Functional replacement of the endogenous tyrosyl-tRNA synthetase–tRNA\textsubscript{Tyr} pair by the archaeal tyrosine pair in \textit{Escherichia coli} for genetic code expansion

Fumie Iraha\textsuperscript{1}, Kenji Oki\textsuperscript{2}, Takatsugu Kobayashi\textsuperscript{1}, Satoshi Ohno\textsuperscript{3}, Takashi Yokogawa\textsuperscript{3}, Kazuya Nishikawa\textsuperscript{3}, Shigeyuki Yokoyama\textsuperscript{1,2,*} and Kensaku Sakamoto\textsuperscript{1,*}

\textsuperscript{1}RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, \textsuperscript{2}Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033 and \textsuperscript{3}Department of Biomolecular Science, Faculty of Engineering, Gifu University, Yanagido 1-1, Gifu 501-1193, Japan

Received November 10, 2009; Revised January 25, 2010; Accepted January 27, 2010

\textbf{ABSTRACT}

Non-natural amino acids have been genetically encoded in living cells, using aminoacyl-tRNA synthetase–tRNA pairs orthogonal to the host translation system. In the present study, we engineered \textit{Escherichia coli} cells with a translation system orthogonal to the \textit{E. coli} tyrosyl-tRNA synthetase (TyrRS)–tRNA\textsubscript{Tyr} pair, to use \textit{E. coli} TyrRS variants for non-natural amino acids in the cells without interfering with tyrosine incorporation. We showed that the \textit{E. coli} TyrRS–tRNA\textsubscript{Tyr} pair can be functionally replaced by the \textit{Methanocaldococcus jannaschii} and \textit{Saccharomyces cerevisiae} tyrosine pairs, which do not cross-react with \textit{E. coli} TyrRS or tRNATyr. The endogenous TyrRS and tRNA\textsubscript{Tyr} genes were then removed from the chromosome of the \textit{E. coli} cells expressing the archaeal TyrRS–tRNA\textsubscript{Tyr} pair. In this engineered strain, 3-iodo-L-tyrosine and 3-azido-L-tyrosine were each successfully encoded with the \textit{E. coli} amber suppressor tRNATyr and a TyrRS variant, which was previously developed for 3-iodo-L-tyrosine and was also found to recognize 3-azido-L-tyrosine. The structural basis for the 3-azido-L-tyrosine recognition was revealed by X-ray crystallography. The present engineering allows \textit{E. coli} TyrRS variants for non-natural amino acids to be developed in \textit{E. coli}, for use in both eukaryotic and bacterial cells for genetic code expansion.

\textbf{INTRODUCTION}

The repertoire of genetically encoded amino acids has been expanded in \textit{Escherichia coli}, yeast, insect and mammalian cells, with the purpose of increasing the chemical and structural diversity in proteins (1–6). The genetically encoded ‘non-natural’ amino acids have unique chemical groups and functionalities not found among the 20 canonical amino acids, such as heavy atoms, photo-reactive linkers, fluorescent groups and chemical groups for site-specific labeling (7). The incorporation of these novel structures into proteins at desired sites promises to facilitate protein science and engineering.

The genetic encoding of a non-natural amino acid requires the development of an aminoacyl-tRNA synthetase (aaRS)–tRNA pair specific to the amino acid. To date, several aaRS species have been engineered to recognize useful non-natural amino acids, and among these, the majority are L-tyrosine or L-phenylalanine derivatives recognized by variants of tyrosyl-tRNA synthetase (TyrRS) (7). In living cells, non-natural amino acids are encoded with the amber codon in most cases, by using the amber suppressor tRNA–aaRS pair from a different organism that does not cross-react with the endogenous tRNA or aaRS (8). The bacterial TyrRS, for example, does not recognize the tRNA\textsubscript{Tyr} molecules from eukaryotes or archaea, while the TyrRS molecules from these two kingdoms do not recognize the bacterial tRNA\textsubscript{Tyr}. In fact, the bacterial TyrRS–tRNA\textsubscript{Tyr} pair does not cross-react with the endogenous tRNA and aaRS species in eukaryotic host cells, and has been used to achieve the site-specific incorporation of L-tyrosine.
derivatives into proteins only at the amber codon (2,9,10).
On the other hand, the genetic encoding of L-tyrosine
derivatives into proteins only at the amber codon (2,9,10).
Thus, the E. coli and archaeal TyrRS–tRNA,Tyr pairs are used in
eukaryotic and E. coli, cells, respectively. Other aaRSs,
such as tryptophanyl-tRNA synthetase (11), phenylalanyl-tRNA
synthetase (5,12), leucyl-tRNA synthetase (13) and lysyl-tRNA
synthetase (14), also take advantage of the orthogonal tRNA specificities
across different kingdoms.

