A Journal of the Gesellschaft Deutscher Chemiker

Angewandte International Edition GDCh Chemie

www.angewandte.org

Accepted Article

Title: Directed evolution of an improved aminoacyl-tRNA synthetase for incorporation of L-3,4-dihydroxyphenylalanine (L-DOPA)

Authors: Ross Thyer, Simon d'Oelsnitz, Molly S. Blevins, Dustin R. Klein, Jennifer S. Brodbelt, and Andrew D. Ellington

To be cited as: *Angew. Chem. Int. Ed.* 10.1002/anie.202100579

Link to VoR: https://doi.org/10.1002/anie.202100579

WILEY-VCH

10.1002/anie.202100579

WILEY-VCH

COMMUNICATION

Directed evolution of an improved aminoacyl-tRNA synthetase for incorporation of L-3,4-dihydroxyphenylalanine (L-DOPA)

Ross Thyer*[ab], Simon d'Oelsnitz[a], Molly S. Blevins[c], Dustin R. Klein[c], Jennifer S. Brodbelt[c] and Andrew D. Ellington*[a]

[a] Dr. R. Thyer, S. d'Oelsnitz, Prof. A. D. Ellington
Department of Molecular Biosciences, The University of Texas at Austin
Austin, Texas, USA
E-mail: ellingtonlab@gmail.com
[b] Dr. R. Thyer
Department of Chemical and Biomolecular Engineering, Rice University
Houston, Texas, USA
E-mail: ross.thyer@rice.edu
[c] M. Blevins, Dr. D. Klein, Prof. J. S. Brodbelt
Department of Chemistry, The University of Texas at Austin
Austin, Texas, USA

Supporting information for this article is given via a link at the end of the document.

Abstract: *The catechol group of 3,4-dihydroxyphenylalanine (L-DOPA) derived from L-tyrosine oxidation is a key post-translational modification (PTM) in many protein biomaterials and has potential as a bioorthogonal handle for precision protein conjugation applications such as antibody-drug conjugates. Despite this potential, indiscriminate enzymatic modification of exposed tyrosine residues or complete replacement of tyrosine using auxotrophic hosts remains the preferred method of introducing the catechol moiety into proteins, which precludes many protein engineering applications. We have developed new orthogonal translation machinery to site-specifically incorporate L-DOPA into recombinant proteins and a new fluorescent biosensor to selectively monitor L-DOPA incorporation in vivo. We show simultaneous biosynthesis and incorporation of L-DOPA and apply this translation machinery to engineer a novel metalloprotein containing a DOPA-Fe chromophore.*

While the standard genetic code is limited to twenty canonical amino acids, nature accesses many additional protein chemistries through post-translational modification and non-canonical incorporation pathways to fine-tune the activity of the proteome. In recent decades, researchers have leveraged the degeneracy of the genetic code and engineered translational machinery comprising of orthogonal aminoacyl-tRNA synthetase:tRNA pairs to repurpose redundant codons (most frequently the UAG stop codon) for site-specific introduction of non-canonical amino acids.

3,4-dihydroxyphenylalanine (L-DOPA) is a non-canonical amino acid accessed via post-translational modification of L-tyrosine. L-DOPA occurs widely in nature and is found in several biological polymers which are the subject of considerable research interest, including mussel foot proteins, tubeworm adhesive, and squid beak polymer[1]. Containing a reactive catechol group, L-DOPA can form redox-mediated crosslinks with proximal nucleophiles, and strong coordination complexes with metals, both of which are key contributors to the strength of these materials. Besides its utility in biomaterials, L-DOPA can serve as a handle for bioorthogonal conjugation[2], and functions as a catalytic radical in some enzymes[3].

