

Structure and Reaction Mechanism of Pyrrolysine Synthase (PylD)**

Felix Quitterer, Philipp Beck, Adelbert Bacher, and Michael Groll*

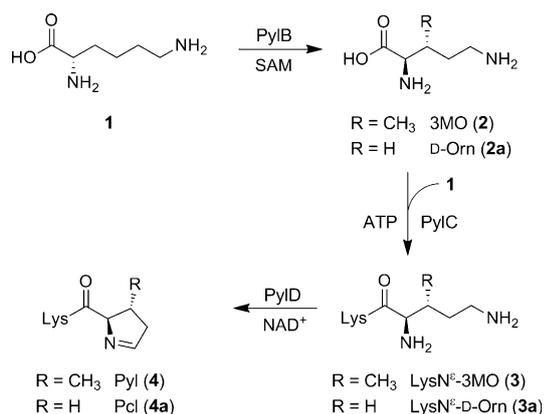
Pyrrolysine (Pyl, **4**), the recently discovered 22nd genetically encoded amino acid, has almost instantaneously become a hotspot of protein biochemistry,^[1] although its natural occurrence appears to be limited to just three proteins that are involved in the breakdown of methylamines in a small subgroup of archaea and bacteria.^[2] The novel amino acid is incorporated by a cognate pair of a tRNA and its synthetase (specified as *pylT* and *pylS*, respectively) through recognition of an amber stop codon (UAG).^[2a,3]

Biosynthesis of **4** is accomplished from two molecules of lysine (**1**) by sequential action of PylB, PylC, and PylD (Scheme 1).^[4] Specifically, the iron-sulfur *S*-adenosylmethionine protein PylB catalyzes the conversion of **1** to (3*R*)-3-methyl-D-ornithine (3MO, **2**),^[5] which is subsequently hooked up ATP-dependently to the ϵ amino group of a second lysine molecule by PylC resulting in **3**.^[6] Dehydrogenation at the C5 position of the methylornithine moiety of the isopeptide and subsequent ring closure catalyzed by PylD completes the

biosynthesis of the unusual amino acid pyrrolysine. The present report on PylD describes the structural investigation and implementation on the reaction mechanism of the enzymes required for the biosynthesis of pyrrolysine and may open novel opportunities for the harnessing of the system for biotechnology purposes.^[7]

We expressed the *pylD* gene of *Methanosarcina barkeri* Fusaro in an *Escherichia coli* strain. The recombinant protein was purified by metal-affinity chromatography and showed in vitro catalytic activity with $K_m = 3.6 \text{ mM} \pm 0.5 \text{ mM}$ (Figure S1 in the Supporting Information) and $k_{\text{cat}} = 0.76 \text{ s}^{-1} \pm 0.04 \text{ s}^{-1}$ using the surrogate L-lysine-N ϵ -D-ornithine (LysN ϵ -D-Orn, **3a**) as substrate. PylD was crystallized together with **3a** and/or a pyridine nucleotide cofactor (NADH or NAD $^+$). Crystals diffracted to a maximum resolution of 1.8 Å and starting phases were obtained by a combination of single-wavelength anomalous diffraction (SAD) methods using a selenomethionine derivative and twofold noncrystallographic symmetry (NCS) averaging. Real-space electron density map averaging was performed with MAIN^[8] in combination with CCP4 routines.^[9] Model building was carried out with MAIN and refinement was completed with REFMAC5^[10] (see Table S1). After structural elucidation of the selenomethionine-labeled PylD holoenzyme (PylD:holo (peak), PDB ID: 4JK3), we crystallized and determined the structure of native PylD in the presence of NAD $^+$ (PylD:holo, PDB ID: 4J43) to 2.2 Å resolution ($R_{\text{free}} = 20.1\%$, Table S1). Its molecular architecture is shown schematically in Figure 1a: the C-terminal segment (residues 139–259) resembles a Rossmann motif of five parallel β strands (S6–S10) with 21345 topology, which is N- and C-terminally flanked by helices H5 and H10, respectively (secondary-structure nomenclature: see Figure S3), yielding in the DALI search^[11] a highest Z-score of 15.6 for the *Rhodospirillum rubrum* Transhydrogenase Domain I (PDB ID: 1L7D). The N-terminal segment (residues 1–138) comprises a β sheet of five strands (S1–S5) whose overall orientation is orthogonal to that of the Rossmann motif of the C-terminal part. Interestingly, the C-terminal helix (H10) of PylD is wedged between the two β sheets supporting the correct fold and orientation of the N-terminal half of the dehydrogenase. In contrast to the Rossmann fold, the DALI search for proteins resembling the topology of the N-terminal segment resulted in only some similarities with the tRNA binding domain of certain tRNA synthetases ($Z\text{-score} < 8$).

