EMBO reports

scientific report

Reconstitution of the *Mycobacterium tuberculosis* pupylation pathway in *Escherichia coli*

Francisca A. Cerda-Maira[1], Fiona McAllister[2]*, Nadine J. Bode[1], Kristin E. Burns[1], Steven P. Gygi[2] & K. Heran Darwin[1+]*

[1]Department of Microbiology, New York University School of Medicine, New York, and [2]Department of Cell Biology, Harvard Medical School, Boston, Massachusetts, USA

Prokaryotic ubiquitin-like protein (Pup) is a post-translational modifier that attaches to more than 50 proteins in *Mycobacteria*. Proteasome accessory factor A (PafA) is responsible for Pup conjugation to substrates, but the manner in which proteins are selected for pupylation is unknown. To address this issue, we reconstituted the pupylation of model *Mycobacterium* proteasome substrates in *Escherichia coli*, which does not encode Pup or PafA. Surprisingly, Pup and PafA were sufficient to pupylate at least 51 *E. coli* proteins in addition to the mycobacterial proteins. These data suggest that pupylation signals are intrinsic to targeted proteins and might not require *Mycobacterium*-specific cofactors for substrate recognition by PafA *in vivo*.

Keywords: Pup; *Mycobacterium tuberculosis*; proteasome; pupylation

EMBO *reports* (2011) **12**, 863–870. doi:10.1038/embor.2011.109

INTRODUCTION

Protein-to-protein post-translational modifiers are essential for several aspects of eukaryotic biology (reviewed in Hochstrasser, 2000; Chen & Sun, 2009). In particular, the small protein ubiquitin (Ub) targets proteins for degradation by a compartmentalized protease called the proteasome (reviewed in Pickart & Cohen, 2004). Post-translational protein modifiers have recently been identified in bacteria (Pearce *et al*, 2008; Burns *et al*, 2009) and archaea (Humbard *et al*, 2010), but have not been characterized as extensively as their eukaryotic counterparts. In *Mycobacterium* species, a prokaryotic Ub-like protein (Pup) modifies proteins to target them for degradation by a bacterial proteasome (Pearce *et al*, 2008; Festa *et al*, 2010; Poulsen *et al*, 2010; Striebel *et al*, 2010; Watrous *et al*, 2010). Thus, Pup is functionally, if not biochemically, related to Ub.

[1]Department of Microbiology, New York University School of Medicine, 550 First Avenue, MSB 236, New York, New York 10016, USA
[2]Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, USA
*These authors contributed equally to this work.
[+]Corresponding author. Tel: +1 212 263 2624; Fax: +1 212 263 8276;
E-mail: heran.darwin@med.nyu.edu

Received 16 February 2011; revised 27 April 2011; accepted 18 May 2011; published online 8 July 2011

Pup can have either a carboxy-terminal glutamine (Pup_{Gln}) or glutamate (Pup_{Glu}) residue. *Mycobacteria* encode Pup_{Gln}, which must be deamidated to Pup_{Glu} by deamidase of Pup (Dop) before substrate conjugation (Striebel *et al*, 2009). Proteasome accessory factor A (PafA) attaches Pup_{Glu} to substrate lysines (Guth *et al*, 2011), and the resulting product is a covalently linked complex with an isopeptide bond between Pup and the substrate (Pearce *et al*, 2008; Burns *et al*, 2009). Similar to ubiquitylation, pupylation is reversible. Dop can hydrolyse the Pup–substrate bond, and thus Dop is a depupylase in addition to functioning as a deamidase (Burns *et al*, 2010a; Imkamp *et al*, 2010b). Depupylation is not required for protein degradation *in vitro* (Striebel *et al*, 2010), but it seems to enable Pup recycling (Cerda-Maira *et al*, 2010). In addition, depupylation might facilitate proteasomal degradation of certain proteins *in vivo* (Burns *et al*, 2010a).

In eukaryotes, it is estimated that more than 1,000 E3 ligases determine which proteins are targeted for ubiquitylation (reviewed in Deshaies & Joazeiro, 2009; Rotin & Kumar, 2009). By contrast, Pup seems to require a single conjugating enzyme for all pupylation in *Mycobacterium tuberculosis* (*Mtb*; Pearce *et al*, 2008; Festa *et al*, 2010). In *Mtb*, pupylation favours 1 of 8 lysines in the proteasome substrate malonyl Co-A acyl carrier protein transacylase (FabD; Pearce *et al*, 2008), and 1 of 23 lysines in myo-inositol-1-phosphate synthetase (Ino1; Burns *et al*, 2009; Festa *et al*, 2010; Watrous *et al*, 2010). *In vitro*, PafA is sufficient to pupylate FabD with Pup_{Glu}, although at least two Pup∼FabD species are produced in this system (Striebel *et al*, 2009).

To understand how proteins are selected for pupylation, we sought to develop a system using an organism that does not normally produce Pup, PafA or Dop. We reconstituted pupylation of *Mtb* proteasome substrates in a laboratory *Escherichia coli* K-12 strain. Unexpectedly, this revealed that pupylation does not require specificity factors that are functionally related to E3 ligases in eukaryotes.

RESULTS AND DISCUSSION

Pupylation of *Mtb* proteasome substrates in *E. coli*

We have previously identified targets of pupylation and proteasomal degradation in *Mtb*, including FabD, ketopantoate hydroxymethyltransferase (PanB) and Ino1 (Pearce *et al*, 2006; Burns

