Supplementary Information for
Designing seryl-tRNA synthetase for improved serylation of selenocysteine tRNAs

Xian Fu, Ana Crnković, Anastasia Sevostyanova and Dieter Söll

This file includes:

Supplementary Text

Figure S1: Nonsense suppression efficiency of $tRNA^{UTuX}$ and $tRNA^{SecUX}$.
Figure S2: Nonsense suppression efficiency using a serine-specific reporter, CAT-S146TAG.
Figure S3: Serylation of $tRNA^{UTuX}$ by SerRS variants with a flexible linker and S55A mutation as monitored by a nonsense suppression assay.
Figure S4: Reexamination of selected SerRS variants in a $tRNA^{UTuX}$-dependent nonsense suppression assay.
Figure S5: Serine 55-to-alanine mutation is essential for SerRS V20 activity *in vivo*.
Figure S6: Stop-codon read-through using evolved variants and $tRNA^{Ser}$ derived suppressor, SupD.

Table S1: Oligonucleotides used in this study.
Supplementary References

Initial design and trials in SerRS evolution

We initially tested two assumptions: (i) that increasing the distance between N- and C-terminal domains of SerRS may help accommodate long acceptor-TψC arm of $tRNA^{Sec}$ and its variants and (ii) that the mutagenesis of an N-domain serine residue (S55), implicated in $tRNA^{Sec}$ vs. $tRNA^{Ser}$ discrimination [1], could help improve $tRNA^{Sec}$ recognition. Because $tRNA^{UTuX}$ allowed a wider range of Cm levels at which we could asses suitability of particular SerRS mutations (Figure S1, S2), and was seemingly a better substrate for wild-type SerRS, we decided to test our SerRS variants using this particular $tRNA^{Sec}$ derivative.

To create a SerRS variant with increased flexibility between its N-and C-domains, a sequence coding for a GGGGS linker was inserted before glycine 114. Because canonical $tRNA^{Sec}$ has a long, atypical, 13-base pair acceptor-TψC arm, this feature was speculated to be the main reason for the lack of efficient charging by *E. coli* SerRS [2,3]. However, *in vivo* suppression assay using CAT-112TAG reporter revealed that DH10B *ΔselAB*::zeo cells expressing the linker-containing variant were slightly more sensitive to Cm than cells expressing wt SerRS (Figure S3). This suggests that the linker-containing SerRS variant serylates $tRNA^{UTuX}$ less proficiently than wt SerRS.

When SerRS-S55A variant was assayed using CAT-112TAG reporter both wt SerRS and SerRS-S55A variant facilitated cell growth at similar Cm concentrations (≤ 75 μg/ml, Fig. S3), revealing that S55A mutation has little if any influence on EcSerRS and $tRNA^{UTuX}$ interaction.

