J|A|C|S
COMMUNICATIONS

Published on Web 03/31/2010

Prokaryotic Ubiquitin-like Protein (Pup) Is Coupled to Substrates via the Side Chain of Its C-Terminal Glutamate

Markus Sutter, Fred F. Damberger, Frank Imkamp, Frédéric H.-T. Allain,* and Eilika Weber-Ban*

Institute of Molecular Biology & Biophysics, ETH Zürich, CH-8093 Zürich, Switzerland

Received December 14, 2009; E-mail: allain@mol.biol.ethz.ch; eilika@mol.biol.ethz.ch

While ubiquitin (Ub) has been known for decades as a post-translationally conjugated protein degradation tag in eukaryotes, prokaryotic ubiquitin-like protein (Pup), the degradation tag in proteasome-harboring actinobacteria, has only recently been discovered.[1,2] Pup (64 residues) and Ub (76 residues) show neither structural nor sequence homology except for a GG motif near or at the C-terminus. Ub forms a compact three-dimensional (3D) structure,[3] whereas Pup is largely unstructured except for an α-helical segment close to its C-terminus.[4−6] Although both Pup and Ub are attached to the amino group of lysine side chains in substrates and target the substrates for degradation by the proteasome, the chemistry of the coupling reaction is different. Ub is coupled to substrates via the carboxy group of its C-terminal glycine in a multistep reaction involving several enzymes.[7] In the mycobacterial pupylation pathway, the C-terminal glutamine of Pup is first deamidated by deamidase of Pup (Dop), after which Pup is coupled by proteasome accessory factor A (PafA) to the ε-amino group of a substrate lysine via an isopeptide bond (Figure 1a).[1,2,8] We refer to the deamidated Pup protein as Pup-GGE. The covalently Pup-modified protein is then recognized by the proteasomal ATPase Mpa and degraded by the proteasome.[6,9]

On the basis of current knowledge, Pup could be coupled to substrates in two ways: at the C-terminal carboxylate, in analogy to Ub (I in Figure 1b), or at the side-chain carboxylate of the C-terminal glutamate (II in Figure 1b). Since the linkage to the substrate was identified by mass spectrometry,[1,2] the two modes of coupling could not be distinguished. Here we have used NMR spectroscopy with isotopically labeled *Mycobacterium tuberculosis* Pup and substrate to show that the coupling of Pup to substrates occurs via the side-chain carboxylate of the terminal glutamate rather than the C-terminal carboxylate of Pup.

To address the coupling specificity by NMR spectroscopy, we first used a model coupling reaction in which an excess of the free amino acid lysine was substituted for a natural protein substrate to allow efficient and complete coupling of Pup-GGE. To identify the atoms from the glutamate and the substrate lysine forming the new amide bond and determine the chemical shift changes resulting from bond formation, we generated unlabeled, uniformly singly labeled ^{15}N-, and doubly labeled ^{13}C,^{15}N-Pup-GGE and recorded 2D and 3D heteronuclear NMR spectra before and after PafA-catalyzed coupling to $^{15}N_2$-L-lysine.

In order to identify the expected additional signal arising from the new amide bond, we measured the [^{15}N,^{1}H] heteronuclear single-quantum correlation (HSQC) spectrum of unlabeled Pup coupled to uniformly ^{15}N-labeled lysine. The spectrum showed only one peak (the circled peak in Figure 2), unambiguously identifying the ^{1}H−^{15}N signal of the newly formed amide bond (NH) from the coupled lysine (K*) and at the same time indicating that Pup is exclusively modified at one position.

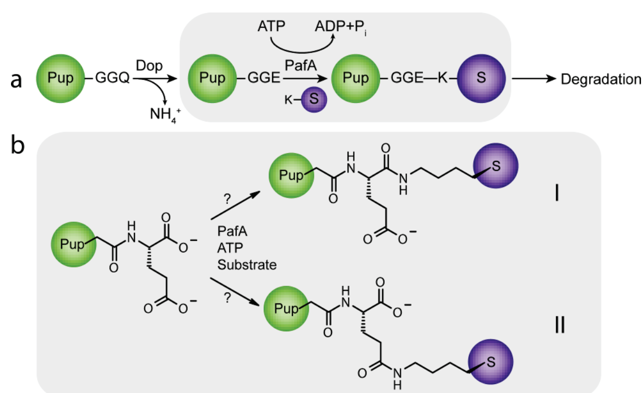

Figure 1. (a) Illustration of the pupylation pathway in mycobacteria. (b) Two possible modes of coupling: (I) C-terminal carboxylate; (II) Carboxylate of the glutamate side chain.

