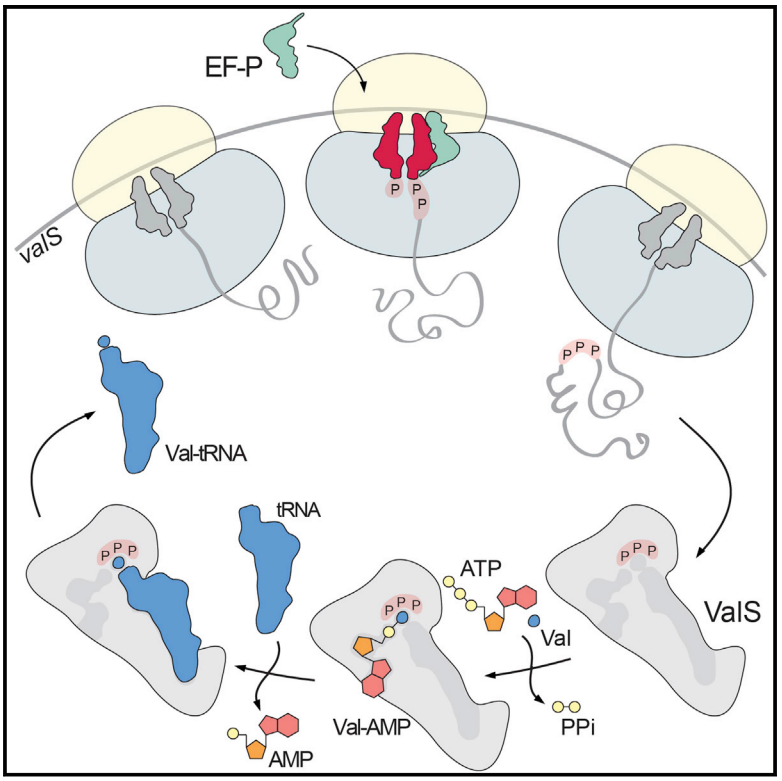
Report

Cell Reports

A Conserved Proline Triplet in Val-tRNA Synthetase and the Origin of Elongation Factor P

Graphical Abstract



Authors

Agata L. Starosta, Jürgen Lassak, ..., Kirsten Jung, Daniel N. Wilson

Correspondence

wilson@lmb.uni-muenchen.de

In Brief

Elongation factor P is required for ribosomes to synthesize polyproline-containing proteins. Here, Starosta et al. reveal that only one polyproline-containing protein is invariant throughout the three domains of life, namely the essential valine tRNA synthetase ValS, suggesting that the invariant proline triplet in ValS may explain the coevolution of EF-P.

Highlights

The only conserved polyproline stretch is present in ValS

The proline triplet in ValS is critical for $tRNA^{Val}$ charging and editing activities

Mutations within the proline triplet of ValS reduce growth and viability of *E. coli*

The invariant proline triplet in ValS may explain the coevolution of EF-P

Starosta et al., 2014, Cell Reports 9, 476–483
October 23, 2014 ©2014 The Authors
http://dx.doi.org/10.1016/j.celrep.2014.09.008

OPEN ACCESS
CellPress

Cell Reports
Report

A Conserved Proline Triplet in Val-tRNA Synthetase and the Origin of Elongation Factor P

Agata L. Starosta,[1] Jürgen Lassak,[2] Lauri Peil,[3,4] Gemma C. Atkinson,[4,5] Christopher J. Woolstenhulme,[6] Kai Virumäe,[7] Allen Buskirk,[6] Tanel Tenson,[4] Jaanus Remme,[7] Kirsten Jung,[2,8] and Daniel N. Wilson[1,8,*]

[1]Gene Center and Department for Biochemistry, University of Munich, Feodor-Lynen-Straße 25, 81377 Munich, Germany
[2]Department of Biology I, Microbiology, Ludwig-Maximilians-Universität München, 82152 Martinsried, Germany
[3]Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH8 9YL, UK
[4]Institute of Technology, University of Tartu, Tartu 50090, Estonia
[5]Department of Molecular Biology and Laboratory for Molecular Infection Medicine Sweden (MIMS), Umeå University, Umeå 901 87, Sweden
[6]Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT 84602, USA
[7]Institute of Molecular and Cell Biology, University of Tartu, Tartu 50090, Estonia
[8]Center for Integrated Protein Science Munich (CiPSM) at the University of Munich, 80539 Munich, Germany
*Correspondence: wilson@lmb.uni-muenchen.de
http://dx.doi.org/10.1016/j.celrep.2014.09.008

SUMMARY

Bacterial ribosomes stall on polyproline stretches and require the elongation factor P (EF-P) to relieve the arrest. Yet it remains unclear why evolution has favored the development of EF-P rather than selecting against the occurrence of polyproline stretches in proteins. We have discovered that only a single polyproline stretch is invariant across all domains of life, namely a proline triplet in ValS, the tRNA synthetase, that charges $tRNA^{Val}$ with valine. Here, we show that expression of ValS in vivo and in vitro requires EF-P and demonstrate that the proline triplet located in the active site of ValS is important for efficient charging of $tRNA^{Val}$ with valine and preventing formation of mischarged Thr-$tRNA^{Val}$ as well as efficient growth of *E. coli* in vivo. We suggest that the critical role of the proline triplet for ValS activity may explain why bacterial cells coevolved the EF-P rescue system.

INTRODUCTION

Polymerization of amino acids by ribosomes to form polypeptide chains is a fundamental process in all cells. Ribosomes can polymerize most polypeptide chains without difficulty, but distinct amino acid combinations pose serious problems. For example, three or more consecutive proline residues induce translational arrest by preventing peptide-bond formation (Doerfel et al., 2013; Ude et al., 2013; Woolstenhulme et al., 2013). Ribosome stalling results from the slow rate of peptide-bond formation between the peptidyl-Pro-Pro-tRNA located in the P site and the Pro-tRNA in the A site (Doerfel et al., 2013). Ribosome stalling in the absence of elongation factor P (EF-P) has also been observed at diprolyl motifs (XPPY), with the efficiency of stalling dependent on the nature of the amino acid located before (X) and after (Y) the diprolyl motif (PP) (Hersch et al., 2013; Peil et al., 2013; Woolstenhulme et al., 2013). The translational arrest at PPP and XPPY motifs was in all cases relieved by the presence of the translation elongation factor EF-P (Peil et al., 2013), which binds to the stalled ribosomes and stimulates peptide-bond formation (Doerfel et al., 2013; Ude et al., 2013). A conserved lysine residue of *Escherichia coli* EF-P is subject to posttranslational modification by YjeA, YjeK, and YfcM (EpmA, EpmB, and EpmC) (Navarre et al., 2010; Peil et al., 2012; Yanagisawa et al., 2010), and this lysinylation modification is required for the rescue activity of EF-P (Doerfel et al., 2013; Peil et al., 2013; Ude et al., 2013). The equivalent lysine of the ortholog of EF-P in archaea and eukaryotes, IF-5A (a/eIF-5A), is also post-translationally modified, but via hypusinylation rather than lysinylation (Park et al., 2010). In yeast, the posttranslational modification of eIF-5A is also critical for rescue of ribosomes stalled on polyproline stretches (Gutierrez et al., 2013). Moreover, the associated archaeal and eukaryotic enzymes, deoxyhypusine synthase (DHS) and deoxyhypusine hydroxylase (DOHH), are evolutionarily unrelated to their bacterial counterparts (Park et al., 2010). Thus, nature has evolved not only specialized translation factors to overcome stalling at polyproline stretches but also independent sets of modification enzymes to activate these factors. This in itself implies that the benefits of retaining polyproline stretches significantly outweigh the cost of implementing and maintaining the EF-P and a/eIF5A rescue systems.

