

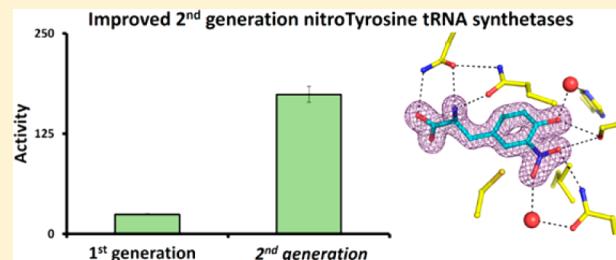
## Structural Basis of Improved Second-Generation 3-Nitro-tyrosine tRNA Synthetases

Richard B. Cooley, Jessica L. Feldman, Camden M. Driggers, Taylor A. Bundy, Audrey L. Stokes, P. Andrew Karplus, and Ryan A. Mehl\*

Department of Biochemistry and Biophysics, Oregon State University, 2011 Agriculture and Life Sciences Building, Corvallis, Oregon 97331, United States

### Supporting Information

**ABSTRACT:** Genetic code expansion has provided the ability to site-specifically incorporate a multitude of noncanonical amino acids (ncAAs) into proteins for a wide variety of applications, but low ncAA incorporation efficiency can hamper the utility of this powerful technology. When investigating proteins containing the post-translational modification 3-nitro-tyrosine (nitroTyr), we developed second-generation amino-acyl tRNA synthetases (RS) that incorporate nitroTyr at efficiencies roughly an order of magnitude greater than those previously reported and that advanced our ability to elucidate the role of elevated cellular nitroTyr levels in human disease (e.g., Franco, M. et al. *Proc. Natl. Acad. Sci. U.S.A.* 2013, 110, E1102). Here, we explore the origins of the improvement achieved in these second-generation RSs. Crystal structures of the most efficient of these synthetases reveal the molecular basis for the enhanced efficiencies observed in the second-generation nitroTyr-RSs. Although Tyr is not detectably incorporated into proteins when expression media is supplemented with 1 mM nitroTyr, a major difference between the first- and second-generation RSs is that the second-generation RSs have an active site more compatible with Tyr binding. This feature of the second-generation nitroTyr-RSs appears to be the result of using less stringent criteria when selecting from a library of mutants. The observation that a different selection strategy performed on the same library of mutants produced nitroTyr-RSs with dramatically improved efficiencies suggests the optimization of established selection protocols could lead to notable improvements in ncAA-RS efficiencies and thus the overall utility of this technology.



Post-translational modifications (PTMs) of proteins are widely involved in biological regulation through the alteration of proteins' properties. The variability in the lifetime and stoichiometry of natural protein PTMs and the extensive crosstalk that can occur between them<sup>1</sup> makes the *in vivo* study of PTMs very challenging. While mass spectrometric studies have significantly advanced our understanding of how extensively PTMs are involved in biological processes (e.g., ref 2), mass spectrometry does not provide information about how PTMs alter a protein's functionality. Adding to the challenge of studying the functional effects of PTMs is that the homogeneously modified protein needed for biochemical characterization can be difficult to produce. Some PTMs can be installed by enzymes or chemical oxidants, but these methods are often hard to control and result in a mixture of protein forms that need to be separated from each other and from unmodified protein before biochemical characterization. Genetic code expansion<sup>3–5</sup> is one of the most powerful approaches to generate protein-containing PTMs because the method is site specific and, in principle, can produce quantities of modified protein sufficient for biochemical and biophysical analyses. It also even has the potential to produce modified proteins in a variety of cells. But the power of these genetic code expansion efforts depends on the efficiency (the ability to incorporate a noncanonical amino acid) and fidelity

(the ability to discriminate against canonical amino acids) of the non-canonical amino acid aminoacyl tRNA synthetases (ncAA-RS) that are developed.

Increased cellular levels of nitrated tyrosine residues have been detected in at least 80 major human diseases.<sup>6,7</sup> Recent bioinformatics analyses have also suggested that approximately 9% of the human proteome is nitratable<sup>8</sup> even though not all surface-exposed tyrosine residues are equally susceptible to such chemical modification. These observations prompted us to use genetic code expansion to encode 3-nitro-tyrosine (nitroTyr) site-specifically into recombinant proteins.<sup>9</sup> We generated an orthogonal ncAA-RS/suppressor tRNA (RS/tRNA<sub>CUA</sub>) pair that could incorporate nitroTyr in response to an amber (TAG) stop codon.<sup>9</sup> Using this first-generation tool, we showed that nitration of Tyr34 in mitochondrial superoxide dismutase is sufficient to inactivate this key antioxidant enzyme.<sup>9</sup> When attempting to use this nitroTyr-RS to establish possible pathological roles of tyrosine nitration of heat shock protein 90 (Hsp90) in amyotrophic lateral sclerosis<sup>10</sup> and of apolipoprotein A1 in atherosclerosis, we discovered that the nitroTyr incorporation

Received: January 29, 2014

Revised: February 28, 2014

Published: March 10, 2014

efficiencies of our first-generation nitroTyr-RSs were too low to produce sufficient yields of single nitroTyr- and multiple nitroTyr-containing Hsp90 and apoA1 for biological evaluation.

We thus sought to obtain a more efficient nitroTyr incorporation system. Several approaches have been used to improve the utility of such systems, depending on the source of the bottleneck, including mutagenesis of ncAA-RS active-site amino acids,<sup>11,12</sup> optimization of the suppressor tRNA recognition,<sup>5</sup> modification of RS proof-reading domains,<sup>13</sup> replacement of endogenous tRNA/RS pairs,<sup>14</sup> modification of elongation factors,<sup>15</sup> and reassignment of chromosomal amber codons.<sup>16</sup> Although identifying the most significant factor hindering expression of ncAA-containing proteins is not always trivial, the most common source of poor expression is suboptimal recognition of the ncAA by the RS. We therefore attempted to extensively mutagenize our first-generation nitroTyr-RSs; however, this did not result in significant functional improvements. Alternatively, we decided to rescreen for nitroTyr-RSs from the same initial synthetase library by modifying the process in which RSs were selected from a library of mutants. This altered selection process produced a group of second-generation nitroTyr-RSs with improved efficiencies that were useful for successfully incorporating nitroTyr at multiple sites in both Hsp90 and apoA1.<sup>10</sup> Here, we characterize the efficiency, fidelity (the ability of the synthetase to discriminate against the 20 canonical amino acids), and permissivity (the ability of a synthetase to incorporate ncAAs structurally similar to the parent ncAA) of these second-generation nitroTyr-RSs. Crystal structures of the most efficient of these new nitroTyr-RSs reveal insight into the molecular origin of their high efficiency compared to the first-generation RSs, and interestingly provide clues regarding why these second-generation RSs were discarded in the original selection process. They also bring to light the previously under recognized complex nature of RS fidelity. The possibility of improving the efficiency of other ncAA-RSs through similarly altered selection processes, and thus the overall utility of this powerful technology, is discussed.