The nicotinamide adenine dinucleotide coenzyme NAD $^+$ is bound to PylD in an extended conformation, inside a groove at the C-terminal pole of the Rossmann β sheet; both furanose rings have C2'-*endo* conformation. A typical VXGXGXXGXXA motif^[12] (residues 146–157) is part of the coenzyme binding site and the backbone elements are predominantly involved in a network of hydrogen bonds with



Scheme 1. Biosynthesis of pyrrolysine. Note: PylB generates only 3MO (**2**); PylC and PylD also catalyze the reactions of **2a** and **3a**, respectively.

[*] F. Quitterer, P. Beck, Prof. Dr. A. Bacher, Prof. Dr. M. Groll
Center for Integrated Protein Science Munich (CIPSM) at the
Department Chemie, Lehrstuhl für Biochemie
Technische Universität München
Lichtenbergstrasse 4, 85747 Garching (Germany)
E-mail: michael.groll@tum.de

[**] We thank the staff of the beamline X06SA at the Paul Scherrer
Institute, Swiss Light Source, Villigen (Switzerland) for their help
with data collection and Katrin Gärtner for excellent technical
assistance. We acknowledge Dr. Anja List who participated in the
PylD project. This work was supported by the Hans-Fischer-
Gesellschaft, by Award No. FIC/2010/07 from the King Abdullah
University of Science and Technology (KAUST), and by the Deutsche
Forschungsgemeinschaft (DFG, grant GR1861/7-1).

Supporting information for this article is available on the WWW
under <http://dx.doi.org/10.1002/anie.201301164>.