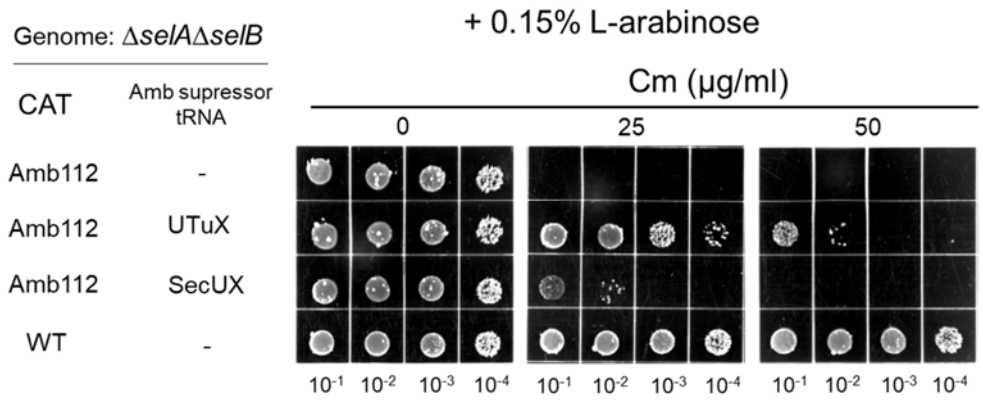

Figure S1. Nonsense suppression efficiency of $tRNA^{UTuX}$ and $tRNA^{SecUX}$. *E. coli* strain DH10B *ΔselAB*::zeo expressing chloramphenicol (Cm) acetyltransferase (CAT) with an amber codon at position 112 and $tRNA^{UTuX}$ or $tRNA^{SecUX}$ was grown on LB with different concentration of Cm (0, 25, 50 μg/ml). Cells not expressing any suppressor tRNAs were used to monitor levels of near-cognate suppression. Cells expressing wild-type (wt) CAT were used to determine maximum expression. Wild-type SerRS was induced with 0.15% L-arabinose.

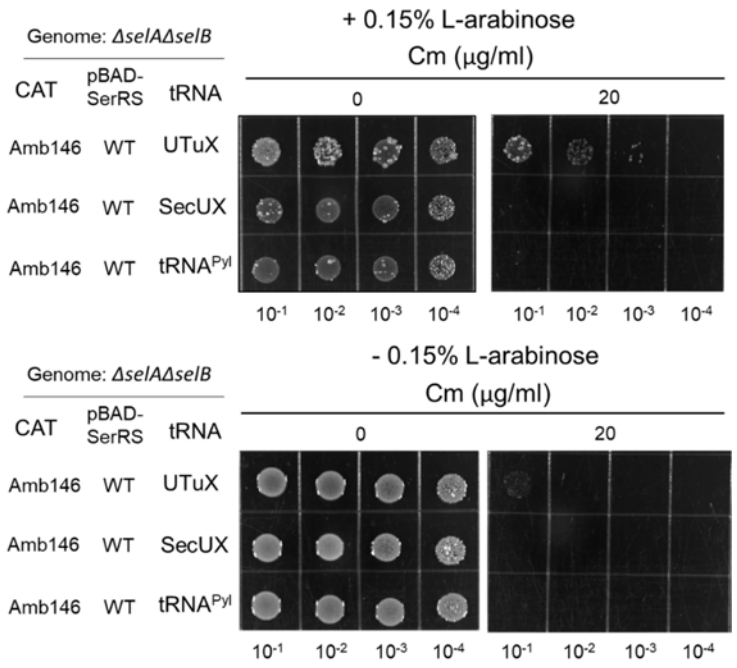

Figure S2. Nonsense suppression efficiency using a serine-specific reporter, CAT-S146TAG. DH10B *ΔselAB*::zeo cells expressing different amber suppressors ($tRNA^{UTuX}$, $tRNA^{SecUX}$ or $tRNA^{Pyl}$) and CAT-S146TAG were grown on LB plates supplemented with various concentration of Cm (0 and 20 μg/ml). tRNA under the control of an *lpp* promoter and the *cat* gene with an amber codon at position 146 were in the same plasmid. SerRS-coding sequence was placed under the P_{BAD} promoter on the second, high-copy plasmid. SerRS expression was induced by 0.15% L-arabinose. Suppressors which are not recognized by *E. coli* SerRS cannot facilitate growth, as demonstrated by attempted use of $tRNA^{Pyl}$.

