JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Fluorescent Probes Reveal a Minimal Ligase Recognition Motif in the Prokaryotic Ubiquitin-like Protein from *Mycobacterium tuberculosis*

Denis Smirnov, Abhinav Dhall, Kalkena Sivanesam, Rudy J. Sharar, and Champak Chatterjee*

Department of Chemistry, University of Washington, Seattle, Washington 98195, United States

Supporting Information

ABSTRACT: The prokaryotic ubiquitin-like protein (Pup)-based proteasomal system in the pathogen *Mycobacterium tuberculosis* (*Mtb*) is essential for its survival in a mammalian host. The Pup ligase enzyme, PafA, conjugates Pup to a suite of proteins targeted for proteasomal degradation and is necessary for persistent infection by *Mtb*. We report the design and application of fluorescent probes for use in elucidating the mechanisms of Pup and substrate recognition by PafA. Our studies revealed that the C-terminal 26 amino acid sequence of Pup is the minimal ligase recognition motif in *Mtb*. Specific hydrophobic residues within this sequence that are known to be important for the interactions of Pup with proteasomes are also critical for the activation of Pup by PafA.

Mycobacterium tuberculosis (*Mtb*) is a human pathogen and the main causative agent of tuberculosis (TB). Globally, tuberculosis and HIV are the leading causes of death due to infection, and about one-third of the world's population carries the opportunistic tuberculosis bacilli.[1] The emergence of extensively drug-resistant forms of TB in recent years has seriously challenged current therapeutic strategies, signaling the need to identify and characterize new drug targets in *Mtb*.[2] Initial investigations of the mycobacterial 20*S* proteasome suggest that along with transcription and translation, inhibition of bacterial protein degradation may also prove to be an effective antibacterial strategy. Indeed, several studies have elucidated the necessity of functional proteasomes for *Mtb* resistance to oxidative challenge in vitro[3] and for persistent infection in mice.[4,5] Furthermore, the treatment of non-replicating *Mtb* with proteasome-inhibiting oxathiazol-2-one compounds was shown to be bactericidal in laboratory cultures.[6] These results have firmly established the proteasomal system as a suitable target for *Mtb* inhibition. The recent discovery of a ubiquitin-like protein modification pathway in *Mtb* that tags proteins for degradation by 20*S* proteasomes has revealed several additional targets for the inhibition of protein turnover.[7,8] The prokaryotic ubiquitin-like protein (Pup) is a short 64 amino acid polypeptide that is conjugated to proteasomal substrates by proteasome accessory factor A (PafA) ligase.[9] Given the high degree of conservation of Pup and PafA homologues within the Actinomycete class of bacteria and the essential role of PafA in maintaining *Mtb* infection,[8] we are interested in elucidating the mechanisms of Pup and substrate recognition by PafA. This understanding may be parlayed into the design of rational inhibitors of PafA and will also shed light on the evolutionary origins of the complex protein ubiquitylation machinery in higher organisms.

Proteasomal substrates in eukaryotes are typically tagged for degradation by conjugation of a lysine side-chain ϵ-amine with the C-terminus of the protein ubiquitin. Ubiquitylation is catalyzed by the E1−E3 family of ligases and begins with E1-ligase-catalyzed activation of the α-carboxylate of the C-terminal Gly in ubiquitin as a ubiquitin adenylate.[10] The activated ubiquitin is transferred to a side-chain thiol in the E1 ligase and subsequently to an E2 ligase. In some instances, the E2−ubiquitin thioester further participates in trans-thioesterification with a side-chain thiol in an E3 ligase. Finally, the ubiquitin C-terminal thioester undergoes nucleophilic attack by a lysine side chain or, in some instances, the N-terminus of a protein substrate to form a stable amide linkage. In contrast with ubiquitin, the small protein Pup is ribosomally synthesized with a C-terminal Gln residue that is deamidated by the deamidase of Pup (Dop) to produce a C-terminal Glu (Figure 1a).[9] The newly formed γ-carboxylate is conjugated with a lysine side-chain ϵ-amine in the substrate (Figure 1b, top row). Another key difference between Pup and ubiquitin is that the