RESULTS

Conservation of Polyproline Stretches

In order to understand which proteins have in fact retained a polyproline stretch (three or more consecutive prolines) throughout evolution, we initially examined the number and conservation of polyproline-containing proteins across 1,273 completely sequenced bacterial genomes. As seen in Figure 1A, the number of polyproline-containing proteins varies across the different bacterial phyla and is generally higher in bacteria with larger genomes, for example, the delta-proteobacterium *Sorangium cellulosum*, which has the largest sequenced genome to date (9,367 protein-coding genes [CDS]) (Schneiker et al.,

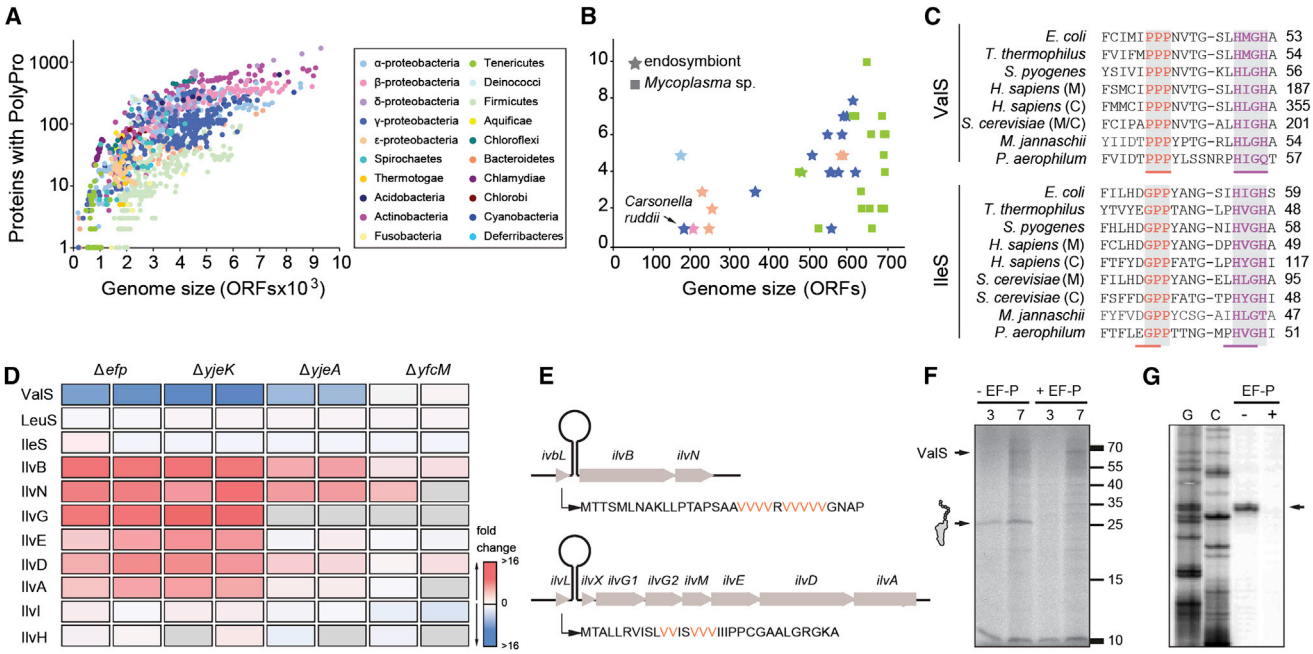


Figure 1. A Conserved Proline Triplet in ValS Confers EF-P Dependence for Expression

(A) Scatterplot illustrating the total number of polyproline-containing proteins relative to genome size (in terms of open reading frame [ORF] number) of species within different bacterial phyla.

(B) Zoom of (A) highlighting bacteria with the smallest genomes.

(C) Sequence alignment illustrating the conservation of PPP and GPP in Val- and Ile-tRNA synthetases, respectively, across the three domains of life, compared with the highly conserved HIGH motif.

(D) Heatmap representation of changes in expression levels in vivo (blue, downregulated; red, upregulated; white, unchanged; gray, not identified or quantified) of selected proteins when comparing Δ*efp*, Δ*yjeA*, Δ*yjeK*, and Δ*yfcM* strains against a wild-type *E. coli* internal standard. Data from two biological replicates were taken from Peil et al. (2013).

(E) Schematic illustration of the gene arrangement of the *ilvL-BN* and *ilvL-GMEDA* operons, with the corresponding sequence of the leader peptides IvbL and IlvL. Valine residues essential for attenuation are shown in red.

(F) Autoradiograph of SDS-PAGE monitoring the in vitro translation at 3 min and 7 min of ValS, in the absence (−EF-P) and presence of EF-P (+EF-P), with full-length (FL) ValS and peptidyl-tRNA (25 kDa) as indicated.

(G) Toeprinting of ValS in the absence (−) and presence (+) of active EF-P. G and C sequence lanes are included.

2007), has the most polyproline-containing proteins (i.e., 2,406 occurrences in 1,779 of the 9,367 proteins, or 19%). By comparison, a typical K12 *E. coli* strain has ~100 polyproline-containing proteins from a total of ~4,100 CDS (2%). Curiously, we identified 19 bacterial genomes ranging in size from 182 to 2,013 CDS encoding only a single polyproline-containing protein (0.05%–0.5%) (Figures 1A and 1B). These include both free-living bacteria of the *Mycoplasma* genus as well as many obligate endosymbiotic bacteria, such as *Carsonella ruddii*, which cohabit with sap-feeding insects, such as aphids, psyllids, and cicadas (Figure 1B). In all cases, the single polyproline-containing CDS was identified as ValS (*E. coli* nomenclature), the Val-tRNA synthetase that aminoacylates $tRNA^{Val}$ with the amino acid valine. The *valS* gene, encoding ValS, is an essential gene (Gerdes et al., 2003) and is correspondingly present in all domains of life. The presence and location of the proline triplet in ValS is invariant not only in bacteria but also in the 205 archaeal and 98 eukaryotic genomes that we analyzed (Figure 1C). In fact, the proline triplet has a higher conservation than the "HIGH" motif that defines ValS as a class I tRNA synthetase (Arnez and Moras, 1997). No conserved polyproline stretches longer than a triplet were identified. The protein with the next-highest conservation of a polyproline stretch was the translational GTPase LepA, which contains a proline triplet in ~65% of the 1,273 sequenced genomes, and, unlike *valS*, the *lepA* gene is not essential for viability (Baba et al., 2006). Given that XPPY motifs other than PPP also cause stalling (albeit at lower levels than PPP), which is relieved by EF-P (Peil et al., 2013), we also analyzed the bacterial genomes for conservation of XPPY motifs. We detected only one diprolyl motif as being correspondingly conserved in all domains of life, namely, the GPP motif found within IleS, the Ile-tRNA synthetase that aminoacylates $tRNA^{Ile}$ with the amino acid isoleucine. The presence and location of the GPP motif in IleS is homologous to the PPP motif in ValS, also being located adjacent to the HIGH motif (Figure 1C). We have previously demonstrated that ribosomes stall only weakly at GPP motifs in the absence of EF-P, whereas much stronger stalling is observed at PPP motifs (Peil et al., 2013).