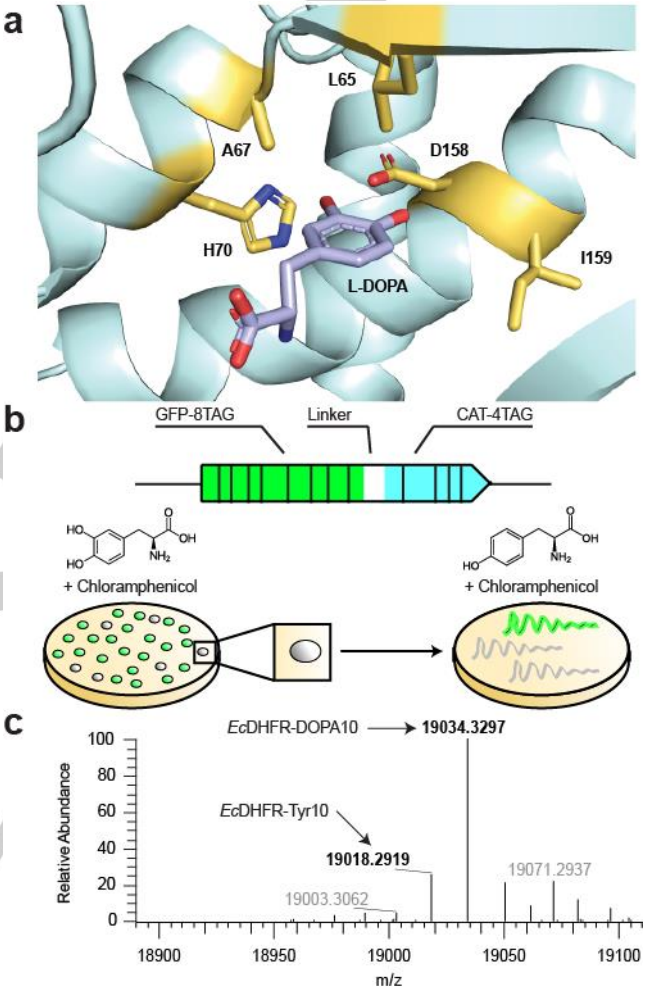


Figure 1: Directed evolution of an aminoacyl-tRNA synthetase for L-DOPA incorporation. a) The substrate binding pocket of *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase indicating the library residues, PDB: 1J1U. b) Schematic showing the GFP-CAT reporter used to isolate L-DOPA incorporating aaRS variants. GFP fluorescence is abolished when most of the tyrosine residues are replaced with L-DOPA. c) Deconvoluted mass spectrum of *Ec*DHFR V10X. Peaks corresponding to incorporation of tyrosine and L-DOPA are marked.

To date, researchers have developed several strategies for producing recombinant proteins containing L-DOPA, including the complete replacement of tyrosine residues using tyrosine-auxotrophic hosts[4], enzymatic modification of exposed tyrosine residues, and the use of orthogonal translation systems[5]. Incorporation of L-DOPA using orthogonal translation machinery provides the significant advantage of specifying precise locations for the catechol group and avoids problems with folding or stability which can occur with complete amino acid replacement. However, due to the small steric differences between L-DOPA and tyrosine,

10.1002/anie.202100579

previously reported aminoacyl-tRNA synthetases (aaRS) engineered to incorporate L-DOPA struggle to discriminate between the two amino acids[5]. In practice, this requires the use of minimal media to ensure a large excess of L-DOPA compared to tyrosine. An elegant attempt to solve this problem was the recently reported incorporation of a photoactivatable amino acid, ONB-DOPA, which converts to L-DOPA upon irradiation with UV light[6]. The bulky ONB-DOPA side chain provides a sterically distinct substrate and enabled evolution of a highly specific aaRS, however, UV irradiation does not uniformly convert ONB-DOPA to L-DOPA and its large size may clash with the local protein microenvironment and cause folding issues.

To engineer an improved aaRS specific for L-DOPA, we randomized five residues which line the amino acid binding pocket of the *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase (MjYRS) (**Figure 1a**) and performed iterative rounds of selection and counter-selection, in the presence or absence of L-DOPA respectively, using a previously described genetic reporter (CAT-pheS)[7]. This selection relies strictly on UAG suppression and is otherwise blind to the identity of the incorporated amino acid. Using this method, we were unable to isolate any MjYRS variants which had acquired selectivity for L-DOPA. Therefore, we hypothesized that a selection scheme based on the biochemical properties of L-DOPA, in addition to UAG suppression, may provide additional stringency and resolve rare aaRS mutants which can discriminate between the two amino acids. Ozawa *et al*. reported that replacement of all native tyrosine residues in GFP with L-DOPA using a cell-free protein expression system yielded a soluble but non-fluorescent protein[8]. Based on this information, we constructed a genetic fusion between sfGFP, in which eight native tyrosine codons were replaced with UAG and chloramphenicol acetyltransferase (CAT) containing four in-frame UAG codons at permissive positions (**Table S1**). This reporter was expected to yield non-fluorescent chloramphenicol resistant cells in response to L-DOPA incorporation, and fluorescent chloramphenicol resistant cells when tyrosine was incorporated (**Figure 1b**). Rescreening our initial library using this reporter yielded multiple clones showing strong sequence convergence (L65I/L/V, A67, H70G/P, D158, I159X), with the exception of a single distinct clone (L65E, A67, H70A, D158A, I159A). A random selection of clones was re-phenotyped by expressing a model protein (*Ec*DHFR) containing a single UAG codon and mass spectrometry was used to measure L-DOPA incorporation. This revealed that only the distinct clone (referred to as Gen 1) had acquired the ability to incorporate L-DOPA (**Figure 1c**, **Figure S1, Table S3**).

