Supplementary Information

Investigation of Amide Bond Formation during Dehydrophos Biosynthesis

Emily C. Ulrich^{1,2}, Despina J. Bougioukou^{2,†}, and Wilfred A. van der Donk^{1,2,3*}

¹Department of Chemistry, ²Carl R. Woese Institute for Genomic Biology, ³Howard Hughes Medical Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States, † Present address: E. I. du Pont de Nemours and Company, Wilmington, Delaware 19803, United States

*Correspondence to: E-mail: vddonk@illinois.edu

Supplementary Methods

Reagents

Molecular Biology: Restriction enzymes (NdeI and DpnI), Phusion® polymerase, Gibson Assembly® master mix, and Taq DNA ligase were purchased from New England Biolabs, Inc. Fail Safe PCR buffers and polymerase and T5 exonuclease were purchased from Epicentre. DTT was purchased from Promega. dNTPs and MgSO4 were purchased from EMB Millipore. Ampicillin (100 μ g/mL) and chloramphenicol (12.5 μ g/mL) were purchased from Fisher Scientific and Sigma Aldrich.

tRNA Generation: 10X NEBufferTM 2, DNA polymerase I large Klenow fragment, rNTP mix, and thermostable inorganic pyrophosphatase (TIPP) were purchased from New England Biolabs, Inc. The dNTP mix was purchased from Novagen, Ambion® UltraPureTM BSA was purchased from Thermo Fisher Scientific. RNasin® RNase inhibitor and RNase-free DNase were purchased from Promega.

Enzymology: Nickel-nitrilotriacetic acid agarose was purchased from Qiagen. Isopropyl-β-Dgalactopyranoside was purchased from IBI Scientific. Amicon ultracentrifugal filters were purchased from EMB Millipore. PD10 desalting columns were purchased from GE Healthcare. Chelex 100 resin, DNase, and (R)-(−)-1-aminoethylphosphonic acid (L-AlaP) were purchased from Sigma Aldrich. L- $[$ ¹⁴C(U)]-amino acids and Ultima Gold cocktail were purchased from Perkin Elmer. Total tRNA from *E. coli* MRE 600 was purchased from Sigma Aldrich/Roche. The FemX(Wv)*¹* and I-TASSER*²* structures were visualized using PyMOL.*³*

Cloning Procedures

New constructs for this study (pET-15b harboring the *E. coli* isoleucyl-tRNA synthetase and phenylalanyl-tRNA synthetase genes) were prepared using DH5α genomic DNA (Table S2) as template and following cloning procedures outlined previously.*⁴*

Upon sequencing pET-15b valRS, we found that it encodes the mutations I520V and Y688H compared to a NCBI reference sequence of *valRS* from DH5α (NZ_JRBB01000028.1). These may be the natural amino acids in the DH5α genomic DNA that we used.

Quikchange Mutagenesis

Reactions contained 1X Pfu Ultra buffer (Agilent Quikchange II Mutagenesis Kit), 0.6 ng/ μ L of pET-15b DhpH-C as template, 0.25 µM of each primer, 0.04 mM dNTPs (Novagen), 5% v/v DMSO, and 0.05 U/µL Pfu Ultra HF polymerase. For R290A, R296A, N325A, and E326A, the reactions contained 1X Buffer G (FailSafe kit), 0.6 ng/µL of pET-15b DhpH-C, 0.5 µM of each primer, and 0.04 U/µL of Phusion® HF polymerase (NEB). All of these reactions were run through the following protocol: 95 ºC for 1 min, 20 cycles of 95 ºC for 1 min, 60 ºC for 1.5 min, 72 ºC for 15 min, then 72 ºC for 5 min. For R233A, W284L, and R292A, the mutation was introduced using Gibson cloning.*⁵* For example, primers DhpHC_W284L_F and pET15b NdeI R were used to amplify around the vector, while DhpHC_W284L_R and NdeIdhpH_C353 were used to amplify the gene with pET-15b DhpH-C as template. The reaction contained 1X of Buffer G (FailSafe kit), 0.5 µM of each primer, 0.9 ng/µL template, and 0.04 U/ μ L Phusion® polymerase. Each fragment was amplified using the following touch down PCR protocol: 98 ºC for 3 min, 30 cycles of 98 ºC for 10 s, 70 ºC for 30 s while dropping 0.5 ºC each round, 72 ºC 8 min (vector) or 2 min (gene), then 72 ºC for 10 min. All samples, regardless of amplification method, were treated with DpnI and purified using the QIAquick PCR purification kit (Qiagen). For Gibson ligation, a 2:1 ratio of insert:vector was combined with assembly master mix from either NEB or prepared as previously described*⁴* (for W284L). Samples were incubated at 50 °C for 1 h. For all samples, an aliquot (\sim 3 μ L) was used to transform 50 μ L of electro-competent DH5α cells. Recovered cells were plated on LB/amp100 plates, which grew overnight at 37 ºC. Colonies were picked for plasmid preparation using the QIAprep Spin Miniprep Kit (Qiagen). Correct mutations were verified by sequencing either at ACGT, Inc. or the UIUC Core Sequencing Facility.