ValS Expression Is Dependent on Modified EF-P

To show that active EF-P is in fact required for ValS expression, we analyzed our previous stable isotope labeling and mass

spectrometry data that monitored the changes in expression levels of proteins in *E. coli* strains lacking the *efp*, *yjeK*, *yjeA*, and *yfcM* genes relative to the parental wild-type *E. coli* strain (Peil et al., 2013). As predicted, ValS was heavily downregulated when EF-P was absent (Δ*efp*) or EF-P was inactive because of lack of lysinylation (Δ*yjeK*/Δ*yjeA*) (Figure 1D). The lack of YfcM (Δ*yfcM*), which hydroxylates the lysinylated EF-P (Peil et al., 2012), did not lead to downregulation of ValS, consistent with previous reports that lysinylated, but not hydroxylated, EF-P still retains the ability to stimulate peptide-bond formation and thus relieve stalling at polyproline stretches (Doerfel et al., 2013; Peil et al., 2013; Ude et al., 2013). The downregulation of ValS in the absence of active EF-P is specific, since downregulation of other tRNA synthetases, such as LeuS or IleS, is not observed (Figure 1D). The lack of dependence of IleS expression on EF-P is consistent with our previous findings that the GPP motif only induces weak translational stalling compared to PPP (2-fold versus 20-fold, respectively) (Peil et al., 2013).

Notably, we also observed upregulation of the proteins encoded in the *ilvBN* and *ilvGEDA* operons involved in the biogenesis of isoleucine, leucine, and valine. This can be explained by transcriptional attenuation of the *ilvBN* and *ilvGEDA* operons (Yanofsky, 1987). When amino acids such as valine are limiting, the reduction in Val-tRNA^{Val} causes ribosomes to stall during translation of the upstream valine-rich leader peptides IvbL and IlvL. This in turn prevents formation of a downstream transcription termination structure, leading to induction of expression of the downstream genes in the operon (Yanofsky, 1987) (Figure 1E). Similarly, a reduction in charged Val-tRNA^{Val} due to the downregulation of ValS would therefore explain the upregulation of *ilvBN* and *ilvGEDA* operons in the absence of active EF-P (Figure 1D). In contrast, the absence of active EF-P had little effect on the expression of the *ilvI* and *ilvH*, which are not attenuated, indicating that the upregulation of the *ilvBN* and *ilvGEDA* operons is specific. Although the *efp* gene is not essential in *E. coli* (Baba et al., 2006; Peil et al., 2012, 2013), the loss of EF-P leads to a reduction in growth rate (Yanagisawa et al., 2010), presumably due to the reduction in levels of important cellular polyproline-containing proteins. Since we have demonstrated that the lack of EF-P leads to a strong downregulation of ValS (Figure 1D), we tested and could show that expression of ValS could partially rescue the growth defect due to lack of the *efp* gene in *E. coli* (Figure S1).

We also demonstrated that ValS expression is dependent on EF-P in vitro (Figure 1F): In the absence of EF-P, translation of ValS in vitro leads to accumulation of a ~25 kDa band, which correlates with that of the ValS nascent polypeptide chain translated up to the proline triplet (~5 kDa), but still remaining attached to a tRNA (~20 kDa tRNA). As expected, the band for the stalled peptidyl-tRNA is absent when EF-P is included in the in vitro translation reaction (Figure 1F). Toeprinting confirmed that when EF-P is absent, the ribosome stalls at the proline triplet of ValS, specifically with the second proline codon of the proline triplet located in the P site (Figure 1G). Collectively, these findings demonstrate that expression of ValS is strictly dependent on the presence of active EF-P in vivo and in vitro.

The Proline Triplet Is Located at the Active Site of ValS

If the EF-P rescue system really evolved to allow the proline triplet in ValS to be retained, this would imply that the proline triplet in ValS is critically important for function, as indicated by its universal conservation. Crystal structures of *T. thermophilus* ValS reveal that the proline triplet (P41–P43, *E. coli* numbering) resides in the active site of ValS, namely where valine is activated by ATP to form valyl-AMP and then transferred to tRNA^{Val} to form Val-tRNA^{Val} (Figure 2A) (Fukai et al., 2000, 2003). The side chain of Pro41 of ValS contacts the γ1-CH_3 group of valyl-AMP, whereas Pro42 and Pro43 appear to also contribute to the stable binding of valyl-AMP by positioning the α-CO group of Pro42 to hydrogen bond with the α-NH_3 group of valyl-AMP (Figure 2A) (Fukai et al., 2000, 2003). In IleS, the presence of Gly45, rather than Pro, allows accommodation of isoleucyl-AMP (Figure 2B) (Nureki et al., 1998; Silvian et al., 1999), consistent with the high conservation of the GPP motif in Ile-tRNA synthetases (Figure 1C). In contrast, accommodation of isoleucyl-AMP in ValS is less favorable due to the close distance between Pro41 and the γ1-CH_3 group of isoleucyl-AMP (Figure 2C), which led to the proposed role of Pro41 in ValS in the discrimination between valine and isoleucine (Fukai et al., 2000). In a simplified model, one might predict that mutation of P41G in ValS should increase the level of misincorporation of isoleucine (Figure 2D); however, this has not been tested.

Critical Importance of the Proline Triplet for ValS Charging Activity

To experimentally validate the role of Pro41 as well as the general functional importance of the proline triplet for ValS activity, we generated the three single ValS Pro-to-Gly mutants (ValS-GPP, -PGP and -PPG) as well as a triple PPP-to-GGG mutant (ValS-GGG). The ability of the ValS mutants to charge tRNA^{Val} with [^{14}C]valine was assessed and compared to wild-type ValS (Figure 2E). Under our reaction conditions at 37°C, maximal charging of tRNA^{Val} with [^{14}C]valine by wild-type ValS occurred within the first minute (Figure 2E). In contrast, all ValS mutants were less efficient than wild-type; the ValS-PPG and -GGG mutants were completely devoid of activity, whereas the ValS-PGP and -GPP mutants retained some activity but at lower levels than the wild-type ValS (Figure 2E).