The strictly qualitative readout, unknown dynamic range and threshold effects, and mediocre ability to identify DOPA incorporating aaRS variants are all serious limitations of the GFP-CAT reporter, so we sought to develop a new fluorescent biosensor to monitor L-DOPA incorporation. In contrast to the loss of fluorescence reported by Ozawa *et al*., another research group reported that replacement of all tyrosine residues with L-DOPA in a different GFP variant resulted in a 19 nm red shift of the emission maximum and yielded a protein which appeared orange to the naked eye[9]. After recoding all native tyrosine codons to UAG, we attempted to express this GFP variant using the Gen 1 DOPA aaRS, but consistently observed very small, non-fluorescent cell pellets, possibly due to toxicity of the protein. Purification of the cell pellets yielded trace amounts of a protein with the correct molecular weight, which did appear orange in

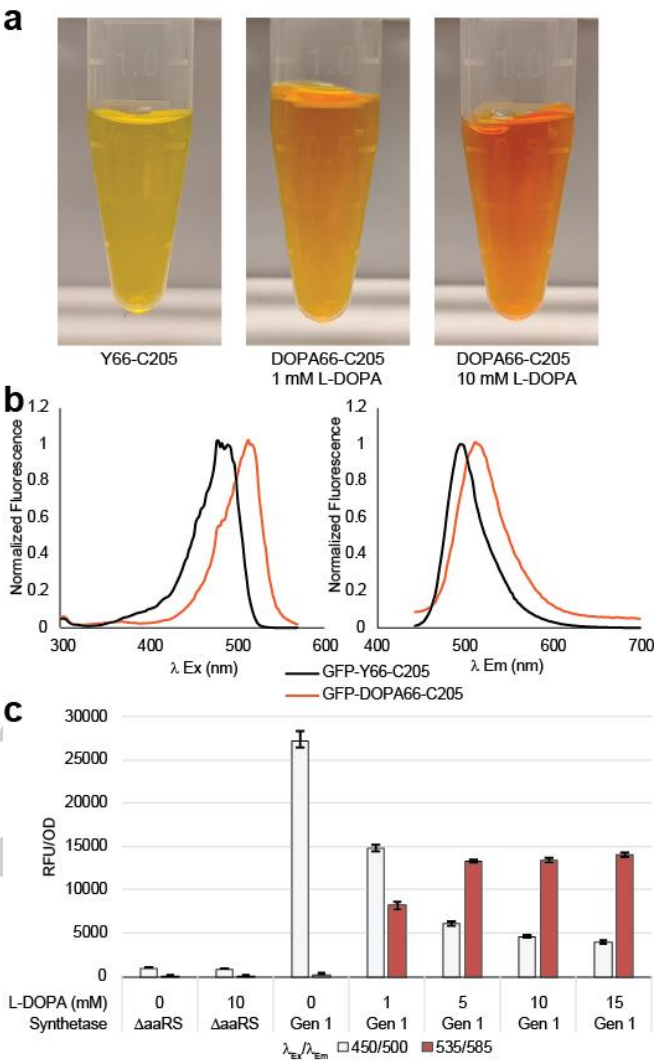


Figure 2: A fluorescent biosensor for L-DOPA incorporation. a) Comparison of GFP Y66 S205C (left) and GFP Y66DOPA S205C with increasing ratios of DOPA:Tyrosine in the chromophore. b) Excitation and emission spectra of purified GFP Y66 S205C and GFP Y66DOPA S205C (0.1 mg.mL^{-1} in PBS pH 7.4). The presence of L-DOPA in the chromophore induces a red-shift in both spectra. c) A fluorescence assay using Ex 535 nm/Em 585 nm yields a DOPA specific signal and enables L-DOPA incorporation to be monitored in living bacterial cells. Fluorescence intensity at 585 nm increases as the concentration of L-DOPA in the growth media increases, while fluorescence at 500 nm decreases.

color, but it could not be purified sufficiently to be resolved by mass spectrometry (**Figure S3d**). We hypothesized that not all tyrosine residues required replacement with L-DOPA; that only direct interactions with the DOPA chromophore produce a red shift in the emission spectrum. However, changing the chromophore tyrosine to L-DOPA and individually mutating the two nearest tyrosine residues (Y92 and Y145) to DOPA was not sufficient to alter the spectral properties (**Figure S4**).

