Supporting information

Discovery and investigation of natural editing function against artificial amino acids in protein translation

Jan-Stefan Völler[†,§], Morana Dulic[‡,§], Ulla I. M. Gerling-Driessen [†], Hernan Biava[‖], Tobias Baumann[‖], Nediljko Budisa[‖,*], Ita Gruic-Sovulj[‡,*], Beate Koksch[†,*].

[†]Institute of Chemistry and Biochemistry – Organic Chemistry, Freie Universität Berlin, Takustrasse 3, 14195 Berlin, Germany.

[‖]Department of Chemistry, Technische Universität Berlin, Müller-Breslau-Strasse 10, 10623 Berlin, Germany.

[‡]Department of Chemistry, Faculty of Science, University of Zagreb, Horvatovac 102a, 10000 Zagreb, Croatia.

Table of Contents

1. Supplementary data

1.1 Kinetic values of the amino acid activations

Table S1. Detailed kinetic values of the amino acid activation of cognate and fluorinated substrates by class Ia AARSs . The amino acid numbers refer to those in the main text.

AARS	AA	k_{cat} (s^{-1})	K_M (µM)	k_{cat}/K_M (s^{-1}/mM)	k_{cat}/K_M (rel)
IleRS	Ile	3.1 ± 0.1	7.1 ± 0.8	437 ± 39	1
	1	1.2 ± 0.1	4454 ± 734	0.28 ± 0.01	1/1567
	2	0.95 ± 0.06	5075 ± 603	0.19 ± 0.01	1/2311
	3	1.7 ± 0.1	7403 ± 1623	0.24 ± 0.04	1/1817
	8	0.32 ± 0.01	764 ± 118	0.427 ± 0.055	1/1023
ValRS	Val	3.4 ± 0.1	12 ± 1	282 ± 25	1
	1	0.56 ± 0.03	639 ± 57	0.89 ± 0.03	1/318
	3	1.36 ± 0.04	7418 ± 814	0.19 ± 0.02	1/1519
	8	0.31 ± 0.01	3917 ± 479	0.079 ± 0.006	1/3579
LeuRS	Leu	3.9 ± 0.2	17.6 ± 0.5	223 ± 4	1
	5	0.41 ± 0.01	514 ± 34	0.79 ± 0.02	1/282
	6	0.11 ± 0.01	1.74 ± 0.21	0.065 ± 0.006	1/3430

1.2 AARS binding pockets

We used the reported crystal structure of *Thermus thermophilus* LeuRS[1], *Thermus thermophilus* IleRS[2], *Thermus thermophilus* ValRS[3] and *Escherichia coli* MetRS[4] to visualize the hydrophobic surface of the class Ia AARS binding pockets (**Figure S1**). We conclude, that the hydrophobic surface in the binding pocket of LeuRS (**Figure S1c**) reflects the shape of its cognate substrate leucine while those of IleRS (**Figure S1a**) and ValRS (**Figure S1b**) do not display such high shape complementarity to their cognate substrates, but rather display an overall hydrophobic cavity. Indeed, the sequences of the *Thermus thermophilus* AARSs (IleRS, LeuRS and ValRS) used for the visualization of the hydrophobic surface in the binding pockets have only around 30-40% sequence identity with the *Escherichia coli* AARSs that were used for determination of the activation capacity of the fluorinated amino acids via the ATP-PP_i exchange assay. However, the synthetic active sites are well conserved showing a homology of around 90%. Moreover all AARS residues that we assigned to contribute to the hydrophobic surface (**Figure S1**) in the binding pocket are identical in both organisms (*Escherichia coli* and *Thermus thermophilus)* In the case of

MetRS, the main determinant for discrimination of non-cognate amino acids is likely the particular induced-fit mechanism of the MetRS binding pocket as described by Serre et al.[5] However, we conclude from the crystal structure of *Escherichia coli* MetRS[4] available in the PDB that this MetRS binding pocket displays a well-defined tunnel like shape strongly resembling the cognate substrate Met additionally contributing to the specificity for Met (**Figure S1d**).

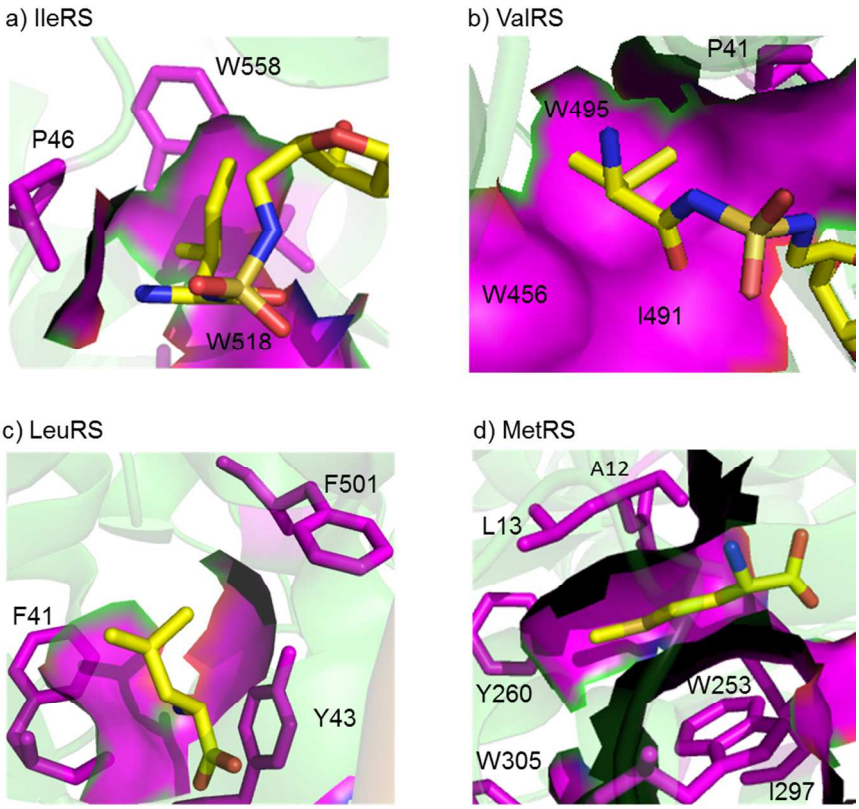

Figure S1: Visualizing the arrangement of the hydrophobic surface (shown in magenta) in the binding pocket of the class I a AARSs, which is a determinant for binding the side chains of the hydrophobic amino acid substrates. a) *Thermus thermophilus* IleRS complexed with isoleucyl-adenylate analogue (N-[isoleucinyl]-N'-[adenosyl]-diaminosufone. b) *Thermus thermophilus* ValRS complexed with valyl-adenylate analogue ((N-[valinyl]-N'-[adenosyl]-diaminosufone. c) *Thermus thermophilus* LeuRS complexed with leucine d) *E. coli* MetRS complexed with methionine. Structures were drawn with PyMOL[6] based on the following PDB entries: 1JZQ (a)), 1GAX (b)), 2V0G (c)), 1PG2 (d)). The given residue numbers are specific to the crystal structures.

1.3. Model Peptide VW18Ile

VW18Ile is a derivative of the previously reported VW18 peptide[7–9], in which the residue valine 14 was changed into the only Ile residue within the sequence (**Figure S2**).

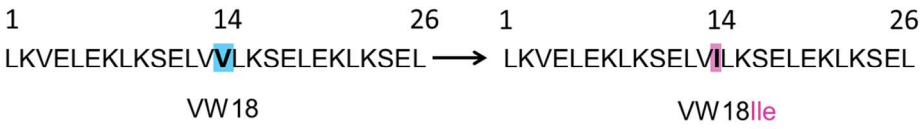

Figure S2: Sequence of VW18 and its derivative VW18Ile.

We have engineered an Ile free SUMO domain, which secures stability and solubility of the VW18 peptide under physiological conditions and can be cleaved off by SUMO protease.

1.4. Expression in the presence of increased cytosolic concentration of IleRS

Artificial amino acids are in general less suitable substrates for the AARSs compared to the cognate substrates. Hence, a unit of the AARS converts less artificial amino acid substrate per time unit compared to the cognate substrate. Increasing the intracellular level of the biocatalyst (AARS), for example, by supplementation of a vector bearing the corresponding gene (**Figure S3a**), can improve the incorporation of artificial amino acids into proteins.[10] We used a lacIPtac promoter to achieve increased IleRS concentration, which was visible in the SDS-PAGE (**Figure S3b**, line: n.i.; +IleRS). However, no translation of TfeGly was observed under these conditions (**Figure S3c**).

