

ARTICLE

Received 27 Jan 2015 | Accepted 22 Jul 2015 | Published 9 Sep 2015

DOI: 10.1038/ncomms9130

OPEN

# A flexible codon in genomically recoded *Escherichia coli* permits programmable protein phosphorylation

Natasha L. Pirman<sup>1,3,\*</sup>, Karl W. Barber<sup>1,3,\*</sup>, Hans R. Aerni<sup>1,3</sup>, Natalie J. Ma<sup>2,3</sup>, Adrian D. Haimovich<sup>2,3</sup>, Svetlana Rogulina<sup>1,3</sup>, Farren J. Isaacs<sup>2,3</sup> & Jesse Rinehart<sup>1,3</sup>

Biochemical investigation of protein phosphorylation events is limited by inefficient production of the phosphorylated and non-phosphorylated forms of full-length proteins. Here using a genomically recoded strain of *E. coli* with a flexible UAG codon we produce site-specific serine- or phosphoserine-containing proteins, with purities approaching 90%, from a single recombinant DNA. Specifically, we synthesize human MEK1 kinase with two serines or two phosphoserines, from one DNA template, and demonstrate programmable kinase activity. Programmable protein phosphorylation is poised to help reveal the structural and functional information encoded in the phosphoproteome.

<sup>1</sup>Department of Cellular & Molecular Physiology, Yale University, New Haven, Connecticut 06520-8114, USA. <sup>2</sup>Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520-8114, USA. <sup>3</sup>Systems Biology Institute, Yale University, New Haven, Connecticut 06520-8114, USA. \* These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.R. (email: jesse.rinehart@yale.edu).

Protein phosphorylation modulates most cellular functions and is controlled by networks of kinases and phosphatases that add or remove phosphate groups at precisely defined positions. Large-scale phosphoproteomics efforts have mapped phosphorylation sites on the majority of human proteins<sup>1,2</sup>. Our knowledge of the phosphoproteome has outpaced our understanding of which protein kinases phosphorylate these important sites. Nature has evolved elaborate mechanisms to ensure proper kinase/substrate pairings<sup>3</sup>. Recombinant phosphoprotein synthesis is impeded by the fact that most of these pairs are unknown. We recently described a solution to this problem using *Escherichia coli* strains engineered to incorporate genetically encoded phosphoserine (Sep) into recombinant proteins by employing an orthogonal translation system (SepOTS)<sup>4</sup>. The SepOTS uses a Sep aminoacyl-tRNA synthetase (SepRS) to charge Sep onto a UAG-decoding tRNA<sup>Sep</sup> and an engineered elongation factor Tu (EF-Sep) that delivers Sep-tRNA<sup>Sep</sup> to the ribosome<sup>4</sup>. The SepOTS can direct phosphoserine incorporation into physiologically relevant positions within proteins without any knowledge of upstream kinases and thus provides a simplified platform for phosphoprotein synthesis.

The first described iteration of the SepOTS consisted of two plasmids housing genes for SepRS, EF-Sep and one gene copy of tRNA<sup>Sep</sup> (here referred to as SepOTS $\alpha$ )<sup>4</sup>. SepOTS $\alpha$  was functional in a standard BL21 *E. coli* strain, but inefficiently encoded Sep at UAG amber codons due to competition with release factor one (RF1). This barrier has long been recognized in the field of non-standard amino-acid (NSAA) incorporation and has compromised the yield and purity of NSAA-containing proteins. We recently demonstrated that recoding<sup>5</sup> the native TAG sites terminating seven essential genes in the *E. coli* genome to TAA enabled RF1 deletion and enhanced Sep incorporation<sup>6</sup>. This partially recoded cell (EcAR7. $\Delta$ A) enhanced UAG read-through, but exhibited severe growth impairment and suppression with natural amino acids. The deletion of RF1 causes ribosome stalling on the non-recoded, UAG-containing mRNAs, and protein instability from proteome-wide UAG read-through is elevated in the presence of the SepOTS<sup>6,7</sup> (Fig. 1a). We found that Sep incorporation at UAG sites was enhanced in EcAR7. $\Delta$ A, but we quantified natural amino-acid incorporation accounting for as much as 60% of UAG decoding<sup>7</sup>, indicating that further advances in strain and SepOTS engineering were required. We recently introduced the first genomically recoded organism (C321. $\Delta$ A) where all 321 TAG sites in the genome were reassigned to TAA<sup>8</sup>. This strain not only tolerated the deletion of RF1 but also eliminated all RF1-knockout-associated growth defects observed in partially recoded strains. Furthermore, C321. $\Delta$ A exhibited muted UAG read-through by natural aminoacyl-tRNAs that compete with NSAA incorporation via OTSs<sup>8</sup> (Fig. 1a). The properties of C321. $\Delta$ A suggested that this strain would provide the ideal setting for Sep incorporation, offering a completely open and assignable UAG codon in a background that does not suffer from the low purity and compromised fitness observed in other RF1-knockout strains<sup>6,7</sup>.

