathy S, Brown JT, Dial RL, et al. Phosphorylation regulates mycobacterial proteasome. Journal of microbiology. 2014;52:743-54.

[67] Barandun J, Damberger FF, Delley CL, Laederach J, Allain FH, Weber-Ban E. Prokaryotic ubiquitin-like protein remains intrinsically disordered when covalently attached to proteasomal target proteins. BMC structural biology. 2017;17:1.

[68] Imkamp F, Striebel F, Sutter M, Ozcelik D, Zimmermann N, Sander P, et al. Dop functions as a depupylase in the prokaryotic ubiquitin-like modification pathway. EMBO reports. 2010;11:791-7.

[69] Hothorn M, Wachter A, Gromes R, Stuwe T, Rausch T, Scheffzek K. Structural basis for the redox control of plant glutamate cysteine ligase. The Journal of biological chemistry. 2006;281:27557-65.

[70] Krishnaswamy PR, Pamiljans V, Meister A. Activated glutamate intermediate in the enzymatic synthesis of glutamine. The Journal of biological chemistry. 1960;235:PC39-40.

[71] Orlowski M, Meister A. Partial reactions catalyzed by -glutamylcysteine synthetase and evidence for an activated glutamate intermediate. The Journal of biological chemistry. 1971;246:7095-105.

[72] Liaw SH, Eisenberg D. Structural model for the reaction mechanism of glutamine synthetase, based on five crystal structures of enzyme-substrate complexes. Biochemistry. 1994;33:675-81.

[73] Guth E, Thommen M, Weber-Ban E. Mycobacterial ubiquitin-like protein ligase PafA follows a two-step reaction pathway with a phosphorylated pup intermediate. The Journal of biological chemistry. 2011;286:4412-9.

[74] Kuberl A, Franzel B, Eggeling L, Polen T, Wolters DA, Bott M. Pupylated proteins in Corynebacterium glutamicum revealed by MudPIT analysis. Proteomics. 2014;14:1531-42.

[75] Regev O, Korman M, Hecht N, Roth Z, Forer N, Zarivach R, et al. An Extended Loop of the Pup Ligase, PafA, Mediates Interaction with Protein Targets. Journal of molecular biology. 2016;428:4143-53.

[76] Regev O, Roth Z, Korman M, Khalaila I, Gur E. A kinetic model for the prevalence of mono- over poly-pupylation. The FEBS journal. 2015;282:4176-86.

[77] Chen X, Li C, Wang L, Liu Y, Li C, Zhang J. The Mechanism of Mycobacterium smegmatis PafA Self-Pupylation. PloS one. 2016;11:e0151021.

[78] Bolten M, Vahlensieck C, Lipp C, Leibundgut M, Ban N, Weber-Ban E. Depupylase Dop requires Inorganic Phosphate in the Active Site for Catalysis. The Journal of biological chemistry. 2017.
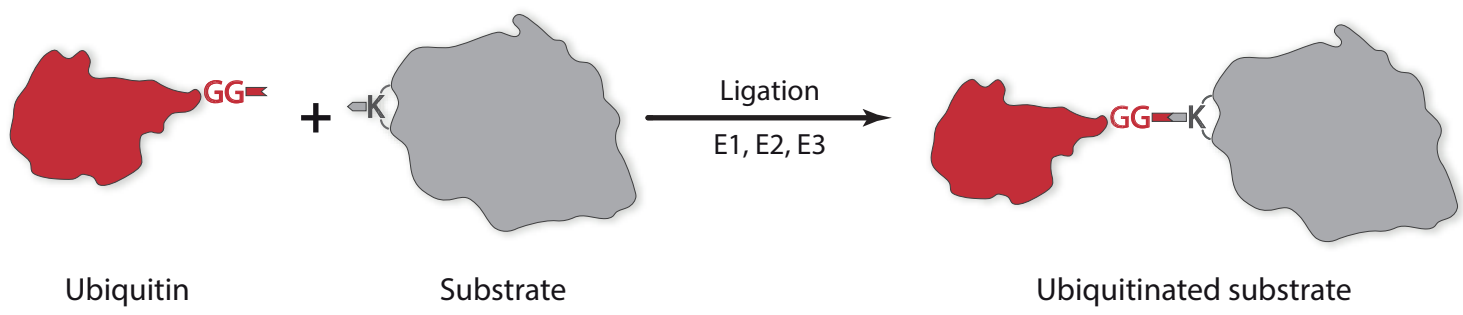
[79] Burns KE, McAllister FE, Schwerdtfeger C, Mintseris J, Cerda-Maira F, Noens EE, et al. Mycobacterium tuberculosis prokaryotic ubiquitin-like protein-deconjugating enzyme is an unusual aspartate amidase. The Journal of biological chemistry. 2012;287:37522-9.