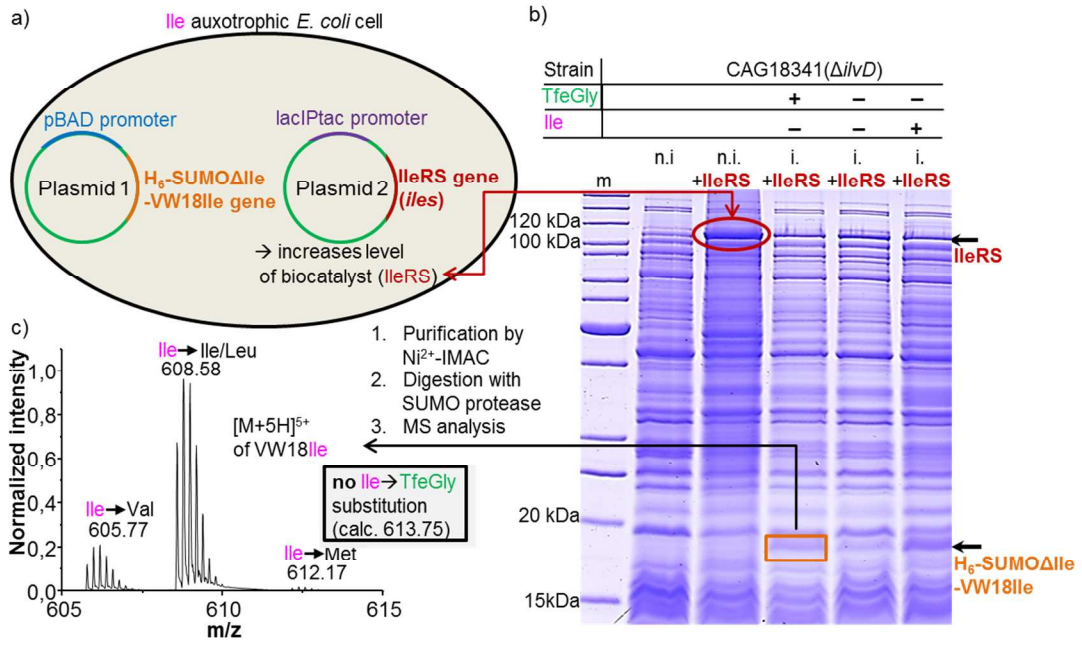

Figure S3: SPI experiments with increased cellular level of AARS (IleRS). a) Schematic representation of a genetic setup in order to increase the amount of the biocatalyst AARS (IleRS). b) SDS-PAGE gel of expression of H_6-SUMOΔIle-VW18Ile in the absence and in the presence of Ile/TfeGly with prior overexpression of IleRS. Mass of IleRS, which was expressed with a StrepTag by induction with IPTG: 105491.0 Da. i.: induced; n.i.: not induced (H_6-SUMOmutΔIle-VW18Ile). c) ESI-MS spectra of VW18Ile expressed in an Ile auxotrophic strain in the presence of 1mM TfeGly in the absence of Ile and subsequently cleaved from the H_6-SUMO domain with SUMO protease. The monoisotopic masses of the 5-fold charged VW18Ile peptide are displayed. Calculated monoisotopic masses for Ile substitution in VW18Ile: Val=605.78, Leu/Ile=608.57, TfeGly=613.75, Met=612.19.

1.5. The post-transfer inactive IleRS mutant IleRSAla10

IleRS post-transfer editing occurs in the CP1 domain, which is distal to the synthetic active site. The IleRS mutant IleRSAla10 was previously reported.[11] In this variant the residues T241-N250 (according to *E. coli* nomenclature) in the CP1 post-transfer editing domain are substituted with alanine residues. The location of the mutation in IleRS CP1 domain is shown in **Figure S4**.

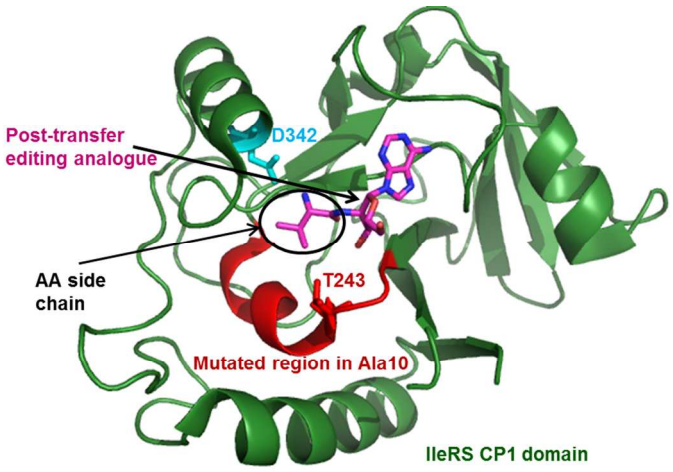

Figure S4: Post-transfer editing analogue 2`-(L-valyl)amino-2´-deoxyadenosine (Val-2AA, magenta) in the CP1 post-transfer editing domain from *Thermus thermophilus* IleRS. The area, in which the residues (T241-N250 according to *E. coli* nomenclature) are located that are substituted by 10 alanines in IleRSAla10[11] is shown in red. Mutation sites of T243R/D342A IleRS are shown as red (T243) and cyan (D342) sticks. Structure was drawn with PyMol[6] based on the PDB entry 1WNZ.

1.6. In vitro aminoacylation and in vivo translation of TfeGly, TfNVal, TfIle, TfVal

In this work we discovered that the post-transfer editing of IleRS hinders TfeGly from being incorporated in place of Ile codons. The post-transfer editing activity towards TfeGly was also reflected by the *in vivo* translation and the increased *in vitro* aminoacylation rate mediated by the post-transfer inactive IleRS mutant IleRSAla10 compared to wild-type IleRS. We further investigated these processes (**Figure S6**, **Figure S7**) for other fluorinated AAs, namely TfNVal, TfVal and TfIle (**Figure S5**) in order to probe if IleRS post-transfer editing may hamper or reduce the translation of those AAs, too.

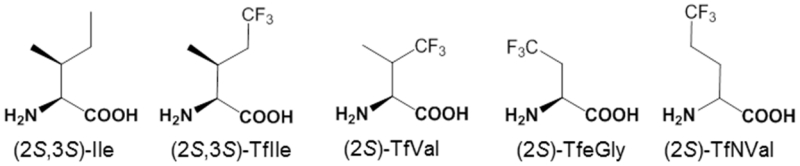

Figure S5: Chemical structure of Ile and fluorinated AA analogues tested in this study regarding *in vitro* aminoacylation and *in vivo* translation using IleRS.

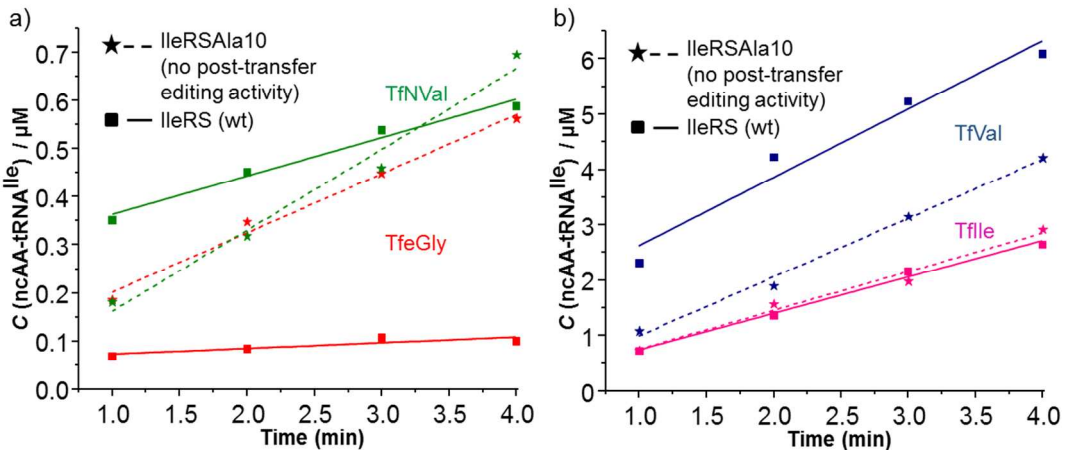

Figure S6: Steady-state aminoacylation of tRNA^{Ile} with a) TfeGly ($k_{aminoacylation}$(IleRS wt)<(0.00007±0.00001) s^{-1}, $k_{aminoacylation}$(IleRSAla10)=(0.0021±0.0001) s^{-1}) andTfNVal ($k_{aminoacylation}$(IleRS wt)=(0.0013±0.0001) s^{-1}, $k_{aminoacylation}$(IleRSAla10)=(0.0028±0.0002) s^{-1}) and b) TfVal ($k_{aminoacylation}$(IleRS wt)=(0.021±0.003) s^{-1}, $k_{aminoacylation}$(IleRSAla10)=(0.018±0.001) s^{-1}) and TfIle ($k_{aminoacylation}$(IleRS wt)=(0.011±0.001) s^{-1}, $k_{aminoacylation}$(IleRSAla10)=(0.012±0.001) s^{-1}) by IleRS and IleRSAla10, respectively.