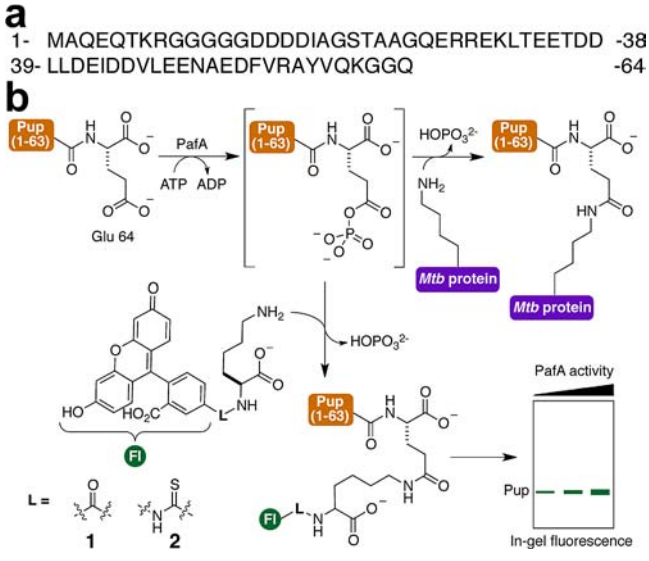

Figure 1. Mechanism-based probes of PafA activity. (a) Sequence of the Pup polypeptide. (b) Scheme depicting PupE conjugation with **1** and **2** catalyzed by PafA.

Received: November 20, 2012
Published: February 12, 2013

   dx.doi.org/10.1021/ja311376h | *J. Am. Chem. Soc.* 2013, 135, 2887−2890

buildup of polymeric chains of Pup is not observed on protein substrates, unlike the polyubiquitin chains that are typically observed on eukaryotic targets and are required for their proteosomal degradation.

Initial mechanistic studies of PafA established a key difference from the family of ubiquitin ligases in that PafA utilizes the terminal phosphate of ATP to activate Pup by generating a γ-carboxyphosphoanhydride at its C-terminal glutamate (Figure 1b).[11] This high-energy intermediate species, which is observable by MALDI-TOF mass spectrometry, is proposed to undergo subsequent nucleophilic attack by the lysine side chain. Several proteomic studies have demonstrated that ~130 different proteins in *Mtb* and the closely related *Mycobacterium smegmatis* are pupylated at internal lysine sites.[12−15] However, there is no known consensus sequence or conserved structure at the sites of pupylation, and the substrates are involved in many different pathways, including metabolism, cell-wall and membrane biosynthesis, transcription regulation, and even proteolysis.[16] The structure of a PafA homologue from the actinomycete *Corynebacterium glutamicum* (*Cglu*) was recently reported,[17] but the absence of bound Pup or a protein substrate precludes knowledge of the precise mechanisms underlying PafA function.

As a first step in our mechanistic studies, we sought to identify a PafA-specific chemical probe that (1) would allow direct and quantitative visualization of its activity and (2) is modular and therefore amenable to structure−activity studies of PafA specificity. In this regard, we noted that the PafA-catalyzed reaction is similar to transglutaminase-mediated amide bond formation between glutamine and lysine side chains. Several fluorescent amines have been employed as substrates for transglutaminases,[18,19] and we envisioned a similar approach for PafA. Therefore, we first tested Lys conjugated with fluorescein-5-carboxylic acid at its α-amine as a substrate for pupylation with purified N-terminally His_6-tagged and deamidated Pup (His_6-PupE) and C-terminally His_6-tagged PafA (PafA-His_6) (Figure 1b and Figures S1 and S2 in the Supporting Information). The amide-linked probe **1** and thiourea-linked analogue **2** were both robust substrates for pupylation in vitro, which was easily detected by in-gel fluorescence following sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Figures 2a and S3). The nature of the linkage between the fluorophore and the amine in our probes did not influence the labeling efficiency, which facilitated the rapid synthesis of multiple probes. Moreover, these probes were highly specific for PafA activity, which permitted the detection of Pup in complex protein mixtures (Figure 2b). Although the response was not saturated, PafA activity was readily detected at probe concentrations as low as 5 μM (Figure S4). The lack of signal saturation is consistent with the previously reported K_M of ~23 mM for free Lys[11] and demonstrates the advantage of employing a highly sensitive fluorescent readout.