Here we demonstrate the utility of genomically recoded *E. coli* to produce phosphoproteins via site-specific incorporation of Sep with an enhanced version of the SepOTS. We take advantage of the flexible UAG codon to genetically program recombinant human MEK1 kinase activity, and we characterize our platform based on phosphoprotein yield, purity and positional bias. Overall, we conclude that our expression system enables robust expression of diverse phosphoproteins for potential biochemical and phosphoproteomic applications.

## Results

**Recoding *E. coli* permits increased phosphoprotein purity.** We aimed to leverage the enhanced properties of C321. $\Delta$ A to further improve the SepOTS and to generate recombinant phosphoproteins with higher yield and purity. To evaluate SepOTS variants, we utilized an established green fluorescent protein (GFP) reporter in which position 17 is permissive to a glutamate-to-Sep substitution<sup>6</sup>. We recently used this reporter and quantitative mass spectrometry to show that using five tRNA<sup>Sep</sup> gene copies in our SepOTS plasmid (here within referred to as SepOTS $\mu$ ) increases the yield and purity of phosphoproteins produced in EcAR7. $\Delta$ A<sup>7</sup>. Our analysis showed that Sep incorporation reached only 40% and UAG read-through by native aminoacyl-tRNAs accounted for 60% of the UAG decoding<sup>7</sup>. To efficiently compare and track this phenomenon across different strains, we adopted a Phos-tag gel shift assay<sup>9</sup> in which phosphoserine-containing proteins are separated from non-phosphorylated proteins and can be quantified on the same western blot<sup>10–12</sup>. The gel shift assays showed that increasing tRNA<sup>Sep</sup> gene copy number from one to five raised Sep incorporation at one UAG in EcAR7. $\Delta$ A from 40 to 60% and was generally in good agreement with our previous studies<sup>7</sup> (Fig. 1b, Supplementary Fig. 1). As a first step, we moved these two SepOTS variants into C321. $\Delta$ A. SepOTS $\alpha$  showed increased Sep incorporation to over 80% in C321. $\Delta$ A compared with only 40% in EcAR7. $\Delta$ A (Fig. 1b). This was the first clear indication that the C321. $\Delta$ A background significantly enhanced the ability of SepOTS $\alpha$  to encode Sep at higher purity. Interestingly, there was little difference in phospho-GFP purity using the SepOTS $\mu$  with 5 tRNA<sup>Sep</sup> gene copies and suggested that simple tRNA gene copy number manipulation would not be sufficient to further enhance SepOTS performance in C321. $\Delta$ A.

## Modifying SepOTS to enhance phosphoserine incorporation.

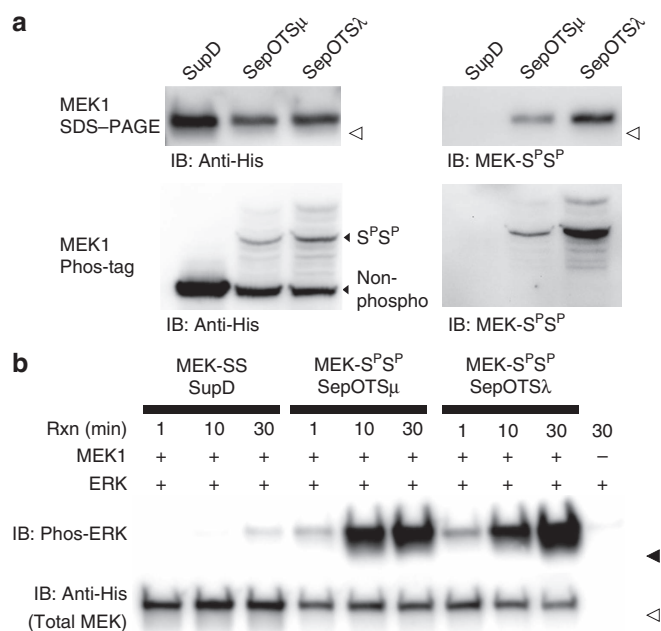
The properties of C321. $\Delta$ A, in particular the increased fitness and reduction of strong UAG read-through observed in partially recoded strains<sup>7,8</sup> that directly competes with Sep insertion, inspired us to perform a more extensive assessment of SepOTS variants in the C321. $\Delta$ A background. To accomplish this, we assembled combinations of known and novel SepOTS components and compared the total phospho-GFP synthesis and purity in C321. $\Delta$ A. We used the enhanced SepRS9 and EF-Sep21 variants and multiple gene copies of tRNA<sup>Sep</sup> known separately to increase SepOTS performance<sup>7,13</sup>. We also introduced a new tRNA<sup>Sep</sup> variant containing a G37A substitution (tRNA<sup>Sep-A37</sup>) known to improve UAG read-through in other natural suppressor tRNAs<sup>14</sup>. We hypothesized that this tRNA<sup>Sep</sup> mutation would enhance Sep decoding by stabilizing anticodon base stacking in the ribosome<sup>15</sup>. We compared relative phospho-GFP expression for 12 different SepOTS combinations (renamed SepOTS $\beta$ –SepOTS $\nu$  for simplicity) to the original two-plasmid system SepOTS $\alpha$  (ref. 4) (Fig. 1c). The most dramatic increase was observed using SepOTS $\lambda$ , which contained enhanced SepRS9, EF-Sep21 and four gene copies of tRNA<sup>Sep-A37</sup> and yielded approximately ninefold elevation in GFP expression compared with SepOTS $\alpha$ . However, the variants containing SepRS9 and EF-Sep21 did not always outperform the original SepOTS $\alpha$  in the combinations tested. We did not observe a consistent correlation between increased tRNA gene copy number and increased phospho-GFP production, but observed that four gene copies of tRNA<sup>Sep-A37</sup> performed better than five gene copies of tRNA<sup>Sep</sup>. This suggests that the G37A mutation increased the performance of the tRNA<sup>Sep</sup> synergistically with a tRNA gene copy number effect. Northern blot analysis of these cells confirmed that tRNA<sup>Sep</sup> levels generally correlate with gene copy number for all SepOTS