To improve the reporter, we explored alternative interactions with the L-DOPA chromophore which might result in novel spectral properties. It has been reported that engineered covalent crosslinks between the chromophore and surrounding residues of fluorescent proteins can alter the spectral properties and improve brightness[10]. L-DOPA can undergo sequential oxidation to yield DOPA-semiquinone and DOPA-quinone, which

10.1002/anie.202100579

are both reactive electrophiles that can form covalent crosslinks with nearby nucleophiles. Based on this information, we mutated several residues with side chains proximal to the chromophore to cysteine, the strongest nucleophile among the standard amino acids. Upon centrifugation of bacterial cultures, we observed that one variant, containing a S205C mutation, resulted in a visibly orange protein (**Figure 2a**). Spectral characterization of the purified protein confirmed a red-shift of both the excitation (490 nm to 518 nm) and emission (510 nm to 528 nm) maxima (**Figure 2b**). Mass spectrometry confirmed the presence of L-DOPA at residue 66, although the consistent backbone fragmentation generated by ultraviolet photodissociation (UVPD) and intact mass did not support the formation of a cysteinyl-DOPA crosslink between DOPA 66 and Cys 205 (**Figure S2-S3a**). Interestingly, these mutations also red-shifted fluorescence in other common GFP variants, including sfGFP and mClover3 (**Figure S3b-c**). Based on the excitation and emission spectra, we tested pairs of red-shifted wavelengths and determined that excitation at 535 nm and emission at 585 nm generated a robust fluorescent signal which is unique to the DOPA-containing species (**Figure 2c**). Green fluorescence (Ex 450 nm/Em 500 nm) was useful for monitoring tyrosine incorporation since L-DOPA-containing molecules only weakly fluoresce in this range.

