Nine Enzymes Are Required for Assembly of the Pacidamycin Group of Peptidyl Nucleoside Antibiotics

Wenjun Zhang, Ioanna Ntai, Megan L. Bolla, Steven J. Malcolmson, Daniel Kahne, Neil L. Kelleher, and Christopher T. Walsh*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Departments

of Chemistry and Molecular Biosciences, Northwestern University, and Department of Chemistry &

Chemical Biology, Harvard University

E-mail: Christopher_Walsh@hms.harvard.edu

SUPPORTING INFORMATION

Table of Contents

Supplementary Methods

Supplementary Tables

Table S1. Primers used in this study.

Table S2. Assay components for pacidamycin analogs production.

Supplementary Figures

Figure S1. Map of pacidamycin gene cluster and SDS-PAGE analysis of the E. coli purified proteins.

Figure S2. Synthesis and ¹H characterization of DABA.

Figure S3. Time course of PacV catalyzed methylation.

- Figure S4. Theoretical MS calculation of PacH-bound biosynthetic intermediates.
- *Figure S5.* ¹H characterization of synthesized 5'-aminouridine.
- *Figure S6.* HR-MS and HR-MS/MS of **1**.
- Figure S7. HR-MS and HR-MS/MS of 2.
- *Figure S8.* HR-MS and HR-MS/MS of **3.**
- *Figure S9.* HR-MS and HR-MS/MS of **4.**
- *Figure S10.* HR-MS and HR-MS/MS of **5**.
- *Figure S11.* HR-MS and HR-MS/MS of **6.**
- *Figure S12.* HR-MS and HR-MS/MS of 7.
- Figure S13. HR-MS and HR-MS/MS of 8.
- *Figure S14.* HR-MS and HR-MS/MS of **9.**
- Figure S15. HR-MS and HR-MS/MS of 10.
- Figure S16. HR-MS and HR-MS/MS of 11.
- *Figure S17.* HR-MS and HR-MS/MS of **12.**
- *Figure S18.* HR-MS and HR-MS/MS of **12**' produced from [2,3-¹³C2]L-Ala.
- Figure S19. Biosynthetic pathway for 1.
- Figure S20. Proposed complete biosynthetic pathway for pacidamycin S.
- References

Supplementary Methods

Cloning, overexpression and purification of proteins. Cloning, expression, and purification of PacJLNOPU have been reported previously.¹ The genes of pacH, W, V, D, I were PCR amplified from genomic DNA extracted from S. coeruleorubidus NRRL 18370, and the gene of PacH' was PCR amplified from genomic DNA extracted from S. roseosporus NRRL 15998 (primers listed in table S1). Purified PCR products were ligated to pET-24b or pET-30 Xa/LIC (Novagen) following the standard protocol and confirmed by DNA sequencing. The resulting expression constructs were transformed into E. coli BL21 or BAP1 cells for protein expression. Expression and purification for all proteins with His6-tag followed the same general procedure and is detailed as follows. In 1 L of liquid culture, the cells were grown at 37 °C in LB medium with 50 µg/mL kanamycin to an OD600 of 0.4. The cells were cooled on ice for 10 min and then induced with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 16 h at 16 °C. The cells were harvested by centrifugation (6000 rpm, 6 min, 4 °C), resuspended in 30 mL lysis buffer (25 mM HEPES pH 8.0, 0.5 M NaCl, 5 mM imidazole) and lysed by sonication on ice. Cellular debris was removed by ultracentrifugation (35000 rpm, 35 min, 4 °C). Ni-NTA agarose resin was added to the supernatant (1 mL/L of culture) and the solution was nutated at 4 °C for 1 h. The protein resin mixture was loaded into a gravity flow column, and proteins were eluted with increasing concentrations of imidazole in Buffer A (50 mM HEPES, pH 8.0, 1 mM EDTA). Purified proteins were concentrated and buffer exchanged into Buffer A + 10% glycerol using Amicon Ultra filters. The final proteins were flash-frozen in liquid nitrogen and stored at -80 °C. The approximate protein yields were 12 mg/L for PacH (10 kDa), 8.3 mg/L for PacH' (10 kDa), 1.1 mg/L for PacD (43 kDa), 1.0 mg/L for PacI (47 kDa), 4.3 mg/L for PacW (56 kDa), and 15 mg/L for PacV (28 kDa).