(Supplementary Fig. 3b). C321.ΔA transformed with SepOTSλ and S68TAG Bla plasmids to encode Sep at the amber codon yielded expression of inactive Bla and no cell growth in the presence of ampicillin (Supplementary Fig. 3b). This experiment also validates the stability of Sep incorporation at position 68 in the Bla protein, since robust *in vivo* Sep dephosphorylation to Ser would have produced an active Bla enzyme. These results show that the amber codon in C321.ΔA can be used to create an active and inactive enzyme from the same gene. Furthermore, this demonstrates that the phosphorylation state and function of a protein can be programmed by simply employing different translation machinery for the amber codon.

**Programmable active kinase production.** We next wanted to use our enhanced SepOTS systems and the flexible UAG to control phosphoserine-related function in a physiologically relevant mammalian protein containing more than one phosphoserine. Mitogen-activated protein kinase kinase 1 (MEK1) directly controls a broad range of cell cycle functions, is implicated in oncogenesis<sup>18</sup>, and like many kinases<sup>19–22</sup> is widely believed to require phosphorylation at two sites for activity. To measure the robustness of our experimental systems we synthesized MEK1 with genetically encoded Ser at positions 218 and 222 (MEK1-S<sup>PS</sup>) in its activation segment. We first used SepOTSμ (refs 7,23) to compare MEK1-S<sup>PS</sup> expression levels in BL21 RF1<sup>+</sup>, EcAR7.ΔA and C321.ΔA. MEK1-S<sup>PS</sup> expression in BL21 RF1<sup>+</sup> was undetectable due to RF1 activity, whereas both EcAR7.ΔA and C321.ΔA RF1-deficient backgrounds easily produced detectable levels of MEK1-S<sup>PS</sup> (Supplementary Fig. 4). C321.ΔA has a growth curve with a steep log phase and final OD<sub>600</sub> closer to wild-type cells<sup>8</sup> and consistently outperforms the poor growth characteristics of EcAR7.ΔA<sup>6</sup> (Supplementary Fig. 5). This substantial boost in fitness, and the enhanced SepOTS performance in the C321.ΔA strain, rendered it the optimal strain for MEK1-S<sup>PS</sup> expression.

To further investigate MEK1 synthesis we used the highest performing SepOTS variant, SepOTSλ, and found expression of MEK1-S<sup>PS</sup> was twofold higher (~2 mg per l culture) than with SepOTSμ (~1 mg per l culture, Fig. 2a). In parallel, we used SupD in C321.ΔA with the same MEK1 gene (TAGs at 218 and 222) to produce inactive MEK1 containing Ser at positions 218 and 222 (Fig. 2a). MEK1-S<sup>PS</sup> production was confirmed with a commercially available phosphospecific antibody for positions 218 and 222 (Fig. 2a) and further characterized by examining singly phosphorylated forms of MEK1 (MEK1-S<sup>PS</sup> and MEK1-SS<sup>P</sup>) by Phos-tag gel shift (Supplementary Fig. 6a). These studies confirmed the synthesis of singly and doubly phosphorylated MEK1 but revealed that the MEK1 phospho-specific antibody could readily detect both MEK1-S<sup>PS</sup> and MEK1-SS<sup>P</sup>, but not MEK1-S<sup>P</sup>S. We also analysed the MEK1-SS, SS<sup>P</sup> and S<sup>PS</sup>P variants by mass spectrometry and verified Sep incorporation (Supplementary Fig. 6b and Supplementary Table 1). Despite our best efforts, we were unable to directly observe the doubly phosphorylated peptide likely due to poor ionization. To overcome this problem, we treated MEK1-S<sup>PS</sup> with calf intestinal alkaline phosphatase to dephosphorylate the protein. Calf intestinal alkaline phosphatase-treated MEK1-S<sup>PS</sup> produced a peptide containing two serine residues at the S218 and S222 positions (absent in untreated MEK1-S<sup>PS</sup> sample), confirming that Sep was originally inserted at both positions within the protein. We next used our established data analysis techniques to identify natural amino acids incorporated at the two different UAG sites in the MEK1 activation loop<sup>7</sup> (Supplementary Table 1). These non-phosphorylated MEK1 peptides explain a portion of the unshifted bands detected in



