Translational incorporation of modified phenylalanines and tyrosines during cell-free protein synthesis†

Zhongqiang Wangabc and Hayden Matthewsabd

Inherent promiscuity of bacterial translation is demonstrated by mass spectrometric quantification of the translational incorporation of ring-substituted phenylalanines and tyrosines bearing fluoro-, hydroxyl-, methyl-, chloro- and nitro-groups in an E. coli-derived cell-free system. Competitive studies using the cell-free system show that the aminoacyl-tRNA synthetases (aaRS) have at least two orders of magnitude higher specificity for the native substrate over these structural analogues, which correlates with studies on the purified synthetase.

Introduction

Protein synthesis is a strictly regulated biosynthetic pathway. The remarkable substrate specificity exhibited by aminoacyl-tRNA synthetases (aaRS), a family of enzymes that catalyze the aminoacylation of amino acids with the cognate tRNA, is important for the high degree of accuracy. However, examples of non-canonical amino acids mistakenly incorporated into proteins via a natural biosynthetic pathway have been reported, indicating that there is a limit to aaRS with regard to distinguishing substrates with structures slightly varied from the canonical amino acid. Here, studies probing the substrate range of wild-type E. coli phenylalanyl-tRNA synthetase (PheRS) and tyrosyl-tRNA synthetase (TyrRS) through the expression of E. coli-derived cell-free system show that the aminoacyl-tRNA synthetases (aaRS) have at least two orders of magnitude higher specificity for the native substrate over these structural analogues, which correlates with studies on the purified synthetase.

Some non-canonical amino acids display quite diverse physiological and pharmacological activity, whereas others are involved in many regards, such as the use for biological markers, altering protein properties and the production of new biological materials. Further, some hydroxylated derivatives are found to be the most abundant modified amino acids bound to proteins that are commonly detected in pathological tissues, and their misincorporation has been associated with a wide variety of pathological conditions such as aging, cataractogenesis, myocardial ischaemia and reperfusion and neurodegenerative disorders such as Parkinson’s diseases.

The aminoacylation reaction catalyzed by aaRSs is the first step in protein synthesis and is accomplished via two stages. The amino acid first uses adenosine triphosphate (ATP) to form an aminoacyl-adenylate along with the release of pyrophosphate, followed by transfer of the amino acid to the terminal adenosine of the cognate tRNA. This has been identified as a rate determining step for both phenylalanyl-tRNA synthetase and tyrosyl-tRNA synthetase, with initial rates of transfer of the amino acid from aminoacyladenylate to cognate tRNA similar to the overall turnover number for the steady state aminoacylation of tRNA. The aminoacylation is a key point in translational quality control as it provides the link between an amino acid, the tRNA anticodon and mRNA codon to enable accurate translation. Under physiological conditions, the estimated average error rate of amino acid misaminoacylation by aaRSs is 10⁻⁴. For some aaRSs, the accurate amino acid selection is accomplished mainly at their active sites, discriminating non-cognate amino acids on steric grounds. In other cases, there is an additional proofreading mechanism assisting the selection, which occurs at a site referred to as an “editing site”. The latter is known as “double-sieve amino acid selection”. While protein synthesis using non-canonical amino acids is an indirect measure of the aaRS activity, work presented here was aimed to investigate the enzyme in a more direct manner, by...
evaluating the interactions between the non-canonical amino acids and the aaRS through an aminoacylation reaction.

Cell-free expression systems use a cell extract, which is partially purified and supplemented with the target DNA, amino acids and other ingredients, to enable *in vitro* protein synthesis in a controlled manner. The cell-free systems allow substitution of a standard amino acid with an unusual one, therefore can bias the competition in favor of the unusual amino acids. By this technique, incorporation of chloro,\(^{15c}\) dehydro\(^{15c}\) and hydroxyl\(^{17a}\) amino acids has been reported. Here, an *E. coli*- derived cell-free system was used to study the modified phenylalanines 1-14 and tyrosines 15-19 (Fig. 1). In these amino acid analogue systems the volume of the substituent on the phenyl ring of non-canonical amino acids is varying, and so is the hydrophobicity. We investigated how the size and hydrophobicity of substituent affect the incorporation of non-canonical amino acids. In earlier reports, translational incorporation of 5a-e into proteins as substitutes for Phe,\(^{16}\) and 15a, 16a, 17a and 19a for Tyr\(^{17}\) have been described. However, the incorporation levels of many of the amino acids and the extent to which they compete with the natural substrate have not been determined.

**Results and discussion**

*E. coli* cell-free protein synthesis with Phe analogues

It is known that PheRS is highly selective for Phe over Tyr.\(^{1}\) However, Tyr isomers m-Tyr (1b) and o-Tyr (1c), and well as 3,4-dihydroxyphenylalanine (3,4-DOPA, 16a) with two OH groups on the ring of the substrate, were reportedly activated\(^{18}\) and incorporated\(^{18d-f}\) into proteins by PheRSs from various organisms. These suggest that, although PheRS has evolved to distinguish between Phe and Tyr primarily through recognition of the OH group at para position, it imparts far lower selectivity against the OH at other positions on the substrate ring. *E. coli* peptidyl-Pro cis-trans isomerase B, which has been extensively studied,\(^{19}\) with a hexahistidine tag attached at the N-terminus (His\(_6\)-PpiB) was expressed as a test protein. It contains 12 and 3 sites for Phe and Tyr, respectively. SDS-PAGE analysis of the synthesized His\(_6\)-PpiB shows that protein was produced when all components required for protein synthesis were supplied (Fig. 2, with DNA and Phe). Mass spectral analysis (Fig. 2d) confirmed that it is wild-type protein, and corresponds to the pair of the predominant peaks. The presence of two signals in mass spectra is due to incomplete deformylation of PpiB by deformylase in the *E. coli* S30 extract.\(^{17a}\) By comparison, only a trace of protein was produced when Phe was excluded from the reaction mixture (Fig. 2, with DNA but no Phe). Mass spectral analysis (data not shown) showed that it is wild-type His\(_6\)-PpiB, suggesting that the production of background protein was due to the presence of traces of Phe during cell-free reaction. This experiment also confirms that Tyr is not incorporated into His\(_6\)-PpiB in place of Phe, as the reaction mixture contained a typical amount of Tyr (1.0 mM) for cell-free synthesis. No protein band corresponding to His\(_6\)-PpiB was observed when the DNA plasmid was not added.