Charging of tRNA^{Val} by ValS occurs in two steps. First, valine is activated with ATP to form Val-AMP, leading to the release of pyrophosphate (Figure 2F). Second, the valine is then transferred to the tRNA^{Val} to form Val-tRNA^{Val}, with a concomitant release of AMP (Figure 2G). Initially, we employed thin-layer chromatography (TLC) and γ[^{32}P]-ATP to monitor the release of pyrophosphate (γ[^{32}P]-PP_i) by wild-type ValS, or the mutant ValS-GPP, in the presence of valine, and in the presence or absence of tRNA^{Val} (Figure 2F). Pyrophosphate (PPi) release is monitored indirectly by treatment with pyrophosphatase that converts the PPi to inorganic monophosphate (Pi), which is then visualized directly by TLC. As seen in Figure 2F, Pi formation is observed only for wild-type ValS in the presence of the pyrophosphatase and, as expected, is stimulated by the presence of tRNA^{Val}. The migration position of Pi on the TLC was determined using a control reaction of γ[^{32}P]-ATP treated with Apyrase, an ATP diphosphatase that hydrolyzes ATP

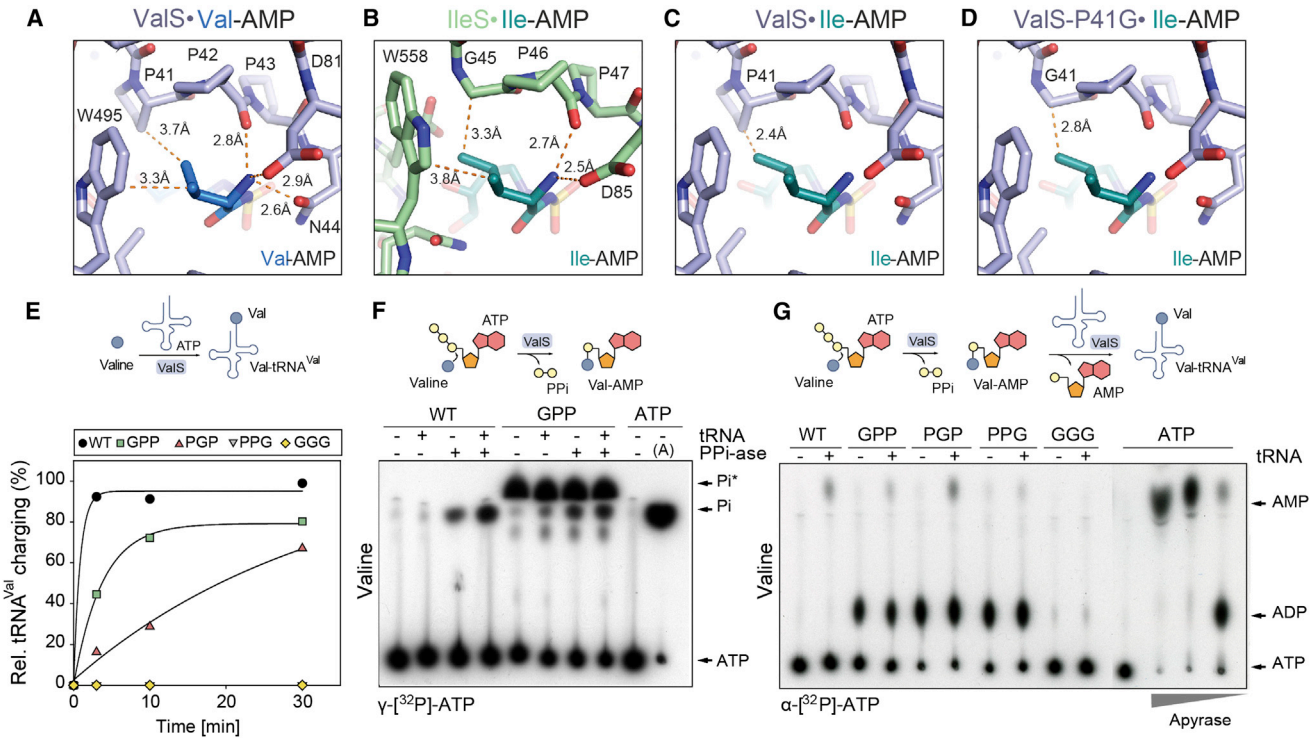


Figure 2. PPP at the Active Site of ValS Is Required for Efficient tRNA Charging

(A) Location of PPP in the active site of ValS relative to Val-AMP (Protein Data Bank [PDB] ID 1GAX) (Fukai et al., 2000).

(B) Active site of IleS with Ile-AMP (1JZQ) (Nakama et al., 2001).

(C) ValS from (A) but with superimposed position of Ile-AMP from (B).

(D) Experiments performed as in (C) but with in silico Pro41Gly mutation.

(E) Charging efficiency of $tRNA^{Val}$ with valine by wild-type (WT) ValS and ValS mutants as a function of time (min).

(F and G) Autoradiograph of TLC separation of (F) $\gamma[^{32}P]$-P_i from $\gamma[^{32}P]$-ATP, and (G) $\alpha[^{32}P]$-AMP and $\alpha[^{32}P]$-ADP from $\alpha[^{32}P]$-ATP, when wild-type ValS (WT), or ValS mutant(s) were incubated with valine, in the absence (−) and presence (+) of pyrophosphatase (PPi-ase) and/or deacylated $tRNA^{Val}$. ATP at the origin indicates where the samples were loaded onto the TLC plate. In (F), the migration position of Pi is determined by treatment of $\gamma[^{32}P]$-ATP with Apyrase (A), and in (G), the migration of AMP and ADP was determined by treatment of $\alpha[^{32}P]$-ATP with increasing concentrations of Apyrase.

sequentially to yield AMP and Pi. Surprisingly, the ValS-GPP mutant was observed to produce large quantities of $\gamma[^{32}P]$-containing material that migrated at a position similar to but distinct from Pi (termed Pi*). Moreover, formation of Pi* by the ValS-GPP mutant was independent of pyrophosphatase treatment or the presence of $tRNA^{Val}$ (Figure 2F). Pi* was also produced by the ValS-PGP and ValS-PPG mutants in a pyrophosphatase- and $tRNA^{Val}$-independent fashion, whereas the ValS-GGG mutant was virtually inactive (Figure S2A). Based on the retention factor (Rf) of Pi* (Rf 0.69), we can conclude that Pi* is not AMP (Rf 0.83), ADP (Rf 0.17), ATP (Rf 0), or monophosphate Pi (Rf 0.56). Furthermore, Pi* cannot be PPi, since PPi is not observable on our TLCs and addition of pyrophosphatase does not influence the migration of Pi* (Figure 2F). PPi generated by tRNA synthetases has been reported to attack ATP and ADP to generate diadenosine polyphosphates, such as Ap^3A and Ap^4A via phosphorolysis (Plateau and Blanquet, 1976); however, we can exclude that Pi* is Ap^3A or Ap^4A, since these compounds are resistant to phosphatase treatment whereas Pi* is not (Figure S2B). Further experiments will be required to determine the exact nature of Pi*.