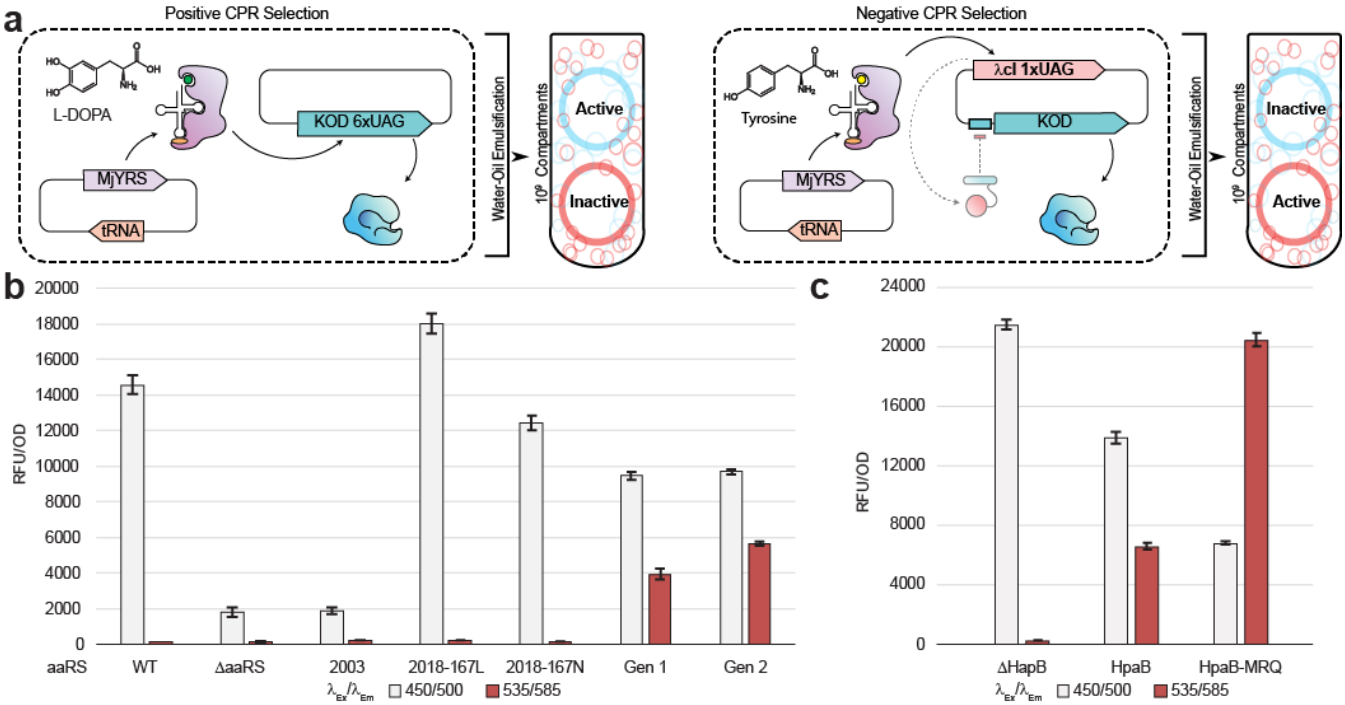


Figure 3: Emulsion PCR selection of aaRS variants with improved specificity for L-DOPA. a) Schematic of genetic circuits for CPR selection and counter-selection. Genes encoding active aaRS variants are enriched by emulsion PCR during positive selection in the presence of L-DOPA. During counter selection, aaRS variants which incorporate tyrosine express a repressor which prevents expression of the DNA polymerase, ensuring only inactive variants are enriched. b) Fluorescence assay measuring the activity of different aaRS variants reported to incorporate L-DOPA, 2003[5], and 2018-167L and 2018-167N[11]. The assay was performed in complex growth medium with 1 mM L-DOPA. c) Co-expression of the Gen 2 DOPA aaRS with the *E. coli hpaBC* genes enables simultaneous biosynthesis and incorporation of L-DOPA into proteins. An engineered HpaB variant with greater activity increases the ratio of DOPA:tyrosine incorporation.

With the ability to easily quantify aaRS performance in high-throughput using our fluorescent biosensor, we systematically mutated additional residues flanking the amino acid binding pocket of MjYRS which have previously been sampled in the literature[12]. Surprisingly, comprehensive screening of individual site-saturation libraries at residues Y32, Y151, Q155, E107, F108, Q109, L162, and A167 for enhanced ratios of DOPA:Tyrosine incorporation did not yield any further improvements. Similarly, rescreening a small library which encompassed the entire pool of converged sequences from our initial selection (L65I/L/V, A67, H70G/P, D158, I159X) did not identify any new clones.

While the amino acid binding pocket has been the focus of most orthogonal aaRS engineering, the anticodon recognition domain also plays an important role and has emerged as a new engineering target[13]. To better interrogate this region of the protein, we adopted a new genetic selection methodology, Compartmentalized Partnered Replication (CPR), which has previously been used to isolate aaRSs with improved specificity for non-canonical amino acids[14]. CPR relies on linking a desired phenotype to the production of a thermostable DNA polymerase and uses this polymerase to enrich the corresponding coding sequences that produced the desired phenotype during emulsion PCR (**Figure 3a**)[14]. To select for improved L-DOPA incorporation, we constructed a CPR selection circuit by introducing six UAG codons in place of native, surface exposed tyrosine residues in KOD DNA polymerase from *Thermococcus kodakaraensis*. To select against tyrosine incorporation, a counter selection circuit was designed with a wild-type KOD gene under the control of the λ phage pR promoter and a constitutively expressed λcI repressor containing a single UAG stop codon.

Five residues in the anticodon recognition domain (F261, H283, P284, M285, and R286) were randomized by site-saturation mutagenesis and the library subjected to three rounds of positive selection interspersed with two rounds of counter selection. In addition, our initial amino acid binding pocket library (L65, A67, H70, D158 and I159) was reconstructed using the Gen 1 plasmid as a template and reselected using the CPR selection. CPR selections yielded two variants (one from each library) with

minor improvements in activity or selectivity (**Figure S5a**). Clone 1.1 was isolated from the anticodon recognition domain library and contained F261S, H283P, P284L, M285R, and R286A at the library residues. In addition, this clone contained two off-target mutations, E220K and L282P, the latter directly adjacent to the library residues and most likely introduced by an imperfect oligonucleotide primer. DNA sequencing revealed that Clone 2.1 from the amino acid binding pocket library retained the parental amino acid sequence but had acquired two off-target mutations (Y119F an N268D). Individually introducing the off-target mutations into the Gen1 aaRS revealed that Y119F enhanced specificity for L-DOPA (**Figure S5b**). Site-saturation of Y119 did not yield any additional mutations at this position. N268D and E220K were not observed to have any effect on aaRS activity or specificity and were not expected to form any new interactions (**Figure S6**). The effects of the library mutations from clone 1.1 and Y119F were observed to be additive and the resulting variant (named Gen 2, **Table S1-S2**) gained moderate improvements to both activity and selectivity towards L-DOPA (**Figure 3b, S5b**).