DhpH-C and aaRS Expression and Purification

All proteins were expressed and purified as previously described.*⁴*

In vitro Transcription of tRNALeu

Sequences for *E. coli* tRNA^{Leu} (UAG) and (GAG) were obtained from the EcoCyc database.⁶ dsDNA templates were prepared based on a previously described method.*⁷* A 500 µL reaction for template generation contained: 1X Buffer 2 (New England Biolabs, Inc.), 4 µM of each primer, 40μ M of dNTPs, and $0.06 \text{ U}/\mu\text{L}$ DNA polymerase I (large Klenow fragment). The reaction was incubated at 25 ºC for 15 min before the addition of EDTA (10 mM final concentration) and inactivation by heating at 75 ºC for 20 min. The DNA was precipitated in EtOH overnight at -20 ºC. The resulting pellet was washed three times with ice cold 75% EtOH and air-dried before resuspension in nuclease-free water. The entire template yield was used in either a 1000 or 1500 µL *in vitro* transcription reaction based on a previously described method.*⁸* The reaction contained 1X transcription buffer (100 mM HEPES pH 7.5, 10 mM $MgCl₂$, 2 mM spermidine, 40 mM DTT, and 0.1 mg/mL RNase-free BSA), 10 mM DTT, 26 mM MgCl2, 7.5 mM of each NTP, 0.8 U/ μ L RNasin® RNase inhibitor, 0.01 U/ μ L TIPP, ~30 ng/ μ L template DNA, and 140 ng/µL T7 RNA polymerase. The reaction was incubated overnight at 37 ºC, then 5 U of RNasefree DNase was added per 500 µL transcription reaction, and incubation continued for 1 h at 37 ºC. The reaction was desalted by filtration using a 10 kDa molecular weight cut-off centrifugal filter (Millipore; 4 wash cycles with 50 mM HEPES pH 7.5). The tRNA was purified using an acidic phenol / 24:1 chloroform:isoamyl alcohol extraction described previously, ending with precipitation in EtOH and resuspension in 2 mM NaOAc pH 5.2.*⁹* The tRNA was precipitated in EtOH overnight at -20 ºC. The resulting pellet was washed with ice cold 75% EtOH and air-dried before resuspension in 2 mM NaOAc pH 5.2. tRNA concentration was determined using a Qubit® 2.0 Fluorometer and the RNA Broad-Range assay kit (Thermo Fisher Scientific).

Aminoacylation Assay for in vitro Transcribed and Total *E. coli* tRNAs

The reactions contained 37.5 mM HEPES pH 7.5, 112.5 mM KCl, 11.2 mM $MgCl_2$, 5 mM ATP, 0.01 U/ μ L TIPP, 8 μ M LeuRS, 50 μ M L-[¹⁴C(U)]-Leucine (328 mCi/mmol), and 16 μ M tRNALeu. Each solution was incubated at room temperature (to replicate conditions for DhpH-C activity assays) for 15 min (1 h for LeuRS*m assays) before quenching and preparation for analysis adapted from previously described protocols.^{4, 10-12} Briefly, an aliquot of 5 μ L was spotted onto 3MM Whatman filter paper squares presoaked in 5% trichloroacetic acid (TCA). The spots were allowed to dry for 5 min before the filter paper squares were washed 3 times in cold 5% TCA and once in cold 75% EtOH for 5 min each. The squares were air-dried and then analyzed by liquid scintillation counting in Ultima Gold scintillation cocktail using a Tri-Carb 2910TR. Experiments were performed in triplicate. Results were plotted using Igor Pro version 6.32A.

Aminoacylation assays with total *E. coli* tRNA were performed as described above, except that 3 mg/mL tRNA was used initially before adjusting the amount of tRNA to achieve equivalent aminoacylation levels for each aaRS (see amounts described for the activity assay). In these assays, 8 μ M of each aaRS was used along with the corresponding amino acid: L- $[14C(U)]$ isoleucine (308 mCi/mmol), L- $I^{14}C(U)$]-valine (271 mCi/mmol), and L- $I^{14}C(U)$]-phenylalanine (477 mCi/mmol).

