

Spontaneous Intermolecular Amide Bond Formation between Side Chains for Irreversible Peptide Targeting

Bijan Zakeri and Mark Howarth*

Department of Biochemistry, Oxford University, South Parks Road, Oxford, OX1 3QU, U.K.

Received December 30, 2009; E-mail: mark.howarth@bioch.ox.ac.uk

Peptides and synthetic peptide-like molecules are powerful tools for analysis and control of biological function.^{1–3} One major problem with the use of peptides is the instability of their interactions with biomolecules, with typically micromolar affinity relating to the limited accessible surface area^{4,5} and the intrinsic flexibility of peptides.⁶ However, appending a short peptide tag is the most common way to allow a protein of interest to be isolated or detected, giving minimum perturbation to protein function.^{7,8} Here we have designed a way to bind a peptide tag irreversibly, by adapting a recently discovered feature of amino acid chemistry: the spontaneous formation of an amide bond between a Lys and an Asn side chain in the appropriate environment.^{9–11}

Amide linkages outside of the protein main chain are termed isopeptide bonds (Figure 1A). Isopeptide bonds are chemically stable and resistant to most proteases.^{12,13} Enzymes such as transglutaminases catalyze isopeptide formation, stabilizing the extracellular matrix and strengthening blood clots,¹² but these enzymes are large and have low sequence specificity.¹² Recently, certain proteins were discovered to autocatalyze single-turnover isopeptide bond formation, yielding ultrathin viral capsid chain mail,⁹ or the proteolytically stable pili of Gram-positive bacteria,^{10,11} through nucleophilic attack of the ϵ -amino group from a Lys to the C γ group of an Asn, promoted by a nearby Glu (Figure 1B).^{9,10}

To apply spontaneous isopeptide bond formation to direct new covalent peptide interactions, we dissected the major pilin protein Spy0128 from *Streptococcus pyogenes*¹⁰ (Figure 1C) and explored whether the two fragments would covalently associate. Split proteins have successfully reconstituted in many cases, including enzymes and fluorescent proteins,¹⁴ albeit through noncovalent interactions. Spy0128 was split at the final β -strand of the C-terminal domain, to give the fragment pilin-C (Spy0128 residues 18–299, with N-terminal His₆) and the isopeptag (Spy0128 residues 293–308: TDKDMTITFTNKKDAE). This placed the reactive Asn on the isopeptag and the reactive Lys on pilin-C. To enhance recombinant expression in *E. coli*, the isopeptag was genetically fused to the N-terminus of maltose binding protein (MBP). To test covalent reaction, we mixed isopeptag-MBP and pilin-C, each at 10 μ M, and boiled the samples in SDS before SDS-PAGE (Figure 1D). A new product formed at ~80 kDa, consistent with reaction between isopeptag-MBP and pilin-C. We verified amide bond formation between isopeptag-MBP and pilin-C by mass spectrometry, demonstrating the loss of NH₃ upon reaction (Figure 2A). Pilin-C K179A, lacking the reactive Lys, did not form a covalent complex with isopeptag-MBP, determined by SDS-PAGE (Figure 1D) and mass spectrometry (Figure S1). Also, pilin-C did not react with MBP fused to an alternative peptide containing four potentially reactive Asn residues (MBP-isopeptag-N, Figure 1D).

Spy0128 contains another isopeptide bond in its N-terminal domain (Figure S2A). We showed the generality of our strategy of designing spontaneous amide bond-forming peptides, by dissecting Spy0128 at its N-terminal β -strand, with this time the reactive Lys