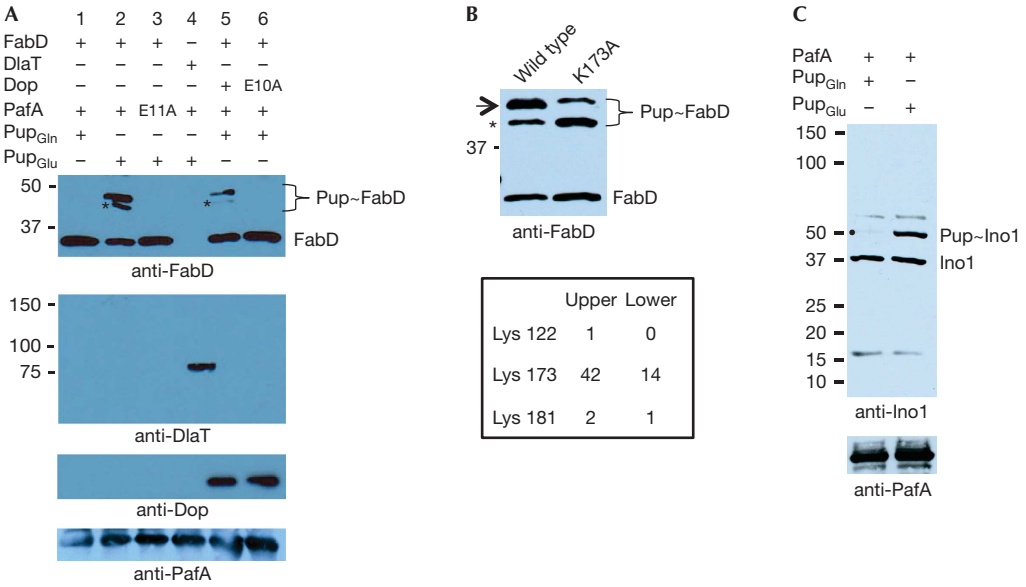

Fig 1 | Pupylation of *Mtb* proteasome substrates in *Escherichia coli*. (**A**) Immunoblot analysis of total cell lysates of *E. coli* harbouring plasmids encoding components of the pupylation system and test substrates. The recombinant proteins produced in these strains are listed on the left. E11A, glutamate 11 to alanine; E10A, glutamate 10 to alanine. These glutamates are predicted to coordinate Mg^{2+}. Asterisks indicate the lower-molecular-weight Pup ~ FabD species similar to that observed *in vitro* (Striebel *et al*, 2009). The polyclonal rabbit antibody used to detect a relevant protein is listed below each panel. (**B**) Top: mutation of Lys 173 in FabD changed the pupylation pattern of FabD. Arrow indicates Pup ~ FabD of the molecular weight observed in mycobacteria. Bottom: summary of MS/MS analysis of wild-type FabD purified from *E. coli*. Numbers represent abundance on the basis of spectral counting. (**C**) The *Mtb* proteasome substrate Ino1 was pupylated in *E. coli* expressing pup_{Glu} and *pafA*. Dot indicates a faint cross-reactive protein in the pup_{Gln} lane (probably PafA–His_6) that co-migrates with Ino1–His_6. For all panels, all proteins, except Pup, have carboxy-terminal His_6 tags. Induction of expression and immunoblot analysis are described in the Methods section. All data are representative of at least two independent experiments. DlaT, dihydrolipoamide acyltransferase; Dop, deamidase of Pup; Ino1, myo-inositol-1-phosphate synthetase; MS/MS, tandem mass spectrometry; PafA, proteasome accessory factor A; Pup, prokaryotic ubiquitin-like protein.

et al, 2010b). Although FabD and PanB can be pupylated *in vitro* (Striebel *et al*, 2009), reconstitution of pupylation in *E. coli* could enable more rapid characterization of Pup, PafA, Dop and their substrates. We expressed pup_{Gln} or pup_{Glu} and *pafA*–his_6 from pET24b(+) using the T7 polymerase system and *fabD*–his_6 from the arabinose-inducible *araBAD* promoter on the compatible plasmid pBAD33 (see supplementary Table S1 online for all plasmids and strains). We also cloned *Mtb* dihydrolipoamide acyltransferase (*dlaT*) into pBAD33; DlaT is not a known target of pupylation or proteasomal degradation in mycobacteria (Pearce *et al*, 2006, 2008). Finally, *dop*–his_6 was expressed from a T7 promoter in a third plasmid, pETDuet-1. As pET24b(+) and pETDuet-1 have the same replication origin, we monitored gene-product synthesis by immunoblotting for all relevant plasmid-encoded proteins produced by these strains.

Pup_{Gln} and PafA could not pupylate FabD in the absence of Dop (Fig 1A, lane 1), presumably because Pup_{Gln} was not deamidated. By contrast, pupylation of FabD was observed after induction of pup_{Glu} and *pafA* (Fig 1A, lane 2). We observed two pupylated species of FabD: a prominent band of the expected molecular weight of approximately 48 kDa and a less-abundant species with a molecular weight slightly below. A point mutation that inactivates PafA in *Mtb* (Cerda-Maira *et al*, 2010) also abrogated FabD pupylation in *E. coli* (Fig 1A, lane 3). We also tested whether *Mtb* DlaT—which is not known to be pupylated in mycobacteria—could be pupylated in *E. coli*. DlaT–His_6 was not pupylated in *E. coli* under the tested conditions (Fig 1A, lane 4). Expression of pup_{Gln} with *dop* and *pafA* resulted in FabD pupylation (Fig 1A, lane 5). The expression of *dop* with an inactivating mutation did not produce Pup ~ FabD (Fig 1A, lane 6), which is consistent with reports from other systems (Cerda-Maira *et al*, 2010; Imkamp *et al*, 2010a).

A previous study reported that two Pup ~ FabD species were observed when pupylation was reconstituted *in vitro*; it was speculated that modification of a different residue resulted in the faster-migrating Pup ~ FabD (Striebel *et al*, 2009). In *Mtb*, FabD is pupylated on Lys 173, and mutagenesis of this lysine stabilizes the protein (Pearce *et al*, 2008). Although it is possible that other FabD residues could be targeted for pupylation in *Mycobacteria*, these secondary sites must be used rarely, because we did not observe pupylated $FabD_{K173A}$. To determine whether Lys 173 was a target for pupylation in *E. coli*, we expressed $fabD_{K173A}$ in pupylation-competent *E. coli*. This mutation reduced the level of the slower-migrating Pup ~ FabD species (Fig 1B, upper panel, arrow) and increased the relative abundance of the faster-migrating species (Fig 1B, upper panel, asterisk). This suggests that in the absence of the preferred residue, PafA pupylated the next most accessible lysine. Similarly to in the Ub system, mutating a preferred lysine can result in another residue in the substrate being ubiquitylated.

scientific report

As the slower-migrating species did not disappear, we hypothesized that a non-Lys 173-modified FabD also migrated at this position. We purified Pup ~ FabD from *E. coli* and analysed both the higher- and lower-molecular-weight species by tandem mass spectrometry (MS/MS). FabD was identified with 96% sequence coverage in the two species. Lys 173 was the predominant modified residue in FabD purified from *E. coli* (Fig 1B, lower panel). Lys 122 and Lys 181 were also identified as pupylation targets. Two of the three pupylated species were found in both the lower and upper bands, probably due to cross-contamination of the closely migrating proteins. Thus, our data show that at least three lysines in *Mtb* FabD can be pupylated in *E. coli*. It also seems that only one residue per FabD polypeptide was pupylated at one time, because we did not observe higher-molecular-weight species (Fig 1A).