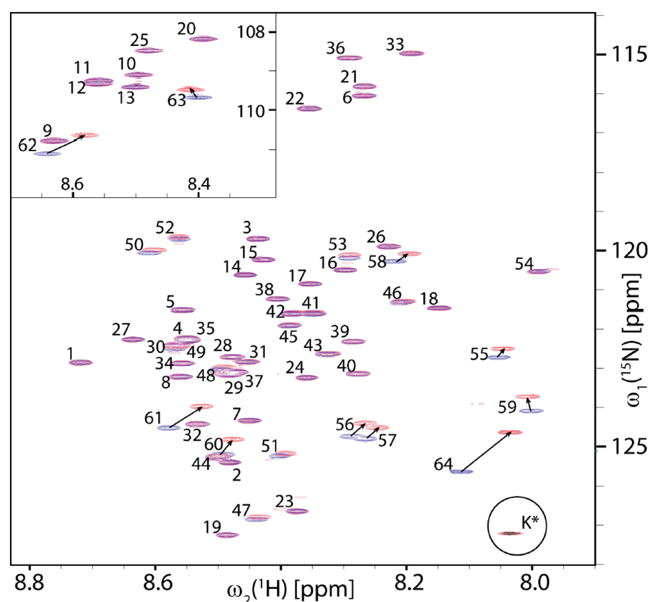

Figure 2. Superposition of [^{15}N,^{1}H]-HSQC spectra of ^{15}N-labeled Pup-GGE (blue), unlabeled Pup-GGE coupled to ^{15}N-labeled L-lysine (gray), and ^{15}N-labeled Pup-GGE coupled to ^{15}N-labeled L-lysine (red). Residue assignments are indicated for backbone amide H−N signals. Arrows connect resonances showing significant chemical shift changes before and after coupling. The position of the new signal originating from the coupling reaction is denoted "K*" and highlighted with a circle. The inset shows the region of ω_1 from 108 to 112 ppm.

Coupling to lysine causes significant changes in the amide ^{1}H and ^{15}N chemical shifts near the C-terminus of Pup-GGE (arrows

10.1021/ja910546x © 2010 American Chemical Society

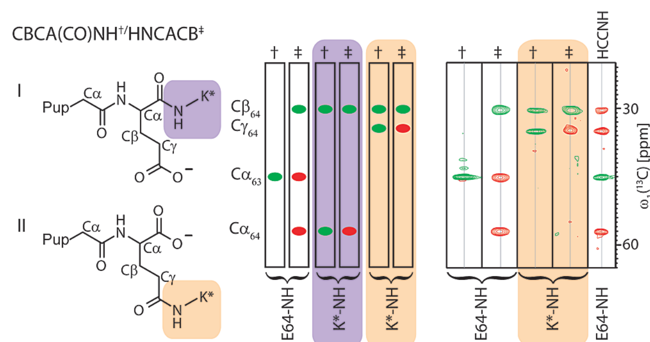

Figure 3. (left) The C-terminal glutamate of Pup could be coupled to the substrate lysine via either the C-terminal (I/purple box) or side-chain carboxylate (II/orange box). (middle) Expected cross-peaks for H−N strips of the C-terminal glutamate E64-NH and the linked ^{15}N-labeled Lys (K*-NH) from the CBCA(CO)NH/HNCACB spectra. (right) Corresponding strip plots from the measured data. In addition, the H−N strip of E64 from the HCCNH spectrum used to identify the Cγ of E64 is shown. Positive and negative contours are shown in red and green, respectively.

in Figure 2; also see Table S1 in the Supporting Information), with the largest effect observed for E64.

