



Accessing isotopically labeled proteins containing genetically encoded phosphoserine for NMR with optimized expression conditions

Received for publication, July 11, 2022, and in revised form, October 11, 2022 Published, Papers in Press, October 17, 2022,

<https://doi.org/10.1016/j.jbc.2022.102613>

Cat Hoang Vesely^{1,2}, Patrick N. Reardon³ , Zhen Yu², Elisar Barbar², Ryan A. Mehl^{1,2}, and Richard B. Cooley^{1,2,*}

From the ¹GCE4All Research Center, ²Department of Biochemistry and Biophysics, and ³Oregon State University NMR Facility, Oregon State University, Corvallis, Oregon, USA

Edited by Wolfgang Peti

Phosphoserine (pSer) sites are primarily located within disordered protein regions, making it difficult to experimentally ascertain their effects on protein structure and function. Therefore, the production of ¹⁵N- (and ¹³C)-labeled proteins with site-specifically encoded pSer for NMR studies is essential to uncover molecular mechanisms of protein regulation by phosphorylation. While genetic code expansion technologies for the translational installation of pSer in *Escherichia coli* are well established and offer a powerful strategy to produce site-specifically phosphorylated proteins, methodologies to adapt them to minimal or isotope-enriched media have not been described. This shortcoming exists because pSer genetic code expansion expression hosts require the genomic *ΔserB* mutation, which increases pSer bioavailability but also imposes serine auxotrophy, preventing growth in minimal media used for isotopic labeling of recombinant proteins. Here, by testing different media supplements, we restored normal BL21(DE3) *ΔserB* growth in labeling media but subsequently observed an increase of phosphatase activity and mis-incorporation not typically seen in standard rich media. After rounds of optimization and adaption of a high-density culture protocol, we were able to obtain ≥10 mg/L homogenously labeled, phosphorylated superfolder GFP. To demonstrate the utility of this method, we also produced the intrinsically disordered serine/arginine-rich region of the SARS-CoV-2 Nucleocapsid protein labeled with ¹⁵N and pSer at the key site S188 and observed the resulting peak shift due to phosphorylation by 2D and 3D heteronuclear single quantum correlation analyses. We propose this cost-effective methodology will pave the way for more routine access to pSer-enriched proteins for 2D and 3D NMR analyses.

Phosphorylation, the most common type of post-translational modification, is an essential protein regulatory mechanism in eukaryotic cells (1–3). Dysregulation of phosphorylation-dependent signaling systems is linked to numerous disease pathologies as these post-translational modifications play key roles in cellular processes such as protein synthesis, signal

transduction, and cell development (4–6). More than two-thirds of proteins in the human proteome undergo reversible phosphorylation and of these proteins, nearly 80% were identified at serine residues (7). Most phosphoserine (pSer) sites are located within flexible or disordered regions of proteins (8, 9), where they serve as regulatory switches to modulate conformational dynamics, function, and allosteric interactions (10–13). NMR spectroscopy is the ideal technique for probing dynamics and structure of intrinsically disordered regions (IDRs)/intrinsically disordered proteins at the molecular level (14). Despite the ubiquitous nature of phosphorylation, few NMR studies focus on understanding the molecular consequences of IDR/intrinsically disordered protein phosphorylation. This is, in large part, caused by a lack of standardized and routine methods for synthesizing isotopically labeled, site-specifically phosphorylated proteins for NMR characterization.

The study of phospho-proteins by NMR requires that they be isotopically enriched with ¹⁵N and/or ¹³C and be homogenously phosphorylated at the targeted site(s). Standard expression strains of *Escherichia coli* (e.g., BL21(DE3)) can biosynthesize all 20 natural amino acids from fundamental carbon and nitrogen building blocks so that protein expression in fully-defined minimal media containing ¹³C-sugars and ¹⁵NH₄Cl provides a convenient strategy to produce isotopically labeled proteins. Site-specific phosphorylation of target proteins, on the other hand, continues to pose a challenge because the required kinase(s) is not always known, they may lack required specificity and they may not be easily isolated in a functional state for *in vitro* utility (14–16). Phosphomimetic mutations (Ser or Thr to Asp/Glu) are commonly introduced at the site of phosphorylation to overcome this issue, however, Asp or Glu do not faithfully recapitulate the geometry nor the charge density (17–19) and so they commonly misinform on the functional effects of authentic phosphorylation (20–25).

Genetic code expansion (GCE) has emerged as a leading technology for the production of phosphorylated proteins because it allows site-specific, homogenous, and efficient translational incorporation of phosphorylated amino acids into any protein (Fig. 1A) (26–29). In 2015, a high-efficiency GCE system was developed to translationally install pSer in response to an amber (TAG) stop codon, yet to date, this

* For correspondence: Richard B. Cooley, rick.cooley@oregonstate.edu.

Methods to access isotope-enriched phosphorylated proteins

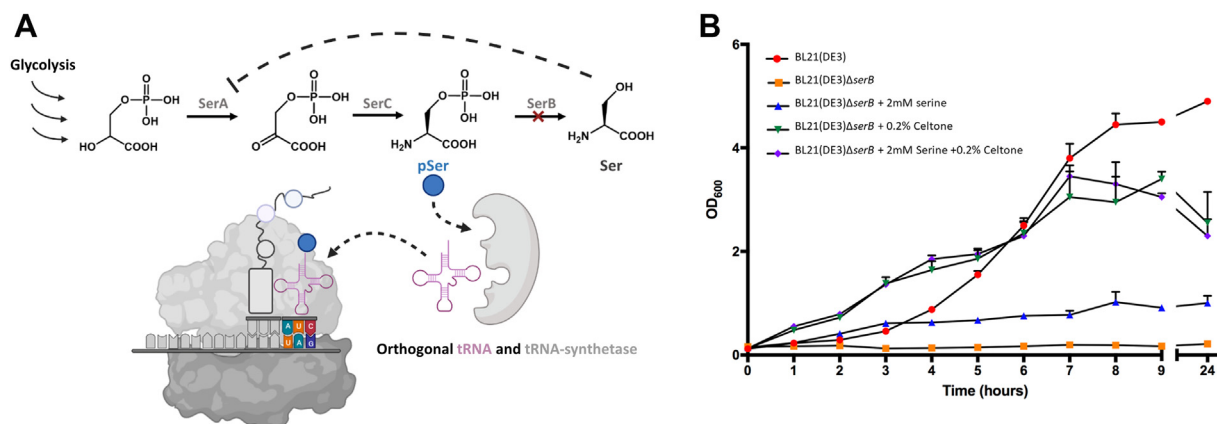


Figure 1. *Escherichia coli* Δ serB expression strains growth in minimal media with key supplements. A, deletion of *serB* allows phosphoserine concentrations to build up inside the cell for sufficient translational incorporation at programmed amber codon(s) with GCE-derived pSer machinery. However, serine feedback inhibition of SerA prevents build-up of pSer inside the cell. B, growth rate comparison of expression strains BL21(DE3) (red) and BL21(DE3) Δ serB (orange) with serine (blue), Celtone (green), or both (purple) over 24 h. Error bars represent SDs of cultures performed in duplicate. GCE, genetic code expansion; pSer, phosphoserine.

system has not been adopted to produce isotopically labeled phosphorylated proteins (26). Reasons for this have not been articulated in the literature, but we hypothesized they were rooted in the fact that pSer GCE expression systems utilize serine auxotrophic strains for phospho-protein expression, which means they cannot grow in minimal media (26, 27, 29, 30). Serine auxotroph expression hosts are used because charged, phosphorylated amino acids like pSer do not traverse from the media into the cell effectively, resulting in low bioavailability of the free amino phospho-amino acid (26, 27, 29–31). By deleting the *serB* gene, which hydrolyzes pSer to produce serine as the last step of serine biosynthesis, pSer accumulates inside the cell, providing a sufficient pool of pSer to feed the GCE machinery (31) (Fig. 1A). Serine supplementation should overcome this issue of auxotrophy, though isotopically labeled serine is expensive and serine feedback inhibition of SerA shuts down production of pSer (Fig. 1A) (32–34). Intracellular levels of pSer in WT *E. coli* strains are remarkably low having an intact SerB (31), and even if pSer media supplementation could overcome this issue, isotopically labeled pSer is not commercially available to our knowledge.