The Gen 1 and Gen 2 aaRS variants were compared to previously reported synthetases for L-DOPA incorporation (**Figure 3b**)[5, 11]. These variants appeared to have low activity and selectivity for L-DOPA in complex growth medium and did not generate a DOPA-specific fluorescent signal under our assay conditions. This is consistent with previous reporting which shows these variants require the use of minimal media to prevent promiscuous incorporation of tyrosine. The efficiency of DOPA incorporation by the Gen 1 and Gen 2 aaRS variants was determined by mass spectrometry using GFP Y66DOPA S205C expressed using two different concentrations of L-DOPA in the growth medium. Mass spectrometry confirmed the Gen 2 aaRS to be more selective for L-DOPA, achieving 67% DOPA incorporation when 1 mM of L-DOPA was provided, compared to 55% for the Gen 1 aaRS (**Figure S7a, c**). Increasing the amount of L-DOPA present during protein expression increased incorporation efficiency to ~83% for both variants (**Figure S7b, d**).

Combining the ability to incorporate a non-canonical amino acid into proteins with its corresponding biosynthetic machinery is an attractive option for producing recombinant proteins at scale. We hypothesized that our novel orthogonal translation machinery with improved selectivity would be sufficient to enable the simultaneous biosynthesis and incorporation of L-DOPA. *E. coli* 4-hydroxyphenylacetate 3-monooxygenase (HpaB) is a promiscuous oxygenase which acts on a broad range of phenolic groups and can hydroxylate tyrosine to form L-DOPA. Co-expression of GFP Y66DOPA S205C, HpaB, and its partner FAD reductase which recycles the necessary flavin cofactor, encoded by the *hpaBC* genes, resulted in a clear fluorescent signal at 585 nm in the absence of exogenous L-DOPA (**Figure 3c**). Consistent with our results in **Figure 2c**, increasing the concentration of L-DOPA by substituting the wild-type HpaB with an engineered variant with greater activity (HpaB-MRQ) significantly increased fluorescence at 585 nm while decreasing the fluorescence intensity at 500 nm[15]. Bypassing costly media supplementation and potentially inefficient cellular import, *in vivo* biosynthesis and incorporation of L-DOPA poses numerous advantages for producing L-DOPA-containing proteins at scale.

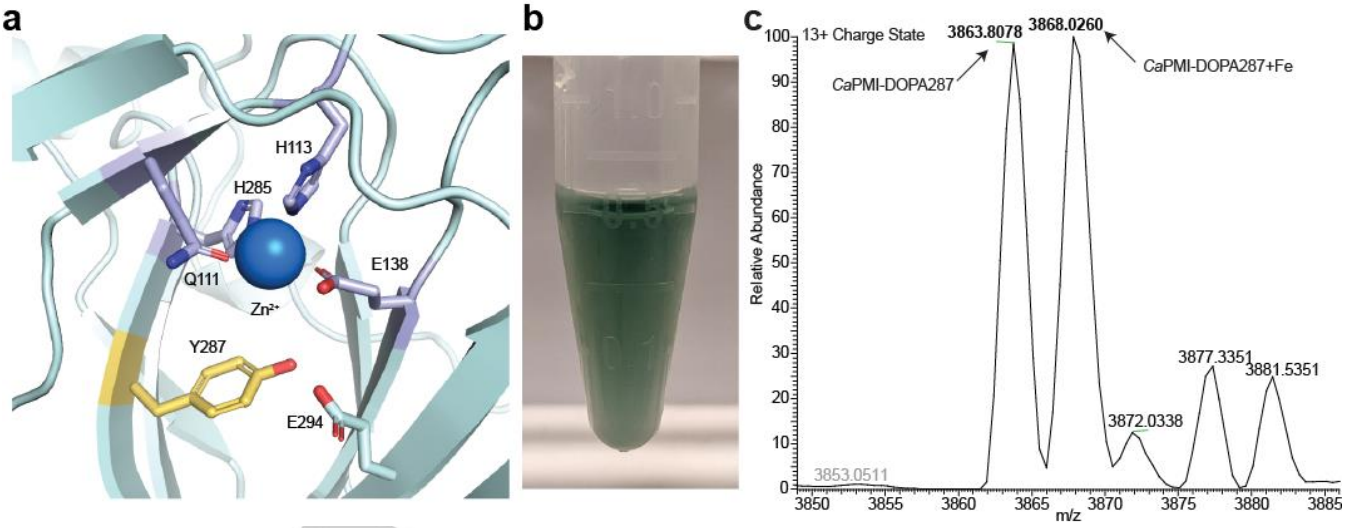


Figure 4: A single L-DOPA residue converts the Zn^{2+} binding pocket of *Ca*PMI into a DOPA-Fe chromophore. a) The Zn^{2+} binding pocket of *Ca*PMI showing the native metal coordinating residues and Tyr 287 which was replaced with L-DOPA. b) Purified *Ca*PMI-DOPA287 appears blue-green due to the formation of a DOPA-Fe^{3+} chromophore. c) UHMR native mass spectrum of intact *Ca*PMI confirms the presence of a single Fe atom. Observed masses are labelled.