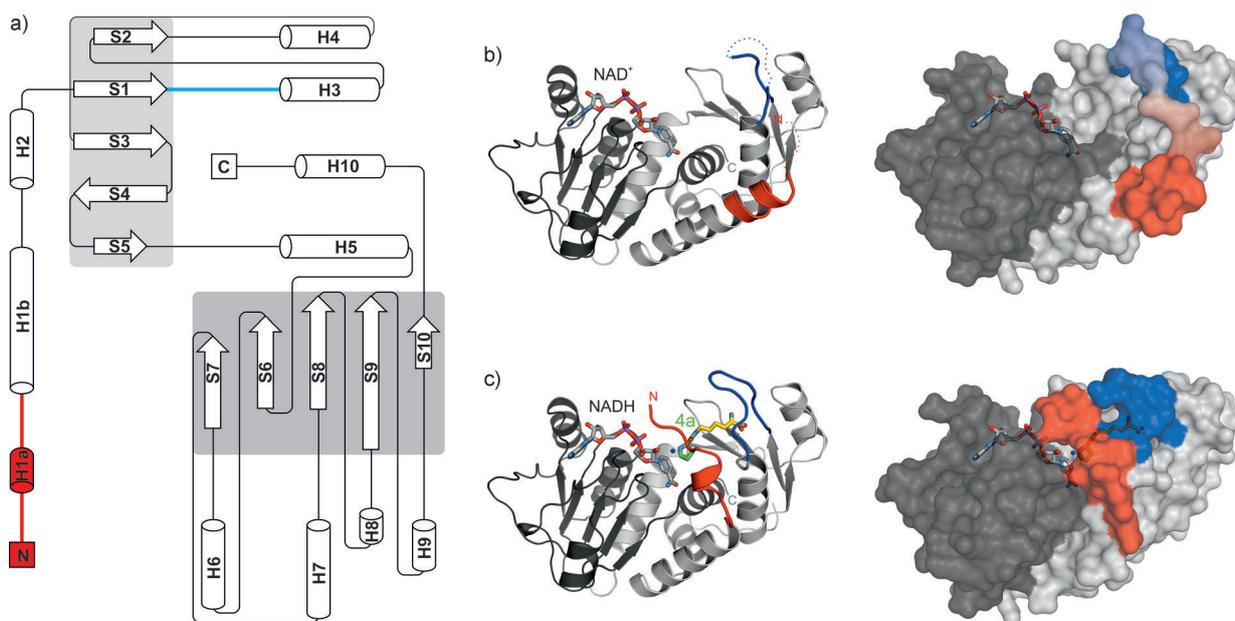


Figure 1. Overall structure of PyID. a) Topology with secondary structure elements shown for PyID in complex with **4a**. The Rossmann motif is highlighted against a dark gray background and the substrate-binding motif against a gray background. The N-terminal part and residues 52–61, which participate in an induced-fit mechanism upon substrate binding, are presented in red and blue, respectively. b) Left panel: Ribbon drawing of the PyID:holo structure (open conformation). The C-terminal Rossmann fold is shown in dark gray and the N-terminal substrate binding segment is depicted in gray. The N-terminus (red) and a loop connecting S1 and H3 (blue) are flexible (indicated by dashed lines). Right panel: Surface representation of the PyID:holoenzyme structure. Disordered regions are shown in accordance with the ribbon drawing. c) Left panel: Ribbon drawing of PyID cocrystallized with **3a** and NAD⁺ (closed conformation). The active site contains the furnished product **4a** (C atoms of the lysine moiety are shown in yellow, C atoms of the pyrroline moiety in green); further color coding is according to (b). Note: all residues of the structure shown in (c) are well defined in the electron density map. Right panel: Surface representation of the PyID:**4a** structure.

NAD⁺ (Figure 2a). Both hydroxy groups of the adenosyl moiety are hydrogen-bonded by the strictly conserved side chain of Asp171. Notably, in the absence of substrate, the four N-terminal amino acids and residues 56–59 of a loop connecting S1 and H3 are structurally disordered and thus are not defined in the electron density map. Therefore, the coenzyme

complex of PyID reveals an open cavity at the interface between the two β sheets that serves as the substrate binding site (Figure 1b).

To investigate the reaction trajectory of PyID, we synthesized the substrate analogue **3a** (for details see the Supporting Information) and determined the structure of PyID cocrystal-