To determine the atoms linked by the coupling reaction, we applied two strategies based upon correlations of ^{13}C resonances with amide ^{1}H and ^{15}N resonances, one involving aliphatic ^{13}C and the other carbonyl ^{13}C. Coupling of ^{15}N-labeled lysine to the carboxy terminus should give rise to a different set of correlations than coupling to the side-chain carboxylate. The HNCACB/CBCA(CO)NH[10] spectra (Figure 3) show correlations of the Cβ and Cγ of glutamate E64 with the NH of K*. This is only possible if K* is coupled to the side-chain carboxylate rather than the carboxy terminus of Pup-GGE. In agreement with this, no correlation between the NH of K* and the Cα of E64 was observed. The HCCNH[11] H−N strip of E64 on the right confirms the assignment of the Cγ of E64 (35.0 ppm). Additionally, the HNCO/HN(CA)CO[10] experiments (Figure S1 in the Supporting Information) clearly showed that the substrate lysine is not part of the peptide main chain, since no correlation between the backbone C′ of E64 and the NH of K* was observed. These data rule out coupling via the C-terminal carboxylate.

We independently verified the linkage between the substrate lysine and the side-chain carboxylate of E64 via the proton shifts (Figure S2 in the Supporting Information). All of the resonance assignments obtained for Pup-GGE−^{15}N-Lys are consistent with the chemical shifts expected for the linkage of the substrate lysine to the side-chain carboxylate of the C-terminal glutamate. Moreover, the atoms of E64 whose resonances show the largest shift changes upon coupling to the substrate are closest in the chemical structure to the site of bond formation (Table S2 in the Supporting Information). Since we employed a tagging assay using free lysine, it is in principle possible that coupling occurs via the amino group at the α-position. To exclude this possibility, we assigned the ^{1}H and ^{13}C resonances of the substrate lysine (Figures S3 and S4 in the Supporting Information) and compared these shifts to random-coil ^{1}H and ^{13}C chemical shifts of lysine (Table S3 in the Supporting Information), confirming that coupling in this assay occurs only through the ε-amino group of the free lysine.

In order to confirm that Pup is also linked to natural protein substrates via the side-chain carboxylate of its C-terminal Glu,