[80] Eisenberg D, Gill HS, Pfluegl GM, Rotstein SH. Structure-function relationships of glutamine synthetases. Biochimica et biophysica acta. 2000;1477:122-45.

[81] Kuberl A, Polen T, Bott M. The pupylation machinery is involved in iron homeostasis by targeting the iron storage protein ferritin. Proceedings of the National Academy of Sciences of the United States of America. 2016;113:4806-11.
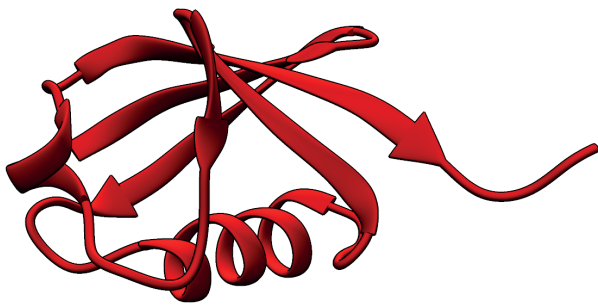
PUPYLATION

PupE + Substrate →(Ligation, Pup ligase PafA)→ Pupylated substrate
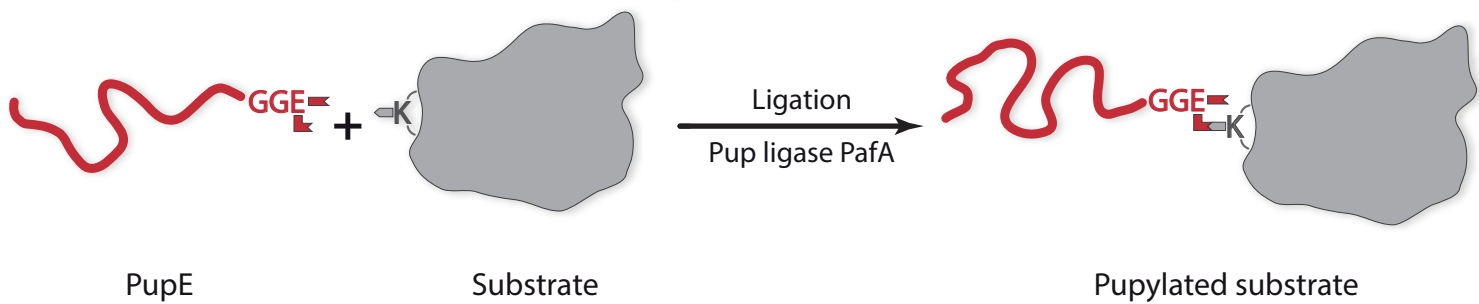
UBIQUITINATION

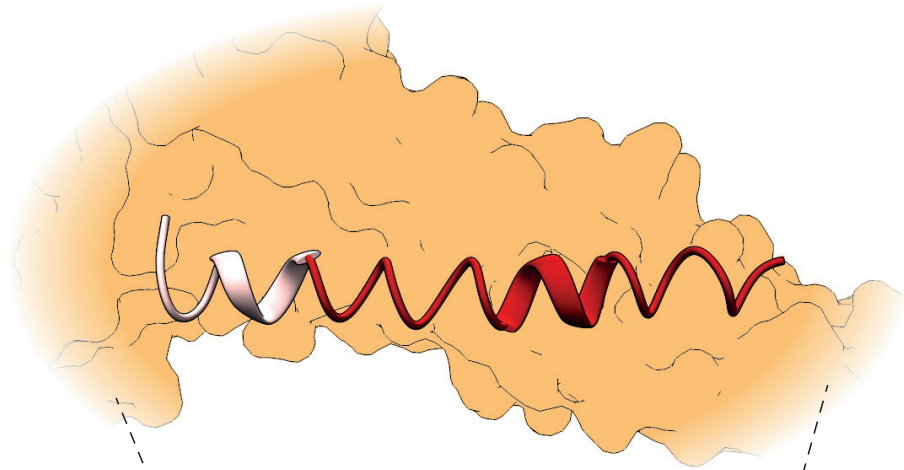Ubiquitin + Substrate →(Ligation, E1, E2, E3)→ Ubiquitinated substrate