Protein expression was efficient when Phe was substituted with 1b,c (Fig. 2a, lanes B and C), reaching almost wild-type levels. Interestingly, the expressed proteins precipitated during purification (Fig. 2b, lanes B and C). This hinted at the possible effects the non-standard amino acid can have on protein structural stability, in particular given the abundance in the protein (12 Phe sites). To ensure that the tested amino acids were indeed incorporated into the protein, the small amounts of isolated protein were analyzed by mass spectrometry (Fig. 2e). Given the presence of peaks mainly corresponding to His\(_6\)-PpiB, we determined that the tested amino acids were incorporated into the protein.

![Fig. 1 Amino acids used in this study.](image-url)
substitution of 1c for Phe was concluded to be greater than 95%, whereas a 90% was given for 1b (ESI Fig. S1†). This incorporation level was assessed semiquantitatively, by following a method described in an earlier report,17 through a comparison of the height of the peaks corresponding to 12, 11 and 10 substitutions in the mass spectrum (Fig. 2e). The amount of His6-PpiB produced with 16a (Fig. 2a and b, lane A) is comparable to that produced in the absence of Phe, suggesting that 16a does not incorporate instead of Phe. The mass spectra of this protein (data not supplied) showed wild-type His6-PpiB. 

Subsequently, Phe derivatives containing Me, Cl and NO2 substituents including mono-substituted 2a–c, 3a–c, 4a–c and di-substituted 8–10, as well as para-substituted derivatives 11–14, were studied in an analogous manner. Of these, protein was only produced with the meta isoforms, including mono-derivatized 2b and 3b (Fig. 2c, lanes H and K), and di-derivatized 9 and 10 (Fig. 2a and b, lanes E and F), suggesting greater flexibility around that position at the synthetase active site. Interestingly, the proteins synthesized with 2b and 3b remained soluble and were successfully isolated without apparent loss of yield, whereas in the cases of 9 and 10, white precipitates were observed after the completion of reaction. Consequently, samples of 9 and 10 for mass spectral analysis were prepared with two methods, by chromatographic purification of the supernatant, which resulted in neither protein bands in SDS-PAGE nor signals in the expecting region in mass spectra, and by solubilizing the insoluble part with buffer containing a high concentration of urea, which showed a high ratio of signal-to-noise in the mass spectra (ESI Fig. S8 and S9†). This suggests that the insoluble part was mainly His6-PpiB containing 9 and 10. The expression was then repeated with derivatives bearing one or more fluorines on the aromatic ring, of which all but the pentafluorophenylalanine (7) were incorporated into the expressed His6-PpiB (Fig. 2c, lanes P–S). In all successful cases soluble proteins were obtained and the incorporation of intact amino acid was verified by mass spectrometry (ESI Fig. S4–S7†). Incorporation of the fluorides 5a–c as substitutes for Phe using whole cell-based methods has been reported previously.16 The results obtained here are consistent with previous studies and confirm that the incorporation will occur in a cell-free system.

The reported crystal structure of the E. coli PheRS–Phe complex (Fig. 3a) indicates that the active site of E. coli PheRS is a mainly hydrophobic pocket with stringent steric requirements for the phenyl ring.27 The Me group of alanine at position 294 (Ala294) provides a steric constraint at the para position, which has been proposed as the primary factor assuring the high degree of selection of Phe over Tyr (Fig. 3a1). Previous mutagenesis studies have shown that substituting alanine with glycine (Ala294Gly) expands the volume of the active site and thus relaxes the substrate specificity of E. coli PheRS. As a consequence, Tyr and derivatives of Phe with para-substituents larger than an OH group, including Cl, Br, I, and ethynyl-, cyano- and azido-groups, are accepted and activated in
aminoacylation reaction. In comparison, substitution with serine (Ala294Ser) decreases the volume and leads to a narrower substrate range for the engineered enzyme, rejecting even 5a, 28 The results of the present work correlate with these previous studies and support the role of synthetic pocket size as a determinant of specificity. The incorporation of 5a, and the lack of incorporation of the para-substituted derivatives 1a, 2a, 3a, 4a, 11–14, indicate slight flexibility at the active site of the wild-type E. coli PheRS near the para position, being able to accommodate only an F group. The requirements for the ortho and meta positions, on the other hand, are not so rigid and the synthetase has been shown to accept larger ring substituents. On one side of the ring the positions are observed projecting toward a small hydrophobic space (Fig. 3a2), surrounded by the side chains of Phe248 and Phe250 arranged in a “edge-to-face” manner for each aromatic pair, whereas the corresponding positions on the opposite side are enclosed with the side chains of Ser171, Gln174, Gln208, Glu210 and Gly296, pointing to a hydrophilic area (Fig. 3a). Incorporation of 5b, c in place of Phe suggests acceptance of an F at each position, which might have been expected given the flexibility of the para position. In addition to an F, an OH was also accepted at the ortho position, as suggested by the results with 1c, and an OH, Me and Cl at the meta position by the results with 1b, 2b and 3b. Further evidence by di-substituted amino acids 6, 9 and 10 reveals difference of the meta positions on the two ring sides, with one being able to accept a group as large as an Me and the other an OH.