The aminoacylation rate of tRNA^{Ile} with TfNVal was increased by employing the post-transfer inactive mutant IleRSAla10 as we observed for TfeGly indicating a post-transfer editing activity against TfNVal that is likely not as effective as against TfeGly. This effect was not observed for TfIle and TfVal, indicating that both amino acids are not proofread by post-transfer editing. The latter observation may result from a possible rejection of TfVal and TfIle from the IleRS CP1 editing domain by steric clashes involving their beta-branched side chains.[12] We further investigated if these observations are reflected by *in vivo* translation rates. *In vivo* translation experiments of the fluorinated AAs were carried out by expressing H_6-SUMOΔIle-VW18Ile in the absence of Ile, but in the presence of 1 mM fluorinated AA and either the wild-type IleRS or the post-transfer editing defective IleRSAla10. Subsequently, H_6-SUMOΔIle-VW18Ile was purified and VW18Ile was cleaved from H_6-SUMOΔIle. ESI-MS was used to determine which AAs had been translated in response to the single Ile codon in VW18Ile. The deletion of the post-transfer editing leads to increased mistranslation of the canonical amino acids Val and Met in place of Ile. Ile competes with TfeGly for translation in place of the single Ile codon- its incorporation should not be affected by post-transfer editing. Our data may indicate that Leu can substitute Ile in the editing-independent manner;[13] this interesting finding needs further clarification. Thus, the signal for the wt-mass (Leu/Ile) can be used as reference mass to quantify the relative incorporation of the artificial amino acid. The effect of *in vivo* post-transfer proofreading can be

observed by the relative incorporation mediated by the post-transfer inactive IleRS mutant in comparison to those mediated by the wild-type enzyme. The observed relative incorporations (**Figure S7**) of the fluorinated amino acids (**Figure S5**) are consistent with the *in vitro* aminoacylation rates (**Figure S6**). TfIle and TfVal are efficiently translated with the help of IleRS and IleRSAla10 without notable differences (**Figure S7**). *In vivo* translation of TfNVal into the model peptide VW18Ile by wild-type IleRS is, in contrast to TfeGly, observed, but in comparison to TfVal and TfIle the level is rather low (**Figure S7**). Hence, post-transfer editing of IleRS likely reduces the ribosomal translation rate of TfNVal in response to Ile codons as we observed that the post-transfer editing defective IleRSAla10 mutant increases translation significantly over those mediated by wild-type IleRS (**Figure S7**).

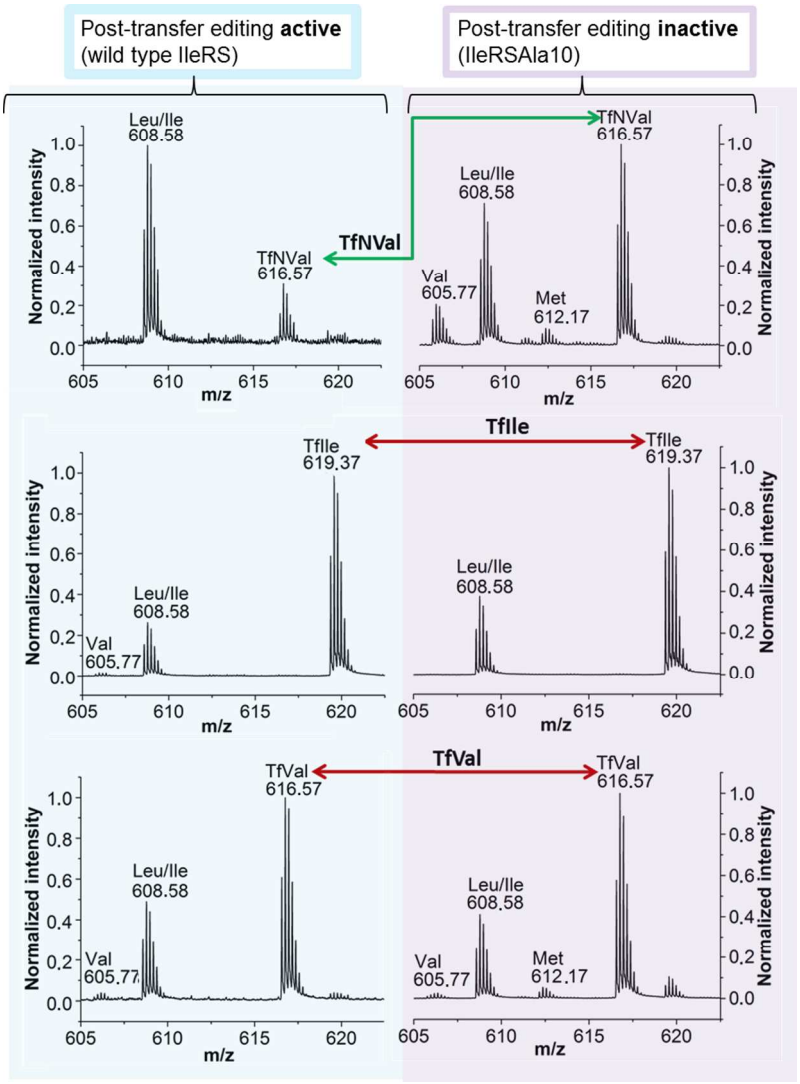

Figure S7: Effect of *in vivo* post-transfer editing versus TfeGly, TfNVal and TfVal. ESI-MS experiment of VW18Ile mutants expressed in the absence of Ile under supplementation of a corresponding fluorinated amino acid. For translation either the wild-type IleRS or the post-transfer editing deficient mutant IleRSAla10 was co-expressed. The monoisotopic masses of the 5-fold charged peptide are displayed. Calculated monoisotopic masses for Ile substitution: Val=605.78, Leu/Ile=608.57, TfVal/TfNVal=616.57, TfIle=619.37, TfeGly=613.75, Met=612.19.

1.7. Aminoacylation and Proofreading of TfeGly with the various CP1 IleRS variants

Table S2: Kinetic data for different reactions of wild-type and editing domain mutants.

	tRNA-independent editing	overall editing	aminoacylation	deacylation
	k_{obs} / s^{-1}	k_{obs} / s^{-1}	k_{obs} / s^{-1}	k_{deacyl} / s^{-1}
wt	0.055 ± 0.004	0.013 ± 0.002	< 0.00007 ± 0.00002	0.048 ± 0.006
Ala10	0.031 ± 0.006	0.002 ± 0.0004	0.002 ± 0.00006	n.d.
T243R	0.057 ± 0.004	0.017 ± 0.004	0.00018 ± 0.00005	0.033 ± 0.005
D342A	0.047 ± 0.005	0.012 ± 0.002	0.00062 ± 0.0001	0.052 ± 0.005
T243R/D342A	0.065 ± 0.003	0.048 ± 0.004	0.015 ± 0.0007 [a]	0.047 ± 0.006

[a] This is the value measured in the absence of activated Ef-Tu. In the presence of Ef-Tu it rises to 0.025 s^{-1}.

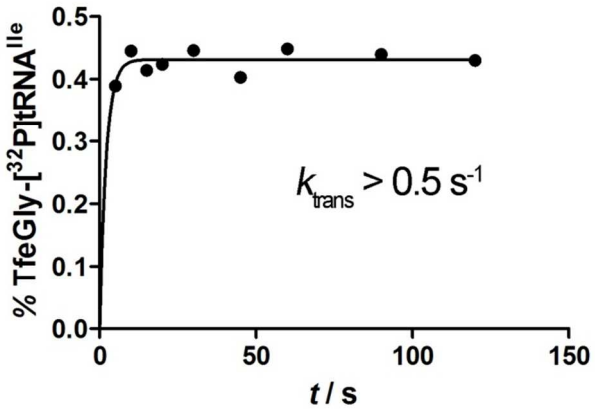

Figure S8: Transfer of TfeGly to the $tRNA^{Ile}$ measured by IleRSAla10.

The time points in **Figure S8** were sampled manually, due to the lack of material. Measurement on the millisecond range would require 50 times more enzyme, which was impossible to obtain. However, the quality of the obtained data is good enough to support our interpretations.

2. Materials and Methods

2.1. Reagents

2.1.1 Chemicals

All reagents and chemicals were purchased from VWR (Darmstadt), Carl Roth (Karlsruhe), Merck (Darmstadt) or Sigma-Aldrich (Taufkirchen) and used without further purification. ^{32}P labeled PP_i was bought from Perkin Elmer (Waltham, Massachusetts, USA).

2.1.2 Artificial amino acids

(2*S,2R*)-4,4,4-trifluoroethylglycine (Synquest Laboratories, FL, USA), (2*S*,2*R*)-4,4,4,4´,4´,4´-hexafluorovaline (Lancaster Synthesis, MA, USA); (2*S*,2*R*)-6,6,6-trifluoronorleucine (Fluorochem, Hadfield, UK); (2*S*,2*R*)-3,3,3-trifluoroalanine (Apollo Scientific, Manchester, UK) were obtained by the above specified companies. (2*S*)-4,4,4-trifluoroethylglycine, (2*S*)-4,4-difluoroethylglycine, (2*S*)-4-monofluoroethylglycine and (2*S*)-4,4-difluoropropylglycine were synthesized according to literature procedures[14–16]; (2*S*,4*S*,4*R*)-5,5,5-trifluoroleucine (TfLeu), (2*S*)-5,5,5,5´,5´,5´-hexafluoroleucine (HfLeu) and (2*S*,3*S*)-5,5,5-trifluoroisoleucine (TfIle) were synthesized as described[17,18]; (2*S*,3*R*)-4,4,4-trifluorovaline (TfVal) and (2*S*,3*S*)-4,4,4-trifluorovaline (TfVal) were synthesized by Holger Erdbrink (Czekelius group, FU Berlin) as published[19]; (2*S*,2*R*)-5,5,5-trifluoronorvaline (TfNVal) was synthesized by Stella Vukelic (Koksch group, FU Berlin) as described[20].