## MATERIALS AND METHODS

**Selection of Second-Generation Aminoacyl-tRNA Synthetases Specific for nitroTyr.** For a detailed description of the positive selection plasmid, pCG, negative selection plasmid, pNEG, and aminoacyl synthetase library plasmid, pBK-3D-Lib see Supporting Information (SI). Scheme 1 depicts a brief comparison of the selection strategies utilized for creating the first- and second-generation synthetases. For the positive

selection of the second-generation synthetases, 2 mL of pCG/pBK-3D-Lib cells were thawed on ice before addition to 1.2 L of room temperature 2 × YT media containing 50 µg/mL kanamycin (Kn) and 25 µg/mL tetracycline (Tet). After incubation (11 h, 250 rpm, 37 °C), a 200 µL aliquot of these cells was plated on eleven 15 cm LB agar plates containing 50 µg/mL Kn, 25 µg/mL Tet, and 40 µg/mL chloramphenicol (Cm). The positive selection agar medium also contained 1 mM nitroTyr. After spreading, the surface of the plates was allowed to dry completely before incubation (37 °C, 15 h). To harvest the surviving library members from the plates, 10 mL of 2 × YT (50 µg/mL Kn, 25 µg/mL Tet) was added to each plate. Colonies were scraped from the plate using a glass spreader. The resulting solution was incubated with shaking (60 min, 37 °C) to wash cells free of agar. The cells were then pelleted, and plasmid DNA was extracted using a Qiagen midiprep kit. The smaller pBK-3D-Lib plasmid was separated from the larger pCG plasmid by agarose gel electrophoresis and extracted from the gel using the Qiagen gel extraction kit.

The purified pBK-3D-Lib was then transformed into pNEG-containing DH10B cells. A 100 µL sample of pNEG electrocompetent cells was transformed with 50 ng of purified pBK-3D-Lib DNA. Cells were rescued in 1 mL of SOC for 1 h (37 °C, 250 rpm) and the entire 1 mL of rescue solution was plated on three 15 cm LB plates containing 100 µg/mL Amp, 50 µg/mL Kn, and 0.2% L-arabinose. Cells were collected from plates, and pBK-3D-Lib plasmid DNA was isolated in the same manner as described above for positive selections.

In order to evaluate the success of the selections based on variation in synthetase efficacy (as opposed to traditional survival/death results), the synthetases resulting from the selection rounds were tested with the pALS plasmid. This plasmid contains the sfGFP reporter with a TAG codon at residue 150 as well as tyrosyl-tRNA<sub>CUA</sub>. When a pBK plasmid with a functional synthetase is transformed with the pALS plasmid and the cells are grown in the presence of the appropriate amino acid on autoinduction agar, sfGFP is expressed, and the colonies are visibly green. One microliter of each library resulting from the negative selection was transformed with 60 µL of pALS-containing DH10B component cells. The cells were rescued for 1 h in 1 mL of SOC (37 °C, 250 rpm). Aliquots of cells (250 and 50 µL) from each library were plated on autoinducing media plates with 25 µg/mL Kn, 25 µg/mL Tet, and 1 mM nitroTyr. Autoinducing agar plates were prepared by combining the reagents in supplemental Table 1 in SI with an autoclaved solution of 40 g of agarose in 400 mL water. Sterile water was added to a final volume of 500 mL. Plates were grown at 37 °C for one day and then grown on the benchtop, at room temperature, for an additional day.

A total of 68 visually green colonies were selected from autoinduction plates and used to inoculate a 96-well plate containing 0.5 mL per well noninducing media (supplemental Table 1 in SI, with sterile water added to a final volume of 500 mL) with 25 µg/mL Kn, 25 µg/mL Tet. After 24 h of growth (37 °C, 250 rpm), 5 µL of these noninducing samples were used to inoculate 96-well plates with 0.5 mL autoinduction media (supplemental Table 1 in SI, with sterile water added to a final volume of 500 mL) containing 25 µg/mL Kn, 25 µg/mL Tet with and without 1 mM nitroTyr. Fluorescence measurements of the cultures were collected 40 h after inoculation using a HORIBA Jobin Yvon FluoroMax-4. The emission from 500 to 520 nm (1 nm bandwidth) was summed with excitation at 488 nm (1 nm bandwidth). Samples were prepared by diluting

Scheme 1

| Selection          | 1 <sup>st</sup> Generation | 2 <sup>nd</sup> Generation |
|--------------------|----------------------------|----------------------------|
| Media              | Defined                    | Rich                       |
| Round 1 positive   | 60 µg/ml Cm*               | 40 µg/ml Cm                |
| Round 1 negative   | Barnase                    | Barnase                    |
| Round 2 positive   | 60 µg/ml Cm                | N/A                        |
| Round 2 negative   | Barnase                    | N/A                        |
| Round 3 positive   | 80 µg/ml Cm                | N/A                        |
| Round 3 negative   | Barnase                    | N/A                        |
| Final RS selection | Lethal 100 µg/ml Cm        | Non-lethal GFP reporter    |

\*Cm: chloramphenicol

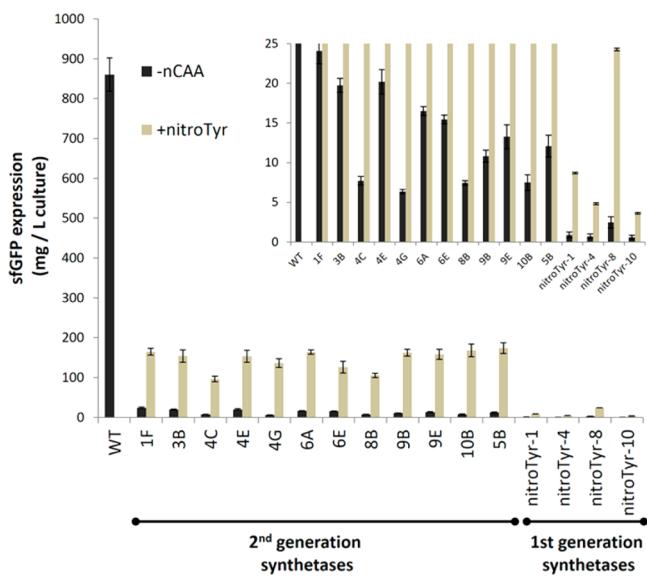
suspended cells directly from culture 100-fold with phosphate buffer saline (PBS). Sequencing of the top 12 performing clones showed that all had unique sequences (Table 1). The top

**Table 1. Sequences for Second- and First-Generation nitroTyr-RSs at the Positions That Were Allowed to Vary in the Library Used for Screening**

| Variant                                          | Residue number |    |    |     |     |     |     |     |     |  |
|--------------------------------------------------|----------------|----|----|-----|-----|-----|-----|-----|-----|--|
|                                                  | 32             | 65 | 70 | 108 | 109 | 155 | 158 | 159 | 162 |  |
| Wild-type                                        | Y              | L  | H  | F   | Q   | Q   | D   | I   | L   |  |
| <i>Second generation synthetases</i>             |                |    |    |     |     |     |     |     |     |  |
| nitroTyr-5B                                      | H              | L  | C  | F   | Q   | Q   | S   | A   | R   |  |
| nitroTyr-1F                                      | L              | V  | A  | R   | P   | Q   | S   | L   | H   |  |
| nitroTyr-3B                                      | S              | T  | L  | F   | Q   | Q   | S   | A   | R   |  |
| nitroTyr-4C                                      | R              | T  | G  | F   | Q   | Q   | S   | E   | H   |  |
| nitroTyr-4E                                      | L              | H  | G  | S   | K   | Q   | S   | L   | H   |  |
| nitroTyr-4G                                      | R              | T  | G  | F   | Q   | Q   | G   | E   | H   |  |
| nitroTyr-6E                                      | V              | V  | C  | F   | Q   | Q   | S   | A   | H   |  |
| nitroTyr-6A                                      | Q              | L  | L  | F   | Q   | Q   | S   | P   | R   |  |
| nitroTyr-8B                                      | H              | L  | A  | F   | Q   | P   | S   | C   | S   |  |
| nitroTyr-9B                                      | A              | T  | M  | F   | Q   | Q   | S   | G   | R   |  |
| nitroTyr-9E                                      | R              | L  | A  | F   | Q   | Q   | S   | P   | S   |  |
| nitroTyr-10B                                     | A              | T  | T  | F   | Q   | Q   | S   | S   | R   |  |
| <i>First generation synthetases</i> <sup>a</sup> |                |    |    |     |     |     |     |     |     |  |
| nitroTyr-1                                       | E              | N  | T  | F   | Q   | Q   | S   | T   | L   |  |
| nitroTyr-4                                       | K              | A  | M  | F   | Q   | Q   | G   | M   | R   |  |
| nitroTyr-8                                       | R              | L  | L  | F   | Q   | M   | G   | L   | H   |  |
| nitroTyr-10                                      | K              | K  | S  | A   | D   | Q   | G   | Y   | S   |  |

