Nat Chem Biol. Author manuscript; available in PMC 2016 April 13.

Published in final edited form as:
Nat Chem Biol. 2015 July ; 11(7): 496–503. doi:10.1038/nchembio.1823.

Efficient genetic encoding of phosphoserine and its non-hydrolyzable analog

Daniel T. Rogerson[1], **Amit Sachdeva**[1], **Kaihang Wang**[1], **Tamanna Haq**[2], **Agne Kazlauskaite**[3], **Susan M. Hancock**[1], **Nicolas Huguenin-Dezot**[1], **Miratul M. K. Muqit**[3,4], **Andrew M. Fry**[2], **Richard Bayliss**[2], and **Jason W. Chin**[1,5]

[1]Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, CB2 0QH

[2]Department of Biochemistry, Henry Wellcome Building, Lancaster Road, Leicester, LE1 9HN

[3]Medical Research Council Protein Phosphorylation and Ubiquitylation Unit, University of Dundee, Dow Street Dundee DD1 5EH, UK

[4]College of Medicine, Dentistry and Nursing, University of Dundee

Abstract

Serine phosphorylation is a key post-translational modification that regulates diverse biological processes. Powerful analytical methods have identified thousands of phosphorylation sites, but many of their functions remain to be deciphered. A key to understanding the function of protein phosphorylation is access to phosphorylated proteins, but this is often challenging or impossible. Here we evolve an orthogonal aminoacyl-tRNA synthetase/$tRNA_{CUA}$ pair that directs the efficient incorporation of phosphoserine into recombinant proteins in *E. coli*. Moreover, combining the orthogonal pair with a metabolically engineered *E. coli* enables the site-specific incorporation of a non-hydrolyzable analog of phosphoserine. Our approach enables quantitative decoding of the amber stop codon as phosphoserine and we purify several milligrams-per-liter of proteins bearing biologically relevant phosphorylations that were previously challenging or impossible to access: including phosphorylated ubiquitin and a kinase (Nek7) that is synthetically activated by a genetically encoded phosphorylation in its activation loop.

The phosphorylation of proteins in eukaryotic cells is a key post-translational modification that regulates essential and diverse biological phenomena. Phosphorylation is installed by protein kinases on serine, threonine, tyrosine and histidine residues, and removed by phosphatases, and serine phosphorylation is the most abundant phosphorylation observed in

[5]Correspondence: chin@mrc-lmb.cam.ac.uk.

Author contributions

D.T.R. and J.W.C. conceived the experimental strategy, analysed the data and wrote the paper with input from other authors. D.T.R. performed all the selections, system characterization and most phospho-protein expressions and purifications. S.M.H. and D.T.R. characterized the starting system. K.W., A.S., D.T.R. and N.H-D. developed and characterized the expression system. T.H. and D.T.R., performed the Nek7 experiments with guidance from A.M.F. and R.B. A.K. performed the Ub assays with guidance from M.M.K.M.

Competing financial interests

The authors declare no competing financial interest.

Europe PMC Funders Author Manuscripts

proteins[1-3]. Accessing the homogeneously, site-specifically and quantitatively phosphorylated form of many proteins remains challenging. This is often because the relevant kinases are unknown - as is increasingly the case for the large numbers of phosphorylation sites of unknown function identified by mass spectrometry[2] - or because the relevant kinase does not modify the desired protein site-specifically and/or completely *in vitro*. The installation of protein phosphorylation by semi-synthetic methods, using native chemical ligation, has provided a powerful approach to address the functional consequences of phosphorylation[4, 5], but is most commonly limited to sites proximal to the termini of proteins. The efficient biosynthesis of proteins bearing site-specific phosphorylation would enable the biochemical, mechanistic and structural characterization of diverse phosphorylated proteins, and provide molecular insight into how phosphorylation dynamically reshapes proteome function[2]; while the efficient biosynthesis of proteins bearing non-hydrolyzable analogs of phosphorylation at known phosphorylation sites would enable a range of additional applications, including the capture of phospho-protein binders in complex mixtures containing phosphatases, and the creation of constitutively active versions of kinases and other proteins. Here we address the challenge of creating a general and efficient method for site-specifically installing phosphoserine, and a non-hydrolyzable analog of phosphoserine into recombinant proteins via genetic code expansion.

Intriguingly, phosphoserine is a metabolic precursor to serine, and in certain methanogenic archea phosphoserine is used in a two-step pathway to direct the incorporation of cysteine into proteins[6]. In the first step of this pathway an unusual aminoacyl-tRNA synthetase (SepRS) aminoacylates phosphoserine on to $tRNA^{Cys}_{GCA}$, to create pSer $tRNA^{Cys}_{GCA}$. A second enzyme (Sep tRNA:Cys tRNA synthase) then converts pSer $tRNA^{Cys}_{GCA}$ to Cys $tRNA^{Cys}_{GCA}$, and this aminoacylated tRNA is taken to the ribosome by archeal elongation factor EF1α (a homolog of EF-Tu), where it directs the incorporation of cysteine in response to cognate codons.

Two orthogonal aminoacyl-tRNA synthetase/$tRNA_{CUA}$ pairs have previously been developed for the efficient, site-specific incorporation of diverse unnatural amino acids into recombinant proteins in response to an amber stop codon introduced into a gene of interest (the PylRS/tRNA pair from *methanosarcina* and the *Methanococcus janaschii* TyrRS/tRNA pair)[7, 8] ; the aminoacylation of $tRNA_{GCA}$ with phosphoserine in the first step of the cysteine incorporation pathway in methanogenic archea raised the possibility of altering the SepRS/ $tRNA_{GCA}$ pair to create a SepRS/$tRNA(Tran)_{CUA}$ pair (in which the CUA anticodon is transplanted in place of the native GCA anticodon) for the site-specific incorporation of phosphoserine into proteins[9].

Despite substantial effort,[10, 11] two challenges have limited the adaptation of the SepRS/ $tRNA_{GCA}$ pair for the site-specific incorporation of phosphoserine into proteins. The first challenge in incorporating phosphoserine results from the prediction that $tRNA^{Cys}_{GCA}$, aminoacylated with phosphoserine, will be a poor substrate for EF-Tu [12]. This prediction was borne out by the observation that mutating the amino acid binding pocket of EF-Tu to create "EF-Sep" enabled the measureable incorporation of phosphoserine into model proteins in *E. coli* using a SepRS/$tRNA(Tran)_{CUA}$/ EF-Sep system[10]. However, this system is substantially less efficient than other systems for incorporating unnatural amino acids into

proteins, and produced only micrograms per L of a protein kinase that was not purified to homogeneity[10], and has not found utility for the synthesis of phosphoproteins. Reducing competition for decoding the amber stop codon by deleting release factor-1 (RF1) was reported to increases protein yields for a model protein[13], but the cells grew very slowly and more than half of the protein that is synthesized mis-incorporated natural amino acids in place of phosphoserine in response to the amber stop codon[14].

A second challenge in adapting the SepRS/$tRNA_{GCA}$ pair results from the consequences of converting the anticodon of the tRNA from GCA to CUA. *In vitro* aminoacylation assays[9] and a SepRS/$tRNA_{GCA}$ structure[9] demonstrated that SepRS recognizes the anticodon (G34, C35, and A36) of $tRNA_{GCA}$ and mutation of the GCA anticodon to CUA to create an amber suppressor led to a tRNA that is a very poor substrate for SepRS[9], and may be poorly decoded on the ribosome. Further work reported an additional mutation in EF-Sep, creating EF-Sep21, and mutations in the anticodon binding region of SepRS, creating SepRS9, and reported a five-fold increase in chloramphenicol resistance from a chloramphenicol acetyl-transferase gene bearing an amber stop codon at position 112 using this system[11]. These authors also reported the synthesis of histone H3 phosphorylated at serine 10 with a yield of 3 mg/L.

We, and others, have demonstrated that the efficiency of unnatural amino acid incorporation with orthogonal synthetase/tRNA pairs, in which the anticodon of the tRNA has been altered to be complementary to a stop codon or a quadruplet codon, can be enhanced by optimizing the sequence surrounding the new anticodon[15, 16]. Selecting optimized sequences surrounding the anticodon of the tRNA may facilitate the presentation and decoding of the tRNA in the ribosome and, in cases where the natural anticodon is an identity element for its synthetase (as for SepRS), it may also facilitate the recognition of the optimized tRNA by the synthetase.

Here we demonstrate that the systematic evolution of the sequence surrounding the anticodon of $tRNA(Tran)_{CUA}$ and subsequent evolution of anticodon recognition region of SepRS enables the creation of evolved SepRS(XX)/$tRNA(XX)_{CUA}$ pairs, where XX defines the selected mutations. These pairs direct the incorporation of phosphoserine eighteen times more efficiently than previously reported pairs[10, 11], with a comparable efficiency to the established orthogonal PylRS/$tRNA_{CUA}$ pair, that is widely used to make modified recombinant proteins for structural and functional studies [7]. Moreover, we demonstrate that, in contrast to previous assumptions[10], EF-Tu engineering is not required for genetically encoded phosphoserine incorporation. We demonstrate the synthesis (with yields of several milligrams per liter of culture), and activity of phosphorylated recombinant proteins, including ubiquitin (Ub) phosphorylated at position 65 (Ub(pSer65))[17-19] and NEK7 kinase[22, 23] phosphorylated in the activation loop. Finally, we manipulate serine biosynthetic enzymes[20] to enable the first route to genetically encoding the site-specific incorporation of a non-hydrolyzable analog of phosphoserine in a recombinant protein.

Results

Evolving pSer tRNA enhances phosphoserine incorporation

We began by testing the efficiency of previously reported systems for phosphoserine incorporation using a chloramphenicol resistance assay. Read through of the amber stop codon in a chloramphenicol acetyl-transferase gene at position 112 (*cat(112TAG))* by EF-Sep/SepRS/ pSer tRNA(Tran)$_{CUA}$, leads to growth on 60 μg/mL chloramphenicol, while in the absence of SepRS cells grow weakly on 20 μg/mL chloramphenicol (Figure 1). Mutating SepRS to SepRS9[11] and EF-Sep[10] to EF-Sep21[11] did not provide an increase in chloramphenicol resistance in our hands (Figure 1 and Supplementary Results, Supplementary Figure 1). We therefore used SepRS and EF-Sep as a starting point for further experiments.

To investigate whether altering the sequence flanking the anticodon of pSer tRNA$_{CUA}$ enables efficient phosphoserine incorporation we first created pSer tRNA(N10)$_{CUA}$, a 5.1×10^7 member library that samples all possible combinations of the ten nucleotides G29-U33 and G37-C41 that flank the anticodon. The library was transformed into cells bearing SepRS and EF-Sep and *cat*(112TAG), and pSer tRNA(N10)$_{CUA}$ library members that enable efficient read-through of the amber stop codon selected on chloramphenicol. This screen identified 96 colonies, and sequencing revealed that these colonies contained 15 distinct tRNA(XX)$_{CUA}$ sequences, indicating a large degree of library convergence (Figure 1).

To quantify the efficiency and specificity of phosphoserine incorporation with the selected pSer tRNA(XX)$_{CUA}$ variants we examined the chloramphenicol resistance of cells containing *cat(112TAG)* with either EF-Sep/SepRS/ pSer tRNA(XX)$_{CUA}$ variants or EF-Sep/pSer tRNA(XX)$_{CUA}$ variants (in which SepRS is omitted). Thirteen out of the fifteen selected EF-Sep/SepRS/pSer tRNA(XX)$_{CUA}$ combinations led to growth on substantially higher concentrations of chloramphenicol (200 μg/mL to 700 μg/mL) than the starting system while combinations using two selected tRNAs (pSer tRNA(A8)$_{CUA}$ and pSer tRNA(C9)$_{CUA}$) were not active. The most active combinations led to approximately ten times the level of chloramphenicol resistance provided by pSer tRNA(Tran)$_{CUA}$(Figure 1b). All 13 functional tRNAs have G37 mutated to A, and conserve C32, A38 and U33. G37A in isolation is known to decrease the efficiency of tRNA$_{GCA}$ aminoacylation by SepRS[21], but the net result of the mutations transplanted into the anticodon and co-selected with G37A from pSer tRNA(N10)$_{CUA}$ on aminoacylation is unknown. We note that A37 is selected[16] in many efficient amber suppressor tRNA and may facilitate the efficient decoding of amber stop codons. Interestingly, tRNAs C2 and C4 are predicted to have an expanded anticodon loop due to mis-pairings between positions 31 and 39. However these tRNAs, like evolved quadruplet decoding PyltRNAs that also lack base pairing at this position[15], are functional.

