www.sciencemag.org/cgi/content/full/333/6046/1151/DC1

Supporting Online Material for

Expanding the Genetic Code of *Escherichia coli* with Phosphoserine

Hee-Sung Park, Michael J. Hohn, Takuya Umehara, Li-Tao Guo, Edith M. Osborne, Jack Benner, Christopher J. Noren,* Jesse Rinehart,* Dieter Söll*

*To whom correspondence should be addressed. E-mail: dieter.soll@yale.edu (D.S.); noren@neb.com (C.J.N.); jesse.rinehart@yale.edu (J.R.)

Published 26 August 2011, *Science* **333**, 1151 (2011)
DOI: 10.1126/science.1207203

This PDF file includes:

Materials and Methods

Construction of strains. To prevent possible enzymatic dephosphorylation of *O*-phospho-L-serine (Sep) *in vivo*, the gene encoding phosphoserine phosphatase (*serB)*, which catalyzes the last step in serine biosynthesis, was deleted from *Escherichia coli* strains Top10 and BL21. Markerless gene deletions were carried out using a λ-red and FLP recombinase-based gene knockout strategy as described earlier (*27*). *E. coli* strains Top10Δ*serB* and BL21*ΔserB* were used for EF-Tu library construction and MEK1 expression experiments.

Construction of plasmids. To construct plasmid pSepT, the full-length gene encoding $tRNA^{Sep}$ was constructed from overlapping oligonucleotides and ligated immediately downstream of the *lpp* promoter in pTECH (*28*) using *Eco*RI and *Bam*HI restriction sites. pCysT, encoding the wild type $tRNA^{Cys}$ gene from *Methanocaldococcus jannaschii* was constructed in the same way.

pKD was derived from pKK223-3 (Pharmacia). The ampicillin (Amp) resistance gene was replaced with a kanamycin (Kan) resistance gene by combining two PCR products generated from pKK223-3 and pET28a. The following PCR primers were used: PKF (5'-TGC AGCA ATG CGG CCG CTT TCA CCG TCA TCA CCG AAA C-3') and PKR (5'-GGG ACG CTA GCA AAC AAA AAG AGT TTG TAG AA-3') for pKK223-3 amplification and PKNF (5'-GGG ACG CTA GCT TTT CTC TGG TCC CGC CGC AT-3') and PKNR (5'-TGC GCA ATG CGG CCG CGG TGG CAC TTT TCG GGG AAA T-3') for Kan^R gene amplification.

The original multiple cloning site (MCS; *Nco*I-*Eco*RI-*Sac*I) was modified by adding an additional ribosome binding site (RBS) and a second MCS (*Nde*I-*Bam*HI-*Sal*I-*Hind*III), thus enabling simultaneous protein expression from two genes, both under the control of the same *tac* promoter. The *Methanococcus maripaludis* SepRS gene was cloned into pKD using *Nco*I and *Sac*I sites to produce pKD-SepRS. The *E. coli* EF-Tu gene (*tufB*) was ligated into pKD-SepRS using *Bam*HI and *Sal*I sites resulting in pKD-SepRS-EFTu. The *M. maripaludis pscS* gene encoding SepCysS was cloned into pKD-SepRS using *Bam*HI and *Sal*I sites to produce pKD-SepRS-SepCysS and the *M. maripaludis* CysRS gene was cloned into pKD using *EcoR*I and *Sal*I to yield pKD-CysRS.

pCAT112TAG-SepT was created from pACYC184. The gene encoding chloramphenicol acetyltransferae (CAT) was modified by quickchange mutagenesis to introduce an amber stop codon at position Asp112 (*29*). The resulting plasmid was PCR amplified using primers PBLAF (5'-TGC GCA ATG CGG CCG CCC GTA GCG CCG ATG GTA GTG T-3') and PBLAR (5'-ACA CGG AGA TCT CTA AAG TAT ATA TGA GTA AAC-3') and ligated with a PCR product containing a $tRNA^{Sep}$ expression cassette from pSepT, created with primers TSEPF (5'-GCA TGC GCC GCC AGC TGT TGC CCG TCT CGC-3') and TSEPR (5'-GCA TAG ATC TTC AGC TGG CGA AAG GGG GAT G-3').

Plasmid pCcdB was created by adding a *ccdB* gene under the control of a *lac* promoter into pTECH (*28*) using *Not*I and *Bgl*II sites (*30*). Two amber stop codons were introduced at positions 13 and 44 based on the crystal structure and mutagenesis study of the CcdB protein (*31, 32*).

pMYO127TAG-SepT was constructed by cloning a codon-optimized and C-terminally His_6-tagged sperm whale myoglobin gene under the control of the *lpp* promoter between *Not*I and *Bgl*II in pSepT. An amber stop codon was introduced to the myoglobin gene at position Asp127 by quickchange mutagenesis. The nucleotide sequence of the codon-optimized myoglobin gene is as follows:

ATGGTTCTGTCTGAAGGTGAATGGCAGCTGGTTCTGCACGTTTGGGCTAAAGTTGAAGCTGAC
GTTGCTGGTCACGGTCAGGACATCCTGATCCGTCTGTTCAAATCTCACCCGGAAACCCTGGA
AAAATTCGACCGTTTCAAACACCTGAAAACCGAAGCTGAAATGAAGGCTTCTGAAGACCTGAA
AAAACACGGTGTTACCGTTCTGACCGCTCTGGGTGCTATCCTGAAGAAAAAGGGTCACCACG
AAGCTGAACTGAAACCGCTGGCTCAGTCTCACGCTACCAAACACAAAATCCCGATCAAATACC
TGGAGTTCATCTCTGAAGCTATCATCCACGTTCTGCACTCTCGTCATCCGGGTAACTTCGGTG
CTGACGCTCAGGGTGCTATGAACAAAGCTCTGGAACTGTTCCGTAAAGACATCGCTGCTAAAT
ACAAAGAACTGGGTTACCAGGGTGGTTCTGGTCATCACCATCACCATCACTAA

Plasmid pL11C-SepT encodes $tRNA^{Sep}$ and the C-terminal domain of the ribosomal protein L11 under control of *lpp* promoters. Part of the *rplK* gene was PCR amplified from genomic *E. coli* DNA using primers L11C-F (5'-GGA ATT CCA TAT GAC CAA GAC CCC GCC GGC AGC AGT T-3') and L11C-R (5'-AGG CGC GCC TTA GTC CTC CAC TAC-3'). The PCR product was digested with *Nde*I and *Asc*I and was ligated into *Nde*I and *Asc*I digested pMYO127TAG-SepT to replace the myoglobin gene.