Ubiquitin (Ub)
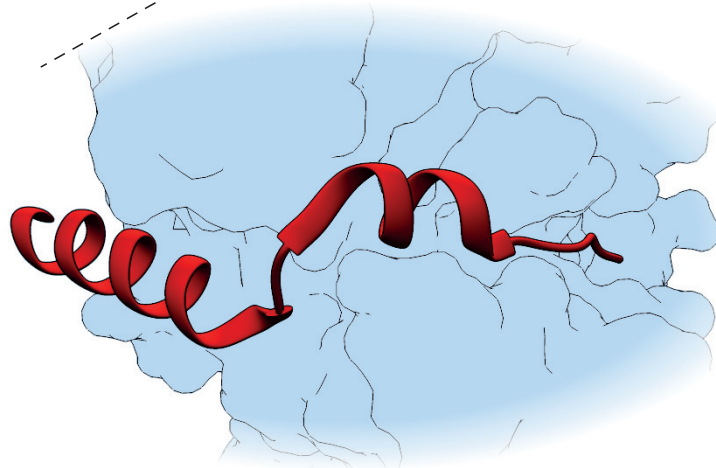
Prokaryotic Ubiquitin-like Protein (Pup)

A

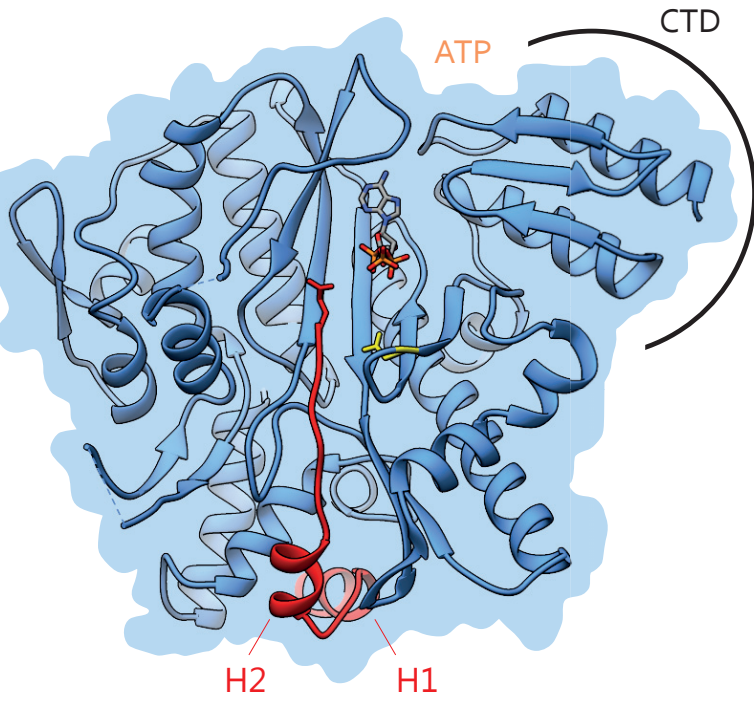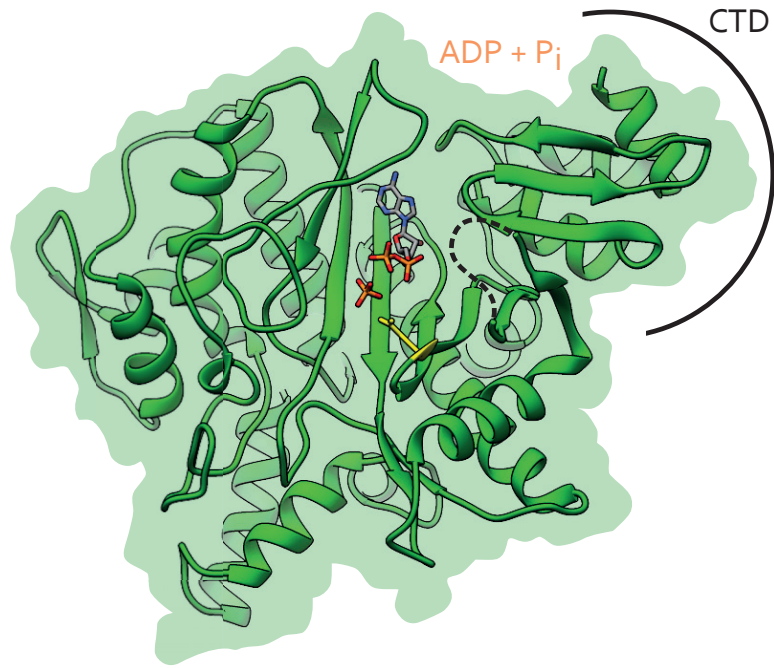MtbPup MAQEQTKRGGGGGDDDDIAGSTAAGQERR EKLTEETDDLLDEIDDVLEENAEDFVRAYVQKGG $\frac{Q}{E}$