2.2. Plasmids

In this work pET (Novagen, Darmstadt), pQE (Qiagen, Hilden) and pSEVA[21–23] based vectors were used. The designated pBU vectors are derivatives of the pSEVA vector and were obtained by inserting an insert (**Figure S9**) into the multiple cloning site of a pSEVA vector via the restriction sites AscI and SpeI. The sequence and features of the insert that was ordered at Thermo Scientifics (Waltham, MA, USA) *GeneArt*® Gene Synthesis is shown:

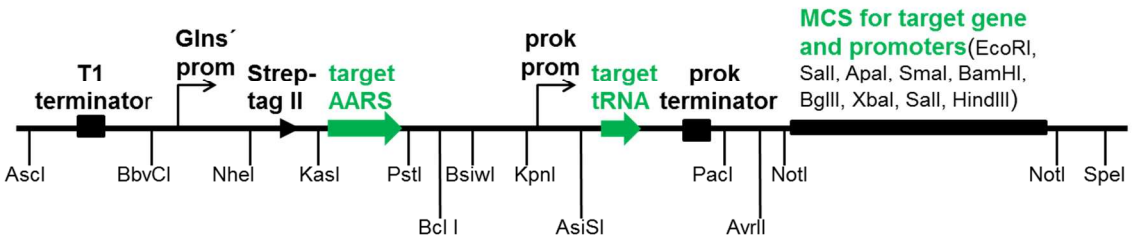

Figure S9: Insert in the multiple cloning site of the pSEVA vector that yields the pSEVA derivative pBU. The GlnS´ promoter was in in this work sometimes exchanged against a lacIPtac or pBAD promotor. The multiple cloning site (MCS) either remained unchanged or the gene encoding the peptide VW18 was inserted downstream of a T5 promotor.

In this work, the following vectors have been used:

Table S3: Plasmid vectors used in this study.

Plasmid name	**Source/Reference**
pQE80+IleRS	This study
pQE80+ValRS	This study
pQE80L+LeuRS	Budisa Laboratory (TU Berlin)
pET28a+MetRS	Budisa Laboratory (TU Berlin)
pBU26'4R(ΔNheI)+IleRS	This study
pBU26'4R(ΔNheI)+ValRS	This study
pSEVA18AraC+H_6-SUMO-VW18Ile	This study
pSEVA18AraC+ H_6-SUMOΔIle-VW18Ile	This study
pSEVA184F- H_6-SUMOΔIle-VW18Ile	This study
pBU26´Ara+IleRS+T5- H_6-SUMOΔIleVW18Ile+proktRNA	This study
pBU26´Ara+IleRS(241-250[Ala10])+ T5- H_6-SUMOΔIleVW18Ile+proktRNA	This study
pSUPER (encoding SUMO protease)	Anne Diehl (FMP Berlin), Patrick Loll (Drexel University College of Medicine, USA[24]

The plasmid sequences have been verified by DNA sequencing analysis. Amino acid sequences of the inserted genes can be found in **Section 2.3 Nucleotide sequences**.

2.3. Nucleotide sequences

2.3.1 pBU insert

GGCGCGCCCAGCTGTCTAGGGCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAAC
AGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTCCTCAGCCAT
CAATCATCCCCATAATCCTTGTTAGATTATCAATTTTAAAAAACTAACAGTTGTCAGCCTGTCCC
GCTTTAATATCATACGCCGTTATACGTTGTTTACGCTTTGAGGAATCCCATATGGCTAGCTGGAG
CCACCCGCAGTTCGAAAAAGGCGCCAGCTAGTAAATTCATGACTGCAGTGATCATCTGACCGTAC
GTTTCAAACGCTAAATTGCCTGATGCGCTACGCTTATCAGGCCTACATGATCTCTGCAATATATT
GAGTTTGCGTGCTTTTGTAGGCCGGATAAGGCGTTCACGCCGCATCCGGCAAGAAACAGCAAACA
ATCCAAAACGCCGCGTTCAGCGGCGTTTTTTCTGCTTTGGTACCGGCTAACTAAGCGGCCTGCTG
ACTTTCTCGCCGATCAAAAGGCATTTTGCTATTAAGGGATTGACGAGGGCGTATCTGCGCAGTAA
GATGCGCCCCGCATTGCGATCGCACGATCGAGGTTTAAACAATTCGAAAAGCCTGCTCAACGAGC
AGGCTTTTTTGCATTTAATTAAAGCGGATAACAATTTCACACAGGAGGCCGCCTAGGCCGCGGCC
GCGCGAATTCGAGCTCCCCGGGGATCCTCTAGAGTCGACAAGCTTGCGGCCGCGTCGTGACTGGG
AAAACCCTGGCGACTAGT

2.3.2 ileV gene (tRNA^{Ile})

AGGCTTGTAGCTCAGGTGGTTAGAGCGCACCCCTGATAAGGGTGAGGTCGGTGGTTCAAGTCCAC
TCAGGCCTACCA

2.4. Amino acid sequences

2.4.1 AARSs used in ATP-PP$_i$ exchange assay

2.4.1.1 IleRS

MRGSHHHHHHGSACELENLYFQGMSDYKSSLNLPETGFPMRGDLAKREPGMLARWTDDDLYGIIR
AAKKGKKTFILHDGPPYANGSIHIGHSVNKILKDIIVKSKGLSGYDSPYVPGWDCHGLPIELKVE
QEYGKPGEKFTAAEFRAKCREYAATQVDGQRKDFIRLGVLGDWSHPYLTMDFKTEANIIRALGKI
IGNGHLHKGAKPVHWCVDCRSALAEAEVEYYDKTSPSIDVAFQAVDQDALKAKFAVSNVNGPISL
VIWTTTPWTLPANRAISIAPDFDYALVQIDGQAVILAKDLVESVMQRIGVTDYTILSTVKGAELE
LLRFTHPFMGFDVPAILGDHVTLDAGTGAVHTAPGHGPDDYVIGQKYGLETANPVGPDGTYLPGT
YPTLDGVNVFKANDIVVALLQEKGALLHVEKMQHSYPCCWRHKTPIIFRATPQWFVSMDQKGLRA
QSLKEIKGVQWIPDWGQARIESMVANRPDWCISRQRTWGVPMSLFVHKDTEELHPRTLELMEEVA
KRVEVNGIQAWWDLDAKEILGDEADQYVKVPDTLDVWFDSGSTHSSVVDVRPEFAGHAADMYLEG
SDQHRGWFMSSLMISTAMKGKAPYRQVLTHGFTVDGQGRKMSKSIGNTVSPQDVMNKLGADILRL
WVASTDYTGEMAVSDEILKRAADSYRRIRNTARFLLANLNGFDPAKDMVKPEEMVVLDRWAVGCA
KAAQEDILKAYEAYDFHEVVQRLMRFCSVEMGSFYLDIIKDRQYTAKADSVARRSCQTALYHIAE
ALVRWMAPILSFTADEVWGYLPGEREKYVFTGEWYEGLFGLADSEAMNDAFWDELLKVRGEVNKV
IEQARADKKVGGSLEAAVTLYAEPELAAKLTALGDELRFVLLTSGATVADYNDAPADAQQSEVLK
GLKVALSKAEGEKCPRCWHYTQDVGKVAEHAEICGRCVSNVAGDGEKRKFA

HisTag, TEV recognition site

2.4.1.2 ValRS

MRGSHHHHHHGSENLYFQGMEKTYNPQDIEQPLYEHWEKQGYFKPNGDESQESFCIMIPPPNVTG
SLHMGHAFQQTIMDTMIRYQRMQGKNTLWQVGTDHAGIATQMVVERKIAAEEGKTRHDYGREAFI
DKIWEWKAESGGTITRQMRRLGNSVDWERERFTMDEGLSNAVKEVFVRLYKEDLIYRGKRLVNWD
PKLRTAISDLEVENRESKGSMWHIRYPLADGAKTADGKDYLVVATTRPETLLGDTGVAVNPEDPR
YKDLIGKYVILPLVNRRIPIVGDEHADMEKGTGCVKITPAHDFNDYEVGKRHALPMINILTFDGD
IRESAQVFDTKGNESDVYSSEIPAEFQKLERFAARKAVVAAVDALGLLEEIKPHDLTVPYGDRGG
VVIEPMLTDQWYVRADVLAKPAVEAVENGDIQFVPKQYENMYFSWMRDIQDWCISRQLWWGHRIP
AWYDEAGNVYVGRNEEEVRKENNLGADVALRQDEDVLDTWFSSALWTFSTLGWPENTDALRQFHP
TSVMVSGFDIIFFWIARMIMMTMHFIKDENGKPQVPFHTVYMTGLIRDDEGQKMSKSKGNVIDPL
DMVDGISLPELLEKRTGNMMQPQLADKIRKRTEKQFPNGIEPHGTDALRFTLAALASTGRDINWD
MKRLEGYRNFCNKLWNASRFVLMNTEGQDCGFNGGEMTLSLADRWILAEFNQTIKAYREALDSFR
FDIAAGILYEFTWNQFCDWYLELTKPVMNGGTEAELRGTRHTLVTVLEGLLRLAHPIIPFITETI
WQRVKVLCGITADTIMLQPFPQYDASQVDEAALADTEWLKQAIVAVRNIRAEMNIAPGKPLELLL
RGCSADAERRVNENRGFLQTLARLESITVLPADDKGPVSVAKIIDGAELLIPMAGLINKEDELAR
LAKEVAKIEGEISRIENKLANEGFVARAPEAVIAKEREKLEGYAEAKAKLIEQQAVIAAL