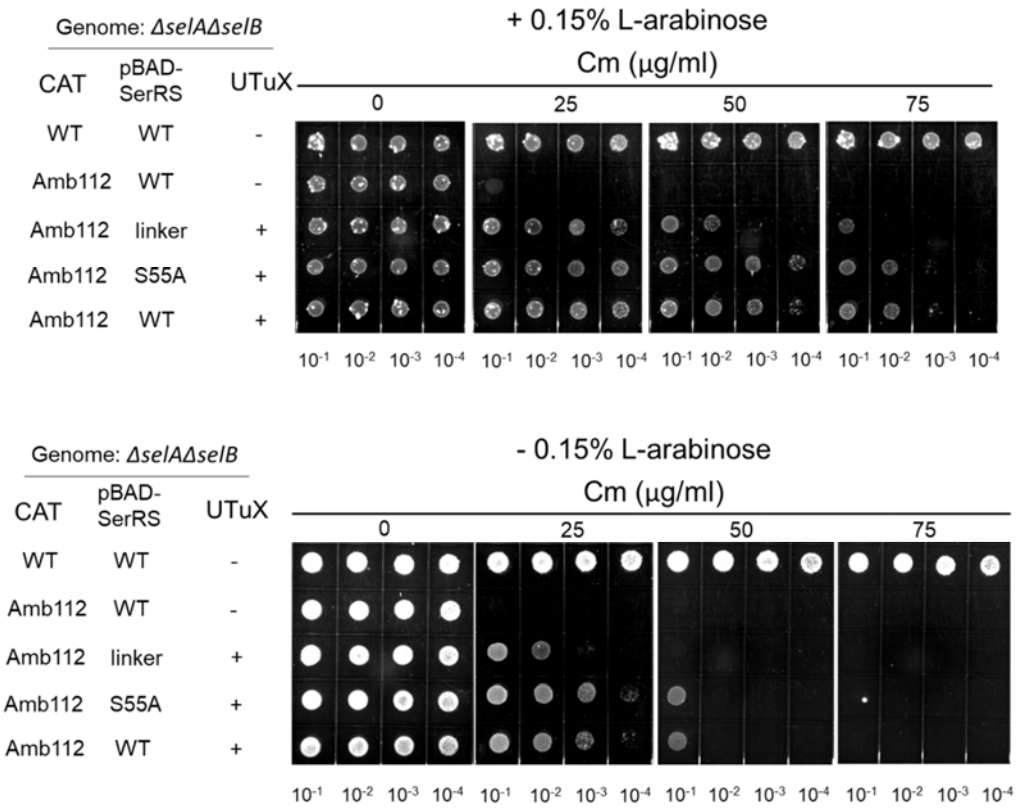

Figure S3. Serylation of $tRNA^{UTuX}$ by SerRS variants with a flexible linker and S55A mutation as monitored by a nonsense suppression assay. *E. coli* cells DH10B *ΔselAB*::zeo expressing $tRNA^{UTuX}$ and SerRS (wt and mutant enzymes as indicated) were grown on LB plates supplemented with different concentrations of Cm (0, 25, 50, 75 μg/ml). $tRNA^{UTuX}$ was under the control of an *lpp* promoter and SerRS was placed under the P_{BAD} promoter; these two genes were on the same high-copy plasmid. The CAT gene with an amber codon at position 112 was on the second plasmid. SerRS expression was induced with 0.15% L-arabinose. Cells not expressing any suppressor tRNAs were used to monitor levels of near-cognate suppression. Cells expressing wild-type (wt) CAT were used to determine maximum expression.