To construct pMAL-EFTu and pMAL-EFSep *E. coli tufB*, or the gene encoding EF-Sep, respectively, were cloned between the *Nde*I and *Bam*HI sites in the pET20b plasmid (Novaven) to add a C-terminal His_6 tag. This fusion construct was then PCR-amplified using primers adding *Mfe*I and *Pst*I restriction sites. The PCR product was cloned in-frame between *Eco*RI and *Pst*I in pMAL-c2x (New England Biolabs) to add an N-terminal maltose binding protein (MBP) tag.

pET15-ERK2 encodes N-terminally His_6-tagged mitogen-activated protein kinase (Erk2) under the control of a T7 promoter. The human Erk2 gene was PCR amplified from plasmid BC017832 (ATCC) using primers ERK2-F (5'-GGA ATT CCA TAT GGC GGC GGC GGC GGC G-3') and ERK2-R (5'-CCG CTC GAG TTA AGA TCT GTA TCC TGG-3'). The PCR product was cloned between *Nde*I and *Xho*I in vector pET15b (Novagen).

pET20-MBPMEK1 encodes a fusion protein consisting of human MEK1 with an N-terminal maltose binding protein (MBP) tag and a C-terminal His_6-tag. The gene encoding human MEK1 which was codon-optimized for *E. coli* and custom-synthesized *in vitro* (Genscript), was cloned between *Eco*RI and *Pst*I into pMAL-c2x (New England Biolabs). The resulting MBP-MEK1 fusion construct was then amplified with primers ET20MEKF (5'-AAG GAA ATT AAT GAA AAT CGA AGA AGG TAA-3') and ET20MEKR (5'-CTA GAG GAT CCG GCG CGC-3') adding *Ase*I and *Bam*HI restriction sites, and the PCR product was ligated between *Nde*I and *Bam*HI into pET20b.

Nucleotide sequence of codon-optimized MEK1:

ATGCCGAAGAAGAAACCGACCCCGATCCAGCTGAACCCGGCTCCGGACGGTTCTGCTGTTAA
CGGCACCTCTTCTGCTGAAACCAACCTGGAAGCTCTGCAAAAGAAACTGGAAGAACTGGAAC
TGGACGAACAGCAGCGTAAACGTCTGGAAGCGTTCCTGACCCAGAAACAGAAAGTTGGTGAA
CTGAAAGACGACGACTTCGAAAAAATCTCTGAACTGGGTGCTGGTAACGGTGGTGTTGTTTTC
AAAGTTTCTCACAAACCGTCCGGTCTGGTTATGGCTCGTAAACTGATCCACCTGGAAATCAAA
CCGGCTATCCGTAACCAGATCATCCGTGAACTGCAAGTTCTGCACGAATGCAACTCTCCGTAC
ATCGTTGGTTTCTACGGTGCTTTCTACTCTGACGGTGAAATCTCTATCTGCATGGAACACATG
GACGGTGGTTCTCTGGACCAGGTTCTGAAAAAAGCTGGTCGTATCCCGGAACAGATCCTGGG
TAAAGTTTCTATCGCTGTTATCAAAGGTCTGACCTACCTGCGTGAAAAACACAAAATCATGCAC
CGTGACGTTAAACCGTCTAACATCCTGGTTAACTCTCGTGGTGAAATCAAACTGTGCGACTTC
GGTGTTTCTGGTCAGCTGATCGACTCTATGGCTAACTCTTTCGTTGGCACCCGTTCTTACATG

TCTCCGGAACGTCTGCAAGGCACCCACTACTCTGTTCAGTCTGACATCTGGTCTATGGGTCTG
TCTCTGGTTGAAATGGCTGTTGGTCGTTACCCGATCCCGCCGCCGGACGCTAAAGAACTGGA
ACTGATGTTCGGTTGCCAGGTTGAAGGTGACGCTGCTGAAACCCCGCCGCGTCCGCGTACTC
CGGGTCGTCCGCTGTCTTCTTACGGTATGGACTCTCGTCCGCCGATGGCTATCTTCGAACTG
CTGGACTACATCGTTAACGAACCGCCGCCGAAACTGCCGTCTGGTGTTTTCTCTCTGGAGTTC
CAGGACTTCGTTAACAAATGCCTGATCAAAAACCCGGCTGAACGTGCTGACCTGAAACAGCTG
ATGGTTCACGCTTTCATCAAACGTTCTGACGCTGAAGAAGTTGACTTCGCTGGTTGGCTGTGC
TCTACCATCGGTCTGAACCAGCCGTCTACCCCGACCCACGCTGCTGGTGTGGCAGCCGCAG
CTGCGCATCATCACCACCATCACTAA

pCG-MBPMEK1SS was generated by the ligation of three PCR products. One PCR product was derived from pGFIB (*33*) using primers GFIB-F (5'-ATA AGA ATG CGG CCG CGC CGC AGC CGA ACG ACC GAG-3') and GFIB-R (5'-CTA GCT AGC GTC TGA CGC TCA GTG GAA CG-3'). The second PCR product was generated from pCDFDuet-1 (Novagen) using primers CDF-F (5'-CTA GCT AGC TCA CTC GGT CGC TAC GCT-3') and CDF-R (5'-ATA AGA ATG CGG CCG CTG AAA TCT AGA GCG GTT CAG-3'). Both PCR products were digested with *Nhe*I and *Not*I and ligated to form plasmid pCG. The third PCR product, encoding an expression cassette for MBP-MEK1-His_6 under the control of T7 promoter and T7 terminator, was generated from pET20-MBPMEK1 using primers ETCDGFF (5'- AAA AGG CGC CGC CAG CCT AGC CGG GTC CTC AAC G-3') and ETCDGFR (5'- AAC TGC AGC CAA TCC GGA TAT AGT TC-3'). This PCR product was cloned between the *Nar*I and *Pst*I sites of pCG.