DABA synthesis. The synthesis of (2R,3R)-diaminobutanoic acid and (2S,3S)-diaminobutanoic acid was accomplished using the procedure developed by Davies *et al* (Figure S2).^{2,3} The experimental details for the complete synthesis of (2S,3S)-diaminobutanoic acid is detailed here.

(*S*)-(-)-*N*-benzyl- α -methyl-benzyl amine (645 mg, 3.05 mmol, 1.4 equiv) was added to a round bottom flask containing THF (16.0 mL) and cooled to –78 °C. A solution of *n*-butyl lithium (1.6M/hexanes) was then added drop-wise to the reaction flask over the 15 minutes. Upon addition the reaction immediately turned pink in color, the solution was maintained at –78 °C. After 30 minutes a solution of *tert*-butyl crotonate (311 mg, 2.18 mmol) in THF (8.00 mL) cooled to –78 °C was canulated drop-wise into the reaction flask. The solution turned orange and then slowly back to pink over 2h at –78 °C. (*S*)-(+)-(10-camphor-sulfonyl)-oxaziridine (1.0 g, 4.36 mmol, 2.0 equiv) was then added and the reaction was warmed to ambient temperature. After 16 h saturated aqueous ammonium chloride (10 mL) was added to the reaction. The mixture was added to a separatory funnel containing brine (10 mL) and then a mixture of CH₂Cl₂:Et₂O (1:1, 10 mL) was also added. The layers were separated and the aqueous layer was extracted with CH₂Cl₂:Et₂O (1:1, 20 mL x 2). The combined organics were dried over magnesium sulfate, filtered, and concentrated *in vacuo*. Purification by flash column chromatography (5% EAc:Hx) afforded **13** as a colorless oil, R_f = 0.35 EAc:Hx (4:2), 76% yield. ¹H and ¹³C



Triphenyphosphine (569 mg, 2.17 mmol, 2.1 equiv) was weighed into a round bottom flask and put in solution with THF (12 mL) at ambient temperature. **13** (402 mg, 1.08 mmol), and diethyl azodicarboxylate (377 mg, 2.17 mmol, 2.0 equiv) were then added. Diphenylphosphorylazide (3.51 mL, 16.2 mmol, 15.0 equiv) was then added drop-wise over 30 minutes. After 40 h the reaction was concentrated *in vacuo* and purified using flash column chromatography 5% EAc:Pet to afford **14** as a colorless oil, $R_f = 0.7$ EAc:Pet (1:6), 64% yield. ¹H and ¹³C NMR matched those previously reported.



Triphenylphosphine (270 mg, 1.03 mmol, 1.5 equiv) was weighed into a round bottom flask and put in solution with THF (2.8 mL) at ambient temperature. **14** (271 mg, 0.687 mmol) was then added followed by water (440 μ L). After 24 h the solution was concentrated *in vacuo* and purified using flash column chromatography Pet:EAc:MeOH (79:20:1) to afford **15** as a white solid, R_f = 0.23 Pet:EAc:MeOH (79:20:1), 90% yield.



15 (224 mg, 0.608 mmol) was put in solution with ethanol (2 mL), palladium hydroxide (102 mg) was then added. The reaction flask was placed in a Parr bomb, purged with hydrogen gas, closed, and then filled to 70 psi with hydrogen. Reaction vessel was then heated to 55 °C. After 20 h the reaction mixture was filtered through celite washing with ethanol (10 mL) and then concentrated *in vacuo*. Product was then triturated with Hex (3 mL) then Et₂O (3 mL) to provide a white/yellow solid. LCMS and ¹H NMR confirmed that the product was **16**.



To a round bottom flask containing **16** and cooled to 0 °C was added trifluoroacetic acid (3 mL). The reaction was then warmed to room temperature. After 19 h the crude solution was concentrated *in vacuo* and then

aqueous 1 M HCl solution was added (3 mL). After 4.5 h the solution was concentrated *in vacuo* to provide **17**. LCMS and ¹H NMR confirmed the product and matched what had been previously published. The yield over the last two steps is 72%.