In the absence of SepRS we observe a range of chloramphenicol resistance levels resulting from mis-aminoacylation of the selected pSer tRNA(XX)$_{CUA}$ variants by natural synthetases that are constitutively present in the cell. In the absence of SepRS, pSer tRNA(Tran)$_{CUA}$ can be aminoacylated with glutamine by GlnRS[10], and the tRNA mutations selected from pSer tRNA(N10)$_{CUA}$ may alter the levels of mis-aminoacylation by GlnRS or other endogenous aminoacyl-tRNAsynthetases. All of the selected active tRNAs show SepRS dependent

chloramphenicol resistance, and two tRNAs (pSer tRNA(B2)$_{CUA}$, pSer tRNA(C2)$_{CUA}$) are approximately ten times more active than pSer tRNA(Tran)$_{CUA}$ in the presence of EF-Sep/SepRS, but have a low activity in the absence of SepRS, that is comparable to that of pSer tRNA(Tran)$_{CUA}$ in the absence of SepRS. This data suggests that these tRNA mutants enable much more efficient incorporation of phosphoserine with EF-Sep/SepRS, but do not increase mis-incorporation of natural amino acids in the absence of SepRS.

Next we aimed i) to directly demonstrate the improvement in phosphoserine incorporation efficiency provided by the evolved pSer tRNA(XX)$_{CUA}$ variants in a recombinant protein and ii) to characterise the fidelity of phosphoserine incorporation by mass spectrometry. We expressed the gene for sperm whale myoglobin bearing a C-terminal His6 tag and an amber codon at position 127 (*myo*(127TAG)-*His6*) in cells that also contain EF-Sep/SepRS and either pSer tRNA(Tran)$_{CUA}$ or one of the fifteen selected pSer tRNA(XX)$_{CUA}$ variants. Expression using each of the thirteen active, selected pSer tRNA(XX)$_{CUA}$ variants led to substantially more myoglobin expression than expression using pSer tRNA(Tran)$_{CUA}$ (Figure 1c and Supplementary Figure 2), consistent with the chloramphenicol resistance assay. Purified myoglobin from each expression was subject to mass spectrometry to demonstrate the incorporation of phosphoserine (Supplementary Figure 3). Despite the variable levels of chloramphenicol resistance observed in the absence of SepRS (Figure 1b), in the presence of SepRS a single protein is produced, for all the pSer tRNA(XX)$_{CUA}$ variants, in which phosphoserine is the only amino acid that is detectably incorporated in response to the amber codon. This is consistent with previous reports that background misacylation of amber suppressor tRNAs may be effectively out-competed in the presence of an efficient cognate synthetase and its substrate[22, 23].

SepRS evolution enhances phosphoserine incorporation

We next created a SepRS anticodon binding domain library (SepRS(AC6)) and selected SepRS variants that function more efficiently with the thirteen selected, active pSer tRNA(XX)$_{CUA}$ variants. From the X-ray crystal structure of the *Archaeoglobus fulgidus* SepRS/tRNA$^{Cys}_{GCA}$ complex[9], we identified six amino acid residues that interact with G34 and C35 of the cognate tRNA's anticodon (Glu412, Glu414, Thr417, Pro495, Ile496 and Phe526) to target for combinatorial mutagenesis to all twenty amino acids (Figure 2a). The resulting library contains two critical residues (Phe526, Thr417) that pack around and between G34 and C35 of the anticodon (Figure 2a), in addition to the anticodon binding domain residues mutated in previous work[11]. To generate the SepRS(AC6) library we first created plasmids recoding Phe526 to every amino acid and then used these plasmids as a template for two rounds of Enzymatic Inverse-PCR mediated randomisation with a "small-intelligent" primer set[24] to encode all 20 amino acids at each of the other five targeted positions. The library was generated with a diversity of 3.1×10^9 c.f.u., surpassing the required diversity of 2.9×10^8.

The SepRS(AC6) library was co-transformed into cells with each of the thirteen selected, active pSer tRNA(XX)$_{CUA}$ variants, EF-Sep, and *cat*(112TAG) in thirteen parallel transformations, and subjected to a positive selection on chloramphenicol. Fourteen new SepRS(XX)/pSer tRNA(XX)$_{CUA}$ pairs, in which the SepRS(XX) variants discovered were

convergent at several positions (Supplementary Table 1), were identified from this selection, and their activity determined by chloramphenicol resistance assays (Figure 2b).

The eight most active, selected SepRS(XX)/ pSer tRNA(XX)$_{CUA}$ pairs survived on 1100 μg/mL of chloramphenicol, indicating that these pairs are greater than eighteen times more efficient than previously reported systems for phosphoserine incorporation[10, 11] (Figure 1 & Supplementary Figure 1). Seven of the eight most active pairs utilize pSer tRNA(XX)$_{CUA}$ variants that exhibit the greatest activity with SepRS (Figure 2b).

Protein expression experiments with the selected SepRS (XX)/ pSer tRNA(XX)$_{CUA}$ pairs, EF-Sep, and *myo*(127TAG*)-His6* confirmed the increase in efficiency provided by the selected pairs, and demonstrate that the pairs we describe are much more efficient than the SepRS/ pSer tRNA(Tran)$_{CUA}$ pair[10], or mutants of SepRS previously reported to function with pSer tRNA(Tran)$_{CUA}$ [11](Figure 2c and Supplementary Figure 2). Moreover, mass spectrometry confirms the quantitative incorporation of phosphoserine in recombinant myoglobin produced from cells containing these systems (Supplementary Figure 4).

Ef-Tu mutation is not required to encode phosphoserine

Next we investigated whether EF-Sep is required for the site-specific incorporation of phosphoserine using the evolved pSer tRNA(XX)$_{CUA}$ variants. Control experiments confirm that the SepRS/pSer tRNA(Tran)$_{CUA}$ pair does not support chloramphenicol resistance in cells bearing *cat*(112TAG) when EF-Sep is omitted. In contrast, cells bearing *cat*(112TAG), and the SepRS, pSer tRNA(XX)$_{CUA}$ pairs, but lacking EF-Sep, survive on concentrations of chloramphenicol up to 100μg/mL (eg: pSer tRNA(A4)$_{CUA}$) (Figure 3a), demonstrating that the selected SepRS/pSer tRNA(XX)$_{CUA}$ pairs are sufficient to direct the incorporation of phosphoserine in the absence of EF-Sep. The level of chloramphenicol resistance provided by selected SepRS/pSer tRNA(XX)$_{CUA}$ pairs exceeds that of the EF-Sep/SepRS/ pSer tRNA(Tran)$_{CUA}$ system (60μg/mL), providing evidence that the selected tRNA(XX)$_{CUA}$s provide a contribution to the efficiency of phosphoserine incorporation that exceeds the contribution provided by EF-Tu engineering to create EF-Sep. Protein expression experiments with the SepRS/ pSer tRNA(XX)$_{CUA}$ pairs and *myo*(127TAG)*His6* confirmed the increase in efficiency provided by pSer tRNA(XX)$_{CUA}$ in the absence of EF-Sep (Figure 3b and Supplementary Figure 2). Moreover, mass spectrometry confirms the quantitative incorporation of phosphoserine in recombinant myoglobin produced from cells containing the SepRS/ pSer tRNA(XX)$_{CUA}$ pairs, but lacking EF-Sep (Supplementary Figure 5).

Next, we asked whether evolving the anticodon binding region of SepRS enables efficient incorporation of phosphoserine in the absence of EF-Sep. To discover SepRS (XX)/ pSer tRNA(XX)$_{CUA}$ pairs that function in the absence of EF-Sep we co-transformed the SepRS(AC6) library with each of the thirteen active pSer tRNA(XX)$_{CUA}$s (Figure 1) and *cat*(112TAG) and selected cells that survive on increasing concentrations of chloramphenicol. We discovered three SepRS (XX)/ pSer tRNA(XX)$_{CUA}$ pairs that direct the incorporation of phosphoserine more efficiently that the progenitor SepRS / pSer tRNA(B2)$_{CUA}$, and SepRS / pSer tRNA(B4)$_{CUA}$ pairs (Figure 3c, Supplementary Table 1). The most efficient of these pairs, SepRS (17)/ pSer tRNA(B2)$_{CUA}$, confers chloramphenicol resistance to approximately 500μg/mL, approximately ten times the resistance conferred by

the SepRS/ pSer tRNA(Tran)$_{CUA}$/EF-Sep system or the SepRS9/ pSer tRNA(Tran)$_{CUA}$/EF-Sep21 system[11] (Figure 3c, Figure 1**,** and Supplementary Figure 1).

Protein expression experiments with the selected SepRS (XX)/ pSer tRNA(XX)$_{CUA}$ pairs, and *myo*(127TAG)*His6* confirmed the substantial increase in efficiency provided by the selected SepRS (XX)/ pSer tRNA(XX)$_{CUA}$ pairs, and further demonstrated that the selected pairs are much more efficient than the SepRS/ pSer tRNA(Tran)$_{CUA}$/EF-Sep, system previously reported[10] (Figure 3d and Supplementary Figure 2). Moreover, mass spectrometry confirmed the quantitative incorporation of phosphoserine in recombinant myoglobin produced from cells containing these systems (Supplementary Figure 5). Taken together the chloramphenicol resistance data, protein expression data and mass spectrometry data demonstrate that pSer tRNA$_{CUA}$ evolution and SepRS evolution lead to much more efficient phosphoserine incorporation than EF-Tu evolution, and demonstrate that, in contrast to previous assumptions[10], alteration of EF-Tu is not required for the incorporation of phosphoserine. Nonetheless, the best EF-Sep/SepRS(XX)/pSer tRNA(XX)$_{CUA}$ combinations we have discovered outperform the best EF-Sep independent SepRS(XX)/pSer tRNA(XX)$_{CUA}$ pairs (Supplementary Figure 1), consistent with the sequential roles of aminoacylation, delivery of the aminoacylated tRNA to the ribosome, and ribosomal decoding in the process of translation.

Optimized site-specific phosphoserine incorporation

Having established the optimal SepRS and pSer tRNA sequences we cloned the SepRS(2)/ pSer tRNA(B4)$_{CUA}$/EF-Sep combination into a single high copy plasmid for convenient use in protein expression experiments when co-transformed with a plasmid encoding the gene of interest. We used a Glutathione-S-transferase-TAG-calmodulin (*gst*(TAG)*cam*) reporter with an amber stop codon at position 1 of calmodulin to characterize the efficiency of phosphoserine incorporation (Supplementary Figure 6) via the ratio of GST (synthesized as a result of termination in response to the amber stop codon) to GST-calmodulin (synthesized as a result of unnatural amino acid incorporation in response to the amber stop codon)[15]. The optimized system led to the synthesis of full-length GST-calmodulin (Supplementary Figure 6), and only a trace amount of GST. To benchmark the efficiency of the phosphoserine incorporation system against the efficiency of a widely used genetic code expansion platform we expressed (*gst*(TAG)*cam*) using the PylRS/tRNA$_{CUA}$ pair in the presence of (Nε-[(tert-butoxy)carbonyl]-L-lysine), an established and efficient substrate for PylRS[25]. This led to a greater fraction of protein synthesis resulting in truncation, demonstrating that the new system we have developed here has an efficiency in this assay that is comparable to or greater than that of a system that has been widely adopted for genetic code expansion.

While the efficiency of the system for phosphoserine incorporation was already very good, we tested whether the efficiency could be improved with an orthogonal translation system that does not recognize RF1, or using an RF1 deletion[15, 23, 26-28] (Supplementary Figure 6). In both cases we see little or no truncated GST indicating near quantitative read through of the amber stop codon. While the use of these systems is unlikely to be necessary for

expressing many proteins, these experiments demonstrate their compatibility with the evolved system for phosphoserine incorporation.