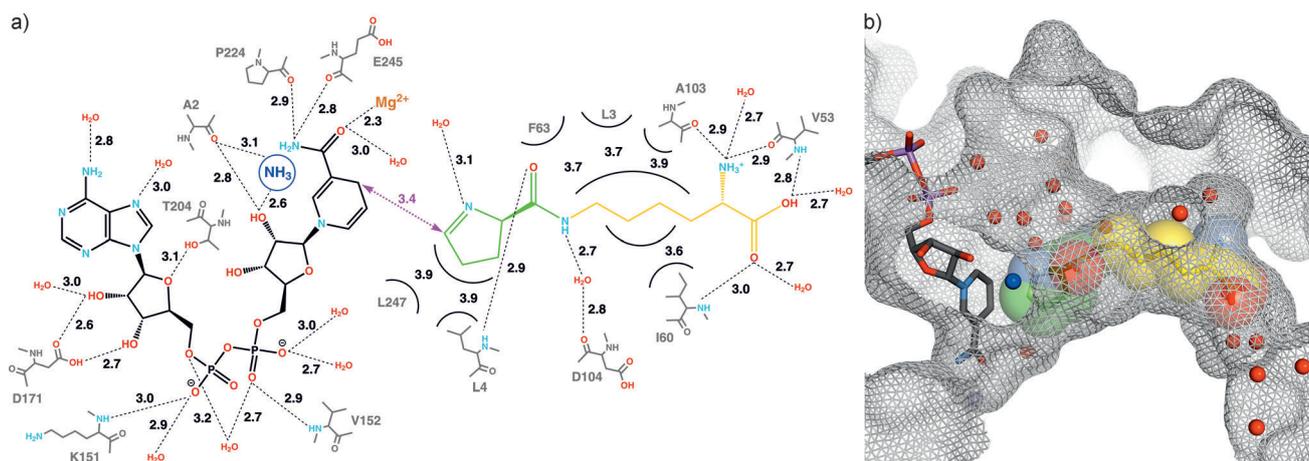


Figure 2. a) Amino acid residues in contact with NADH (black) and **4a** (pyrroline moiety: green, lysine moiety: yellow). The interaction distances are shown in Å. b) Connolly surface representation of the active site of PyID. The atoms of ligand **4a** are presented as spheres with their corresponding van der Waals radii. Solvent molecules are drawn as smaller spheres (water: red, ammonia: blue). Note: the ammonia is located inside a channel leading to the protein surface. Orientation of the structure is according to Figure 1. The stereoview of Figure 2b is presented in the Supporting Information (Figure S2).

lized with NADH and **3a** (PylD:**3a**, PDB ID: 4J4B) to 1.9 Å resolution ($R_{\text{free}} = 24.0\%$, Table S1). Remarkably, substrate binding is accompanied by breaking of helix H1, resulting in a major topological reorganization of the N-terminal segment that is conducive to closure of the substrate cleft by the proximal part of the broken helix (residues 1–11, colored in red in Figure 1). As a consequence, the bound substrate is almost entirely enclosed in a narrow, hydrophobic tunnel formed by residues Leu3, Leu4, Ile60, Phe63, and Leu247 as shown in Figure 2a. The C- and N-terminal end of the substrate and the isopeptide bond are hydrogen-bonded to backbone elements and water molecules (Figure 2a). The ornithine motif of the substrate analogue adopts a sickle-shaped conformation, with the amino and amide nitrogen atoms in close proximity (2.0 Å). At a distance of 2.6 Å to C4 of the pyridine ring of the cofactor, the pro-*R* hydrogen atom of the aminomethylene group appears poised for abstraction as hydride; thus, the orientation of NADH in the PylD:**3a** complex structure towards the substrate suggests that PylD acts as an *R*-type dehydrogenase (Figure 3a).

A further approach was to confirm that PylD crystallized in an active state and that the induced-fit mechanism is part of the catalytic reaction. We therefore soaked a PylD/NAD⁺ cocrystal with **3a** (PylD:soak, PDB ID: 4J49); we observed visible changes of the crystal morphology and transient formation of bubbles inside the crystal. The substrate-treated crystal still diffracted to a resolution of 2.2 Å ($R_{\text{free}} = 20.1\%$, Table S1). Elucidation of the structure revealed two molecules in the asymmetric unit; whereas one molecule shows the open conformation without bound substrate or product, the electron density for the second molecule indicates that **3a** had been converted to the product pyrroline-carboxy-lysine (Pcl, **4a**), though the ligand displays only partial occupancy as indicated by an increased Debye–Waller factor. We thus cocrystallized PylD in the presence of **3a** and NAD⁺ and determined the complex structure (PylD:**4a**, PDB ID: 4J4H) to 1.8 Å resolution ($R_{\text{free}} = 21.8\%$, Table S1). The overall architecture of the structure does not differ from that of the PylD:**3a** complex (Figure 3b); however, the differences between the substrate and the furnished product **4a** are clearly depicted in the omit electron density maps (Figure 3c).^[13] Furthermore, there is a most notable change in the appearance of a fixed small molecule that is tentatively interpreted as ammonia produced as the second reaction product besides **4a** (Figure 3b), whereas all defined solvent molecules embracing the ligands are present in both the substrate and product PylD crystal structures with a root-mean-square deviation (r.m.s.d.) of less than 0.2 Å (Figure 3c).

