Genetic Code Expansion in *Rhodobacter sphaeroides* to Incorporate Noncanonical Amino Acids into Photosynthetic Reaction Centers

Jared Bryce Weaver* and Steven G. Boxer*

Department of Chemistry, Stanford University, Stanford, California 94305, United States

Supporting Information

**ABSTRACT:** Photosynthetic reaction centers (RCs) are the membrane proteins responsible for the initial charge separation steps central to photosynthesis. As a complex and spectroscopically complicated membrane protein, the RC (and other associated photosynthetic proteins) would benefit greatly from the insight offered by site-specifically encoded noncanonical amino acids in the form of probes and an increased chemical range in key amino acid analogues. Toward that goal, we developed a method to transfer amber codon suppression machinery developed for *E. coli* into the model bacterium needed to produce RCs, *Rhodobacter sphaeroides*. Plasmids were developed and optimized to incorporate 3-chlorotyrosine, 3-bromotyrosine, and 3-iodotyrosine into RCs. Multiple challenges involving yield and orthogonality were overcome to implement amber suppression in *R. sphaeroides*, providing insights into the hurdles that can be involved in host transfer of amber suppression systems from *E. coli*. In the process of verifying noncanonical amino acid incorporation, characterization of this membrane protein via mass spectrometry (which has been difficult previously) was substantially improved. Importantly, the ability to incorporate noncanonical amino acids in *R. sphaeroides* expands research capabilities in the photosynthetic field.

**KEYWORDS:** amber suppression, noncanonical amino acids, *Rhodobacter sphaeroides*, photosynthetic reaction centers, orthogonality

Photosynthetic reaction centers (RCs) are the membrane proteins responsible for the initial charge separation steps central to photosynthesis. They capture solar energy in a highly efficient combination of energy and electron transfer steps. RCs possess two branches of chromophores arranged in nearly identical paths, L and M (or A and B). A pair of excitonically coupled bacteriochlorophylls, known as the special pair (P), absorbs a photon directly or by energy transfer, and the excited state transfers an electron to neighboring chromophores as illustrated in Figure 1. Following photoexcitation, however, electrons are only transferred down the L branch of chromophores. Despite intense effort, the origin(s) of this unidirectional electron transfer is not fully understood, and even the role of intermediary electron acceptors is unclear.

In this manuscript, we introduce site-specific noncanonical amino acid (ncAA) incorporation as a method to analyze protein design features in this complicated membrane protein complex from *R. sphaeroides*, consisting of three protein subunits (denoted L, M and H) and 9 prosthetic groups (6 chromophores, 2 quinones, a nonheme Fe, and a carotenoid). Conventional site-directed mutagenesis has been extensively applied to perturb photosynthetic charge separation, but to date, tools for introducing ncAAs into the RC have not been described. Their inclusion would allow greater chemical control than the often-large chemical changes which ensue in canonical amino acid mutations. For instance, the tyrosine at M210 (Figure 1) has long been indicated as a key factor for L-branch electron transfer because it is a clear deviation in symmetry between the L and M branch protein environments, and it is the only 1 out of the 28 tyrosines in the RC whose hydroxyl

**Figure 1.** Chromophore arrangement within the RC as determined by the wild-type *R. sphaeroides* crystal structure (Protein Data Bank: 2J8C; quinones and accessory carotenoid not shown for simplicity). Chromophore phytol side chains are removed for clarity. Blue arrow indicates direction of electron flow along L branch. Tyr at M210, the target of the current work, is labeled in blue and its symmetry-related counterpart, Phe at L181, is shown labeled in gold.
group is not hydrogen-bonded. Past simulation results have indicated that the orientation and dipole of the tyrosine hydroxyl play an important role in stabilizing initial charge transfer intermediates. Recent results from our lab demonstrated the impact of introducing a putative hydrogen bond partner near Y(M210), a change which could reorient the tyrosine −OH dipole away from maximally stabilizing interactions for charge transfer indicated by these simulation results. This change dramatically slowed electron transfer from 3–4 to 50 ps. However, by using ncAAs such as halogenated tyrosines, for example, the effect of hydroxyl dipole strength could be tuned more precisely. While the tyrosine at M210 is an important initial target, ncAA incorporation in the form of tyrosines, for example, the e

Figure 2. Constructs and ncAAs used in this work. (A) Promoter, RBS, and/or terminator layout of modified tRNA and aaRS gene cassettes (see Figure S1 for an example of a plasmid map containing the aaRS and tRNA gene cassettes). Promoters, RBS, and terminators are either derived from or known to be functional in R. sphaeroides unless otherwise stated. The specific aaRS or tRNA contained in the general gene cassettes depicted above varied depending on plasmid construct (see Table 2; i.e., [8] contained pAFRS and tRNA as the aaRS and tRNA in cassettes P