HisTag, TEV recognition site

2.4.1.3 LeuRS

MRGSHHHHHHGSMQEQYRPEEIESKVQLHWDEKRTFEVTEDESKEKYYCLSMLPYPSGRLHMGHV
RNYTIGDVIARYQRMLGKNVLQPIGWDAFGLPAEGAAVKNNTAPAPWTYDNIAYMKNQLKMLGFG
YDWSRELATCTPEYYRWEQKFFTELYKKGLVYKKTSAVNWCPNDQTVLANEQVIDGCCWRCDTKV
ERKEIPQWFIKITAYADELLNDLDKLDHWPDTVKTMQRNWIGRSEGVEITFNVNDYDNTLTVYTT
RPDTFMGCTYLAVAAGHPLAQKAAENNPELAAFIDECRNTKVAEAEMATMEKKGVDTGFKAVHPL
TGEEIPVWAANFVLMEYGTGAVMAVPGHDQRDYEFASKYGLNIKPVILAADGSEPDLSQQALTEK
GVLFNSGEFNGLDHEAAFNAIADKLTAMGVGERKVNYRLRDWGVSRQRYWGAPIPMVTLEDGTVM
PTPDDQLPVILPEDVVMDGITSPIKADPEWAKTTVNGMPALRETDTFDTFMESSWYYARYTCPQY
KEGMLDSEAANYWLPVDIYIGGIEHAIMHLLYFRFFHKLMRDAGMVNSDEPAKQLLCQGMVLADA
FYYVGENGERNWVSPVDAIVERDEKGRIVKAKDAAGHELVYTGMSKMSKSKNNGIDPQVMVERYG
ADTVRLFMMFASPADMTLEWQESGVEGANRFLKRVWKLVYEHTAKGDVAALNVDALTENQKALRR
DVHKTIAKVTDDIGRRQTFNTAIAAIMELMNKLAKAPTDGEQDRALMQEALLAVVRMLNPFTPHI
CFTLWQELKGEGDIDNAPWPVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATEEQVRERAGQEH
LVAKYLDGVTVRKVIYVPGKLLNLVVG

HisTag

2.4.1.4 MetRS

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEFMTQVAKKILVTCALPYANGSIHLGHMLEH
IQADVWVRYQRMRGHEVNFICADDAHGTPIMLKAQQLGITPEQMIGEMSQEHQTDFAGFNISYDN
YHSTHSEENRQLSELIYSRLKENGFIKNRTISQLYDPEKGMFLPDRFVKGTCPKCKSPDQYGDNC
EVCGATYSPTELIEPKSVVSGATPVMRDSEHFFFDLPSFSEMLQAWTRSGALQEQVANKMQEWFE
SGLQQWDISRDAPYFGFEIPNAPGKYFYVWLDAPIGYMGSFKNLCDKRGDSVSFDEYWKKDSTAE
LYHFIGKDIVYFHSLFWPAMLEGSNFRKPSNLFVHGYVTVNGAKMSKSRGTFIKASTWLNHFDAD
SLRYYYTAKLSSRIDDIDLNLEDFVQRVNADIVNKVVNLASRNAGFINKRFDGVLASELADPQLY
KTFTDAAEVIGEAWESREFGKAVREIMALADLANRYVDEQAPWVVAKQEGRDADLQAICSMGINL
FRVLMTYLKPVLPKLTERAEAFLNTELTWDGIQQPLLGHKVNPFKALYNRIDMRQVEALVEASKE

EVKAAAAPVTGPLADDPIQETITFDDFAKVDLRVALIENAEFVEGSDKLLRLTLDLGGEKRNVFS
GIRSAYPDPQALIGRHTIMVANLAPRKMRFGISEGMVMAAGPGGKDIFLLSPDAGAKPGHQVK

HisTag

2.4.2 AARSs for in vivo translation of artificial amino acids

2.4.2.1 IleRS

MASWSHPQFEKGASDYKSSLNLPETGFPMRGDLAKREPGMLARWTDDDLYGIIRAAKKGKKTFIL
HDGPPYANGSIHIGHSVNKILKDIIVKSKGLSGYDSPYVPGWDCHGLPIELKVEQEYGKPGEKFT
AAEFRAKCREYAATQVDGQRKDFIRLGVLGDWSHPYLTMDFKTEANIIRALGKIIGNGHLHKGAK
PVHWCVDCRSALAEAEVEYYDKTSPSIDVAFQAVDQDALKAKFAVSNVNGPISLVIWTTTPWTLP
ANRAISIAPDFDYALVQIDGQAVILAKDLVESVMQRIGVTDYTILSTVKGAELELLRFTHPFMGF
DVPAILGDHVTLDAGTGAVHTAPGHGPDDYVIGQKYGLETANPVGPDGTYLPGTYPTLDGVNVFK
ANDIVVALLQEKGALLHVEKMQHSYPCCWRHKTPIIFRATPQWFVSMDQKGLRAQSLKEIKGVQW
IPDWGQARIESMVANRPDWCISRQRTWGVPMSLFVHKDTEELHPRTLELMEEVAKRVEVNGIQAW
WDLDAKEILGDEADQYVKVPDTLDVWFDSGSTHSSVVDVRPEFAGHAADMYLEGSDQHRGWFMSS
LMISTAMKGKAPYRQVLTHGFTVDGQGRKMSKSIGNTVSPQDVMNKLGADILRLWVASTDYTGEM
AVSDEILKRAADSYRRIRNTARFLLANLNGFDPAKDMVKPEEMVVLDRWAVGCAKAAQEDILKAY
EAYDFHEVVQRLMRFCSVEMGSFYLDIIKDRQYTAKADSVARRSCQTALYHIAEALVRWMAPILS
FTADEVWGYLPGEREKYVFTGEWYEGLFGLADSEAMNDAFWDELLKVRGEVNKVIEQARADKKVG
GSLEAAVTLYAEPELAAKLTALGDELRFVLLTSGATVADYNDAPADAQQSEVLKGLKVALSKAEG
EKCPRCWHYTQDVGKVAEHAEICGRCVSNVAGDGEKRKFA

Streptag II

2.4.2.2 ValRS

MASWSHPQFEKGAEKTYNPQDIEQPLYEHWEKQGYFKPNGDESQESFCIMIPPPNVTGSLHMGHA
FQQTIMDTMIRYQRMQGKNTLWQVGTDHAGIATQMVVERKIAAEEGKTRHDYGREAFIDKIWEWK
AESGGTITRQMRRLGNSVDWERERFTMDEGLSNAVKEVFVRLYKEDLIYRGKRLVNWDPKLRTAI
SDLEVENRESKGSMWHIRYPLADGAKTADGKDYLVVATTRPETLLGDTGVAVNPEDPRYKDLIGK
YVILPLVNRRIPIVGDEHADMEKGTGCVKITPAHDFNDYEVGKRHALPMINILTFDGDIRESAQV
FDTKGNESDVYSSEIPAEFQKLERFAARKAVVAAVDALGLLEEIKPHDLTVPYGDRGGVVIEPML
TDQWYVCADVLAKPAVEAVENGDIQFVPKQYENMYFSWMRDIQDWCISRQLWWGHRIPAWYDEAG
NVYVGRNEEEVRKENNLGADVALRQDEDVLDTWFSSALWTFSTLGWPENTDALRQFHPTSVMVSG
FDIIFFWIARMIMMTMHFIKDENGKPQVPFHTVYMTGLIRDDEGQKMSKSKGNVIDPLDMVDGIS
LPELLEKRTGNMMQPQLADKIRKRTEKQFPNGIEPHGTDALRFTLAALASTGRDINWDMKRLEGY
RNFCNKLWNASRFVLMNTEGQDCGFNGGEMTLSLADRWILAEFNQTIKAYREALDSFRFDIAAGI
LYEFTWNQFCDWYLELTKPVMNGGTEAELRGTRHTLVTVLEGLLRLAHPIIPFITETIWQRVKVL
CGITADTIMLQPFPQYDASQVDEAALADTEWLKQAIVAVRNIRAEMNIAPGKPLELLLRGCSADA
ERRVNENRGFLQTLARLESITVLPADDKGPVSVAKIIDGAELLIPMAGLINKEDELARLAKEVAK
IEGEISRIENKLANEGFVARAPEAVIAKEREKLEGYAEAKAKLIEQQAVIAAL