Expression and characterization of Ub(pSer65)

To further demonstrate the generality of our approach we produced ubiquitin bearing phosphoserine at position 65, a phosphorylation that has been identified in cells [17-19] and is installed by PINK1, a kinase that also phosphorylates the E3 ubiquitin ligase Parkin[29]. Mutations in PINK1 or Parkin are the most common cause of autosomal recessive Parkinson's disease[30]. Upon stress, the phosphorylation of both Parkin and ubiquitin at Ser 65 by PINK1 appear to be required for the activation of Parkin and the ubiquitylation and clearance of damaged mitochondria from cells [17-19]. Pure ubiquitin phosphorylated at serine 65 is challenging to generate enzymatically *in vitro*, and is typically obtained in sub-milligram quantities, though very recently a more efficient route to enzymatic phosphorylation of serine 65 has been reported and used for detailed biochemical characterization of this phosphorylation[31].

We expressed ubiquitin specifically phosphorylated at position 65 (Ub (pSer65)) in cells containing Ub (65TAG)-His6 and the optimized phosphoserine incorporation system. Ub (pSer65) was purified with a yield of 3 mg per L of culture (Figure 4a,b,c). We investigated whether the Ub(pSer65) we have generated can directly activate Parkin E3 ligase activity to accelerate discharge of ubiquitin from a loaded E2 and ubiquitylation of substrates, as previously demonstrated for enzymatically derived ubiquitin that is phosphorylated at serine 65 [17-19], which we refer to as Ub(pSer65*) for clarity. Addition of Ub(pSer65) led to near maximal ubiquitin discharge from a ubiquitin loaded E2 (UbcH7) by Parkin (Figure 4d and Supplementary Figure 7) and had a similar effect to Ub(pSer65*) (Fig 4d). Consistent with previous results we did not observe any discharge upon addition of un-phosphorylated ubiquitin (Fig 4d). Parkin activation by Ub(pSer65) led to the formation of polyubiquitin chains, increased auto-ubiquitylation of Parkin and ubiquitylation of a Parkin substrate, Miro 1 (Figure 4e and Supplementary Figure 2). The activity conferred by Ub(pSer65) was comparable to that observed with Ub(pSer65*) (Figure 4e). In control experiments un-phosphorylated ubiquitin did not activate full-length Parkin (Figure 4e), as expected.

Expression of synthetically-activated Nek7 kinase

To demonstrate that our system enables the efficient incorporation of phosphoserine at biologically relevant sites in a protein kinase we expressed Nek7, an essential kinase that is activated by phosphorylation at Serine 195 in mammalian cells[32, 33]. Nek7 plays an important role in mitotic progression, and is required for robust mitotic spindle formation and cytokinesis[32, 33].

We expressed and purified Nek7 (Ser195TAG)-His6, leading to the production of Nek7 (pSer195)-His6 with a purified yield of 2 mg per L of culture (Figure 5a). ESI-MS and ESI MS/MS demonstrates the synthesis of Nek7 phosphorylated at position 195 (Figure 5b,c and Supplementary Figure 8). A small fraction of the purified protein has a mass consistent with the autophosphorylation of the synthetically activated kinase at an additional site (or sites) (Figure 5b**).** Indeed, we observe autophosphorylation of Nek7 (Supplementary Figure 9), as

previously observed for the closely related kinase Nek6[32]. To further confirm that the minor species results from autophosphorylation we purified a catalytically inactive variant, Nek7 (D179N) (Supplementary Figure 10). Expression of Nek7 (D179N, Ser195TAG)-His6 led to production of Nek7 (D179N, pSer195)-His6 with a purified yield of 2 mg per L of culture. ESI-MS and MS/MS confirmed the site-specific incorporation of phosphoserine with no other phosphorylations.

We compared the activity of wild-type (wt) Nek7, Nek7(pSer195) and Nek7(D179N, pSer195) in phosphorylation of a model substrate (myelin basic protein, MBP)[32] (Figure 5d and Supplementary Figure 11). While Nek7 (pSer195) robustly phosphorylates the substrate, neither wtNek7 nor Nek7(D179N, pSer195) efficiently phosphorylates the substrate. The addition of the C-terminal domain of Nek9 (Nek9-CTD) activates Nek7, as expected[34], but does not further activate Nek7 (pSer195), indicating that Nek7 (pSer195) bypasses the requirement for activation by Nek9 *in vitro*.

These experiments provide a first route to expressing and purifying Nek7 with a quantitative and site-specific phosphorylation in the activation loop, and demonstrate that milligram quantities of phosphorylated kinases can be obtained to enable structural and functional studies. Moreover, they demonstrate that synthetic phosphorylation of Nek7 is sufficient to activate the kinase and bypass trans-activation by the Nek9-CTD *in vitro*.

Encoding a non-hydrolyzable analog of phosphoserine

The genetically directed incorporation of the non-hydrolyzable analog of phosphoserine, 2-amino-4-phosphonobutyric acid (**2**)[35, 36], into recombinant proteins may facilitate the isolation of phospho-protein binders in complex mixtures that contain phosphatases, the creation of constitutively active kinases, as well as the characterization of complexes formed between phosphatase and their substrates.

No unnatural substrates have been incorporated with the SepRS/tRNA system, and a key challenge associated with the incorporation of unnatural substrates with this system results from the abundance of phosphoserine in the cell that will outcompete unnatural substrates at the active site of SepRS (Figure 6a). We asked whether we could manipulate the cell's phosphoserine biosynthesis and destruction pathways[20] to facilitate the incorporation of **2** using the evolved SepRS(XX)/tRNA(XX)$_{CUA}$ pairs (Figure 6b). Since phosphoserine is synthesized from 3-phosphohydroxypyruvate by phosphoserine aminotransferase (encoded by *serC*), and converted to serine by phosphoserine phosphatase (encoded by *serB*) [20]we reasoned that the intracellular levels of phosphoserine may be decreased by either overexpressing *serB*, deleting *serC* from the genome or by overexpressing *serB* in a *serC* deletion strain.

We found that the addition of 2mM **2** to wild-type *E. coli* provided with the SepRS(2)/ tRNA(B4)$_{CUA}$ /EF-Sep system did not increase the levels of myoglobin synthesized from *myo (*127TAG*)-His6 (*Figure 6c*),* and mass spectrometry confirmed the incorporation of phosphoserine rather than **2** (data not shown). Overexpression of *serB* led to an increase in myoglobin level, but the level of expression was not dependent upon the addition of **2**. These observations are consistent with *serB* overexpression being insufficient to deplete

phosphoserine to enable the specific incorporation of **2**. Deletion of *serC* led to a decrease in myoglobin levels, consistent with a depletion of intra-cellular phosphoserine, but myoglobin levels were not rescued by addition of **2**, suggesting that a *serC* deletion is not sufficient to enable the incorporation of **2**. Overexpression of *serB* in the *serC* deletion led to an increase in myoglobin levels, as observed for wild-type *E.coli*. However, the addition of **2** to the *serC* deletion overexpressing *serB* led to a clear increase in myoglobin levels, consistent with the specific incorporation of **2** in this strain.

We note that it is currently unclear why *serB* overexpression leads to an increase in myoglobin levels, or why the levels of myoglobin increase upon addition of **2** to Δ*serC* +*serB,* but not upon addition of **2** to Δ*serC*. Nonetheless, these experiments prompted us to investigate whether **2** can be site-specifically incorporated into proteins produced in *E.coli* deleted for *serC* and overexpressing *serB.* In Δ*serC* + *serB* cells containing *myo(*127TAG*)-His6*/EF-Sep and the SepRS(2)/tRNA(B4)$_{CUA}$ pair we observe expression of full-length myoglobin that is dependent on the addition of the phosphoserine analog **2** (Figure 6d). As expected, there is still some full-length protein produced in the absence of the added unnatural amino acid **2**. The yield of full-length myoglobin incorporating **2** from Δ*serC* + *serB* cells was comparable to the yield for myoglobin incorporating phosphoserine in a *serB* deletion strain (Figure 6c), suggesting that when phosphoserine levels are reduced, the non-hydrolyzable amino acid (**2)** is an efficient substrate for the SepRS(2)/tRNA(B4)$_{CUA}$ pair.

Mass spectrometry and MS/MS of myoglobin expressed in the presence of **2** demonstrates the site-specific incorporation of the analog (Figure 6e,f). This demonstrates that the analog effectively competes with any amino acids that may be incorporated in response to the amber codon in its absence, a common observation in many unnatural amino acid incorporation experiments[23].

Discussion

In summary, by evolving the nucleic acid sequence surrounding the anticodon of tRNA(Tran)$_{CUA}$ and the complementary amino acid sequence in the SepRS anticodon binding region we have discovered new orthogonal aminoacyl-tRNA synthetase/tRNA$_{CUA}$ pairs, including pairs that do not require mutations in EF-Tu, that direct the efficient incorporation of phosphoserine into recombinant proteins expressed in *E. coli*. These pairs enable the genetically encoded synthesis of several phospho-proteins with yields of several milligrams per liter, as exemplified by our syntheses of phosphorylated ubiquitin and a synthetically activated kinase, Nek7. Finally, defining a re-wired serine biosynthesis machinery enables the first incorporation of an unnatural substrate, a non-hydrolyzable analog of phosphoserine, into a recombinant protein with this synthetase/tRNA pair.

The approach we have developed provides a step-change in the ability of scientists to study the effects of serine phosphorylation, and we are currently investigating the structural and functional consequences of previously in-accessible phosphorylations. We are also exploring the evolution of the orthogonal pairs we have discovered for the incorporation of diverse unnatural amino acids. Moreover, we are investigating the mutual orthogonality[16, 37] of the

new pairs with respect to established orthogonal synthetase/tRNA pairs[7], as a route to further expand the number of unnatural monomers that may be genetically encoded in cells.

Online methods

Plasmids generation

The *Methanocaldococcus jannaschii* (*Mj*) $tRNA_{CUA}$, containing a C20U mutation shown to improve acylation by SepRS[38], was cloned into (p15A vector) between NotI and PstI to give $pSertRNA_{CUA}$. A DNA fragment encoding the *Methanococcus maripaludis SepRS* was cloned into NdeI/StuI digested pBK-Amp by Gibson cloning to give pBK-SepRS. For testing synthetase dependent incorporation, SepRS was removed from the pBK-SepRS plasmid using NEB Q5 SDM kit (New England Biolabs) and primers DTR634 and DTR635 to give pBK-NoRS. The *E.coli* EF-Tu *gene (tufA)* was amplified from DH10B genomic DNA and cloned into pCDFDuet-1 (Novagen) under control of the tac promoter. The phosphoserine permissive EF-Tu mutations (*H67R, E216N, D217G, F219Y, T229S and N274W)* were introduced by Overlap Extension PCR to give pCDF-EF-Sep. The gene for Chloramphenicol Acetyl-Transferase containing an amber stop codon at position 112 was amplified cloned into pCDF-EF-Sep to give pCDF-EF-Sep-CAT. pCDF-CAT was generated by deletion of EF-Sep gene using the NEB Q5 SDM kit (New England Biolabs) and primers DTR623 and DTR624. The gene for myoglobin $(D127TAG)_{His6}$ under the control of the arabinose promoter was amplified by PCR using DTR674 and DTR675. This fragment was inserted by Gibson cloning into pCDF-EF-Sep which had been amplified with DTR676 and DTR678 to give pCDF-EF-Sep-Myo$(D127TAG)_{His6}$. pCDF-Myo$(D127TAG)_{His6}$ was generated by deletion of EF-Sep from pCDF-EF-Sep-Myo$(D127TAG)_{His6}$ using the NEB Q5 SDM kit and primers DTR623 and DTR 624. *S*epRS mutant 2 and 17, along with the WT, were amplified from pBK-SepRS(2), pBK-SepRS(17) and pBK-SepRS using primers DTR733/734. pKW (Cm^R) was digested using PstI and NdeI, and the amplified SepRS fragment were inserted by Gibson cloning to yield pKWRS2, pKWRS17 and pKWRSWT. Gene blocks encoding tRNA(B4), tRNA(B2) and tRNA(WT) were ordered and Gibson cloned into BglII/SepI digested pKWRS2, pKWRS17 and pKWRSWT respectively, to yield pKW2, pKW17 and pKWWT. EF-Sep was amplified from pCDF-EF-Sep including its promoter and terminator using DTR823/824 and this fragment was Gibson cloned with the backbone of pKW2, which was amplified with DTR837 andDTR838 to yield pKW2-EF-Sep. The pRSF_ribo-Q1_O-$gst\text{-}cam_{1TAG}$ and pRSF_ribo-wt_$gst\text{-}cam_{1TAG}$ plasmids contain the following components, (i) a gene encoding ribosomal RNA (rRNA), former containing ribo-Q1 and latter with wildtype rRNA, (ii) a gene encoding *glutathione-S-transferase* and *calmodulin* (*gst-cam*) which is downstream of either an orthogonal ribosomal binding site (O-*gst-cam*) or a wildtype ribosomal binding site, and (iii) both the plasmids contain kanamycin resistance marker [15]. pCDF-EF-Sep-*serB* was generated by PCR amplifying the *serB* gene from the genome of MegaX DH10B (Invitrogen) using DTR793 and DTR794, and Gibson cloning the fragment into pCDF-EF-Sep-CAT in place of CAT gene by amplification of the backbone with DTR795 and DTR796. All primer sequences can be found in Supplementary Table 2**.**