While amber codon suppression has been developed for well over a decade, the range of bacterial hosts for which this technique is available is very limited. This is due to multiple factors—as in any heterologous gene expression, the aminoacyl tRNA synthetase (aaRS) and tRNA must have transcriptional and translational regulatory control compatible with their host organism. This involves choosing appropriate promoters, ribosome-binding sites (RBSs), and terminators for different aaRS/tRNA pairs, along with other genetic features. The relative strength or ability of these regulatory elements to up/downregulate aaRS and tRNA expression in a new host is also of importance both to avoid negative effects on growth and to obtain sufficient yields of the ncAA-incorporated proteins. While these issues can generally be managed by screening different aaRS and tRNA regulatory control features and plasmid vectors, perhaps the greatest challenge is the potential lack of orthogonality of the new aaRS or tRNA in the given host organism. While other types of orthogonality are involved in a successful amber suppressor aaRS/tRNA pair, we will focus on the orthogonality, which is perhaps more changeable and detrimental if lost upon amber suppression host transfer (see SI Section A for discussion). Orthogonality in this case refers to the property by which the introduced aaRS binds only to its cognate amber suppressor tRNA in the presence of the other endogenous host tRNAs. Similarly, an orthogonal tRNA must be recognized only by the aaRS evolved for ncAA incorporation and none of the other endogenous aaRSs. While any given aaRS/tRNA amber suppressor pair is evolved in E. coli under positive and negative pressure to be orthogonal to endogenous E. coli aaRSs and tRNAs, orthogonality deviations in evolved pairs can still remain following selection or can occur upon transfer to organisms other than E. coli. They can result in either misincorporation at the amber-encoded-ncAA site of interest or greater stop codon read through at amber stop codon-terminating genomic genes, as nonorthogonal amber suppressor tRNAs are aminoacylated with not just ncAAs but also canonical amino acids. Both issues are detrimental to yield and obtaining desired protein constructs and can result in decreased growth rate and toxicity in less optimal cases. Notwithstanding the challenges in transferring amber suppressor aaRS/tRNA pairs from the organism in which they were originally developed, referred to here as amber suppression host transfer, many proteins cannot be expressed in model hosts like E. coli. The RC is an example due to the number of proteins and unique prosthetic groups required for
its assembly and function. The following describes how we transferred a functional, orthogonal aaRS/tRNA pair to R. sphaeroides, the challenges we encountered, and how we addressed them, further informing on obstacles generally encountered in aaRS/tRNA host transfer.

**RESULTS AND DISCUSSION**

Plasmid Construction. To encode site-specific ncAA incorporation in R. sphaeroides we initially used the *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase and its associated amber suppressor tRNA (MjTyrrs/tRNA<sup>Tyr</sup>).<sup>21</sup> Incorporation of ncAAs with the MjTyrrs/tRNA<sup>Tyr</sup> pair has been implemented in multiple species, and it is widely used in *E. coli*, which we thought might be closely enough related to the *R. sphaeroides* tyrosyl-tRNA synthetase and its associated amber suppressor tRNA<sup>Tyr</sup> for pAF<sup>22</sup> We chose to use the pAF<sup>21</sup> expression cassette because it did not require the use of *E. coli*-specific promoters, RBS, and terminators<sup>22</sup> and together to determine if metabolic stress from two strongly promoted genes due to new regulatory control was detrimental to cell viability and growth.

The pIND4-RC plasmid<sup>1</sup>, a pMG160 derivative, was chosen as a vector for pAF aaRS (pAFRS)/tRNA constructs (Table 1, Group 1 and Figure S1) due to its higher copy number in both *E. coli* and *R. sphaeroides*<sup>18–23</sup> in pIND4-RC vectors as opposed to 4–7 in commonly used pRK vectors in *R. sphaeroides* cells) making it more amenable to plasmid construction.<sup>27–29</sup> The pIND4-RC plasmid also contained the H, M, and L genes for RC production under control of the IPTG-inducible lac operon which facilitated coordination of RC induction with ncAA addition for amber suppression.<sup>27–29</sup> The pIND4-RC plasmid and the *R. sphaeroides* RCx strain (an electroportable cell strain with genes RC and light harvesting complexes chromosomally removed) described below were kindly provided by Drs. Thomas Beatty and Daniel Jun.

Initially the *E. coli* promoters, RBS, and terminators associated with the pAFRS/tRNA pair were left unchanged in its insertion into the *R. sphaeroides* vector pIND4-RC. This generated the plasmid pIND4-RC-pAF0<sup>7</sup>, a plasmid which also contained the genes for RC production (Table 1, Group 2). The pIND4-RC vector already contained the genes for the RC H, L, and M subunits, so to test ncAA incorporation, the *R. sphaeroides* RCx strain (an electroportable cell strain with genes RC and light harvesting complexes chromosomally removed) described below was kindly provided by Drs. Thomas Beatty and Daniel Jun.

### RESULTS AND DISCUSSION

Plasmid Construction. To encode site-specific ncAA incorporation in *R. sphaeroides* we initially used the *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase and its associated amber suppressor tRNA<sub>Tyr</sub> (MjTyrs/tRNA<sub>Tyr</sub>).<sup>21</sup> Incorporation of ncAAs with the MjTyrs/tRNA<sub>Tyr</sub> pair has been implemented in multiple species, and it is widely used in *E. coli*, which we thought might be closely enough related to the *Rhodobacter* genus (both Gram-negative bacteria). Importantly it has been evolved to incorporate multiple ncAAs, among them several tyrosine and phenylalanine analogues (3-chlorotyrosine, p-cyanophenylalanine, p-azidophenylalanine, etc.).<sup>22–24</sup> The first MjTyrs/tRNA<sub>Tyr</sub> pair we implemented was one in which the aaRS had been evolved for p-aminophenylalanine (pAF, Figure 2B) incorporation encoded on the plasmid pDule-pAF<sup>2</sup> (Table 1, Group 1).<sup>25,26</sup>

### RESULTS AND DISCUSSION

Plasmid Construction. To encode site-specific ncAA incorporation in R. sphaeroides we initially used the *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase and its associated amber suppressor tRNA<sub>Tyr</sub> (MjTyrs/tRNA<sub>Tyr</sub>).<sup>21</sup> Incorporation of ncAAs with the MjTyrs/tRNA<sub>Tyr</sub> pair has been implemented in multiple species, and it is widely used in E. coli, which we thought might be closely enough related to the *Rhodobacter* genus (both Gram-negative bacteria). Importantly it has been evolved to incorporate multiple ncAAs, among them several tyrosine and phenylalanine analogues (3-chlorotyrosine, p-cyanophenylalanine, p-azidophenylalanine, etc.).<sup>22–24</sup> The first MjTyrs/tRNA<sub>Tyr</sub> pair we implemented was one in which the aaRS had been evolved for p-aminophenylalanine (pAF, Figure 2B) incorporation encoded on the plasmid pDule-pAF<sup>2</sup> (Table 1, Group 1).<sup>25,26</sup>