**Figure 2 | A single-gene system for robust production of phosphorylated and non-phosphorylated human kinases.** (a) Analysis of MBP-MEK1 S218TAG/S222TAG expression in the fully recoded C321.ΔA strain using plasmids that enable incorporation of Ser (one-plasmid system encoding SepRS, EF-Sep and 2 × tRNA<sup>SupD</sup> denoted SupD) or Sep (one-plasmid system encoding SepRS, EF-Sep and either 5 × tRNA<sup>Sep</sup> denoted SepOTSμ or 4 × tRNA<sup>Sep-A37</sup> denoted SepOTSλ) at amber codons. Total expression was evaluated by western blot using an antibody against the C-terminal 6xHis tag on MBP-MEK1. Loading was normalized by OD<sub>600</sub> measurement. Presence of phosphorylated MBP-MEK1 was verified using a phosphospecific antibody (extended data shown in Supplementary Fig. 6a). Phosphoprotein purity was determined by Phos-tag gel shift and western blot by comparing the relative intensity of the upper bands compared with the lower band, corresponding to the phosphorylated and non-phosphorylated forms of MBP-MEK1, respectively. (b) Assay of MBP-MEK1 kinase activity monitored by phosphorylation of a kinase-dead ERK2 substrate *in vitro*. Representative blots showing phosphorylated (anti-Phos-ERK) and total MEK (anti-His) are shown. Assays and westerns were performed in triplicate and representative results are shown. The open and solid triangles denote 75- and 37-kDa molecular weight markers, respectively. Phos-tag gels are run without molecular weight standards.

our Phos-tag analysis and, consistent with our previous studies<sup>7,8</sup>, demonstrate that near-cognate suppression of the amber codon can lead to natural amino-acid incorporations that interfere with SepOTS activity and phosphoprotein purity.

A side-by-side comparison of MEK1 production showed that SepOTSμ and SepOTSλ yielded ~12% and ~30% phosphorylated MEK1-S<sup>PS</sup>P, respectively (Fig. 2a). While SepOTSλ increased the yield of MEK1-S<sup>PS</sup>P as predicted by the GFP reporter studies, the yield of phosphorylated MEK1 was lower than phospho-GFP. This could be due to instability of phosphorylated MEK or context-specific effects of UAG suppression in the MEK activation segment. To examine protein stability we conducted time course expression studies that showed higher MEK1-S<sup>PS</sup>P phosphoprotein yield with shorter expression times (Supplementary Fig. 6c). We assessed positional effects of Sep incorporation or stability using several singly and doubly phosphorylated forms of GFP (Supplementary Fig. 6d). The differences in phosphoserine incorporation between the GFP variants led us to conclude that UAG context and

phosphoprotein stability are both contributing factors to the overall purity of recombinant phosphoproteins.

To demonstrate the utility of the flexible UAG codon in synthesizing active mammalian kinases, we carried out MEK1 *in vitro* phosphorylation assays. MEK1-S<sup>PS</sup>P produced robust *in vitro* ERK2 phosphorylation while non-phosphorylated MEK1 (produced with SupD in C321.ΔA) was completely inactive (Fig. 2b). This demonstrates that the activity of a protein kinase can be programmed by simply using different translational machinery for the UAG codon in C321.ΔA cells. We envision that this flexibility could be leveraged in experiments with large panels or arrays of recombinant enzymes designed to explore the phosphoproteome at greater depth.