To assist the selection for Phe, there is also an editing site located in the B3/4 domain of the β-subunit, about 35 Å from the synthetic active site, providing a post-transfer proofreading mechanism through hydrolysis of the tRNA\textsuperscript{Phe} mistakenly charged with non-cognate amino acid. However, this proof-reading activity appears to require a phenol OH for hydrogen bond formation with the side chain of Glu334 and the main chain amide of Gly315. Previous reports have shown that E. coli PheRS hydrolyzes the tRNA\textsuperscript{Phe} misacylated with Tyr\textsuperscript{1b} and 1b, 22 but does not cleave the tRNA\textsuperscript{Phe} charged with para-substituted (fluoro-, chloro-, bromo- and amino-) Phe derivatives, further evidence comes from incorporation of 5a via wild-type machinery, 16, 28 as discussed above, and of p-Cl-Phe (3a) using a Ala294GlyPheRS mutant-containing E. coli strain. Consistent with these, results of the current work support the requirement for hydrogen bond interactions and showed translational incorporation of a series of amino acid derivatives. Of particular interest was a mutant with significantly reduced editing activity, by which formation of the m-Tyr-tRNA\textsuperscript{Phe} was successful in aminoacylation. This indicates that 1b is a substrate of both the active and editing sites of E. coli PheRS, but the rate of hydrolysis is far lower than the rate of activation. This may account for the incorporation of 1b, and that of the 1c, 9 and 10 with at least one OH group on the substrate ring.

**E. coli cell-free protein synthesis with Tyr analogues**

Analogously, the substrate scope of E. coli TyrRS was studied through protein expression with Tyr analogues. To assess the requirement for the ring location of the OH, Tyr isomers 1b, c were studied. To assess the importance of the p-OH group, derivatives bearing a range of para-substituents, including 2a, 3a, 4a, 5a, 11–14 were studied. None of these amino acids were able to support protein expression (Fig. 4a).

Having established that p-OH is not replaceable for substrate recognition, Tyr derivatives with substituents at other ring positions, including both ortho and meta positions, were investigated. Derivatives studied involve the fluorides 15a, b, and hydroxyl-, chloro-, methyl- and nitro-substituted 16a, b, 17a, b, 18a, b and 19a, b. Apart from 19a, soluble proteins were obtained and successfully isolated with all derivatives (Fig. 4b), even with 16a, b that bear two OH groups. Compared to protein
loss observed for 1b,c, successful isolation with 16a,b is prob-
ably attributable to a low level of abundance in the protein (3 Tyr sites). With a similar cell-free system, incorporation of 19a in place of Tyr was reported. It was accomplished by consid-
erably enhancing the concentrations of E. coli TyrRS, which was prepared separately and supplemented to the system exoge-
nously, and the substrate 19a. The necessity for this suggests a poor activation by E. coli TyrRS, which is most probably the cause of the lack of incorporation in the present study.

The reported crystal structure of the E. coli TyrRS–tyrosyl-
adenylate complex (Fig. 3b) shows that the enzyme forms hydrogen bonds with the p-OH group of Tyr by the side chains of Tyr37 and Asp182 (Fig. 3b). The p-OH group is a key element for the substrate specificity of E. coli TyrRS as it offers comple-
mentary hydrogen bonds for the recognition of the amino acid, that of the twenty normal amino acids only Tyr can form. Studies of the Tyr isomers and para-substituted derivatives show that this OH group must locate at the para position and it is not replaceable. In analogy to the E. coli PheRS active site, spaces near ring positions are observed for E. coli TyrRS. The ortho and meta positions on one side of the substrate ring, enclosed by the side chains of Tyr175, Thr76 and Gln179, project to a small hydrophilic space (Fig. 3b), whereas the corresponding positions on the other side point to an open area (Fig. 3b), in particular the ortho position. Consistent with earlier findings results obtained with 15a, 16a and 17a indicate the space near the meta position at the E. coli TyrRS active site can accommodate an F, OH or Cl. In addition, this space has been shown to accommodate a Me (18a). The ortho position displayed the most flexibility, being able to accept not only an F, OH, Cl and Me but also a NO2.

**Competition between the cognate and derivatized amino acids**

A major advantage of the cell-free expression system over whole cell-based techniques is to precisely control the concentration of substrate. This allows to determine the extent to which the non-canonical amino acids compete with the natural counter-
part. Competitive experiments between Phe and Tyr and their related derivatives were therefore performed. Although many amino acid analogues were found to be efficient substrates of E. coli PheRS or TyrRS, the natural amino acids were highly fav-
oured against all analogues investigated (Table 1). Inclusion of a small amount of Phe in the reaction mixture precluded the incorporation of most Phe analogues. The inefficient competi-
tion of only 1c (2.1%) and the fluorides 5a–c (3.6%, 2.1% and 2.1%...
Table 1  Summary of incorporation levels and competition efficiencies of Phe and Tyr derivatives, and kinetics of E. coli TyrRS

| Derivative (2.0 mM) | Incorporation levela (%) | | | | |
|---|---|---|---|---|---|
| | No Phe | 0.2 mM Phe | Competition efficiency (%) | \(k_{cat}\)b (s⁻¹) | \(K_M\) (mM) | \(k_{cat}/K_M\) (s⁻¹ M⁻¹) |
| Phe | — | — | — | — | 0.05(ref. 20) | — |
| 1b | 90 | N.D. | N.D. | 2.1 (ref. 22) | 0.25 (ref. 22) | 0.8 x 10⁴ |
| 1c | >95 | 21 ± 0 | 2.1 | — | — | — |
| 2b | >95 | N.D. | N.D. | — | — | — |
| 3b | >95 | N.D. | N.D. | — | — | — |
| 5a | >95 | 36 ± 0.9 | 3.6 | — | 0.56 (ref. 20) | — |
| 5b | >95 | 21 ± 1.7 | 2.1 | — | 1.4 (ref. 20) | — |
| 5c | >95 | 14 ± 2 | 1.4 | — | 5.0 (ref. 20) | — |
| 6 | >95 | N.D. | N.D. | — | — | — |
| 9 | >95 | N.D. | N.D. | — | — | — |
| 10 | >95 | N.D. | N.D. | — | — | — |