Pyrolysyl-tRNA synthetase (PylRS) differs from these
aaRS species, in that the PylRS–tRNA,Pyl pair from
Methanosarcina mazei is reportedly orthogonal to both the
E. coli and eukaryotic translation systems (15–17).
Archaeal PylRS variants can be developed in E. coli,
cells, and are then used for incorporating L-lysine derivatives
into proteins in E. coli, yeast and mammalian cells
(16–20). With this advantage, the PylRS–tRNA,Pyl pair is
a promising system for genetic code expansion.
The usefulness of E. coli TyrRS variants would also be
increased, if the E. coli translation system could be engi-
neered to be orthogonal to the E. coli TyrRS–tRNA,Tyr
pair. This modification to E. coli cells may be achieved
by substituting an archaeal or eukaryal tyrosine pair for
the endogenous tyrosine pair.

In the present study, the genes encoding TyrRS and
tRNA,Tyr were disrupted in the chromosome of the
E. coli cells expressing the M. jannaschii TyrRS–
tRNA,Tyr pair. The engineered cells then hosted the
E. coli tyrRS variant and the amber suppressor
tRNA,Tyr, to genetically encode either 3-iodo-L-tyrosine or
3-azido-L-tyrosine, depending on the presence of the cor-
responding amino acid in the growth medium. This TyrRS
variant, designated as iodoTyrRS-ec, was originally developed for 3-iodo-L-tyrosine (21), and has been used for
the site-specific incorporation of this amino acid into proteins in mammalian cells (2,22). We found that
iodoTyrRS-ec also recognizes 3-azido-L-tyrosine, and the structural basis for its recognition by this variant was
elucidated by X-ray crystallography.

MATERIALS AND METHODS

Strains, growth conditions, growth rate, reagents and
non-natural amino acids

Escherichia coli strains TOP10, BL21(DE3) and DH5α
were purchased from Invitrogen, Novagen and Toyobo
(Tokyo, Japan), respectively. Kanamycin (Km) (30 µg/ml),
chloramphenicol (Cm) (34 µg/ml), ampicillin
(Amp) (10 µg/ml), Zeocin (25 µg/ml), t-arabinose (0.2%,
w/v) and d-glucose (0.4%, w/v) were added to LB media
as indicated, except that Cm was added at a concentration
of 25 µg/ml when non-natural amino acids were incorpo-
rated into proteins. The growth rate was monitored as the
change in the natural logarithm of the optical density at
600 nm of the culture per hour. Zeocin was purchased from
InvivoGen (San Diego, USA). 3-Iodo-L-tyrosine
was purchased from Sigma-Aldrich. 3-Bromo-L-tyrosine,
3-azido-L-tyrosine and the chemical conjugate between
triarylphosphine and fluorescein were commercially
synthesized by Shinsei Chemical Company Ltd (Osaka,
Japan).

Plasmids

The E. coli, M. jannaschii and Saccharomyces cerevisiae
TyrRS genes, and the variant gene encoding iodoTyrRS-
cec were cloned with the tyr,S promoter into pACYC184
(Shionogi & Nippon Gene Co., Ltd., Japan) to create pECYS, pMJYS,
pScYS and pECtYS, respectively. The wild-type TyrRS
genes with the promoter were also cloned in an
ampicillin-resistant ColIb-P9 plasmid, pApI02 (23), to
create pApECYS, pApMJYS and pApScYS, respectively. The
archaeal and yeast tRNA,Tyr genes were cloned with the
tyr,T promoter into pACYC184, to create pMjYR and
pScYR, respectively. These tRNA,Tyr genes with the
promoter were also cloned in pMJYS, pScYS and pECtYS,
respectively. Three copies of these tRNA,Tyr genes were cloned after the
tyr,T promoter in pApI02, to create pApMJYS and
pApScYS, respectively. The archaeal and yeast TyrRS
genes with the tyr,S promoter were cloned in these
plasmids, pApMJYS and pApScYS, to create
pApMJYSYR and pApScYSYR, respectively. The kanamycin resistance (Km) gene was cloned in the ‘amber suppressor control’ plasmid from the Interchange
amber suppressor in vivo mutagenesis system (Promega),
to create the plasmid pACamK. The resulting plasmid
carries the replication origin of pACYC184 and
an amber mutant chloramphenicol acetyltransferase (CAT)
gen, in addition to the cloned Km gene. The E. coli
amber suppressor tRNA,Tyr gene was cloned with the
tyr,T promoter into pACamK, to create the plasmid
pACamKsupF.