Streptag II

2.4.2.3 LeuRS

MASWSHPQFEKGAQEQYRPEEIESKVQLHWDEKRTFEVTEDESKEKYYCLSMLPYPSGRLHMGHV
RNYTIGDVIARYQRMLGKNVLQPIGWDAFGLPAEGAAVKNNTAPAPWTYDNIAYMKNQLKMLGFG
YDWSRELATCTPEYYRWEQKFFTELYKKGLVYKKTSAVNWCPNDQTVLANEQVIDGCCWRCDTKV
ERKEIPQWFIKITAYADELLNDLDKLDHWPDTVKTMQRNWIGRSEGVEITFNVNDYDNTLTVYTT
RPDTFMGCTYLAVAAGHPLAQKAAENNPELAAFIDECRNTKVAEAEMATMEKKGVNTGFKAVHPL
TGEEIPVWAANFVLMEYGTGAVMAVPGHDQRDYEFASKYGLNIKPVILAADGSEPDLSQQALTEK
GVLFNSGEFNGLDHEAAFNAIADKLTAMGVGERKVNYRLRDWGVSRQRYWGAPIPMVTLEDGTVM
PTPDDQLPVILPEDVVMDGITSPIKADPQWAKTTVNGMPALRETDTFDTFMESSWYYARYTCPEY
KEGMLDSEAANYWLPVDIYIGGIEHAIMHLLYFRFFHKLMRDAGMVNSDEPAKQLLCQGMVLADA
FYYVGENGERNWVSPVDAIVERDEKGRIVKAKDAAGHELVYTGMSKMSKSKNNGIDPQVMVERYG
ADTVRLFMMFASPADMTLEWQESGVEGANRFLKRVWKLVYEHTAKGDVAALNVDALTEDQKALRR
DVHKTIAKVTDDIGRRQTFNTAIAAIMELMNKLAKAPTDGEQDRALMQEALLAVVRMLNPFTPHI
CFTLWQELKGEGDIDNAPWPVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATEEQVRERAGQEH
LVAKYLDGVTVRKVIYVPGKLLNLVVG

Streptag II

2.4.3 SUMO-VW18

2.4.3.1 H_6-SUMOΔIleVW18Ile

MHHHHHHGGSDSEVNQEAKPEVKPEVKPETHVNLKVSDGSSEVFFKVKKTTPLRRLMEAFAKRQG
KEMDSLRFLYDGQRLQADQTPEDLDMEDNDTLEAHREQTGGLKVELEKLKSELVILKSELEKLKS
EL

HisTag; SUMOΔIle; VW18Ile

2.4.3.2 H_6-SUMO-VW18Ile

MGHHHHHHGGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQ
GKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGGLKVELEKLKSELVILKSELEKLK
SEL

HisTag; SUMO; VW18Ile

2.5. Bacterial strains

2.5.1 Strains

Table S4: Bacterial strains used in this study.

Strain name	Genotype	Source/Reference
E. coli CAG18341	*E. coli* K12 C> F- λ- *ilvG*$^-$ *rfb*$^-$*50 rph-1 nadA3052::Tn10kan*1 *rph-1 ilvD*	Coli Genetic Stock Center (CGSC).[25]
E. coli BL21 Gold (DE3)	*E. coli B* *C> F- ompT hsdSB(rB- mB-) dcm+ Tetr gal λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) endA Hte*	Strain database of Budisa Group /Life Technologies, Darmstadt

In order to use the IleRS translational system for efficient incorporation of artificial amino acids (e.g. fluorinated AAs) in peptides and proteins, the absence of Ile in the cell has to be secured during translation. This can be achieved using strains with genetic knockouts that cannot produce Ile and are thus auxotrophic for this AA. CAG18341[25] is auxotrophic for Ile, Leu and Val, which is based on the genetic knockout of *ilvD* that encodes the dihydroxy-acid dehydratase belonging to the Ile, Val and Leu biosynthetic pathway.[26]

2.5.2. Ile auxotrophy of CAG18341

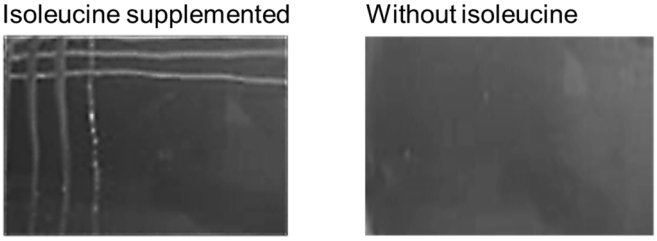

Figure S10: *Escherichia coli* CAG 18431(Δ*ilvD*) streaked on plates (MS glucose minimal media, 0.5 mM Leu/Val) in the presence (left) or in the absence (right) of Ile.

As reported[25], the strain CAG18341 is not able to grow in the absence of Ile (**Figure S10**) indicating that it is not able to produce this amino acid, thus highlighting its Ile auxotrophy.

2.6. Protein expression

2.6.1. Standard protein expression

The plasmids pQE80+IleRS, pQE80+ValRS, pQE80L+LeuRS and pET28a+MetRS were used for the Expression of the AARSs for the ATP-PP_i exchange assay (IleRS, ValRS, LeuRS, MetRS). The plasmid with the target gene under control of a T5 promotor was transformed into the *E. coli* expression strain BL21 Gold (DE3). 10 mL LB medium containing ampicillin (100 µg/mL) were inoculated with a single colony and grown overnight. This pre-culture was used to inoculate 1 L of LB medium containing ampicillin (100 µg/mL) and cells were grown under shaking (160 rpm) at 37 °C. When the cell culture reached an OD_{600} between 0.5 and 0.8 expression of the AARS gene was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Protein expression was performed for 16 h at 28 °C. The cells were subsequently harvested by centrifugation (20 min, 4 °C, 8,000 x *g*) and stored at -80 °C until further usage. The SUMO protease (ULP1) was expressed according to Weeks et al.[24]

2.6.2. Incorporation of artificial amino acids via the Selective-pressure incorporation (SPI)-method

In most cases, competent cells of the strain CAG18341(Δ*ilvD*) were transformed with a plasmid containing the heterologous gene (H_6-SUMOΔIle-VW18Ile) under control of an IPTG inducible T5 promotor, *iles* (gene encoding IleRS or IleRS variant) under arabinose inducible pBAD promotor and *ileV* (gene encoding tRNA(Ile)) under control of constitutive proK promotor. A single colony was used to inoculate 5 mL NMM(glycerol) containing 0.2 mM of all cAAs and kanamycin (50 µg/mL) and cells were incubated over day at 37 °C until OD_{600}=0.3. 250 mL NMM(glycerol) containing 0.2 mM of all cAAs and kanamycin (50 µg/mL) were inoculated with 25 µL of the overday culture and incubated over night at 30 °C, 180 rpm. At OD_{600}=0.2-0.3 expression of IleRS was induced with 0.2% arabinose and the cells were grown until OD_{600}=0.7-0.8. Subsequently the culture was centrifuged (6 min, 3,000 x *g*). After washing the pellet three times with 120 mL NMM(glycerol), the cells were resuspended in 250 mL NMM(glycerol) without Ile, supplemented with 0.075 mM Val, 0.075 mM Leu and 0.1 mM of all other 17 standard cAAs. Before addition of 1 mM artificial amino acid, cells were incubated for 30 min at 37 °C. Target gene expression was induced with 1 mM IPTG after 10 min incubation with the artificial amino acid. Cells were harvested by centrifugation (8,000 *g*, 4 °C for 20 min) and the pellet was stored at -80 °C after 4

h protein expression at 37 °C. For experiments that desired a strong overexpression of IleRS or ValRS we used a co-transformation of one plasmid bearing the heterologous gene (H_6-SUMOΔIle-VW18Ile) under control of a pBAD promoter and another plasmid bearing the AARS gene under control of a lacIPtac promotor. In such cases the following modifications were done: The AARS (IleRS or ValRS) was first expressed in LB medium overnight in the presence of 0.1 mM IPTG and kanamycin (50 µg/mL). Expression of target gene was induced with 0.2% arabinose after cells were shifted to NMM(glycerol) containing ampicillin (100 µg/mL) and all cAAs (0.2 mM) except Ile, when cells reached OD_{600}=0.7-0.8.

2.7. Cell lysis

Cell pellets were resuspended in lysis buffer (pH=8.0) (10mL per 1g of cell pellet) containing 50 mM TRIS, 150 mM NaCl, 20 mM imidazole, DNase and RNase. Cell lysis was carried out by sonification (Sonoplus HD3200, KE76 electrode) in 5 cycles of 2 min each, 60% amplitude at an interval of 1 s pulse and 4 s pause. After centrifugation (18,000 g, 30 min at 4 °C), the supernatant was filtered (0.45 µm) and the lysate stored on ice until purification.

2.8. Protein purification

IMAC purification was carried out for all H_6-tagged proteins used in this work (AARSs used in ATP-PP_i assay, SUMO protease, H_6-SUMOΔIle-VW18Ile variants).