We speculate that the selection of FabD Lys 173 for pupylation in mycobacteria might be partly due to its accessibility to Pup and PafA. FabD is part of the multi-enzyme fatty acid synthesis II (FASII) pathway and has eight surface-exposed lysines (Kremer *et al*, 2001; Ghadbane *et al*, 2007). We propose that other proteins or macromolecules mask lysines or binding domains in FabD, thus preventing their pupylation under typical culture conditions in mycobacteria. For example, other FASII enzymes might prevent PafA from having access to most of the lysines in FabD, and Lys 173 is therefore preferred because it is the most exposed. In addition, Lys 173 might not always be accessible, which could determine the degree of FabD pupylation. This might suggest that pupylation is stochastic. These hypotheses could be tested in *E. coli* by examining FabD pupylation in the context of other *Mtb* FASII enzymes. Interestingly, every protein encoded in the *Mycobacterium fabD*/FASII operon is a target for pupylation, although they are not all robust proteasome substrates under routine culture conditions (Festa *et al*, 2010; Watrous *et al*, 2010).

We also examined Ino1, a target for pupylation in both *Mycobacterium smegmatis* (*Msm*) and *Mtb* (Festa *et al*, 2010; Watrous *et al*, 2010). Expression of *Mtb ino1*, *pafA* and pup_{Glu}, but not pup_{Gln}, resulted in pupylated Ino1 in *E. coli* (Fig 1C). Only one pupylated species of Ino1 was observed, similar to that observed in *Mycobacteria* (Burns *et al*, 2010b). MS/MS analyses of Ino1 from *Mtb* and *Msm* identified a single, specific lysine for pupylation (Burns *et al*, 2009; Festa *et al*, 2010; Watrous *et al*, 2010). Although it is possible that different lysines are pupylated when Ino1 is produced in *E. coli*, it seems likely that the same residue is modified in all bacterial species.

Pupylation of *E. coli* proteins

We next wondered if any *E. coli* proteins could be pupylated by this reconstituted system. Surprisingly, several proteins were pupylated in *E. coli* producing Pup_{Glu} and PafA (Fig 2A), or Pup_{Gln}, Dop and PafA (Fig 2B). Mutations that abrogate PafA or Dop activity in mycobacteria (Cerda-Maira *et al*, 2010) also inactivated function in *E. coli* (Fig 2A,B). To determine when pupylation was maximal in the reconstituted *E. coli* system, we examined the appearance of the pupylome over time. Pupylation in *E. coli* was apparent within 20 min of pup_{Glu}–*pafA* induction and reached a plateau at around 100 min (Fig 2C, top panel). Under the conditions tested, the viability of pupylation-competent *E. coli* was not different from that of *E. coli* expressing pup_{Gln}–*pafA* (Fig 2C, lower panel).

We next tested if other mutations in PafA or Dop affected pupylation in *E. coli*. Modelling and mutational analysis of PafA and Dop predict that the active sites of these proteins are at the amino-termini and are similar to those found in glutamine synthetases (Iyer *et al*, 2008; Cerda-Maira *et al*, 2010). By contrast, the C-termini do not resemble any protein of known function. We previously showed that the deletion of 38 residues from the C-terminus of PafA (PafA414) resulted in no pupylation in *Mtb* (Cerda-Maira *et al*, 2010). This mutation also eliminated pupylation in *E. coli* (Fig 2D, left panels). We therefore tested smaller C-terminal deletions of PafA, as well as four truncations of Dop (supplementary Table S1 online). All but one of the truncated proteins were detected in *E. coli*; the largest deletion in *dop* resulted in no protein, suggesting that this mutant protein was unstable or that the use of two incompatible plasmids affected the production of recombinant protein (Fig 2D, top right panel). None of the truncated proteins catalysed pupylation in *E. coli* (Fig 2D, lower panels). Most importantly, the truncated forms of *pafA* and *dop* were unstable and failed to complement transposon mutations in *Mtb pafA* or *dop*, respectively (Fig 2E).

Taken together, these results suggest that the C-termini of PafA and Dop are required for the folding, stability or activity of these enzymes. This *E. coli* system also demonstrated that additional *Mycobacterium*-specific cofactors are not required for pupylation *in vivo*. Importantly, these results demonstrate that *E. coli* could be used as a surrogate system in which to study *Mtb* pupylation.

Identification of the *E. coli* pupylome

We hypothesized that the *E. coli* pupylome could provide insight into how proteins are selected for pupylation. We affinity purified pupylated proteins from *E. coli* expressing pup_{Glu} and *pafA* (see Methods section; Festa *et al*, 2010). Incubation of the *E. coli* pupylome with *Mtb* Dop resulted in the disappearance of the Pup signal when assayed by immunoblotting, whereas overall protein levels were unchanged (Fig 3A). This is consistent with the idea that Pup formed isopeptide linkages with *E. coli* proteins, similarly to in mycobacteria (Festa *et al*, 2010; Watrous *et al*, 2010; Burns *et al*, 2010a).

Proteomic analysis identified over 400 proteins in the *E. coli* pupylome, 51 of which had one or more lysines linked to Pup_{Glu} (Table 1; supplementary Table S2 online). Several *E. coli* proteins (AdhC, adenylate kinase (Adk), ProS, GroL and SodB) have homologues in *Mtb* that are also targets of pupylation (Festa *et al*, 2010). However, when the equivalent lysines were present they were not modified in both bacterial species. Furthermore, comparison of the peptides did not identify a consensus sequence surrounding the modified lysines.