Liquid Chromatography-Mass Spectrometry (LCMS) Analysis of aaRS Endpoint Assays To verify that the masses of the proposed dipeptides are present in the samples with various amino acids, samples of DhpH-C with IleRS, ValRS, PheRS, and LeuRS*m that had been quenched after 3 h were analyzed by LCMS. A sample with no DhpH-C and the Leu/LeuRS aatRNA generation system was used as background. Samples were analyzed by using the Q-Exactive MS system (Thermo. Bremen, Germany) in the Metabolomics Laboratory of Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign. Software Xcalibur 4.1.31.9 was used for data acquisition and analysis. The Dionex Ultimate 3000 series HPLC system (Thermo, Germering, Germany) used includes a degasser, an autosampler, and a binary pump. The LC separation was performed on a Phenomenex RP Fusion column (4.6 x 150 mm, 4 μm) with mobile phase A (water with 0.1% formic acid) and mobile phase B (acetonitrile with 0.1% formic acid). The flow rate was 1 mL/min. The linear gradient was as follows: 0 min, 100% A; 5-15 min, 5% A; 16-22 min, 100% A. The autosampler was set to 15°C. The injection volume was 10 μL. Mass spectra were acquired with electrospray ionization under both positive and negative mode: sheath gas flow rate, 65; aux gas flow rate: 20; sweep gas flow rate, 4; spray voltage, 4.0 kV; capillary temp, 300 °C; Aux gas heater temp, 500 °C. The resolution was set to 70,000 with the scan range of m/z 50 - 600. The AGC target was 1E6 with a maximum injection time of 200 ms.

The primers were parents on from the grace DTMT Teenhologies, the.			
Name	Sequence (mutation is in bold)		
DhpHC R233A F	geogeoagegeoacegeoaceggetegeocaag		
DhpHC_R233A R	cttgggcgagccggtggcggtggcgctggcggc		
DhpHC W284L F	caagcgcctgcaggcactcgccggagggtatgtg		
DhpHC W284L R	cacataccctccggcgagtgcctgcaggcgcttg		
DhpHC_R290A F	gccggagggtatgtggccggccgcgccgacctg		
DhpHC R290A R	caggtcggcgggccggccacataccctccggc		
DhpHC R292A F	gggtatgtgcgcggcgccgccgacctgcggttc		
DhpHC R292A R	gaaccgcaggtcggcggcgcgcgcacataccc		
DhpHC R296A F	ggccgcgccgacctggcgttcagcccgtactac		
DhpHC R296A R	gtagtacgggctgaacgccaggtcggcgcggcc		
DhpHC N325A F	gtgcggccggctcgccgaatccttcaaggagaag		
DhpHC N325A R	cttctccttgaaggattcggcgagccggccgcac		
DhpHC E326A F	gtgcggccggctcaacgcatccttcaaggagaag		
DhpHC E326A R	cttctccttgaaggatgcgttgagccggccgcac		
DhpHC F328L F	ccggctcaacgaatccctcaaggagaagatgctc		

Table S1. Primer Sequences

All primers were purchased from Integrated DNA Technologies, Inc.