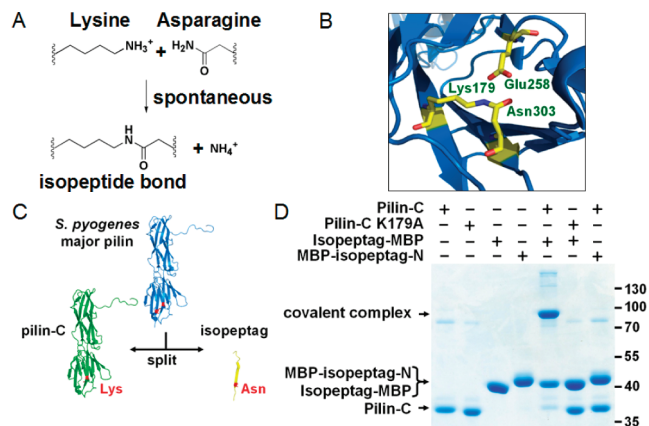


Figure 1. Spontaneous intermolecular amide bond formation. (A) Amide bond formation between Lys and Asn side chains. (B) Key residues for amide bond formation in the C-domain of the major pilin, from PDB 3B2M. (C) Cartoon of isopeptag construction. The pilin was dissected into a large N-terminal fragment (pilin-C in green) and a small C-terminal fragment (isopeptag in yellow). Reactive residues are highlighted in red. (D) Isopeptag and pilin-C associated covalently. Isopeptag-MBP and pilin-C were mixed at 10 μ M and analyzed by SDS-PAGE with Coomassie staining, alongside unreactive controls.

on the peptide (isopeptag-N) and the reactive Asn on the protein fragment (pilin-N) (Figure S2B): these partners also formed a covalent bond to each other when mixed (Figure S2C).

We determined the exact features of pilin-C and isopeptag important for reaction: truncating pilin-C earlier or later in the final β -strand did not substantially change the reactivity, but reaction was dramatically reduced upon truncating the isopeptag by the 5 residues of the loop preceding the final β -strand (Figure S3).

We tested the speed of pilin reconstitution: with each partner at 10 μ M, reaction was clearly detectable at 1 h and at later time points reached ~60% yield (Figure 2B). With a 2-fold excess of isopeptag, 98% of pilin-C was able to react in 24 h (Figure 2C). We further tested the concentration-dependence of the reaction, incubating both partners at 1, 5, or 10 μ M: the extent of reconstitution increased according to concentration over this range (Figure S4A). Surprisingly, the yield and speed of reaction were largely temperature-independent at 4–37 $^{\circ}$ C (Figure 2D). Reaction was also largely independent of pH at pH 6–8 but was reduced by 15% at pH 5 after 24 h (Figure S4B). Bond formation proceeded to a similar extent in a range of biological buffers, including with detergent, and with no requirement for any particular monovalent or divalent ions (Figure S4C). We have not yet found conditions that prevent spontaneous amide bond formation. The rate of *intramolecular* Lys-Asn bond formation has not been determined, because the reaction had gone to completion when the pilin was isolated,¹⁰ but it is likely to be substantially faster than the ~25 min generation time of *S. pyogenes*; future screening of phage-