### RESULTS AND DISCUSSION

Plasmid Construction. To encode site-specific ncAA incorporation in R. sphaeroides we initially used the *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase and its associated amber suppressor tRNA<sub>Tyr</sub> (MjTyrs/tRNA<sub>Tyr</sub>).<sup>21</sup> Incorporation of ncAAs with the MjTyrs/tRNA<sub>Tyr</sub> pair has been implemented in multiple species, and it is widely used in *E. coli*, which we thought might be closely enough related to the *Rhodobacter* genus (both Gram-negative bacteria). Importantly it has been evolved to incorporate multiple ncAAs, among them several tyrosine and phenylalanine analogues (3-chlorotyrosine, p-cyanophenylalanine, p-azidophenylalanine, etc.).<sup>22–24</sup> The first MjTyrs/tRNA<sub>Tyr</sub> pair we implemented was one in which the aaRS had been evolved for p-aminophenylalanine (pAF, Figure 2B) incorporation encoded on the plasmid pDule-pAF<sup>2</sup> (Table 1, Group 1).<sup>25,26</sup>

### RESULTS AND DISCUSSION

Plasmid Construction. To encode site-specific ncAA incorporation in R. sphaeroides we initially used the *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase and its associated amber suppressor tRNA<sub>Tyr</sub> (MjTyrs/tRNA<sub>Tyr</sub>).<sup>21</sup> Incorporation of ncAAs with the MjTyrs/tRNA<sub>Tyr</sub> pair has been implemented in multiple species, and it is widely used in *E. coli*, which we thought might be closely enough related to the *Rhodobacter* genus (both Gram-negative bacteria). Importantly it has been evolved to incorporate multiple ncAAs, among them several tyrosine and phenylalanine analogues (3-chlorotyrosine, p-cyanophenylalanine, p-azidophenylalanine, etc.).<sup>22–24</sup> The first MjTyrs/tRNA<sub>Tyr</sub> pair we implemented was one in which the aaRS had been evolved for p-aminophenylalanine (pAF, Figure 2B) incorporation encoded on the plasmid pDule-pAF<sup>2</sup> (Table 1, Group 1).<sup>25,26</sup>

### RESULTS AND DISCUSSION

Plasmid Construction. To encode site-specific ncAA incorporation in R. sphaeroides we initially used the *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase and its associated amber suppressor tRNA<sub>Tyr</sub> (MjTyrs/tRNA<sub>Tyr</sub>).<sup>21</sup> Incorporation of ncAAs with the MjTyrs/tRNA<sub>Tyr</sub> pair has been implemented in multiple species, and it is widely used in *E. coli*, which we thought might be closely enough related to the *Rhodobacter* genus (both Gram-negative bacteria). Importantly it has been evolved to incorporate multiple ncAAs, among them several tyrosine and phenylalanine analogues (3-chlorotyrosine, p-cyanophenylalanine, p-azidophenylalanine, etc.).<sup>22–24</sup> The first MjTyrs/tRNA<sub>Tyr</sub> pair we implemented was one in which the aaRS had been evolved for p-aminophenylalanine (pAF, Figure 2B) incorporation encoded on the plasmid pDule-pAF<sup>2</sup> (Table 1, Group 1).<sup>25,26</sup>

### RESULTS AND DISCUSSION

Plasmid Construction. To encode site-specific ncAA incorporation in R. sphaeroides we initially used the *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase and its associated amber suppressor tRNA<sub>Tyr</sub> (MjTyrs/tRNA<sub>Tyr</sub>).<sup>21</sup> Incorporation of ncAAs with the MjTyrs/tRNA<sub>Tyr</sub> pair has been implemented in multiple species, and it is widely used in *E. coli*, which we thought might be closely enough related to the *Rhodobacter* genus (both Gram-negative bacteria). Importantly it has been evolved to incorporate multiple ncAAs, among them several tyrosine and phenylalanine analogues (3-chlorotyrosine, p-cyanophenylalanine, p-azidophenylalanine, etc.).<sup>22–24</sup> The first MjTyrs/tRNA<sub>Tyr</sub> pair we implemented was one in which the aaRS had been evolved for p-aminophenylalanine (pAF, Figure 2B) incorporation encoded on the plasmid pDule-pAF<sup>2</sup> (Table 1, Group 1).<sup>25,26</sup>
Table 2. Amber Suppression Constructs and Their Toxicity

<table>
<thead>
<tr>
<th>Construct</th>
<th>Transferrable to R. sphaeroides&lt;sup&gt;+&lt;/sup&gt;</th>
<th>Transferrable to E. coli</th>
</tr>
</thead>
<tbody>
<tr>
<td>pIND4-RC-pAF0&lt;sup&gt;7&lt;/sup&gt;</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>✓</td>
</tr>
<tr>
<td>pIND4-RC-pAF1</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>X</td>
</tr>
<tr>
<td>pIND4-RC-pAF2</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>✓</td>
</tr>
<tr>
<td>pIND4-RC-pAF3</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>✓</td>
</tr>
<tr>
<td>pIND4-RC-GFP1</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>✓</td>
</tr>
<tr>
<td>pIND4-RC-GFP2</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>✓</td>
</tr>
<tr>
<td>pIND4-RC-GFP3</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>✓</td>
</tr>
<tr>
<td>pIND4-RC-pAFRNA1</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>✓</td>
</tr>
<tr>
<td>pIND4-RC-pAFRNA2</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>✓</td>
</tr>
<tr>
<td>pIND4-RC-pAFRNA3</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>X</td>
</tr>
<tr>
<td>pIND4-RC-pAFRS</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>✓</td>
</tr>
<tr>
<td>pIND4-RC-pCNFRS</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>X</td>
</tr>
<tr>
<td>pIND4-RC-HaloY1</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>✓</td>
</tr>
<tr>
<td>pIND4-RC-HaloY2</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>✓</td>
</tr>
<tr>
<td>pIND4-RC-HaloY3</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td>✓</td>
</tr>
</tbody>
</table>

<sup>+</sup>Transferrable, in this context, refers to the ability of a plasmid to produce colonies on agar plates under antibiotic selection following electroporation and outgrowth of transformed cells and is used as a measure of plasmid toxicity, as transformed cells do not produce growth on agar plates in contrast to positive controls attempted simultaneously. Plasmids then that were not transferrable were designated toxic. All constructs that did not yield colonies following transformation had successful positive controls (pIND4-RC transformation) with at least $1.0 \times 10^2$ cfu/μg. While pIND4-RC-pAF0<sup>7</sup> was not toxic in *R. sphaeroides* cells, amber suppression was not functional (Figure 4).