## Discussion

These results clearly demonstrate that the physiological properties of protein phosphorylation can be controlled by a programmable genetic code. The flexible UAG codon in C321.ΔA can be manipulated with different OTS systems to produce proteins encoding either Sep or Ser at amber codons. We show that C321.ΔA is an optimal host for phosphoprotein expression with SepOTS<sub>ε</sub> and SepOTS<sub>λ</sub> spanning a ninefold expression range without compromising purity. These systems provide a platform to reveal the functional consequences of serine phosphorylation. As a proof of principle, we reproduced the phosphorylation events underlying the MEK1 activity switch that controls a human signalling cascade. Our programmable UAG can be used to model other important kinase regulatory switches to enable applications such as kinase substrate discovery and screens for novel kinase inhibitors. More broadly, different pairs of natural suppressor tRNAs and OTSs could be introduced into C321.ΔA to explore more types of post-translational modification<sup>24</sup>. Certainly, exploring the function of protein phosphorylation in more enzymatic classes and across different species should now be possible. We recently showed that our C321.ΔA-based phosphoserine systems can be used for *in vitro* protein synthesis platforms, which may further extend the experimental landscape<sup>25</sup>. Guided by the large number of physiological protein phosphorylation sites<sup>2</sup>, systematic investigations into the structure and function of any phosphoproteome could be within reach.

## Methods

**Transformations and strain storage.** All *E. coli* strains used in this study were made chemically competent using a standard RbCl<sub>2</sub> method. EcAR7.ΔA was co-transformed with E17TAG GFP in the modified pCR-Blunt II-TOPO vector with the signified SepOTS version. C321.ΔA was co-transformed with PCRT7 NT Topo tetR pLtetO plasmid containing S2TAG, E17TAG, Q157TAG, S2TAG/E17TAG, E17TAG/Q157TAG GFP, maltose binding protein-MEK1 (MBP-MEK1) (S218TAG/S222), MBP-MEK1 (S218/S222TAG) or MBP-MEK1 (S218TAG/S222TAG) cassettes with the designated SepOTS. BL21 RF1 + was co-transformed with MBP-MEK1 and the signified SepOTS. All strains were stored as frozen glycerol stocks and restreaked on selective Luria-Bertani (LB) agar plates before protein expression. Strains harbouring the following plasmids were grown with the indicated antibiotic concentrations: E17TAG GFP in the modified pCR-Blunt II-TOPO plasmids, 25 µg per ml Zeocin (Zeo); all pCRT7 NT Topo tetR pLtetO plasmids, 100 µg per ml ampicillin (Amp); SupD and all SepOTS plasmids, 25 µg per ml kanamycin (Kan) with the exception of the two plasmid SepOTS<sub>ε</sub> which requires 6 µg per ml tetracycline (Tet) and 25 µg per ml Kan; β-lactamase S68TAG plasmid, 10 µg per ml chloramphenicol (Cam) or 100 µg per ml Amp, as indicated. All cultures were started from a freshly streaked glycerol stock on LB agar plates with the appropriate combination of antibiotics and 0.08% glucose. Detailed information about strain and plasmid generation can be found in the Supplementary Methods. All primer sequences can be found in Supplementary Table 2. Sequences of tRNA-related synthesized genes can be found in Supplementary Table 3.

**GFP variant protein expression.** *E. coli* strains were streaked from -80 °C frozen glycerol stocks on LB agar selective plates and grown ~20 (BL21-based strains), ~24 (C321.ΔA-based strains), or ~48 h (EcAR7.ΔA-based strains) at 30 °C. Five

colonies were inoculated in 5 ml of LB media supplemented with appropriate antibiotic and 0.08% glucose. Precultures were grown ~16 h at 30 °C shaking at 230 r.p.m. Cells were diluted to OD<sub>600</sub> of 0.15 a.u. in 20 ml of LB with antibiotics, 0.08% glucose, 2 mM *O*-phospho-L-serine (Sep), and protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside and 100 ng per ml anhydrotetracycline. Protein was expressed for 20 h at 30 °C shaking at 230 r.p.m. unless otherwise noted. An equivalent number of cells as 1 ml of OD<sub>600</sub> 2.5 a.u. was collected and spun down at 4,000g for 5 min at 4 °C. Media was aspirated and dry cell pellets were stored at -80 °C. Transformation and retention of the correct SepOTS variant (including tRNA gene copy number) were confirmed by PCR from frozen cell pellets. Cells were resuspended in 40 µl bacterial lysis buffer (50 mM Tris/HCl (pH 7.4), 150 mM NaCl, 1 mM dithiothreitol (DTT), 50 mM NaF, 1 mM NaVO<sub>4</sub> and 5% glycerol) supplemented with cOmplete mini-EDTA-Free protease inhibitor cocktail tablets (Roche) and 1 × BugBuster detergent (Novagen), and kept on ice for 10 min. Samples were then spun down at 21,000g for 7 min at 4 °C and the supernatant (lysate) was transferred to a new tube and stored at -80 °C.