| Derivative (2.0 mM) | Incorporation levela (%) | | | | |
|---|---|---|---|---|---|
| | No Tyr | 0.2 mM Ty | Competition efficiency (%) | \(k_{cat}\)b (s⁻¹) | \(K_M\) (mM) | \(k_{cat}/K_M\) (s⁻¹ M⁻¹) |
| 1a (Tyr) | — | — | — | 11.3 ± 0.15 | 0.006 (ref. 21) | 1.85 x 10⁶ |
| 15a | >95 | 29 ± 0.6 | 2.9 | 6.5 ± 0.14 | 0.13 (ref. 21) | 5.02 x 10⁴ |
| 15b | >95 | 23 ± 1.2 | 2.3 | 4.8 ± 0.12 | — | — |
| 16a | >95 | N.D. | N.D. | 3.3 ± 0.01 | 1.4 (ref. 21) | 2.33 x 10³ |
| 16b | >95 | N.D. | N.D. | 0.89 ± 0.01 | — | — |
| 17a | >95 | N.D. | N.D. | 0.34 ± 0.07 | — | — |
| 17b | >95 | N.D. | N.D. | 0.17 ± 0.08 | — | — |
| 18a | >90 | N.D. | N.D. | 0.09 ± 0.07 | — | — |
| 18b | >95 | N.D. | N.D. | 0.13 ± 0.06 | — | — |
| 19a | <10 | N.D. | N.D. | N.D. | — | — |
| 19b | >80 | N.D. | N.D. | 0.06 ± 0.01 | — | — |

a Incorporation levels and rates of reaction determined in duplicate experiments varied by less than 20%. Data are the mean of duplicate experiments. b N.D. = not determined.

1.4%, respectively, Table 1) shows that the selectivity for Phe is at least two orders of magnitude higher than for all surrogates. The \(K_M\) value of E. coli PheRS for Phe (0.05 mM) is reported to be approximately 100-, 30- and 10-fold lower than that for 5c (5.0 mM), 5b (1.4 mM) and 5a (0.56 mM), respectively (Table 1). These ratios correlate with the competition efficiencies obtained. In an analogous manner, preclusion of most Tyr analogues was also observed with a small amount of Tyr in the reaction mixture. The inefficient competition of 15a,b (2.9% and 2.3%, respectively, Table 1) suggests that the selectivity for Tyr is at least two orders of magnitude higher. The ratio of the \(K_M\) value of E. coli TyrRS for Tyr (0.0061 mM) relative to that of 15a (0.13 mM) (4.7%) also correlates with the competition efficiencies obtained for 15a.

Monitoring activity of E. coli TyrRS

Apart from the protein expression experiments, the interactions of the tested analogues with the E. coli TyrRS were assessed directly through adenosine monophosphate (AMP) production assays. Of the two synthetases indirectly investigated by protein expression, E. coli TyrRS was selected as it only has one active site. From Table 1 it is evident that E. coli TyrRS catalyzes the reactions of Tyr analogues at different rates, suggesting varying degrees of binding of the synthetase toward each analogue.
as efficiently as Tyr by *E. coli* TyrRS. This ratio correlates with the 2.9% competition efficiency observed in competitive experiment (Table 1), indicating that competing for incorporation is essentially a competition for the activation by *E. coli* TyrRS. These show a direct relationship between the ability of the synthetase to bind different substrates and catalyze the aminoacylation reactions at varying rates with the incorporation of non-canonical amino acids.

Incorporation of hydroxylated amino acids during protein synthesis has been associated with deposition and accumulation of oxidized proteins, as in biological systems the aromatic amino acids Phe and Tyr are prone to modification by cellular oxidants such as reactive oxygen species to give the corresponding products 1b,c and 16a. Elevated levels of the three amino acids have been found bound to oxidatively damaged proteins that are commonly detected in pathological tissues. Although the involvement of aaRSs in this process has been hypothesized, the specific pathway by which oxidized amino acids are incorporated remains poorly understood. While both whole-cell and cell-free as well as aminoacylation studies have shown that 1b,c could be incorporated via PheRS and 16a via TyrRS, the incorporation may occur by the other way around, with 16a via PheRS and 1b,c via TyrRS, since activation of 16a by both bacterial and eukaryotic PheRSs has been observed, as mentioned above and, as isomers of Tyr, 1b,c could in principle replace Tyr. In this work, we have clarified that, in an *E. coli* system, 1b,c can only be incorporated at the Phe location, not at the Tyr location, and that 16a would only be incorporated at the Tyr location (Fig. 5). This may aid the study of age-related pathologies associated with oxidized protein accumulation.

**Conclusions**

The work described here shows inherent substrate promiscuity of the *E. coli* PheRS and TyrRS. They can activate a range of structural analogues and subsequently charge them onto the cognate tRNA for protein expression. To the best of our knowledge, it is the first time incorporation of the amino acid analogues 2b, 3b, 6, 9, 10, and 15b, 16b, 17b, 18, and 19b into protein via wild-type translational machinery has been observed.

The inherent substrate promiscuity of aaRS may prove useful as new tools for biotechnology, as it allows the production of proteins engineered with nonproteinogenic amino acids in a direct manner, without the need for aaRSs with relaxed substrate specificity. Unlike the case of expanded genetic codes, this system does not accommodate an additional noncanonical amino acid, instead replacing a Phe or Tyr partially or completely. There is no site specificity with regard to partial replacement and hence a plasmid cannot be generated to direct modifications in a predetermined position of the protein it

---

**Fig. 5** Proposed oxidized protein synthetic pathways by incorporation of oxidized amino acid via *E. coli* PheRS and TyrRS. Intracellular oxidation of Phe and Tyr produces noncanonical hydroxylated derivatives that are substrates of *E. coli* PheRS and TyrRS. 3,4-DOPA (16a) is a substrate for both PheRS and TyrRS, while ρ- and m-tyrosine (1b,c) are both substrates of PheRS. The proofreading domain of PheRS is likely to prevent incorporation of 16a into protein in place of Phe, however misincorporation of 16a in place of Tyr may occur along with misincorporation of 1b or 1c in place of Phe.
encodes. In this work, the incorporation levels of most amino acids are greater than 95% and thus enabled a high degree of amino acid modification, including fluorination, oxidation, methylation, chlorination and nitration, which can not only be used for applications such as isotopic labeling for use in spectroscopic studies, but also for investigation of general rather than specific effects of amino acid derivatives on protein structure and function. In addition, this study also identifies the oxidized protein synthetic pathways by incorporation of oxidized amino acids involving aaRSs. Further, the substrate promiscuity of bacterial PhεRS and TyrRS may be different from that of eukaryotic enzymes, making these two synthetases potential targets for the development of antibacterial agents.