We next tested whether *E. coli* proteins could be pupylated in *Mycobacteria*. We chose two proteins: phosphoenolpyruvate-protein phosphotransferase I (PtsI)—one of the most highly represented proteins in the *E. coli* pupylome (supplementary Table S3 online)—and Adk, which has an *Mtb* homologue that is a target for pupylation (Festa *et al*, 2010). In addition, *E. coli* Adk had four lysines that could be pupylated (Table 1). Before examining pupylation in *Mycobacteria*, we cloned *adk* and *ptsI* with his_6 tags into pBAD33 to assess whether these recombinant proteins could still be pupylated in *E. coli*. Adk–His_6 seemed to be pupylated when purified from *E. coli* expressing pup_{Glu}–*pafA*, but not pup_{Gln}–*pafA* (Fig 3B). Additional higher-molecular-weight

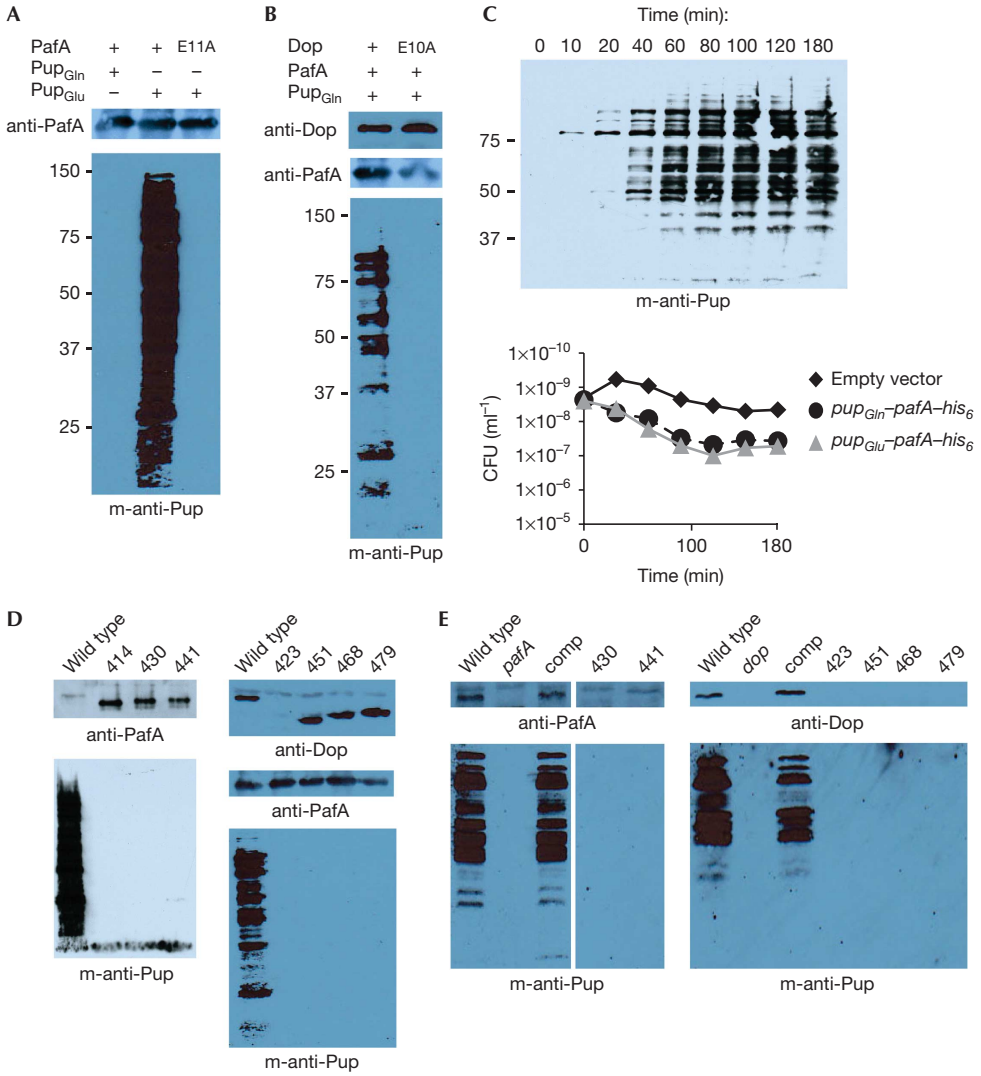

Fig 2 | Pupylation of *E. coli* proteins. (**A**,**B**) Immunoblot analysis of total cell lysates of *E. coli* strains harbouring plasmids encoding various components of the pupylation system. Mouse monoclonal antibodies to Pup (m-anti-Pup) were used to detect the *E. coli* pupylome. (**C**) Detection of pupylation over time in *E. coli* expressing pup_{Glu}–*pafA* (top). Pupylation of *E. coli* proteins was not detrimental to bacterial viability up to 3 h after induction of *pup–pafA* expression (bottom). The addition of IPTG to cultures of strains containing the empty vector typically slows growth but does not kill the bacteria; therefore, we also enumerated CFUs from a culture of the same *E. coli* strain carrying empty pET24b(+). (**D**) Truncating PafA and Dop abrogated pupylation in *E. coli*. Left: immunoblot analysis of total cell lysates of *E. coli* expressing pup_{Glu} and wild-type or mutant *pafA*. Wild-type PafA comprises 452 amino acids plus a His_6 tag. 414, 430 and 441 indicate the last residue of the truncated PafA protein. Right: immunoblot analysis of total cell lysates of *E. coli* expressing pup_{Gln}–*pafA* and wild-type or mutant *dop*. Wild-type Dop comprises 505 amino acids plus a His_6 tag. 423, 451, 468 and 479 indicate the last residue of the truncated Dop protein. None of the truncated proteins was epitope tagged. (**E**) Immunoblot analysis of total cell lysates of *Mtb* expressing different *pafA* or *dop* truncations. Induction of expression, lysate preparation and immunoblot analysis are described in the Methods section. These data are representative of two independent experiments. comp, complemented; CFU, colony-forming unit; Dop, deamidase of Pup; IPTG, isopropyl-β-D-thiogalactopyranoside; *Mtb*, *Mycobacterium tuberculosis*; Myc; PafA, proteasome accessory factor A; Pup, prokaryotic ubiquitin-like protein; WT, wild type.

bands were also detected using antibodies to Pup, which might indicate that Adk–His_6 was polypupylated in *E. coli*. It is also possible that these larger Pup-reactive species are different pupylated proteins that copurified with Adk–His_6. PtsI was also pupylated in *E. coli* expressing pup_{Glu}–*pafA*, as shown by immunoblotting with antibodies to His_5 or Pup (Fig 3C). Additional higher-molecular-weight species were also detected with both antibodies, suggesting that PtsI could have more than one Pup attached. Taken together, these results suggest that addition of the His_6 tag did not hinder pupylation of these proteins in *E. coli*.