B

	10	20	30	40	50	60	70
Mtb	MAQEQ - - - - - - - - TKRGGGGGDDDDIAGSTAAGQERREKLTEETDDLLDEIDDVLEENAEDFVRAYVQKGGQ
Msm	MAQEQ - - - - - - - - TKRGGGGGEDDDLPGASAAGQERREKLTEETDDLLDEIDDVLEENAEDFVRAYVQKGGQ
Cglu	MNAKQ - - - - - - - - TQIMGGGGRDEDNAEDSAQASGQVQINTEGVDSLLDEIDGLLENNAEEFVRSYVQKGGE
Rery	MAQEQ - - - - - - - - TQRAGGGEDDETTGGDGSAGQERREKLAAETDDLLDEIDDVLEENAEDFVRAYVQKGGQ
Aaur	MAAQEQQQP - - - - QSRETETEVDVPEAP - PAAPEAQASEATQGVDDLLDEIDGVLESNAEEFVRAFVQKGGQ
Scoe	MATKDTGGGQQKATRSTEEVEEQAQDAQASEDLKERQEKLSDDVDSVLDEIDDVLEENAEDFVRSFVQKGGE
Bado	MPQKFEQ - - - - MQSAEQKHDED - - - - ETIAQAGTQIDDTVDALDAVLDDIESVLESNAEEYVGSFVQKGGE
Tfus	MATKE - TGGQKHATRRNQEVEEI - - - - EVTTETSVRNEKLAEDVDDILDEIDEVLESNAEDFVRQFVQKGGE
Rsal	MAGQEQQSS - - - - SPREEHEVA - - DAPVPVPSSPQASAHTDGVDDLLDEIDGVLESNAEEFVRGFVQKGGQ
Stro	MATRD - SGGQSQTGRSQQGEEIEDVTTEASAEAAERHAEITEDVDDLLDEIDSVLEENAEEFVRGYVQKGGE
Krhi	MAQERIFGT - - - GSRREDEPDTPAPVD - PPVSGAAQAQRDMQGTDDLLAEIDGVLETNAEAFVKGFVQKGGQ
Mlut	MSQEQIRPR - - - PSGGDDETGAAA - - - - - GRAHVKPAAQDQGLDSLLDEIDSVLESNAEEFVKGFVQKGGQ
Acel	MPEKD - TGGQHRATRRTEEHDETIDEATATSDVQERREKLDADVDAILDEIDDVLEENAEEFVRSYIQKGGE

pupylated
substrate

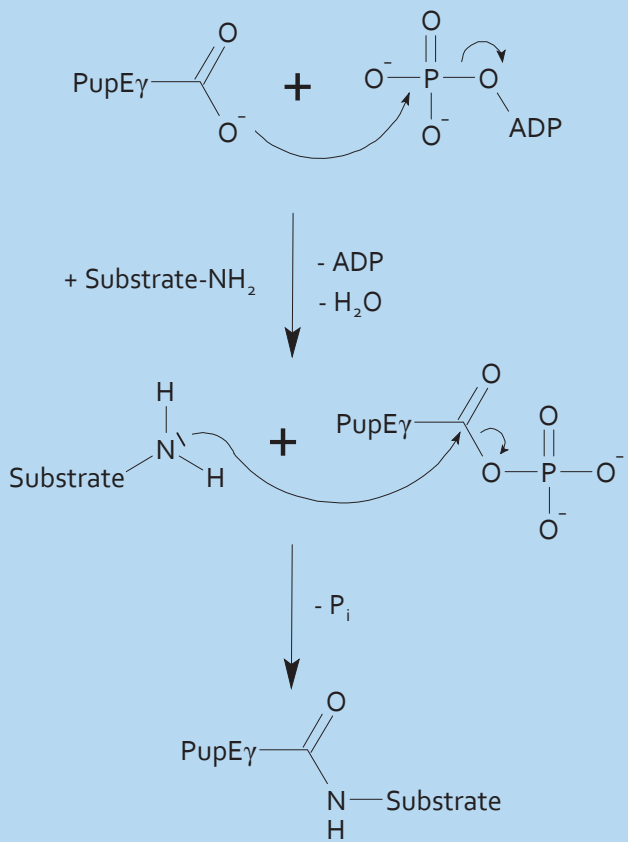Mpa

20S CP

ATP   ADP + P$_i$

recognition

threading
degradation

PafA

Dop

PUPYLATION

DEPUPYLATION

PupQ

Dop

deamidation

PupE + Substrate

PafA

pupylation

depupylation

Dop

Pupylated substrate

Mpa-20S CP

degradation