Our crystallographic data illustrate that PylD is a member of the large FAD/NAD(P)-binding Rossmann-fold superfamily. Nevertheless, no assignment to a specific subgroup can be made on the basis of structural arguments. With a length of 259 amino acids, PylD is in the 250–300 residue range of the short-chain dehydrogenase/reductase family (SDR). However, the Rossmann motif of SDR members is typically located at the N-terminal end,^[14] as opposed to PylD where it extends from positions 143–243 and is followed by helix H10 which interacts closely with the N-terminal part housing the substrate-binding site. Even more importantly, SDR members whose structures so far have been determined contain a highly conserved tyrosine residue which serves as a base for the abstraction of a proton and thus facilitates the hydride transfer.^[15]

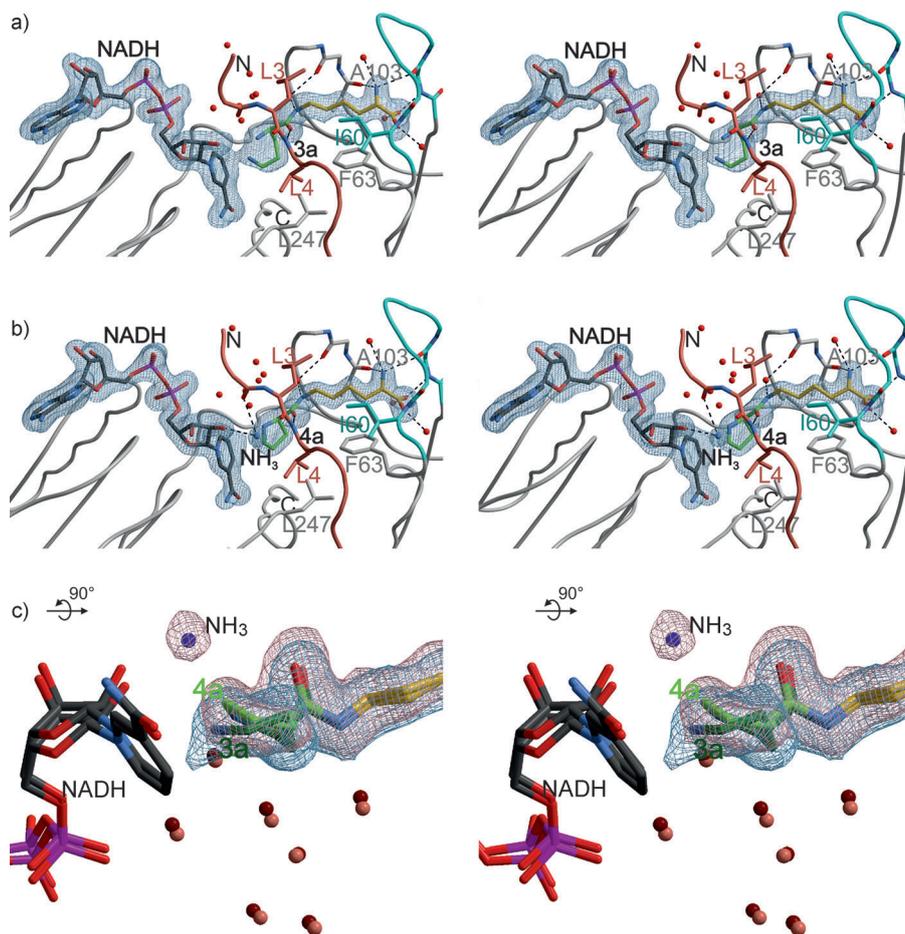
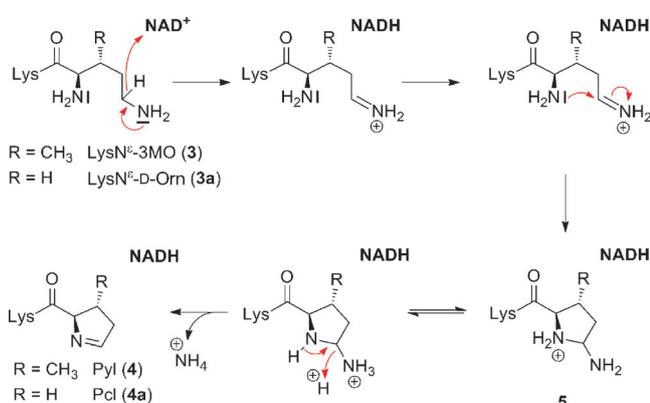