5'-aminouridine synthesis. The synthesis of 5'-aminouridine was accomplished using the procedure developed by Winans *et al.*⁴

Under an inert atmosphere of N₂, an oven-dried (160 °C) 100-mL round-bottom flask, equipped with a magnetic stir bar, was charged with uridine (2.50 g, 10.2 mmol), *p*-toluenesulfonic acid monohydrate (325 mg, 1.71 mmol), and crushed activated 4 Å molecular sieves (980 mg). Dry DMF (30 mL) was added and the mixture allowed to stir. Once the solids had all dissolved, 2,2-dimethoxypropane (5.0 mL, 41 mmol) was added and the mixture heated to 40 °C. After 1.5 h, the mixture was allowed to cool to ambient temperature and was neutralized with Amberlyst A-21 free base resin (ca. 0.5 g). The solids were filtered through a pad of Celite and the supernatant concentrated to a viscous light yellow oil, which could be solidified by concentration from a MeOH/EtOAc/hexanes mixture. The off-white solid was then purified by silica gel chromatography (dry-load method, 15:1 to 9:1 CH₂Cl₂:MeOH) to yield acetonide **18** (720 mg, 2.53 mmol, 24.8% yield) as a white solid. Physical and spectral data matched that previously reported.



Under an inert atmosphere of N₂, an oven-dried (160 °C) 25-mL round-bottom flask, equipped with a magnetic stir bar, was charged with **18** (357 mg, 1.26 mmol), *p*-toluenesulfonic anhydride (656 mg, 2.01 mmol), and CH_2Cl_2 (5 mL). Pyridine (1.00 mL, 12.6 mmol) was added and the mixture allowed to stir; the solids dissolved and the mixture warmed slightly. The flask was fitted with an oven-dried (160 °C) reflux condenser and the mixture heated to a gentle reflux. After 2.5 h, the mixture was allowed to cool to ambient temperature and was poured into CH_2Cl_2 (25 mL). The mixture was then washed with 0.5 M aqueous HCl (4 X 15 mL) and saturated aqueous NaHCO₃ (40 mL). The organic layer was dried over Na₂SO₄, filtered, and concentrated to a yellow solid, which was clean enough to be carried forward without purification. Physical and spectral data matched that previously reported.



Under an inert atmosphere of N₂, an oven-dried (160 °C) 25-mL round-bottom flask, equipped with a magnetic stir bar, was charged with sodium azide (246 mg, 3.78 mmol; **WARNING:** sodium azide may be explosive – handle with care). A solution of **19**, obtained directly from the tosylation reaction above, in 2.5 mL DMF was transferred to the reaction vessel by cannula; the flask containing **19** was rinsed with 0.5 mL DMF and similarly transferred. The mixture was allowed to stir and was heated to 45 °C. After 18 h, the mixture was allowed to cool to ambient temperature and the solids were filtered through a plug of cotton, eluting with copious CH_2Cl_2 . The solvent was removed by distillation, leaving behind a yellow solid. The material was then purified by silica gel chromatography (dry-load method, 20:1 CH_2Cl_2 :MeOH) to deliver azide **20** (189 mg, 0.613 mmol, 48.7% yield over two steps) as an off-white solid (**WARNING:** alkyl azides may be explosive – handle with care). Physical and spectral data matched that previously reported.



Under an atmosphere of air, to the 10-mL vial containing azide **20** (189 mg, 0.613 mmol; **WARNING:** alkyl azides may be explosive – handle with care) was added a magnetic stir bar and a 9:1 solution of TFA:H₂O (3 mL total volume). The mixture was allowed to stir for 40 min, at which time the solvent was removed *in vacuo*. Residual water was removed by azeotropic vacuum distillation with benzene (2 mL). The resulting viscous yellow oil was purified by silica gel chromatography (dry-load method, 15:1 to 9:1 CH₂Cl₂:MeOH) to afford diol **21** (151 mg, 0.561 mmol, 91.5% yield) as a white solid (**WARNING:** alkyl azides may be explosive – handle with care). Physical and spectral data matched that previously reported.