The catechol moiety of L-DOPA can form strong coordination complexes with metals, particularly ferric iron, which are a key feature of many interesting biological polymers. To harness this powerful interaction for protein engineering applications using our evolved aaRS, we demonstrate controlled formation of a DOPA-Fe^{3+} complex within the metalloenzyme phosphomannose isomerase (CaPMI) from the yeast *Candida albicans*. This enzyme natively binds a Zn^{2+} ion in the catalytic site, however, under Zn deficient conditions the native metal can be replaced by Fe^{2+}. If the ferrous iron is oxidized to Fe^{3+} the enzyme can undergo autocatalytic hydroxylation of Tyr 287 which flanks the catalytic site to form L-DOPA (**Figure 4a**). This DOPA residue then chelates the ferric iron to form a blue-green chromophore[16]. To explore whether autocatalytic formation of DOPA can be replaced by translational incorporation and direct chelation of ferric iron (without requiring the absence of zinc), we mutated the codon for Tyr 287 to UAG. CaPMI-DOPA287 expressed in the presence of 10 mM exogenous L-DOPA and 10 µM iron(III) citrate yielded a dark blue-green protein (**Figure 4b, S9c**) with a UV-Vis absorption spectrum consistent with previous reports for this protein[16]. A large fraction of the protein was confirmed by mass spectrometry to contain one iron atom (**Figure**

10.1002/anie.202100579

4c, S9a). No masses were detected which corresponded to a protein containing Tyr 287 (**Figure S8**); we hypothesize this may be due to a combination of autocatalytic conversion of Tyr 287 to L-DOPA and poor stability of CaPMI-Y287 in the absence of coordinated Zn^{2+}. CaPMI-DOPA serves as an interesting example of a metalloprotein containing an engineered DOPA-Fe motif introduced using an orthogonal translation system.

Here we report a new orthogonal aaRS which enables site-specific incorporation of the catechol amino acid L-DOPA into proteins. In contrast to previously reported synthetases for DOPA incorporation, the Gen 2 aaRS retains both high activity and selectivity for L-DOPA and enables the introduction of this versatile chemistry into recombinant proteins without the requirement for tyrosine deficient media and/or tyrosine auxotrophic hosts. Despite extensive mutagenesis of the MjYRS scaffold over the past two decades, we have identified a novel residue (Y119) outside of the primary amino acid binding pocket which contributes to substrate specificity. This residue lies at the dimer interface and has no direct contacts with amino acid substrates but may act indirectly via repositioning Met 154 (**Figure S6**), another residue which has not been sampled during previous mutagenesis studies. The full mutational landscape of this well studied protein remains to be explored and many cryptic loci which influence activity and substrate selectivity could yet be discovered.

With our significantly improved L-DOPA incorporation machinery, new protein engineering applications are now within reach. These include our demonstration a DOPA-Fe motif within CaPMI, which may also be used to strengthen protein-based biomaterials, site-specific incorporation of L-DOPA to provide a novel handle for small-molecule conjugation, or enzymes engineered to mimic nature's use of L-DOPA as an active site catalytic radical. In addition, the DOPA-specific biosensor we describe allows for high-throughput monitoring of L-DOPA incorporation *in vivo* and will be an invaluable tool for further engineering efforts. Finally, simultaneous *in vivo* biosynthesis and incorporation of L-DOPA circumvents the need for media supplementation and establishes the feasibility of producing recombinant proteins containing L-DOPA at scale.

Acknowledgements

Funding from the NIH NCI (5K99CA207870-02 to R.T.), NIH NIGMS (R01 GM121714 to J.S.B.), the Welch Foundation (F1155 to J.S.B. and F-1654 to A.D.E), the Air Force Office of Scientific Research (FA9550-14-1-0089 to A.D.E), and the US Army Research Office (W911NF-18-1-0169, and SP0036191-PROJ0009952 via MURI to A.D.E) is acknowledged.

Conflict of Interest: The authors declare no conflicts of interest.