* mU, mG: 2'-*O*-methylated nucleotide

Name	Features	Source
E. coli DH5α λpir	sup E44, ΔlacU169	$^{(13)}$
	$(\phi$ lacZ $\Delta M15)$, recA1, endA1,	
	hsdR17, thi-1, gyrA96, relA1,	
	λpir phage lysogen.	
E. coli Rosetta 2 (DE3)pLysS	\overline{CAM}^R ; provides seven rare, in	Novagen
	E. coli, tRNAs for the codons	
	CGG, AUA, AGG, AGA,	
	CUA, CCC, and GGA in the	
	same plasmid that habors the	
	T7 lysozyme	
$pET-15b$	$AMPR$; N-terminal 6xHis-	Novagen
	tag® of protein, includes a	
	thrombin recognition sequence	
pET-15b dhpH-C	AMP ^R (Ndel/XhoI)	(4)
pET-15b dhpH-C R233A	$\overline{\text{AMP}^R}$ (Ndel/XhoI)	This study
pET-15b dhpH-C W284L	AMP ^R (Ndel/XhoI)	This study
pET-15b dhpH-C R290A	$\overline{\text{AMP}^R}$ (Ndel/XhoI)	This study
pET-15b dhpH-C R292A	$\overline{\text{AMP}}^R(\text{Ndel/XhoI})$	This study
pET-15b dhpH-C R296A	$\overline{\text{AMP}^R}$ (Ndel/XhoI)	This study
pET-15b dhpH-C N325A	AMP ^R (Ndel/Xhol)	This study
pET-15b dhpH-C E326A	AMP ^R (Ndel/XhoI)	This study
pET-15b dhpH-C F328L	AMP ^R (Ndel/Xhol)	This study
pET-15b dhpH-C K329M	AMP ^R (Ndel/XhoI)	This study
pET-15b dhpH-C K329R	$\overline{\text{AMP}^R}$ (Ndel/XhoI)	This study
pET-14b leuRS	$AMP^{\overline{R}}$	Gift from the Martinis lab
		(UIUC) $(^{4, 14})$
pET-15b leuRS	AMP ^R	Gift from the Martinis lab
Y330A/D342A/D345A		(UIUC)
		$^{(15)}$
pET-15b ileRS	AMP ^R (Ndel/XhoI)	This study
pET-15b valRS	$\overline{AMP}^R(Ndel/Xhol)$	(16)
pET-15b pheRS	AMP ^R (Ndel/XhoI)	This study

Table S2. Strains and Plasmids

Supplementary Figures

Figure S1. Example 31P NMR spectra collected to follow the DhpH-C reaction over time.

The phosphonate spectral region of an example series of array data depicting product formation over roughly 80 min for the reaction with DhpH-C E326A. Spectra were collected roughly every 4 min.

 9.0

Figure S2. Product formation after 3 h for DhpH-C variants. Experiments were performed in duplicate.

Figure S3. No product is observed with DhpH-C K329M/R. ³¹P NMR spectra after a 12 h reaction with a) DhpH-C WT, b) DhpH-C K329R, or c) DhpH-C K329M. Reactions were performed in duplicate with $30 \mu M$ enzyme.

Figure S4. Reaction rate does not increase when doubling the aaRS concentration. Reactions were set up with 3 mg/mL total tRNA and 35 μ M DhpH-C WT (10 μ M for the LeuRS sample). Spectra were collected every 4 min for 2 h. Reactions contained 8 μ M synthetase (filled markers) or 16 μ M synthetase (open markers).

Figure S5. Determination of conditions to achieve equivalent aminoacylation levels for each synthetase. Aminoacylation experiments were performed with different amounts of total *E. coli* tRNA to achieve comparable amounts of aminoacylated tRNA. The final conditions identified are shown. Amount of total *E. coli* tRNA used: 3 mg/mL (LeuRS), 5.1 mg/mL (IleRS), 3.9 mg/mL (ValRS), 9.6 mg/mL (PheRS), and 5.4 mg/mL (LeuRS*m). For experimental details, see the section on the aminoacylation assay for total *E.coli* tRNAs. The bar graph represents the average of three replicates, with error bars indicating the standard deviation.

Figure S6. The aminoacylation reaction with various synthetases has gone to completion. Aminoacylation reaction aliquots were quenched after 15 min (striped) and 30 min (solid), or 1 h (striped) and 2 h (solid) for LeuRS*m to determine if the reaction had gone to completion.

Figure S7. Analysis of dipeptide product formation by LCMS. The mass of the predicted dipeptide products with IleRS, ValRS, PheRS, and LeuRS*m was confirmed for samples that had been incubated for 3 h. Each extracted ion chromatogram of the indicated dipeptide is compared to a sample with Leu/LeuRS and no DhpH-C.

Figure S8. Product formation after 3 h with different amino acid/aaRS pairs. Experiments were performed in duplicate with total *E. coli* tRNA, 35 μ M DhpH-C, and 8 μ M aaRS.

Figure S9. Reaction rate with various tRNALeu variants does not increase when the LeuRS concentration is doubled. Reactions were set up with 16 μ M tRNA^{Leu} variants (8 μ M for GAG (AU)) and 10 μ M DhpH-C WT. Reactions contained 8 μ M LeuRS (filled markers) or 16 μ M LeuRS (open markers).

Figure S10. Determination of conditions to achieve equivalent aminoacylation levels with each tRNA^{Leu} variant. Aminoacylation experiments were performed with 16 µM tRNA^{Leu}, except 8 μ M was used for GAG (AU). The bar graph represents the average of three replicates, with error bars indicating the standard deviation.