Needed therefore is an expression methodology that merges existing pSer GCE systems with methods for expressing isotopically labeled proteins. Here, we overcome these challenges to formulate an efficient and low-cost expression strategy that is sufficient for production of ¹⁵N-labeled proteins with site-specific pSer incorporated using serine auxotroph BL21(DE3) Δ serB as the expression host. This method can be easily adapted for ¹³C labeling and should accelerate access to phosphorylated proteins for NMR structural biology projects.

Results

Optimization of BL21(DE3) Δ serB growth in labeling media

We use here the *E. coli* expression host BL21(DE3) Δ serB, with the pSer phosphatase *serB* deleted to increase the intracellular levels of pSer (31) and thus, improve pSer

incorporation into recombinant isotopically labeled protein (26, 28). Our first goal was to compare the growth rate between BL21(DE3) WT and BL21(DE3) Δ serB hosts in unlabeled minimal media without antibiotics and identify key supplements required for optimal BL21(DE3) Δ serB growth (Fig. 1B). We tracked optical density (OD) at 600 nm (OD₆₀₀) over 24 h and confirmed WT BL21(DE3) cells grew robustly in the minimal media, while BL21(DE3) Δ serB would not grow unless supplemented with 2 mM serine, 0.2% (w/v) Celtone base powder, or both (Fig. 1B). Celtone base powder (referred to hereafter as Celtone) is an algal hydrolysate containing a mix of amino acids and used here because it is available with ¹³C and/or ¹⁵N enrichment. Interestingly, serine supplementation alone to BL21(DE3) Δ serB did not fully rescue growth phenotype to those of WT BL21(DE3), as indicated by an impaired growth rate and a low final cell density (final OD₆₀₀ ~1 versus ~5, respectively). On the other hand, the addition of Celtone alone or in combination with serine improved growth rates and final cell densities that were comparable growth to BL21(DE3) WT. With these data, we reasoned that a minimal media supplemented with either both 2 mM serine and 0.2% (w/v) Celtone, or just Celtone alone, could be used for phospho-protein production with further expression optimization.

Optimization of expression conditions for phosphorylated protein production

Having identified media for robust BL21(DE3) Δ serB growth, we set out to evaluate pSer protein expression using the high efficiency GCE system created by Chin *et al.* (26). This system utilizes the pKW2-EFsep plasmid which expresses the pSer tRNA synthetase variant SepRS-2, the amber codon suppression tRNA Sep-tRNA_{CUA}(B4), and an EFTu variant enabling efficient delivery of pSer-aminoacylated tRNA to the ribosome, EF-Sep. To easily evaluate phospho-protein expression parameters, we expressed the fluorescent reporter super folder GFP (sfGFP) containing an amber TAG codon at

Table 1**Components for 50 mL of ZY non-inducing (ZY-NIM) and ZY auto-inducing media (ZY-AIM)**

Component	For 50 ml of ZY-NIM	For 50 ml of ZY-AIM
ZY media ^a	45 ml	45 ml
1 M MgSO ₄ ^b	0.1 ml	0.1 ml
25x M-Salts-1 ^c	2.0 ml	2.0 ml
25 g/L NH ₄ Cl ^b	2.0 ml	2.0 ml
40% (w/v) Glucose ^d	0.625 ml	-
50x 5052 ^e	-	1.0 ml
5000x Trace Metals ^f	10 µl	10 µl

Components were sterilized separately and combined as follows after they are cooled to room temperature and immediately before use.

^a ZY Media, per liter: 10 g Tryptone, 5 g Yeast Extract. Autoclave.

^b Unlabeled. Sterile by autoclaving or filtering.

^c 25x M-Salts-1: 0.625 M Na₂HPO₄, 0.625 M KH₂PO₄, 0.125 M Na₂SO₄. No pH adjustment. Autoclave.

^d 40% (w/v) glucose: 20 g α-D-glucose in total volume of 50 mL. Autoclave or sterile filter.

^e 50x 5052, per 500 mL: 12.5 g α-D-glucose, 50 g α-lactose, 125 mL glycerol (or 250 mL of 50% glycerol, which is easier to measure). Note: to get the lactose to dissolve, heat gently in microwave. Once dissolved, lactose will stay in solution indefinitely.

^f Not essential for ZY-NIM, but can help cell growth, as previously described (28).

position N150 (sfGFP-150TAG) from the pRBC plasmid (28). Previously, we found that rich auto-induction media (called ZY-AIM, Table 1) provided maximal levels of homogeneously phosphorylated protein with this pSer GCE system (28), and so here, we first tested auto-induction expression strategies with Celtone- and serine-supplemented media amenable to isotopic labeling (Minimal AIM, Table 2, Method 1 in Fig. 2). In parallel, we also tested manual induction expression strategies in a minimal media similarly supplemented with Celtone and serine (MIM-1, Table 2, Method 2 in Fig. 2). For these expressions, starter cultures were grown overnight (~16–18 h) in a rich non-inducing media (ZY-NIM, Table 1). Cells were then pelleted, resuspended in their respective expression minimal media. Protein expression yields were quantified by whole-cell fluorescence. The homogeneity of phosphorylation was assessed by Phos-tag gel electrophoresis of purified sfGFP proteins. With Phos-tag gels, phosphorylated proteins migrate slower than nonphosphorylated proteins permitting near-quantitative evaluation of protein phosphorylation status (35). For comparison, phosphorylated sfGFP was also expressed in rich ZY-AIM (28).

All expression cultures grown at 37 °C grew to a similar OD₆₀₀ of ~4 to 5 and with total cell fluorescence about 30% higher in the auto-induction expression, corresponding to 1 to 2 mg of protein per liter culture (Fig. 3A). However, only ~40% of purified sfGFP expressed from auto-induction and ~60% from manual-induction were phosphorylated, in stark contrast to >95% of the protein being phosphorylated when expressed in and purified from rich ZY-AIM media (Fig. 3C lanes 1, 7 and 5 respectively). We assumed the source of nonphosphorylated protein could come from the following: (1) near-cognate suppression by endogenous tRNAs at encoded amber codons, (2) protein dephosphorylation from phosphatase activity, and (3) mis-aminoacylation of the Sep-tRNA_{CUA}(B4) by endogenous aminoacyl-tRNA synthetases. Near cognate suppression of amber stop codons is generally insignificant when using expression hosts such as BL21(DE3) Δ*serB* that contain Release Factor 1 (RF1), the *E. coli* protein responsible for terminating translation at amber codons (28, 36). We next considered the possibility that the our GCE tRNA, Sep-tRNA_{CUA}(B4), was mis-aminoacylated by endogenous synthetases in these

conditions where overall incorporation efficiency was low (27). We addressed this by replacing Sep-tRNA_{CUA}(B4) with the Sep-tRNA_{CUA}^{v2} (Fig. 3B), a variant containing mutations in its acceptor stem that minimize mis-acylation by endogenous synthetases (27). Using this strategy, we re-expressed sfGFP-150TAG in the same auto- and manual-induction media at 37 °C. While expression yields were comparable with Sep-tRNA_{CUA}^{v2}, measurable improvements in incorporation fidelity were observed, with ~80% of the purified sfGFP being phosphorylated (Fig. 3C, lanes 2 and 8). The remaining non-phosphorylated protein was identified by mass spectrometry to be exactly 80 Da lighter, consistent with hydrolysis of pSer by cellular phosphatases during expression (Fig. 3D). We found that by conducting the expressions at 18 °C instead of 37 °C, cellular phosphatase activity was notably decreased, with >90% of the purified sfGFP being phosphorylated using Methods 1 and 2 with the Sep-tRNA_{CUA}v2 (Fig. 3B, lanes 4 and 10). Having improved homogeneity of phospho-protein expression, we next set out to optimize expression yield.