With a modular and specific probe in hand, we first focused on understanding the substrate specificity of PafA. In this regard, we noted that a significant difference between pupylation and ubiquitylation is that ubiquitin may be attached to side-chain ϵ-amines of lysines as well as to the N-terminal α-amines of proteins.[20] However, pupylation has been observed only at lysine side chains.[15,21] To understand the basis for the amine specificity of PafA, we synthesized fluorescent amine substrates **3**−**5** (Figure 2c) in which the distance of the amine from the bulky fluorophore was varied. Activity assays with

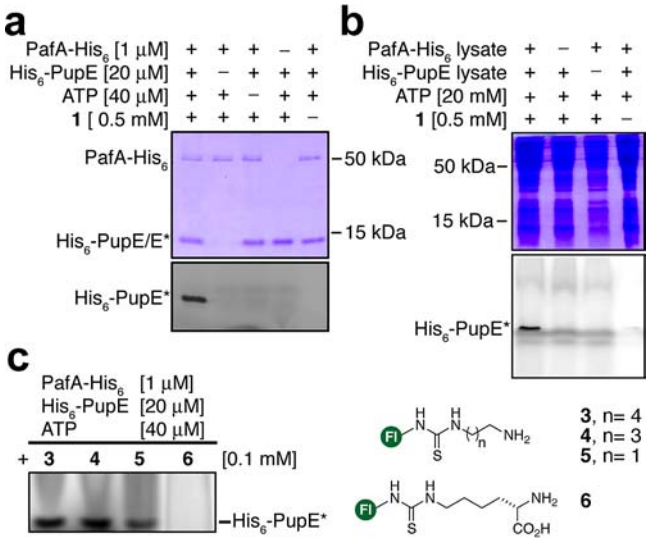

Figure 2. Application of fluorescent substrates to probe PafA activity. (a) 15% SDS-PAGE showing labeling of His_6-PupE by probe **1** in an ATP- and PafA-His_6-dependent manner: (top) Coomassie-stained gel; (bottom) gel slice showing in-gel fluorescence. (b) 15% SDS-PAGE showing the specificity of **1** for labeling of His_6-PupE in cellular lysates: (top) Coomassie-stained gel; (bottom) gel showing in-gel fluorescence. (c) In-gel fluorescence of His_6-PupE modified by probes **3**−**6**, whose structures are shown at the right. His_6-PupE* indicates the probe-labeled fluorescent peptide and Fl denotes fluorescein.

PafA and Pup revealed that longer-chain amines were better substrates and that the extent of pupylation decreased as the distance of the amine from the bulky fluorophore was shortened (Figures 2c and S5). However, pupylation was observed only at amines attached to primary carbons. Thus, the α-amine of N^{ϵ}-FITC-L-Lys (**6**) was not significantly pupylated (Figures 2c and S5). The free amino acids Ala and Gly were also tested in pupylation assays followed by liquid chromatography (LC) and electrospray ionization mass spectrometry (ESI-MS). Similar to our results with **6**, Ala was not measurably pupylated. However, free Gly, in which the α-amine is attached to a primary carbon, was pupylated by PafA (Figure S6). Importantly, the PafA specificity for the ϵ-amine of Lys did not change at higher pH (Figure S7), suggesting that the selectivity arises from structural requirements in the ligase and not from the chemical step. The recently reported structure of *Cglu* PafA with bound ADP and Mg^{2+} revealed a shallow open surface where substrates may bind (Figure S8).[17] Our results indicate that the reactive phosphoanhydride of Pup is accessible to linear amines and that branching at the carbon adjacent to the nucleophilic amine may interfere with the favored Bürgi−Dunitz angle of nucleophilic attack.[22] A subset of ubiquitin E2 ligases have also been shown to inherently select against amines attached to secondary carbons,[23] and it is likely that the selectivity observed in PafA is an evolutionary precursor to that observed in E2 ligases.