TAG. When the pIND4-RC-pAF0 Y(M210)TAG and wild-type (WT) variants were then compared in side-by-side screens to which pAF was added or withheld, no RC production was observed in Y(M210)TAG containing cells (see experimental details in Amber Suppression Screening below). This indicated aRS/tRNA expression was absent or impaired due to improper regulatory control.

To ensure translation of the aRS/tRNA pair, one aRS gene cassette was generated where the aRS was placed under control of the strong constitutive *P*m<sub>mb</sub> Promoter, RBS, and *T*<sub>omega</sub> terminator (*P*m<sub>mb</sub>−*pAFRS−*T*<sub>omega</sub>). All three had been shown to function in *R. sphaeroides* and were chosen due to their characterized nature and their regulatory strength in promoting high gene expression.<sup>30</sup>−<sup>32</sup> Due to the incompletely characterized nature of post-transcriptional tRNA processing in *R. sphaeroides*, transcriptional control of the tRNA was varied. Three tRNA gene cassettes were utilized (Figure 2A): one with the *R. sphaeroides* *P*m<sub>mb</sub> promoter and *T*<sub>mc</sub> terminator originally associated with the amber suppressor tRNA<sup>Tyr</sup> in *E. coli* plasmid pDule-pAF (<sup>P</sup>*m<sub>mb</sub>−*tRNA<sup>Tyr</sup>*E. coli* <sub>T*mc*</sub>), one with the *R. sphaeroides* *P*m<sub>mb</sub> promoter and *T*<sub>omega</sub> terminator (<sup>P</sup>*m<sub>mb</sub>−*tRNA<sup>Tyr</sup>−*T*<sub>omega</sub>), and the last with the promoter and terminator found on the strongly transcribed native *R. sphaeroides* tRNA (<sup>P</sup>*m<sub>leu2</sub>−*tRNA<sup>Tyr</sup>−*T*<sub>leu2</sub>).<sup>34</sup> The combination of the pIND4-RC vector, the aRS gene cassette (Figure 2A), and the three tRNA gene cassettes (Figure 2A), resulted in the construction of three plasmids: pIND4-RC-pAF1<sup>8</sup>, pIND4-RC-pAF2<sup>9</sup>, and pIND4-RC-pAF3<sup>10</sup> (Table 1, Group 3).

Amber Suppression Toxicity. In our initial attempts to carry out amber suppression in *R. sphaeroides*, we transformed strain RCx with pIND4-RC-pAF1<sup>8</sup>, pIND4-RC-pAF2<sup>9</sup>, and pIND4-RC-pAF3<sup>10</sup> by electroporation and plated on RLB/agar plates containing the appropriate antibiotic (kan- mycin (Kan) for pIND4-based constructs, see Methods). Here RLB is an LB-based media, modified for *R. sphaeroides* growth.<sup>35</sup> Despite repeated plasmid transformation attempts (more than 3 attempts), no cell growth was observed following transformation of Group 3 constructs for pAF incorporation under *R. sphaeroides* transcriptional control. In contrast, positive controls performed in each round of electroporation consisting of pIND4-RC<sup>1</sup> transformations consistently produced $1 \times 10^2$ to $1 \times 10^3$ cfu/μg (in general *R. sphaeroides* competency is low), and negative controls, where no plasmid had been added, consistently produce no *R. sphaeroides* colonies.<sup>37</sup> While it is likely that multiple factors (cell shock, low competency) contributed to the inability to transfer plasmids into viable *R. sphaeroides* cells, given the consistent and unambiguous success of positive controls, it seemed apparent that some attribute of the current constructs was toxic to transformed cells. This toxicity continued when transformed cells were supplemented with 1 mM pAF in RLB/pAF media during the outgrowth step and during antibiotic selection on RLB/Kan/pAF agar plates. To check that the aRS and tRNA in amber suppression constructs were functional and assembled properly, the amber suppression plasmid with aRS/tRNA genes under *E. coli* transcriptional control, pIND4-RC-pAF0<sup>7</sup>, was tested in *E. coli* for ncAA incorporation. The plasmid pIND4-RC-pAF0<sup>7</sup> was cotransformed with pBad-sfGFP-150TAG<sup>6</sup>, a reporter plasmid encoding GFP with an inserted stop codon at residue 150, into Thermo DH10-B cells.<sup>26</sup> Site 150 is a surface-exposed residue amenable to mutagenesis; in the absence of ncAA incorporation at position 150, only truncated non-fluorescent peptide is produced. In separate, parallel experiments, a WT pBad-sfGFP<sup>5</sup> was cotransformed with pIND4-RC-pAF0<sup>7</sup> as a positive control (similar to WT and Y(M210)TAG screening with RCs). By providing or withholding pAF (1 mM) from cotransformed cells, amber suppression was not functional (Figure 4).
suppression function was verified by GFP fluorescence present only when pAF had been added or in positive controls (Figure S2).