Figure 3. Stereoview of the active site of PylD. The $2F_o - F_c$ omit electron density map is contoured at 1.0σ . The N-terminus is shown in red and the loop connecting S1 and H3 is depicted in blue. The active-site residues and ligands are presented as stick models and solvent molecules are drawn as balls (red: water, blue: ammonia). Hydrogen bonds are indicated by dashed lines. a) PylD in complex with **3a**; b) PylD in complex with **4a**; c) superposition of **3a** (dark green, blue mesh) and **4a** (green, red mesh). Orientation of the structure is according to Figure 1.

PylD has neither a conserved tyrosine nor any other amino acid in its active site cavity that could act as an auxiliary base to support proton release from the amino nitrogen. We also note that to our knowledge dehydrogenases have not been described to undergo the type of induced-fit motility observed with PylD. Therefore, PylD appears to be without known precedent beyond the general assignment to the Rossmann-fold superfamily nor can any analogy to a specific family be based on its unique primary structure. Rather, it must be emphasized that the allowed sequence space for PylD's functionality is, per se, very large. In a pairwise comparison of the 22 hypothetical PylD sequences in the Universal Protein Resource (UniProt) database, identities as low as 22% are observed (Table S2).

Based on the crystal structures of PylD:holo, PylD:3a, and PylD:4a, we propose the reaction mechanism shown in Scheme 2. The observed topology of the substrate analogue



Scheme 2. Proposed reaction mechanism of PylD.

3a and the pyridine nucleoside coenzyme (in its reduced, that is, nonreactive form) suggests that the pro-*R* hydrogen at the terminal aminomethylene group of **3a** is well positioned for hydride abstraction by the coenzyme in the oxidized form. The resulting, protonated and positively charged imine carbon is perfectly suitable for a nucleophilic attack by the α amino group of the ornithine moiety, thus enabling the closure of the five-membered pyrrolidine ring. Alternatively, the abstraction of the hydride ion and the attack by the α amino group could be achieved in a concerted process, thus explaining the absence of a base for proton abstraction from the N-terminal nitrogen of the substrate. Notably, the absence of a proton acceptor in the active site sets PylD apart from other alcohol dehydrogenases and amine dehydrogenases. Subsequent to the formation of the 2-aminopyrrolidine intermediate **5**, the transfer of a proton from the ring nitrogen to the amino group at C5 of **5** could be mediated by a network of hydrogen-bonded water molecules (Figure 2b and Figure 3). Thus, the proton transponders could set the stage for the elimination of the amino group affording the pyrroline ring. In conclusion, our structural data provide snapshots of different steps of the PylD reaction trajectory and strongly support the proposed short mechanism rather than a more

complex pathway, for example through the formation of an aldehyde intermediate.

Overall, the PyIBCD biosynthetic machinery essentially encompasses three chemical reactions, each representing a challenge for the synthetic chemist. Whereas PylB and PylC catalysis is strictly dependent on their substrates **1**, **2**, and **2a**, active PylD literally harbors enough room for innovation beyond nature's scope of substrates and products. In particular, its binding cavity for the ornithine moiety is large enough to accommodate a variety of different substrate analogues. The presented results provide new directions for future investigations on pyrrolysine derivatives synthesized and incorporated into a defined primary sequence by means of the PylDST machinery in vivo.

Received: February 8, 2013

Revised: March 9, 2013

Published online: May 29, 2013

Keywords: amino acids · dehydrogenases · *Methanosarcina barkeri* · PylD · pyrrolysine