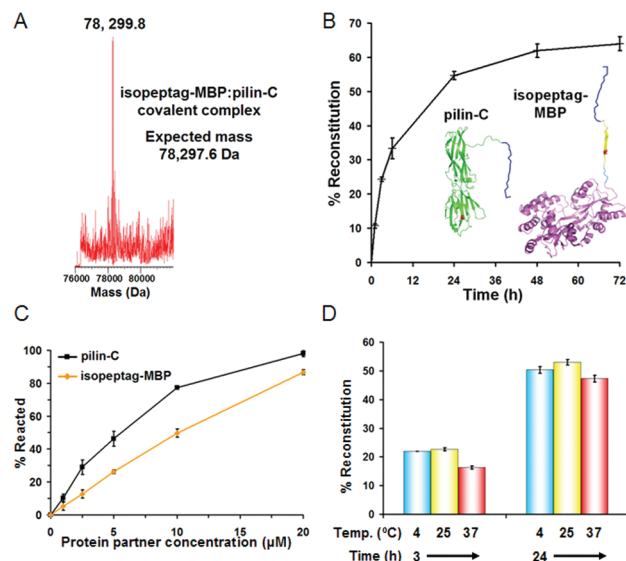


Figure 2. Characterization of spontaneous amide bond formation. (A) Mass spectrometry of reconstitution between isopeptag-MBP and pilin-C. (B) Time course of isopeptag-MBP:pilin-C covalent complex formation, with each partner at 10 μ M at 25 $^{\circ}$ C, pH 7.0 determined by SDS-PAGE. Illustration of the reacting species, with the His-tags and linker in blue. (C) Disappearance of 10 μ M pilin-C (black) or isopeptag-MBP (orange), upon incubation with varying concentrations of its partner, after 24 h at 25 $^{\circ}$ C, pH 7.0. (D) Temperature-dependence of isopeptag-MBP reconstitution with pilin-C at pH 7.0. (All graphs mean of triplicate \pm 1 s.d.).

display peptide libraries may identify isopeptag variants that associate rapidly and approach the intramolecular rate of reaction.

To demonstrate that this spontaneous amide bond formation would occur within living cells, we made a bicistronic construct, with pilin-C and isopeptag-MBP expressed from the same promoter. Inside the cytosol of *E. coli*, pilin-C but not the pilin-C K179A control efficiently reconstituted with isopeptag-MBP (Figure 3A). To test the specificity of the pilin-C:isopeptag interaction in a complex environment, we targeted the isopeptag to the surface of mammalian cells. Isopeptag-CFP-TM (illustrated in Figure S5) was labeled by pilin-C, but no binding was detected by the control pilin-C K179A, indicating good specificity of isopeptide formation on cells (Figure 3B).

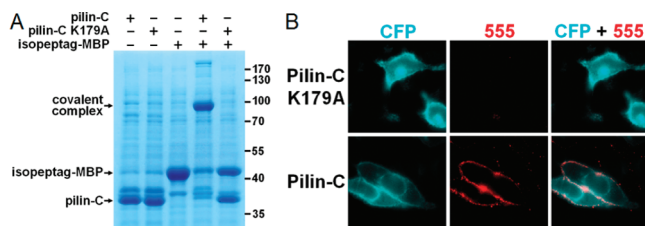


Figure 3. Specificity of isopeptag labeling. (A) Reconstitution of coexpressed isopeptag-MBP and pilin-C in the *E. coli* cytosol, shown by SDS-PAGE with Coomassie staining. (B) HeLa cells expressing isopeptag-CFP-TM (CFP, cyan) were incubated with pilin-C or pilin-C K179A control. Reconstitution was detected with anti-His tag antibody and Alexa Fluor 555-secondary antibody (555, red). CFP and 555 images are overlaid in the right panel.

Here we show how to harness autocatalytic side chain amide bond formation to provide a new possibility for a genetically encoded covalent reaction between a peptide and a protein. This reaction proceeded with similar efficiency at 4 and 37 $^{\circ}$ C; since there must be an activation barrier to the reaction, we hypothesize that the limiting step is association of the isopeptag with pilin-C in

a conformation suitable for reaction, and that such a conformation is less stable at elevated temperature.⁶ However, this temperature-independence opens up the possibility of isolation of isopeptag-containing proteins from cell lysates at 4 $^{\circ}$ C, to minimize sample degradation. Some split proteins do not reconstitute or remain soluble at 37 $^{\circ}$ C,¹⁴ but we obtained reaction at 37 $^{\circ}$ C and observed solubility >200 μ M for pilin-C and isopeptag-MBP. Interestingly, we observed a small amount of side products of pilin-C reaction (Figures 1D, S1, and S6), which may point to alternative conformations where amide bond formation can occur. It will be valuable to explore the behavior upon splitting of several of the other domains known to contain spontaneous isopeptides.^{10,11} Spontaneous amide bond formation proceeded over a pH range from 5 to 8, indicating that it could be applied even in low pH cellular compartments such as endosomes. Pilin-C and the isopeptag do not contain cysteines and so the redox status of the compartment should not matter for reaction, in contrast to bisarsenicals¹⁵ and most split inteins.¹⁶