To test the general toxicity of non-aaRS/tRNA sequence features in aaRS/tRNA gene cassettes, namely regulatory elements of the tRNA and aaRS (promoters/RBS for the aaRS/terminators), aaRS and tRNA orientation, and contextual aaRS- or tRNA-specific linker sequences, control plasmids were constructed derived from the toxic plasmid, pIND4-RC-pAF2 [9]. This was done to ensure all aspects of the plasmid, even linking sequences and the orientation of the aaRS and tRNA genes (which should both have no function) were nontoxic. Here GFP was inserted in the place of the pAFRS and tRNA genes in aaRS and tRNA gene cassettes respectively. Then one or both cassettes were inserted into pIND4-RC as performed in Group 3 constructs. This was done with the $P_{ma}$-aaRS-$T_{omega}$ cassette to create pIND4-RC-GFP1 [11], the $P_{ma}$-tRNA-$T_{omega}$ cassette to create pIND4-RC-GFP2 [12], and both the $P_{ma}$-aaRS-$T_{omega}$ and the $P_{ma}$-tRNA-$T_{omega}$ cassettes to create pIND4-RC-GFP3 [13] (Table 1, Group 4). While transformation efficiency was lowered ($2.0 \times 10^{5}$ cfu/μg to $3.8 \times 10^5$ cfu/μg) it was not enough to impede successful plasmid transfer (Table 2). Growth in liquid culture was also typical with final turbidity following induction (OD700 = 2.9 ± 0.1) near equivalent to other pIND4-RC construct-containing RCx cultures (OD700 = 2.8 ± 0.2; OD700 measured due to 600 nm absorbance of photosynthetic chromophores produced by cells). Positive controls likewise produced $1.4 \times 10^5$ cfu/μg and negative controls produced no colonies (see Methods). Additionally, all transformed cells displayed constitutive GFP expression (Figure 3) while cells lacking GFP plasmids only exhibited weak autofluorescence.

Given that all other features of inserted aaRS/tRNA gene cassettes (contextual gene linker sequences and regulatory elements) and the pIND4-RC vector backbone were shown to be nontoxic and functional at expressing GFP, and amber suppression was shown to be functional in $E. coli$, we reasoned that the toxicity of pIND4-RC-pAF1 [8], pIND4-RC-pAF2 [9], and pIND4-RC-pAF3 [10] was due to a lack of orthogonality in the aaRS and/or tRNA. To better identify the cause of cell growth inhibition, control plasmids were made where only the aaRS or one of each of the three tRNA gene cassettes had been introduced: pIND4-RC-pAFtRNA1 [14], pIND4-RC-pAFtRNA2 [15], pIND4-RC-pAFtRNA3 [16], and pIND4-RC-pAFRS [17] (Table 1, Group 5). While all three transformations of tRNA insert constructs had typical transformation efficiencies ($1.1 \times 10^3$ to $2.4 \times 10^5$ cfu/μg), the construct containing only the aaRS cassette produced no colonies upon plating (Table 2). Again, our positive control had typical transformation efficiency ($1.4 \times 10^5$ cfu/μg) and our negative control produced no growth. The inability to transfer pIND4-RC-pAFRS [17] into $R. sphaeroides$ cells led us to conclude that the pAFRS is toxic and therefore not orthogonal in $R. sphaeroides$. Furthermore, it suggests that the toxicity of Group 3 plasmid constructs was due, at least in part, to this lack of orthogonality in the pAFRS (for discussion on the potential mechanism of toxicity see SI Section C).

Due to the apparent toxicity of the aaRS evolved for pAF incorporation, other strategies for reducing toxicity and/or choosing other possible nontoxic aaRS/tRNA pairs were pursued. Initially, the same $R. sphaeroides$-regulated aaRS/tRNA cassettes from three Group 3 plasmids were inserted into the lower copy pRK vector, given past reports that lower-copy plasmids ameliorated toxicity, but pRK construct transformations also failed to yield cell growth (SI Section D). Attempts to place the pAFRS under an inducible promoter, while successful in yielding colonies in the absence of pAFRS induction by IPTG, failed to produce RCs. This was perhaps due to the interplay of identical induction for RC and pAFRS genes (both lac-inducible promoters, SI Section E).
MJTyrs/tRNA^Tyr^ pair, which had been developed for p-cyanophenylalanine (pCNF) incorporation, was then investigated. To test for toxicity, the pCNF aaRS (pCNFRS) from the plasmid pDule-pCNF [3] was inserted into the pIND4-RC [1] backbone alone (since the tRNA for pCNF incorporation was the same as that used in pAF incorporation) making pIND4-RC-pCNFRS [18] (Table 1, Group 6). Like the MJTyrs evolved for pAF, however, the pCNFRS also displayed toxicity upon transformation (Table 2).

The toxicity induced by the M. jannaschii TyrRSs evolved for pAF and pCNF incorporation was a surprising result. More often in current literature for amber suppressor aaRS/tRNA pairs where a lack of orthogonality has been observed, the tRNA has been focused on as the source of impaired cellular fitness. Also, the absence of orthogonality in the host transfer of an archaeal-derived aaRS/tRNA pair already evolved in E. coli, to another Gram-negative bacterial host, was unexpected and demonstrates the potential for species-specific toxicity in amber suppressor aaRSs.

Though it is possible that with optimization (proper combination of inductible promoters for the pAFRS and RC genes, lower-strength promoters, or different media composition) an MJTyrs-derived aaRS/tRNA pair could function in R. sphaeroides, other aaRS/tRNA pairs of different origins were pursued. We hypothesized that an aaRS with a different mechanism for substrate (tRNA and/or amino acid) recognition might have a higher chance of success in achieving incorporation of tyrosine derivatives into R. sphaeroides. To this end, an aaRS/tRNA pair was examined which had been isolated from another organism and evolved from a different aaRS, the Methanosarcina barkeri pyrrolylsyl tRNA synthetase (MbPylRS). The particular aaRS/tRNA pair selected was one that has recently been evolved for incorporation of 3-chlorotyrosine (ClY) and other halogenated tyrosines (HaloY) and was generously provided by Dr. Ryan Mehl at Oregon State University (Table 1, plasmid [4] see footnote a). As with pAF, the MbPylRS/tRNA^ClY^ pair for HaloY incorporation was inserted into the tRNA and aaRS gene cassettes in the pIND4-RC vector yielding three plasmids: pIND4-RC-HaloY1 [19], pIND4-RC-HaloY2 [20], and pIND4-RC-HaloY3 [21] (Table 1, Group 7). Encouragingly, all constructs displayed transformability efficiency similar to that of pIND4-RC positive controls (1 × 10^2 to 1 × 10^3 cfu/µg).