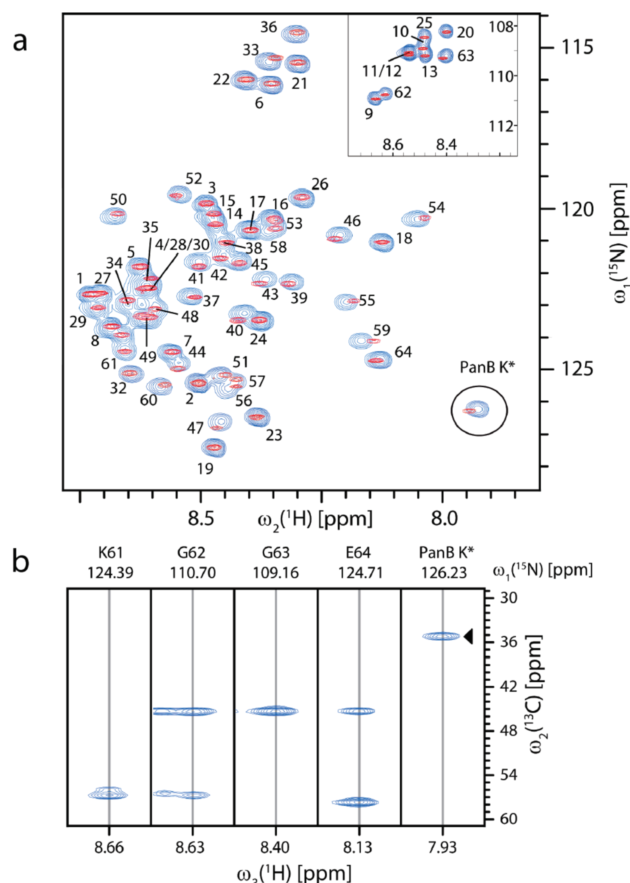

Figure 4. The C-terminal glutamate of Pup-GGE is coupled to the protein substrate PanB lysine residue via the side-chain carboxylate. (a) ^{1}H−^{15}N projections of the 3D HNCA spectrum obtained for a sample containing ^{13}C,^{15}N-Pup-GGE linked to ^{15}N-labeled PanB in 8 M urea (blue) and the spectrum obtained for ^{15}N-Pup-GGE linked to ^{15}N-labeled Lys in 8 M urea (red). Residue assignments are indicated for backbone amide H−N signals of Pup-GGE. The position of the signal matching the resonance position of K* in Figure 2 is circled and denoted "PanB K*". The inset shows the region of ω_1 from 108 to 112 ppm. (b) HNCA strips obtained from the C-terminal tetrapeptide of Pup and from the resonance labeled PanB K* in (a). The single upfield ^{13}C resonance in the strip of PanB K* marked with a triangle is consistent with the formation of an isopeptide bond linking the lysine of PanB and the side-chain carboxylate of E64 from Pup-GGE.

we coupled ^{13}C,^{15}N-labeled Pup to ^{15}N-labeled PanB, a proteasomal substrate previously identified in *M. tuberculosis*.[12] Because of the unfavorable size of the PanB decamer (300 kDa) for NMR spectroscopy under native conditions, we added 8 M urea to the NMR samples prior to the measurements to disassemble and unfold the substrate while preserving the covalent linkage. [^{15}N,^{1}H]-HSQC spectra showed the presence of both signals from Pup and the linked PanB (286 residues), as expected (Figure S5 in the Supporting Information). To obtain the backbone resonance assignments of Pup-GGE linked to PanB, we carried out triple resonance experiments. The ^{1}H−^{15}N projection of the 3D HNCA spectrum shows only resonances due to Pup-GGE (Figure 4a). The resonance positions show a pattern similar to that observed for Pup linked to free lysine (see Figure 2), and the resonances closely coincide when the [^{15}N,^{1}H]-HSQC spectrum of Pup linked to free lysine is measured in 8 M urea (red contours). In addition to the resonances from the backbone of Pup, we detected a signal in the lower-right

corner of the ^{1}H−^{15}N HNCA projection whose position closely matches that of the signal from the isopeptide bond formed between free lysine and the side-chain carboxylate of E64 (marked PanB K*).

The 3D HNCA ^{1}H−^{13}C strips obtained at the ^{15}N resonance positions of the residues K61 through E64 show a clear sequential walk, which is confirmed by the 3D HN(CA)CO and 3D HNCACB spectra (Figures S6 and S7 in the Supporting Information), whereas the strip from the resonance labeled PanB K* shows a single correlation to a ^{13}C at a chemical shift of 35.5 ppm in the HNCA spectrum (Figure 4b). This shift is far outside the range of random-coil shifts exhibited by the ^{13}Cα resonances of Pup and closely matches the chemical shift determined for the ^{13}Cγ of E64 linked to free lysine (Figure 3). Its assignment is confirmed by the H−N strip of PanB K* in a 3D HCCNH spectrum. The 3D HNCACB and CBCA(CO)NH data confirm the link between the side-chain carboxylate of E64 and the amide of PanB K* (Figure S7 in the Supporting Information). The unusually large ^{15}N shift combined with the small ^{1}HN shift of the side-chain isopeptide makes this resonance a unique identifier for pupylated substrate, at least under denaturing conditions.

The NMR data obtained with ^{13}C,^{15}N-Pup coupled to ^{15}N-PanB are consistent with the conclusions drawn from the measurements performed with ^{13}C,^{15}N-Pup coupled to ^{15}N-Lys. They indicate that Pup is linked to substrate proteins by the formation of an isopeptide bond between the side-chain carboxylate of E64 and a lysine side chain of the substrate protein.