To create the plasmid pMJY, the coding sequence of the
Zeocin resistance gene was amplified by polymerase chain
reaction (PCR) from the vector pcDNA4/TO (Invitrogen).
The sequence 5'-ACGACTCTATAGGAGGCGCC-3',
in which the ribosome-binding site (RBS) is underlined,
was added at the 5'-end of the coding sequence. The amplified gene was cloned between the Sppl–PstI sites of
the pBR322 vector, downstream of the promoter of the
ampicillin resistance gene and in place of this gene, to
create the plasmid pBRZeo. The M. jannaschii TyrRS
and tRNA,Tyr genes with the tyr,S and tyr,T promoters,
respectively, were cloned in pBRZeo to obtain pMJY.

For incorporating non-natural amino acids into proteins, the E. coli amber suppressor tRNA,Tyr gene with
the tyr,T promoter was cloned in pApI02 to create
pApYS. The replication origin and the ampicillin resis-
tance gene in the plasmid pGEX-4T-3 (GE Healthcare)
were replaced with the origin and Km gene from the
vector pRSF-1b (Novagen), to create the plasmid
pRGexGST. The resulting plasmid retains the glutathione
S-transferase (GST) gene under the control of the tac
promoter. Then, the tyrosine codon at position 84 of
GST was mutated to the amber codon, to create
pRGexGST (Am84). The wild-type gst gene and its
mutant with the amber codon at position 25 (24) were
cloned after the T7 promoter in pRSF-1b to create
pRSFGST and pRSFGST(\text{Am25}), respectively. The rat calmodulin gene and its mutant with the amber codon at position 80 (25) were also cloned after the T7 promoter in pRSF-1b, to create pRSFCAM and pRSFCAM(\text{Am80}), respectively.

**Construction of the conditional tyrS mutant by chromosome engineering**

All of the modifications to the *E. coli* chromosome were made using the RT/ET kit (Gene Bridges GmbH, Germany), which utilizes the bacteriophage λ recombination system to promote homologous recombination (26,27). The DNA encoding the RBS, the *araBAD* promoter and the *araC* gene—designated as the *ara* block—was introduced into the chromosome in place of the *tyrS* promoter. For this engineering, the *ara* block was amplified by PCR from the pBAD/His vector (Gene Bridges GmbH, Germany), which utilizes the bacteriophage λ recombination system to promote homologous recombination (26,27). The DNA encoding the RBS, the *ara* block— which was introduced into the chromosome in place of the *tyrS* promoter. For this engineering, the *ara* block was amplified by PCR from the pBAD/His vector (Gene Bridges GmbH, Germany), which utilizes the bacteriophage λ recombination system to promote homologous recombination (26,27).

The knocked-in *cat* gene was removed using a linear DNA fragment consisting of the lhm and rhm sequences linked directly to each other. Cells deprived of *cat* were selected in the growth media supplemented with Cm (10 \mu g/ml) and ampicillin (200 \mu g/ml). The cells that are resistant to Cm and are able to grow in this medium should undergo lysis in the presence of ampicillin, while only sensitive cells should survive. The colonies formed by the surviving cells were replicated on an LB plate containing Cm (10 \mu g/ml), to check for the sensitivity to Cm.

The *E. coli* amber suppressor tRNA$_{\text{Tyr}}$ was introduced into the chromosome in place of the knocked-in *cat* in the *tyrS* locus, using a linear DNA consisting of lhm-U, the tRNA coding sequence and rhm-U in this order. For the introduction of the suppressor tRNA$_{\text{Tyr}}$ in the *tyrT* and *tyrV* loci, the *cat* gene was introduced into the chromosome in place of the knocked-in *cat* in the *tyrS* locus, using a linear DNA consisting of lhm-U, the tRNA coding sequence and rhm-U in this order.

**Site-specific incorporation of non-natural amino acids into proteins**

FT3 cells were transformed with pEcIYS, pApS\text{YR} and pRG\text{ExGST}(\text{Am84}), and were grown in medium containing Cm, Amp and Km. The expression of the *gst* gene was induced by the addition of IPTG (1 mM, final), and at the same time, 3-iodo-L-tyrosine (0.3 mg/ml, final) was added to the medium. FB3 cells were transformed with pEcIYS and either pRSFGST(\text{Am25}) or pRSFCAM(\text{Am80}), and were grown in medium containing Kam and Cm. The expression of the recombinant proteins was induced with IPTG (1 mM, final), and at the same time, 3-iodo-L-tyrosine or 3-azido-L-tyrosine (0.3 mg/ml, final) was added to the medium. Labeling with the triarylphosphine-fluorescein conjugate was performed as described previously (20,25). Fluorescence from an SDS-gel was detected using an image analyzer, LAS-1000 (Fuji Film, Japan). Mass spectrometric analysis was commercially performed by Shimadzu Biotech (Japan).