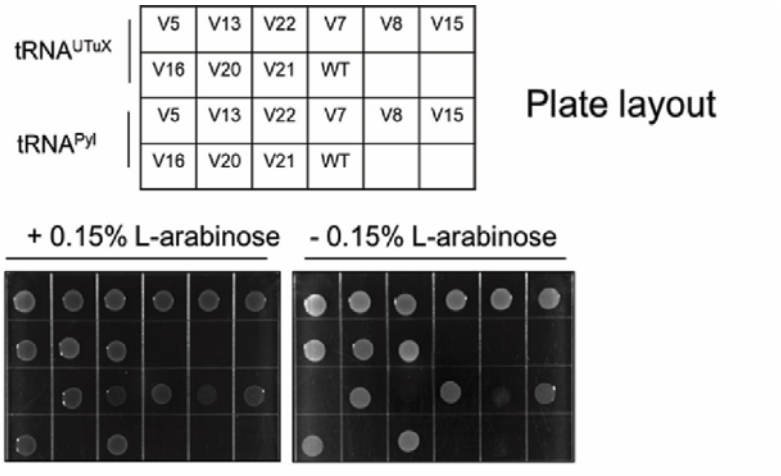

Figure S4. Reexamination of selected SerRS variants in a $tRNA^{UTuX}$-dependent nonsense suppression assay. Using a serine-specific CAT-S146TAG reporter in combination with either $tRNA^{UTuX}$ or $tRNA^{Pyl}$, a total of 22 selected SerRS clones were assayed in the DH10B *ΔselAB*::zeo strain. Cell survival was examined at 65 μg/ml Cm. Four clones (SerRS V5, V8, V20, and V22) were found to allow much higher Cm resistance than the wt SerRS; later experiments showed that the activity of variant V5 is not stable. The activity of other SerRS clones was not dependent on $tRNA^{UTuX}$ expression and those clones were removed from the study. Although SerRS-encoding sequences were placed under the P_{BAD} promoter, no difference in growth was observed when arabinose was omitted from the screen. This arabinose-independent expression occurred due to an in-frame deletion of the AraC N-terminus. The N-terminal arm of AraC is involved in the light switch mechanism where a significant structural change occurs upon arabinose binding [4]. Because AraC mutations allow constitutive expression from P_{BAD} promoter [5] the four SerRS clones were expressed even when arabinose was absent from the media, thereby allowing CAT biosynthesis and subsequently Cm resistance.

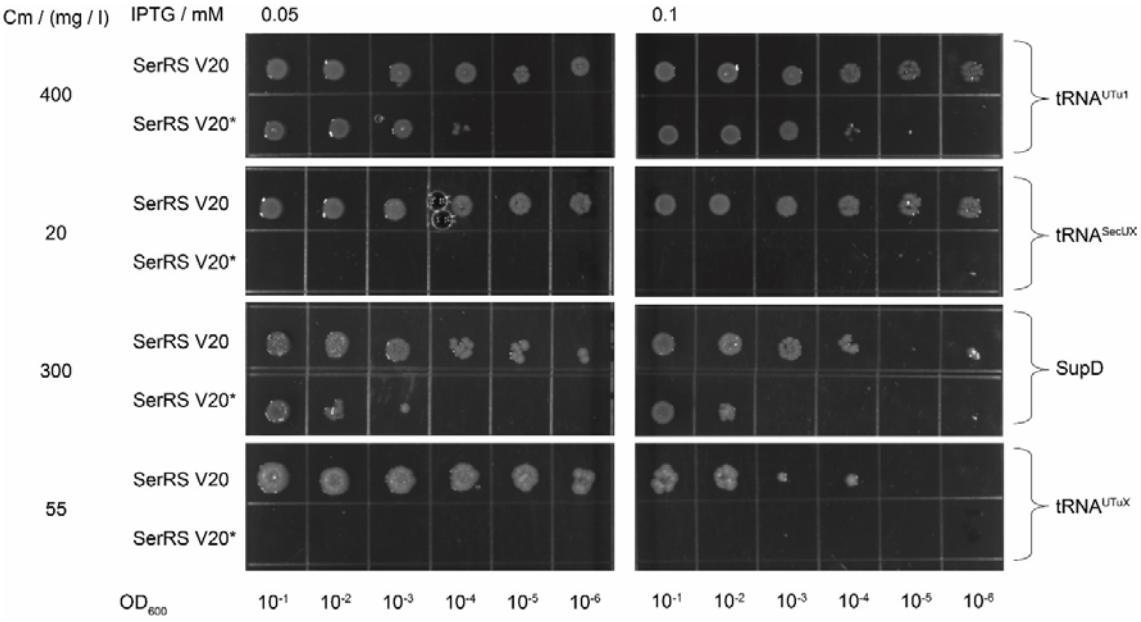

Figure S5. Serine 55-to-alanine mutation is essential for SerRS V20 activity *in vivo*. *E. coli* strain WL30153 (MC4100 Δ*selAB*, [6]) expressing a suppressor tRNA ($tRNA^{UTu1}$, $tRNA^{SecUX}$, SupD or $tRNA^{UTuX}$) and SerRS V20 or SerRS V20-A55S (denoted as SerRS V20*) was grown on LB plates supplemented with different concentrations of Cm (400, 20, 300, 55 mg/l). The *cat* gene with an amber codon at position 146 and the tRNA gene under the control of an *lpp* promoter were on the same plasmid. SerRS V20 or V20-A55S were placed under the *trc* promoter on a separate plasmid. SerRS expression was induced with 0.05 or 0.1 mM IPTG.