Europe PMC Funders Author Manuscripts

Strain generation—The *serB* gene of MegaX DH10B cells (Invitrogen) was disrupted using the genebridges E. coli quick KO kit and confirmed by sequencing to generate DH10B Δ*serB*. C321(DE3) was generated using the initial strain C321.ΔA.exp (Addgene strain 49018) and a λDE3 Lysogenization Kit (EMD Millipore), and confirmed by sequencing. C321(DE3) Δ*serB* was generated using the genebridges E. coli quick KO kit and confirmed by sequencing. The Keio WT (BW25113), Keio Δ*serB* and Keio Δ*serC* strains were obtained from GE-Dharmacon and the BL21 Δ*serB*(DE3) strain from Addgene (34929).

pSertRNA(10)$_{CUA}$ library generation

To create the pSertRNA(10)$_{CUA}$ library, the 5 bases either side of the anticodon of pSer tRNA$_{CUA}$ in pSertRNA$_{CUA}$ were randomised by Enzymatic Inverse PCR (EI-PCR)[39] using KOD polymerase (Novagen) and primers DTR195/DTR198. The 3.6kb PCR product was gel extracted, digested with SapI, re-circularised using T4 ligase and electroporated into MegaX DH10B (Invitrogen) cells to afford a library with a diversity of ≈5.1×10^7 colony forming units (c.f.u), surpassing the required theoretical diversity of 4.8×10^6 c.f.u. The pool of DNA clones was subsequently extracted by DNA maxiprep and frozen.

SepRS(AC6) library generation

Guided by the structure of the *Archaeoglobus fulgidus* SepRS–tRNA^{Cys}–O-phosphoserine ternary complex, residues E412, E414, T417, P495, I496 and F526 were chosen for mutation due to their proximity to G34 and C35 in the tRNA anticodon. Initially pBK-SepRS was *mutated u*sing the NEB Q5 site directed mutagenesis kit (New England Biolabs), wherein F526 was mutated to every amino acid using the primers in Supplementary Table 3. Each plasmid, including the WT F526 plasmid, was used as a template for individual EI-PCR mediated mutagenesis reactions using Q5 polymerase. To maximise sequence coverage and minimize degeneracy bias, “small-intelligent” oligos were used [24]. The first round of mutation randomised P495, I496 (primers in Supplementary Table 4); the resulting PCR reactions were normalised for concentration, pooled, digested with BsaI, re-circularised by ligation and transformed. A miniprep of this transformation was then used to randomise E412, E414 and T417 (primers in Supplementary Table 4) which was digested with BsaI, ligated and transformed into MegaX DH10B (Invitrogen) to give the SepRS(AC6) library with a diversity of ≈3.1×10^9 c.f.u, with a theoretical requirement of 2.9×10^8 c.f.u.

pSertRNA(XX)$_{CUA}$ selection

The pSertRNA(10)$_{CUA}$ anticodon library was co-transformed into DH10B Δ*serB* with pBK-SepRS and pCDF-EF-Sep-CAT to a diversity of 8×10^9 c.f.u. The pool of transformed cells was grown in LB + 2mM phosphoserine to a density of OD$_{600}$=1.0. 10mL of this culture (approximately 8×10^9 cells) was transferred to 500mL fresh LB + 2mM phosphoserine + 1mM IPTG and to grown to OD$_{600}$=0.6. 2mL of this culture, equivalent to approximately 1×10^9 cells, was plated on LB + 2mM phosphoserine + 1mM IPTG and either 100, 250 or 500μg/ml chloramphenicol. From these plates 960 colonies were picked and phenotyped directly on increasing concentrations of chloramphenicol. 96 clones showing improved chloramphenicol resistance over pSertRNA(Tran)$_{CUA}$ were selected and the tRNA plasmids were isolated by digestion of the pBK-SepRS and pCDF-EF-Sep-CAT. The tRNA plasmids

were retransformed individually into chemically competent DH10B and plated to single colony density, then cultured and miniprepped to yield isolated tRNA plasmids. Sequencing of these clones revealed 15 unique tRNA sequences. These 15 tRNA plasmids and the pSertRNA(Tran)$_{CUA}$ plasmid were transformed into DH10B ΔserB with pCDF-EF_Sep-CAT along with pBk-SepRS or pBK-NoRS, for observing synthetase dependent and independent incorporation respectively, and re-phenotyped on increasing concentrations of chloramphenicol. To test incorporation in the absence of EF-Sep, the 15 mutant tRNAs and pSertRNA(Tran)$_{CUA}$ were co transformed and re-phenotyped as in the presence of EF-Sep but with the pCDF-CAT plasmid in place of pCDF-EF-Sep-CAT.

SepRS(XX) selection

In separate transformations, electrocompetent DH10B Δ*serB* cells containing each of the 13 functional mutant tRNAs (excluding the non-functional mutants A8 and C9) and either pCDF-CAT or pCDF-EF-Sep-CAT were transformed with the SepRS(AC6) anticodon library to a diversity of at least 1×10^9c.f.u, giving full library coverage for every tRNA, and grown overnight. 1mL overnight culture of each pCDF-CAT transformant was used to inoculate 500ml LB + 2mM phosphoserine + 1mM IPTG. 1mL of each overnight culture was used to inoculate 500ml of the LB + 2mM phosphoserine + 1mM IPTG and grown to OD_{600}=0.6. Of these cultures 8×10^9 cells, covering the mutant diversity requirement of 3.7× $\times10^9$ c.f.u, was plated separately on LB + 2mM phosphoserine + 1mM IPTG and either 50, 100, 200, 300 and 400 μg/ml chloramphenicol and 600, 700, 800, 900 and 1000 μg/ml chloramphenicol for pCDF-CAT and pCDF-EF-Sep-CAT cultures respectively. 288 colonies were picked for each set and phenotyped directly on increasing chloramphenicol concentration plates. From this phenotyping, 14 EF-Sep dependent and 3 EF-Sep independent tRNA-synthetase pairs were found. These clones were picked, the plasmid DNA isolated, and the tRNA/synthetase for each was sequenced. The tRNA and CAT reporter plasmids were digested away and the synthetase plasmids were isolated by retransformation of each to single colony density, then miniprepping DNA from these colonies and plasmid isolation. The isolated synthetase plasmids were then co-transformed with their respective reporters, in the presence and absence of the corresponding tRNAs, for re-phenotyping on increasing chloramphenicol concentration plates.

Myoglobin expressions

The 15 selected tRNA(XXX)$_{CUA}$ variants were transformed with pBK-SepRS and either *pCDF-EF-Sep-MyoD127TAG*$_{His6}$ or *pCDF-MyoD127TAG*$_{His6}$ for EF-Sep dependent and independent incorporation. Similarly, the 17 SepRS(XX) variants identified from the SepRS(AC6) library were transformed with their corresponding tRNA and either *pCDF-EF-Sep-MyoD127TAG*$_{His6}$ or *pCDF-EF-Sep-MyoD127TAG*$_{His6}$. 250ul of each overnight culture was inoculated into 50ml LB, grown to OD_{600}=0.6 and induced with 1mM IPTG, 0.2% arabinose and 2mM phosphoserine. The culture were grown for 4 hours and pelleted. Samples were boiled in 4× LDS, subject to SDS-PAGE, transferred to nitrocellulose and full-length protein detected using an anti-his HRP conjugate antibody (Novagen, 1:1000 dilution). To purify the protein for mass spectrometry, the pellet was lysed in 1× lysis buffer (1× BugBuster (Novagen), 50mM Tris, 300mM NaCl, 20mM Imidazole, 0.5mg/mL lysozyme, 50μg/mL DNase, 0.2mM PMSF and 1 protease inhibitor cocktail tablet (Roche)

Europe PMC Funders Author Manuscripts

per 50ml lysis buffer) and the clarified lysate incubated overnight at 4°C with magnetic Ni-NTA beads (Qiagen). Bound protein was eluted in 50mM Tris, 300mM NaCl and 300mM Imidazole and analysed by Electrospray ionization mass spectrometry.

Expression and purification of recombinant GST-CaM incorporating phosphoserine

For overexpression of GST-CaM incorporating phosphoserine, DH10B Δ*serB* or C321(DE3) Δ*serB E. coli* cells were transformed by electroporation with pRSF and pKW vectors, and recovered in 1 ml SOB medium for one hour at 37°C prior to aliquoting to 100 ml LB-KC (LB media with 25 μg/ml kanamycin, and 17.5 μg/ml chloramphenicol) and incubated overnight (37°C, 250 r.p.m., 16 h). The overnight culture was diluted to OD_{600}=0.1 in LB-KC. At OD_{600}=0.5, IPTG (1 mM final concentration) and phosphoserine (2 mM) were added to the culture. The culture was incubated (37°C, 250 r.p.m.) for 5 h, pelleted (3,000g, 10 min, 4°C) and washed twice with 1 ml PBS. The overexpressed proteins were purified using the glutathione affinity chromatography. Cell pellets were resuspended in 1 ml of BugBuster Protein Extraction Reagent (Novagen) (supplemented with 1× protease inhibitor cocktail tablet (Roche), 1 mg ml^{-1} lysozyme (Sigma), 1 mg ml^{-1} DNase I (Sigma)) and lysed (25°C, 250 r.p.m., 1 h). The lysate was clarified by centrifugation (25,000g, 30 min, 4°C). GST containing proteins from the lysate were bound in batch (1 h, 4°C) to 70 μl of glutathione sepharose beads (GE Healthcare). Beads were washed 4 times with 1 ml PBS prior to elution by heating in 1× NuPAGE LDS sample buffer (Invitrogen, 95°C, 5 min) supplemented with 100 mM DTT. All samples were analysed on 4-12% Bis-Tris gels (Invitrogen) with BIO-RAD Low Range Molecular Weight Standard as marker. The gels were subsequently stained with Coomassie Blue (InstantBlue, Expedeon).

Expression and purification of Ubiquitin phosphorylated at serine 65

BL21 Δ*serB*(DE3) cells containing pKW2-*EF*-Sep (a chloramphenicol resistant plasmid containing SepRS 2, pSer $tRNAB4_{CUA}$ and EF-Sep) and pNH-Ub(S65TAG)-His_6 (a tetracycline resistant plasmid containing LacI and Ub(S65TAG)-His_6) were grown overnight (37°C, 220 r.p.m., in TB+CT: TB media containing 25μg/mL chloramphenicol and 25μg/mL tetracycline). The culture was diluted 1:50 into 4L of fresh TB+CT and incubated at 37°C, 220 r.p.m. At OD_{600}=0.6 phosphoserine and IPTG were added to a final concentration of 2mM and 1mM respectively. After incubation (37°C, 220 r.p.m., 4hr), cells were harvested by centrifugation and re-suspended in ice-cold lysis buffer (PBS pH 7.4, 20mM imidazole, 0.5mg/mL lysozyme, 50μg/mL DNase, 0.2mM PMSF), then lysed by sonication. The lysate was clarified by centrifugation (39000 × g, 30 min) and purified using a 5mL HisTrap HP (AKTA explorer), with elution in PBS pH 7.4 + 400mM imidazole. Fractions were determined by SDS-PAGE and were pooled and concentrated to <9mL with a Amicon Ultra-15 3kDa MWCO centrifugal filter device (Millipore). The sample was then dialyzed against 10mM Tris pH 7.6 overnight at 4°C. 1mM DTT was added to the Ub(pSer65)-His_6 sample, followed by UCH-L3 to a final concentration of 15μg/mL. The sample was incubated at (37°C, 1 hr) to remove the C-terminal His_6 tag. Ub (pSer65) was then dialyzed against 50mM Ammonium acetate pH 4.5 overnight at 4°C and purified by cation exchange chromatography using a HiTrap SP HP column (ATKA explorer), with elution in 50mM Ammonium acetate pH 4.5 + 400mM NaCl. Pure fractions were confirmed by SDS-PAGE and were pooled and concentrated as before. The sample was then dialyzed against 10mM

Tris 7.6, 150 mM NaCl at 4°C. Protein yield was quantified by SDS-PAGE gel relative to a commercial ubiquitin standard and then aliquoted and flash frozen at −80°C. Correct masses and site specific phosphoserine incorporation was confirmed by whole protein ESI-MS and LC-MS/MS.