The codon for Ser 222 in MEK1 was then replaced by a GAA codon (encoding Glu) using Quickchange mutagenesis (Stratagene). In the same way, codon Ser 218 was either changed to GAA to generate pCG-MBPMEK1EE, or to an amber stop codon, resulting in pCG-MBPMEK1XE. In pCG-MBPMEK1XS only the codon for Ser218 was changed to UAG and in pCG-MBPMEK1XX both codons for Ser218 and Ser222 were changed to amber.

Library construction and selection of Sep-tRNA specific EF-Tu. Six residues, His67, Asp216, Glu217, Phe219, Thr229, and Asn274, located in the amino acid binding pocket of the *E. coli* elongation factor EF-Tu were selected for randomization based on the crystal structure of the *E. coli* EF-Tu:Phe-$tRNA^{Phe}$ complex (protein data base accession number 1OB2). Multiple rounds of overlap PCR were carried out to incorporate random codons (NNK) at these positions (*5*) by using the following primers: 67XF, 5'-GT ATC ACC ATC AAC ACT TCT NNK GTT GAA TAC GAC ACC CCG-3'; H67R, 5'-AGA AGT GTT GAT GGT GAT AC-3'; 216XF, 5'-CCG TTC CTG CTG CCG ATC NNK NNK GTA NNK TCC ATC TCC GGT CGT GGT-3'; 216R, 5'-GAT CGG CAG CAG GAA CGG-3'; 229XF, 5'-GGT CGT GGT ACC GTT GTT NNK GGT CGT GTA GAA CGC GG-3'; 229R, 5'-AAC AAC GGT ACC ACG ACC-3'; 274XF, 5'-GAA GGC CGT GCT GGT GAG NNK GTA GGT GTT CTG CTG CG-3'; 274R, 5'-CTC ACC AGC ACG GCC TTC-3'. The final PCR products were purified and digested with *Bam*HI and *Sal*I, and ligated into pKD-SepRS to generate the EF-Tu library. The ligated vectors were transformed into *E. coli* Top10Δ*serB* containing pCAT112-SepT to generate a library of 3×10^8 mutants. The unbiased mutation of the library was confirmed by selecting twenty random clones and sequencing each mutant *tufB* insert.

The mutant EF-Tu library was subjected to a first round of selection, in which clones suppressing the amber stop codon in the CAT gene can survive on LB plates supplemented with 10

µg/ml tetracycline (Tc), 25 µg/ml Kan, 50 µg/ml chloramphenicol (Cm), 2 mM Sep, and 0.05 mM isopropyl-β-D-thiogalactopyranoside (IPTG). After 48h incubation at 30°C, a pool of ca. 10^4 colonies was collected from the plates for plasmid preparation. The pKD-SepRS-EFTu plasmids were separated from the reporter plasmid by agarose gel electrophosis and isolated using the Qiagen gel purification kit.

To assure that UAG codon suppression in the CAT gene was dependent on the presence of EF-Sep and Sep-$tRNA^{Sep}$ and did not arise from a hypothetical insertion of a canonical amino acid at this site, EF-Tu mutants were tested in a negative selection experiment in the absence of $tRNA^{Sep}$. The pKD-SepRS-EFTu plasmids from the first round of positive selection were transformed into *E. coli* Top10Δ*serB* cells harboring pCcdB. Suppression of both UAG codons in the pCcdB-encoded *ccdB* gene will result in expression of the toxic *ccdB* gene product and will therefore eliminate EF-Tu mutants potentially mediating UAG read through with an aminoacylated non-suppressor tRNA. Transformants were plated onto LB agar supplemented with 25 µg/ml Kan, 25 µg/ml Cm, and 0.1 mM IPTG. After 48h incubation at 30°C, twenty individual clones were picked and subjected to plasmid purification to isolate pKD-SepRS-EFTu as described above.

Recovered pKD-SepRS-EFTu plasmids were transformed into *E. coli* Top10Δ*serB* containing pCAT112-SepT for a third round of selection which was carried out under the same conditions as the first. This time, individual colonies were isolated from agar plates and clones were tested for their ability to grow on Cm over a concentration range from 5 to 100 µg/ml. Total plasmid was purified from isolates showing strong Cm resistance, and pKD-SepRS-EFTu plasmids were subjected to sequencing.

To confirm that the observed Cm resistance is dependent on the presence of both mutant EF-Tu and SepRS, EF-Tu mutant genes were excised from their plasmids, recloned into pKD, and retransformed into *E. coli* Top10*ΔserB* containing pCAT112-SepT. Cells were then tested for Cm resistance as described above.

Protein expression and purification

Expression and purification of M. maripaludis SepRS and CysRS. SepRS and CysRS were produced in *E. coli* and purified as described elsewhere (*5*).

Expression and purification of EF-Tu and EF-Sep*.* pMAL-EFTu or pMAL-EFSep were transformed into *E. coli* BL21 (DE3) codon plus (Stratagene). A pre-culture was used to inoculate 1000 ml of LB broth with 100 µg/ml of Amp, 34 µg/ml Cm, 5052 solution, and phosphate buffer for autoinduction as described previously (*34*). The cells were grown for 6 h at 37°C and continued at 20°C for 18 h.