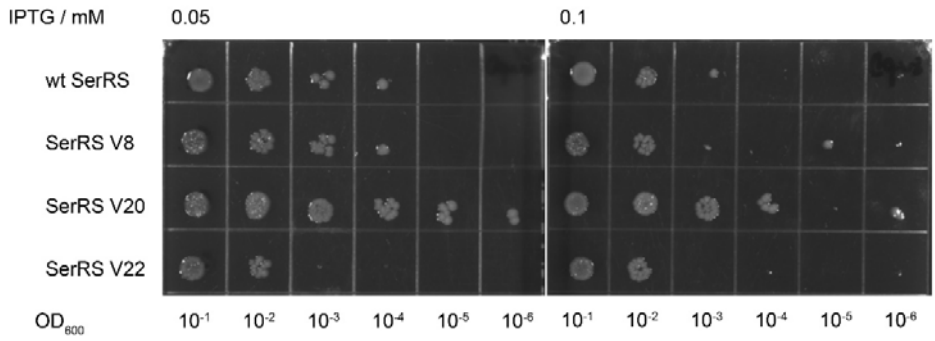

Figure S6. Stop-codon read-through using evolved variants and $tRNA^{Ser}$ derived suppressor, SupD. *E. coli* strain WL30153 [6] in which tRNA SupD and SerRS variants were expressed was grown on LB plates supplemented with 300 mg/l of Cm. The *cat* gene with an amber codon at position 146 and the tRNA gene under the control of an *lpp* promoter were on the same plasmid. SerRS variants were placed under the *trc* promoter on a separate plasmid. SerRS expression was induced with 0.05 or 0.1 mM IPTG.

Table S1. Oligonucleotides used in this study.

Primer name	Sequence	Application
pACYC.UTu1F	gttcccctccgaattcagcgttacaagtattacac	pACYC184 backbone amplification for $tRNA^{UTu1}$ gene insertion
pACYC.UTu1R	ccgccaggatcctctagagtcgacctg	
UTu1.F	cgctgaattcggaggggaacttctatctgg	$tRNA^{UTu1}$ gene amplification
UTu1.R	tctagaggatcctggcggaggggaagggaatc	
pACYC.selCF	cgatcttccgaattcagcgttacaagtattacac	pACYC184 backbone amplification for selC(amb) insertion
pACYC.selCR	ttccgccaggatcctctagagtcgacctg	
SelC.F	gctgaattcggaagatcgtcgtctccgg	selC(amb) amplification
SelC.R	ctagaggatcctggcggaagatcacaggag	
pACYC.supDF	atctctccgaattcagcgttacaagtattacacaaag	pACYC184 backbone amplification for supD insertion
pACYC.supDR	ctccgccaggatcctctagagtcgacctgcagatccttag	
supD.F	acgctgaattcggagagatgccggagcg	supD amplification
supD.R	ctagaggatcctggcggagagaggggg	
A55S.F	agcgtaactcccgatcgaaatccattggccaggcgaaagc	SerRS V20 mutagenesis
A55S.R	caatggatttcgatcgggagttacgctccgcttgc	
EcSer.Fi	cagcagattgggatcgagcat	pET15 backbone amplification
EcSer.Ri	gaacggactggaatatattggc	
EcSer.F	atgctcgatcccaatctgctg	serS gene amplification to generate pET15-serS
EcSer.R	gccaatatattccagtccgttc	
p18UTu1.F	gttcccctcctatagtgagtcgtattaggatccccg	pUC18 backbone amplification for pUC18-tRNA(UTu1)
p18UTu1.R	ctccgccaggcataagcttggcgtaatc	
18UTu1.F	cactataggaggggaacttctatctgg	tRNA(Utu1) amplification for pUC18-tRNA(UTu1)
18UTu1.R	tgcctggcggaggggaagg	
p18supD.F	ctccggcatctctcctatagtgagtcgtattaggatccccg	pUC18 backbone
p18supD.R	ctctctccgccaggcataagcttggcgtaatc	