Expression and purification of Nek7 phosphorylated at serine 195

BL21 Δ*ser*(DE3) cells containing pKW2-EF-Sep and pNH-Nek7(D179N/S195TAG)-His_6 (a tetracycline resistant plasmid containing LacI and Nek7(D179N/S195TAG)-His_6) were grown overnight (37°C, 220 r.p.m., in TB+CT: TB media containing 25μg/mL chloramphenicol and 25μg/mL tetracycline). The culture was diluted 1:50 into 4L of fresh TB+CT and incubated at 37°C, 220 r.p.m. At OD_{600}=0.6 phosphoserine and IPTG were added to a final concentration of 2mM and 1mM respectively. After incubation (18°C, 220 r.p.m., 16hr), cells were harvested by centrifugation and re-suspended in ice-cold lysis buffer (50mM HEPES pH 7.5, 300mM NaCl, 5% Glycerol, 20mM Imidazole, 1mM $MgCl_2$, 0.2mM $MnCl_2$ and 50μg/mL DNase), then lysed by sonication. The lysate was clarified by centrifugation (39000 × g, 30 min) and purified using a 5mL HisTrap HP (AKTA explorer), with elution in 50mM HEPES pH 7.5, 300mM NaCl, 5% Glycerol, 250mM Imidazole. D179N-pSer195containing fractions were determined by SDS-PAGE and were pooled and concentrated to <5mL with a Vivaspin 20 (GE Healthcare) with a molecular weight cut off (MWCO) of 10kDa. The protein was then further purified by size exclusion chromatography using a HiLoad Superdex S200 16/60 (GE Healthcare) gel filtration column and eluted in 50mM HEPES pH 7.5, 300mM NaCl, 5% Glycerol, 5mM DTT, the protein was then aliquoted and flash frozen at −80°C until required. Correct masses and site specific phosphoserine incorporation was confirmed by whole protein ESI-MS and LC-MSMS. Nek7(S195TAG)-His_6 was expressed and characterized as described above.

Phosphoserine analog incorporation

Phosphoserine analog incorporation was performed and analysed as for myoglobin $(D127TAG)_{His6}$ expression with phosphoserine, with the exception of using Keio WT strain (BW25113), Keio Δ*serB* or Keio Δ*serC* with plasmids pKW2, pCDF-EF-Sep-*serB* and p15A-Myo$(D127TAG)_{His6}$ for the reporter, as described in the text. DL-AP4 sodium salt (CH_2 substituted phosphoserine analogue, **2**) was obtained from Abcam and used in place of phosphoserine where indicated at a final concentration of 2mM.

Ub(pSer65) *in vitro* E2-discharge assay

Wild-type Parkin (His6–SUMO cleaved) and enzymatic derived Ub(pSer65), hereby referred to as Ub(pSer65*), was expressed and purified as previously described [19]. The E2-charging reaction was performed in a 5μl volume containing 0.5 μg Ube1, 2 μg UbcH7, 50 mM Hepes pH 7.5 and 10 μM FLAG-ubiquitin, in the presence of 2 mM magnesium acetate and 0.2 mM ATP. After an initial incubation of 60 min at 30°C, full-length wild-type Parkin was added in the presence of Ub(pSer65), Ub(pSer65*) and non-phospho-ubiquitin. The reaction was allowed to continue for a further 10 min at 30°C. Reactions were terminated by the addition of 5μl of LDS sample loading buffer and were subjected to SDS/PAGE (4–12 % gel) analysis in the absence of any reducing agent. Gels were stained using InstantBlue.

Ub(pSer65) *in vitro* ubiquitylation assay

Wild-type Parkin (2μg) was added to ubiquitylation assay component mastermix (50mM Tris pH 7.5, 5mM MgCl2, 0.12μM ubiquitin E1, 1μM UbcH7, 2μg of His6–SUMO-Miro1 and 2mM ATP) to a final volume of 50μl. Each reaction contained 0.05mM ubiquitin comprising 25μg of FLAG-ubiquitin (Boston Biochem) mixed with 5μg of Ub(pSer65), Ub(pSer65*) or non-phospho-ubiquitin. Ubiquitylation reactions were carried out at 30°C for 1hr and terminated by the addition of LDS sample buffer. For all assays, reaction mixtures were resolved by SDS/PAGE (4–12% gel). Ubiquitylation reactions were subjected to analysis by immunoblotting as follows: ubiquitin (anti-FLAG antibody; Sigma, 1:10000), Parkin (anti-Parkin antibody; Santa Cruz, 1:5000) and Miro-1 (anti-SUMO1 antibody; 1:2000).

Nek7-His_6 Kinase assay

Nek7-His_6 or the indicated variant and GST-Nek9-CTD were incubated in 50mM HEPES pH7.4, 5mM NaF, 5mM $MnCl_2$, 5mM β-glycerolphosphate, 1mM DTT and Myelin Basic Protein (MBP) (Room temperature, 1h). After incubation, 1mCi radiolabelled ^{32}P-ATP and 4μM unlabelled ATP were added and incubated (30°C, 30min, 200r.p.m). LDS sample loading buffer was added to the reactions and they were subjected to SDS/PAGE (4–12% gel). The gel was stained with Coomassie and then analysed by autoradiography. Efficiency of ^{32}P incorporation into both Nek7(D179N-pSer195)-His_6 by autophosphorylation and MBP substrate was determined by scintillation counting.

Electrospray ionization mass spectrometry

Mass spectra for protein samples were acquired on an Agilent 1200 LC-MS system that employs a 6130 Quadrupole spectrometer. The solvent system used for liquid chromatography (LC) was 0.2 % formic acid in H_2O as buffer A, and 0.2 % formic acid in acetonitrile (MeCN) as buffer B. Samples were injected into Phenomenex Jupiter C4 column (150 × 2 mm, 5 μm) and subsequently into the mass spectrometer using a fully automated system. Spectra were acquired in the positive mode and analyzed using the MS Chemstation software (Agilent Technologies). The deconvolution program provided in the software was used to obtain the mass spectra. Theoretical average molecular weight of proteins with unnatural amino acids was calculated by first computing the theoretical molecular weight of wild-type protein using an online tool (http://www.peptidesynthetics.co.uk/tools/), and then manually correcting for the theoretical molecular weight of unnatural amino acids.

Electrospray ionization tandem mass spectrometry

Polyacrylamide gel slices (1-2 mm) containing the purified proteins or proteins in solution were prepared for mass spectrometric analysis by manual *in situ* enzymatic digestion. Briefly, the excised protein gel pieces were placed in a well of a 96-well microtitre plate and destained with 50% v/v acetonitrile and 50 mM ammonium bicarbonate, reduced with 10 mM DTT, and alkylated with 55 mM iodoacetamide. After alkylation, proteins were digested with 6 ng/μL Trypsin (Promega, UK) overnight at 37 °C. The resulting peptides were extracted in 2% v/v formic acid, 2% v/v acetonitrile. The digest was analysed by nano-scale capillary LC-MS/MS using a Ultimate U3000 HPLC (ThermoScientific Dionex, San

Jose, USA) to deliver a flow of approximately 300 nL/min. A C18 Acclaim PepMap100 5 μm, 100 μm × 20 mm nanoViper (ThermoScientific Dionex, San Jose, USA), trapped the peptides prior to separation on a C18 Acclaim PepMap100 3 μm, 75 μm × 150 mm nanoViper (ThermoScientific Dionex, San Jose, USA). Peptides were eluted with a gradient of acetonitrile. The analytical column outlet was directly interfaced via a modified nano-flow electrospray ionisation source, with a hybrid dual pressure linear ion trap mass spectrometer (Orbitrap Velos, ThermoScientific, San Jose, USA). Data dependent analysis was carried out, using a resolution of 30,000 for the full MS spectrum, followed by ten MS/MS spectra in the linear ion trap. MS spectra were collected over a m/z range of 300–2000. MS/MS scans were collected using an threshold energy of 35 for collision induced dissociation. LC-MS/MS data were then searched against an in house protein sequence database, containing Swiss-Prot and the protein constructs specific to the experiment, using the Mascot search engine programme (Matrix Science, UK). Database search parameters were set with a precursor tolerance of 5 ppm and a fragment ion mass tolerance of 0.8 Da. Two missed enzyme cleavages were allowed and variable modifications for oxidized methionine, carbamidomethyl cysteine, pyroglutamic acid and phosphorylated serine, threonine, tyrosine, were included. MS/MS data were validated using the Scaffold programme (Proteome Software Inc., USA). All data were additionally interrogated manually.

Statistical methods

Bar graphs show the mean of the indicated number of trials. All error bars represent the s.d.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by grants to JWC from the Medical Research Council, UK (U105181009 and UD99999908) and the European Research Council. We are grateful to MRC-LMB Mass Spectrometry for extensive assistance. Mark Richards (Leicester) for the Nek7 plasmid, Thomas Elliott (MRC-LMB) and Mohan Mahesh (MRC-LMB) for assistance. M.M.K.M. is supported by the Wellcome Trust (101022/Z/13/Z), J Macdonald Menzies Charitable Trust and Tenovus (Scotland).

References

1. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. Science. 2002; 298:1912–1934. [PubMed: 12471243]
2. Olsen JV, et al. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell. 2006; 127:635–648. [PubMed: 17081983]
3. Cohen P. The structure and regulation of protein phosphatases. Annual review of biochemistry. 1989; 58:453–508.
4. Ottesen JJ, Huse M, Sekedat MD, Muir TW. Semisynthesis of phosphovariants of Smad2 reveals a substrate preference of the activated T beta RI kinase. Biochemistry. 2004; 43:5698–5706. [PubMed: 15134444]
5. Hejjaoui M, et al. Elucidating the role of C-terminal post-translational modifications using protein semisynthesis strategies: alpha-synuclein phosphorylation at tyrosine 125. J Am Chem Soc. 2012; 134:5196–5210. [PubMed: 22339654]