Finally, we expressed *E. coli adk–his$_6$* and *ptsI–his$_6$* in *Msm* (supplementary Table S1 online) and purified the recombinant proteins using nickel affinity chromatography. Although Adk–His_6 was robustly produced in *Msm*, we could not detect a pupylated

scientific report

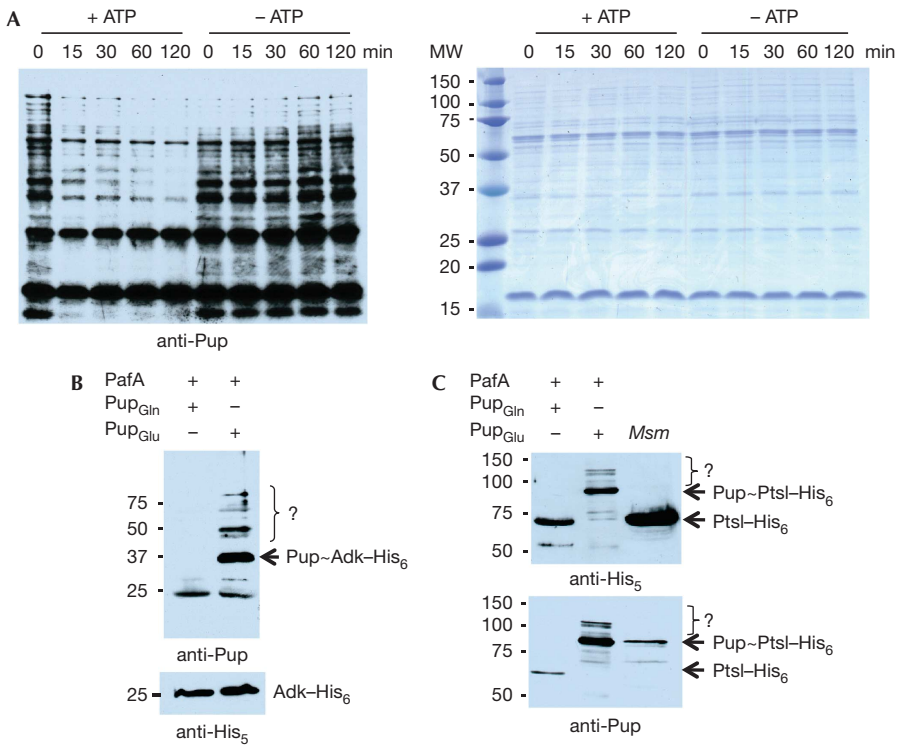

Fig 3 | Analysis of the *Escherichia coli* pupylome. (**A**) The purified *E. coli* pupylome was sensitive to depupylation by *Mtb* Dop–His_6. Samples were analysed by 12% SDS–PAGE and immunoblotting against Pup (left) or Coomassie blue staining (right). (**B**) *E. coli* Adk–His_6 was pupylated in *E. coli* expressing pup_{Glu}–*pafA*–his_6. Ni-NTA-purified Adk–His_6 was analysed by 12% SDS–PAGE and immunoblotting with polyclonal antibodies to Pup (top) or His_5 (bottom). Question mark indicates possible species of Adk–His_6 with several Pup molecules attached, or other pupylated proteins that copurified with Adk–His_6. (**C**) *E. coli* PtsI–His_6 was pupylated in *Msm*. The recombinant protein was purified from *Msm* using Ni-NTA chromatography and analysed by 10% SDS–PAGE. Molecular-weight markers and 15 μl of each sample were loaded onto the same gel twice. Proteins were transferred to a nitrocellulose membrane and the membrane was cut in half for immunoblotting with antibodies to His_5 (upper panel) or Pup (lower panel). Question mark indicates possible species of PtsI–His_6 with several Pup molecules attached. Adk, adenylate kinase; Dop, deamidase of Pup; *Msm*, *Mycobacterium smegmatis*; *Mtb*, *Mycobacterium tuberculosis*; MW, molecular weight; Ni-NTA; nickel-nitrilotriacetic acid; PafA, proteasome accessory factor A; PtsI, phosphoenolpyruvate-protein phosphotransferase I; Pup, prokaryotic ubiquitin-like protein; SDS–PAGE, SDS–polyacrylamide gel electrophoresis.

species (data not shown). Interestingly, *Mtb* Adk—a target of pupylation as detected by MS/MS in *Mtb*—could not be purified in the pupylated form from *Mycobacteria* (Festa *et al*, 2010). In contrast to Adk, PtsI–His_6 was pupylated in *Msm*, as detected with antibodies to Pup (Fig 3C, lower panel, last lane). We detected Pup∼PtsI–His_6 using antibodies to His_5 only if the film was overexposed (not shown); however, we have previously observed that His_5 antibodies are not always sensitive enough to detect pupylated proteins (Festa *et al*, 2010). PtsI does not have a homologue in mycobacteria, and thus a completely foreign protein could be targeted by the native *Mycobacterium* pupylation system.

Perhaps it is not surprising that *E. coli* proteins could be pupylated, as a motif is lacking in over 50 pupylated proteins purified from either *Mtb* or *Msm* (Festa *et al*, 2010; Watrous *et al*, 2010). Although pupylation seems to be widespread, not just any protein or lysine can be pupylated. For example, *Mtb* DlaT has never been demonstrated to be a target of pupylation, despite having 29 lysines (Fig 1; Pearce *et al*, 2008). This suggests that proteins require additional chemical signatures, such as localized changes in charge, phosphorylation, oxidation, misfolding, or even other proteins such as chaperones, to target them for pupylation.