The cells were pelleted and lysed by shaking for 20 min. in BugBuster (Novagen) reagent supplemented with 50 mM Tris–HCl (pH 7.6), 60 mM NH_4Cl, 7 mM $MgCl_2$, 14.3 mM 2-mercapto-ethanol, 50 µM GDP, 10% glycerol, 25 U ml^{-1} Benzonase, 1 mg ml^{-1} lysozyme, and Protease inhibitor cocktail (Roche).

The extract was clarified by ultracentrifugation and applied to a Ni^{2+}-NTA resin (Qiagen) and purified according to the manufacturer's instructions.

The eluted enzymes were dialyzed into 20 mM HEPES–KOH (pH 7.0), 40 mM KCl, 1 mM $MgCl_2$, 5 mM DTT, 50 µM GDP, and 30% glycerol. SDS–PAGE electrophoresis followed by staining with Coomassie blue revealed greater than 95% purity.

Expression and purification of Myoglobin. To express mutant myoglobin, pKD-SepRS-EF-Sep and pKD-SepRS were transformed into *E. coli* Top10*ΔserB* containing pMYO127TAG-SepT. *E. coli* Top10*ΔserB* with pMYO, encoding the wild type myoglobin gene was used as a control. Cultures were grown in LB medium supplemented with 2 mM Sep. When A600 reached 0.6 protein expression was induced with 0.05mM IPTG for 12h at 25°C. The cells were harvested, resuspended in lysis buffer (50mM Tris-HCl (pH 7.8), 300mM NaCl, 14.3 mM 2-mercaptoethanol) supplemented with protease inhibitor cocktail (Roche), and subjected to sonication. The lysate was centrifuged at 10,000×g for 30 min and the supernatant was applied to Ni^{2+}-NTA agarose (Qiagen) purification according to the manufacturer's instruction. The yield of purified myoglobin was 2mg/L of culture.

Expression and purification of MEK1. To express MEK1 (as a maltose binding protein fusion-protein) *E. coli* BL21Δ*serB* was transformed with plasmids pKD-SepRS-EFSep, pCAT112TAG-SepT, and pCG-MBPMEK1SS, pCG-MBPMEK1EE, pCG-MBPMEK1XE, pCG-MBPMEK1XS, or pCG-MBPMEK1XX, respectively. Plasmid pCAT112TAG-SepT was replaced by pL11C-SepT in the strain used to produce MBP-MEK1(Sep218,Ser222)-His_6 for mass spectrometry analysis.

Cells were grown at 30°C in 1 liter of LB supplemented with 100 µg/ml of Amp, 50 µg/ml Kan, 12 µg/ml Tc, 2mM Sep, 5052 solution, and phosphate buffer for autoinduction as described previously (*34*). When A600 reached 0.6, temperature was changed to 16°C and incubation continued for 18h. After harvesting, cells were lysed in 20 ml BugBuster reagent containing 50 mM Tris-HCl (pH 7.8), 500 mM NaCl, 0.5 mM EGTA, 0.5 mM EDTA, 14.3 mM 2-mercapto-ethanol, 10% glycerol, 0.03% Brij-35, protease inhibitors, 25 U ml^{-1} Benzonase, and 1 mg ml^{-1} lysozyme. The lysate was clarified by ultracentrifugation, and applied to a 0.4 ml Ni^{2+}-NTA agarose column. The column was washed with 15 ml wash buffer (50 mM Tris-HCl (pH 7.8), 150 mM NaCl, 0.5 mM EGTA, 0.5 mM EDTA, 14.3 mM 2-mercaptoethanol, 10% glycerol, 0.03% Brij-35, and 20 mM imidazole). Proteins were eluted in 0.8 ml of wash buffer supplemented with 300 mM imidazole, dialyzed against 50 mM Tris-HCl (pH 7.8), 150 mM NaCl, 0.1 mM EGTA, 5 mM DTT, 30% glycerol, and 0.03% Brij-35, and stored at -20°C. Purified proteins were analyzed by SDS-PAGE. Western blot analysis was carried out using monoclonal antibodies directed against the phosphorylated active site of MEK1 (anti-P-MEK1/2 (S217/221; Cell Signaling Technology; Cat# 9154) or against the maltose binding protein domain (anti-MBP; New England Biolabs; Cat# E8038).

Expression and purification of Erk2. *E. coli* BL21 (DE3) codon plus cells were transformed with pET15-ERK2 and grown at 37°C in 1 liter LB broth supplemented with 100 µg/ml Amp and 34 µg/ml Cm. When the cultures reached A600 of 0.6, 0.2 mM IPTG was added and expression was induced for 19 h at 16°C.

Cell lysis, Ni^{2+} purification, and dialysis of Erk2 were carried out as described for MEK1. Erk2 was 99% pure, as judged by Coomassie brilliant blue staining after SDS-PAGE.

Preparation and aminoacylation of tRNA. Total tRNA from *E. coli* Top10 or from *E. coli* Top10 complemented with pCysT or pSepT, respectively, was purified by standard procedures and acylated with [^{14}C]Sep by *M. maripaludis* SepRS as described previously (*5*). We used *in vivo* synthesized tRNA for this experiment to ensure that nucleoside modifications introduced into tRNA by *E. coli* modifying enzymes do not affect tRNA recognition by SepRS. *M. jannaschii* $tRNA^{Cys}$

contains m^1G37 when isolated from *M. jannaschii* (*35*). Since the *E. coli* methylase TrmD is known to methylate G37 of archaeal $tRNA^{Pro}$ (*36*), we assume that the *in vivo* expressed $tRNA^{Sep}$ also carries the m^1G37 modification. *In vitro* transcript of *M. jannaschii* $tRNA^{Cys}$ was prepared and acylated with [^{14}C]Sep or [^{35}S]Cys using recombinant *M. maripaludis* SepRS or CysRS as described previously (*5*). *M. jannaschii* $tRNA^{Cys}$ transcript was chosen for these experiments because of the poor folding properties of *in vitro* transcribed *M. maripaludis* $tRNA^{Cys}$ (*5*).