Optimizing phospho-protein yield for NMR applications

To this point, all sfGFP protein expressions were cultured in minimal media supplemented with 2 mM serine and 0.2% (w/v) Celtone, so we hypothesized that removing serine would improve protein yield without compromising protein phosphorylation by removing feedback inhibition of SerA (Fig. 1A) (32–34). For auto-induction expression (Method 1), we observed no notable defect in cell growth (Fig. 4A) nor increase in protein production when exogenous serine was omitted (Fig. 4B). On the other hand, with manual induction (Method 2), cell growth was similar (Fig. 4A) but phosphorylated protein yield increased 2-fold (Fig. 4B) compared to the same cultures containing serine and ~30% higher than auto-induction expression (Method 1). In all cases, >90% of the purified protein was phosphorylated (Fig. 4C). Given the expense of isotopically labeled (¹⁵N and/or ¹³C) serine and that its omission does not have adverse side effects, serine was left out of subsequent expressions.

Methods to access isotope-enriched phosphorylated proteins

Table 2
Components for 50 mL of minimal auto-induction and minimal manual-induction media

Component	For 50 ml		
	Minimal AIM	Minimal MIM-1 (low density)	Minimal MIM-2 (high density)
Sterile Water	45 ml	45 ml	45 ml
1 M MgSO ₄	0.1 ml	0.1 ml	0.1 ml
25x M-Salts-1	2.0 ml	2.0 ml	-
25x M-Salts-2 (high pH) ^a	-	-	2.0 ml
25 g/L NH ₄ Cl OR ¹⁵ NH ₄ Cl	2.0 ml	2.0 ml	2.0 ml
40% (w/v) Glucose (Unlabeled OR ¹³ C)	-	0.625 ml	1.25 ml
50x 5052	1.0 ml	-	-
10% (w/v) CELTONE (Unlabeled, ¹⁵ N OR ¹⁵ N/ ¹³ C) ^b	1.0 ml	1.0 ml	1.0 ml
5000x Trace Metals	10 µl	10 µl	10 µl

Components were sterilized separately and combined as follows after they cooled to room temperature and immediately before use.

^a 25x M-Salts 2 (high pH): Same as 25x M-salts-1 (see Table 1), pH adjusted to 8 to 8.2 with 10 M KOH. Autoclave.

^b Dissolve Celtone powder in sterile water to 10% (w/v) final concentration. Do not autoclave.

For these expressions, cells were grown to mid-log phase from an overnight non-inducing culture, at which time protein expression was induced and the final OD₆₀₀ reached ~3 to 5. To maximize protein yields per liter culture, we tested methods in which protein expression is performed at higher cell densities (Method 3 and 4 in Fig. 2) (37). We first tested a strategy in which cells were grown to stationary phase in ZY-NIM media overnight, resuspended in a Celtone supplemented minimal media, grown for a short period, and then induced manually at an OD₆₀₀ ~5 to 10 (Method 3 in Fig. 2). While the final OD₆₀₀ at harvest was substantially higher than low-density expressions (Methods 1 and 2), total protein culture fluorescence was only modestly improved (Method 3, Fig. 4B), indicating protein production per cell was compromised, perhaps because the cells had not recovered from being in stationary phase when induction began. Alternatively, we inoculated fresh ZY-NIM media with the overnight culture, allowed it to grow to mid/late-log phase (OD₆₀₀ ~3–4) at which

time the freshly grown cells were pelleted and resuspended in a Celtone supplemented minimal media and grown for a short period prior to IPTG induction (Method 4 in Fig. 2, adapted from (37)). The final cell density reached ~8 to 12 like Method 3, however, overall protein production was >3-fold improved compared to the low-density cultures (Fig. 4B). The fluorescence values of these cultures correspond to approximately ~8 to 10 mg of sfGFP-pSer150 per liter culture. Accurate incorporation of pSer was confirmed by Phos-tag gels and whole protein mass spectrometry (Fig. 4, C and D). Thus, we selected Method 4 for future isotopically labeled expressions.

Production of ¹⁵N sfGFP-pSer150 and ¹H-¹⁵N HSQC spectra

Using Method 4, we produced ¹⁵N-labeled, >90% phosphorylated sfGFP at site N150, as well as WT sfGFP, to confirm uniform isotopic enrichment and subsequent utility for NMR

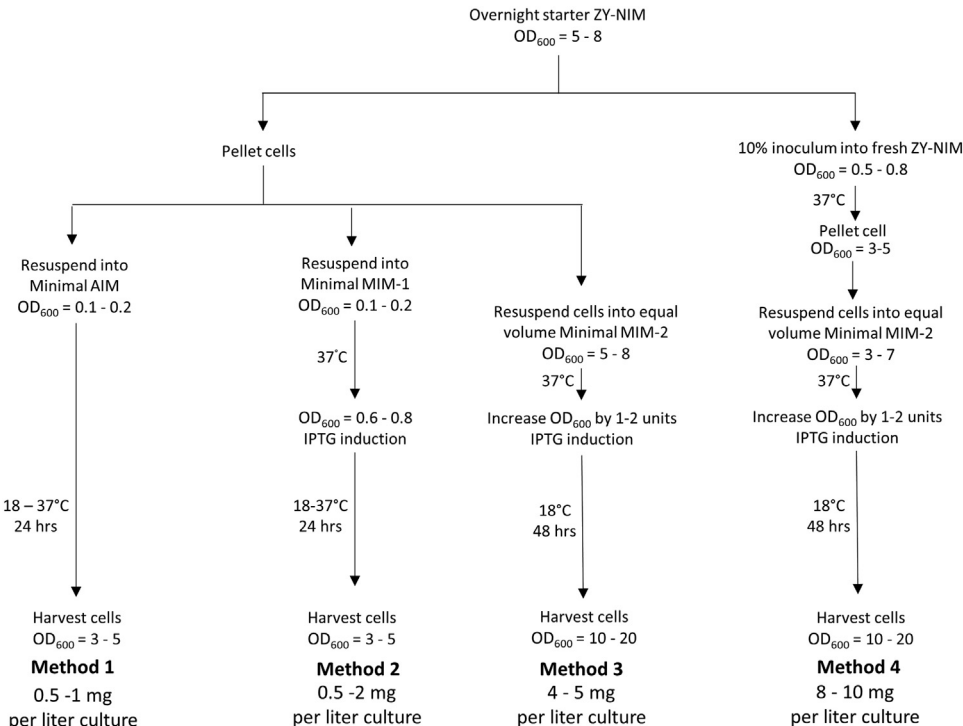
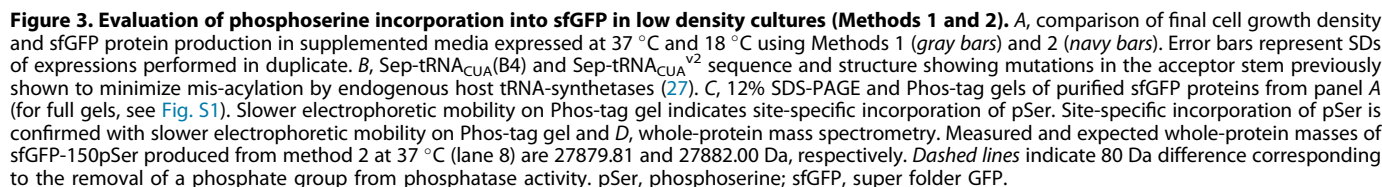


Figure 2. Schematic of the four primary experimental expression strategies evaluated in this study.



We did not observe a clear down-field shift of the site 150 resonance in response to phosphorylation, however, which would be expected if a hydrogen bond was formed between the phosphate group and amide backbone (40). But being located within a β -sheet, the backbone amide of residue 150 is not expected to form such a hydrogen bond while also maintaining the native fold (41) (Fig. 5C). Despite lacking a clear assignment of the pSer150 cross peak, Phos-tag gel electrophoresis confirmed >90% phosphorylation both before and after HSQC data collection at 42 °C (Fig. 5A) (38, 39).