The reconstitution of pupylation in *E. coli* is important because it will simplify the structure–function analysis of enzymes of a less-genetically-tractable organism. Researchers in the *Mtb* field can now easily test if a protein is a target of pupylation. It is also striking that so many proteins can be pupylated in *E. coli*, in comparison to ubiquitylation in eukaryotes. Several Ub ligases determine which proteins are modified at a particular place and time in the cell. Despite this fact, there is precedent for the modification of proteins by heterologous systems. Proteins injected from pathogenic Gram-negative bacteria can be ubiquitylated in mammalian cells (Knodler *et al*, 2009; Patel *et al*, 2009). The Ub-like protein interferon-stimulated gene 15 targets viral as well as host proteins (Durfee *et al*, 2010). Thus, protein modifiers might sometimes rely on overproduction to increase their chances of finding their targets. Pupylation is enzymatically, if not functionally, distinct from ubiquitylation, and it now seems that new differences will be identified in the manner in which substrates are selected for modification.

Table 1 | *Escherichia coli* proteins with pupylated lysines*

Locus	Gene	Position of modified lysines (protein MW)	Locus	Gene	Position of modified lysines (protein MW)
b0014	*dnaK*	635, 637 (638)	b2185	*rplY*	34 (94)
b0026	*ileS*	814 (938)	b2400	*gltX*	44 (471)
b0115	*aceF*	161 (630)	b2416	*ptsI*	174 (575)
b0118	*acnB*	77 (865)	b2507	*guaA*	437 (525)
b0166	*dapD*	100 (274)	b2614	*grpE*	43, 66 (197)
b0194	*proS*	426, 495, 571 (572)	b2687	*luxS*	163 (171)
b0356	*adhC*	13 (369)	b2891	*prfB*	162 (365)
b0474	*adk*	47, 50, 136, 141 (214)	b2935	*tktA*	316, 347 (663)
b0492	*ybbN*	184, 268 (296)	b2954	*rdgB*	3 (197)
b0525	*ppiB*	60 (164)	b3011	*yqhD*	43 (387)
b0578	*nfnB*	21, 62 (217)	b3296	*rpsD*	156, 167 (206)
b0812	*dps*	105 (167)	b3301	*rplO*	141 (144)
b0965	*yccU*	97 (164)	b3341	*rpsG*	56, 131 (179)
b1059	*solA*	357 (372)	b3559	*glyS*	624 (689)
b1107	*nagZ*	283 (341)	b3640	*dut*	15 (151)
b1136	*icd*	4, 273 (416)	b3644	*yicC*	86 (287)
b1215	*kdsA*	79 (284)	b3708	*tnaA*	161 (476)
b1656	*sodB*	44 (193)	b3866	*yihI*	158 (169)
b1667	*ydhR*	60, 74 (101)	b3924	*fpr*	135 (248)
b1749	*xthA*	141 (268)	b3960	*argH*	95 (457)
b1779	*gapA*	249 (331)	b4113	*basR*	70, 164 (222)
b1823	*cspC*	9 (69)	b4143	*groL*	65, 168, 321 (548)
b1854	*pykA*	60, 93 (480)	b4200	*rpsF*	35 (131)
b1864	*yebC*	192 (246)	b4207	*fklB*	146 (259)
b2040	*rfbD*	245 (299)	b4280	*yjhC*	79 (377)
b2146	*yeiT*	364 (412)			

MW, molecular weight.
*Four additional proteins with pupylated peptides but no locus identifiers were observed by tandem mass spectrometry (supplementary Table S2 online).

METHODS

Strains, plasmids, primers and culture conditions. Strains and plasmids are listed in supplementary Table S1 online. Primer sequences are available on request. *E. coli* strains were grown in Luria–Bertani broth or on Luria–Bertani-agar plates (Difco). *Mycobacteria* strains were grown in Middlebrook 7H9 broth or on 7H10 agar (Difco) with supplementation (0.2% glycerol, 0.05% Tween-80, 0.5% bovine serum albumin, 0.2% dextrose and 0.085% sodium chloride for broth; or BBL Middlebrook OADC for agar) when necessary. Antibiotic concentrations were as follows for *E. coli* and *Mycobacteria*, respectively: 150 or 50 μg ml^{-1} hygromycin and 100 or 50 μg ml^{-1} kanamycin. Chloramphenicol was used at 25 μg ml^{-1} for *E. coli*. Isopropyl-β-D-thiogalactopyranoside was used at 0.1 mM.

pup and *pafA* derivatives were sequentially cloned into pET24b(+). Briefly, *pafA* (wild-type or mutant alleles) were first cloned into the *Nde*I–*Hind*III sites of pET24b(+). *pup* was amplified from pET24b + *pup*-STOP, digested with *Bgl*II and *Nhe*I, and cloned into the *Bgl*II and *Xba*I sites of the pET24b(+)-*pafA* plasmids. The *dop* derivatives were cloned into pETDuet-1 multiple cloning site 2 *Nde*I and *Pac*I sites. *dlaT, ino1, adk, ptsI and fabD* were cloned into the *Kpn*I or *Xba*I and *Sph*I sites of pBAD33. For *Mtb* complementation analysis, *pafA* and *dop* alleles were cloned with their native promoters into pMV306, using *Kpn*I and *Cla*I, or *Xba*I and *HindIII*, respectively. For expression in *Msm*, *adk* and *ptsI* were cloned into pMN402 using *Pac*I and *Bam*HI. All plasmids were sequenced to confirm the veracity of the cloned sequences (GENEWIZ Inc.). Transformation of bacteria was performed as described previously (Sambrook *et al*, 1989; Hatfull & Jacobs, 2000).

For analysis of *Mtb* lysates, 5 ml cultures were grown to an optical density at 580 nm of approximately 1.0–1.5. Bacteria were collected and processed as described previously (Cerda-Maira

et al, 2010). Total cell lysates were analysed by immunoblotting using antibodies to PafA, Dop or Pup.