Under an atmosphere of air, to the 10-mL vial containing diol **21** (151 mg, 0.561 mmol; **WARNING:** alkyl azides may be explosive – handle with care) was added a magnetic stir bar and MeOH (3.5 mL). 10% Pd/C (30 mg) was then added and the vial fitted with a rubber septum; the mixture was allowed to stir. The atmosphere was then purged with a balloon of hydrogen gas and the reaction vessel then fitted with a fresh balloon of hydrogen gas. After 75 min, the mixture was fitted through a plug of Celite, eluting with MeOH. The solution was concentrated *in vacuo* to afford an off-white solid. The material was then triterated by dissolving in minimal MeOH, addition of one-half volume of EtOAc, and then rapid addition of 10 volumes of hexanes. The

solid was filtered and allowed to dry, delivering 5'-aminouridine **22** (101.5 mg, 0.417 mmol, 74.3% yield) was a white solid. Physical and spectral data matched that previously reported.



ATP-PP_i **Exchange Assays.** The assays were performed in 100 μ L of reaction buffer (50 mM Tris-HCl/2 mM MgCl₂, pH 7.8) containing 5 mM ATP, 1 mM Na₄[³²P]PP_i (~4 X 10⁶ cpm/mL), 1 mM TCEP, 5 mM substrate, and 1 μ M enzyme. Reactions were incubated at 25 °C for 1 h then quenched by the addition of charcoal suspension (1.6% w/v activated charcoal, 0.1 M Na₄PP_i, 3.5% HClO₄). Free [³²P]PP_i was removed by centrifugation of the sample followed by washing twice with wash solution (0.1 M Na₄PP_i and 3.5% HClO₄). Charcoal-bound radioactivity was measured on a Beckman LS 6500 scintillation counter.

Loading Assays with ¹⁴C-labeled Substrates. A typical assay contained, in a total volume of 25 μ L, 5 mM ATP, 2 mM MgCl₂, 1 mM TCEP, 20-100 μ M amino acids and SAM, 10 μ M enzymes, and 50 mM HEPES, pH 8.0. ¹⁴C-labeled substrate was added to each reaction accordingly (L-Ala [0.25 μ Ci], L-Phe [0.12 μ Ci], or SAM [0.07 μ Ci]. After 2 h incubation at 25 °C, samples were quenched by adding 1× SDS sample buffer. Following SDS-PAGE, radiolabeled protein was detected using a BAS-III imaging plate (Fuji Film, 48–96 h exposure) and a Typhoon 9400 phosphorimager (GE Healthcare).

Methylation time course assays with ³H-labeled SAM. The assays were performed in 150 μ L of reaction buffer (50 mM HEPES/2 mM MgCl₂, pH 8) containing 5 mM ATP, 1 mM TCEP, 5 mM DABA or DAP, 20 μ M PacP, 1 μ M PacV and 2.75 μ Ci SAM (0.3 nmoles). Reactions were incubated at 25 °C and 25 μ L samples were quenched at different time points by adding 0.5 mL of 10% TCA (with 50 μ g of BSA for visualization of precipitated protein). Protein precipitate was pelleted by centrifugation, washed twice with 10% TCA, and dissolved in 80% formic acid for liquid scintillation counting. A "zero" time point was taken by processing a reaction prior to the addition of PacV. In a parallel set of assays, DABA/DAP was incubated with PacP first for 1 h, buffer exchanged using Amicon centrifugal filter unit (3 kDa MWCO) to remove ATP and DABA/DAP before adding PacV and SAM.

LC-FTMS analysis of PacH-bound biosynthetic intermediates. Assays were performed in 100 μ L of 50 mM HEPES (pH 8.0) containing 5 mM ATP, 2 mM MgCl₂, 1 mM TCEP, 5 mM amino acids, 0-10 μ M PacPUJLOND and 50 μ M PacH. After ~4 h incubation at 25 °C, 0.1 M Tris and trypsin (1:5 w:w trypsin:total protein) were added and further incubated at 30 °C for 15 min. The reactions were quenched with 25% formic acid and analyzed by nano-capillary LC-MS using a 100 mm x 75 μ m C18 column in-line with a LTQ-FT (7 T). All MS methods included the following events: 1) FT scan, *m*/*z* 400–2,000, 2) data-dependent MS/MS on the top 3 peaks in each spectrum from scan event 1 using collision-induced dissociation (CID) with the following parameters: detection of all ions in the ion trap MS in profile mode, isolation width 5 *m*/*z*, activation q value 0.25, activation time 30 ms, NCE 35, and 3) FT scan, source-induced dissociation (SID) = 75, detect *m*/*z* 200–760 (Ppant ejection assay). All data were analyzed using QualBrowser, part of the Xcalibur software packaged with the ThermoFisher LTQ-FT. All mass values reported are for the neutral monoisotopic peaks.