As a biologically relevant example, we produced the Ser/Arg-rich IDR of the SARS-CoV-2 Nucleocapsid (N) protein

We expressed the SR-Linker region of N as a fusion with cleavable tags at its N- and C-terminus to improve solubility (*bds*SUMO and TEV-sfGFP-His6, respectively) first using Method 4 in unlabeled media. Purification and removal of the *bds*SUMO and sfGFP-His6 tags yielded pure SR-Linker with ~80 to 90% phosphorylation at site S188 as confirmed by Phos-tag electrophoresis and whole-protein mass spectrometry (Fig. 6, B and C). Having confirmed production of SR-Linker in sufficient quality, we repeated the expression in an isotopically enriched culture medium. ^1H - ^{15}N HSQC spectra at 10 °C of the purified proteins are characteristic of an IDR where signals cluster in the 8.0 to 8.5 ppm region due to poor dispersion in ^1H spectra and matched well with previously

Methods to access isotope-enriched phosphorylated proteins

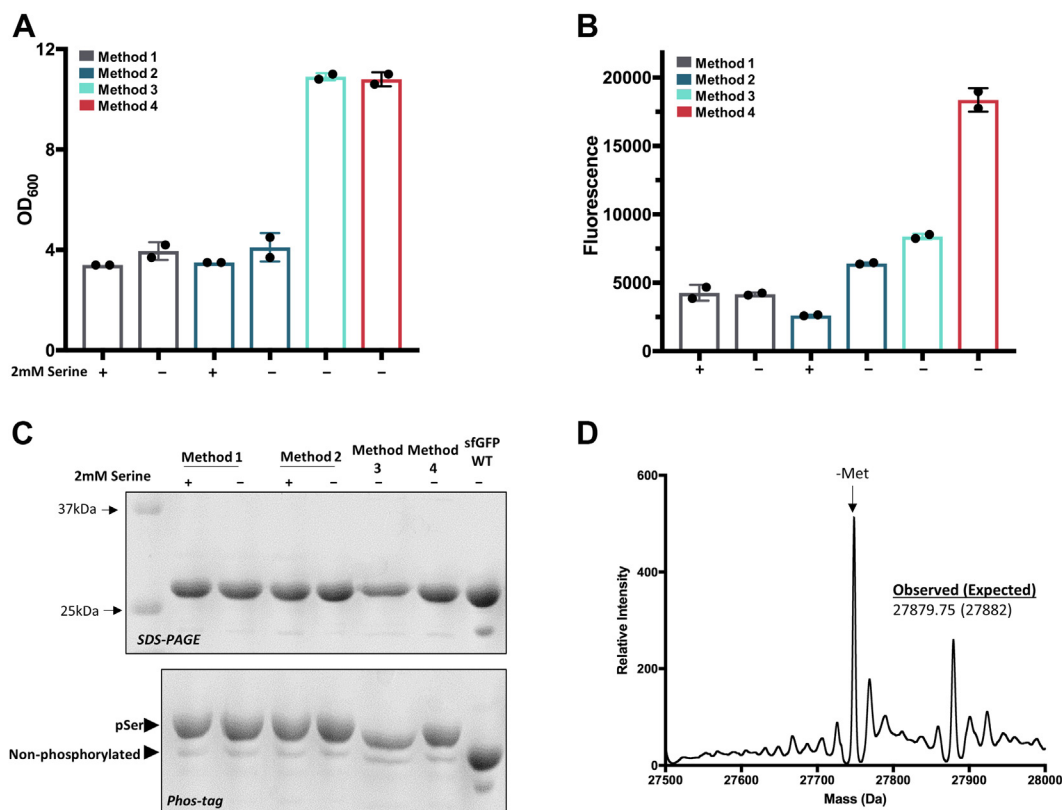


Figure 4. Improving protein production with high-density culture protocols. Final cell densities (panel A) and total culture fluorescence (panel B) of cultures expressing sfGFP-150pSer show that omission of free serine amino acid improved protein production for Method 2 (manual induction, low density). Expression at high-density further increased phospho-protein yield by ~3-fold (Method 4). Error bars represent SDs of expressions performed in duplicate. C, purified sfGFP-150pSer proteins were >90% phosphorylated from all expressions (for full gels, see Fig. S1) and D, whole-protein mass spectrometry of sfGFP-150pSer produced with Method 4 further confirms homogenous phosphorylation. sfGFP, super folder GFP.

published spectrum of a similar SR-Linker N construct (Fig. 6D) (45). In these spectra, a clear ^1H - ^{15}N peak is observed at 9.1 ppm/119.6 ppm in the pSer188 protein sample that is not present in the WT sample, consistent with where a phosphorylated serine amide peak would be expected (Fig. 6D) (46, 47). To further confirm the assignment of the resonance at

9.1 ppm to pSer188, we collected a 3D- ^{15}N -nuclear overhauser effect spectroscopy (NOESY)-HSQC (Fig. 6E). We observed sequential backbone amide nuclear overhauser effects for residues ± 2 from the pSer188 (Fig. 6E). These residues could be assigned based on the side chain proton chemical shifts, which were consistent with the expected amino acid sequence

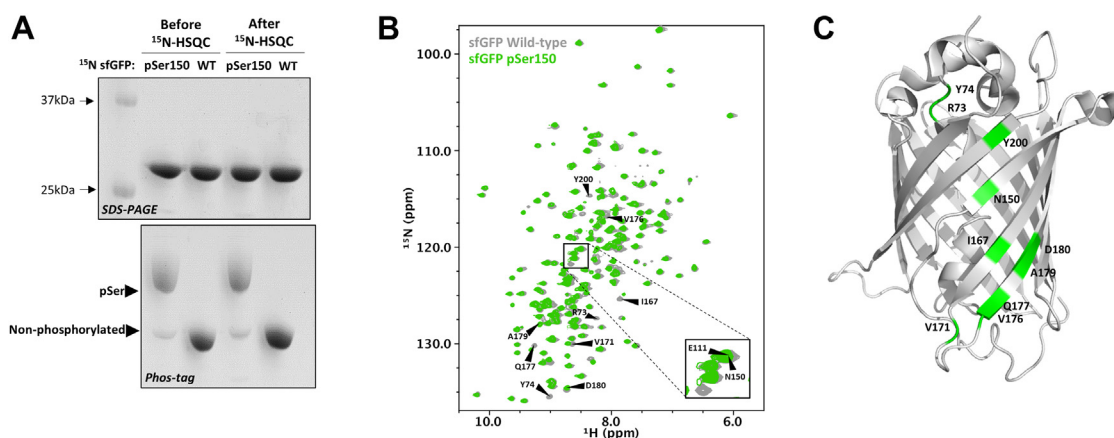


Figure 5. Production and ^1H - ^{15}N HSQC analysis of ^{15}N -sfGFP proteins. A, SDS-PAGE and Phos-tag gel electrophoresis of purified ^{15}N -sfGFP proteins before and after data collection demonstrate stability of phosphorylation even at 42 °C (for full gels, see Fig. S1) (38, 39). B, ^1H - ^{15}N -HSQC spectra of wild-type sfGFP (gray) and 150pSer (green) in panel A were visualized and overlaid in NMRViewJ. Inset shows zoom in of peaks corresponding to E111 and N150. Resonances with chemical shift perturbation in 150pSer HSQC spectrum are labeled and C, highlighted on wild-type sfGFP 3D structure (highlighted green). sfGFP, super folder GFP.