SDS–PAGE and immunoblot analysis. For immunoblotting analysis, proteins were separated on 7% (anti-PafA and anti-Dop), 10 or 12% (anti-His_5), or 15% (anti-Pup) SDS–polyacrylamide gel electrophoresis (PAGE) gels (Ausubel *et al*, 2002). Proteins were transferred onto 0.22 μm nitrocellulose (Schleicher and Schuell) using a semi-dry transfer system (Bio-Rad), and incubated with antibodies to PafA-His_6, Dop-His_6, FabD-His_6, Ino1-His_6, Pup-His_6 or His_5 (Qiagen), as described previously (Festa *et al*, 2007, 2010; Pearce *et al*, 2008; Cerda-Maira *et al*, 2010). Horseradish peroxidase-coupled rabbit secondary antibodies were used according to the manufacturer's instructions (Thermo Scientific). Horseradish peroxidase was detected using either the SuperSignal West Pico or West Femto Chemiluminescent Substrate (Thermo Scientific).

Pupylation in *E. coli* and *Msm*. To examine the pupylation of proteins in *E. coli*, strains were incubated at 37 °C with shaking until an optical density of 0.6 at 600 nm was reached. When strains with pBAD33 derivatives were used, arabinose was added to a final concentration of 0.02–0.2% w/v and incubated for 1 h. Cultures were then induced with 0.1 mM IPTG and incubated for one more hour. If *dop* activity was tested, cultures were incubated for 2 h. A volume of 0.5 ml of the induced cultures was collected and bacteria were isolated by centrifugation for 1 min at top speed in microfuge tubes. Bacteria were resuspended in 400 μl of lysis buffer, mixed with 4 × sample buffer and boiled for 10 min before analysis by SDS–PAGE and immunoblotting.

To determine whether *E. coli* proteins were pupylated in *Msm*, 100 ml saturated overnight cultures of *Msm* producing recombinant protein were processed, as described previously for *Mtb* (Festa *et al*, 2010). Nickel-nitrilotriacetic acid purification (Qiagen) was used to isolate His_6-tagged proteins.

Proteomic analysis. A 250-ml culture of EHD1295 (supplementary Table S1 online) was grown as described above. Cells were collected by centrifugation, resuspended in denaturing lysis buffer B (The QIAexpressionist manual), and TAP was performed as described in detail elsewhere (Festa *et al*, 2010). Purified proteins were visualized on a 12% SDS–PAGE gel and excised from the gel for MS/MS analysis. The gel slice was divided into 10 bands, each of which was cut into cubes of approximately 1 mm and transferred into 1.5 ml Eppendorf tubes. In-gel digestion was done as described previously (Li *et al*, 2007). Peptides were eluted from the StageTips (Rappsilber *et al*, 2003) using 40% acetonitrile and 5% formic acid, followed by 80% acetonitrile and 5% formic acid into glass inserts, and dried in a Speed-Vac. The resulting peptides were analysed as described previously (Festa *et al*, 2010). Protein hits were filtered at the peptide level to a 0.1% false discovery rate using the target-decoy strategy (Elias & Gygi, 2007) using in-house software and filtering on the basis of ΔCn, XCorr, p.p.m., charge state and peptides per protein. Proteins were also filtered at the protein level to a 1% false discovery rate using linear discriminate analysis based on X_{corr}, ΔCn, precursor mass error and charge state, as described previously (Huttlin *et al*, 2010). For all pupylome MS/MS data, see supplementary Tables S2 and S3 online.

To identify the pupylated sites in *Mtb* FabD produced in *E. coli*, FabD–His_6 was purified under native conditions from 250 ml EHD1288 cultured as described above. Protein was separated on 10% SDS–PAGE and the Pup ~ FabD bands were excised for MS/MS analysis, as described previously.

Depupylation assay. The depupylation reaction was done essentially as described previously (Burns *et al*, 2010a). Briefly, 15 μg of Ni-NTA-purified proteins were incubated with 2.5 mM ATP, 20 mM $MgCl_2$, 50 mM NaCl, 1 mM dithiothreitol and 2.5 μg *Mtb* Dop–His_6 in 50 mM Tris–HCl, pH 8, in a final volume of 100 μl. Dop was purified using Ni-NTA from *Msm*, as described previously (Burns *et al*, 2010a). Aliquots were removed at the indicated time points, and the reaction was stopped with the addition of protein sample buffer.

Supplementary information is available at EMBO *reports* online (http://www.emboreports.org).

ACKNOWLEDGEMENTS

We thank A. Darwin and R. Festa for critical review of draft versions of this manuscript. We are grateful to T. Huang and K. Hofmann for helpful discussions. This work was supported by National Institutes of Health grants AI065437 and HL92774 awarded to K.H.D and HG3456 to S.P.G. K.H.D. is supported by the Irma T. Hirschl Trust and holds an Investigators in the Pathogenesis of Infectious Disease Award from the Burroughs Wellcome Fund.

Author contributions F.C.-M., F.M. and K.H.D. designed and conducted the experiments, and wrote the manuscript. N.J.B. and K.E.B. conducted the experiments. S.P.G. designed the experiments.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

Ausubel F, Brent R, Kingston R, Moore D, Seidman J, Smith J, Struhl K (2002) *Short Protocols in Molecular Biology*, Vol. 1. New York, NY, USA: Wiley

Burns KE, Liu WT, Boshoff HI, Dorrestein PC, Barry CE III (2009) Proteasomal protein degradation in *Mycobacteria* is dependent upon a prokaryotic ubiquitin-like protein. *J Biol Chem* **284:** 3069–3075

Burns KE, Cerda-Maira FA, Wang T, Li H, Bishai WR, Darwin KH (2010a) 'Depupylation' of prokaryotic ubiquitin-like protein from mycobacterial proteasome substrates. *Mol Cell* **39:** 821–827

Burns KE, Pearce MJ, Darwin KH (2010b) Prokaryotic ubiquitin-like protein provides a two-part degron to *Mycobacterium* proteasome substrates. *J Bacteriol* **192:** 2933–2935