We showed that spontaneous amide bond formation is specific at the surface of mammalian cells; there are many other approaches for labeling of cellular proteins with fluorophores,¹⁵ but few precedents for covalent labeling of a genetically encoded peptide with a genetically encoded protein partner on cells. Alternative approaches to form covalent bonds to peptides include sortase-catalyzed reaction of N-terminal oligoglycine with C-terminal LPXTG, which has the advantage of the small tags required but requires millimolar calcium¹⁷ (disruptive in the cytosol and nucleus) and is only applicable at termini. Disulfide bonds can also be used for covalent peptide binding but are reversible and prone to nonspecific interactions. Covalent bond formation will be particularly advantageous either when peptide attachment must be stable over long periods, such as for bioassembly or imaging, or when proteins are subject to high forces, such as from the shear in the blood stream¹⁸ or from the firing of molecular motors.¹⁹

Acknowledgment. Funding provided by the Clarendon Fund, the Department of Biochemistry, and Worcester College Oxford.

Supporting Information Available: Additional characterization and experimental protocols. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Sato, A. K.; Viswanathan, M.; Kent, R. B.; Wood, C. R. *Curr. Opin. Biotechnol.* **2006**, *17*, 638.
- (2) Bautista, A. D.; Craig, C. J.; Harker, E. A.; Schepartz, A. *Curr. Opin. Chem. Biol.* **2007**, *11*, 685.
- (3) Horne, W. S.; Gellman, S. H. *Acc. Chem. Res.* **2008**, *41*, 1399.
- (4) Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9997.
- (5) Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X. *Angew. Chem., Int. Ed.* **2003**, *42*, 4872.
- (6) Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Chem.* **1991**, *20*, 519.
- (7) Huh, W. K.; Falvo, J. V.; Gerke, L. C.; Carroll, A. S.; Howson, R. W.; Weissman, J. S.; O'Shea, E. K. *Nature* **2003**, *425*, 686.
- (8) Jarvik, J. W.; Telmer, C. A. *Annu. Rev. Genet.* **1998**, *32*, 601.
- (9) Wikoff, W. R.; Liljas, L.; Duda, R. L.; Tsuruta, H.; Hendrix, R. W.; Johnson, J. E. *Science* **2000**, *289*, 2129.
- (10) Kang, H. J.; Coulibaly, F.; Clow, F.; Proft, T.; Baker, E. N. *Science* **2007**, *318*, 1625.
- (11) Budzik, J. M.; Marraffini, L. A.; Souda, P.; Whitelegge, J. P.; Faull, K. F.; Schneewind, O. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 10215.
- (12) Lorand, L.; Graham, R. M. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 140.
- (13) Arocha-Pinango, C. L.; Marchi, R.; Carvajal, Z.; Guerrero, B. *Blood Coagul. Fibrinolysis* **1999**, *10*, 43.
- (14) Ozawa, T. *Anal. Chim. Acta* **2006**, *556*, 58.
- (15) Marks, K. M.; Nolan, G. P. *Nat. Methods* **2006**, *3*, 591.
- (16) Schwarzer, D.; Cole, P. A. *Curr. Opin. Chem. Biol.* **2003**, *4*, 561.
- (17) Tsukiji, S.; Nagamune, T. *ChemBioChem* **2009**, *10*, 787.
- (18) Pierres, A.; Touchard, D.; Benoliel, A. M.; Bongrand, P. *Biophys. J.* **2002**, *82*, 3214.
- (19) Morris, P. D.; Tackett, A. J.; Raney, K. D. *Methods* **2001**, *23*, 149.

JA910795A