Since the ValS mutants produce Pi*, rather than PPi, we reasoned that the ValS mutants may also catalyze a reaction other than canonical ATP to AMP, as catalyzed by a ValS-tRNA synthetase. To address this, charging assays were performed as before but using $\alpha[^{32}P]$-ATP instead of $\gamma[^{32}P]$-ATP. Each reaction contained the amino acid valine, together with either wild-type ValS or one of the ValS mutants, and was performed in the absence or presence of $tRNA^{Val}$ (Figure 2G). As expected, wild-type ValS produced AMP only in the presence of $tRNA^{Val}$. AMP formation was also observed for the ValS-GPP, PGP, and, to a lesser extent, the PPG mutant, consistent with some degree of charging activity observed in Figure 2E, whereas ValS-GGG was virtually inactive. Strikingly, unlike wild-type ValS, the ValS-GPP, -PGP, and -PPG mutants converted large amounts of $\alpha[^{32}P]$-ATP to $\alpha[^{32}P]$-ADP (Figure 2G). The migration position on the TLC of $\alpha[^{32}P]$-ADP (and $\alpha[^{32}P]$-AMP) was determined using a control reaction where $\alpha[^{32}P]$-ATP was treated with increasing concentrations of Apyrase that hydrolyzes sequentially ATP to ADP and then AMP and Pi. Furthermore, the generation of $\alpha[^{32}P]$-ADP by the ValS-GPP, -PGP and -PPG mutants was independent of the presence of $tRNA^{Val}$

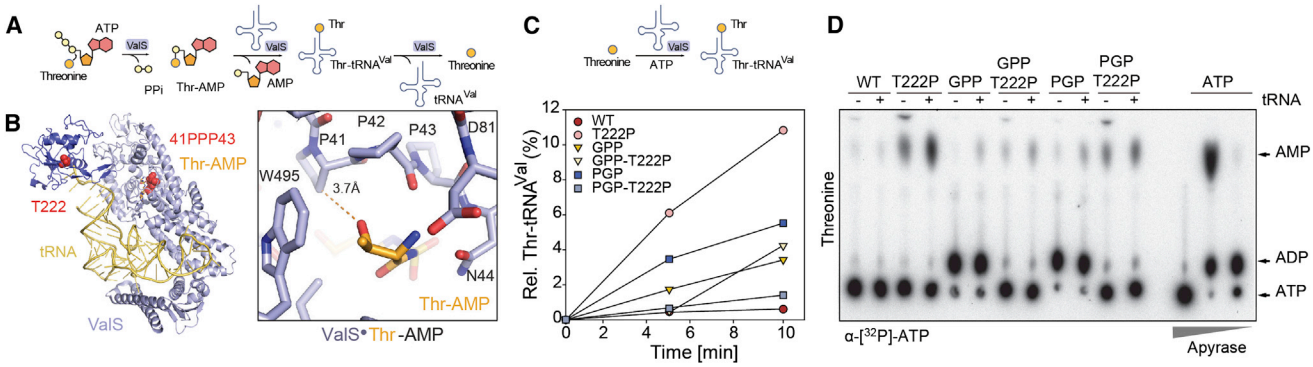


Figure 3. PPP at the Active Site of ValS Is Required to Prevent Thr-tRNA^{Val} Formation

(A) Scheme for misactivation and mischarging of tRNA^{Val} with threonine and subsequent deacylation by ValS.

(B) Overview of ValS bound with tRNA (yellow) and Thr-AMP (orange) (PDB 1GAX) (Fukai et al., 2000). PPP and T222 are indicated by red spheres. Inset shows the active site of ValS with PPP relative to Thr-AMP.

(C) Mischarging of tRNA^{Val} with threonine by wild-type (WT) ValS and ValS mutants as a function of time (min).

(D) Autoradiograph of TLC separation of α[^{32}P]-AMP and α[^{32}P]-ADP from α[^{32}P]-ATP, when wild-type ValS (WT) or ValS mutant(s) were incubated with threonine, in the absence (−) and presence (+) of deacylated tRNA^{Val}. ATP at the origin indicates where the samples were loaded onto the TLC plate and apyrase treatment of ATP provides markers for ADP and AMP positions.

and also occurred when no amino acid was present (Figure S2C). Moreover, the α[^{32}P]-ADP generated by the ValS-GPP, -PGP, and -PPG mutants could be converted to α[^{32}P]-AMP by Apyrase (Figure S2D). In summary, these assays suggest that mutations within the conserved proline triplet of ValS not only reduce the efficiency of tRNA charging but also cause the ValS mutants to nonproductively hydrolyze ATP to ADP.

The Proline Triplet Is Important for the Editing Activity of ValS

To test the hypothesis that mutation of Pro41 in ValS distinguishes between valine and isoleucine (Figure 2D), we performed mischarging assays where [^{14}C]isoleucine replaced [^{14}C]valine. No mischarging of tRNA^{Val} with [^{14}C]isoleucine by wild-type ValS, or any of the ValS mutants, was observed (data not shown). Furthermore, introduction of T222P mutation that disables the ValS editing function did not promote Ile-tRNA^{Val} formation (data not shown), consistent with Ile not being a substrate for editing by ValS (Tardif et al., 2001). Similarly, isoleucine did not stimulate AMP formation by ValS-GPP (Figure S2E), collectively suggesting that Pro41 is not the only residue in ValS involved in discrimination of valine from isoleucine. On the other hand, threonine, which is isosteric with valine, is readily activated by ValS and transferred to tRNA^{Val}, but the mischarged Thr-tRNA^{Val} that forms is rapidly deacylated by the editing domain of ValS (Fersht and Kaethner, 1976) (Figures 3A and 3B). Indeed, in mischarging assays where [^{14}C]threonine replaced [^{14}C]valine, we observed a low rate of Thr-tRNA^{Val} formation by wild-type ValS, which was enhanced when the editing domain was disabled by the T222P mutation (Döring et al., 2001) (Figures 3B and 3C). Formation of Thr-tRNA^{Val} was also enhanced by the ValS-GPP and -PGP mutations. The introduction of the T222P mutation in ValS-GPP did not further enhance mischarging, indicating that the GPP mutation may disable the editing function of ValS (Figure 3C). Surprisingly, mischarging was suppressed when the T222P was introduced into the ValS-PGP mutant (Figure 3C), suggesting an intricate communication exists between the activation and editing domains of ValS. This interplay was also observed in the TLC analyses measuring the misactivation of threonine, namely, in that the nonproductive ADP formation by the ValS-GPP and -PGP mutants was suppressed in the presence of the additional T222P mutation (Figure 3D). Collectively, the in vitro assays indicate that the conserved proline triplet of ValS is important not only for efficient tRNA charging but also for communication with the editing domain to ensure efficient deacylation of mischarged Thr-tRNA^{Val}.