We next turned our attention to investigating how *Mtb* PafA binds Pup. Unlike the well-folded ubiquitin, Pup is disordered in buffered solutions, with no structural motifs seen in circular dichroism spectra and minimal helicity inferred from NMR experiments.[24−26] Darwin and co-workers demonstrated that an N-terminally truncated Pup(31−64) peptide is sufficient for pupylation in vivo.[27] Given the strong conservation of the Pup C-terminal sequence in Actinobacteria (Figure S9), we wondered what minimal sequence of Pup is sufficient for

PafA-mediated conjugation with substrates. Therefore, we tested the C-terminal 34 amino acids of Pup and N-terminal truncants thereof in in vitro pupylation assays with PafA and probe **1**. To facilitate soluble expression of the short fragments (MW < 4000 Da) and simplify visualization of the assay products by SDS-PAGE, an N-terminal ubiquitin (Ub) fusion tag was employed.[28] We first confirmed that full-length His_6-Ub-PupE(1−64) was efficiently labeled with probe **1** at Glu64 by PafA (Figure S10) and then proceeded to test Pup fragments in pupylation assays. Starting from His_6-Ub-PupE-(31−64), we truncated five N-terminal amino acids at a time. This allowed us to narrow down the residues critical for pupylation to a region between amino acids 36 and 41 in Pup (Figure S10). With this knowledge in hand, four additional Pup N-terminal truncants starting from His_6-Ub-PupE(37−64) were prepared and tested with PafA and probe **1** (Figure 3a). Because of the high specificity of probe **1**, each of the His_6-Ub-Pup fragments could be expressed in *Escherichia coli* and directly tested in cell lysates containing PafA and probe **1** without additional purification steps. Interestingly, we noted a gradual and significant decrease in labeling by **1** as the Pup sequence was truncated from His_6-Ub-PupE(39−64) to His_6-Ub-PupE(41−64) (Figure 3a). Furthermore, assays with synthetic PupE(40−64) and PupE(41−64) peptides followed by LC/ESI-MS analysis confirmed that the former is labeled to some degree while the latter is not (Figure S11). These results demonstrated that the C-terminal 26 amino acid PupE(39−64) sequence is the minimal recognition motif for PafA in *Mtb*.