Experimental

Syntheses

Most of the amino acid derivatives used in this study were commercially available, except for 8-10, 15b, 16b, 17b, 18 and 19b, which were synthesized in the laboratory. Apart from 3-hydroxy-5-methylphenylalanine (10), the preparation of the rest of the derivatives was described previously.

Diethyl 2-acetamido-2-(3-methoxy-5-methylbenzyl)malonate.

To a solution of metallic sodium (0.235 g, 0.01 mol) in dry ethanol (20.0 mL), diethyl acetaminomalonate (2.17 g, 0.01 mol) was added. The resultant mixture was stirred at room temperature for 30 minutes, mixed with 3-bromomethyl-5-methylanisole (2.0 g, 0.0093 mol), heated under reflux for two hours, concentrated, extracted with ethyl acetate (3 x 5 mL). The combined extracts were washed with saturated brine, dried (sodium sulfate), filtered and concentrated to give diethyl 2-acetamido-2-(3-methoxy-5-methylbenzyl)malonate as a white solid (2.42 g, 74% yield). Mp 109.0–110.1 °C; 1H NMR (400 MHz, CDCl3): δ = 6.75–7.69 (m, 1H; Ar H), 6.67–6.69 (m, 1H; Ar H), 6.53 (s, 1H; Ar H), 4.25 (q, J = 7.2 Hz, 2H; (CH2CH3)2), 3.77 (s, 3H; OCH3), 3.53 (s, 2H; Ar-CH2), 2.18 (s, 3H; OCH3), 2.01 (s, 3H; (C==O)CH3), 1.28 (t, J = 7.2 Hz, 6H; (CH2CH3)2). 13C NMR (100 MHz, CDCl3): δ = 169.03 (NHC==O), 167.69 (2C, (C==O)OEt), 156.97 (Ar C), 132.26 (Ar C), 128.07 (Ar C), 126.29 (Ar C), 109.74 (Ar C), 67.43 (quaternary C), 62.60 (2C, CH2), 55.27 (OCH3), 37.03 (Ar-CH2), 23.09 ((C==O)CH3), 16.25 (Ar-CH3), 14.10 (2C, CH2CH3). HRMS (ESI) calcd for C10H13NO3: C 63.14, H 6.71, N 7.17; found: C 63.13, H 6.71, N 7.17. 

(R,S)-3-Hydroxy-5-methylphenylalanine (10).

A solution of diethyl 2-acetamido-2-(3,5-dimethoxybenzyl)malonate (2.0 g, 0.0057 mol) in 62% aqueous hydrobromic acid (11.5 mL) and acetic acid (7.0 mL) was heated in a sealed tube at 150 °C for two hours. The mixture was diluted with water (10.0 mL), decolorized with charcoal and concentrated, affording a residue which was recrystallized from water and ethanol to give (R,S)-3-hydroxy-5-methylphenylalanine as a white solid (0.89 g, 80% yield). Mp 250.0–253.8 °C. 1H NMR (400 MHz, D2O): δ = 6.36 (s, 1H; Ar H), 6.31–6.32 (m, 2H; Ar H), 3.44–3.46 (m, 1H; CH); 2.86–2.90 (m, 1H; CHHCH), 2.56–2.62 (m, 1H; CHHCH), 2.19 (s, 3H; CH3). 13C NMR (100 MHz, D2O): δ = 171.21 (C==O), 153.17 (Ar-C), 131.93 (Ar-C), 127.87 (Ar-C), 125.60 (Ar-C), 125.51 (Ar-C), 115.47 (Ar-C), 53.99 (CH3), 34.60 (CH3), 15.01 (CH3). HRMS (ESI) calcd for C10H13NO3Na [M + Na]+ m/z 321.0793; found 321.0793. Anal. calcd for C10H13NO3; C 61.53, H 6.71, N 7.17; found: C 61.87, H 6.96, N 7.26.

Construction of plasmids

The gene in pND706 vector encoding the E. coli Hisε-TyrRS was kindly provided by the Dixon group at the University of Wollongong (Australia). The genes in pND1098 vector encoding E. coli Hisε-PplIβ and in pND706 vector encoding the E. coli Hisε-TyrRS were transformed into the respective E. coli DH5α and E. coli AN1459 for amplification and were isolated using Qiagen® Mini kit following procedures recommended by the manufacturer. The gene encoding the E. coli TyrRS was then cloned in pETMCSIII vector within the Ndel-HindIII sites to generate the plasmid pET-TyrRS-His. The sequence of pET-TyrRS-His was confirmed using ABI 3730 Genetic Analyzer. The concentration of the DNA was measured using a NanoDrop spectrophotometer.

Preparation of E. coli S30 extract

The E. coli S30 extract was prepared from E. coli star BL21(DE3) by following a protocol established earlier.15,17 The natural amino acids were used at 1.0 mM. The unnatural amino acids were used at 2.0 mM and the corresponding natural amino acid was not added. The reactions were proceeded for 6 h at 30 °C, with shaking at 200 rpm, unless otherwise stated. The competitive experiments with unnatural amino acids and their corresponding natural counterparts were conducted in a system identical to that used for usual cell-free protein synthesis, with the cognate amino acid supplied to the expression at 0.2 mM concentration to compete with the foreign substrate at 2 mM. In cases where racemic mixtures were used, this concentration refers to the L-enantiomer content. The levels of unnatural amino acid incorporation in competitive experiments were concluded from the mass spectra. The competition efficiencies are given as percentages and were calculated based on the ratio of the concentrations between the natural amino acid and its corresponding competitor, and the level of unnatural amino acid incorporation (incorporation level (%) x [natural amino acid]/[derivative]). The amino acid derivatives that did not incorporate were excluded from this assay, except for 19α, which was studied for the purpose of comparison.