		amplification for pUC18-supD
18supD.F	ctataggagagatgccggagcggc	supD amplification for pUC18-supD
18supD.R	cttatgcctggcggagagaggg	
p18UTuX.F	gcaccatcttcctatagtgagtcgtattaggatccccg	pUC18 backbone amplification for pUC18-tRNA(UTuX)
p18UTuX.R	gcatcttccgccaggcataagcttggcgtaatc	
18UTuX.F	cactataggaagatggtgccgtcc	tRNA(UtuX) amplification for pUC18-tRNA(UTuX)
18UTuX.R	tatgcctggcggaagatgcagg	
p18SecUX.F	cgaccatcttcctatagtgagtcgtattaggatccccg	pUC18 backbone amplification for pUC18-tRNA(SecUX)
p18SecUX.R	cttgcatcttccgccaggcataagcttggcgtaatc	
18SecUX.F	ctcactataggaagatggtcgtctccg	tRNA(SecUX) amplification for pUC18-tRNA(SecUX)
18SecUX.R	gcctggcggaagatgcaaggag	
tRNA_184_FW	caacatgaatggtcttcgggaaagcaaattcgacccgggcg	*lpp*-tRNA-*rrnC* cassette amplification for pXF113/114/115
tRNA_184_RV	gtttaccggtttattgactacgatcgtgctcctgtcgttg	
inV184_FW	gtagtcaataaaccggtaaaccag	pACYC184 backbone amplification for pXF113/114/115
inV184_RV	ccgaagaccattcatgttgttg	
Linker FW	ggtggtggtggcagcggtaaagacgaaaatgacaac	add GGGGS linker before G114 of SerRS
Linker RV	gctgccaccaccacctaccggcacttcatctgcag	
S55A FW	gcggagcgtaactcccgagcgaaatccattggccaggcg	mutate Ser55 of SerRS to Ala
S55A RV	cgcctggccaatggatttcgctcgggagttacgctccgc	
tRNA_426_FW	tccccgaaaagtgccaccgccaccctggcgccgcttcttt	*lpp*-tRNA-*rrnC* cassette amplification for pBAD-SerRS plasmids
tRNA_426_RV	catttgacaggcacattatgttgaggacccggctaggctagatc	
inV426_FW	cataatgtgcctgtcaaatggacgaagc	pACYC184 backbone amplification
inV426_RV	gcggtggcacttttcggggaaa	