**Structure determination**

The iodoTyrRS-ec catalytic domain (residues 1–322) was prepared as described previously (28). The samples were dialyzed against 20 mM Tris–Cl (pH 7.5) containing 50 mM NaCl, 10 mM 2-mercaptoethanol and 1 mM 3-azido-L-tyrosine, and then they were concentrated to 4–6 mg/ml with an Amicon Ultra 4 (Millipore). Crystallization of the complex between iodoTyrRS-ec and 3-azido-L-tyrosine was performed as described (28). The samples and crystals were shaded to avoid the
photolysis of 3-azido-L-tyrosine. The X-ray diffraction experiment of the crystals was performed at the beamline BL26B1 at SPring-8 (Harima, Japan). The data were processed with HKL2000 (29). The iodoTyrRS-ec–3-azido-L-tyrosine structure was determined by molecular replacement with the program Molrep (30), using the structure of the complex between iodoTyrRS-ec and 3-iodo-L-tyrosine (28) as a search model. The refinement and model building were performed as described (28). The model of 3-azido-L-tyrosine was generated by the Dundee PRODRG2 server (31).

RESULTS AND DISCUSSION

Complementation of a conditional E. coli tyrS mutant by the archaeal and eukaryal TyrRS–tRNA^{Tyr} pairs

We examined whether the translational function of E. coli TyrRS can be replaced with that of a eukaryal or archaeal TyrRS–tRNA^{Tyr} pair in E. coli cells. The expression system for the arabinose operon was used to control the expression of the chromosomal tyrS gene encoding E. coli TyrRS. The araBAD promoter, together with the araC gene, was introduced into the chromosome in place of the tyrS promoter, to allow the induction of TyrRS expression with L-arabinose. Transcription from the araBAD promoter is tightly repressed in the presence of D-glucose (32); the created conditional growth was complemented by the conditional growth (Supplementary Figure S1B). This observation is consistent with the reports that these archaeal and eukaryal TyrRS–tRNA Tyr pairs from multicopy plasmids are necessary for the complementation, we used the low-copy plasmid pApEcYS to introduce the yeast TyrRS–tRNA Tyr pairs into multicopy plasmids is necessary. In these plasmids, the tRNA Tyr genes were expressed from the low-copy plasmid pApScYS, with the tRNA Tyr expressed from the multicopy plasmid pMjYR, but again, no complementation was detected (Supplementary Figure S1D). These observations indicated that the complementation requires the overproduction of both the archaeal TyrRS and tRNA^{Tyr} molecules.

We also expressed S. cerevisiae TyrRS from either the multicopy plasmid pScYS, with the yeast tRNA^{Tyr} being expressed from the low-copy plasmid pApScYS, or the low-copy plasmid pApEcYS, with the tRNA^{Tyr} being expressed from the multicopy plasmid pScYS. Neither combination achieved the complementation (Supplementary Figure S1D), indicating that both the yeast TyrRS and tRNA^{Tyr} need to be overproduced. Similar complementation tests were performed for the yeast TyrRS–tRNA^{Tyr} pair at 30°C, the normal growth temperature for yeast. However, no complementation was observed (data not shown), indicating that the optimal temperature for these molecules is probably not relevant to the requirement of the overproduction of the TyrRS–tRNA^{Tyr} pair. Although the reason why the exogenous tyrosine pairs must be overproduced is not clear, this requirement is consistent with a recent report showing that the overproduction of the M. jannaschii TyrRS–tRNA^{Tyr} pair in E. coli cells is a prerequisite for the efficient incorporation of non-natural amino acids into proteins (35).

Replacement of the E. coli TyrRS–tRNA^{Tyr} pair by the M. jannaschii pair in an E. coli strain

The TyrRS and tRNA^{Tyr} genes were removed from the chromosome in an E. coli K12 strain expressing the M. jannaschii tyrosine pair from the multicopy plasmid pMJY, constructed as described in the ‘Materials and Methods’ section (Figure 1B). First, the CAT gene was introduced into the chromosome in place of the tyrS gene by homologous recombination. Then, the cat gene was removed from the chromosome to create the strain FT1 [TOP10 ΔtyrS pMJY]. A DNA fragment including the tyrS locus was amplified from this strain and analyzed replication origin and its copy number is 1.7 per cell (34). In a control experiment, the vector pAp102 did not complement the tyrS mutant, while a pAp102-derived plasmid carrying the tyrS gene (pApEcYS) displayed the complementing activity (Supplementary Figure S1B). The M. jannaschii and S. cerevisiae TyrRS genes were then each cloned in this low-copy vector, together with three copies of their cognate tRNA Tyr genes, to create the plasmids pApMjYSYR and pApScYSYR, respectively. Neither of these plasmids showed the complementing activity (Supplementary Figure S1C), indicating that these exogenous TyrRS or tRNA^{Tyr} molecules must be expressed from a multicopy plasmid.