EF-Tu hydrolysis protection assays. To assay hydrolysis protection of acylated $tRNA^{Cys}$ by EF-Tu, Mmp $tRNA^{Cys}$ *in vitro* transcripts acylated with [^{14}C]Sep or [^{35}S]Cys, respectively, were phenol/chloroform extracted, and the aqueous phase was passed over Sephadex G25 Microspin columns (GE Healthcare) equilibrated with water. Protection of aminoacylated tRNA by EF-Tu was assayed as described earlier with slight modifications (*14*). Briefly, EF-Tu or EF-Sep (both purified as maltose binding protein fusion proteins) were activated for 20 min. at 37°C in buffer containing 100 mM Tris-HCl (pH 8.2), 120 mM NH_4Cl, 7 mM $MgCl_2$, 5 mM DTT, 5 mM phosphoenolpyruvate, 1.5 mM GTP, and 0.12 µg/µl pyruvate kinase. Hydrolysis of 2 µM [^{14}C]Sep-$tRNA^{Cys}$ was then monitored at 25°C in the presence of 40 µM EF-Tu (wt), EF-Sep, or BSA, respectively. Aliquots were taken from the reaction mix at indicated time points and spotted on 3MM filter discs presoaked with 10% trichloroacetic acid. Filters were washed with 5% trichloroacetic acid, dried, and radioactivity was measured by liquid scintillation counting.

MEK activity assays. Recombinant MEK1 variants were assayed (as maltose binding protein (MBP) fusion-proteins) by a slightly modified established method (*37*). Briefly, in a first reaction, various amounts (2.5 – 5000 ng) of recombinant MBP-MEK1 variants were used to phosphorylate (and activate) bacterially expressed MAP kinase (Erk2) for 15 min. at 30°C in 35 µl kinase assay buffer containing 12 mM MOPS pH 7.2, 20 mM $MgCl_2$, 3 mM EGTA, 15 mM β-glycerol phosphate, 0.6 mM DTT, 140 µM ATP, and 1 µg Erk2.

After 15 min, a 5 µl aliquot was transferred to a second reaction in which activated Erk2 phosphorylates myelin basic protein (MBP; 570 µg ml^{-1}) in kinase assay buffer in the presence of [γ-^{32}P]ATP. After 15 min. incubation at 30°C 25 µl aliquots were transferred onto p81 phospho-cellulose filters (Whatman). The filters were washed three times with 180 mM phosphoric acid and then rinsed with acetone. Phosphorylation was quantitated by scintillation counting and the specific activity of MEK1 was calculated from the amount of [^{32}P]phosphate incorporated into MBP.

MS TOF Analysis of Intact Myoglobin. Myoglobin-His_6 and Myoglobin-His_6 (Asp127Sep) proteins were subjected to chromatography on a PLRP-S reverse phase column (0.5 x 150 mm, 5 µm particles using an Agilent 1100 LC system at a flow rate of 10 µl/min with 15 % ACN with 0.1 % Formic acid. The column was developed with a linear gradient over 12 minutes to 100 % ACN with 0.1 % Formic acid and the effluent was directly introduced by electrospray ionization to an Agilent 6210 TOF mass spectrometer. Protein peaks were selected, extracted and deconvoluted using Bioconfirm A 2.0 software.

Ion Trap MS/MS and Data Analysis of Trypsin-Digested Myoglobin. Myoglobin-His_6 and Myoglobin-His_6 (Asp127Sep) were digested overnight with trypsin and subjected to analyses on an Agilent 6330 Ion Trap mass spectrometer with an integrated ChipCube C18 nano-column electrospray ionization. The digests were injected (8 μl) and the C18 chip (75 μm x 150 mm) with integrated 40 ηl trap was developed at 400 ηl/min using a gradient from 5% ACN to 50 % ACN with 0.1% Formic acid over 45 min. The four largest ions were selected for fragmentation by collision-induced-disassociation and excluded after two occurrences for 30 seconds. The ion trap was scanned at a rate of 25,000 Da/sec and a mass range of 150-1800 Da. Files were searched using Spectrum Mill version 3.03 with a Swissprot database that had two Myoglobin sequences concatenated to it. Phosphorylated S and T, and oxidation of M were variable modifications. Peptide and fragment mass tolerance were 1.5 and 0.7 Da, respectively, with 2 missed cleavages.

LC and MS/MS conditions for Multiple Reaction Monitoring (MRM). Purified MEK1 proteins were separated by SDS-PAGE, visualized with Comassie stain, excised, washed in 50% acetonitrile (ACN)/50 mM NH_4HCO_3, crushed, and digested at 37°C in a 20 µg/ml trypsin (Promega) solution in 10 mM NH_4HCO_3. Digested peptides in solution were dried and dissolved in 3 µl of 70% formic acid (FA), and then diluted to 10 µl with 0.1% TFA. Peptides for MRM were synthesized at the KECK peptide synthesis facility at Yale. The human MEK peptide LCDFGVSGQLIDS*MANSFVGTR (*phospho-Ser; YPED peptide ID, SOL14075) was synthesized to permit the development of a specific method for quantitative MRM. Crude synthetic peptides were direct infused at a concentration of ~10 pmol/µl and Collision Energy and Declustering Potentials of the transitions were optimized. LC-MRM was performed on an ABI 5500 QTRAP triple quadrupole mass spectrometer inter-faced with a Waters nanoAcquity UPLC system running Analyst 1.5 software. Peptides were resolved for MRM (LC step) by loading 4 µl of sample onto a Symmetry C18 nanoAcquity trapping column (180µm x 20mm 5µm) with 100% water at 15 µl per minute for 1 minute. After trapping, peptides were resolved on a BEH130 C18 nanoAcquity column (75µmx50mm 1.7 µm) with a 30 minute, 2-40% ACN/0.1% FA linear gradient (0.5 µl/min flow rate). MRM scanning was carried out with 18 transitions and a cycle time of 1.44 seconds with a 40 millisecond dwell time per transition. An MRM Initiated Detection and Sequencing (MIDAS) was performed. The IDA method consisted of the most intense peak using rolling collision energy. The target ions were excluded after 3 occurrences for 30 seconds. The EPI scan had a scan rate of 20,000 Da/sec with a sum of 3 scans and mass range of 100-1000 Da and a cycle time of 1.4 msec. Files were searched using Mascot version 2.3 with the Swissprot database (08/2010) selected (human taxonomic restriction,). Phosphorylated S and T, and propionamide C were variable modifications. Peptide and fragment mass tolerance is 0.6 Da, with 1 missed cleavage. Quantification was performed using MultiQuant 2.0.