<sup>a</sup>First-generation nitroTyr-RS sequences as reported by Neumann et al.<sup>9</sup>

performing clones were moved from the *pBK-RS* plasmid to the *pDule* plasmid (*pDule-RS*) and evaluated with the pBad-GFP150 reporter plasmid (Figure 1). *pDule* plasmid was generated by



**Figure 1.** *In vivo* expression yields of nitroTyr-150-sfGFP using the selected nitroTyr *Mj* synthetases. The total yield of sfGFP containing nitroTyr at position 150 expressed in autoinducing cultures (see Materials and Methods) is shown in milligrams of protein per liter of culture for each RS in media lacking (black) and supplemented with (gray) 1 mM nitroTyr. First-generation nitroTyr-RSs are those reported previously,<sup>9</sup> while the second-generation nitroTyr-RSs are characterized here. Inset contains identical data except the y-axis has been adjusted to more easily observe values for the low-yielding expression cultures.

amplifying the *MjYRS* gene from the *pBK* plasmid isolated from the library using primers RSmovef (5'-CGCGCGCCATGGACGAATTGAAATG-3') and RSmovev (5'-GACTCAGTCTAGGTACCGTTGAAACTGCAGTTATA-3'). The amplified DNA fragments were cloned into the respective sites on the *pDule* plasmids using the incorporated *Nco*I and *Kpn*I sites.

**Expression, Purification and Crystallization the of nitroTyr-5B RS.** The DNA fragment containing nitroTyr-5B-RS was amplified by PCR from the *pDule* plasmid and ligated into the *Nco*I/*Xba*I sites of the expression vector *pET28a* (forward primer: 5'-CGCGCGCCATGGACGAATTGAAATG-3', reverse primer: 5'-GGCGCTCGAGTAATCTCTTCTATTGGCTCTAAAATC-3'). This vector was transformed into DH10B cells, purified with a QIAprep spin mini kit (Qiagen), and transformed into BL21(DE3) cells for expression. Protein was expressed and purified similar to previous descriptions<sup>17</sup> with only minor modifications (see SI).

Crystals of the apo-open and apo-closed nitroTyr-5B RS were grown using the hanging-drop vapor diffusion method at room temperature against a reservoir containing 22–23% polyethylene glycol (PEG) 300, 5% PEG 8000, 10% glycerol, and 100 mM Tris pH 8.0–8.3 by mixing 2 μL of protein (16 mg/mL in 20 mM Tris, 50 mM NaCl, 10 mM β-mercaptoethanol, pH 8.5) with 2 μL of reservoir. Crystals grew to full size (~300 × 100 × 100 μm) within 4–6 days, at which time they were harvested by directly submerging in liquid nitrogen. Both apo-open and apo-closed crystals were present in the same individual drops and were indistinguishable by eye.

For the nitroTyr structure, crystals were grown in an identical manner except that the nitroTyr-5B RS was preincubated with 2 mM nitroTyr. A preliminary data set of a crystal grown under these conditions resulted in a structure identical to that of the apo-closed crystals with Cys70 similarly modified (see Results and Discussion); thus, all subsequent crystals were soaked in mother liquor supplemented with 10 mM DTT and 100 mM nitroTyr for 1 h at room temperature prior to freezing in liquid nitrogen.

**Data Collection, Structure Determination and Refinement of the nitroTyr-5B RS.** Data were collected using beamlines 5.0.2 and 4.2.2 at the Advanced Light Source (Lawrence Berkeley National Laboratory) at a temperature of 100 K. Data were processed and scaled in space group *P4*<sub>3</sub>*2*<sub>1</sub>*2* using iMosflm v1.0.040 and SCALA, and 5% of the data were randomly flagged for use in *R*<sub>free</sub>. Data collection statistics are reported in Table 2.

Structures were determined by molecular substitution<sup>18</sup> using previously solved structures (PDBs 1u7d for the apo-open structure and 2zp1 with substrate omitted for the apo-closed and substrate-bound structures). The majority of the apo crystals adopted the apo-open form. All model building and refinements were conducted using Coot v0.6.1<sup>19</sup> and Phenix v1.8,<sup>20</sup> respectively. Standard criteria were used for modeling water molecules (>1  $\rho_{rms}$  intensity in the  $2F_o - F_c$  map, >2.4 Å distance from nearest contact). Translation/libration/screw (TLS)<sup>21</sup> refinement of B-factors was performed for each structure using two groups which roughly correspond to the N- and C-terminal domains (residues 1–194 and 195–307). For the closed-nitroTyr structure, substrate was omitted from the model until the final rounds of refinement. Molecular restraints for nitroTyr were created using Phenix-Elbow.<sup>20</sup> Molprobity<sup>22</sup> was used to monitor model geometry. For all three structures independently, water molecules were sorted by their peak  $2F_o - F_c$  electron density, and numbered so that Wat1 is the water with the

Table 2. Data Collection and Refinement Statistics for the nitroTyr-5B RS

|                                                          | apo-open                         | apo-closed                       | closed-nitroTyr                  |
|----------------------------------------------------------|----------------------------------|----------------------------------|----------------------------------|
| Data Collection <sup>a</sup>                             |                                  |                                  |                                  |
| space group                                              | P4 <sub>3</sub> 2 <sub>1</sub> 2 | P4 <sub>3</sub> 2 <sub>1</sub> 2 | P4 <sub>3</sub> 2 <sub>1</sub> 2 |
| unit cell axes <i>a</i> , <i>b</i> , <i>c</i> (Å)        | 100.14, 100.14, 73.29            | 101.77, 101.77, 72.28            | 102.1, 102.1, 71.72              |
| resolution limits (Å)                                    | 50.57–2.00 (2.09–2.00)           | 51.00–2.00 (2.10–2.00)           | 23.99 – 1.70 (1.79–1.70)         |
| unique observations                                      | 25,254 (3549)                    | 26,261 (3737)                    | 41,621 (5634)                    |
| completeness                                             | 96.6 (94.9)                      | 100.0 (100.0)                    | 98.7 (98.7)                      |
| multiplicity                                             | 12.9 (11.4)                      | 16.1 (13.3)                      | 9.3 (5.4)                        |
| average <i>I</i> / <i>σ</i>                              | 13.2 (0.8)                       | 12.7 (1.2)                       | 13.7 (1.5)                       |
| <i>R</i> <sub>meas</sub> <sup>b</sup> (%)                | 9.3 (172)                        | 12 (225)                         | 9.4 (142)                        |
| CC <sup>1/2c</sup>                                       | 1.00 (0.38)                      | 1.00 (0.42)                      | 1.00 (0.25)                      |
| Refinement                                               |                                  |                                  |                                  |
| <i>R</i> <sub>cryst</sub> / <i>R</i> <sub>free</sub> (%) | 19.9/24.9                        | 22.3/27.8                        | 18.0/22.3                        |
| no. protein molecules                                    | 1                                | 1                                | 1                                |
| no. protein residues                                     | 306                              | 310                              | 308                              |
| no. water molecules                                      | 125                              | 134                              | 300                              |
| total number atoms                                       | 2573                             | 2596                             | 2914                             |
| rmsd bond angles (deg)                                   | 1.08                             | 1.05                             | 1.04                             |
| rmsd bond lengths (Å)                                    | 0.007                            | 0.007                            | 0.007                            |
| <B> protein (Å <sup>2</sup> )                            | 36.5                             | 36.6                             | 20.7                             |
| <B> water (Å <sup>2</sup> )                              | 39.9                             | 35.0                             | 30.4                             |
| <B> nitroTyr (Å <sup>2</sup> )                           | N/A                              | N/A                              | 9.9                              |
| Ramachandran Plot (%) <sup>d</sup>                       |                                  |                                  |                                  |
| favored                                                  | 97.7                             | 98.4                             | 99.1                             |
| outliers                                                 | 0.3                              | 0.0                              | 0.3                              |
| PDB code                                                 | 4ND6                             | 4ND7                             | 4NDA                             |