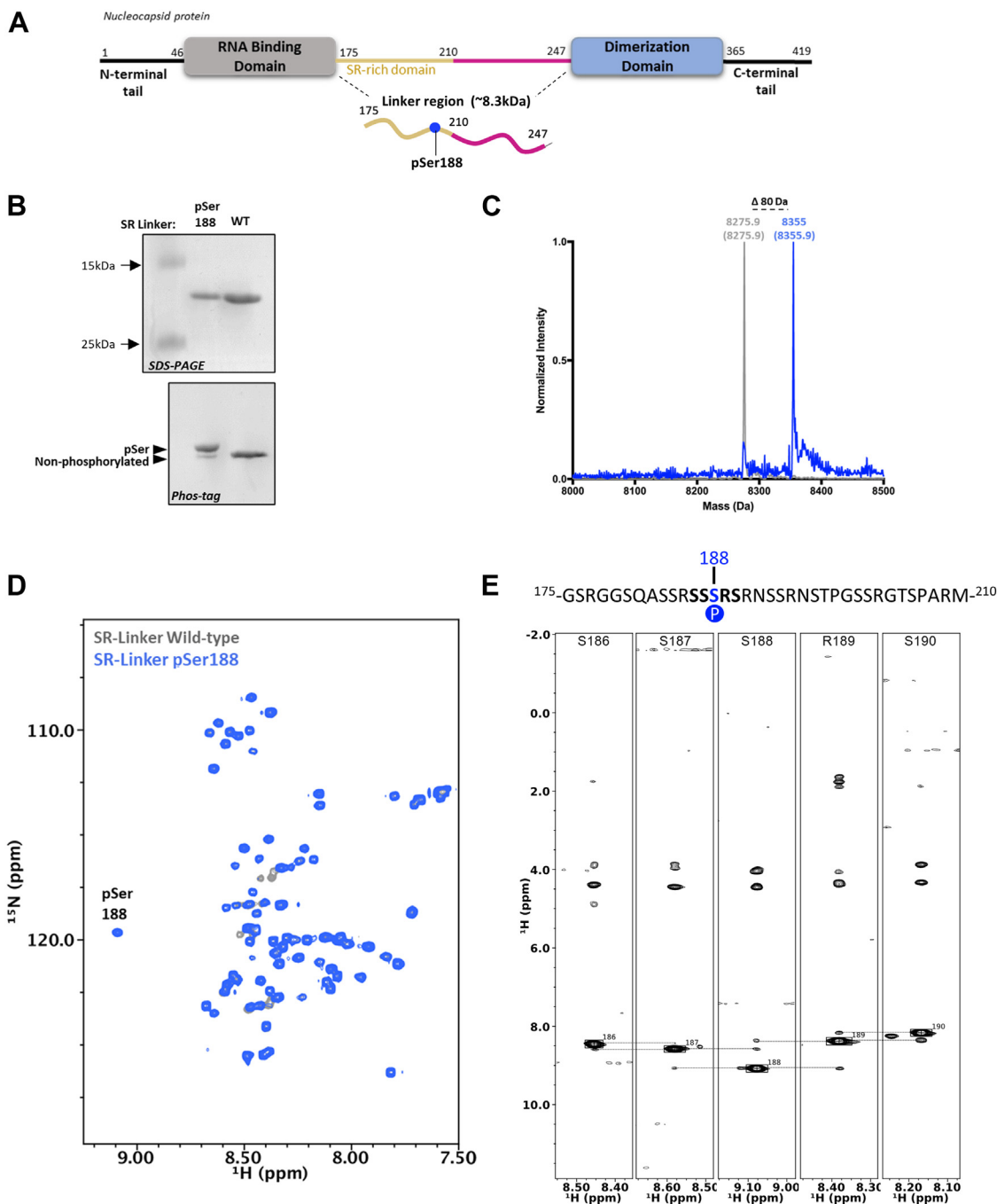


Figure 6. Analysis of the SR Linker region of the N protein phosphorylated at S188. A, the linker region of SARS-CoV-2 N phosphoprotein contains a serine/arginine (SR) rich region that is phosphorylated in virions. This SR-Linker (175–247) is genetically fused to a TEV cleavable C-terminal sfGFP protein. B, purified unlabeled SR-Linker 188pSer are ~80 to 90% phosphorylated as confirmed by 15% SDS-PAGE, Phos-tag gel electrophoresis (for full gels, see Fig. S1) and C, whole-protein mass spectrometry. Dashed line indicates +80 Da increments from wild-type SR-Linker (gray) to 188pSer (blue). D, ^1H - ^{15}N HSQC spectra of wild-type SR-Linker (gray) and 188pSer (blue) is typical of an intrinsically disordered protein with pSer 188 chemical shift ^1H - ^{15}N at 9.1 ppm and 119.6 ppm. E, amino acid sequence of the Ser/Arg-rich intrinsically disordered region of SARS-CoV-2 with pSer188 highlighted in blue (top). The sequential assignment of backbone amide NOE's for residues 186 to 190 visualized in SR-Linker pSer 188 3D- ^{15}N -NOESY-HSQC strip plot matches with amino acid sequence of SR-Linker (bottom). This confirms 9.1 ppm/119.6 ppm cross-peak observed in D belongs to pSer188. HSQC, heteronuclear single quantum coherence; NOE, nuclear overhauser effect; NOESY, nuclear overhauser effect spectroscopy; pSer, phosphoserine.

(Fig. 6E). Collectively, these data demonstrate the facile ability to generate sufficient quantities of a biologically relevant, phosphorylated IDR for NMR applications. Detailed structural and dynamics analyses of these proteins will be described in a subsequent article.

Discussion

Here, we have described an optimized protocol for the generation of site-specifically phosphorylated, isotopically enriched proteins suitable for NMR analysis. The adoption of GCE pSer expression system imposed several challenges,

Methods to access isotope-enriched phosphorylated proteins

including how to overcome growth deficiencies of the serine auxotroph BL21(DE3) $\Delta serB$ in minimal media, as well as enhanced phosphatase activity not typically seen when expressing phospho-proteins in standard rich media. Through rounds of optimization, we found three important parameters that facilitated homogenous phospho-protein production in labeling media: (i) Celtone as a key additive needed for healthy cell growth to overcome serine auxotrophy, (ii) Sep-tRNA-^{CUA}^{v2} for reduced mis-acylation by endogenous synthetases to minimize mis-incorporation of natural amino acids, and (iii) protein expression at lower temperatures to minimize phosphatase activities. Subsequent adaption of high-density methods allowed us to produce ≥ 10 mg purified sfGFP per liter culture with $>90\%$ phosphorylation, costing approximately $\sim \$300$ per liter of uniformly ^{15}N -labeled proteins. Although not demonstrated it here, this method can be adapted for ^{13}C labeling as well by using ^{13}C -labeled Celtone and glucose.

The ability to produce isotopically labeled phospho-proteins with GCE has been reported previously, but only in three instances to date. Two of these reports used the BL21(DE3) $\Delta serB$ strain but their methods did not describe a mechanism to overcome serine auxotrophy, and they also added unlabeled pSer amino acid to the culture media so that the expressed protein presumably has a mixture of labeled and unlabeled pSer (48, 49). In a third instance, recent work by Scheffner *et al.* (50) circumvented serine auxotrophy by using BL21(DE3) with intact SerB, and so the expression required the addition of high concentrations of unlabeled pSer to drive incorporation and downstream work benefited from the ability to purify away unphosphorylated populations by anion exchange chromatography, which may not always be possible. In the methodologies reported here, all pSer incorporated into the protein is biosynthesized from isotopically enriched media components, and thus, all pSer residues are expected to be isotopically labeled and will be visible in NMR spectra as demonstrated with ^{15}N -HSQC data of phosphorylated SR Linker at site 188 (Fig. 6D).

The methods described here were optimized to express phosphorylated sfGFP and SR-Linker proteins, but we anticipate different proteins will require additional adjustments in expression protocols for specific applications. We found Celtone used at relatively low concentrations (0.2% (w/v)) to be an economically viable supplement to support phospho-protein expression; however, other commercially available isotopically labeled media additives such as BioExpress (Cambridge Isotope Laboratories, Inc), SILEX Media (Silantes), and ISO-GRO (Sigma-Aldrich) may also be tractable. We believe that multi-site pSer incorporation should be possible, though feasibility will depend on the protein of interest and sites of incorporation. Adoption of other $\Delta serB$ strains of *E. coli*, such as the RF1-deficient B95(DE3) $\Delta A \Delta fabR \Delta serB$ (28), may prove helpful in this regard. However, our attempts to express sfGFP-150pSer in this strain using the methods reported here resulted in undesirable quantities of near-cognate suppression at the intended site of phosphorylation (data not shown), and so additional optimizations will be required for expressing

isotopically labeled pSer proteins in this “truncation free” expression host. Nevertheless, the work here provides an important framework by which isotopically labeled, site-specific pSer-containing proteins can be expressed efficiently in *E. coli* and opens the door to more routine analyses of phosphorylated proteins with two-dimensional and three-dimensional NMR experiments.