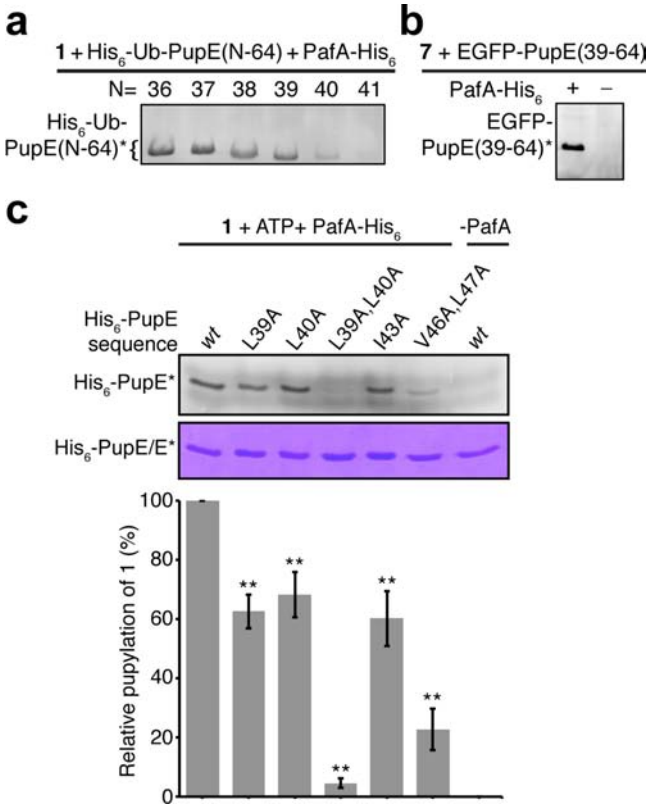

Figure 3. Identification of a minimal sequence and residues in Pup critical for pupylation. (a) In-gel fluorescence from a 15% SDS-PAGE gel showing PafA-His_6-catalyzed labeling of the indicated His_6-Ub-PupE fragments by probe **1** in cellular lysates. (b) In-gel fluorescence from a 15% SDS-PAGE gel showing the PafA-His_6-dependent labeling of EGFP-PupE(39−64) by Nα-TAMRA-L-Lys (**7**) in cellular lysates. (c) 15% SDS-PAGE gel of PafA-mediated labeling of wild-type (*wt*) and mutant full-length His_6-PupE polypeptides by probe **1**: (top) gel slice showing in-gel fluorescence of labeled proteins; (bottom) Coomassie-stained gel slice as a loading control. The bar graph below shows the quantitation of the in-gel fluorescence of each mutant relative to *wt* Pup after normalization for protein loading. Error bars are standard deviations (n = 3). ** indicates $P < 0.05$ (Student's two-tailed t test). Asterisks indicate the probe-labeled fluorescent peptides/proteins in each gel.

An additional and surprising result from our experiments was the observation that adding ubiquitin to the N-terminus of the minimal recognition motif, PupE(39−64), did not inhibit PafA activity. To test this motif as a general tag for protein labeling by PafA, we appended it to the C-terminus of the ∼27 kDa enhanced green fluorescent protein (EGFP). To our delight, the EGFP-PupE(39−64) fusion protein was labeled by Lys conjugated at its α-amine with tetramethylrhodamine (N^{α}-TAMRA-L-Lys, **7**) in a PafA-dependent manner (Figure 3b). Protein-labeling strategies for cell-surface imaging have been extensively developed with short-peptide-modifying enzymes such as the mammalian transglutaminases[19] and Sfp phosphopantetheinyl transferase.[29] The ability to append the PupE(39−64) sequence to proteins genetically without inhibiting PafA function is promising for its application as an orthogonal cell-surface protein-labeling strategy that will complement and expand the repertoire of currently available techniques.

Having identified the PupE(39−64) sequence as the minimal recognition motif for PafA, we turned our attention to identifying residues within this sequence that are critical for pupylation. We first focused on Leu39 and Leu40 that are present in the minimal sequence but are absent in the poor substrate PupE(41−64) (Figures 1a and 3a). Site-directed Ala mutagenesis of Leu39 and Leu40 in full-length His_6-PupE, either individually or in combination, showed dramatic effects on the extent of labeling with probe **1**, with the His_6-PupE(L39A,L40A) double mutant showing the least labeling (Figures 3c, S12, and S13). This indicates a key role for Leu39 and Leu40 in catalysis by PafA. To understand the role of these residues in the first chemical step (phosphorylation of PupE by PafA), we employed a previously reported thin-layer chromatography-based radioassay that measures the production of α-^{32}P-ADP from α-^{32}P-ATP.[11] Assays with PafA, α-^{32}P-ATP, and a tagless version of Pup, G-PupE(1−64), or the minimal sequence G-PupE(39−64) (Figure S14) revealed that the two substrates were phosphorylated at similar rates (Figure S15). On the other hand, the rate of phosphorylation of the G-PupE(L39A,L40A) mutant was indistinguishable from the background hydrolysis of ATP by PafA in the absence of Pup (Figure S15). This indicates that Leu39 and Leu40 are critical for recognition and efficient phosphorylation of PupE by PafA. Several hydrophobic residues in PupE, including Leu39 and Leu40, were also proposed to make key van der Waals contacts with the coiled-coil domain of the proteasomal ATPase Mpa.[30] Although PafA does not share significant homology with Mpa, our results clearly show that the two proteins engage overlapping regions of PupE (Mpa interacts with residues 21−51 and PafA interacts with residues 39−64). Therefore, we wondered whether the same hydrophobic residues in PupE, namely, Leu39, Leu40, Ile43, Val46, and Leu47, play roles in both Mpa binding and PafA activity. Site-directed Ala mutations of each of the hydrophobic residues also led to decreased labeling with probe **1** (Figures 3c and S16), indicating that they are important for PafA activity. However,