Amber Suppression Screening and Verification. Amber suppression was then tested on RCs for site-specific 3-chlorotyrosine, 3-bromotyrosine (BrY), and 3-iodotyrosine (IY) incorporation as was done with pIND4-RC-pAF0. As with pIND4-RC-pAF0, Y(M210)TAG variants were created for each of the nontoxic HaloY incorporating plasmids: pIND4-RC-HaloY1 [19], pIND4-RC-HaloY2 [20], and pIND4-RC-HaloY3 [21]. These plasmids and their wild-type at M210 counterparts were then compared side-by-side in screens to which halogenated tyrosine was added or withheld to Y(M210)TAG RC variants. Expression of RCs in Y(M210)-TAG variants was compared to RC expression in the corresponding aaRS/tRNA plasmid containing wild-type RC. RC production was monitored by the characteristic near-IR (NIR) RC absorption (Figure 4, Figure S3). For each cell culture to which ncAA had been supplemented to the growth media (1 mM), RC production was observed, while in corresponding negative controls to which no ncAA had been added, no RC production occurred. This contrasted with the pIND4-RC-pAF0 plasmid, where the pAFRS and tRNA^Tyr^ were still under E. coli regulatory control. It was likely only transferrable to viable R. sphaeroides cells due to the lack of expression of amber suppressor genes due to improper regulatory control, as it did not produce any RCs in identical amber suppression screening upon addition of pAF. For all HaloY ncAAs screened here, none appeared to change final cell turbidity relative to samples where ncAA was withheld, indicating no toxicity was present for any ncAA.

To verify ncAA incorporation, liquid chromatography–mass spectrometry (LC–MS) was performed on isolated RCs to yield an intact subunit mass characterization. As a hydrophobic, three-subunit, ~100 kDa membrane protein complex, the RC has been difficult to analyze via mass spectrometry and has not been completely characterized in the literature, especially at the level of single-Dalton mass accuracy for intact subunits. In this study, LC separation of the three RC subunits and subsequent MS determined subunit mass with 10 ppm mass accuracy (approximately ±1 Da for a 30 kDa subunit, Figures S4–S7). LC–MS spectra supported halogenated tyrosine incorporation.
with expected 34, 78, and 126 Da mass increases on the M-subunit peak for the ClY, BrY, and IY RC variants, respectively (Figure 5). The peak for the M-subunit showed little to no detectable amount of WT M-subunit from a tyrosine mis- incorporation at position M210 (Figure S8). Nonspecific CIY incorporation into the RC (detected via CID fragmentation) occurred at only 4 of the potential 27 off-target sites and was less than 10% (within experimental error) in each case (SI Section F, Tables S4 and SS).

The M210 tyrosine chemical variants produced only minor changes in RC UV—vis—NIR absorption spectrum (Figure 6), relative to wild-type RC. This is in line with past results in the literature where perturbative mutations (ex. tyrosine mutated to tryptophan) caused only small (<10 nm) shifts in the NIR (or Q) region. Further structural and electron transfer kinetic characterization of these RC variants is ongoing.

Amber Suppression Efficiency Comparison. With a functional and nontoxic aaRS in hand (MbPylRS evolved for HaloY), we returned to the question of which regulatory elements could optimize the expression of this synthetase/tRNA in R. sphaeroides, and as a result, the efficiency of amber suppression. To do this, we screened the amber suppression plasmids mentioned earlier, pIND4-RC-HaloY1 [19], pIND4-RC-HaloY2 [20], and pIND4-RC-HaloY3 [21]. The ability of the three different constructs to express RCs with CIY at M210 was assessed by comparing RC screens performed as listed below. For each plasmid two samples contained the Y(M210)- TAG mutation and the third was wild-type at M210. Again, by testing protein expression in the presence and absence of 1 mM ncAA, amber suppression efficiency was assayed. Experiments were performed in triplicate, and, using NIR RC absorbance on detergent-solubilized RCs in cell lysates, crude yields were measured for each amber suppression plasmid where tRNA regulatory elements had been varied. A significant difference in amber suppression efficiency was observed. The most efficient construct, pIND4-RC-HaloY1 [19], had a yield (2.6 mg/L or 30 ± 2% WT yield), over an order of magnitude higher than the least efficient (pIND4-RC-HaloY2 [20] produced 0.16 mg/
L or 1.8 ± 0.5% WT yield, Figure S9 and Figure S3). Surprisingly, the construct in which the tRNA had been placed under the same regulatory control as that of a native highly expressed tRNA, produced 5 times less RC than the most efficient construct (0.5 mg/L or 11 ± 2% WT yield). For the other variants, pIND4-RC-HaloY (with TAG at M210) produced RCs containing BrY at M210 with similar yields (2 mg/L) as CIY variants, while IY containing variants were produced with less than half that yield (0.7–0.9 mg/L). Even for the lowest yielding RC variant, IY at M210, yields were tractable for transient absorption, X-ray crystallography, and other biophysical characterization of RC variants under typical RC production conditions (3–6 L cultures).