Experimental procedures

Strains and plasmids

The BL21(DE3) $\Delta serB$ strain was a gift from Jesse Rinehart (Addgene #34929). BL21(DE3) and DH10b strains of *E. coli* were purchased from ThermoFisher Scientific. The pRBC-sfGFP WT and pRBC-sfGFP-150TAG plasmids were as previously described (Addgene #174075 and 174076, respectively) (28) (Fig. S2). The pKW2-EFsep was a generous gift from Jason Chin (Addgene # 173897) (Fig. S2). Genes for *bdSUMO*, *bdSENP1*, Sep-tRNA^{v2}, and SARS-CoV-2 SR-Linker were synthesized by Integrated DNA Technologies. The SARS-CoV-2 SR-Linker protein (residues 175–24) expression plasmids were made by fusing a *bdSUMO* fusion protein (lacking the first 19 residues) at its N-terminus (51) and a TEV cleavable sfGFP-His₆ at its C-terminus. All cloning steps were performed by SLiCE (52). The PPY strain used to generate SLiCE cloning extract was a gift from Yongwei Zhang (Albert Einstein College of Medicine) (52).

Molecular biology reagents

Oligonucleotide primers were synthesized by Integrated DNA Technologies. Molecular biology reagents including restriction enzymes, T4 ligase, and polymerases were purchased either from Thermo Fisher Scientific or New England Biolabs. DNA Miniprep, Midiprep, PCR cleanup, and gel extraction kits were purchased from Machery Nagel. L-Serine, Celtone base powder (#1030P-U and 1030-N for unlabeled and ^{15}N labeled, respectively), and 100X MEM Vitamin were purchased from Sigma-Aldrich, Cambridge Isotope Laboratories, Inc, and Thermo Fisher Scientific, respectively. Phos-tag Acrylamide for gel electrophoresis was purchased from NARD Institute, Ltd.

Cell growth assessment

BL21(DE3) and BL21(DE3) $\Delta serB$ cells were streaked on LB/agar without antibiotics and grown overnight at 37 °C. A single colony was used to inoculate a buffered, glucose-rich ZY-NIM (Table 1) overnight with shaking at 250 rpm. The overnight starter was diluted to a starting optical density (OD₆₀₀) of 0.15 into 50 ml of fresh minimal manual-inducing media (NIM-1) containing no additives, 2 mM serine, 0.2% Celtone, or both. Celtone was prepared by resuspending base powder with sterile water to a final concentration of 0.2% (w/v). Cells were grown for 24 h with OD₆₀₀ measurements taken every hour for 9 h and once after 24 h. Cultures were grown in duplicate.

Protein expression in minimal media

Fresh transformations were performed for all expressions in this study. BL21(DE3) $\Delta serB$ cells were cotransformed with either pKW2-EFsep containing either Sep-tRNA_{CUA}(B4) or Sep-tRNA_{CUA}^{v2} and the appropriate pRBC plasmid. Approximately, a dozen colonies were used to inoculate overnight ZY-NIM (Table 1) and grown at 37 °C. All ZY-NIM cultures contained 100 µg/ml ampicillin and 25 µg/ml chloramphenicol, while all minimal cultures contained 50 µg/ml ampicillin and 15 µg/ml chloramphenicol. Where indicated, L-serine was added at 2 mM final concentration and Celtone at 0.2% (w/v). To produce ¹⁵N-labeled proteins, NH₄Cl was replaced with ¹⁵NH₄Cl and Celtone replaced with ¹⁵N-Celtone.

Method 1: Minimal auto-induction media

Overnight ZY-NIM cells (OD₆₀₀ ~5–8) were pelleted by centrifugation at 5000 rcf and then resuspended into minimal AIM (Table 2) (starting OD₆₀₀ ~0.1–2). Cultures were grown at either 37 °C or 18 °C with shaking at 250 rpm in baffled flasks and harvested 24 or 48 h later, respectively.

Method 2: Minimal manual induction media, low density

Overnight ZY-NIM cells (OD₆₀₀ ~5–8) were pelleted by centrifugation at 5000 rcf and then resuspended into minimal manual induction media 1 (MIM-1, Table 2) (starting OD₆₀₀ ~0.1–0.2). Cultures were grown at 37 °C until OD₆₀₀ reached ~0.6 to 0.8 and then induced with 1 mM IPTG. Cultures were grown at either 37 °C or 18 °C and harvested 24 or 48 h after IPTG addition, respectively.

Method 3: Minimal MIM, high density

ZY-NIM overnight starters (OD₆₀₀ ~5–8) were pelleted by centrifugation at 5000 rcf and resuspended into an equal volume of minimal MIM 2 (Table 2) (starting OD₆₀₀ ~5–8). Cultures were grown at 37 °C in baffled flasks until the OD₆₀₀ increased by 1 to 2 units (~1–2 h) and then induced with 1 mM IPTG. Cultures were grown at 18 °C for 48 h after IPTG addition.

Method 4: Minimal MIM, high density, with freshly grown cells

Cells from a ZY-NIM overnight starter culture (OD₆₀₀ ~5–8) were used to inoculate a fresh ZY-NIM culture (10% inoculum, e.g., 5 ml into 50 ml fresh ZY-NIM). Cultures were grown at 37 °C with shaking at 250 rpm in baffled flasks until OD₆₀₀ reached ~3 to 4, at which point cells were pelleted by centrifugation at 5000 rcf and resuspended into an equal volume of minimal MIM 2 (Table 2) (starting OD₆₀₀ ~3–4). Cultures were grown at 37 °C in baffled flasks until the OD₆₀₀ increased by 1 to 2 units (~1–2 h) and then induced with 1 mM IPTG. Cultures were grown at 18 °C for 48 h after IPTG addition.

Protein purification

sfGFP purification

Cell pellets containing sfGFP proteins were resuspended in Lysis Buffer (50 mM Tris pH 7.5, 500 mM NaCl, 5 mM

imidazole, and phosphatase inhibitors: 50 mM sodium fluoride, 5 mM sodium pyrophosphate, 1 mM sodium orthovanadate) and lysed by microfluidization. Soluble cell lysate was obtained by centrifugation at 28,000 rcf for 45 min, to which TALON metal affinity resin was added. His₆-tagged protein was allowed to bind to the TALON resin for 30 to 60 min with gentle rocking. Resin was collected and extensively washed with Lysis Buffer, and then protein was eluted with Lysis Buffer supplemented with 300 mM imidazole. After elution, proteins were further purified by gel-filtration on a 10/300 Superdex S75 column (Cytiva Life Sciences) in NMR buffer (30 mM sodium phosphate (pH 6.8), 100 mM NaCl) and then concentrated to 600 µM using a 10,000 Da cutoff filter prior to NMR analysis. SDS-PAGE and Phos-tag gels were poured immediately before use and run according to manufacturer recommendation.