dx.doi.org/10.1021/ja311376h | J. Am. Chem. Soc. 2013, 135, 2887−2890

the most dramatic effect was seen for the His_6-PupE-(L39A,L40A) double mutant. Thus, PupE uses the same hydrophobic residues to bind Mpa and PafA, with Leu39 and Leu40 contributing most significantly to the latter interaction.

In conclusion, we have demonstrated the successful design and application of specific and modular fluorescent probes for use in studying the mechanism of function of PafA, the sole Pup ligase in *Mtb*. Our studies have revealed mechanistic similarities between PafA and ubiquitin E2 ligases and provide a rationale for the selective pupylation of Lys side-chain ϵ-amines. Fluorescent probe **1** was also used to identify residues 39−64 in PupE as the minimal recognition motif for PafA. Surprisingly, this minimal motif could be appended to the C-termini of small and large proteins, such as ubiquitin and EGFP, respectively, without compromising the activity of PafA. Finally, similar to the ubiquitin system, where a hydrophobic patch composed of Leu8, Ile44, and Val70 is a docking site for many ubiquitin-binding proteins,[31] Leu39 and Leu40 are part of a docking site for Pup binding to both Mpa and PafA. This is the first identification of a common hydrophobic protein interaction surface on Pup that is similar to what is already known for ubiquitin in eukaryotes. Studies are currently underway to identify the structure of PupE(39−64) peptide bound to PafA, which will guide the future design of specific inhibitors of the Pup−PafA interaction.

■ ASSOCIATED CONTENT

Supporting Information

Supporting figures and detailed experimental procedures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

chatterjee@chem.washington.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was generously supported by the Department of Chemistry at the University of Washington, Seattle. We are grateful to Professors Michael H. Gelb and Rachel E. Klevit for insightful suggestions. We thank Professor Wilfred A. van der Donk, Dr. Leah M. Miller and members of the Chatterjee lab for critical reading of the manuscript. Denis Smirnov is thankful for a Mary Gates Research Scholarship and a UW Amgen Scholarship. Finally, we thank Professors Dustin J. Maly and Pradipsinh K. Rathod for allowing us to use their phosphor-imaging equipment.

■ REFERENCES

(1) Lawn, S. D.; Zumla, A. I. *Lancet* **2011**, *378*, 57.

(2) Zignol, M.; Hosseini, M. S.; Wright, A.; Lambregts-van Weezenbeek, C.; Nunn, P.; Watt, C. J.; Williams, B. G.; Dye, C. *J. Infect. Dis.* **2006**, *194*, 479.

(3) Darwin, K. H.; Ehrt, S.; Gutierrez-Ramos, J.-C.; Weich, N.; Nathan, C. F. *Science* **2003**, *302*, 1963.

(4) Gandotra, S.; Schnappinger, D.; Monteleone, M.; Hillen, W.; Ehrt, S. *Nat. Med.* **2007**, *13*, 1515.

(5) Gandotra, S.; Lebron, M. B.; Ehrt, S. *PLoS Pathog.* **2010**, *6*, No. e1001040.

(6) Lin, G.; Li, D.; de Carvalho, L. P. S.; Deng, H.; Tao, H.; Vogt, G.; Wu, K.; Schneider, J.; Chidawanyika, T.; Warren, J. D.; Li, H.; Nathan, C. *Nature* **2009**, *461*, 621.

(7) Burns, K. E.; Liu, W.-T.; Boshoff, H. I. M.; Dorrestein, P. C.; Barry, C. E. *J. Biol. Chem.* **2008**, *284*, 3069.

(8) Pearce, M. J.; Mintseris, J.; Ferreyra, J.; Gygi, S. P.; Darwin, K. H. *Science* **2008**, *322*, 1104.