For western analysis, equal volumes of each sample were diluted in 1 × Laemmli buffer (Bio-Rad). Samples were run on a 15-well 4–15% acrylamide gels (Bio-Rad) or on a handcast 15-well 12% acrylamide gel containing 100 µM Phos-tag Acrylamide (Waco Pure Chemical Industries, Inc., AAL-107). Transferred polyvinylidene difluoride (PVDF) membranes were blotted with 1:2,500 Anti-His (6xHis Epitope TAG, PA1-983B, Thermo Fisher Scientific), followed by 1:10,000 DAR-HRP (Peroxidase-conjugated AffiniPure Donkey Anti-Rabbit IgG, 711-035-152, Jackson ImmunoResearch). Signal was detected by enhanced chemiluminescence (Bio-Rad) imaged on a ChemiDoc XRS + CCD camera (Bio-Rad). Densitometry was performed using the Bio-Rad Image Lab software. Three biological replicates (starting from glycerol stocks) were performed for both the total expression and phospho-shifted westerns. Uncropped western blot images corresponding to text and Supplementary Figures are included in Supplementary Figs 7–12.

**MBP-MEK1 recombinant protein expression and purification.** The 5-ml precultures were inoculated with 5–20 colonies and grown overnight to confluency in LB media containing 0.08% glucose and antibiotics. The precultures were diluted to OD<sub>600</sub> 0.15 a.u. in 100 ml LB media containing 0.08% glucose, antibiotics and 2 mM Sep, and were incubated at 30 °C, 230 r.p.m. for ~3 h to an OD<sub>600</sub> of 0.8 a.u. Protein expression was then induced with 100 ng per ml anhydrotetracycline and 1 mM isopropyl β-D-1-thiogalactopyranoside, and cultures were grown at 20 °C, 230 r.p.m. for ~20–22 h unless otherwise noted. An equivalent number of cells as 1 ml of OD<sub>600</sub> 2.5 a.u. was collected and spun down at 4,000g for 5 min at 4 °C. The remaining cells were collected at 4,000g, 20 min, 4 °C. Pellets were resuspended in ~30 ml of the used LB media and transferred to 50-ml centrifuge tubes and centrifuged under the same conditions again. All media was decanted, and pellets were stored at -80 °C.

The 1-ml OD<sub>600</sub> 2.5 a.u. cell pellets were resuspended in 40 µl bacterial lysis buffer (with protease inhibitors and detergent, as for E17TAG GFP expression above) and kept on ice for 10 min. Samples were then spun down at 21,000g for 7 min at 4 °C and the supernatant (lysate) was transferred to a new tube and stored at -80 °C.

For western analysis, equal volumes of OD<sub>600</sub> normalized lysates were diluted in 1 × Laemmli buffer (Bio-Rad). Samples were run on a 15-well 4–15% acrylamide gels (Bio-Rad) or on a handcast 10-well 7.5% acrylamide gel containing 25 µM Phos-tag Acrylamide. Transferred PVDF membranes were blotted with either 1:2,500 anti-His (6xHis Epitope TAG, PA1-983B, Thermo Fisher Scientific) for total MEK1 or 1:1,000 Anti-MEK-S<sup>PS</sup>P (Phospho-MEK1/2 (Ser217/221), 9154, Cell Signaling Technology), followed by 1:10,000 DAR-HRP. Signal was detected by enhanced chemiluminescence (Bio-Rad) imaged on a ChemiDoc XRS + CCD camera. Densitometry was performed using the Bio-Rad Image Lab software.

Samples for MBP-MEK1 (S218TAG/S222TAG) in C321.ΔA harbouring either the SepOTS<sub>λ</sub> or SupD plasmids expression time course were expressed as described above, with exception that after induction 1-ml aliquots were removed at denoted time points and spun down at 4,000g for 5 min at 4 °C. All media was decanted, and pellets were stored at -80 °C. For western analysis samples were lysed in 80 µl of 2 × Laemmli buffer and ran on a handcast 15-well 7.5% acrylamide gel containing 25 µM Phos-tag Acrylamide. Membranes were blotted with Anti-His as described above.

Cell pellets for purification were thawed in a 37 °C water bath for ~20 s and then put on ice. The pellets were resuspended in 5 ml of bacterial lysis buffer for sonication (50 mM Tris/HCl (pH 7.4), 500 mM NaCl, 0.5 mM EDTA, 0.5 mM EGTA, 1 mM DTT, 1 mg per ml lysozyme, 50 mM NaF, 1 mM NaVO<sub>4</sub>, 10% glycerol and complete mini-EDTA-Free protease inhibitor cocktail tablets) and incubated on ice for 30 min, followed by sonication. The lysates were centrifuged at 22,000g, 15 min, 4 °C and the clarified lysate was transferred to a 15-ml tube, and centrifuged under the same conditions again to remove all remaining insoluble material. The 200 µl bed volumes of Ni-NTA agarose resin (Qiagen Valencia, CA) were transferred to Pierce spin columns (Thermo Scientific Waltham, MA) and resin was equilibrated with 5 ml of Ni-NTA equilibration buffer (50 mM Tris/HCl (pH 7.4), 500 mM NaCl, 0.5 mM EDTA, 0.5 mM EGTA, 1 mM DTT, 50 mM NaF, 1 mM NaVO<sub>4</sub> and 10% glycerol). The clarified lysate was loaded onto the column via syringe then the column was washed with 10 ml of Ni-NTA wash buffer