<sup>a</sup>Numbers in parentheses correspond to values in the highest resolution bin. <sup>b</sup>*R*<sub>meas</sub> is the multiplicity-weighted merging *R*-factor.<sup>30</sup> <sup>c</sup>Correlation coefficient between two randomly chosen subsets containing the average intensities of each unique reflection.<sup>31</sup> <sup>d</sup>Ramachandran plot generated using Molprobity.<sup>22</sup>

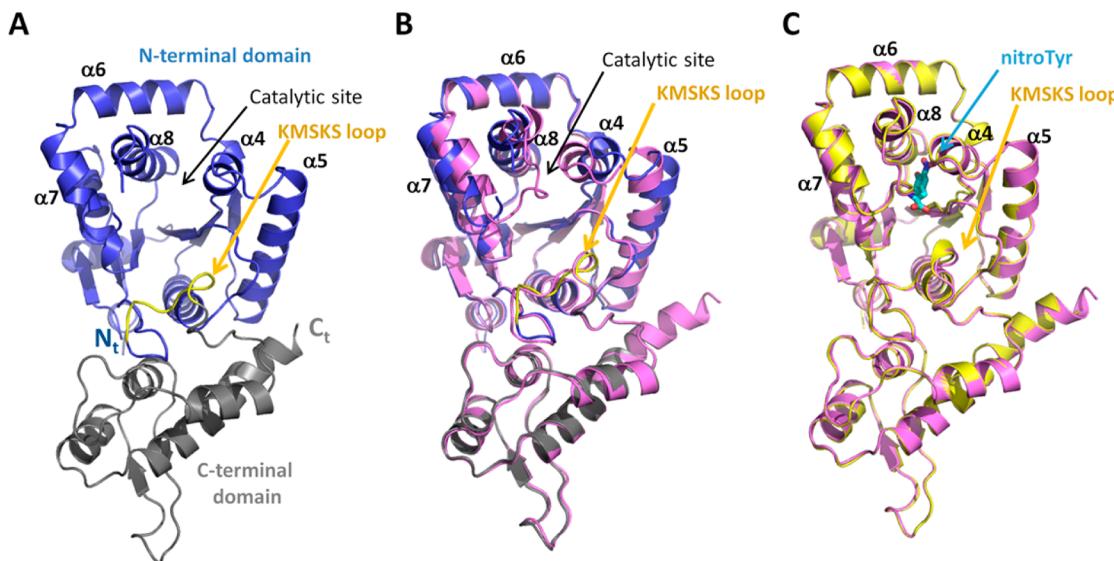
strongest density. The apo-open (2.0 Å resolution), apo-closed (2.0 Å resolution) and closed-nitroTyr structures (1.7 Å) were refined to final *R*/*R*<sub>free</sub> values of 19.9/24.9, 22.3/27.8, and 18.0/22.3%, respectively. We suspect that the ~2.5% higher *R*-factors in the apo-closed compared to those in the apo-open structure are due to it having slightly higher crystal lattice disorder as is reflected in its unit cell dimensions being between those of the apo-open and closed-nitroTyr, rather than matching closed-nitroTyr, and for this reason it may not be as well modeled by a single conformation. We also suggest that, since the electron density for Cys70 is unmodified in the apo-open form and modified in the apo-closed form (supplemental Figure S1 in SI), the Cys70-modified and unmodified proteins largely partition into the respective crystal forms (see Results and Discussion). Additional refinement statistics are reported in Table 2.

**Assessment of nitroTyr-RS Efficiency, Fidelity and Permissivity.** Each ncAA used in this study was purchased from Peptech (Burlington, MA), Bachem (Torrance, CA), or Sigma (St. Louis, MO). To measure the efficiency of nitroTyr incorporation, a *pBad* expression plasmid containing superfolder GFP (sfGFP) with an amber (TAG) codon at position 150 (*pBad*-sfGFP-150TAG) was cotransformed with *pDule* plasmids containing first- and second-generation nitroTyr-RSs (see Table 1) into DH10B cells. Five milliliters of autoinducing medium (see supplementary Table 1 in SI) containing 1 mM nitroTyr was inoculated with 50 μL of overnight culture grown in noninducing medium. Cells were incubated with shaking at 37 °C for 40 h, at which time sfGFP fluorescence of the cultures was measured in a 96-well plate (excitation: 488 nm, emission: 510 nm). The amount nitroTyr-incorporated sfGFP was determined by comparing total fluorescence of each culture with a standard

curve of purified sfGFP. No fluorescence over that of phosphate buffer saline was detected in cells grown in noninducing medium. To measure the absolute fidelity of each nitroTyr-RS, the same procedure was performed except that nitroTyr was omitted from the autoinduction media. Permissivity measurements were conducted similarly except that nitroTyr was substituted for each of the ncAAs shown in supplemental Figure 2 in SI. Mutations were introduced into the nitroTyr-5B RS gene using established overlap extension PCR protocols,<sup>23</sup> which were confirmed by sequencing (Center for Genome Research and Biocomputing, Oregon State University), and their functional effects tested using the same efficiency, fidelity, and permissivity assay described above.

## RESULTS AND DISCUSSION

**Selection Strategy Leading to the Second-Generation nitroTyr tRNA Synthetases.** Noncanonical amino acid tRNA synthetases are typically generated using multiple alternating rounds of lethal positive and negative selection and they have been used for a diverse assortment of scientific exploration in many different cell types and animals.<sup>1,3</sup> Similarly, the first-generation nitroTyr-RSs were selected from a library of mutant *Methanocaldococcus jannaschii* (*Mj*) Tyr-RSs by three rounds of alternating positive and negative selection steps<sup>9</sup> (Scheme 1). The three rounds of positive selection required amber suppression in a chloramphenicol acetyltransferase gene in *E. coli* on minimal media plates containing 1 mM nitroTyr and chloramphenicol at concentrations of 60, 60, and 80 μg/mL, respectively, while the negative selection step required suppression of an amber disrupted barnase gene in the absence of nitroTyr. After these selections, library members surviving a



**Figure 2.** Structures of nitroTyr-SB RS. (A) Cartoon representation of the apo-open conformation of the nitroTyr-SB RS depicting the N- and C-terminal domains in blue and gray, respectively, and the KMSKS loop in yellow. The amino acid substrate binding pocket is labeled along with its surrounding helices. N<sub>t</sub>: N-terminus, C<sub>t</sub>: C-terminus. (B) Overlay of the apo-open structure (same coloring as panel A) with the apo-closed structure (magenta) demonstrating the inward collapse of helix pairs  $\alpha 4/\alpha 5$  and  $\alpha 7/\alpha 8$ . (C) Overlay of the apo-closed structure (same coloring as panel B) with the nitroTyr bound structure (yellow). Carbon atoms of the nitroTyr substrate are colored cyan.

fourth positive selection at 100  $\mu\text{g}/\text{mL}$  chloramphenicol were evaluated for efficiency and fidelity. This resulted in four distinct first-generation synthetases with good efficiency and strict fidelity, among which ‘nitroTyr-8’ was the top performer (Figure 1, Table 1).<sup>9</sup>

In an attempt to improve their efficiency, we performed extensive mutagenesis on these first-generation nitroTyr-RSs; however, no significant improvements in efficiency were observed (supplemental Table 2 in SI). Alternatively, we redesigned the selection strategy (Scheme 1) to use only a single round of positive selection with reduced antibiotic concentrations (at 40  $\mu\text{g}/\text{mL}$  chloramphenicol) and a single negative selection, and all selections were performed on rich media plates rather than plates with minimal media. After the single positive and negative selection step, the remaining library was directly assessed for nitroTyr containing protein expression using the sfGFP-ncAA fluorescence assay on autoinduction media plates<sup>24</sup> as opposed to ramping up antibiotic concentrations to select for the “best” ncAA-RS. Green fluorescing colonies were selected (68), and their efficiency and fidelity in producing sfGFP under protein expression conditions were measured. For further evaluation, the top 12 nitroTyr-sfGFP producing RS hits (Table 1) were ligated into the *pDule* expression plasmid<sup>25</sup> containing its cognate amber suppressing tRNA<sub>CUA</sub>.