(9) Striebel, F.; Imkamp, F.; Sutter, M.; Steiner, M.; Mamedov, A.; Weber-Ban, E. *Nat. Struct. Mol. Biol.* **2009**, *16*, 647.

(10) Hershko, A.; Ciechanover, A. *Annu. Rev. Biochem.* **1998**, *67*, 425.

(11) Guth, E.; Thommen, M.; Weber-Ban, E. *J. Biol. Chem.* **2011**, *286*, 4412.

(12) Watrous, J.; Burns, K.; Liu, W.-T.; Patel, A.; Hook, V.; Bafna, V.; Barry, C. E.; Bark, S.; Dorrestein, P. C. *Mol. BioSyst.* **2010**, *6*, 376.

(13) Poulsen, C.; Akhter, Y.; Jeon, A. H.-W.; Schmitt-Ulms, G.; Meyer, H. E.; Stefanski, A.; Stühler, K.; Wilmanns, M.; Song, Y.-H. *Mol. Syst. Biol.* **2010**, *6*, 386.

(14) Festa, R. A.; McAllister, F.; Pearce, M. J.; Mintseris, J.; Burns, K. E.; Gygi, S. P.; Darwin, K. H. *PLoS One* **2010**, *5*, No. e8589.

(15) Tung, C. W. *BMC Bioinf.* **2012**, *13*, 40.

(16) Delley, C. L.; Striebel, F.; Heydenreich, F. M.; Ozcelik, D.; Weber-Ban, E. *J. Biol. Chem.* **2012**, *287*, 7907.

(17) Ozcelik, D.; Barandun, J.; Schmitz, N.; Sutter, M.; Guth, E.; Damberger, F. F.; Allain, F. H.; Ban, N.; Weber-Ban, E. *Nat. Commun.* **2012**, *3*, 1014.

(18) Pasternack, R.; Laurent, H. P.; Ruth, T.; Kaiser, A.; Schon, N.; Fuchsbauer, H. L. *Anal. Biochem.* **1997**, *249*, 54.

(19) Lin, C.-W.; Ting, A. Y. *J. Am. Chem. Soc.* **2006**, *128*, 4542.

(20) Ciechanover, A. *Methods Mol. Biol.* **2005**, *301*, 255.

(21) Sutter, M.; Damberger, F. F.; Imkamp, F.; Allain, F. H.; Weber-Ban, E. *J. Am. Chem. Soc.* **2010**, *132*, 5610.

(22) Bürgi, H. B.; Dunitz, J. D.; Lehn, J. M.; Wipff, G. *Tetrahedron* **1974**, *30*, 1563.

(23) Pickart, C. M.; Rose, I. A. *J. Biol. Chem.* **1985**, *260*, 1573.

(24) Liao, S.; Shang, Q.; Zhang, X.; Zhang, J.; Xu, C.; Tu, X. *Biochem. J.* **2009**, *422*, 207.

(25) Sutter, M.; Striebel, F.; Damberger, F. F.; Allain, F. H.-T.; Weber-Ban, E. *FEBS Lett.* **2009**, *583*, 3151.

(26) Chen, X.; Solomon, W. C.; Kang, Y.; Cerda-Maira, F.; Darwin, K. H.; Walters, K. J. *J. Mol. Biol.* **2009**, *392*, 208.

(27) Burns, K. E.; Pearce, M. J.; Darwin, K. H. *J. Bacteriol.* **2010**, *192*, 2933.

(28) Butt, T. R.; Jonnalagadda, S.; Monia, B. P.; Sternberg, E. J.; Marsh, J. A.; Stadel, J. M.; Ecker, D. J.; Crooke, S. T. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 2540.

(29) Yin, J.; Lin, A. J.; Golan, D. E.; Walsh, C. T. *Nat. Protoc.* **2006**, *1*, 280.

(30) Wang, T.; Darwin, K. H.; Li, H. *Nat. Struct. Mol. Biol.* **2010**, *17*, 1352.

(31) Hicke, L.; Schubert, H. L.; Hill, C. P. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 610.