Cerda-Maira FA, Pearce MJ, Fuortes M, Bishai WR, Hubbard SR, Darwin KH (2010) Molecular analysis of the prokaryotic ubiquitin-like protein (Pup) conjugation pathway in *Mycobacterium tuberculosis*. *Mol Microbiol* **77:** 1123–1135

Chen ZJ, Sun LJ (2009) Nonproteolytic functions of ubiquitin in cell signaling. *Mol Cell* **33:** 275–286

Deshaies RJ, Joazeiro CA (2009) RING domain E3 ubiquitin ligases. *Annu Rev Biochem* **78:** 399–434

Durfee LA, Lyon N, Seo K, Huibregtse JM (2010) The ISG15 conjugation system broadly targets newly synthesized proteins: implications for the antiviral function of ISG15. *Mol Cell* **38:** 722–732

Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* **4:** 207–214

Festa RA, Pearce MJ, Darwin KH (2007) Characterization of the proteasome accessory factor (paf) operon in *Mycobacterium tuberculosis*. *J Bacteriol* **189:** 3044–3050

Festa RA, McAllister F, Pearce MJ, Mintseris J, Burns KE, Gygi SP, Darwin KH (2010) Prokaryotic ubiquitin-like protein (Pup) proteome of *Mycobacterium tuberculosis*. *PLoS ONE* **5:** e8589

Ghadbane H, Brown AK, Kremer L, Besra GS, Futterer K (2007) Structure of *Mycobacterium tuberculosis* mtFabD, a malonyl-CoA:acyl carrier protein transacylase (MCAT). *Acta Crystallogr Sect F Struct Biol Cryst Commun* **63:** 831–835

Guth E, Thommen M, Weber-Ban E (2011) Mycobacterial ubiquitin-like protein ligase PafA follows a two-step reaction pathway with a phosphorylated pup intermediate. *J Biol Chem* **286:** 4412–4419

Hatfull GF, Jacobs WR Jr (2000) *Molecular Genetics of Mycobacteria*. Washington, DC, USA: ASM Press

Hochstrasser M (2000) Evolution and function of ubiquitin-like protein-conjugation systems. *Nat Cell Biol* **2:** E153–E157

Humbard MA, Miranda HV, Lim JM, Krause DJ, Pritz JR, Zhou G, Chen S, Wells L, Maupin-Furlow JA (2010) Ubiquitin-like small archaeal modifier proteins (SAMPs) in *Haloferax volcanii*. *Nature* **463:** 54–60

Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, Beausoleil SA, Villen J, Haas W, Sowa ME, Gygi SP (2010) A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **143:** 1174–1189

Imkamp F, Rosenberger T, Striebel F, Keller PM, Amstutz B, Sander P, Weber-Ban E (2010a) Deletion of dop in *Mycobacterium smegmatis* abolishes pupylation of protein substrates *in vivo*. *Mol Microbiol* **75:** 744–754

Imkamp F, Striebel F, Sutter M, Ozcelik D, Zimmermann N, Sander P, Weber-Ban E (2010b) Dop functions as a depupylase in the prokaryotic ubiquitin-like modification pathway. *EMBO Rep* **11:** 791–797

Iyer LM, Burroughs AM, Aravind L (2008) Unraveling the biochemistry and provenance of pupylation: a prokaryotic analog of ubiquitination. *Biol Direct* **3:** 45

Knodler LA, Winfree S, Drecktrah D, Ireland R, Steele-Mortimer O (2009) Ubiquitination of the bacterial inositol phosphatase, SopB, regulates its biological activity at the plasma membrane. *Cell Microbiol* **11:** 1652–1670

Kremer L, Nampoothiri KM, Lesjean S, Dover LG, Graham S, Betts J, Brennan PJ, Minnikin DE, Locht C, Besra GS (2001) Biochemical characterization of acyl carrier protein (AcpM) and malonyl-CoA:AcpM transacylase (mtFabD), two major components of *Mycobacterium tuberculosis* fatty acid synthase II. *J Biol Chem* **276:** 27967–27974

Li X, Gerber SA, Rudner AD, Beausoleil SA, Haas W, Villen J, Elias JE, Gygi SP (2007) Large-scale phosphorylation analysis of α-factor-arrested *Saccharomyces cerevisiae*. *J Proteome Res* **6:** 1190–1197

Patel JC, Hueffer K, Lam TT, Galan JE (2009) Diversification of a *Salmonella* virulence protein function by ubiquitin-dependent differential localization. *Cell* **137:** 283–294

Pearce MJ, Arora P, Festa RA, Butler-Wu SM, Gokhale RS, Darwin KH (2006) Identification of substrates of the *Mycobacterium tuberculosis* proteasome. *EMBO J* **25:** 5423–5432

Pearce MJ, Mintseris J, Ferreyra J, Gygi SP, Darwin KH (2008) Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science* **322:** 1104–1107

Pickart CM, Cohen RE (2004) Proteasomes and their kin: proteases in the machine age. *Nat Rev Mol Cell Biol* **5:** 177–187

Poulsen C, Akhter Y, Jeon AH, Schmitt-Ulms G, Meyer HE, Stefanski A, Stuhler K, Wilmanns M, Song YH (2010) Proteome-wide identification of mycobacterial pupylation targets. *Mol Syst Biol* **6:** 386

Rappsilber J, Ishihama Y, Mann M (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* **75:** 663–670

Rotin D, Kumar S (2009) Physiological functions of the HECT family of ubiquitin ligases. *Nat Rev Mol Cell Biol* **10:** 398–409

Sambrook J, Maniatis T, Fritsch E (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, New York, USA: Cold Spring Harbor Laboratory Press

Striebel F, Imkamp F, Sutter M, Steiner M, Mamedov A, Weber-Ban E (2009) Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to substrates by distinct but homologous enzymes. *Nat Struct Mol Biol* **16:** 647–651

Striebel F, Hunkeler M, Summer H, Weber-Ban E (2010) The mycobacterial Mpa-proteasome unfolds and degrades pupylated substrates by engaging Pup's N-terminus. *EMBO J* **29:** 1262–1271

Watrous J, Burns K, Liu WT, Patel A, Hook V, Bafna V, Barry CE III, Bark S, Dorrestein PC (2010) Expansion of the mycobacterial 'PUPylome'. *Mol Biosyst* **6:** 376–385