The Proline Triplet of ValS Is Important for Viability of *E. coli*

Given the defects of the ValS mutants in vitro, we assessed their functionality in vivo using a genetic complementation system (Figure 4A). The Δ*valS E. coli* strain was complemented with a plasmid bearing ValS containing an N-terminal 6xHistidine (6xHis) tag and small ubiquitin-like modifier (SUMO) tag (Hay, 2005) followed by a WFCWS linker. The resulting JL001 strain (Table S1) was viable but exhibited a slightly reduced growth rate, presumably because of the lower expression level of the plasmid-encoded ValS (Figure 4B). Next a plasmid expressing Ulp1 (pUlp1) was introduced into JL001. Ulp1 encodes a SUMO-specific protease that cleaves the SUMO tag from ValS exposing the N-terminal WFCWS sequence motif that, in accordance with the N-end degradation rule, leads to recognition and degradation of the ValS protein by the ClpAP protease (Dougan et al., 2010; Wang et al., 2007) (Figure 4A). Under the conditions used, western blotting against the 6xHis tag indicated a rapid loss of ValS protein in the cell within minutes of induction of Ulp1 expression (Figure S3). Consistently, induction of Ulp1 expression also led to strong reduction in growth rate of the JL001 strain (Figure 4B). In the next step, we repeated the growth experiments with the JL001 strain bearing pUlp1 (or a control plasmid) together with an additional plasmid bearing an IPTG-inducible copy of either the wild-type ValS (pValS-PPP; Figure 4C)

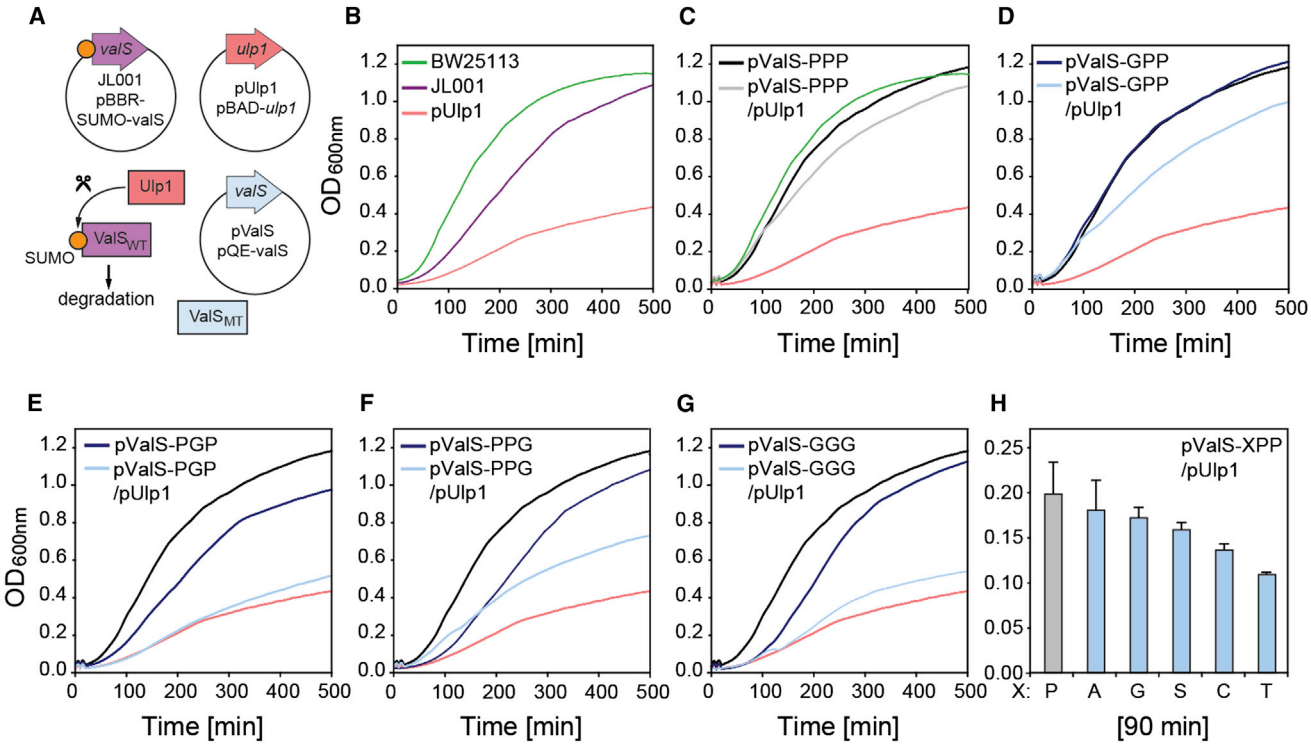


Figure 4. The Proline Triplet of ValS Is Required for Viability of *E. coli*

(A) Scheme for the ValS complementation system.

(B–G) Growth curves (OD_{600nm}) of *E. coli* strains (B) BW25113, JL001, JL001 + pUlp1, or (C–G) JL001 bearing additional plasmids expressing pUlp1, and/or (C) wild-type ValS (pValS-PPP) or ValS mutants (D) pValS-GPP, (E) pValS-PGP, (F) pValS-PPG, or (G) pValS-GGG.

(H) Comparison of growth (after 90 min) of JL001 bearing plasmids expressing pUlp1 and pValS/XPP mutants normalized by subtraction of background growth in the presence of pUlp1 alone. The error bars represent the SD from the mean of three biological replicates.

or one of the ValS mutants (ValS-GPP, -PGP, -PPG, or -GGG; Figures 4D–4G). As expected, the presence of an additional copy of the wild-type ValS (pValS-PPP) could rescue the growth of JL001 following the degradation of the SUMO-ValS that occurs upon induction of Ulp1 expression (Figure 4C). Similarly, pValS-GPP was also able to rescue growth upon Ulp1 expression (Figure 4D), albeit not as efficiently as pValS-PPP. By contrast, the pValS-PGP, -PPG, and -GGG displayed little or no ability to rescue growth in the absence of wild-type ValS (Figures 4E–4G). Unfortunately, attempts to transform plasmid libraries of pValS-XXX mutants to select for viable mutations of the proline triplet failed; however, given the ability of the ValS-GPP mutant to rescue, we therefore generated all possible pValS-XPP mutants and monitored their ability to rescue growth in the absence of wild-type ValS (Figure 4H). Of the 20 pValS-XPP mutants tested, six mutants (PPP, GPP, APP, SPP, CPP, and TPP) rescued growth in the absence of wild-type ValS, with only APP and SPP being able to sustain growth at levels comparable to PPP and GPP (Figure 4H).