(50 mM Tris/HCl (pH 7.4), 500 mM NaCl, 0.5 mM EDTA, 0.5 mM EGTA, 1 mM DTT, 50 mM NaF, 1 mM NaVO<sub>4</sub>, 10% glycerol and 20 mM imidazole). Protein was eluted using Ni-NTA elution buffer (50 mM Tris/HCl (pH 7.4), 500 mM NaCl, 0.5 mM EDTA, 0.5 mM EGTA, 1 mM DTT, 50 mM NaF, 1 mM NaVO<sub>4</sub>, 10% glycerol and 250 mM imidazole) and 400 µl elutions were collected in a clean Eppendorf tubes. Each fraction was assessed by SDS–polyacrylamide gel electrophoresis (SDS–PAGE).

The eluates were pooled, concentrated and buffer exchanged into the protein storage buffer (50 mM Tris/HCl (pH 7.4), 150 mM NaCl, 1 mM DTT and 20% glycerol) using a 0.5-ml Amicon Ultra centrifugal filter (Millipore Billerica, MA) and the protein was stored at –20 °C. The protein concentration was estimated by UV280 and by comparing known quantities of BSA standards on an SDS–PAGE gel.

K54R ERK2 was expressed and purified as described for MEK1 above.

**MEK1-S<sup>P</sup> kinase activity assay.** Kinase activity of purified MBP-MEK1 (S218/S222 or S<sup>P</sup>218/S<sup>P</sup>222) was evaluated by measuring ERK2 phosphorylation. In all, 1.0-µM MBP-MEK1 variants were pre-incubated in kinase activity buffer (50 mM Tris/HCl (pH 7.4), 150 mM NaCl, 1 mM DTT, 20% glycerol, 10 mM MgCl<sub>2</sub> and 1 mM ATP) at 30 °C for 5 min, then 2.5 µM ERK2 substrate was added to the reaction and further incubated at 30 °C. A 7.5 µl volume of the reaction was removed at 1, 10 and 30 min time points and quenched with 7.5 µl of 2 × Laemmli sample buffer (Bio-Rad), then heated to 55 °C for 5 min. A negative control was run with only ERK2 substrate for 30 min and quenched in the same manner as kinase-containing samples. The quenched reactions were run on 15-well 4–15% acrylamide SDS–PAGE gels and transferred to a PVDF membrane. Each membrane was cut between the 50–75-kDa protein markers and the bottom portion of the membrane was blotted with 1:1,000 anti-Phos-Erk antibody (Phospho p44/42 MAPK (Erk1/2) (Thr 202/Tyr204), 9101, Cell Signaling Technology) and the top portion of the membrane was blotted with 1:2,500 Anti-His antibody (6xHis Epitope TAG, PA1-983B, Thermo Fisher Scientific) for total MBP-MEK1, followed by 1:10,000 DAR–HRP. Signal was detected by enhanced chemiluminescence (Bio-rad) imaged on a ChemiDoc XRS + CCD camera. The kinase activity assay was run in triplicate using the same purified preparation of MBP-MEK1 kinase and K54R ERK2.