6. Sauerwald A, et al. RNA-dependent cysteine biosynthesis in archaea. Science. 2005; 307:1969–1972. [PubMed: 15790858]
7. Chin JW. Expanding and Reprogramming the Genetic Code of Cells and Animals. Annu Rev Biochem. 2014; 83:379–408. [PubMed: 24555827]
8. Liu CC, Schultz PG. Adding New Chemistries to the Genetic Code. Annu Rev Biochem. 2010; 79:413–444. [PubMed: 20307192]
9. Fukunaga R, Yokoyama S. Structural insights into the first step of RNA-dependent cysteine biosynthesis in archaea. Nat Struct Mol Biol. 2007; 14:272–279. [PubMed: 17351629]
10. Park HS, et al. Expanding the Genetic Code of Escherichia coli with Phosphoserine. Science. 2011; 333:1151–1154. [PubMed: 21868676]
11. Lee S, et al. A facile strategy for selective incorporation of phosphoserine into histones. Angew Chem Int Ed Engl. 2013; 52:5771–5775. [PubMed: 23533151]
12. Eargle J, Black AA, Sethi A, Trabuco LG, Luthey-Schulten Z. Dynamics of Recognition between tRNA and elongation factor Tu. J Mol Biol. 2008; 377:1382–1405. [PubMed: 18336835]
13. Heinemann IU, et al. Enhanced phosphoserine insertion during Escherichia coli protein synthesis via partial UAG codon reassignment and release factor 1 deletion. FEBS letters. 2012; 586:3716–3722. [PubMed: 22982858]
14. Aerni HR, Shifman MA, Rogulina S, O'Donoghue P, Rinehart J. Revealing the amino acid composition of proteins within an expanded genetic code. Nucleic acids research. 2015; 43:e8. [PubMed: 25378305]
15. Wang K, et al. Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. Nature chemistry. 2014; 6:393–403.
16. Chatterjee A, Xiao H, Schultz PG. Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in Escherichia coli. P Natl Acad Sci USA. 2012; 109:14841–14846.
17. Koyano F, et al. Ubiquitin is phosphorylated by PINK1 to activate parkin. Nature. 2014; 510:162–166. [PubMed: 24784582]
18. Kane LA, et al. PINK1 phosphorylates ubiquitin to activate Parkin E3 ubiquitin ligase activity. The Journal of cell biology. 2014; 205:143–153. [PubMed: 24751536]
19. Kazlauskaite A, et al. Parkin is activated by PINK1-dependent phosphorylation of ubiquitin at Ser65. The Biochemical journal. 2014; 460:127–139. [PubMed: 24660806]
20. Pizer LI. The Pathway and Control of Serine Biosynthesis in Escherichia Coli. J Biol Chem. 1963; 238:3934–3944. [PubMed: 14086727]
21. Zhang CM, Liu C, Slater S, Hou YM. Aminoacylation of tRNA with phosphoserine for synthesis of cysteinyl-tRNA(Cys). Nat Struct Mol Biol. 2008; 15:507–514. [PubMed: 18425141]
22. Swanson R, et al. Accuracy of in vivo aminoacylation requires proper balance of tRNA and aminoacyl-tRNA synthetase. Science. 1988; 242:1548–1551. [PubMed: 3144042]
23. Wang K, Neumann H, Peak-Chew SY, Chin JW. Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. Nature biotechnology. 2007; 25:770–777.
24. Tang L, et al. Construction of "small-intelligent" focused mutagenesis libraries using well-designed combinatorial degenerate primers. Biotechniques. 2012; 52:149–158. [PubMed: 22401547]
25. Nguyen DP, Garcia Alai MM, Kapadnis PB, Neumann H, Chin JW. Genetically Encoding Nε-Methyl- l-lysine in Recombinant Histones. Journal of the American Chemical Society. 2009; 131:14194–14195. [PubMed: 19772323]
26. Neumann H, Wang K, Davis L, Garcia-Alai M, Chin JW. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. Nature. 2010; 464:441–444. [PubMed: 20154731]
27. Mukai T, et al. Codon reassignment in the Escherichia coli genetic code. Nucleic acids research. 2010; 38:8188–8195. [PubMed: 20702426]
28. Lajoie MJ, et al. Genomically recoded organisms expand biological functions. Science. 2013; 342:357–360. [PubMed: 24136966]

29. Kondapalli C, et al. PINK1 is activated by mitochondrial membrane potential depolarization and stimulates Parkin E3 ligase activity by phosphorylating Serine 65. Open Biol. 2012; 2:120080. [PubMed: 22724072]
30. Valente EM, et al. Hereditary early-onset Parkinson's disease caused by mutations in PINK1. Science. 2004; 304:1158–1160. [PubMed: 15087508]
31. Wauer T, et al. Ubiquitin Ser65 phosphorylation affects ubiquitin structure, chain assembly and hydrolysis. The EMBO journal. 2015; 34:307–325. [PubMed: 25527291]
32. Belham C, et al. A mitotic cascade of NIMA family kinases. Nercc1/Nek9 activates the Nek6 and Nek7 kinases. J Biol Chem. 2003; 278:34897–34909. [PubMed: 12840024]
33. O'Regan L, Fry AM. The Nek6 and Nek7 protein kinases are required for robust mitotic spindle formation and cytokinesis. Mol Cell Biol. 2009; 29:3975–3990. [PubMed: 19414596]
34. Richards MW, et al. An autoinhibitory tyrosine motif in the cell-cycle-regulated Nek7 kinase is released through binding of Nek9. Mol Cell. 2009; 36:560–570. [PubMed: 19941817]
35. Liu F, et al. Serendipitous alkylation of a Plk1 ligand uncovers a new binding channel. Nat Chem Biol. 2011; 7:595–601. [PubMed: 21765407]
36. Klingberg R, et al. Analysis of phosphorylation-dependent protein-protein interactions of histone h3. ACS Chem Biol. 2015; 10:138–145. [PubMed: 25330109]
37. Neumann H, Slusarczyk AL, Chin JW. De NovoGeneration of Mutually Orthogonal Aminoacyl-tRNA Synthetase/tRNA Pairs. Journal of the American Chemical Society. 2010; 132:2142–2144. [PubMed: 20121121]
38. Hohn MJ, Park HS, O'Donoghue P, Schnitzbauer M, Soll D. Emergence of the universal genetic code imprinted in an RNA record. P Natl Acad Sci USA. 2006; 103:18095–18100.
39. Stemmer WP, Morris SK. Enzymatic inverse PCR: a restriction site independent, single-fragment method for high-efficiency, site-directed mutagenesis. Biotechniques. 1992; 13:214–220. [PubMed: 1327007]

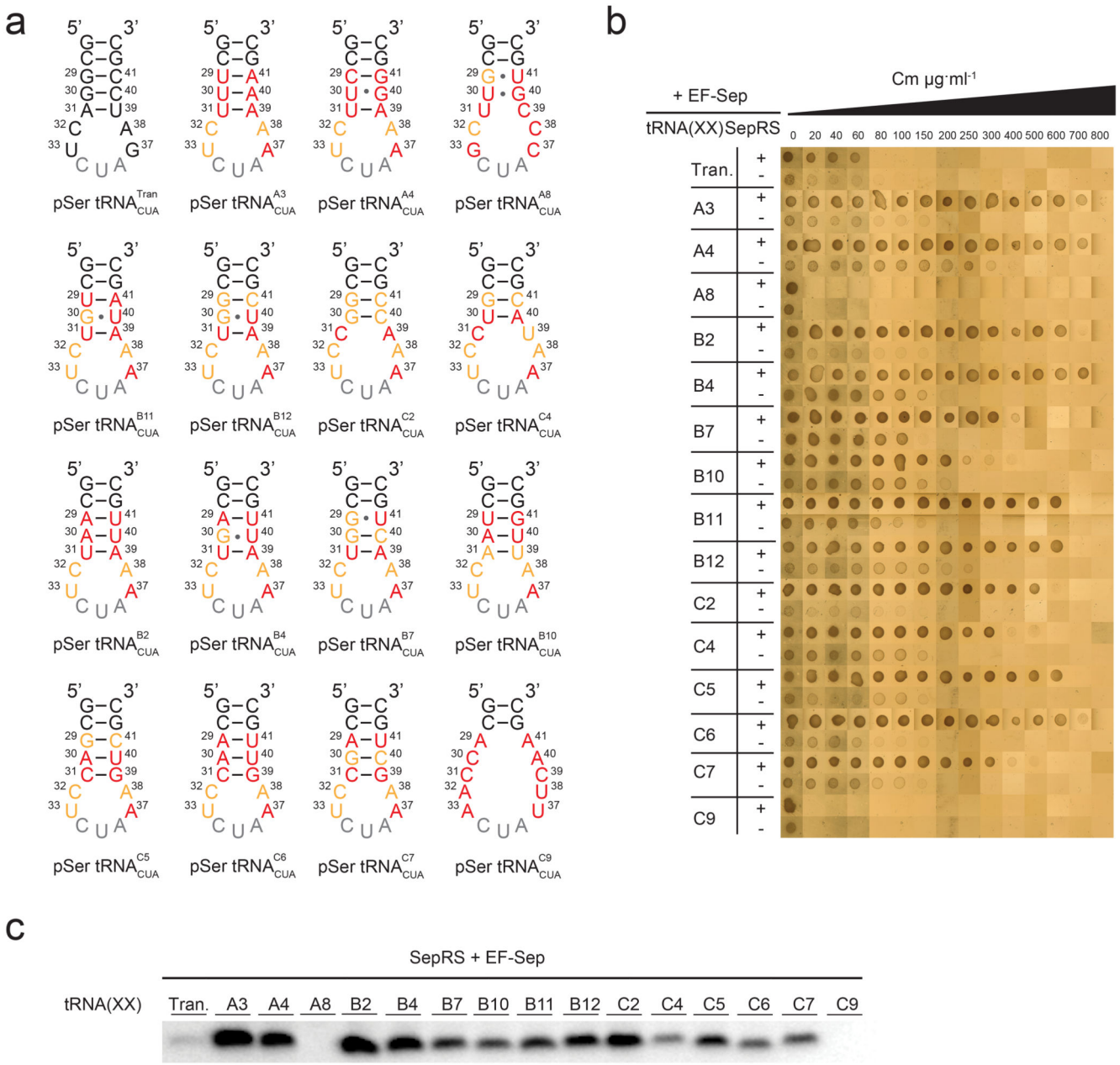

Figure 1.

Phosphoseryl-tRNA anticodon stem and loop evolution enables dramatically improved phosphoserine incorporation in response to the amber stop codon. **A.** pSer tRNA $(N10)_{CUA}$ variants selected for efficient incorporation in response to the amber stop codon. The anticodon sequence transplanted into pSer tRNA is shown in grey. pSer $tRNA(Tran)_{CUA}$ has the CUA anticodon in place of GCA, but is otherwise wildtype. Bases included in the N10 library that were mutated in the selected sequence are shown in red. Bases that were targeted in the N10 library but remained the same after selection are shown in yellow. **B**. Phenotyping evolved pSer tRNA $(XX)_{CUA}$ directed phosphoserine incorporation reveals the dramatic improvement in phosphoserine incorporation resulting from tRNA evolution. Cells

contained the relevant pSer tRNA(XX)$_{CUA}$ variant, EF-Sep and *cat*(112TAG) (expressed from pCDF EF-Sep *cat*112TAG), and either SepRS (encoded by pBKSepRS) or pBKNoRS. **C**. The site-specific incorporation of phosphoserine in recombinant myoglobin is dramatically enhanced using the evolved pSer tRNA(XX)$_{CUA}$s. The western blot detects full-length myoglobin-His6 in cell lysate with an anti-His 6 antibody. Cells contained pCDF EF-Sep myo(127TAG), pBKSepRS and the indicated pSer tRNA (XX)$_{CUA}$ variant. Full blot shown in Supplementary Figure 2.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

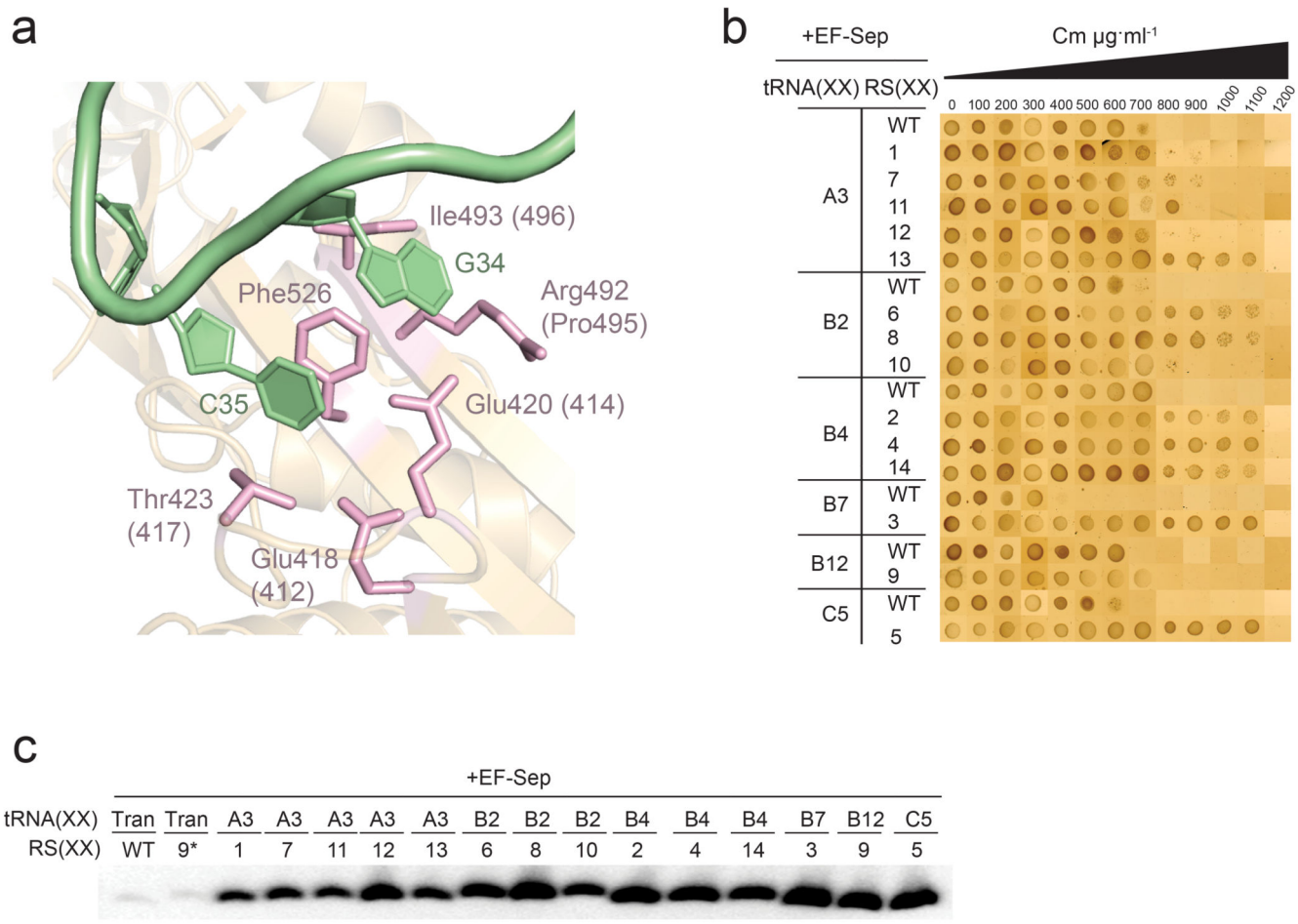

Figure 2.