pTRC_R	catggtctgtttcctgtgtg	pTrc99A backbone amplification
pTRC_F	tgttttggcggatgagagaag	
Ser_pTRC_F	caggaaacagaccatgctcgatcccaatctgctgc	serS gene amplification for pTrc99A
Ser_pTRC_R	ctcatccgccaaaacattagccaatatattccagtccgttc	
Frag1 FW	gtgtctataatcacggcagaaaagtccac	To generate SerRS inserts fully randomized at given positions (residues K86, A87, D90, S160 and D228 and T229)
Frag1 RV	cgatatcgcgaatttcagcctgtaaagcmnncagctcmnnmnntgctgcatccagctcttcgcc	
Frag2 FW	gctttacaggctgaaattcgcgatatcg	
Frag2-1 RV	caatctgccctttcattaccacaaagcgAHNaccagtcagcttaactgcagctgc	
Frag2-2 RV	caatctgccctttcattaccacaaagcgCDBaccagtcagcttaactgcagctgc	
Frag2-3 RV	caatctgccctttcattaccacaaagcgCCAaccagtcagcttaactgcagctgc	
Frag3 FW	cgctttgtggtaatgaaagggcagattg	
Frag3-1 RV	ccgttgggatcagcgcatagttactAHNAHNtgcttcttcttccagcggacgag	
Frag3-2 RV	ccgttgggatcagcgcatagttactCDBCDBtgcttcttcttccagcggacgag	
Frag3-3 RV	ccgttgggatcagcgcatagttactCCACCAtgcttcttcttccagcggacgag	
Frag3-4 RV	ccgttgggatcagcgcatagttactCDBAHNtgcttcttcttccagcggacgag	
Frag3-5 RV	ccgttgggatcagcgcatagttactCCAAHNtgcttcttcttccagcggacgag	
Frag3-6 RV	ccgttgggatcagcgcatagttactCCACDBtgcttcttcttccagcggacgag	
Frag3-7 RV	ccgttgggatcagcgcatagttactAHNCDBtgcttcttcttccagcggacgag	
Frag3-8 RV	ccgttgggatcagcgcatagttactAHNCCAtgcttcttcttccagcggacgag	
Frag3-9 RV	ccgttgggatcagcgcatagttactCDBCCAtgcttcttcttccagcggacgag	
Frag4 FW	agtaactatgcgctgatcccaacgg	
Frag4 RV	ctgtgccgggatccatacttcc	
EM7-zeo FW	ggcttgatgcgcgatatgtcctcctgacccatctcacgttacaatccgtgcggatctgatcagcacgtgttgac	EM7-zeocin cassette amplification for DH10B Δ*selAB*::zeo strain
EM7-zeo RV	ccttatttttccggaaataataatgcgtcgcgcaataaatgatcatttcctcagtcctgctcctcggcc	
EcSerSf	cagaccatggaaatgctcgatcccaatctgctgc	amplification of the serS gene to generate pBAD-serS plasmids
EcSerSr	agccggtaccttagccaatatattccagtccgttcatatacg	
supF 113 FW	ctctaaatctgccgtcatcgacttcgaaggttcgaatccttcccccaccaccaggatcctctagagtcgacctg	Replace $tRNA^{UTuX}$ in pXF113 by supF
supF 113 RV	ctttgtgtaatacttgtaacgctgaattcggtggggttcccgagcggccaaagggagcagactctaaatctgccgtcat	

Table 2. Steady-state kinetic parameters for SupD in case of wild-type SerRS and SerRS variants V8 and V20. All reactions were performed at 37 °C. Concentration of SupD was corrected for plateau charging (29.7 %). In aminoacylation with SupD enzyme concentrations were: 30 nM wild-type SerRS, 15 nM SerRS V8, 15 nM SerRS V20.

enzyme	K_M (μM)	k_{cat} (min^{-1})	k_{cat} / K_M (min^{-1} μM^{-1})	relative loss
wt	0.76	95	130	1
V8	1.6	47	29	4.3
V20	1.3	31	23	5.5

Supplementary references:

[1] Wang, C., Guo, Y., Tian, Q., Jia, Q., Gao, Y., Zhang, Q., Zhou, C. and Xie, W. (2015). SerRS-tRNA^{Sec} complex structures reveal mechanism of the first step in selenocysteine biosynthesis. Nucleic Acids Res 43, 10534-45.

[2] Baron, C. and Böck, A. (1991). The length of the aminoacyl-acceptor stem of the selenocysteine-specific tRNA^{Sec} of *Escherichia coli* is the determinant for binding to elongation factors SELB or Tu. J Biol Chem 266, 20375-9.

[3] Holman, K.M., Puppala, A.K., Lee, J.W., Lee, H. and Simonović, M. (2017). Insights into substrate promiscuity of human seryl-tRNA synthetase. RNA 23, 1685-1699.

[4] Schleif, R. (2010). AraC protein, regulation of the *L*-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. FEMS Microbiol Rev 34, 779-96.

[5] Ross, J.J., Gryczynski, U. and Schleif, R. (2003). Mutational analysis of residue roles in AraC function. J Mol Biol 328, 85-93.

[6] Leinfelder, W., Forchhammer, K., Zinoni, F., Sawers, G., Mandrand-Berthelot, M.A. and Böck, A. (1988). *Escherichia coli* genes whose products are involved in selenium metabolism. J Bacteriol 170, 540-6.