To demonstrate the incorporation of Sep at position 218 we used our MRM assay that was designed to detect an intact tryptic phosphopeptide ion (m/z 823.4^{+3}) derived from MBP-MEK1(Sep218,Ser222) and 4 fragment ions produced by collision-induced dissociation of this

intact phosphopeptide (Figure 4B and Table S1). The MRM method included an Information Dependent Acquisition (IDA) step that triggered a full MS/MS scan once the 823.4^{+3} ion, and associated fragment ions, were detected. The IDA MS/MS spectrum confirmed the incorporation of Sep at position 218 and Ser at 222 in MBP-MEK1 (Sep218, Ser222) (Fig. 4B). We acknowledge the efforts of C. Colangelo, T. Abbott and M. Berberich at the W.M. Keck Biotechnology Resource Laboratory at Yale.

Supporting Figures

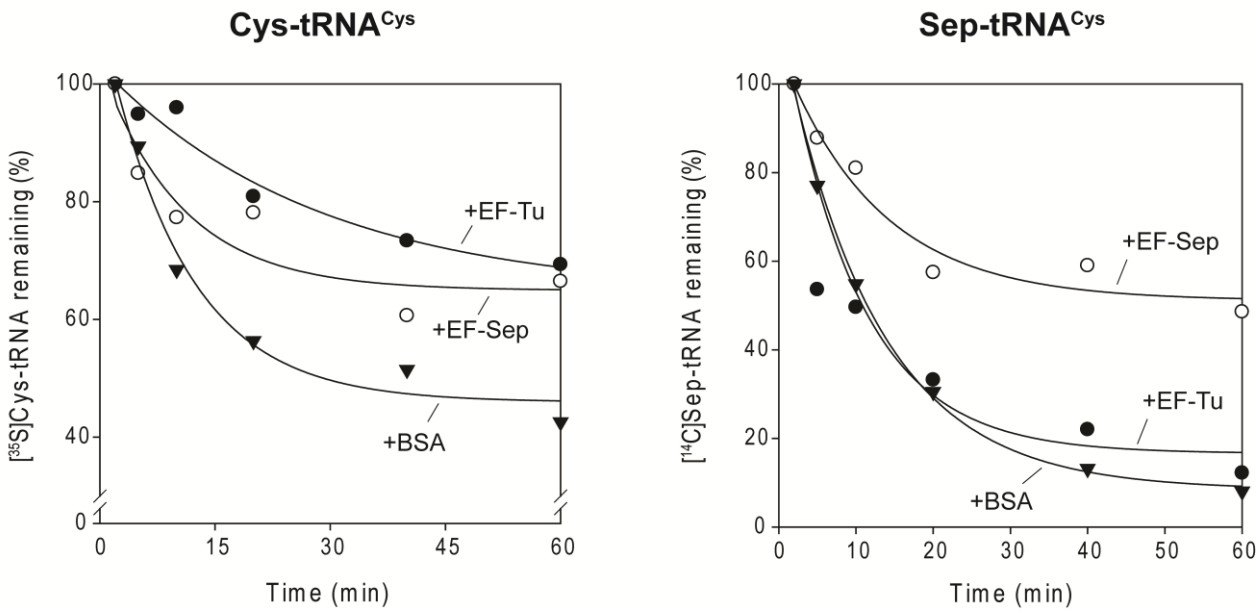

Fig. S1 EF-Tu protects Cys-tRNA^{Cys} but not Sep-tRNA^{Cys} from deacylation. Hydrolysis of *M. jannaschii* [^{35}S]Cys-tRNA^{Cys} or [^{14}C]Sep-tRNA^{Cys} was determined at pH 8.2 and room temperature in the presence and absence of EF-Tu or EF-Sep, respectively.