## References

- Olsen, J. V. & Mann, M. Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol. Cell. Proteomics* **12**, 3444–3452 (2013).
- Hornbeck, P. V. *et al.* PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* **40**, D261–D270 (2012).
- Mok, J. *et al.* Deciphering protein kinase specificity through large-scale analysis of yeast phosphorylation site motifs. *Sci. Signal* **3**, ra12 (2010).
- Park, H. S. *et al.* Expanding the genetic code of *Escherichia coli* with phosphoserine. *Science* **333**, 1151–1154 (2011).
- Isaacs, F. J. *et al.* Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science* **333**, 348–353 (2011).
- Heinemann, I. U. *et al.* Enhanced phosphoserine insertion during *Escherichia coli* protein synthesis via partial UAG codon reassignment and release factor 1 deletion. *FEBS Lett.* **586**, 3716–3722 (2012).
- Aerni, H. R., Shifman, M. A., Rogulina, S., O'Donoghue, P. & Rinehart, J. Revealing the amino acid composition of proteins within an expanded genetic code. *Nucleic Acids Res.* **43**, e8 (2014).
- Lajoie, M. J. *et al.* Genomically recoded organisms expand biological functions. *Science* **342**, 357–360 (2013).
- Kinoshita, E., Kinoshita-Kikuta, E., Takiyama, K. & Koike, T. Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Mol. Cell. Proteomics* **5**, 749–757 (2006).
- Takeya, K., Loutzenhiser, K., Shiraishi, M., Loutzenhiser, R. & Walsh, M. P. A highly sensitive technique to measure myosin regulatory light chain phosphorylation: the first quantification in renal arterioles. *Am. J. Physiol. Renal Physiol.* **294**, F1487–F1492 (2008).
- Yamada, S. *et al.* Separation of a phosphorylated histidine protein using phosphate affinity polyacrylamide gel electrophoresis. *Anal. Biochem.* **360**, 160–162 (2007).
- Kinoshita-Kikuta, E., Aoki, Y., Kinoshita, E. & Koike, T. Label-free kinase profiling using phosphate affinity polyacrylamide gel electrophoresis. *Mol. Cell. Proteomics* **6**, 356–366 (2007).
- Lee, S. *et al.* A facile strategy for selective incorporation of phosphoserine into histones. *Angew. Chem. Int. Ed.* **52**, 5771–5775 (2013).
- Eggertsson, G. & Soll, D. Transfer ribonucleic acid-mediated suppression of termination codons in *Escherichia coli*. *Microbiol. Rev.* **52**, 354–374 (1988).
- Demeshkina, N., Jenner, L., Westhof, E., Yusupov, M. & Yusupova, G. A new understanding of the decoding principle on the ribosome. *Nature* **484**, 256–259 (2012).
- Steege, D. A. A nucleotide change in the anticodon of an *Escherichia coli* serine transfer RNA results in supD-amber suppression. *Nucleic Acids Res.* **11**, 3823–3832 (1983).
- Normanly, J., Ogden, R. C., Horvath, S. J. & Abelson, J. Changing the identity of a transfer RNA. *Nature* **321**, 213–219 (1986).
- Scholl, F. A., Dumesic, P. A. & Khavari, P. A. Effects of active MEK1 expression in vivo. *Cancer Lett.* **230**, 1–5 (2005).
- Zheng, C. F. & Guan, K. L. Activation of MEK family kinases requires phosphorylation of two conserved Ser/Thr residues. *EMBO J.* **13**, 1123–1131 (1994).
- Alessi, D. R. *et al.* Mechanism of activation of protein kinase B by insulin and IGF-1. *EMBO J.* **15**, 6541–6551 (1996).
- Kyriakis, J. M., Brautigan, D. L., Ingebritsen, T. S. & Avruch, J. pp54 microtubule-associated protein-2 kinase requires both tyrosine and serine/threonine phosphorylation for activity. *J. Biol. Chem.* **266**, 10043–10046 (1991).
- Raingaud, J. *et al.* Pro-inflammatory cytokines and environmental stress cause p38 mitogen-activated protein kinase activation by dual phosphorylation on tyrosine and threonine. *J. Biol. Chem.* **270**, 7420–7426 (1995).
- Sawyer, N. *et al.* Designed phosphoprotein recognition in *Escherichia coli*. *ACS Chem. Biol.* **9**, 2502–2507 (2014).
- Davis, L. & Chin, J. W. Designer proteins: applications of genetic code expansion in cell biology. *Nat. Rev. Mol. Cell Biol.* **13**, 168–182 (2012).
- Oza, J. *et al.* Robust production of recombinant phosphoproteins using cell-free protein synthesis. *Nat. Commun.* **6**, 8168 (2015).

## Acknowledgements

This research was supported by National Institutes of Health (NIDDK-K01DK089006 and P01DK01743341 to J.R. and NIH-MSTP-TG-T32GM07205 to A.D.H.); Defense Advanced Research Projects Agency (N66001-12-C-4211 to J.R. and F.J.I.); DuPont, Inc. (F.J.I.); Gen9, Inc. (F.J.I.); The Arnold and Mabel Beckman Foundation (F.J.I.); K.W.B. is supported by NIH-T32GM100884 and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1122492. We would also like to thank Noah Reynolds, Hubert Salvail and Eduardo Groisman for assistance with northern blot experiments.

## Author contributions

N.L.P. and K.W.B. assisted in cloning, expression, performing experiments, analysing data, preparing figures and writing the manuscript; H.R.A. performed the mass spectrometry experiments; N.J.M. and A.D.H. designed and constructed the modified C321.ΔA strain; S.R. assisted in cloning and expression; F.J.I. supervised the construction of the recoded strains, and provided advice on experimental design and manuscript content and J.R. supervised the research, assisted in analysing data, preparing the figures and writing the manuscript.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Pirman, N. L. *et al.* A flexible codon in genomically recoded *Escherichia coli* permits programmable protein phosphorylation. *Nat. Commun.* **6**:8130 doi: 10.1038/ncomms9130 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>