SR-Linker purification

The SR-Linker of N was genetically fused to a *bdSUMO* cleavable N-terminal tag and a TEV cleavable C-terminal sfGFP protein for enhanced solubility (pRBC-bdSUMO-SR-Linker-sfGFP-His₆). Cell pellets containing SR-Linker were resuspended in Lysis Buffer (50 mM Tris pH 7.5, 500 mM NaCl, 5 mM imidazole, and phosphatase inhibitors: 50 mM sodium fluoride, 5 mM sodium pyrophosphate, 1 mM sodium orthovanadate) and lysed by microfluidization. Soluble cell lysate was obtained by centrifugation at 28,000 rcf for 45 min, to which TALON metal affinity resin and 50 nM untagged *bdSEN1* protease was added. The SR-sfGFP-His₆ protein was allowed to bind to the TALON resin as the *bdSEN1* protease cleaved the *bdSUMO* tag. Resin was collected and extensively washed with Lysis Buffer, and then SR-sfGFP-His₆ was eluted with Lysis Buffer supplemented with and 300 mM imidazole. Purified protein was buffer exchanged into 50 mM Tris, 350 mM NaCl with phosphatase inhibitors using PD-10 desalting columns (Cytiva Life Sciences). The sfGFP-His₆ tag was cleaved by TEV protease (1:20 TEV to SR) overnight at 4 °C, and the mixture was flowed through fresh TALON resin. The flow-through fraction contained SR-Linker protein while the TEV protease and sfGFP-His₆ tag bound to the resin. For NMR analysis, SR-Linker protein was dialyzed overnight at 4 °C in 50 mM sodium phosphate, 150 mM NaCl, pH 6.5 buffer (without phosphatase inhibitors) and using a 3000 Da cutoff filter and concentrated to 100 µM prior to analysis.

Quantification of sfGFP expression in cultures

Yield of sfGFP expressed per liter culture was calculated by measuring in-cell fluorescence of sfGFP and subtracting the contribution of cell auto-fluorescence (measured from the same density of cells not expressing any sfGFP construct). Fluorescence values were converted to mass of sfGFP per liter culture based on a standard curve of purified sfGFP. All values reported are the average of at least two independent replicate cultures, and error bars represent SDs.

Methods to access isotope-enriched phosphorylated proteins

Whole protein mass spectrometry

Purified sfGFP proteins were exchanged into LC-MS grade water with PD-10 desalting columns. The SR-Linker proteins were buffer exchanged into 200 mM ammonium acetate by repeated concentration and dilutions using a 3000 Da cut-off centrifugal filter units. Mass spectra were obtained with a Waters Synapt G2 Mass Spectrometer at the Mass Spectrometry Facility at Oregon State University. The deconvoluted masses were obtained by using Waters MassLynx MaxEnt1 software.

NMR analysis

NMR experiments were carried out on an 800-MHz Bruker Advance III HD NMR spectrometer equipped with a 5-mm triple resonance (HCN) cryogenic probe. Data collection for ^{15}N -sfGFP proteins was carried out at 42 °C in 30 mM sodium phosphate, 100 mM NaCl (pH 6.8) buffer at a final concentration of 0.3 mM (38), while for ^{15}N -SR Linker proteins, data were collected at 10 °C in 50 mM sodium phosphate, 150 mM NaCl, pH 6.5 buffer (53) at a final concentration of approximately 80 to 100 μM . All samples contained 10% D_2O , 1 mM sodium azide, protease inhibitor mixture (Roche Applied Science), and 0.2 mM 2–2 dimethylsilapentane-5-sulfonic acid for ^1H chemical shift referencing. All two-dimensional spectra were processed using NMRPipe (54) and visualized with NMRViewJ (55). To confirm the assignments of the pSer in the SR-Linker sample, we also collected a 3D- ^{15}N -NOESY-HSQC. The NOESY data were collected with a mixing time of 120 ms, 256 complex points in each indirect dimension, and 35% random nonuniform sampling. The three-dimensional data were processed using NMRPipe (54), SMILE (56), and visualized using NMRViewJ (55).

Data availability

All data and reagents generated during this study are available from the corresponding author (rick.cooley@oregonstate.edu) upon request.

Supporting information—This article contains supporting information (26, 28).

Acknowledgments—We are grateful for the assistance from Jeff Moore for mass spectrometry data acquisition and processing. Figures 1A and 3B were created with BioRender.com. This work was supported in part by the GCE4All Biomedical Technology Development and Dissemination Center supported by National Institute of General Medical Science grant RM1-GM144227. We acknowledge the funding from National Science Foundation EAGER MCB 2034446, as well the support of the Oregon State University NMR Facility funded in part by the National Institutes of Health, HEI Grant 1S10OD018518, and by the M. J. Murdock Charitable Trust grant #2014162.

Author contributions—C. H. V. and R. B. C. conceptualization; C. H. V. methodology; C. H. V. and P. N. R. validation; C. H. V., P. N. R., and Z. Y. investigation; C. H. V. visualization; C. H. V. and R. B. C.

writing—original draft; P. N. R., Z. Y., E. B., and R. A. M. writing—review and editing; R. A. M. and R. B. C. supervision.

Funding and additional information—This work was supported in part by the National Institute of Health [5R01GM131168-02 to R. A. M., 1S10OD020111-01 to the Oregon State University Mass Spectrometry Facility], the Medical Research Foundation at Oregon Health Sciences University [to R. B. C.], and the Collins Medical Trust [to R. B. C.]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—The authors declare that they have no conflict of interest.

Abbreviations—The abbreviations used are: AIM, auto-induction media; GCE, genetic code expansion; HSQC, heteronuclear single quantum coherence; IDR, intrinsically disordered region; MIM, manual induction media; NIM, non-inducing media; NOESY, nuclear overhauser effect spectroscopy; OP, optical density; pSer, phosphoserine; sfGFP, super folder GFP.

References

- Walsh, C. T., Garneau-Tsodikova, S., and Gatto, G. J. (2005) Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew. Chem. Int. Ed. Engl.* **44**, 7342–7372
- Humphrey, S. J., James, D. E., and Mann, M. (2015) Protein phosphorylation: a major switch mechanism for metabolic regulation. *Trends Endocrinol. Metab.* **26**, 676–687
- Ochoa, D., Jarnuczak, A. F., Viéitez, C., Gehre, M., Soucheray, M., Mateus, A., et al. (2020) The functional landscape of the human phosphoproteome. *Nat. Biotechnol.* **30**, 365–373
- Lahiry, P., Torkamani, A., Schork, N. J., and Hegele, R. A. (2010) Kinase mutations in human disease: interpreting genotype-phenotype relationships. *Nat. Rev. Genet.* **11**, 60–74
- Brognard, J., and Hunter, T. (2011) Protein kinase signaling networks in cancer. *Curr. Opin. Genet. Dev.* **21**, 4–11
- Cicenas, J., Zalyte, E., Bairoch, A., and Gaudet, P. (2018) Kinases and cancer. *Cancers (Basel)* **10**, 63
- Sharma, K., D'Souza, R. C. J., Tyanova, S., Schaab, C., Wiśniewski, J. R., Cox, J., et al. (2014) Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell Rep.* **8**, 1583–1594
- Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z., et al. (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32**, 1037–1049
- Bah, A., and Forman-Kay, J. D. (2016) Modulation of intrinsically disordered protein function by post-translational modifications. *J. Biol. Chem.* **291**, 6696–6705
- McDowell, G. S., Hindley, C. J., Lippens, G., Landrieu, I., and Philpott, A. (2014) Phosphorylation in intrinsically disordered regions regulates the activity of Neurogenin2. *BMC Biochem.* **15**, 24
- Travers, T., Shao, H., Joughin, B. A., Lauffenburger, D. A., Wells, A., and Camacho, C. J. (2015) Tandem phosphorylation within an intrinsically disordered region regulates ACTN4 function. *Sci. Signal.* **8**, ra51
- Yang, P., Mathieu, C., Kolaitis, R.-M., Mittag, T., Kim, H. J., Paul, J., et al. (2020) G3BP1 is a tunable switch that triggers phase separation to assemble stress granules. *Cell* **181**, 325–345
- Ju, X., Zhu, Y., Wang, Y., Li, J., Zhang, J., Gong, M., et al. (2021) A novel cell culture system modeling the SARS-CoV-2 life cycle. *PLoS Pathog.* **17**, e1009439
- Canman, C. E., Lim, D. S., Cimprich, K. A., Taya, Y., Tamai, K., Saka-guchi, K., et al. (1998) Activation of the ATM kinase by ionizing radiation and phosphorylation of p53. *Science* **281**, 1677–1679
- Zhu, G., Liu, Y., and Shaw, S. (2004) Protein kinase specificity: a strategic collaboration between kinase peptide specificity and substrate recruitment. *Cell Cycle* **4**, 52–56