Evolution of SepRS anticodon binding further improves phosphoserine incorporation with selected pSer tRNA $(XX)_{CUA}$ variants. **A.** The structure of the *A. fulgidus* SepRS/pSer $tRNA_{GCA}$ pair. The tRNA is shown in green, and the anticodon bases indicated. Amino acid residues in SepRS that interact with G34 and C35 in the tRNA, and are targeted for mutagenesis in the SepRS (AC6) library are shown in pink. *A. fulgidus* Sep RS numbering is used in the figure, the *M. maripaludis* numbering and amino acid differences are shown in parenthesis. The figure was created using pymol (www.pymol.org). **B**. Phenotyping evolved SepRS (XX) /pSer tRNA $(XX)_{CUA}$ pair directed phosphoserine incorporation reveals the substantial increase in incorporation efficiency resulting from synthetase evolution. Cells contained the relevant pSer tRNA$(XX)_{CUA}$ variant, EF-Sep and *cat*(112TAG) (expressed from pCDF EF-Sep *cat*112TAG), and SepRS (encoded by pBKSepRS) or the indicated SepRS(XX) variant. RS indicates SepRS, tRNA(XX) indicates the pSertRNA(N10)CUA library member. **C**. The site-specific incorporation of phosphoserine in recombinant myoglobin is dramatically enhanced by evolved SepRS (XX)/pSer tRNA$(XX)_{CUA}$ pairs. The western blot detects full-length myoglobin-His6 in cell lysate with an anti-His 6 antibody. Cells contained pCDF EF-Sep *myo*(127TAG), and the indicated variant of pBKSepRS(XX)

and pSer tRNA $(XX)_{CUA}$. *9** indicates a variant of SepRS previously reported to enhance phosphoserine incorporation[11]. Full blot shown in Supplementary Figure 2.

Europe PMC Funders Author Manuscripts

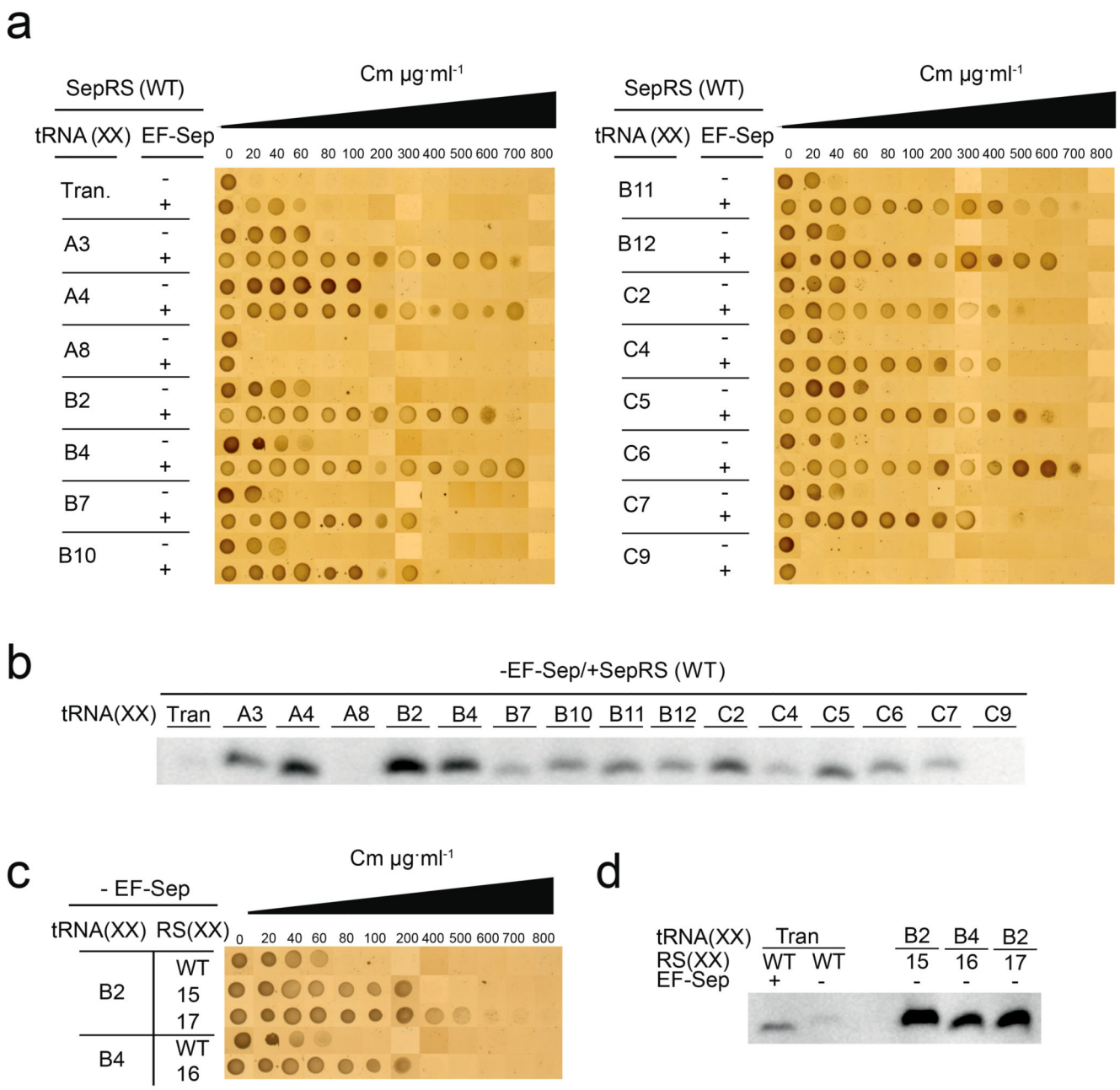

Figure 3.

EF-Sep is not required for site-specific incorporation of phosphoserine using SepRS(XX)/tRNA(XX)$_{CUA}$ pairs. **A**. SepRS/Sep tRNA (XX)$_{CUA}$ variants incorporate phosphoserine using endogenous EF-Tu. tRNA(XX) indicates a selected pSertRNA(N10)$_{CUA}$ library member. Cm is chloramphenicol. Cells contained the relevant pSer tRNA(XX)$_{CUA}$ variant, +/− EF-Sep and *cat*112TAG (expressed from pCDF EF-Sep *cat*112TAG or pCDF EF-Sep *cat*112TAG), and SepRS (encoded by pBKSepRS) **B**. The western blot detects full-length myoglobin-His6 in cell lysate with an anti-His6 antibody. Cells contained pCDF *myo*(127TAG)His6, pBKSepRS, and the indicated pSer tRNA (XX)$_{CUA}$ variant. “Trans” indicates the tRNA in which the CUA anticodon is simply transplanted in place of the native

GCA anticodon. Full blot shown in Supplementary Figure 2. **C**. SepRS(XX) variants selected from SepRS(AC6) in the absence of EF-Sep have substantially enhanced activity with their cognate pSertRNA(XX). **D**. The selected SepRS(XX)/tRNA(XX)$_{CUA}$ pairs are more active with wild-type EF-Tu than the SepRS/pSertRNA$_{CUA}$ pair with EF-Sep. Full blot shown in Supplementary Figure 2.

Europe PMC Funders Author Manuscripts

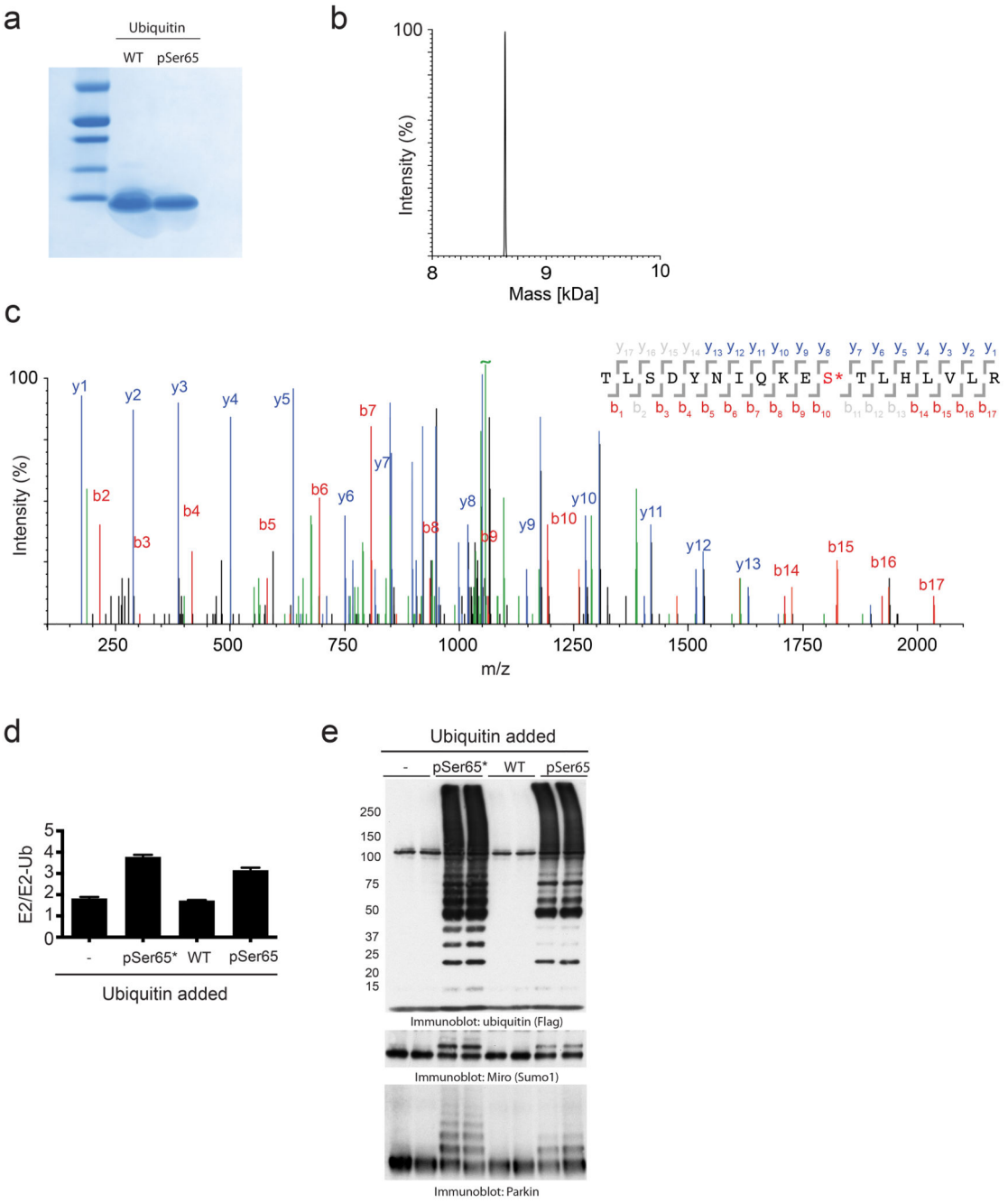

Figure 4.