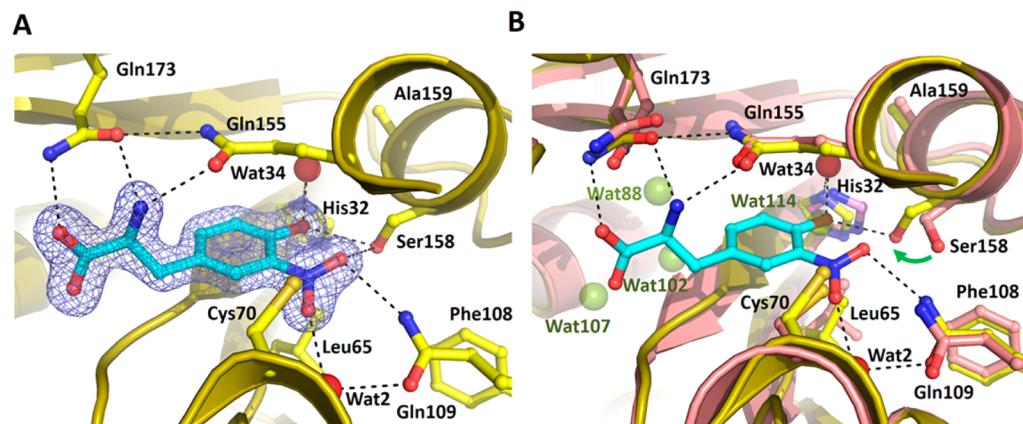
**Efficiency and Fidelity of Second-Generation nitroTyr-RSs.** The efficiency and fidelity of the 12 nitroTyr-RS variants in *pDule* were evaluated using a similar fluorescent expression reporter, but now the truncated sfGFP gene was encoded on a standard protein expression plasmid. Efficiencies for each of the 12 second-generation RSs ranged between 100 and 200 mg of nitroTyr-incorporated sfGFP per liter of culture, a substantial improvement over the 10–30 mg per liter of culture with the first-generation nitroTyr-RSs, and only ~5-fold less than native sfGFP production (Figure 1).

When assessed in the absence of nitroTyr, fidelities of the second-generation nitroTyr-RSs were about 10-fold worse than

those of the first-generation nitroTyr-RSs (Figure 1). However, when fidelities were assessed in the presence of 1 mM nitroTyr, no evidence for Tyr incorporation (or any other canonical amino acid) was observed in tryptic digests of purified sfGFP and Hsp90 proteins analyzed by ESI-Q-ToF mass spectrometry analysis and tandem mass spectrometry (MS/MS) fragmentation.<sup>10</sup> These results indicate that the relaxed fidelity of the second-generation RSs in the absence of nitroTyr compromises neither the integrity nor the utility of the new nitroTyr-RSs and is therefore of no functional consequence as fidelity only needs to be high enough to avoid incorporation of canonical amino acids when in competition with the targeted ncAA. Important to note is that rescreening did not lead to the lucky discovery of a single rare high efficiency version or even to a few variants that had incrementally higher efficiency, but in a single new selection run with reduced stringency, we obtained at least a dozen variants that were all nearly 10 times more efficient.

**The nitroTyr 5B Synthetase Is Minimally Permissive.** Permissivity is a desirable property of ncAA-RSs because relaxation of substrate specificity would eliminate the need to select or engineer a unique ncAA-RS variant for each new ncAA.<sup>26,27</sup> For instance, a permissive nitroTyr synthetase that could incorporate other related meta-substituted analogues of nitroTyr modification would facilitate work to uncover the molecular mechanisms by which the nitroTyr PTM confers altered protein functionality, and/or to probe its physiological fates. Several such ncAAs (supplemental Figure 2 in SI) were tested for their ability to be incorporated by the nitroTyr-RS SB, the most efficient of the second-generation nitroTyr-RSs; however, no incorporation into sfGFP was detected for any. Rescreening libraries or rebuilding libraries from scratch based on the nitroTyr-SB RS structure (discussed below) will therefore be necessary to expand the number of nitroTyr analogues capable of being incorporated into proteins with this *Mj* Tyr-RS/tRNA system.

**Structural Basis for nitroTyr Recognition by the Most Efficient nitroTyr-RS.** To investigate the structural features that



**Figure 3.** Active site of nitroTyr 5B synthetase. (A) The 1.7 Å structure of the substrate-bound nitroTyr-5B RS active site reveals key protein–substrate interactions. Carbon atoms of nitroTyr are colored in cyan, and protein carbon atoms are colored in yellow. Oxygen and nitrogen atoms are colored in red and blue, respectively. Hydrogen bonds are depicted with black dashed lines. Blue mesh represents  $2F_o - F_c$  omit map density contoured at  $1.5 \rho_{rms}$ . (B) Overlay of the substrate-bound active site (same coloring as panel A) with that of the apo-closed structure (carbon atoms in pink). Waters present in the apo-closed structure that are displaced by nitroTyr binding are shown as transparent green spheres. Green arrow represents the change in positioning of the Ser158 side chain when nitroTyr is bound to the active site.

account for the improved efficiency of the second-generation nitroTyr-RSs, we crystallized the most efficient of the second-generation nitroTyr-RSs, variant 5B (Figure 1 and Table 1). Fortuitously, this synthetase crystallized in three different conformations, an “open” state without substrate bound (referred to as the apo-open structure), a “closed” state without substrate bound (referred to as the apo-closed structure), and a “closed” state with nitroTyr bound in the active site (referred to as the closed-nitroTyr structure). All three crystal forms were similar to that used for solving the structure of the parent *Mj*Tyr-RS,<sup>28</sup> making the structure solution trivial. The three structures were solved at resolutions of 1.7–2.0 Å with final  $R_{free}$  values ranging from 22 to 28% (Table 2). All structures adopt the expected architecture seen for other *Mj* RS structures which includes an N-terminal Rossmann-fold catalytic domain (residues 1–202), within which the connective polypeptide 1 (CP1) region (residues 110–148) and the acceptor stem binding region (133–143) are located, a KMSKS loop (residues 202–209), which links the N- and C-terminal domains, and the C-terminal domain (residues 210–308) (Figure 2).