LC-HRMS product assays. Assays were performed in 100 μ L of 50 mM HEPES (pH 8.0) containing 5 mM ATP, 2 mM MgCl₂, 1 mM TCEP, 5 mM amino acids and uridine, 0-10 μ M PacPUJLOND and 50 μ M PacH (Table S2). After 4 h incubation at 25 °C, the proteins were removed by 3 kDa MWCO filter tubes, and the filtered reaction mixture was subjected to LC-HRMS and MS/MS analysis using an Agilent Technologies 6520 Accurate-Mass Q-TOF LC-MS instrument and a 75 mm x 4.6 mm Luna C₁₈ column. A linear gradient of 2 to 80% CH₃CN (v/v) over 15 min in H₂O supplemented with 0.1% (v/v) formic acid at a flow rate of 0.5 mL/min was used.

Supplementary Tables

Table S1. Primers used in this study.

Primer	Sequence	Note
111a_NdeI	5'- AAAAAACATATGAATCTACAGGATCAGAA-3'	Used for PacH expression
111a_EcoRI	5'- AAAAAAGAATTCGGTTTCCTGGTGCCGTTCGCAC-3'	
Hc_NdeI	5'- AAAAAACATATGAGTCTGCAGAACGCAAA-3'	Used for PacH' expression
Hc_EcoRI	5'- AAAAAAGAATTCGGCTTCTCGCCGCCGTGCACCG-3'	
PacW_PET30F	5'- GGTATTGAGGGTCGCATGTCTCTCACATTGGTCGA-3'	Used for PacW expression
PacW_PET30R	5'- AGAGGAGAGTTAGAGCCTTAGCCAGGCAACTCCTCCGCC-3'	
126_PET30F	5'- GGTATTGAGGGTCGCATGTCGGACAATGATGCTCG-3'	Used for PacV expression
126_PET30R	5'- AGAGGAGAGTTAGAGCCTTAGGCCTGGGGAGCGCCGTT-3'	
108_PET30F	5'- GGTATTGAGGGTCGCATGACCGACAGGCTTCCCTT-3'	Used for PacD expression
108_PET30R	5'- AGAGGAGAGTTAGAGCCTTAGGATTTCCGCTCGGGAAAG-3'	
112_PET30F	5'- GGTATTGAGGGTCGCATGACATTGTCTCAGGTCAA-3'	Used for PacI expression
112_PET30R	5'- AGAGGAGAGTTAGAGCCTTAGCTGCCGGCGGACAGTGC-3'	

Table S2. Assay components for pacidamycin analogs production. ATP, Ala, PacPHJLONDI were included inall assays. See Methods "LC-HRMS product assays" for detailed protocol.

	1	2	3	4	5	6	7	8	9	10	11	12
Ar aa	F	W	<i>m</i> -Y	<i>m</i> -Y	<i>m</i> -Y	F	F	W	<i>m</i> -Y	<i>m</i> -Y	F	W
DABA	+	+	+	+	+	-	+	+	+	+	+	+
DAP	-	-	-	-	-	+	-	-	-	-	-	-
uridine	-	-	-	+	+	+	+	+	-	-	-	-
3'-deoxyuridine	-	-	-	-	-	-	-	-	+	+	+	+
5'-aminouridine	+	+	+	-	-	-	-	-	-	-	-	-
PacU	+	+	-	-	+	+	+	+	-	+	+	+
PacW	-	-	+	+	-	-	-	-	+	-	-	-



Figure S1. Map of pacidamycin gene cluster and SDS-PAGE analysis of the *E. coli* purified proteins. Miniprotein TGX gel (4-15% precast, Biorad) was used. PacH' is a homolog of PacH from *S. roseosporus*.