Expression purification and characterization of Ubiquitin phosphorylated at Serine 65. **A**. Purified ubiquitin bearing site-specific phosphorylation at serine 65. Protein was expressed from BL21*ΔSerB* transformed with pT7 Ub(65TAG)-His6 and pKW2-EF-Sep. The protein was purified by Ni-NTA chromatography, the His6 tagged cleaved with UCHL3 and purified by size exclusion chromatography. The purified yield was 2 mg per L of culture. Full gel shown in Supplementary Figure 2. **B**. Electrospray ionization mass spectrometry demonstrates the quantitative incorporation of phosphoserine into ubiquitin (Found 8639.4

Da, expected 8639.6 Da). **C**. ESI-MS/MS demonstrates the incorporation of phosphoserine at the genetically encoded site (S* = Ser + 80 Da). **D**. Ub (PSer65) stimulates Parkin mediated ubiquitin discharge from loaded E2. An E2-discharge assay was performed by incubating full-length (WT), Parkin in the presence or absence of Ub(PSer65), enzymatically derived Ub(PSer65*) or Ub with the E2 ligase UbcH7 that had been pre-loaded with Ub. The data show the mean of two trials, the error bars represent the s.d. A representative assay gel is shown in Supplementary Figure 7. **E**. Ub(PSer65) activates Parkin E3 ligase mediated ubiquitylation. Full-length Parkin was incubated with the ubiquitylation assay components (E1 and UbcH7) in the presence of His6-SUMO-Miro1 and ubiquitin (comprising 5 μg of Ub(PSer65), Ub(PSer65*), or Ub mixed with 25 μg of FLAG–ubiquitin). Ubiquitin, Parkin, and Miro1 were detected using anti-FLAG, anti-Parkin and anti-SUMO antibodies respectively. Full blots are shown in Supplementary Figure 2.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

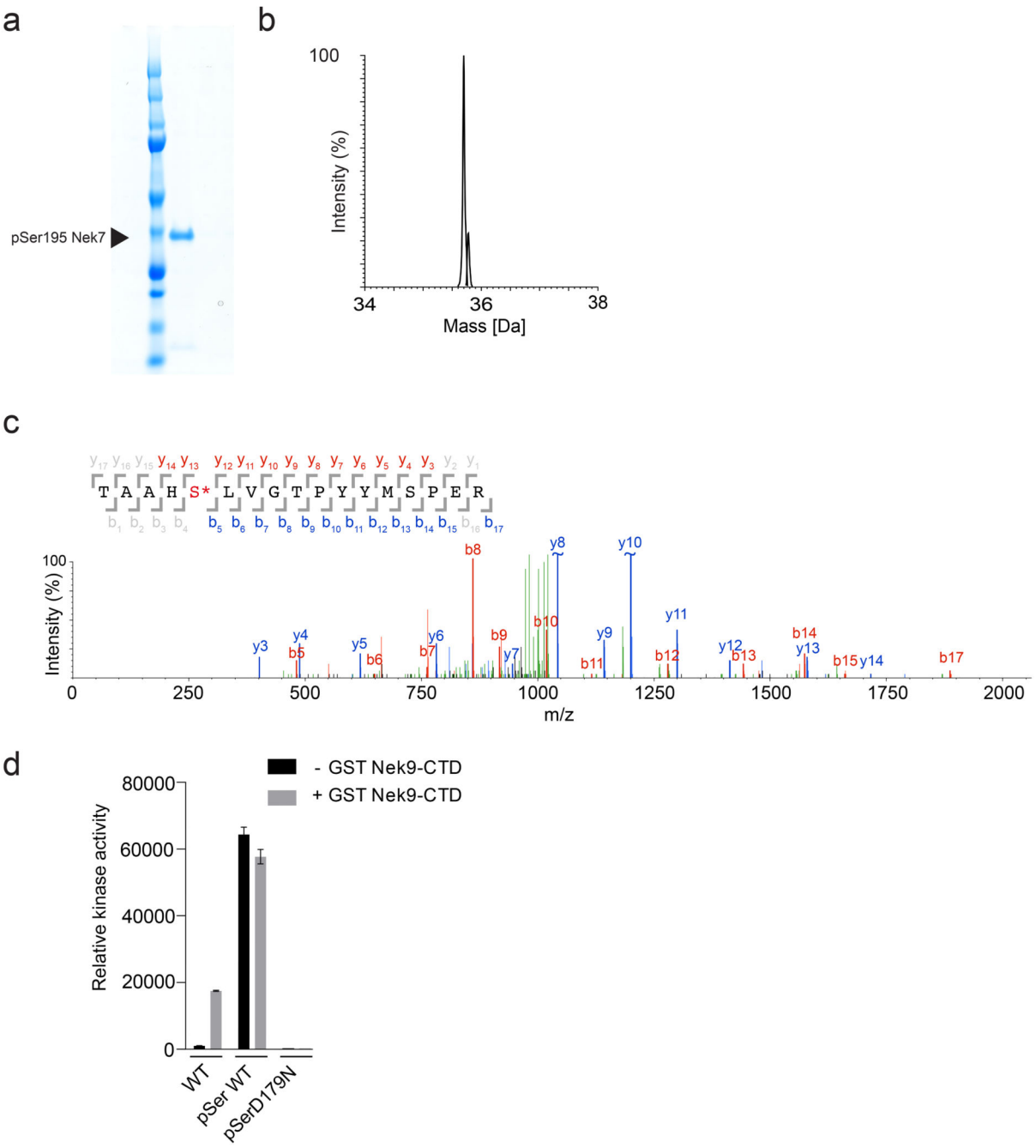

Figure 5.

Synthetic activation of Nek7 bypasses Nek9-CTD mediated activation. **A**. Coomassie stained SDS-PAGE of purified Nek7(pSer195)-$_{His6}$. Full gel in Supplementary Figure 2. **B**. Electrospray ionization mass spectrometry of purified Nek7(pSer195)-$_{His6}$ demonstrates specific phosphorylation of Nek7, found 35,696 Da, expected 35,696 Da. A minor peak, 35,774 Da , corresponds to autophosphorylation. Raw spectra in Supplementary Figure 8 **C**. ESI MS/MS identifies the site of encoded phosphorylation as position 195 S*= (Ser+80) Da). **D**. Kinase assay demonstrating constitutive activity of Nek7(pSer195)-$_{His6}$ compared to

wild type Nek7 with Myelin Basic Protein (MBP) as substrate. Addition of the activating GST-Nek9-CTD protein fragment results in no additional Nek7 activity, indicating that Nek7-S195 phosphorylation alone is sufficient to fully activate the kinase *in vitro*. The data show the mean of three independent trials and the error bar represents the s.d. Representative gels for this experiment are shown in Supplementary Figure 11.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

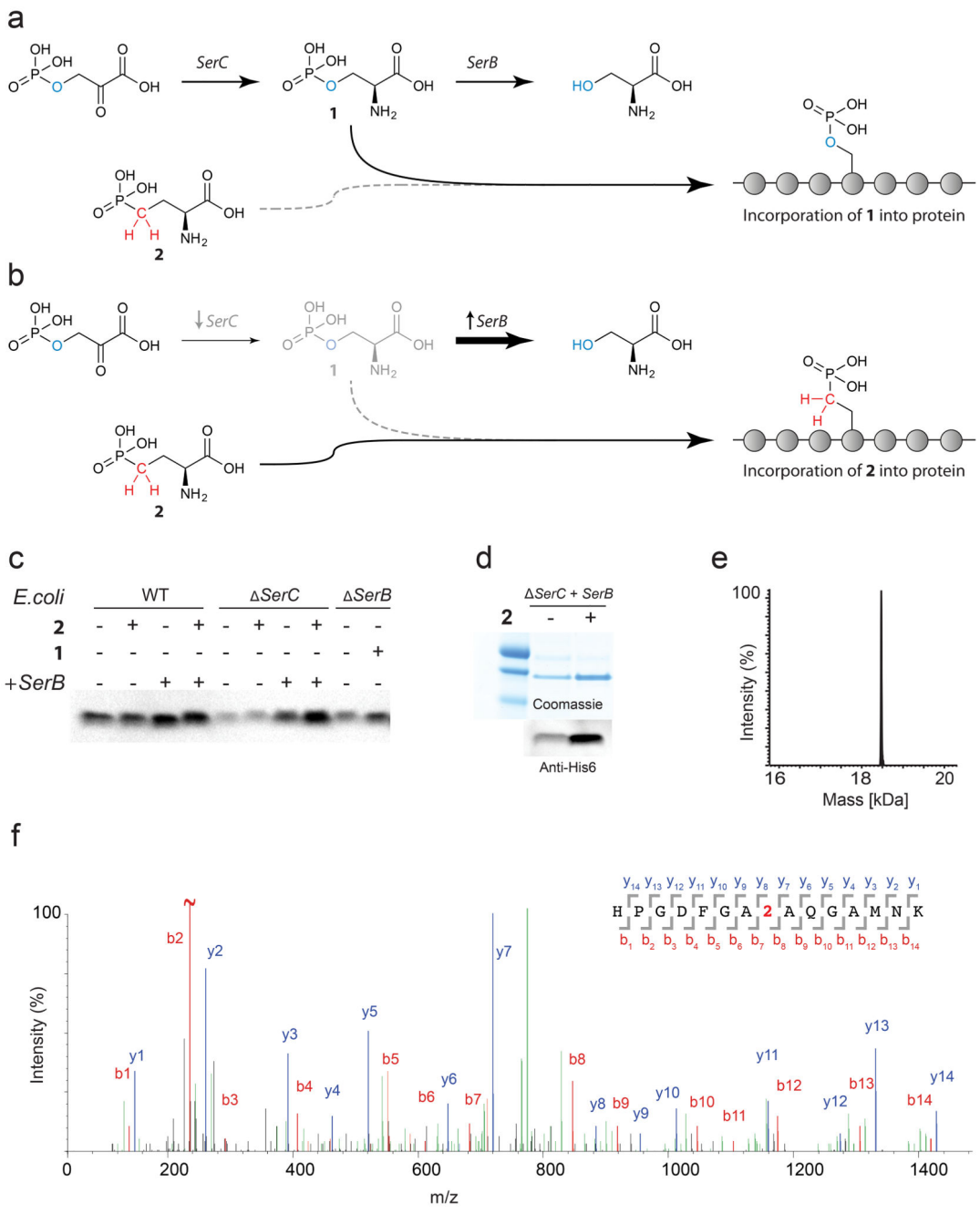

Figure 6.

Genetically encoding a non-hydrolyzable analog of phosphoserine (**2**). **A**. The last two steps in serine biosynthesis. Phosphoserine (**1**) may be site specifically incorporated into proteins. **B**. *serC* deletion and overexpression of *serB* enables the site-specific incorporation of **2** into proteins. **C**. Amino acid dependent expression of *myo(127TAG)His6* expression with SepRS(2)/pSertRNA(B4), EF-Sep requires both *serC* deletion and *serB* overexpression. *E. coli* of the indicated genotype were transformed with *myo(127TAG)His6,* EF-Sep, SepRS(2)/pSertRNA(B4), with or without *serB* overexpression. The full blot is shown in

Supplementary Figure 2. **D**. Expression and purification of *myo(127TAG)His6* in *E.coli* DH10B Δ*serC*, that express *serB*, EF-Sep, SepRS(2)/pSertRNA(B4) in the presence and absence of added **2** (2mM). The full gel and blot is shown in Supplementary Figure 2. **E**. Electrospray ionization mass spectrometry demonstrates the quantitative incorporation of **2** into myoglobin His6 (Found 18405 Da, expected 18405 Da). **F**. ESI-MS/MS demonstrates the incorporation of **2** (Ser + 78 Da) at the genetically encoded site.

Europe PMC Funders Author Manuscripts