Cell-free protein synthesis

The E. coli cell-free protein synthesis was conducted by following a procedure, as described in earlier reports.15,17 The natural amino acids were used at 1.0 mM. The unnatural amino acids were used at 2.0 mM and the corresponding natural amino acid was not added. The reactions were proceeded for 6 h at 30 °C, with shaking at 200 rpm, unless otherwise stated. The competitive experiments with unnatural amino acids and their corresponding natural counterparts were conducted in a system identical to that used for usual cell-free protein synthesis, with the cognate amino acid supplied to the expression at 0.2 mM concentration to compete with the foreign substrate at 2 mM. In cases where racemic mixtures were used, this concentration refers to the L-enantiomer content. The levels of unnatural amino acid incorporation in competitive experiments were concluded from the mass spectra. The competition efficiencies are given as percentages and were calculated based on the ratio of the concentrations between the natural amino acid and its corresponding competitor, and the level of unnatural amino acid incorporation (incorporation level (%) x [natural amino acid]/[derivative]). The amino acid derivatives that did not incorporate were excluded from this assay, except for 19α, which was studied for the purpose of comparison.

Protein analysis

The crude contents of the inner reaction was analyzed by SDS-PAGE, to confirm the production of soluble protein. The inner reaction contents were then purified using the HisGraviTrap® Kit following the native conditions recommended by the
the activity of the prepared enzyme determines the catalytic rate constant \( k_{\text{cat}} \) through HPLC monitoring the formation of AMP by analyzing aliquots taken from the reaction mixture and quenched with 0.1% SDS every 15 seconds (Fig. 6c). It gave Tyr a \( k_{\text{cat}} \) of 11.3 s\(^{-1}\) (Fig. 6c and d), which is in excellent agreement with reported value of 12 s\(^{-1}\).\(^{24}\)

### Aminoacylation assay

The aminoacylation of amino acid derivative was conducted in a manner identical to Tyr. Stock solutions were prepared in Tris–HCl buffer containing 100.0 mM Tris, 15.0 mM MgCl\(_2\), 40.0 mM KCl and 1.0 mM DTT, at pH 7.6. The reaction was performed at 37 \(^\circ\)C in a 0.5 mL microcentrifuge, with a total volume of 60 \( \mu \)L containing 0.05 or 0.15 \( \mu \)M \( E.\) coli His\(_6\)TyrRS, 1.0 mM ATP, 16.7 or 60.0 \( \mu \)M tRNA\(^{\text{Tyr}}\), and 0.1 mg mL\(^{-1}\) bovine serum albumin (BSA). After pre-incubation for one minute, the reaction was initiated by adding in the \( E.\) coli His\(_6\)TyrRS. A 10 \( \mu \)L sample was taken out by hand from the proceeding reaction every 15 or 40 seconds and quenched into 5.0 \( \mu \)L 0.3% (w/v) aqueous sodium dodecylsulfate (SDS). The SDS-treated samples were either immediately analyzed by HPLC or stored at –20 \(^\circ\)C for later analysis. In the latter case, the storage was never longer than 12 hours. The solvent system used for HPLC analysis was as follows:

| Time (min) | Buffer A (%) | Buffer B (%) |
|------------|---------------|---------------|
| 0          | 87            | 13            |
| 18         | 87            | 13            |
| 19         | 70            | 30            |
| 23         | 70            | 30            |
| 24         | 87            | 13            |

In this solvent system, Buffer A contains 60.0 mM ammonium phosphate and 5.0 mM tetrabutylammonium phosphate, at pH 5.0, while Buffer B contains 5.0 mM tetrabutylammonium phosphate in methanol. Both buffers were filtered and degased before use.

### Conflicts of interest

There are no conflicts to declare.

### Acknowledgements

We acknowledge financial support by the Doctoral Research Startup Fund Project of the Zunyi Medical University (F-946). We express our sincere thanks to M. Liutkus and D. Fankhauser for helpful discussions regarding the preparation of the manuscript. We are also indebted to J. E. Hennessy and A. Alissandratos for assistance in recording protein mass spectra.

### References

1. (a) M. Ibba and D. Söll, Aminoacyl-tRNA synthesis, *Annu. Rev. Biochem.*, 2000, **69**, 617–650; (b) O. Kotik-Kogan, N. Moor, D. Tworowski and M. Safro, Structural basis for discrimination of \( \delta \)-phenylalanine from \( \delta \)-tyrosine by
phenylalanyl-tRNA synthetase, Structure, 2005, 13, 1799–1807.

2 (a) S. A. Fraser and C. J. Easton, Biosynthetic incorporation of fluorinated amino acids into peptides and proteins, Aust. J. Chem., 2015, 68, 9–12; (b) N. Budisa, Prolegomena to future experimental efforts on genetic code engineering by expanding its amino acid repertoire, Angew. Chem., Int. Ed., 2004, 43, 6426–6463.

3 (a) T. H. Johnston, A. M. Lacoste, N. P. Visanji, A. E. Lang, S. H. Fox and J. M. Brodie, Repurposing drugs to treat L-DOPA-induced dyskinesia in Parkinson’s disease, Neuropharmacology, 2019, 147, 11–27; (b) R. A. Ruggiero, J. Bruzzo, P. Chiarella, P. Di Gianna, M. A. Isturiz, S. Linskens, N. Speziale, R. P. Meiss, O. D. Bustuabad and C. D. Pasqualini, Tyrosine isomers mediate the classical phenomenon of concomitant tumor resistance, Cancer Res., 2011, 71, 7113–7124.

4 (a) B. R. Ipson and A. L. Fisher, Roles of the tyrosine isomers meta-tyrosine and ortho-tyrosine in oxidative stress, Ageing Res. Rev., 2016, 27, 93–107; (b) G. Gueron, N. Anselmino, P. Chiarella, E. G. Ortiz, S. L. Vickers, A. V. Paez, J. Giudice, M. D. Contin, D. Leonardi, F. Jaworski, V. Manzano, A. Straza, D. R. Montagna, E. Labanca, J. Cotignola, N. D’Accorso, A. Woloszynska-Read, N. Navone, R. P. Meiss, R. Ruggiero and E. Vazquez, Game-changing restraint of Ros-damaged phenylalanine, upon tumor metastasis, Cell Death Dis., 2018, 9, 140.