In conclusion, we report here the first example of site-specific ncAAs in R. sphaeroides via amber suppression. The difficulty initially faced in transferring the MjTyrRS/tRNA\textsubscript{Tyr} pair was unexpected, given the pair was derived from archaea and then evolved to be orthogonal in another Gram-negative bacterium (E. coli). A potential explanation is that organisms in the same domain (bacteria) may have distinct tRNA pools so an aaRS that is orthogonal in one organism may not be in another (SI Section C). Through use of the more orthogonal M. barkeri Pyr/\textsubscript{C}U/\textsubscript{A}U pair, these problems with toxicity were overcome. The MbPylRS/tRNAPyl\textsuperscript{Pyf} pair has been found to be orthogonal across a wide range of organisms, a point attributed to the exotic sequence features of the Mb tRNAPyl\textsuperscript{Pyf}, such as its unusually short variable arm, and as found in recent structural studies, the N- and C-terminal domains of its cognate aaRS which possess more interfacial surface area than any of the known 20 canonical aaRSs.\textsuperscript{35,44} It is possible the combination of these features allows pyrrolyl-tRNA synthetases to recognize features which are truly distinct and not likely to be found on tRNAs from many organisms, and in our case R. sphaeroides tRNAs, whereas the orthogonality of MjTyrRS to E. coli tRNAs may be more incidental.\textsuperscript{35} Regulation of tRNA transcription seen in the three combinations of promoter/terminators used here had a major influence on amber suppression efficiency, with over an order of magnitude difference in yields. These improvements highlight the importance of optimizing tRNA regulation.

In the process of verification of ncAAs incorporation, LC–MS and LC–MS/MS characterization of this membrane complex was dramatically enhanced. These methods provide valuable tools for additional ncAA incorporation and general protein engineering verification of this important photoactive protein. Introduction of CIY, BrY, and IY into M210 in the RC has shown minimal perturbation in the absorption spectrum, and studies of electron transfer kinetics are underway to test the role of tyrosine M210.

**METHODS**

**Plasmid Construction and Cell Strains.** All plasmids were constructed through use of the NEBuilder HiFi DNA Assembly Master Mix. All aaRS and tRNA gene cassettes were ordered as G-blocks from IDT, along with GFP inserts flanked by aaRSs or tRNA gene cassettes. The pIND4-RC vector was linearized with the KpnI restriction enzyme and G-blocks for aaRSs, tRNAs, and GFP inserts were ordered with necessary homologous sequences for vector/insert assembly. All aaRSs and tRNAs were designed based on sequence information from pDule-pAF, pDule-pCNF, and pDule-HaloY (see SI Section B). E. coli-derived promoters and terminators used in this paper were also derived from promoters and terminators originally located on the pDule-pAF and pDule-HaloY aaRS/tRNA genes developed for E. coli. Plasmids were transformed into NEB 5-alpha Competent E. coli cells. E. coli strains were cultured at 37°C and 225 rpm in LB broth. Mutagenesis to insert the amber stop codon at M210 was performed with the QuickChange Lightning Site-Directed Mutagenesis Kit on RC-containing pIND4-RC vectors prior to HiFi assemblies. All plasmids generated in this study and mutations made were verified by DNA sequencing (see SI for further detail on vector and insert sequences and plasmid construction).

Plasmids were transformed into the R. sphaeroides rsh\textsubscript{L}, LHI, LHI\textsubscript{II}, and RC- (H, L, and M) deletion strain, RCx; here LHI and LHII are light harvesting complexes I and II. Due to the chromosomal deletion of the rshI gene from the RCx strain produced by the Beatty lab (and in a similar deletion strain produced by Philip Laible and Deborah Hansen) plasmids could be transferred via electroporation, greatly facilitating plasmid transfer.\textsuperscript{35,45} Electroporation cell preparation was similar to previously published literature: 2 × 80 mL RCx cells were grown in semi-aerobic growth in 2 × 125 mL Erlenmeyer flask until at early to mid log phase growth (OD\textsubscript{600} = 0.4) at 33°C.\textsuperscript{35,37,45} Cells were then pelleted at 4600g and resuspended at 4°C three times with 2 × 40 mL, 2 × 20 mL, and 1 × 20 mL of ice cold sterile Millipore water. Cells were then resuspended in 10 mL of 10% glycerol, pelleted again at 4600g and finally resuspended in 1 mL 10% glycerol. Aliquots of 40 μL were electroporated with 100 ng of plasmid at 2500 V, 25 μF, and 200 ohms, and then 900 μL of RLB media (see below) was added. Cells were then allowed to recover by culturing at 35°C and 150 rpm for 4 h and then plated on RLB agar plates supplemented with 25 μg/mL Kan. In each transformation testing toxicity of plasmid constructs (Table 2), pIND4-RC [1] plasmids were transformed as positive controls (every plasmid constructed in this study had a pIND4-RC vector backbone). Each transformation testing toxicity also had a negative control where plasmid was withheld and electroporated cells were allowed to recover and plated as in other construct and positive control transformations. In each transformation of plasmids containing a PAFRS (Table 1, Group 3–5 Plasmids), pAF was not added to RLB/Kan agar plates. Transformations repeated on Group 3 plasmids where pAF was supplemented to RLB/Kan agar plates also had identical results, demonstrating pAF addition did not alleviate toxicity.

For liquid culture, R. sphaeroides cells were grown at 30°C and 200 rpm in RLB, LB broth which has been supplemented with 810 μM MgCl\textsubscript{2} and 510 μM CaCl\textsubscript{2}.\textsuperscript{33} Kanamycin selection concentrations were 25 μg/mL for R. sphaeroides and 30 μg/mL for E. coli.

**Amber Suppression Screening and Efficiency.** Amber suppression function was verified by the presence of RC expression. R. sphaeroides cell cultures were made in groups of 3 × 50 mL in 250 mL Erlenmeyer flasks to allow for aerobic growth. In two of the 50 mL cultures the plasmid of interest (possessing aaRSs and tRNA gene cassettes) contained the amber codon at M210, while the third 50 mL culture was wild-type at M210. Normally aerobic growth conditions for production of RCs are avoided because the RC is generally under control of the puf or puc promoter and the presence of O	extsubscript{2} down-regulates genes under puf or puc control.\textsuperscript{35} RC induction is then produced by naturally decreasing oxygen concentrations as cell turbidity increases during cell culture. The RCx deletion strain and the pIND4-RC plasmid developed...
by Jun et al. and others, remove O2 suppression and allowed the RC genes to be placed under control of the IPTG-inducible lac operon.\textsuperscript{35} This feature not only allowed aerobic culture, which sped up cell growth rate, but was also helpful in allowing RC production to be coordinated with addition of the ncAA (CIY, BrY, IY) for initiating amber suppression.