**The Apo-Open Structure.** Unexpectedly, multiple data sets collected on crystals grown in the absence of nitroTyr showed variations in unit cell dimensions of up to ~2%. The solved structures showed that, in crystals with unit cell dimensions  $a = b = \sim 100.0$  Å,  $c = \sim 73$  Å, the protein adopted an apo-open conformation, whereas in those with unit cell dimensions  $a = b = \sim 102$  Å,  $c = 72$  Å, it adopted a more closed conformation (described below). Screening of several crystals grown under these conditions indicated the majority adopted the apo-open structure. The best resolved apo-open conformation structure was refined at 2.0 Å (Table 2, Figure 2A), and in this structure, only residues 139–141 were without interpretable density and were not modeled. These residues are part of the tRNA acceptor stem binding loop that is similarly disordered in the apo wild-type *Mj*Tyr-RS structure (PDB 1u7d;<sup>28</sup>) but not in the apo *o*-methyl-tyrosine variant structure (PDB 1u7x;<sup>28</sup>). The KMSKS loop, which is involved in ATP-binding, is also ordered in this nitroTyr-RS structure despite the absence of ATP, just as seen in the *o*-methyl-Tyr-RS but in contrast to the apo wild-type *Mj*Tyr-RS structure. In well-ordered portions of the structure, little difference exists between this apo-open structure and the two

other known apo *Mj*RS structures (0.66 Å  $C\alpha$  RMSD over 306 residues for both wild-type and the *o*-methyl-Tyr-RSs).

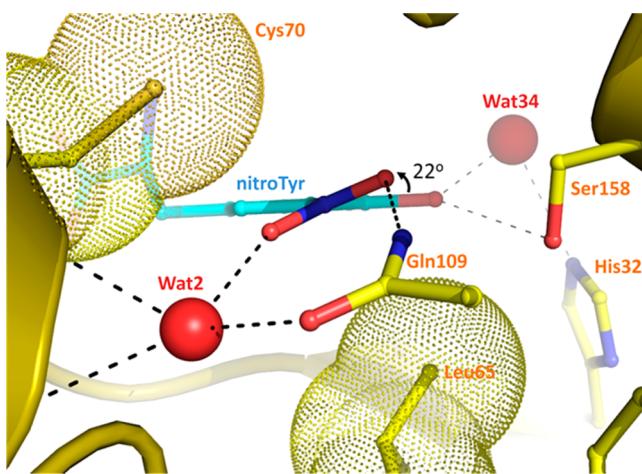
**The Apo-Closed Structure.** Crystals of the apo-closed structure grew in the same drops as for the above-described apo-open structure and were distinguishable only by measurement of their unit cells. The best diffracting crystal with the larger unit cell dimensions was refined at 2.0 Å resolution (Table 2). Although no nitroTyr is bound, this conformation approximates that observed for *Mj*Tyr-RS when substrate is bound (e.g., PDB 1zh6;<sup>29</sup>). In this conformation, helix pairs  $\alpha 4/\alpha 5$  (residues 66–88) and  $\alpha 7/\alpha 8$  (residues 127–155) collapse inwardly toward the core of the N-terminal domain to form the amino acid substrate binding pocket (Figure 2B). Residues 139–141 of the acceptor stem loop region, which help form the active site pocket, were sufficiently ordered to model. In the amino acid binding pocket electron density for four water molecules was present as well as additional density extending from the sulfur atom of Cys70 (supplemental Figure 1 in SI), which we hypothesize is a mixed disulfide with a partially disordered  $\beta$ -mercaptoethanol molecule—a reagent present during purification and crystallization. The additional hydrophobic burial associated with a modified Cys70 could be responsible for stabilizing this apo-closed conformation, and consistent with this, the crystals containing the apo-open structure appear not to have this modification (supplemental Figure 1 in SI). This would explain why this conformation has not been seen for wild-type enzyme which has a more polar His residue at position 70. It also provides a striking reminder that substrate affinity and specificity is not simply a matter of the interactions made by the bound ligand, but they depend on how much better those interactions are compared with the interactions made with solvent in the unliganded enzyme. Thus, changes in the energetic difference between the open and closed conformations of this apo enzyme could be a part of what enhances the overall kinetic properties of these second-generation nitroTyr-RSs. In any case, this is the first structure to our knowledge of an *Mj* synthetase variant in an apo-closed state, and its identification provides interesting insight into factors that influence the conformational dynamics and catalytic action of these synthetases.

**The Closed-nitroTyr Structure.** The initial diffraction data set from a crystal grown in the presence of nitroTyr showed no

nitroTyr bound with the RS adopting the apo-closed conformation described above. As for the apo-closed structure, Cys70 was modified, leading us to speculate this modification was impeding the ability of nitroTyr to bind the RS. Indeed, reduction of the crystals with dithiothreitol (DTT) in the presence of nitroTyr led to a structure, refined at 1.7 Å resolution (Figure 2C and Table 2), with unambiguous electron density for nitroTyr at the active site (Figure 3A).

Comparison of the closed-nitroTyr and the apo-closed structures reveals minimal changes occur in the active site upon substrate binding. Notable differences include the displacement of four water molecules by the nitroTyr substrate and a conformational adjustment in the side chains of Ser158, Gln173, and Leu65 (Figure 3B). The four displaced water molecules had been placed at positions roughly corresponding to those filled by the amino, carboxyl, and hydroxyl groups as well as the  $C\beta$  of nitroTyr.

The structure of the closed-nitroTyr-RS reveals how its active site recognizes nitroTyr. Although the nitroTyr-SB RS possesses five mutations compared to wild-type Tyr-RS (Tyr32His, His70Cys, Asp158Ser, Ile159Ala, and Leu162Arg, Table 1), only three of these residues are directly involved in protein–substrate interactions. Specifically, the hydroxyl group of nitroTyr is positioned by hydrogen bonds with the side chain of Ser158 and Wat34 (waters with lower numbers have stronger electron density; see Materials and Methods), which is held in place by the side chain of His32. The  $\text{NO}_2$  moiety hydrogen bonds with both Gln109 and Wat2 (Figures 3A and 4). To satisfy

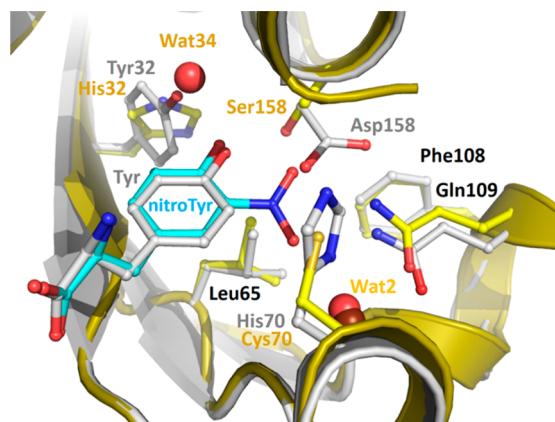


**Figure 4.** Interactions of the nitro group of nitroTyr with the active site of the nitroTyr-SB RS. Binding of nitroTyr to the nitroTyr-SB RS induces a 22° rotation of the nitro group out of the plane of the phenolic ring, which is stabilized by hydrogen bonds with Gln109 and Wat2, as well as close-packing interactions with Cys70 and Leu65. Coloring is the same as in Figure 3. Dots represent space filled by the side chains of Cys70 and Leu65.

this hydrogen-bonding network, the nitro group is rotated out of the plane of the aromatic ring by 22° (Figure 4). This out-of-plane rotation is also partly stabilized by Cys70 in which the sulfur atom lies slightly above and in close contact with the plane of the nitro group. The size and positioning of residue 70 is therefore important to accommodate the nitro group. Although the pH of the crystals (~8.0) is above the expected  $pK_a$  of nitroTyr in solution (~7.1), we suggest that the rotation of the nitro group decreases its resonance with the phenolate which,

together with the local H-bonding environment, stabilizes the protonated form of nitroTyr upon binding.

Comparisons of the nitroTyr-SB active site with that of the wild-type Tyr-RS reveal why the wild-type RS is incompatible with nitroTyr binding. Specifically, the Asp158-Gln109-His70 interaction seen in the wild type active site fills space at the meta position of the phenolic ring, preventing the binding of substrates with substituents at this position, such as nitroTyr (Figure 5).