DISCUSSION

We identified only a single polyproline stretch that is invariant across all bacterial genomes, namely a proline triplet present in ValS, the Val-tRNA synthetase (Figures 1B and 1C), and demonstrated that efficient expression of ValS in vivo and in vitro requires the presence of EF-P. Since the PPP triplet in ValS is invariant in bacteria, archaea, and eukaryotes (Figure 1C), this leads us to propose that EF-P and a/eIF-5A may have initially (co)evolved to facilitate primarily the expression of ValS. We cannot, however, exclude that other polyproline-containing proteins contributed to the evolution of EF-P or that EF-P and a/eIF-5A evolved to facilitate expression of another protein that contained a polyproline-stretch at the time but that was subsequently lost from some or all of the currently available bacterial genomes.

Regardless of whether EF-P and a/eIF-5A evolved to facilitate expression of ValS or not, our study nevertheless identified a highly conserved PPP motif located in the active site of ValS (Figure 2A) that is critical for the tRNA charging activity (Figure 2E) and productive ATPase activity (Figures 2F and 2G) of ValS. Although comparisons of ValS and IleS led to the suggestion that Pro41 of ValS is important for the discrimination of valine from isoleucine (Fukai et al., 2000), we find no evidence to support this. Instead, we find that mutations within the PPP motif of ValS enhance formation of mischarged Thr-tRNA^{Val} (Figure 3C). Moreover, we could also demonstrate that the T222P mutation within the editing site of ValS could suppress the nonproductive formation of ADP by the ValS mutants (Figure 4D) and in the case of ValS-PGP was even able to rescue the editing

activity (Figure 4C). Collectively, these findings suggest an intimate communication between the aminoacylation and editing sites on ValS that warrants further study.

Consistent with our in vitro assays, the ValS-GPP mutant was able to rescue growth and viability of an *E. coli* strain lacking wild-type *valS* gene, but not as efficiently as ValS-PPP, whereas ValS-PGP, -PPG, and -GGG displayed little or no rescue phenotype. Testing of all possible ValS-XPP mutants demonstrated that it is possible to select for alternative motifs, such as APP or SPP, that support viability of *E. coli* under optimal conditions; however, it remains to be tested whether such mutants are also impaired in tRNA charging activities as observed for the ValS-GPP mutant.

The critical importance of the proline triplet of ValS for tRNA charging and viability of *E. coli* would provide an explanation for why nature needed to coevolve EF-P and a/eIF-5A to facilitate expression during translation of the ValS.

EXPERIMENTAL PROCEDURES

Oligonucleotides, Plasmids, and Bacterial Strains

Primers, plasmids, and strains used in this study are listed in Table S1.

Bioinformatic Analysis

To search for ValRS protein sequences and determine the conservation of the PPP motif, 1,273 bacterial, 205 archaeal, and 98 eukaryotic genomes were searched with a hidden Markov model (HMM) profile (Eddy, 1998) with an E value cutoff of e-70 as described previously. The HMM was generated from a MAFFT (Katoh et al., 2005) alignment of ValRS homologs identified in an initial BlastP search (Altschul et al., 1990). As the HMM also hits the close relative Ile-tRNA synthetase, phylogenetic analysis using FastTree (Price et al., 2010) was carried out to extract sequences from the ValRS clade of orthologs. The resulting sequences were aligned, and the PPP motif region was inspected for conservation.

$tRNA^{Val}$ Charging Assays

Charging reactions and assays to monitor aminoadenylate formation were performed as described previously (Splan et al., 2008).

Growth Conditions

Lysogeny broth (LB) was used as complex medium (Bertani, 1951) and modified with NaCl. When indicated, LB was supplemented with 0.4% (w/v) glucose as a carbon source to suppress Ulp1 expression before induction with L-arabinose 0.2% (w/v). A total of 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) (Sigma-Aldrich) was used to induce expression of pQE70-ValS (pValS-PPP) and GPP, PGP, PPG, and GGG variants. Antibiotics were used when necessary with the following concentrations: 100 μg/ml ampicillin sodium salt, 50 μg/ml kanamycin sulfate, 34 μg/ml chloramphenicol, 20 μg/ml gentamycin sulfate.

ValS Silencing

E. coli BW2113 was used to delete *valS* via pRED/ET recombination technology. Due to the fact that *valS* is an essential gene, gene deletion was performed in a strain containing plasmid pBBR1MCS-5-PT7-SUMO-WFCWS-valS encoding SUMO-ValS (ValS*). The resultant strain JL001 was cotransformed with pBAD33 or pBAD33-ulp1, respectively, and pQE70 encoding either wild-type ValS or ValS-GPP, -PGP, -PPG, or -GGG mutants. Cells were plated on LB agar containing chloramphenicol and ampicillin for transformant selection as well as IPTG to induce pQE70-ValS expression and glucose to suppress pBAD activity by catabolite repression. Transformants were precultured overnight in LB + glucose + chloramphenicol (pBAD33) + gentamycin sulfate (pBBR1MCS-5) + ampicillin sodium salt (pQE70) + kanamycin sulfate (ΔvalS::kan). A second preculture was started by diluting the overnight culture 1:100. Cells were grown to an optical density of 600 nm (OD_{600nm}) = 0.3–0.7. Subsequently, cells were washed twice in LB + 0.2% (w/v) arabinose and adjusted to an OD_{600nm} = 0.1. The main culture was started by a further dilution step to OD_{600nm} = 0.01 in a microtiter plate in a final volume of 200 μl LB + 0.2% arabinose + chloramphenicol. Growth was recorded in a Tecan reader (Infinite M200 Pro) by measuring OD_{600nm} every 10 min for 8 hr. In parallel, Ulp1 production and ValS degradation was monitored by western blot analysis. Cells of the JL001 strain harboring pBAD33-ulp1 and pBBR1MCS-5-PT7-SUMO-WFCWS-valS were harvested at different times after arabinose induction. As a control, JL001 cells harboring pBAD33 and pBBR1MCS-5-PT7-SUMO-WFCWS-valS were analyzed.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.09.008.

ACKNOWLEDGMENTS

We would like to thank Ingrid Weitl for technical assistance and Michael Graf for help with purification of the ValS mutants. This research was supported by grants from the Deutsche Forschungsgemeinschaft (FOR1805 and WI3285/4-1 to D.N.W., and Exc114/2 to K.J.), the Estonian Science Foundation (ETF9020 to G.C.A), and IUT 14021 (JR). L.P. and G.A. received support from the European Social Fund program Mobilitas grant MJD144 and MJD99, respectively. A.L.S. is funded by an AXA Research Fund postdoctoral fellowship, and L.P. is supported by the Marie Curie FP7-PEOPLE-2011-IEF postdoctoral fellowship.