We then determined which molecule, TyrRS or tRNA^{Tyr}, needed to be overproduced. First, for the M. jannaschii tyrosine pair, the enzyme was expressed from the multicopy plasmid pMjYS, with the tRNA^{Tyr} being expressed from the low-copy plasmid pApMjYS. No complementing activity was observed for this pair of plasmids (Supplementary Figure S1D). Next, TyrRS was expressed from the low-copy plasmid pApMjYS, with the tRNA^{Tyr} expressed from the multicopy plasmid pMjYR, but again, no complementation was detected (Supplementary Figure S1D). These observations indicated that the complementation requires the overproduction of both the archaeal TyrRS and tRNA^{Tyr} molecules.

The TyrRS and tRNA Tyr genes were then placed after the promoter of the araBAD promoter, to allow the induction of TyrRS expression with L-arabinose. The iodoTyrRS–ec–3-azido-L-tyrosine structure was determined by molecular replacement with the program Molrep (30), using the structure of the complex between iodoTyrRS-ec and 3-iodo-L-tyrosine (28) as a search model. The model of 3-azido-L-tyrosine was generated by the Dundee PRODRG2 server (31).
to confirm the disruption of the tyrS gene. The absence of the E. coli TyrRS activity in this strain was also confirmed by expressing the E. coli amber suppressor tRNA\textsubscript{Tyr}, together with a cat amber mutant gene, from the plasmid pACamKsupF. This plasmid conferred chloramphenicol resistance to the parent strain TOP10, but not to FT1 cells (Figure 1C), showing that the E. coli TyrRS activity was absent in FT1 cells, and that the E. coli suppressor tRNA\textsubscript{Tyr} is orthogonal to the translation system of the cell.

Next, we removed all three of the E. coli tRNA\textsubscript{Tyr} genes from the chromosome of the FT1 cells. Two of them, tyrT and tyrV, are located close to each other on the E. coli chromosome, and share the tyrT promoter for expression. The cat gene was introduced into the chromosome of FT1 in place of these genes and their promoter, and the knocked-in cat gene was then removed to create the strain FT2 [TOP10 tyrS tyrT tyrV pMJY]. A DNA fragment encompassing the tyrT and tyrV loci was amplified from this strain and analyzed to confirm the disruption of these tRNA\textsubscript{Tyr} genes and the tyrT promoter. Then, to remove the tyrU gene, which is the third copy of the tRNA\textsubscript{Tyr} gene and is co-transcribed with two tRNA\textsubscript{Thr} species and a tRNA\textsubscript{Gly}, the cat gene without a promoter was introduced into the chromosome of FT2 in place of tyrU, for transcription together with the other tRNAs. The cat gene was then removed to create the strain FT3 [TOP10 tyrS tyrT tyrU tyrV pMJY]. The absence of the tRNA\textsubscript{Tyr} gene in the tyrU locus was confirmed by a sequence analysis.

In order to show that the strain FT3 lacks E. coli tyrS activity, we expressed iodoTyrRS-ec, an E. coli TyrRS variant that recognizes 3-iodo-L-tyrosine, and investigated growth inhibition. When 3-bromo-L-tyrosine was added to the growth media, the growth rate of DH5\textalpha{} cells, an E. coli K-12 strain, overproducing the variant was significantly reduced, as compared to that of the cells overproducing the wild-type E. coli TyrRS (Figure 2A). This observation suggested that iodoTyrRS-ec is able to efficiently attach 3-bromo-L-tyrosine to the E. coli tRNA\textsubscript{Tyr}, and to inhibit the growth by randomly incorporating this amino acid into proteins at tyrosine positions. On the other hand, strain FT3 is immune to this growth inhibition (Figure 2C), which requires the interaction between iodoTyrRS-ec and E. coli tRNA\textsubscript{Tyr}, and this observation indicates the absence of the E. coli tRNA\textsubscript{Tyr} in the cell. Thus, the E. coli translation system was successfully engineered to be orthogonal to the E. coli TyrRS–tRNA\textsubscript{Tyr} pair.