- [1] C. Hertweck, *Angew. Chem.* **2011**, *123*, 9712–9714; *Angew. Chem. Int. Ed.* **2011**, *50*, 9540–9541.
- [2] a) G. Srinivasan, C. M. James, J. A. Krzycki, *Science* **2002**, *296*, 1459–1462; b) B. Hao, W. Gong, T. K. Ferguson, C. M. James, J. A. Krzycki, M. K. Chan, *Science* **2002**, *296*, 1462–1466; c) J. A. Soares, L. Zhang, R. L. Pitsch, N. M. Kleinholz, R. B. Jones, J. J. Wolff, J. Amster, K. B. Green-Church, J. A. Krzycki, *J. Biol. Chem.* **2005**, *280*, 36962–36969.
- [3] a) A. Mahapatra, A. Patel, J. A. Soares, R. C. Larue, J. K. Zhang, W. W. Metcalf, J. A. Krzycki, *Mol. Microbiol.* **2006**, *59*, 56–66; b) S. K. Blight, R. C. Larue, A. Mahapatra, D. G. Longstaff, E. Chang, G. Zhao, P. T. Kang, K. B. Green-Church, M. K. Chan, J. A. Krzycki, *Nature* **2004**, *431*, 333–335; c) C. Polycarpo, A. Ambrogelly, A. Berube, S. M. Winbush, J. A. McCloskey, P. F. Crain, J. L. Wood, D. Söll, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12450–12454.
- [4] a) M. A. Gaston, L. Zhang, K. B. Green-Church, J. A. Krzycki, *Nature* **2011**, *471*, 647–650; b) S. E. Cellitti, W. Ou, H. P. Chiu, J. Grunewald, D. H. Jones, X. Hao, Q. Fan, L. L. Quinn, K. Ng, A. T. Anfora, S. A. Lesley, T. Uno, A. Brock, B. H. Geierstanger, *Nat. Chem. Biol.* **2011**, *7*, 528–530.
- [5] F. Quitterer, A. List, W. Eisenreich, A. Bacher, M. Groll, *Angew. Chem.* **2012**, *124*, 1367–1370; *Angew. Chem. Int. Ed.* **2012**, *51*, 1339–1342.
- [6] F. Quitterer, A. List, P. Beck, A. Bacher, M. Groll, *J. Mol. Biol.* **2012**, *424*, 270–282.
- [7] a) E. Kaya, M. Vrabel, C. Deiml, S. Prill, V. S. Fluxa, T. Carell, *Angew. Chem.* **2012**, *124*, 4542–4545; *Angew. Chem. Int. Ed.* **2012**, *51*, 4466–4469; b) M. J. Gattner, M. Vrabel, T. Carell, *Chem. Commun.* **2013**, *49*, 379–381; c) B. H. Geierstanger, W. Ou, S. Cellitti, T. Uno, T. Crossgrove, H. P. Chiu, J. Grunewald, X. Hao, International patent application WO/2010/048582, **2010**; d) P. R. Chen, D. Groff, J. Guo, W. Ou, S. Cellitti, B. H. Geierstanger, P. G. Schultz, *Angew. Chem.* **2009**, *121*, 4112–4115; *Angew. Chem. Int. Ed.* **2009**, *48*, 4052–4055; e) M. G. Hoesl, N. Budisa, *Angew. Chem.* **2011**, *123*, 2948–2955; *Angew. Chem. Int. Ed.* **2011**, *50*, 2896–2902.
- [8] D. Turk, *PhD Thesis*, Technische Universität München **1992**.
- [9] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A.

- Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr. Sect. D* **2011**, *67*, 235–242.
- [10] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. Sect. D* **1997**, *53*, 240–255.
- [11] L. Holm, P. Rosenström, *Nucleic Acids Res.* **2010**, *38*, W545–549.
- [12] G. Kleiger, D. Eisenberg, *J. Mol. Biol.* **2002**, *323*, 69–76.
- [13] The asymmetric unit contains four PylD molecules, each in complex with the product. The r.m.s.d. of C_α atoms among subunits is less than 0.2 Å.
- [14] B. Persson, Y. Kallberg, U. Oppermann, H. Jornvall, *Chem.-Biol. Interact.* **2003**, *143*, 271–278.
- [15] U. Oppermann, C. Filling, M. Hult, N. Shafqat, X. Wu, M. Lindh, J. Shafqat, E. Nordling, Y. Kallberg, B. Persson, H. Jornvall, *Chem.-Biol. Interact.* **2003**, *143–144*, 247–253.
-