5 F. Agostini, J. S. Völler, B. Koksch, C. G. Acevedo-Rocha, V. Kubyshkin and N. Budisa, Biocatalysis with unnatural amino acids: enzymology meets xenobiology, Angew. Chem., Int. Ed., 2017, 56, 9680–9703.

6 P. J. Nowatcki, C. Franck, S. A. Maskarinec, G. Ravichandran and D. A. Tirrell, Mechanically tunable thin films of photosensitive artificial proteins: preparation and characterization by nanoindentation, Macromolecules, 2008, 41, 1839–1845.

7 E. R. Stadtman, Protein oxidation and aging, Science, 1992, 257, 1220–1224.

8 (a) C. Amatore, S. Arbault, M. Guille and F. Lemaitre, Electrochemical monitoring of single cell secretion: vesicular exocytosis and oxidative stress, Chem. Rev., 2008, 108, 2585–2621; (b) H. Ou, Z. Huang, Z. Mo and J. Xiao, The characteristics and roles of advanced oxidation protein products in atherosclerosis, Cardiovasc. Toxicol., 2017, 17, 1–12.

9 G. A. Molnár, V. Nemes, Z. Biró, A. Ludány, Z. Wagner and I. Wittmann, Accumulation of the hydroxyl free radical markers meta-, ortho-tyrosine and DOPA in cataractous lenses is accompanied by a lower protein and phenylalanine content of the water-soluble phase, Free Radical Res., 2005, 39, 1359–1366.

10 J. Z. Sun, H. Kaur, B. Halliwell, X. Y. Li and R. Bolli, Use of aromatic hydroxylation of phenylalanine to measure production of hydroxyl radicals after myocardial ischemia in vivo. Direct evidence for a pathogenetic role of the hydroxyl radical in myocardial stunning, Circ. Res., 1993, 73, 534–549.

11 C. A. O’Neill, L. W. Fu, B. Halliwell and J. C. Longhurst, Hydroxyl radical production during myocardial ischemia and reperfusion in cats, Am. J. Physiol.: Heart Circ. Physiol., 1996, 271, H660–H667.

12 S. W. Chan, R. A. Dunlop, A. Rowe, K. L. Double and K. J. Rodgers, 3-DOPA is incorporated into brain proteins of patients treated for Parkinson’s disease, inducing toxicity in human neuroblastoma cells in vitro, Exp. Neurol., 2012, 238, 29–37.

13 J. T. Ngo and D. A. Tirrell, Noncanonical amino acids in the interrogation of cellular protein synthesis, Acc. Chem. Res., 2011, 44, 677–685.

14 (a) A. R. Fersht and M. M. Kaethner, Enzyme hyperspecificity. Rejection of threonine by the valyl-tRNA synthetase by misacylation and hydrolytic editing, Biochemistry, 1976, 15, 3342–3346; (b) A. R. Fersht, J. S. Shindler and W. C. Tsui, Probing the limits of protein-amino acid side chain recognition with the aminoacyl-tRNA synthetases. Discrimination against phenylalanine by tyrosyl-tRNA synthetases, Biochemistry, 1980, 19, 5520–5524; (c) F. Fasiolo and A. R. Fersht, The aminoacyl adenylate mechanism in the aminoacylation reaction of yeast phenylalanyl-tRNA synthetase, Eur. J. Biochem., 1978, 85, 85–88.

15 (a) D. J. Stigers, Z. I. Watts, J. E. Hennessy, H. K. Kim, R. Martini, M. C. Taylor, K. Ozawa, J. W. Keillor, N. E. Dixon and C. J. Easton, Incorporation of chlorinated analogues of aliphatic amino acids during cell-free protein synthesis, Chem. Commun., 2011, 47, 1839–1841; (b) I. N. Arthur, J. E. Hennessy, D. Padmakshan, D. J. Stigers, S. Lesturgez, S. A. Fraser, M. Liutkus, G. Otting, J. G. Oakeshott and C. J. Easton, In situ deprotection and incorporation of unnatural amino acids during cell-free protein synthesis, Chem. – Eur. J., 2013, 19, 6824–6830; (c) M. Liutkus, S. A. Fraser, K. Caron, D. J. Stigers and C. J. Easton, Peptide synthesis through cell-free expression of fusion proteins incorporating modified amino acids as latent cleavage sites for peptide release, ChemBioChem, 2016, 17, 908–912.

16 (a) N. Voloshchuk, A. Y. Zhu, D. Snydacker and J. K. Montclare, Positional effects of monofluorinated phenylalanines on histone acetyltransferase stability and activity, Bioorg. Med. Chem. Lett., 2009, 19, 5449–5451; (b) K. R. Mehta, C. Y. Yang and J. K. Montclare, Modulating substrate specificity of histone acetyltransferase with unnatural amino acids, Mol. BioSyst., 2011, 7, 3050–3055.

17 (a) K. Ozawa, M. J. Headlam, D. Mouradov, S. J. Watt, J. L. Beck, K. J. Rodgers, R. T. Dean, T. Huber, G. Otting and N. E. Dixon, Translational incorporation of 3,4-dihydroxyphenylalanine into proteins, FEBS J., 2005, 272, 3162–3171; (b) N. Ayyadurai, N. S. Prabhu, K. Deepankumar, A. Kim, S. G. Lee and H. Yun, Biosynthetic substitution of tyrosine in green fluorescent protein with its surrogate fluorotyrosine in Escherichia coli, Biotechnol. Lett., 2011, 33, 2201; (c) N. Ayyadurai, K. Deepankumar, N. S. Prabhu, N. Budisa and H. Yun, Evaluation and biosynthetic incorporation of
chlorotyrosine into recombinant proteins, Biotechnol. Bioprocess Eng., 2012, 17, 679–686; (d) S. J. Oh, K. H. Lee, H. C. Kim, C. Catherine, H. Yun and D. M. Kim, Translational incorporation of multiple unnatural amino acids in a cell-free protein synthesis system, Biotechnol. Bioprocess Eng., 2014, 19, 426–432.