**Figure 5.** Comparison of nitroTyr-RS active site with wild-type Tyr-RS. Overlay of the nitroTyr-SB RS amino acid substrate binding pocket (same coloring as Figure 3) with that of the wild-type *Mj*Tyr-RS (PDB 1j1u). For the wild-type Tyr structure, carbon, oxygen, and nitrogen atoms are colored in gray, red, and blue respectively. Also, the shown amide conformation of Gln109 in the wild-type structure is flipped 180° from the deposited coordinates, because the electron density and refined B-factors of that structure imply that it had been incorrectly fit. This corrected orientation of Gln109 matches that of the closed-nitroTyr structure.

This space is enlarged in the nitroTyr-SB RS to accommodate the additional nitro group by reducing the size of side chains of Glu158 and His70 in the wild-type enzyme to Ser and Cys in the nitroTyr-SB RS, respectively. In doing so, this pocket is not only enlarged compared to that of the wild-type synthetase, but a suitable hydrogen-bonding network for interaction with the nitro group is also formed (Figure 4).

**Features Distinguishing the Second- from the First-Generation nitroTyr-RSs.** To understand the key features of the second-generation nitroTyr-RSs, we looked for elements that were common to the 12 second-generation enzymes and different from the first-generation versions. One of the most conserved residues among the second-generation enzymes is Gln155, which is involved in interactions with the backbone portion of the amino acid substrate (Figure 3A). Also conserved in 11 of the 12 second-generation RSs is Gln109 (making Phe108 similarly conserved since alterations in these two residues were linked by design of the library; see Materials and Methods), which hydrogen bonds with the nitro group. Two of the four first-generation RSs (nitroTyr-8 and -10) have a different residue at one of these positions—nitroTyr-8 has a hydrophobic Met at position 155, abolishing important protein–substrate hydrogen bonds, while nitroTyr-10 has a Glu at position 109, which would interact unfavorably with the partially negative oxygen atoms of the nitro group (Figure 4). In both cases, protein–substrate interactions (and therefore ncAA incorporation efficiency) would be expected to be significantly compromised in these two first-generation RSs.

Also well conserved in the second-generation RSs is a Ser at position 158, which is observed in 11 of the 12 variants and contributes an important hydrogen bond to the hydroxyl group of nitroTyr (Figure 3A). Interestingly, three of the four first-generation RSs possess a Gly at this position, rendering them incapable of forming such a hydrogen bond. Compromising this interaction in nitroTyr-5B RS by mutation of the Ser158 to Ala reduced its ability to incorporate nitroTyr by 85% (supplemental Figure 3 in SI), demonstrating the importance of this residue in substrate recognition.

Interestingly, nitroTyr-1 of the first-generation RSs possess all three of the most conserved residues found in the second-generation RSs (Gln155, Gln109, and Ser158), yet it still displays limited ability to incorporate nitroTyr. Although the reasons for its poor nitroTyr incorporation efficiency are not immediately obvious, it is the only variant among first- and second-generation nitroTyr-RSs to possess a negatively charged side chain at position 32 (Table 1).

**Structural Basis for the Fidelity of the nitroTyr-RSs.** In terms of rationalizing the differences in fidelity between the first- and second-generation nitroTyr-RSs, we note that the nitroTyr-5B RS preserves all hydrogen-bonding interactions required for Tyr binding and has a larger binding pocket in order to accommodate the NO<sub>2</sub> moiety, suggesting that nitroTyr-5B should still be able to bind Tyr (though less favorably than nitroTyr). In the first-generation RSs, however, these interactions with Tyr would not be expected to be preserved given that most have either Gly at position 158 (which would abolish important interactions with this side chain) or compromised interactions with the amino acid substrate backbone.

These observations are consistent with the relaxed fidelity observed for the second-generation nitroTyr-RSs and the near perfect fidelity observed for the first-generation nitroTyr-RSs (when expression is conducted in the absence of nitroTyr). They are also consistent with the main differences by which the first and second-generation RSs were generated, whereby selection pressures, particularly those for maintaining fidelity, were relaxed. It seems plausible, therefore, that during the original screening process the higher efficiency second-generation nitroTyr-RS variants were lost due to selection criteria that maintained unnecessarily stringent fidelity in the absence of nitroTyr, leaving behind only nitroTyr-RSs with strict fidelity but poor efficiency. Given that in the application of ncAA technology, the expression of protein is always conducted in the presence of ncAA, the relaxed fidelity of the second-generation RSs in the absence of ncAA is of no consequence as long as fidelity is maintained when nitroTyr is present in the media at 1 mM concentration during expression. Thus, for the broad application of ncAA incorporation technology, only the “functional fidelity” of the RS, that is the fidelity measured in the presence of ncAA, is of any consequence. Yet historically, all ncAA-RS selection strategies are largely designed to maintain fairly strict “absolute fidelity”, that is fidelity in the absence of ncAA. While the correlation between enhanced efficiency and relaxed ‘absolute’ fidelity seen for the nitroTyr-RSs could be the consequence of the structural similarity between nitroTyr and Tyr, it is important to note that many ncAAs capable of being incorporated into proteins through *in vivo* recombinant expression systems are structural analogues of the amino acid substrate for the parent natural RS. These results highlight the potential value of optimizing future selection strategies to focus on enforcing ‘functional’ rather than ‘absolute’ fidelity.

**Structural Basis for the Lack of Permissivity of the Second-Generation nitroTyr-RSs.** The exact reasons for the observed lack of permissivity of nitroTyr-5B are not immediately obvious. However, its crystal structure reveals a relatively tightly packed active site, and mutagenesis analyses show that the enzyme is rather sensitive to mutations that might be expected to promote incorporation of alternative substrates (supplemental Figure 3 in SI). Also, given that the nitroTyr substrate is bound through an intricate network of hydrogen bonds with the phenolic alcohol and the nitro group (which must be rotated out of the plane of the aromatic ring), it can be understood why alternative substrates missing one or more of these groups would bind significantly more weakly and their incorporation would therefore be limited.

## CONCLUSIONS

We have described here the development of second-generation amino-acyl tRNA synthetases capable of incorporating nitroTyr with efficiencies roughly an order of magnitude greater than those previously reported. Biochemical and structural characterizations of these highly efficient second-generation nitroTyr-RSs reveal key protein–substrate interactions that form the basis for this improvement. Intriguing to ask is why were these second-generation nitroTyr-RSs not identified in the original<sup>9</sup> screening process? On the basis of the crystal structure of the nitroTyr-5B RS, we suggest this is the result of the second-generation nitroTyr-RSs preserving more of the protein–substrate interactions associated with Tyr binding—a feature consistent with the relaxed selection strategy leading to these second-generation RSs. These observations reveal the complexities associated with ncAA-RS fidelity and highlight a possible need to rethink the way in which fidelity is implemented as a selection criteria during the selection process. Because we did not do any work to try to optimize the new selection strategy, it cannot be ruled out that additional efforts to explicitly optimize screening strategies might lead to yet further major gains in efficiencies, not only for nitroTyr-RSs but also other ncAA-RSs. This is particularly relevant because, just as we experienced in our work on nitroTyr modifications, the production of sufficient quantities of pure site-specifically modified proteins via suitably efficient *in vivo* expression systems is often a major methodological roadblock to applying ncAA technologies for the study of human physiology and disease and for the development of novel probes and nanomaterials.

## ASSOCIATED CONTENT

### Supporting Information

Description of the plasmids used in the selection of nitroTyr tRNA synthetases and the process by which the nitroTyr 5B synthetase was expressed and purified. Supporting Tables 1 and 2 as well as Supporting Figures 1–3 with their associated legends. This material is available free of charge via the Internet at <http://pubs.acs.org>.