The iodoTyrRS-ec–E. coli suppressor tRNA\textsubscript{Tyr} pair site-specifically incorporates 3-iodo-L-tyrosine into proteins in FT3 cells

We used iodoTyrRS-ec to incorporate 3-iodo-L-tyrosine into proteins at the amber position in the strain FT3. This TyrRS variant was expressed from the tyrS promoter on the multicopy plasmid pEcIYS, which is compatible with pMJY expressing the archaeal tyrosine pair. Since the expression of the E. coli amber suppressor tRNA\textsubscript{Tyr} from a multicopy plasmid negatively affected the growth of the E. coli cells, the tRNA gene was expressed from the low-copy plasmid pApSYR. The gene encoding GST with an interrupting amber codon was carried on the third plasmid pRGexGST(Am84). Thus, it was necessary...
to maintain four different plasmids in FT3 cells for the site-specific incorporation of non-natural amino acids into proteins. The \( \text{gst} \) amber mutant gene produced the full-length product only when 3-iodo-\( L \)-tyrosine was supplemented in the growth media (Figure 3A, lanes 1 and 2), indicating that 3-iodo-\( L \)-tyrosine was incorporated into GST at the amber position, although the suppression efficiency was low.

The growth rate of FT3 was found to be just half that of its parent strain TOP10; the growth rates were determined as 0.7 and 1.4 \( h^{-1} \) for FT3 and TOP10, respectively, in LB medium with no antibiotics. Due to this slow growth rate, the maintenance of four different plasmids, and the low yield of suppression products, FT3 cells were apparently not useful for yielding large amounts of proteins containing non-natural amino acids. We tried to circumvent this problem by further engineering, starting with another \( E. coli \) strain.

**Replacement of the \( E. coli \) TyrRS–tRNA\( ^{Tyr} \) pair by the \( M. jannaschii \) tyrosine pair in strain BL21(DE3)**

The \( E. coli \) TyrRS and tRNA\( ^{Tyr} \) genes were replaced by the \( M. jannaschii \) pair in strain BL21(DE3), an \( E. coli \) B strain. This strain carries the gene encoding the bacteriophage T7 RNA polymerase in its chromosome, for the high-level expressions of recombinant proteins. The order of gene disruptions was not the same as that with TOP10. In addition, the \( E. coli \) amber suppressor tRNA\( ^{Tyr} \) gene was introduced into the chromosome of BL21(DE3), instead of on a plasmid, to reduce the number of required plasmids (Figure 1B). First, the cat gene was substituted for the tyrU gene, and was then replaced by the sequence encoding the \( E. coli \) suppressor tRNA\( ^{Tyr} \). This step effectively mutated tyrU to the suppressor tRNA\( ^{Tyr} \) gene, thus creating the strain FB1 [BL21 \( \lambda \)DE3 tyrU]. This strain was transformed with pMJY, and the tyrS gene was then removed from the chromosome to create FB2 [BL21 \( \lambda \)DE3 tyrU \( \Delta \)tyrS pMJY]. The tyrT and tyrV genes in FB2 were replaced with one copy of the suppressor tRNA\( ^{Tyr} \) gene. This step created the strain...
FB3 [BL21 λDE3 tyrU tyrV ΔtyrS ΔtyrT pMJY], in which tyrT was removed and tyrV was mutated to another copy of the suppressor tRNA\textsubscript{Tyr} gene, expressed from the tyrT promoter. Thus, the created FB3 strain expresses the archaean TyrRS–tRNATyr pair and the \textit{E. coli} suppressor tRNA\textsubscript{Tyr}, while it lacks \textit{E. coli} TyrRS and wild-type tRNA\textsubscript{Tyr}. The growth rates of the FB3 and BL21(DE3) strains were determined as 1.1 and 1.3 h\(^{-1}\), respectively; the FB3 cells grow almost as fast as the parent strain, and much faster than FT3. The modifications to the BL21(DE3) chromosome had no serious effects on the viability of the cells.

To demonstrate the site-specific incorporation of 3-iodo-L-tyrosine into proteins in the strain FB3, iodoTyrRS-ec was expressed from pEcLYS, and the \textit{gst} gene with the amber codon at position 25, \textit{gst} (Am25), was expressed from a separate multicopy plasmid, pRSFGST(Am25), under the control of the T7 promoter. The FB3 strain bearing these plasmids produced the full-length GST only in the presence of 3-iodo-L-tyrosine, and its yield was significantly higher than that in FT3 cells (Figure 3A, lanes 5 and 6). We previously used the amber suppressor tRNA\textsubscript{Tyr} and the 3-iodo-L-tyrosine-specific variant of TyrRS (iodoTyrRS-\textit{mj}) from \textit{M. jannaschii}, in order to site-specifically incorporate 3-iodo-L-tyrosine into proteins in \textit{E. coli} cells (36). The yield of the full-length product of \textit{gst}(Am25) in the presence of 3-iodo-L-tyrosine was then compared between FB3 cells expressing the iodoTyrRS-ec–\textit{E. coli} suppressor tRNA\textsubscript{Tyr} pair and the parent BL21(DE3) cells expressing the iodoTyrRS-\textit{mj}–archaean suppressor tRNA\textsubscript{Tyr} pair. Comparable amounts of the full-length product were detected in the extracts from these cells (Figure 3B).