Figure S11. The aminoacylation reaction with tRNALeu variants has gone to completion. Aminoacylation experiments were performed with 16 μ M tRNA^{Leu}, except 8 μ M was used for GAG (AU). Reaction aliquots were quenched after 15 min (striped) and 30 min (solid) to determine if the reaction had gone to completion.

Supplementary References

- 1. Fonvielle, M., Li de La Sierra-Gallay, I., El-Sagheer, A. H., Lecerf, M., Patin, D., Mellal, D., Mayer, C., Blanot, D., Gale, N., Brown, T., van Tilbeurgh, H., Etheve-Quelquejeu, M., and Arthur, M. (2013) The structure of FemX(Wv) in complex with a peptidyl-RNA conjugate: mechanism of aminoacyl transfer from Ala-tRNA(Ala) to peptidoglycan precursors. *Angew. Chem. Int. Ed. Engl. 52*, 7278-7281.
- 2. Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics 9*, 40.
- 3. DeLano, W. L. (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos, CA.
- 4. Bougioukou, D. J., Mukherjee, S., and van der Donk, W. A. (2013) Revisiting the biosynthesis of dehydrophos reveals a tRNA-dependent pathway. *Proc. Natl. Acad. Sci. U. S. A. 110*, 10952-10957.
- 5. Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods. 6*, 343-345.
- 6. Keseler, I. M., Mackie, A., Peralta-Gil, M., Santos-Zavaleta, A., Gama-Castro, S., Bonavides-Martinez, C., Fulcher, C., Huerta, A. M., Kothari, A., Krummenacker, M., Latendresse, M., Muniz-Rascado, L., Ong, Q., Paley, S., Schroder, I., Shearer, A. G., Subhraveti, P., Travers, M., Weerasinghe, D., Weiss, V., Collado-Vides, J., Gunsalus, R. P., Paulsen, I., and Karp, P. D. (2013) EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res. 41*, D605-612.
- 7. Sherlin, L. D., Bullock, T. L., Nissan, T. A., Perona, J. J., Lariviere, F. J., Uhlenbeck, O. C., and Scaringe, S. A. (2001) Chemical and enzymatic synthesis of tRNAs for highthroughput crystallization. *RNA 7*, 1671-1678.
- 8. Rio, D. C., Ares, M. J., Hannon, G. J., and Nilsen, T. W. (2011) Synthesis, purification, labeling, and substitution of transcripts synthesized in vitro. Protocol 2: High-yield synthesis of RNA using T7 RNA polymerase and plasmid DNA or oligonucleotide templates, In *RNA: A Laboratory Manual*, pp 216-219, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- 9. Walker, S. E., and Fredrick, K. (2008) Preparation and evaluation of acylated tRNAs. *Methods 44*, 81-86.
- 10. Ortega, M. A., Hao, Y., Walker, M. C., Donadio, S., Sosio, M., Nair, S. K., and van der Donk, W. A. (2016) Structure and tRNA specificity of MibB, a lantibiotic dehydratase from actinobacteria involved in NAI-107 biosynthesis. *Cell Chem. Biol. 23*, 370-380.
- 11. Hoben, P., and Soll, D. (1985) Glutaminyl-tRNA synthetase of *Escherichia coli*. *Methods. Enzymol. 113*, 55-59.
- 12. Zhao, H., Palencia, A., Seiradake, E., Ghaemi, Z., Cusack, S., Luthey-Schulten, Z., and Martinis, S. (2015) Analysis of the resistance mechanism of a benzoxaborole inhibitor reveals insight into the leucyl-tRNA synthetase editing mechanism. *ACS Chem. Biol. 10*, 2277-2285.
- 13. Grant, S. G., Jessee, J., Bloom, F. R., and Hanahan, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U. S. A. 87*, 4645-4649.
- 14. Mursinna, R. S., Lincecum, T. L., Jr., and Martinis, S. A. (2001) A conserved threonine within *Escherichia coli* leucyl-tRNA synthetase prevents hydrolytic editing of leucyltRNALeu. *Biochemistry 40*, 5376-5381.
- 15. Martinis, S. A., Briggs, J. M., Mursinna, R. S., Lee, K. W., Lincecum, T. L., Williams, A. M., and Zhai, Y. Method and composition for leucyl-tRNA synthetases and derivatives thereof that activate and aminoacylate non-leucine amino acid to tRNA adaptor molecules. U. S. Patent 7,785,827 B2, August 31, 2010.
- 16. Huang, Z., Wang, K. A., and van der Donk, W. A. (2016) New insights into the biosynthesis of fosfazinomycin. *Chem. Sci. 7*, 5219-5223.