* indicates truncated domain or domain predicted to be catalytically inactive.







Figure S2. Synthesis and ¹H characterization of synthesized *2S,3S*-DABA. The synthesis was based on the

published methods.^{2,3}



Figure S3. Time course of PacV catalyzed SAM-dependent methylation. See Methods "Methylation time course assays with ³H-labeled SAM" for detailed protocol.



g)

Active site peptide	Mass	Mass	Error	Shift from apo-PacH	Shift from apo-PacH	
	(Da, Theor)	(Da, Exp)	(ppm)	(Da, Theor)	(Da, Exp)	
Apo-PacH (a)	1371.7470	not detected	-	-	-	
Holo-PacH (b)	1711.8328	1711.8346	1.0	340.0858	340.0876	
PacH-S-DABA (c)	1811.8965	1811.8990	1.4	440.1495	440.1529	
PacH-S-DABA ₃ -Ala ₂ (d)	1882.9336	1882.9360	1.3	511.1866	511.1932	
PacH-S-DABA ₃ -Ala ₄ -CO-Phe ₅ (e)	2073.9919	2073.9976	2.7	702.2449	702.2506	
PacH-S-DABA ₃ (Ala ₂)-Ala ₄ -CO-Phe ₅ (f)	2145.0290	2145.0319	1.3	773.2820	773.2849	

Figure S4. Theoretical MS calculation of PacH-bound biosynthetic intermediates. Panel a) shows the amino acid sequence of the PacH tryptic peptide with the active site serine highlighted in red. Panel b) shows the structure of the active site peptide after phosphopantetheinylation, while panels c)-f) show the PacH-bound biosynthetic intermediates. Panel g) includes calculations of the theoretical and experimental mass of the PacH active site peptides illustrated in panels a)-f). All mass values are for the neutral monoisotopic peaks.



Figure S5. ¹H characterization of synthesized 5'-aminouridine. The synthesis was based on the published methods.⁴



Figure S6. HR-MS and HR-MS/MS of 1 measured during LC-MS.



Figure S7. HR-MS and HR-MS/MS of 2 measured during LC-MS.



Figure S8. HR-MS and HR-MS/MS of 3 measured during LC-MS.



Figure S9. HR-MS and HR-MS/MS of 4 measured during LC-MS.



Figure S10. HR-MS and HR-MS/MS of 5 measured during LC-MS.



Figure S11. HR-MS and HR-MS/MS of 6 measured during LC-MS.



Figure S12. HR-MS and HR-MS/MS of 7 measured during LC-MS.







Figure S15. HR-MS and HR-MS/MS of 10 measured during LC-MS.



Figure S16. HR-MS and HR-MS/MS of 11 measured during LC-MS.



Figure S17. HR-MS and HR-MS/MS of 12 measured during LC-MS.





Figure S19. Biosynthetic pathway for 1.



Figure S20. Proposed complete biosynthetic pathway for pacidamycin S. The timing of 4',5'-ene formation in the uridine moiety is not yet examined: an enamino-deoxyuridine shown in the figure is only one possibility. The 4',5'-alkene could be installed by a dehydrogenation after condensation with the tetrapeptide.

References:

[1] Zhang, W.; Ostash, B.; Walsh, C. T. Proc. Natl. Acad. Sci. U S A. 2010, 107, 16828-16833.

[2] Bunnage, M. E.; Burke, A. J.; Davies, S. G.; Millican, N. L.; Nicholson, R. L.; Roberts, P. M.; Smith, A. D. *Org. Biomol. Chem.* **2003**, *1*, 3708-3715.

[3] Davies, S. G.; Garrido, N. M.; Kruchinin, D.; Ichihara, O.; Kotchie, L. J.; Price, P. D.; Price Mortimer, A. J.; Russell, A. J.; Smith, A. D. *Tet. Asymm.* **2006**, *17*, 1793-1811.

[4] Winans, K. A.; Bertozzi, C. R. Chem. Biol. 2002, 9, 113-129.