Finally, the full-length GST from FB3 cells was analyzed by mass spectrometry. This amber suppression product was subjected to a trypsin digestion, for comparison with the peptide data previously obtained for the wild-type GST (36). The wild-type protein generates a peptide (Tyr25 peptide) with an average mass corresponding with the azido group, the shape of the space snugly enzyme residues appear to make firm contacts or interactions with the amino-acid binding pocket (Figure 4B). Although no enzyme residues appear to make firm contacts or interactions with the amino-acid binding pocket, the shape of the space snugly fits the substituent. By contrast, in the reported structure, iodoTyrRS-ec recognizes 3-iodo-L-tyrosine by making van der Waals contacts between the mutated residues, Val37 and Cys195, and the iodine atom (Figure 4B). The binding of L-tyrosine to iodoTyrRS-ec, on the other hand, is weakened by the Tyr37Val substitution, which removes the hydrogen bond between the phenolic hydroxyl groups of L-tyrosine and Tyr37 (28) (Figure 4A).

The \textit{M. jannaschii} iodoTyrRS\textit{mj} suppressor tRNA\textsubscript{Tyr} pair does not incorporate 3-iodo-L-tyrosine into proteins (data not shown), indicating that this archaean TyrRS efficiently, it has not been used for genetically encoding L-tyrosine derivatives in \textit{E. coli} cells. We found that iodoTyrRS-ec also recognizes 3-azido-L-tyrosine efficiently; this amino acid supplemented in the growth media inhibits the growth of normal \textit{E. coli} cells expressing iodoTyrRS-ec, as observed for 3-bromo-L-tyrosine (Figure 2B). Therefore, iodoTyrRS-ec was used to incorporate 3-azido-L-tyrosine into proteins in FB3 cells. The full-length product of \textit{gst}(Am25) was obtained only in the presence of 3-azido-L-tyrosine, and its yield was comparable to that of the product with 3-iodo-L-tyrosine (Figure 3A, lanes 7 and 8). Thus, FB3 cells expressing iodoTyrRS-ec and the \textit{E. coli} suppressor tRNA\textsubscript{Tyr} produced the proteins site-specifically containing either 3-azido-L-tyrosine or 3-iodo-L-tyrosine, depending on which amino acid is supplemented in the growth media.

To demonstrate the site-specific protein modification, rat calmodulin containing 3-azido-L-tyrosine was produced in FB3 cells expressing the iodoTyrRS-ec–suppressor tRNA\textsubscript{Tyr} pair. The full-length product was expressed from the amber mutant gene, carried on plasmid pRSFCAM(Am80), only in the presence of 3-azido-L-tyrosine, and was modified with a fluorescent probe by the Staudinger–Bertozzi reaction, while the wild-type protein expressed from plasmid pRSFCAM was not labeled (Figure 3C). p-Azido-L-phenylalanine and a L-lysine derivative with the azido group have already been genetically encoded in \textit{E. coli} cells (20,38). Tyrosines tend to reside on the protein surface (39), and such tyrosines on the surface could be safely replaced with 3-azido-L-tyrosine for site-specific protein labeling.

**Structural basis for 3-azido-L-tyrosine recognition by iodoTyrRS-ec**

\textit{E. coli} TyrRS was previously engineered into iodoTyrRS-ec by introducing a double substitution, Tyr37Val and Gln195Cys (Figure 4A) (21,28). To reveal the structural basis for the additional recognition of 3-azido-L-tyrosine by iodoTyrRS-ec, the crystal structure of a complex between these molecules was determined at 1.8-Å resolution (PDB ID: 2YXN). The statistics of the data set are listed in Table 1. The overall structure of the variant superposed well on the reported crystal structure of iodoTyrRS-ec complexed with 3-iodo-L-tyrosine (28).