RC expression was induced by addition of 50 μL of 1 M IPTG once cell culture reached early log phase (OD600 = 0.2) to achieve a final concentration of 1 mM IPTG in cell media. At the same time 0.5 mL of 100 mM of the ncAA of interest was added, taking the final ncAA final concentration to 1 mM. While CIY and BrY were solubilized to 100 mM in water (with the help of gentle heating below 55°C for less than an hour and sonication), IY required the addition of 160 μL of 8 M NaOH for 10 mL of 100 mM IY to fully solubilize similar to Mehl et al. 2007.\textsuperscript{46} Induction was allowed to proceed for 36 h at 30°C and 200 rpm.

Cells were pelleted and resuspended in 5 mL Lysis Buffer (10 mM Tris, 150 mM NaCl) then frozen in liquid nitrogen. To proceed with the preparation, cells were thawed, minimal mM Tris, 150 mM NaCl) then frozen in liquid nitrogen. To

Agilent Pursuit 5 diphenyl 150 μm column in an Agilent 1260 HPLC and Bruker MicroTOF-QII.\textsuperscript{48} Temperature on the column was 80°C, the flow rate was 0.3 mL/min and the injection volume was 5 μL. Samples were injected at a concentration of 3.5 μM RC in 10 mM Tris, 0.03% LDAO, pH = 8.0 buffer. To elute peptides a gradient containing 0.1% formic acid was run from water (with 0.05% trifluoroacetic acid) to acetonitrile (both also contained formic acid, Table S1). Data was collected in full scan MS mode with a mass scan range of 400–4000 Da. Samples were ionized with ESI and the collision RF setting was 800 Vpp.

LC–MS/MS was used to verify site-specificity of CIY incorporation at M210 in doubly enzymatically digested RC samples. Sample containing 20 μg of RCs in 4.3 μL of 0.03% LDAO, 10 mM Tris, pH = 8.0 was taken to 50 μL with 50 mM ammonium bicarbonate. The sample was then reduced by taking the solution to 10 mM dithiothreitol for 30 min at 55°C and reduced cysteines capped by addition of acrylamide to 30 mM overall concentration and incubation for 30 min at room temperature. After reduction and alkylation proteins were digested with AspN for 2 h at 37°C following by Trypsin/LysC overnight at 37°C in the presence of 0.02% ProteaseMax (Promega). Formic acid was then added to a final concentration of 2.5% to terminate digestion. Samples were then analyzed by a Thermo Scientific Orbitrap Fusion with a nanoLC column. Peptides were eluted in an 80 min gradient and were ionized by ESI (Table S3). Fragmentation for MS2 spectra was obtained by three different methods, electron-transfer dissociation (ETD), higher-energy collision-induced dissociation (HCD), and collision-induced dissociation (CID) to maximize sequence coverage.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.8b00100.

Orthogonality discussions, Plasmid map and sequence information, lower plasmid copy and IPTG-induction attempts to decrease pAFRS toxicity, amber suppression in E. coli, mass spectrometry spectra and experimental details, tRNA transcriptional regulation effect on yield and NIR absorption screens (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: sboxer@stanford.edu.

ORCID

Jared Bryce Weaver: 0000-0002-3823-422X

Steven G. Boxer: 0000-0001-9167-4286

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Drs. Thomas Beatty and Daniel Jun at the University of British Columbia for their generous gifts of the plasmid pIND4-RC and the RCx strain, along with many helpful discussions. We likewise thank Dr. Ryan Mehl, Rachel Henson, and Dr. Joseph Porter at Oregon State University for the maps and sequence information on the pDule-pAF, pDule-pCNF, and pDule-HaloY, many helpful discussions and their generous provision of pDule-HaloY prior to publication. We thank Dr. Stephen Fried for his advice to trouble-shooting toxicity in amber suppression plasmid constructs and critical reading of
the manuscript, and Drs. Philip Laible and Deborah Hanson of the Argonne National Laboratory for their help in working with R. sphaeroides and assessing early challenges in this study. We also thank Theresa McLaughlin, Dr. Ryan Leib, Dr. Christopher Adams, and the Vincent Coates Foundation Mass Spectrometry Laboratory, Stanford University Mass Spectrometry for all their support performing the LC−MS and LC−MS/MS in this study. This work is supported by a grant from the NSF Biophysics Program (MCB-1408785).

**ABBREVIATIONS**

RC, reaction center; P, special pair; ncAA, noncanonical amino acid; aaRS, aminoacyl tRNA synthetase; MtTyrRS, Methanocaldococcus jannaschii tyrosyl-tRNA synthetase; pAF, p-aminophenylalanine; RBS, ribosome binding site; Kan, kanamycin; WT, wild-type; IPTG, isopropyl β-D-1-thiogalactopyranoside; pCNF, p-cyano-phenylalanine; MbPytRS, Methanosarcina barkeri pyrrolyl-tRNA synthetase; pAFRS, p-aminophenylalanyl-tRNA synthetase; pCNFRS, p-cyano-phenylalanyl-tRNA synthetase; CIY, 3-chlorotyrosine; BrY, 3-bromotyrosine; IY, 3-iodotyrosine; HaloYRS, halotyrosyl-tRNA synthetase; LC−MS, liquid chromatography−mass spectrometry; LC−MS/MS, liquid chromatography−tandem mass spectrometry; LDAO, dimethyldecylammonium N-oxide; ETD, electron-transfer dissociation; HCD, higher-energy collision-induced dissociation; CID, collision-induced dissociation; LHI, light harvesting complex I; LHII, light harvesting complex II

**REFERENCES**


plasmid for Rhodobacter sphaeroides and Paracoccus denitrificans. 