Composition of buffers used for the purification:

Binding buffer: 50 mM Tris (pH=7.0-8.0), 150 mM NaCl, 20-30 mM imidazole, 1 mM DTT

Elution buffer: 50 mM Tris (pH=7.0-8.0), 150 mM NaCl, 500 mM imidazole, 1 mM DTT

1 mL and 5 mL FF HisTrap columns (GE Healthcare, Munich) attached to an ÄKTA FPLC system were used. The Ni-NTA column was equilibrated with 5 column volumes binding buffer the cell lysate was applied onto the column. The bound proteins were washed with 15 column volumes binding buffer. The target protein was subsequently eluted either with linear or step gradient (1.Step: 5% elution buffer, 2.Step: 70% elution buffer) of increasing imidazole concentration. The fractions containing the desired protein were selected based on the UV signal (λ=280), pooled together and subjected to further analysis (SDS-PAGE (**Figure S11**) and/or mass spectrometry).

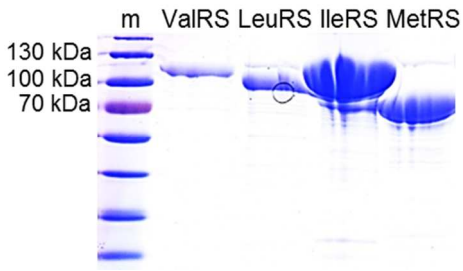

Figure S11: SDS-PAGE gel of purified (Ni^{2+}-IMAC) AARSs that were used for the ATP-PP_i exchange assay. m: PageRuler™ Prestained Protein Ladder. Calculated masses: ValRS=110 kDa, LeuRS=97 kDa, IleRS=107 kDa, MetRS=80 kDa.

2.9. Protein concentration

The concentration of IleRS, LeuRS, ValRS, MetRS and SUMO protease was determined spectrophotometrically by measuring their absorbance at 280 nm according to Lambert-Beer law applying molar extinction coefficients that were calculated using the web based software ProtParam (http://web.expasy.org/protparam). Concentrations of all VW18Ile peptide and protein variants were determined with the Bradford assay[27].

2.10. Digestion of H_6-SUMOΔIle-VW18Ile with SUMO protease

VW18Ile peptides were cleaved after expression from the attached Ile-free solubility and stability domain (H_6-SUMOΔIle) by incubation of H_6-SUMOΔIle-VW18Ile with SUMO protease. Digestion was carried out in the presence of 5 mM DTT, 1 mM EDTA, 250 mM NaCl, 50 mM Tris, 3% acetonitrile (ACN), 20 µM H_6-SUMO-VW18 and 5 µM SUMO protease. Additionally, 3% ACN was supplemented to prevent amyloid formation of VW18Ile. The reaction was performed at 30 °C, pH=8.0, for 40 min and a protease to substrate ratio of 1:4. The digestion was quenched with 50% ACN and 0.1% formic acid (CH_2O_2).

2.11. HPLC-ESI-MS

MS data was acquired using either FTMS (Exactive, Thermo Fisher Scientific GmbH, Dreieich, Germany) or TOFMS (G6530B Accurate-Mass Q-TOF Agilent Technologies GmbH, Waldbronn, Germany) instruments applying the positive ion mode. A C5 reverse phase column (Supelco Analytical, Sigma-Aldrich, Taufkirchen) attached to Agilent 1200 (Exactive) or Infinity 1200 (G6530B) series HPLC system was used for chromatographic separation. A mobile phase of water (A) and acetonitrile (B) in the presence of 0.1% formic acid was used. A gradient with 20%

to 90% of B in 20 min with a flow of 0.2 mL/min (Exactive) or with 20% to 80% of B in 20 min with a flow of 0.3 mL/min (G6530B) was applied. The chromatographic system was coupled to the mass spectrometer via Heated Electrospray Ionization (HESI) (Exactive) or Dual Agilent Jet Stream Electrospray Ionization (Dual AJS ESI) (G6530B) interface. The obtained mass spectra were deconvoluted with maximum entropy using either the program MagTran (Exactive) or the Agilent Mass Hunter Analysis software (G6530B) to obtain the molecular mass of the analyzed proteins.

2.12. ATP-PP$_i$ exchange assay

The assay was conducted at 37 °C in a 600 µL reaction mix containing 100 mM Tris/HCl (pH=8.0), 80 mM $MgCl_2$, 5 mM KF, 700 mM β-mercaptoethanol, 2.2 mM NaP_2O_7, 5.5 mM ATP, 0.1 mg/mL BSA, ~10 µCi ^{32}P labeled PP_i and 1 µM AARS. The amount of artificial amino acid (0.1 mM to 40 mM) and natural substrate (0.001-1 mM) were varied, respectively. Reactions were quenched with 600 µL of 240 mM sodium pyrophosphate containing 70% (v/v) perchloric acid and 7.5% (w/v) activated charcoal. The mixture was vortexed and subsequently filtered through a filter paper. After two rounds of washing with 10 mL of distilled water and the filter paper was put into 4 mL of scintillation solution and the absorbed ^{32}P labeled ATP was analyzed by scintillation counting. The initial reaction rates were determined (slope (ATP formation over time), plotted through x=0 at y=0) and plotted against the substrate concentrations. The Michaelis-Menten model was applied and the kinetic values k_{cat} and K_M were determined. Measurements were performed in triplicate and the standard error for k_{cat} and K_M was determined as follows: ($s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}$). The error for k_{cat}/K_M was obtained by applying error propagation ($\Delta\left(\frac{k_{cat}}{K_M}\right) = \left(\frac{1}{K_M}\right) s \times k_{cat} - \left(\frac{k_{cat}}{{K_M}^2}\right) s \times K_M$).

2.13. ATP-ase assay

The assay was performed at 37 °C in 50 mM HEPES (pH=7.5), 20 mM $MgCl_2$, 5 mM DTT, 0.1 mg/ml BSA, 0.004 units/µl inorganic pyrophosphatase and 0.5 mM [α-^{32}P]ATP (0.01–0.1 mCi/ml). Steady-state parameters for tRNA-independent pre-transfer editing were determined at a concentration of 40 mM D, L-TfeGly and 2 µM AARS. The overall editing was followed in the same reaction conditions, with addition of 10 µM $tRNA^{Ile}$. Reactions were initiated by addition of amino acids and stopped by quenching in formic acid (final concentration 1 M). The mixture was

spotted onto polyethyleneimine (PEI)-cellulose plates (Fluka) that were prewashed with water. Separation of AA-[^{32}P]AMP, [^{32}P]AMP, and [^{32}P]ATP was achieved using TLC in 0.1 M ammonium acetate, 5% acetic acid. Visualization and quantization of the signal were performed on a Typhoon PhosphorImager, using ImageQuant.

2.14. Preparation of tRNA^{Ile}

$tRNA_{GAU}^{Ile}$ was overexpressed in *E. coli* BL21(DE3), isolated, and purified by reverse phase chromatography as described previously[28] yielding $tRNA^{Ile}$ capable of 90% aminoacylation. tRNA was radiolabeled by exchange of endogenous A76 with [α-^{32}P]ATP[29] by incubation of 10 µM $tRNA^{Ile}$ with 5 µM tRNA-nucleotidyltransferase, 1 µM [α-^{32}P]ATP (specific activity of 3000 Ci/mmol), and 5 mM Na-PP_i in a buffer containing 200 mM Tris (pH 8.0), 20 mM $MgCl_2$, and 0.5 mM DTT at 37 °C. After 1 min, 0.1 unit/µL thermostable inorganic pyrophosphatase was added to enhance [α-^{32}P]ATP incorporation, and the mixture was incubated at 37 °C for an additional 2 min followed by phenol/ chloroform extraction. The sample was passed through two consecutive Bio-Spin P-30 columns (Bio-Rad) and dialyzed against 5 mM HEPES (pH=7.0).

2.15. Aminoacylation assay

Aminoacylation reactions were performed at 37 °C in the presence of 20 mM HEPES (pH=7.5), 0.1 mM EDTA, 150 mM NH_4Cl, 0.01 mg/ml BSA, 10 mM $MgCl_2$, 2 mM ATP, 40 mM artificial amino acid and 15 µM tRNA. AARSs were present at 1 µM, unless stated otherwise. Some reactions additionally contained 40 µM EF-Tu, activated as described previously[30]. The reaction was stopped with sodium acetate (pH=5.0) and SDS (final concentrations of 0.4 M and 0.1%, respectively). tRNA was degraded with P1 nuclease and AA-AMP was separated from AMP by TLC on PEI cellulose plates. The latter were developed in a solution containing 100 mM ammonium acetate and 5% acetic acid. The ratio of AA-AMP to AMP is equivalent to the ratio of aminoacylated to nonaminoacylated tRNA in the reaction mixture. Visualization and quantization of the signal were performed on a Typhoon PhosphorImager, using ImageQuant.

The transfer step was measured by mixing the 20 µM of *in situ* preformed enzyme:aa-AMP complex with 1 µM [^{32}P]tRNA in the same assay conditions as described previously for Aminoacylation. Reactions were manually quenched with sodium acetate (pH 5.0) and SDS (final concentrations of 0.4 M and 0.1%, respectively) and further treated as described above for

Aminoacylation. The amount of aminoacylated tRNA was plotted versus time and fit to the first-order exponential equation $y = Y_0 + A * (1 - e^{-k_{\text{trans}} * t})$, where Y_0 is the *y* intercept, *A* is a scaling constant, k_{trans} is the apparent transfer rate constant, and *t* is time.