18 (a) L. Klipcan, N. Moor, N. Kessler and M. G. Safro, Eukaryotic cytosolic and mitochondrial phenylalanyl-tRNA synthetases catalyze the charging of tRNA with the meta-tyrosine, Proc. Natl. Acad. Sci. U. S. A., 2009, 106, 11045–11048; (b) N. Moor, L. Klipcan and M. G. Safro, Bacterial and eukaryotic phenylalanyl-tRNA synthetases catalyze misaminoacylation of tRNAphe with 3,4-dihydroxy-l-phenylalanine, Chem. Biol., 2011, 18, 1221–1229; (c) E. A. First, l-DOPA ripes in tRNA^phe, Chem. Biol., 2011, 18, 1201–1202; (d) O. Popp, V. Larrailet, H. Kettenberger, I. H. Gorr, M. Hilger, F. Lipsmeyer, A. Zeck and N. Beaucamp, Molecular polygamy: the promiscuity of l-phenylalanyl-tRNA-synthetase triggers misincorporation of meta- and ortho-tyrosine in monoclonal antibodies expressed by Chinese hamster ovary cells, Biotechnol. Bioeng., 2015, 112, 1187–1199; (e) K. J. Rodgers and N. Shiozawa, Misincorporation of amino acid analogues into proteins by biosynthesis, Int. J. Biochem. Cell Biol., 2008, 40, 1452–1466; (f) H. Gurer-Orhan, N. Ercal, S. Mare, S. Pennathur, H. Orhan and J. W. Heinecke, Misincorporation of free m-tyrosine into cellular proteins: a potential cytotoxic mechanism for oxidized amino acids, Biochem. J., 2006, 395, 277–284.

19 K. J. Edwards, D. L. Ollis and N. E. Dixon, Crystal structure of cytoplasmic Escherichia coli peptidyl-prolyl isomerase: evidence for decreased mobility of loops upon complexation, J. Mol. Biol., 1997, 271, 258–265.

20 D. V. Santi and P. V. Danenberg, Phenylalanyl transfer ribonucleic acid synthetase from Escherichia coli. Analysis of the phenylalanine binding site, Biochemistry, 1971, 10, 4813–4820.

21 R. Calendar and P. Berg, The catalytic properties of tyrosyl ribonucleic acid synthetases from Escherichia coli and Bacillus subtilis, Biochemistry, 1966, 5, 1690–1695.

22 T. J. Bullwinkle, N. M. Reynolds, M. Raina, A. Moghal, E. Matsa, A. Rajkovic, H. Kayadibi, F. Fazlollahi, C. Ryan, N. Howitz, K. F. Faull, B. A. Lazazzera and M. Ibbba, Oxidation of cellular amino acid pools leads to cytotoxic mistranslation of the genetic code, eLife, 2014, 3, e02501.

23 T. Kobayashi, T. Takimura, R. Sekine, K. Vincent, K. Kamata, K. Sakamoto, S. Nishimura and S. Yokoyama, Structural snapshots of the KMSKS loop rearrangement for amino acid activation by bacterial tyrosyl-tRNA synthetase, J. Mol. Biol., 2005, 346, 105–117.

24 D. Kiga, K. Sakamoto, K. Kodama, T. Kigawa, T. Matsuda, T. Yabuki, M. Shirouzu, Y. Harada, H. Nakayama, K. Takio, Y. Hasegawa, Y. Endo, I. Hirao and S. Yokoyama, 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., 2002, 99, 9715–9720.

25 M. C. Hartman, K. Josephson and J. W. Szostak, Enzymatic aminoacylation of tRNA with unnatural amino acids, Proc. Natl. Acad. Sci. U. S. A., 2006, 103, 4356–4361.

26 (a) K. J. Rodgers, Non-protein amino acids and neurodegeneration: the enemy within, Exp. Neurol., 2014, 253, 192–196; (b) W. M. Garrison, Reaction mechanisms in the radiolysis of peptides, polypeptides, and proteins, Chem. Rev., 1987, 87, 381–398.

27 I. Mermershtain, I. Finarov, L. Klipcan, N. Kessler, H. Rozenberg and M. G. Safro, Idiosyncreasy and identity in the prokaryotic Phe-system: crystal structure of E. coli phenylalanyl-tRNA synthetase complexed with phenylalanine and AMP, Protein Sci., 2011, 20, 160–167.

28 (a) P. Kast and H. Hennecke, Amino acid substrate specificity of Escherichia coli phenylalanyl-tRNA synthetase altered by distinct mutations, J. Mol. Biol., 1991, 222, 99–124; (b) M. Ibbba, P. Kast and H. Hennecke, Substrate specificity is determined by amino acid binding pocket size in Escherichia coli phenylalanyl-tRNA synthetase, Biochemistry, 1994, 33, 7107–7112; (c) N. Sharma, R. Furter, P. Kast and D. A. Tirrell, Efficient introduction of aryl bromide functionality into proteins in vivo, FEBS Lett., 2000, 467, 37–40; (d) K. Kirshenbaum, I. S. Carrico and D. A. Tirrell, Biosynthesis of proteins incorporating a versatile set of phenylalanine analogues, ChemBioChem, 2002, 3, 235–237; (e) D. Datta, P. Wang, I. S. Carrico, S. L. Mayo and D. A. Tirrell, A designed phenylalanyl-tRNA synthetase variant allows efficient in vivo incorporation of aryl ketone functionality into proteins, J. Am. Chem. Soc., 2002, 124, 5652–5653.

29 A. R. Fersht, J. P. Shi, J. Knill-Jones, D. M. Lowe, A. J. Wilkinson, D. M. Blow, P. Brick, P. Carter, M. M. Y. Waye and G. Winter, Hydrogen bonding and biological specificity analyzed by protein engineering, Nature, 1985, 314, 235.