The azido group of 3-azido-L-tyrosine is accommodated in the space created by the double substitution at the amino-acid binding pocket (Figure 4B). Although no enzyme residues appear to make firm contacts or interactions with the azido group, the shape of the space snugly fits the substituent. By contrast, in the reported structure, iodoTyrRS-ec recognizes 3-iodo-L-tyrosine by making van der Waals contacts between the mutated residues, Val37 and Cys195, and the iodine atom (Figure 4B). The binding of L-tyrosine to iodoTyrRS-ec, on the other hand, is weakened by the Tyr37Val substitution, which removes the hydrogen bond between the phenolic hydroxyl groups of L-tyrosine and Tyr37 (28) (Figure 4A).

\textit{M. jannaschii} iodoTyrRS\textit{mj} suppressor tRNA\textsubscript{Tyr} pair does not incorporate 3-azido-L-tyrosine into proteins (data not shown), indicating that this archaean TyrRS
variant does not recognize 3-azido-L-tyrosine. In the reported crystal structure of iodoTyrRS-ec bound with 3-azido-L-tyrosine, the mutated residues are denoted in red. The carbon atoms of the ligand are shown in green. The nitrogen, oxygen, and sulfur atoms are in blue, red, and yellow, respectively. The hydrogen bonds are shown as pink broken lines. (B) The surface structures of the binding pockets of iodoTyrRS-ec and iodoTyrRS-mj accommodating 3-iodo-L-tyrosine or 3-iodo-L-tyrosine, as indicated below each panel. The enzyme surface and section are shown in light blue. The iodine atom is in purple.

Table 1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Data collection</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P3_1,21</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>83.16, 83.16, 93.67</td>
</tr>
<tr>
<td>α, β, γ(°)</td>
<td>90, 90, 120</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.0000</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50–1.8</td>
</tr>
<tr>
<td>Measured reflections</td>
<td>382,492</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>35,502</td>
</tr>
<tr>
<td>Redundancy</td>
<td>10.8</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100.0</td>
</tr>
<tr>
<td>I/σ (I)</td>
<td>23.0</td>
</tr>
<tr>
<td>Refinement statistics</td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>39.26–1.8</td>
</tr>
<tr>
<td>Protein atoms</td>
<td>2536</td>
</tr>
<tr>
<td>Substrates (per subunit)</td>
<td>1</td>
</tr>
<tr>
<td>Water oxygens</td>
<td>295</td>
</tr>
<tr>
<td>Mean B value (Å²)</td>
<td>21.3</td>
</tr>
<tr>
<td>Rwork (%)</td>
<td>19.9</td>
</tr>
<tr>
<td>Rfree (%)</td>
<td>24.7</td>
</tr>
<tr>
<td>Bond length (Å)</td>
<td>0.011</td>
</tr>
<tr>
<td>Bond angle (°)</td>
<td>1.40</td>
</tr>
</tbody>
</table>

CONCLUSION

A variety of tyrosine or phenylalanine derivatives have been genetically encoded in mammalian cells and E. coli to facilitate protein science and technology. In mammalian cells, for example, photo-reactive cross-linking amino acid was used for analyzing protein–protein interactions (9), and the incorporated 4-azido-L-phenylalanine allowed the probing of conformational changes in a membrane protein (40). E. coli cells have been useful for producing large amounts of proteins containing non-natural amino acids. The incorporation of 4-azido-L-phenylalanine allowed the site-specific protein modification with polyethylene glycol (41), and iodine-atom-containing amino acids have been utilized in X-ray crystallography (36,42). For each of these non-natural amino acids, specific archaeal and bacterial TyrRS variants both needed to be developed, when the amino acid was used in the bacterial and eukaryotic cells. Our engineered E. coli strains promise to reduce the labor for variant development, because they can be engineered in the E. coli cells, and then used in both eukaryotic cells and bacteria. Our strategy could be applied to other aaRS species, and would facilitate genetic code expansion in living cells.

ACCESSION NUMBER

2YXN.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Dr Nobuhiro Hayashi for kindly providing the rat calmodulin gene, Dr Takamitsu Hosoya and Dr Masaaki Suzuki for kindly providing the first lot of the triarylphosphine-fluorescein conjugate used in this study,
and Ms Azusa Ishii and Ms Tomoko Nakayama for clerical assistance.

FUNDING
Targeted Proteins Research Program (TPRP): RIKEN Structural Genomics/Proteomics Initiative (RSGI) in the National Project on Protein Structural and Functional Analyses, and in part, by a Grant-in-Aid for Scientific Research (B) 19380195 to K.S. from the Ministry of Education, Culture, Sports, Science and Technology of Japan; Special Postdoctoral Researchers Program at RIKEN to T.K. Funding for open access charge: RIKEN.

Conflict of interest statement. None declared.

REFERENCES