Received: April 25, 2014
Revised: August 5, 2014
Accepted: September 5, 2014
Published: October 9, 2014

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Arnez, J.G., and Moras, D. (1997). Structural and functional considerations of the aminoacylation reaction. Trends Biochem. Sci. *22*, 211–216.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli K-12* in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. *2*, 0008.

Bertani, G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. J. Bacteriol. *62*, 293–300.

Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M.V. (2013). EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. Science *339*, 85–88.

Döring, V., Mootz, H.D., Nangle, L.A., Hendrickson, T.L., de Crécy-Lagard, V., Schimmel, P., and Marlière, P. (2001). Enlarging the amino acid set of Escherichia coli by infiltration of the valine coding pathway. Science *292*, 501–504.

Dougan, D.A., Truscott, K.N., and Zeth, K. (2010). The bacterial N-end rule pathway: expect the unexpected. Mol. Microbiol. *76*, 545–558.

Eddy, S.R. (1998). Profile hidden Markov models. Bioinformatics *14*, 755–763.

Fersht, A.R., and Kaethner, M.M. (1976). Enzyme hyperspecificity. Rejection of threonine by the valyl-tRNA synthetase by misacylation and hydrolytic editing. Biochemistry *15*, 3342–3346.

Fukai, S., Nureki, O., Sekine, S., Shimada, A., Tao, J., Vassylyev, D.G., and Yokoyama, S. (2000). Structural basis for double-sieve discrimination of L-valine from L-isoleucine and L-threonine by the complex of tRNA(Val) and valyl-tRNA synthetase. Cell *103*, 793–803.

Fukai, S., Nureki, O., Sekine, S., Shimada, A., Vassylyev, D.G., and Yokoyama, S. (2003). Mechanism of molecular interactions for tRNA(Val) recognition by valyl-tRNA synthetase. RNA *9*, 100–111.

Gerdes, S.Y., Scholle, M.D., Campbell, J.W., Balázsi, G., Ravasz, E., Daugherty, M.D., Somera, A.L., Kyrpides, N.C., Anderson, I., Gelfand, M.S., et al. (2003). Experimental determination and system level analysis of essential genes in Escherichia coli MG1655. J. Bacteriol. *185*, 5673–5684.

Gutierrez, E., Shin, B.S., Woolstenhulme, C.J., Kim, J.R., Saini, P., Buskirk, A.R., and Dever, T.E. (2013). eIF5A promotes translation of polyproline motifs. Mol. Cell *51*, 35–45.

Hay, R.T. (2005). SUMO: a history of modification. Mol. Cell *18*, 1–12.

Hersch, S.J., Wang, M., Zou, S.B., Moon, K.M., Foster, L.J., Ibba, M., and Navarre, W.W. (2013). Divergent protein motifs direct elongation factor P-mediated translational regulation in Salmonella enterica and Escherichia coli. MBio *4*, e00180–e13.

Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. *33*, 511–518.

Nakama, T., Nureki, O., and Yokoyama, S. (2001). Structural basis for the recognition of isoleucyl-adenylate and an antibiotic, mupirocin, by isoleucyl-tRNA synthetase. J. Biol. Chem. *276*, 47387–47393.

Navarre, W.W., Zou, S.B., Roy, H., Xie, J.L., Savchenko, A., Singer, A., Edvokimova, E., Prost, L.R., Kumar, R., Ibba, M., and Fang, F.C. (2010). PoxA, yjeK, and elongation factor P coordinately modulate virulence and drug resistance in *Salmonella enterica*. Mol. Cell *39*, 209–221.

Nureki, O., Vassylyev, D.G., Tateno, M., Shimada, A., Nakama, T., Fukai, S., Konno, M., Hendrickson, T.L., Schimmel, P., and Yokoyama, S. (1998). Enzyme structure with two catalytic sites for double-sieve selection of substrate. Science *280*, 578–582.

Park, M.H., Nishimura, K., Zanelli, C.F., and Valentini, S.R. (2010). Functional significance of eIF5A and its hypusine modification in eukaryotes. Amino Acids *38*, 491–500.

Peil, L., Starosta, A.L., Virumäe, K., Atkinson, G.C., Tenson, T., Remme, J., and Wilson, D.N. (2012). Lys34 of translation elongation factor EF-P is hydroxylated by YfcM. Nat. Chem. Biol. *8*, 695–697.

Peil, L., Starosta, A.L., Lassak, J., Atkinson, G.C., Virumäe, K., Spitzer, M., Tenson, T., Jung, K., Remme, J., and Wilson, D.N. (2013). Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. Proc. Natl. Acad. Sci. USA *110*, 15265–15270.

Plateau, P., and Blanquet, S. (1976). Synthesis of Np_nN' (n = 3 or 4) *in vitro* and *in vivo*. In Ap4A and Other Dinucleoside Polyphosphates, A.G. McLennan, ed. (Boca Raton: CRC Press), pp. 81–107.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS ONE *5*, e9490.

Schneiker, S., Perlova, O., Kaiser, O., Gerth, K., Alici, A., Altmeyer, M.O., Bartels, D., Bekel, T., Beyer, S., Bode, E., et al. (2007). Complete genome sequence of the myxobacterium Sorangium cellulosum. Nat. Biotechnol. *25*, 1281–1289.

Silvian, L.F., Wang, J., and Steitz, T.A. (1999). Insights into editing from an ile-tRNA synthetase structure with tRNAile and mupirocin. Science *285*, 1074–1077.

Splan, K.E., Musier-Forsyth, K., Boniecki, M.T., and Martinis, S.A. (2008). In vitro assays for the determination of aminoacyl-tRNA synthetase editing activity. Methods *44*, 119–128.

Tardif, K.D., Liu, M., Vitseva, O., Hou, Y.M., and Horowitz, J. (2001). Misacylation and editing by Escherichia coli valyl-tRNA synthetase: evidence for two tRNA binding sites. Biochemistry *40*, 8118–8125.

Ude, S., Lassak, J., Starosta, A.L., Kraxenberger, T., Wilson, D.N., and Jung, K. (2013). Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. Science *339*, 82–85.

Wang, K.H., Sauer, R.T., and Baker, T.A. (2007). ClpS modulates but is not essential for bacterial N-end rule degradation. Genes Dev. *21*, 403–408.

Woolstenhulme, C.J., Parajuli, S., Healey, D.W., Valverde, D.P., Petersen, E.N., Starosta, A.L., Guydosh, N.R., Johnson, W.E., Wilson, D.N., and Buskirk, A.R. (2013). Nascent peptides that block protein synthesis in bacteria. Proc. Natl. Acad. Sci. USA *110*, E878–E887.

Yanagisawa, T., Sumida, T., Ishii, R., Takemoto, C., and Yokoyama, S. (2010). A paralog of lysyl-tRNA synthetase aminoacylates a conserved lysine residue in translation elongation factor P. Nat. Struct. Mol. Biol. *17*, 1136–1143.

Yanofsky, C. (1987). Operon specific control by transcription attenuation. Trends Genet. *3*, 356–360.