2.16. Deacylation assay

Misaminoacylated $tRNA^{Ile}$ was prepared by mixing 10 µM $tRNA^{Ile}$ with 1 µM T243R/D342A IleRS and 40 µM activated EF-Tu in the same buffer conditions as described under Aminoacylation. The reaction was allowed to proceed for 45 minutes after which it was phenol/chloroform extracted, purified on two consecutive Bio-Spin P-30 columns (Bio-Rad) and dialyzed against 5 mM sodium acetate (pH=5.0).

Single-turnover deacylation assay was carried out by mixing equal volumes of 100-500 nM TfeGly-[^{32}P]$tRNA^{Ile}$ with 20 µM AARS in 200 mM HEPES (pH 7.5), 75 mM NH_4Cl, 20 mM $MgCl_2$, 5 mM DTT, 10 ng/µL BSA. Data were fit to the single exponential equation: $y = Y_0 + A * e^{k_{deacyl} * t}$. ($Y_0$=*y* intercept, *A*=amplitude, k_{deacyl}=the apparent deacylation rate constant, *t*=time). Reactions were sampled manually by quenching in sodium acetate (pH 5.0) and SDS (final concentrations of 0.6 M and 0.1%, respectively), followed by degradation with P1 nuclease, thin-layer chromatography and analysis, as described under **Section 2.15. Aminoacylation.**

3. References

(1) Rock, F. L.; Mao, W.; Yaremchuk, A.; Tukalo, M.; Crepin, T.; Zhou, H.; Zhang, Y.-K.; Hernandez, V.; Akama, T.; Baker, S. J. *et al.* An Antifungal Agent Inhibits an Aminoacyl-tRNA Synthetase by Trapping tRNA in the Editing Site. *Science* **2007**, *316*, 1759–1761.

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

(3) Fukai, S.; Nureki, O.; Sekine, S.; Shimada, A.; Tao, J.; Vassylyev, D. G.; Yokoyama, S. 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* **2000**, *103*, 793–803.

(4) Crepin, T.; Schmitt, E.; Mechulam, Y.; Sampson, P. B.; Vaughan, M. D.; Honek, J. F.; Blanquet, S. Use of analogues of methionine and methionyl adenylate to sample conformational changes during catalysis in Escherichia coli methionyl-tRNA synthetase. *J. Mol. Biol.* **2003**, *332*, 59–72.

(5) Serre, L.; Verdon, G.; Choinowski, T.; Hervouet, N.; Risler, J. L.; Zelwer, C. How methionyl-tRNA synthetase creates its amino acid recognition pocket upon L-methionine binding. *J. Mol. Biol.* **2001**, *306*, 863–876.

(6) Schrödinger, L. L. *The PyMOL Molecular Graphics System, Version 1.3r1*, **2010**.

(7) Gerling, Ulla I. M. Evaluating Fluorine's Impact on Amyloid Formation – a Systematic Study Using a Coiled-Coil Based Model Peptide, *PhD thesis*, **2013**.
(8) Pagel, K.; Wagner, S. C.; Samedov, K.; Berlepsch, H. von; Böttcher, C.; Koksch, B. Random coils, beta-sheet ribbons, and alpha-helical fibers: one peptide adopting three different secondary structures at will. *J. Am. Chem. Soc.* **2006**, *128*, 2196–2197.
(9) Brandenburg, E.; Berlepsch, H. von; Gerling, Ulla I M; Böttcher, C.; Koksch, B. Inhibition of amyloid aggregation by formation of helical assemblies. *Chemistry* **2011**, *17*, 10651–10661.
(10) Wang, P.; Fichera, A.; Kumar, K.; Tirrell, D. A. Alternative translations of a single RNA message: an identity switch of (2S,3R)-4,4,4-trifluorovaline between valine and isoleucine codons. *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 3664–3666.
(11) Pezo, V.; Metzgar, D.; Hendrickson, T. L.; Waas, W. F.; Hazebrouck, S.; Döring, V.; Marlière, P.; Schimmel, P.; Crécy-Lagard, V. de. Artificially ambiguous genetic code confers growth yield advantage. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8593–8597.
(12) Fukunaga, R.; Yokoyama, S. Structural basis for substrate recognition by the editing domain of isoleucyl-tRNA synthetase. *J. Mol. Biol.* **2006**, *359*, 901–912.
(13) Boniecki, M. T.; Martinis, S. A. Coordination of tRNA synthetase active sites for chemical fidelity. *J. Biol. Chem.* **2012**, *287*, 11285–11289.
(14) Tsushima, T.; Kawada, K.; Ishihara, S.; Uchida, N.; Shiratori, O.; Higaki, J.; Hirata, M. Fluorine containing amino acids and their derivatives. 7. Synthesis and antitumor activity of α- and γ-substituted methotrexate analogs. *Tetrahedron* **1988**, *44*, 5375–5387.
(15) Osipov, S. N.; Lange, T.; Tsouker, P.; Spengler, J.; Hennig, L.; Koksch, B.; Berger, S.; El-Kousy, S. M.; Burger, K. Hexafluoroacetone as a Protecting and Activating Reagent: Synthesis of New Types of Fluoro-Substituted α-Amino, α-Hydroxy and α-Mercapto Acids. *Synthesis* **2004**, 1821–1829.
(16) Winkler, D.; Burger, K. Synthesis of Enantiomerically Pure D- and L-Armentomycin and Its Difluoro Analogues from Aspartic Acid. *Synthesis* ***1996***, 1419–1421.
(17) Biava, H.; Budisa, N. Obtention of enantiomerically pure 5,5,5-trifluoro-l-isoleucine and 5,5,5-trifluoro-l-alloisoleucine. *J. Fluor. Chem.* **2013**, *156*, 372–377.
(18) Biava, H.; Budisa, N. Biocatalytic synthesis of (2S)-5,5,5-trifluoroleucine and improved resolution into (2S,4S) and (2S,4R) diastereoisomers. *Tetrahedron Lett.* **2013**, *54*, 3662–3665.
(19) Erdbrink, H.; Peuser, I.; Gerling, Ulla I. M.; Lentz, D.; Koksch, B.; Czekelius, C. Conjugate hydrotrifluoromethylation of α,β-unsaturated acyl-oxazolidinones: synthesis of chiral fluorinated amino acids. *Org. Biomol. Chem.* **2012**, *10*, 8583–8586.
(20) Vukelić, S.; Ushakov, D. B.; Gilmore, K.; Koksch, B.; Seeberger, P. H. Flow Synthesis of Fluorinated α-Amino Acids. *Eur. J. Org. Chem.* ***2015***, 3036–3039.
(21) Martínez-García, E.; Aparicio, T.; Goñi-Moreno, A.; Fraile, S.; Lorenzo, V. de. SEVA 2.0: an update of the Standard European Vector Architecture for de-/re-construction of bacterial functionalities. *Nucleic Acids Res.* **2015**, *43*, 1183-1189.
(22) Durante-Rodríguez, G.; Lorenzo, V. de; Martínez-García, E. The Standard European Vector Architecture (SEVA) plasmid toolkit. *Methods Mol. Biol.* **2014**, *1149*, 469–478.
(23) Silva-Rocha, R.; Martínez-García, E.; Calles, B.; Chavarría, M.; Arce-Rodríguez, A.; Las Heras, A. de; Páez-Espino, A. D.; Durante-Rodríguez, G.; Kim, J.; Nikel, P. I. *et al.* The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* **2013**, *41*, 666-675.
(24) Weeks, S. D.; Drinker, M.; Loll, P. J. Ligation independent cloning vectors for expression of SUMO fusions. *Protein Expr. Purif.* **2007**, *53*, 40–50.
(25) Singer, M.; Baker, T. A.; Schnitzler, G.; Deischel, S. M.; Goel, M.; Dove, W.; Jaacks, K. J.; Grossman, A. D.; Erickson, J. W.; Gross, C. A. A collection of strains containing genetically linked

alternating antibiotic resistance elements for genetic mapping of Escherichia coli. *Microbiol. Rev.* **1989**, *53*, 1–24.
(26) Pátek, M. *Amino acid biosynthesis: Pathways, regulation, and metabolic engineering*; Volume 5 of the series Microbiology monographs ; Springer: Berlin, 2007; pp 129-162.
(27) Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254.
(28) Dulic, M.; Cvetesic, N.; Perona, J. J.; Gruic-Sovulj, I. Partitioning of tRNA-dependent editing between pre- and post-transfer pathways in class I aminoacyl-tRNA synthetases. *J. Biol. Chem.* **2010**, *285*, 23799–23809.
(29) Wolfson, A. D.; Uhlenbeck, O. C. Modulation of tRNAAla identity by inorganic pyrophosphatase. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5965–5970.
(30) Cvetesic, N.; Akmacic, I.; Gruic-Sovulj, I. Lack of discrimination against non-proteinogenic amino acid norvaline by elongation factor Tu from Escherichia coli. *Croat. Chem. Acta* **2013**, *86*, 73–82.