### Accession Codes

Coordinates and structure factors for the apo-open, apo-closed and nitroTyr-bound structures have been deposited in the Protein Data Bank under accession numbers 4ND6, 4ND7 and 4NDA, respectively.

## AUTHOR INFORMATION

### Corresponding Author

\*Telephone: (541) 737-4511. Fax (541) 737-0481. E-mail: ryan.mehl@oregonstate.edu.

### Funding

This work was supported by F&M Hackman and Eyler funds, NSF-MCB-0448297, HHMI Undergraduate Science Program and the services provided by the Cell Imaging and Analysis Facilities and Services Core of the Environmental Health Sciences Center, Oregon State University, Grant Number P30 ES00210, National Institute of Environmental Health Sciences, National Institutes of Health.

### Notes

The authors declare no competing financial interest.

## REFERENCES

- (1) Davis, L., and Chin, J. W. (2012) Designer proteins: applications of genetic code expansion in cell biology. *Nat. Rev. Mol. Cell. Biol.* 13, 168–182.
- (2) Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834–840.
- (3) Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* 79, 413–444.
- (4) Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of Escherichia coli. *Science* 292, 498–500.
- (5) Kobayashi, T., Nureki, O., Ishitani, R., Yaremcuk, A., Tukalo, M., Cusack, S., Sakamoto, K., and Yokoyama, S. (2003) Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat. Struct. Biol.* 10, 425–432.
- (6) Abello, N., Kerstjens, H. A., Postma, D. S., and Bischoff, R. (2009) Protein tyrosine nitration: selectivity, physicochemical and biological consequences, denitration, and proteomics methods for the identification of tyrosine-nitrated proteins. *J. Proteome Res.* 8, 3222–3238.
- (7) Aulak, K. S., Koeck, T., Crabb, J. W., and Stuehr, D. J. (2004) Proteomic method for identification of tyrosine-nitrated proteins. *Methods Mol. Biol.* 279, 151–165.
- (8) Ng, J. Y., Boelen, L., and Wong, J. W. (2013) Bioinformatics analysis reveals biophysical and evolutionary insights into the 3-nitrotyrosine post-translational modification in the human proteome. *Open Biol.* 3, 120148.
- (9) Neumann, H., Hazen, J. L., Weinstein, J., Mehl, R. A., and Chin, J. W. (2008) Genetically encoding protein oxidative damage. *J. Am. Chem. Soc.* 130, 4028–4033.
- (10) Franco, M. C., Ye, Y., Refakis, C. A., Feldman, J. L., Stokes, A. L., Basso, M., Melero Fernandez de Mera, R. M., Sparrow, N. A., Calingasan, N. Y., Kiaei, M., Rhoads, T. W., Ma, T. C., Grumet, M., Barnes, S., Beal, M. F., Beckman, J. S., Mehl, R., and Estevez, A. G. (2013) Nitration of Hsp90 induces cell death. *Proc. Natl. Acad. Sci. U.S.A.* 110, E1102–E1111.
- (11) Wang, Y. S., Fang, X., Wallace, A. L., Wu, B., and Liu, W. R. (2012) A rationally designed pyrrolysyl-tRNA synthetase mutant with a broad substrate spectrum. *J. Am. Chem. Soc.* 134, 2950–2953.
- (12) Kiga, D., Sakamoto, K., Kodama, K., Kigawa, T., Matsuda, T., Yabuki, T., Shirouzu, M., Harada, Y., Nakayama, H., Takio, K., Hasegawa, Y., Endo, Y., Hirao, I., and Yokoyama, S. (2002) An engineered Escherichia coli tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9715–9720.
- (13) Oki, K., Sakamoto, K., Kobayashi, T., Sasaki, H. M., and Yokoyama, S. (2008) Transplantation of a tyrosine editing domain into a tyrosyl-tRNA synthetase variant enhances its specificity for a tyrosine analog. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13298–13303.
- (14) Iraha, F., Oki, K., Kobayashi, T., Ohno, S., Yokogawa, T., Nishikawa, K., Yokoyama, S., and Sakamoto, K. (2010) Functional replacement of the endogenous tyrosyl-tRNA synthetase-tRNATyr pair by the archaeal tyrosine pair in Escherichia coli for genetic code expansion. *Nucleic Acids Res.* 38, 3682–3691.
- (15) Park, H. S., Hohn, M. J., Umehara, T., Guo, L. T., Osborne, E. M., Benner, J., Noren, C. J., Rinehart, J., and Soll, D. (2011) Expanding the genetic code of Escherichia coli with phosphoserine. *Science* 333, 1151–1154.
- (16) Heinemann, I. U., Rovner, A. J., Aerni, H. R., Rogulina, S., Cheng, L., Olds, W., Fischer, J. T., Soll, D., Isaacs, F. J., and Rinehart, J. (2012) Enhanced phosphoserine insertion during Escherichia coli protein synthesis via partial UAG codon reassignment and release factor 1 deletion. *FEBS Lett.* 586, 3716–3722.
- (17) Turner, J. M., Graziano, J., Spraggon, G., and Schultz, P. G. (2006) Structural plasticity of an aminoacyl-tRNA synthetase active site. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6483–6488.
- (18) Tronrud, D. E. (1997) TNT refinement package. *Methods Enzymol.* 277, 306–319.
- (19) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* 60, 2126–2132.
- (20) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* 66, 213–221.
- (21) Painter, J., and Merritt, E. A. (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D* 62, 439–450.
- (22) Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* 35, W375–383.
- (23) Heckman, K. L., and Pease, L. R. (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protoc.* 2, 924–932.
- (24) Seitchik, J. L., Peeler, J. C., Taylor, M. T., Blackman, M. L., Rhoads, T. W., Cooley, R. B., Refakis, C., Fox, J. M., and Mehl, R. A. (2012) Genetically encoded tetrazine amino acid directs rapid site-specific in vivo bioorthogonal ligation with trans-cyclooctenes. *J. Am. Chem. Soc.* 134, 2898–2901.
- (25) Farrell, I. S., Toroney, R., Hazen, J. L., Mehl, R. A., and Chin, J. W. (2005) Photo-cross-linking interacting proteins with a genetically encoded benzophenone. *Nat. Methods* 2, 377–384.
- (26) Miyake-Stoner, S. J., Refakis, C. A., Hammill, J. T., Lusic, H., Hazen, J. L., Deiters, A., and Mehl, R. A. (2010) Generating permissive site-specific unnatural aminoacyl-tRNA synthetases. *Biochemistry* 49, 1667–1677.
- (27) Stokes, A. L., Miyake-Stoner, S. J., Peeler, J. C., Nguyen, D. P., Hammer, R. P., and Mehl, R. A. (2009) Enhancing the utility of unnatural amino acid synthetases by manipulating broad substrate specificity. *Mol. Biosyst.* 5, 1032–1038.
- (28) Zhang, Y., Wang, L., Schultz, P. G., and Wilson, I. A. (2005) Crystal structures of apo wild-type *M. jannaschii* tyrosyl-tRNA synthetase (TyrRS) and an engineered TyrRS specific for o-methyl-L-tyrosine. *Protein Sci.* 14, 1340–1349.
- (29) Turner, J. M., Graziano, J., Spraggon, G., and Schultz, P. G. (2005) Structural characterization of a p-acetylphenylalanyl aminoacyl-tRNA synthetase. *J. Am. Chem. Soc.* 127, 14976–14977.
- (30) Diederichs, K., and Karplus, P. A. (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat. Struct. Biol.* 4, 269–275.
- (31) Karplus, P. A., and Diederichs, K. (2012) Linking crystallographic model and data quality. *Science* 336, 1030–1033.