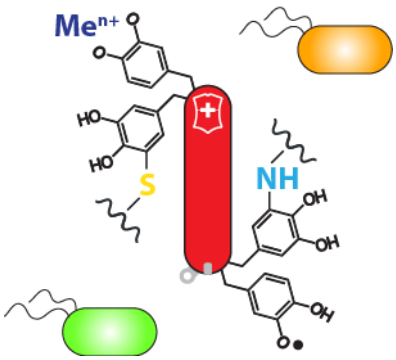
Keywords: Protein engineering • Protein Design • Synthetic Biology • Enzymes • Metalloproteins

10.1002/anie.202100579

WILEY-VCH

COMMUNICATION

Entry for the Table of Contents



Catechols are a multifunctional chemistry capable of diverse interactions with metals, radical chemistry, the ability to form redox-mediated crosslinks, and play an important role in many biological polymers. Here we report new genetic tools which enable site-specific incorporation of the catechol amino acid 3,4-dihydroxyphenylalanine (L-DOPA) into proteins in bacteria, and a new DOPA-specific fluorescent biosensor for monitoring incorporation in living cells.

Institute and/or researcher Twitter usernames: @BiomolEngineer

References

[1] J. Yang, M. A. Cohen Stuart, M. Kamperman, *Chem Soc Rev* **2014**, *43*, 8271-8298.

[2] E. Montanari, A. Gennari, M. Pelliccia, L. Manzi, R. Donno, N. J. Oldham, A. MacDonald, N. Tirelli, *Bioconjugate chemistry* **2018**, *29*, 2550-2560.

[3] J. Stubbe, M. R. Seyedsayamdost, *Biochemistry* **2019**, *58*, 435-437.

[4] B. Yang, N. Ayyadurai, H. Yun, Y. S. Choi, B. H. Hwang, J. Huang, Q. Lu, H. Zeng, H. J. Cha, *Angew Chem Int Ed Engl* **2014**, *53*, 13360-13364.

[5] L. Alfonta, Z. Zhang, S. Uryu, J. A. Loo, P. G. Schultz, *Journal of the American Chemical Society* **2003**, *125*, 14662-14663.

[6] M. Hauf, F. Richter, T. Schneider, T. Faidt, B. M. Martins, T. Baumann, P. Durkin, H. Dobbek, K. Jacobs, A. Möglich, N. Budisa, *Chembiochem* **2017**, *18*, 1819-1823.

[7] R. Thyer, A. Filipovska, O. Rackham, *Journal of the American Chemical Society* **2013**, *135*, 2-5.

[8] K. Ozawa, M. J. Headlam, D. Mouradov, S. J. Watt, J. L. Beck, K. J. Rodgers, R. T. Dean, T. Huber, G. Otting, N. E. Dixon, *The FEBS journal* **2005**, *272*, 3162-3171.

[9] N. Ayyadurai, N. S. Prabhu, K. Deepankumar, Y. J. Jang, N. Chitrapriya, E. Song, N. Lee, S. K. Kim, B. G. Kim, N. Soundrarajan, S. Lee, H. J. Cha, N. Budisa, H. Yun, *Bioconjugate chemistry* **2011**, *22*, 551-555.

[10] Z. Xiang, H. Ren, Y. S. Hu, I. Coin, J. Wei, H. Cang, L. Wang, *Nat Methods* **2013**, *10*, 885-888.

[11] S. Kim, B. H. Sung, S. C. Kim, H. S. Lee, *Chemical communications (Cambridge, England)* **2018**, *54*, 3002-3005.

[12] M. Amiram, A. D. Haimovich, C. Fan, Y. S. Wang, H. R. Aerni, I. Ntai, D. W. Moonan, N. J. Ma, A. J. Rovner, S. H. Hong, N. L. Kelleher, A. L. Goodman, M. C. Jewett, D. Söll, J. Rinehart, F. J. Isaacs, *Nature biotechnology* **2015**, *33*, 1272-1279.

[13] R. Gan, J. G. Perez, E. D. Carlson, I. Ntai, F. J. Isaacs, N. L. Kelleher, M. C. Jewett, *Biotechnol Bioeng* **2017**, *114*, 1074-1086.

[14] J. W. Ellefson, A. J. Meyer, R. A. Hughes, J. R. Cannon, J. S. Brodbelt, A. D. Ellington, *Nature biotechnology* **2014**, *32*, 97-101.

[15] C. Wang, J. E. Powell, V. Pandarinath, S. d'Oelsnitz, N. A. Moran, A. D. Ellington, R. Thyer, *Unpublished* **2021**.

[16] J. J. Smith, A. J. Thomson, A. E. Proudfoot, T. N. Wells, *Eur J Biochem* **1997**, *244*, 325-333.