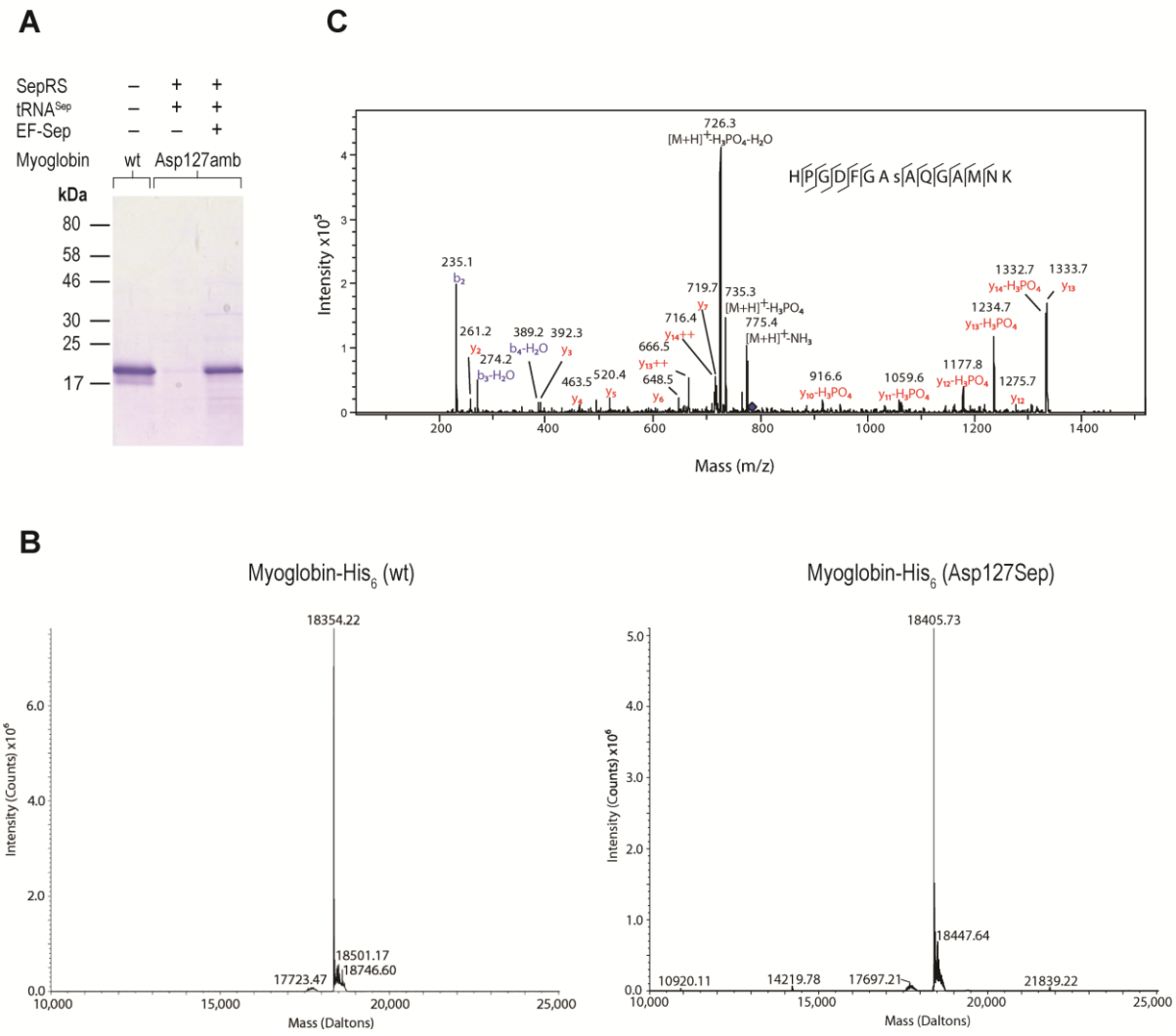

Fig. S2 Site-specific insertion of Sep in response to a TAG codon in the myoglobin gene requires tRNA^{Sep}, SepRS and EF-Sep. (**A**) Wild-type Myoglobin-His6 (lane 1) and Myoglobin-His6 (Asp127Sep) produced in the presence of Sep (2mM), SepRS and tRNA^{Sep}, with or without EF-Sep (lanes 2 and 3, respectively). Proteins were purified by affinity chromatography on Ni^{2+}-NTA and separated by SDS-PAGE.

(**B**) Electrospray ionization mass spectrometry of purified Myoglobin-His_6 (left) and purified Myoglobin-His_6 (Asp127Sep) (right) was performed as described in the methods. The found and expected masses are as follows: Myoglobin-His_6 (Found= 18354.22; expected = 18354.20 Da); Myoglobin-His_6 (Asp127Sep) (Found= 18405.73 Da; expected= 18406.17 Da). The mass difference between the two proteins 51.51 Da compares well with the expected mass difference of 51.97 Da.

(**C**) A representative spectrum from a doubly charged peptide-ion from a trypsin digest of *E. coli* produced phospho-myoglobin protein separated using an Agilent C18 reverse phase Chip (0.75

mm x150 mm) and fragmented by CID in an Agilent 6330 ion trap mass spectrometer with integrated Nano-Electrospray ionization. The observed b and y ions are indicated. Multiple spectra were identified as described that matched the peptide HPGDFGAsAQGAMNK where s was consistent with a phosphoserine residue. The precursor peptide had an m/z of 784.5 at eluted at 10.05 min. and was identified by Spectrum Mill (Agilent Technologies). Y ions are indicated in red, b ions in blue and neutral loss fragments are indicated as well as the charge state of each.

Supporting Table

Table S1: Peptide information for MRM

peptide	precursor/product ion	CE	DP
LC*DFGVSGQLIDS^{P}MANSFVGTR	823.4 $^{(3+)}$/ 333.2$^{(1+)}$[y3]	30.85	160.9
LC*DFGVSGQLIDS^{P}MANSFVGTR	823.4 $^{(3+)}$/ 666.35 $^{(1+)}$[y6]	38.26	160.9
LC*DFGVSGQLIDS^{P}MANSFVGTR	823.4 $^{(3+)}$/ 780.4 $^{(1+)}$[y7]	38.62	160.9
LC*DFGVSGQLIDS^{P}MANSFVGTR	823.4 $^{(3+)}$/ 851.4 $^{(1+)}$[y8]	38.12	160.9

S^{P}, phosphoserine; C*, propionamide; CE, Collision energy; DP, Dilution Potential

References and Notes

1. F. Lipmann, The bonding of phosphoric acid in phosphoproteins. I. Note: Isolation of some phosphoric amino acids (Serin phosphoric acid) out of casein. *Biochem. Z.* **262**, 3 (1933).

2. J. V. Olsen *et al*., Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127**, 635 (2006).

3. G. Manning, D. B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, The protein kinase complement of the human genome. *Science* **298**, 1912 (2002).

4. A. Sauerwald *et al*., RNA-dependent cysteine biosynthesis in archaea. *Science* **307**, 1969 (2005).

5. M. J. Hohn, H. S. Park, P. O'Donoghue, M. Schnitzbauer, D. Söll, Emergence of the universal genetic code imprinted in an RNA record. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 18095 (2006).