Since Pup-GGE is ligated via the glutamyl side-chain carboxylate rather than the carboxy terminus, we tested whether the coupling could also occur at an internal GGE motif. For this purpose, we generated a variant of Pup that was extended by a glycine at the C-terminus (Pup-GGEG). We performed a previously described coupling assay with both wild-type Pup-GGE and the variant. In this assay, coupling of Pup-GGE to free lysine results in a mobility shift of the product in SDS gel electrophoresis. Our results demonstrated that the variant could no longer be coupled under conditions where the coupling reaction with Pup-GGE is complete (Figure 5). These data suggest that the C-terminus is involved in the binding to the active site of PafA or that steric occlusion prevents the reaction in the presence of the additional glycyl residue.

The reason why two proteins involved in a functionally homologous but otherwise very different pathway both have a characteristic GG motif is unclear. A possible explanation is that the double glycine motif provides flexibility that is needed to access the active site. Our results establish unambiguously the chemical identity of the linkage between the marker and the substrate protein in the prokaryotic ubiquitin-like pathway, which differs from that observed in the eukaryotic degradation pathway. This information will be useful for future studies on the enzymes involved in coupling Pup to substrates and for developing inhibitors that may interfere with this regulatory process in pathogenic organisms such as *M. tuberculosis*.

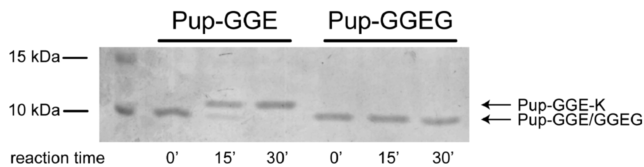

Figure 5. Time course of a coupling assay for Pup-GGE and Pup-GGEG with lysine followed by analysis of the samples by SDS-PAGE. The variant of Pup in which an additional glycine is present after the glutamate cannot react with lysine under the same conditions. The entire gel is shown in Figure S8 in the Supporting Information.

Acknowledgment. This work was supported by the Swiss National Science Foundation (SNF), the Center for Excellence in Research (NCCR) Structural Biology Program of the SNF, and an ETH research grant.

Supporting Information Available: Experimental details, Figures S1−S8, and Tables S1−S3. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Burns, K. E.; Liu, W. T.; Boshoff, H. I.; Dorrestein, P. C.; Barry, C. E. *J. Biol. Chem.* **2009**, *284*, 3069–3075.
(2) Pearce, M. J.; Mintseris, J.; Ferreyra, J.; Gygi, S. P.; Darwin, K. H. *Science* **2008**, *322*, 1104–1107.
(3) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531–544.
(4) Chen, X.; Solomon, W. C.; Kang, Y.; Cerda-Maira, F.; Darwin, K. H.; Walters, K. J. *J. Mol. Biol.* **2009**, *392*, 208–217.
(5) Liao, S.; Shang, Q.; Zhang, X.; Zhang, J.; Xu, C.; Tu, X. *Biochem. J.* **2009**, *422*, 207–215.
(6) Sutter, M.; Striebel, F.; Damberger, F. F.; Allain, F. H.; Weber-Ban, E. *FEBS Lett.* **2009**, *583*, 3151–3157.
(7) Kerscher, O.; Felberbaum, R.; Hochstrasser, M. *Annu. Rev. Cell. Dev. Biol.* **2006**, *22*, 159–180.
(8) Striebel, F.; Imkamp, F.; Sutter, M.; Steiner, M.; Mamedov, A.; Weber-Ban, E. *Nat. Struct. Mol. Biol.* **2009**, *16*, 647–651.
(9) Striebel, F.; Hunkeler, M.; Summer, H.; Weber-Ban, E. *EMBO J.* [Online early access]. DOI: 10.1038/emboj.2010.23. Published Online: March 4, 2010.
(10) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Skelton, N. J.; *Protein NMR Spectroscopy: Principles and Practice*, 1st ed.; Academic Press: London, 1996.
(11) Lyons, B. A.; Montelione, G. T. *J. Magn. Reson., Ser. B* **1993**, *101*, 206–209.
(12) Pearce, M. J.; Arora, P.; Festa, R. A.; Butler-Wu, S. M.; Gokhale, R. S.; Darwin, K. H. *EMBO J.* **2006**, *25*, 5423–5432.

JA910546X