6. S. Kamtekar *et al*., Toward understanding phosphoseryl-tRNA^{Cys} formation: The crystal structure of *Methanococcus maripaludis* phosphoseryl-tRNA synthetase. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 2620 (2007).

7. R. Fukunaga, S. Yokoyama, Structural insights into the first step of RNA-dependent cysteine biosynthesis in archaea. *Nat. Struct. Mol. Biol.* **14**, 272 (2007).

8. C. C. Liu, P. G. Schultz, Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **79**, 413 (2010).

9. B. L. Wanner, W. W. Metcalf, Molecular genetic studies of a 10.9-kb operon in *Escherichia coli* for phosphonate uptake and biodegradation. *FEMS Microbiol. Lett.* **79**, 133 (1992).

10. F. J. LaRiviere, A. D. Wolfson, O. C. Uhlenbeck, Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science* **294**, 165 (2001).

11. D. M. Rothman *et al*., Caged phosphoproteins. *J. Am. Chem. Soc.* **127**, 846 (2005).

12. T. Dale, L. E. Sanderson, O. C. Uhlenbeck, The affinity of elongation factor Tu for an aminoacyl-tRNA is modulated by the esterified amino acid. *Biochemistry* **43**, 6159 (2004).

13. J. Eargle, A. A. Black, A. Sethi, L. G. Trabuco, Z. Luthey-Schulten, Dynamics of Recognition between tRNA and elongation factor Tu. *J. Mol. Biol.* **377**, 1382 (2008).

14. J. Ling *et al*., Pathogenic mechanism of a human mitochondrial tRNA^{Phe} mutation associated with myoclonic epilepsy with ragged red fibers syndrome. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15299 (2007).

15. Y. Doi, T. Ohtsuki, Y. Shimizu, T. Ueda, M. Sisido, Elongation factor Tu mutants expand amino acid tolerance of protein biosynthesis system. *J. Am. Chem. Soc.* **129**, 14458 (2007).

16. T. Ohtsuki, H. Yamamoto, Y. Doi, M. Sisido, Use of EF-Tu mutants for determining and improving aminoacylation efficiency and for purifying aminoacyl tRNAs with non-natural amino acids. *J. Biochem.* **148**, 239 (2010).

17. P. Nissen *et al*., Crystal structure of the ternary complex of Phe-tRNA^{Phe}, EF-Tu, and a GTP analog. *Science* **270**, 1464 (1995).

18. J. S. Sebolt-Leopold, R. Herrera, Targeting the mitogen-activated protein kinase cascade to treat cancer. *Nat. Rev. Cancer* **4**, 937 (2004).

19. D. R. Alessi *et al*., Identification of the sites in MAP kinase kinase-1 phosphorylated by p74raf-1. *EMBO J.* **13**, 1610 (1994).

20. J. M. Schrader, S. J. Chapman, O. C. Uhlenbeck, Tuning the affinity of aminoacyl-tRNA to elongation factor Tu for optimal decoding. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5215 (2011).

21. S. Yoshizawa, A. Böck, The many levels of control on bacterial selenoprotein synthesis. *Biochim. Biophys. Acta* **1790**, 1404 (2009).

22. K. Nozawa *et al*., Pyrrolysyl-tRNA synthetase-$tRNA^{Pyl}$ structure reveals the molecular basis of orthogonality. *Nature* **457**, 1163 (2009).

23. T. Mukai *et al*., Codon reassignment in the *Escherichia coli* genetic code. *Nucleic Acids Res.* **38**, 8188 (2010).

24. F. J. Isaacs *et al*., Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science* **333**, 348 (2011).

25. J. D. Scott, T. Pawson, Cell signaling in space and time: where proteins come together and when they're apart. *Science* **326**, 1220 (2009).

26. M. B. Yaffe *et al*., The structural basis for 14-3-3:phosphopeptide binding specificity. *Cell* **91**, 961 (1997).

27. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6640 (2000).

28. S. Bunjun *et al*., A dual-specificity aminoacyl-tRNA synthetase in the deep-rooted eukaryote *Giardia lamblia*. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12997 (2000).

29. L. Wang, A. Brock, B. Herberich, P. G. Schultz, Expanding the genetic code of *Escherichia coli*. *Science* **292**, 498 (2001).

30. P. Bernard, P. Gabarit, E. M. Bahassi, M. Couturier, Positive-selection vectors using the F plasmid *ccdB* killer gene. *Gene* **148**, 71 (1994).

31. K. Bajaj, P. Chakrabarti, R. Varadarajan, Mutagenesis-based definitions and probes of residue burial in proteins. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16221 (2005).

32. M. H. Dao-Thi *et al*., Molecular basis of gyrase poisoning by the addiction toxin CcdB. *J. Mol. Biol.* **348**, 1091 (2005).

33. J. Normanly, R. C. Ogden, S. J. Horvath, J. Abelson, Changing the identity of a transfer RNA. *Nature* **321**, 213 (1986).

34. F. W. Studier, Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207 (2005).

35. C. M. Zhang, C. Liu, S. Slater, Y. M. Hou, Aminoacylation of tRNA with phosphoserine for synthesis of cysteinyl-$tRNA^{Cys}$. *Nat. Struct. Mol. Biol.* **15**, 507 (2008).

36. T. Christian, Y. M. Hou, Distinct determinants of tRNA recognition by the TrmD and Trm5 methyl transferases. *J. Mol. Biol.* **373**, 623 (2007).

37. S. Traverse, N. Gomez, H. Paterson, C. Marshall, P. Cohen, Sustained activation of the mitogen-activated protein (MAP) kinase cascade may be required for differentiation of

PC12 cells. Comparison of the effects of nerve growth factor and epidermal growth factor. *Biochem. J.* **288**, 351 (1992).