16. del Peso, L., González-García, M., Page, C., Herrera, R., and Nuñez, G. (1997) Interleukin-3-induced phosphorylation of BAD through the protein kinase Akt. *Science* **278**, 687–689
17. Chen, Z., and Cole, P. (2015) Synthetic approaches to protein phosphorylation. *Curr. Opin. Chem. Biol.* **28**, 115–122
18. Hunter, T. (2012) Why nature chose phosphate to modify proteins. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 2513–2516
19. Cerulli, R., and Krizter, J. (2020) Phosphotyrosine isosteres: past, present and future. *Org. Biomol. Chem.* **18**, 583–605
20. Kampourakis, T., Ponnamp, S., Sun, Y. B., Sevrieva, I., and Irving, M. (2018) Structural and functional effects of myosin-binding protein-C phosphorylation in heart muscle are not mimicked by serine-to-aspartate substitutions. *J. Biol. Chem.* **293**, 14270–14275
21. Haj-Yahya, M., Gopinath, P., Rajasekhar, K., Mirbaha, H., Diamond, M. I., and Lashuel, H. A. (2020) Site-specific hyperphosphorylation inhibits, rather than promotes, tau fibrillization, seeding capacity, and its microtubule binding. *Angew. Chem. Int. Ed. Engl.* **59**, 4059–4067
22. Kast, D., and Dominguez, R. (2019) Mechanism of IRSp53 inhibition by 14-3-3. *Nat. Commun.* **10**, 483
23. Somale, D., di Nardo, G., di Blasio, L., Puliafito, A., Vara-Messler, M., Chia-verina, G., *et al.* (2020) Activation of RSK by phosphomimetic substitution in the activation loop is prevented by structural constraints. *Sci. Rep.* **10**, 591
24. Balasuriya, N., Kunkel, M., Liu, X., Biggar, K., Li, S.-C., Newton, A., *et al.* (2018) Genetic code expansion and live cell imaging reveal that Thr-308 phosphorylation is irreplaceable and sufficient for Akt1 activity. *J. Biol. Chem.* **293**, 10744–10756
25. Paleologou, K. E., Oueslati, A., Shakked, G., Rospigliosi, C. C., Kim, H. Y., Lamberto, G. R., *et al.* (2010) Phosphorylation at S87 is enhanced in synucleinopathies, inhibits α -synuclein oligomerization, and influences synuclein-membrane interactions. *J. Neurosci.* **30**, 3184–3198
26. Rogerson, D. T., Sachdeva, A., Wang, K., Haq, T., Kazlauskaitė, A., Hancock, S. M., *et al.* (2015) Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. *Nat. Chem. Biol.* **11**, 496–503
27. Zhang, M. S., Brunner, S. F., Huguenin-Dezot, N., Liang, A. D., Schmied, W. H., Rogerson, D. T., *et al.* (2017) Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. *Nat. Methods* **14**, 729–736
28. Zhu, P., Gafken, P. R., Mehl, R. A., and Cooley, R. B. (2019) A highly versatile expression system for the production of multiply phosphorylated proteins. *ACS Chem. Biol.* **14**, 1564–1572
29. Park, H. S., Hohn, M. J., Umehara, T., Guo, L. T., Osborne, E. M., Benner, J., *et al.* (2011) Expanding the genetic code of *Escherichia coli* with phosphoserine. *Science* **333**, 1151–1154
30. Luo, X., Fu, G., Wang, R. E., Zhu, X., Zambaldo, C., Liu, R., *et al.* (2017) Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. *Nat. Chem. Biol.* **13**, 845–849
31. Steinfeld, J. B., Aerni, H. R., Rogulina, S., Liu, Y., and Rinehart, J. (2014) Expanded cellular amino acid pools containing phosphoserine, phosphothreonine, and phosphotyrosine. *ACS Chem. Biol.* **9**, 1104–1112
32. Thompson, J. R., Bell, J. K., Bratt, J., Grant, G. A., and Banaszak, L. J. (2005) Vmax regulation through domain and subunit changes. The active form of phosphoglycerate dehydrogenase. *Biochemistry* **44**, 5763–5773
33. Deng, H., Chen, C., Sun, C., and Wei, C. (2016) Construction and characterization of *Escherichia coli* D-3-phosphoglycerate dehydrogenase mutants with feedback-inhibition relief. *Sheng Wu Gong Cheng Xue Bao* **32**, 468–477
34. Mundhada, H., Schneider, K., Christensen, H. B., and Nielsen, A. T. (2016) Engineering of high yield production of L-serine in *Escherichia coli*. *Biotechnol. Bioeng.* **113**, 807–816
35. Kinoshita, E., Kinoshita-Kikuta, E., Takiyama, K., and Koike, T. (2006) Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Mol. Cell. Proteomics* **5**, 749–757
36. Beyer, J. N., Hosseinzadeh, P., Gottfried-Lee, I., van Fossen, E. M., Zhu, P., Bednar, R. M., *et al.* (2020) Overcoming near-cognate suppression in a release factor 1-deficient host with an improved nitro-tyrosine tRNA synthetase. *J. Mol. Biol.* **432**, 4690–4704
37. Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., and Li, Q. (2009) Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Sci.* **18**, 936–948
38. Andrews, B. T., Gosavi, S., Finke, J. M., Onuchic, J. N., and Jennings, P. A. (2008) The dual-basin landscape in GFP folding. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 12283–12288
39. Khan, F., Stott, K., and Jackson, S. (2003) ¹H, ¹⁵N and ¹³C backbone assignment of the green fluorescent protein (GFP). *J. Biomol. NMR* **26**, 281–282
40. Du, J. T., Li, Y. M., Wei, W., Wu, G. S., Zhao, Y. F., Kanazawa, K., *et al.* (2005) Low-barrier hydrogen bond between phosphate and the amide group in phosphopeptide. *J. Am. Chem. Soc.* **127**, 16350–16351
41. Nagano, N., Gail Hutchinson, E., and Thornton, J. M. (1999) Barrel structures in proteins: automatic identification and classification including a sequence analysis of TIM barrels. *Protein Sci.* **8**, 2072–2084
42. Carlson, C., Asfaha, J., Ghent, C., Howard, C., Hartooni, N., Safari, M., *et al.* (2020) Phosphoregulation of phase separation by the SARS-CoV-2 N protein suggests a biophysical basis for its dual functions. *Mol. Cell* **80**, 1092–1103
43. Savastano, A., de Opakua, A. I., Rankovic, M., and Zweckstetter, M. (2020) Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates. *Nat. Commun.* **11**, 6041
44. Lu, S., Ye, Q., Singh, D., Cao, Y., Diedrich, J., Yates, J. I., *et al.* (2021) The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. *Nat. Commun.* **12**, 502
45. Guseva, S., Perez, L. M., Camacho-Zarco, A., Bessa, L. M., Salvi, N., Malki, A., *et al.* (2021) ¹H, ¹³C and ¹⁵N backbone chemical shift assignments of the n-terminal and central intrinsically disordered domains of SARS-CoV-2 nucleoprotein. *Biomol. NMR Assign.* **15**, 255–260
46. Julien, M., Bougouchtoui, C., Alik, A., Ghoul, R., Zinn-Justin, S., and Theillet, F. X. (2020) Multiple site-specific phosphorylation of IDPs monitored by NMR. *Methods Mol. Biol.* **2141**, 793–817
47. Conibear, A. C., Rosengren, K. J., Becker, C. F. W., and Kaehlig, H. (2019) Random coil shifts of posttranslationally modified amino acids. *J. Biomol. NMR* **73**, 587–599
48. Vieweg, S., Mulholland, K., Bräuning, B., Kachariya, N., Lai, Y. C., Toth, R., *et al.* (2020) PINK1-dependent phosphorylation of Serine111 within the SF3 motif of Rab GTPases impairs effector interactions and LRRK2-mediated phosphorylation at Threonine72. *Biochem. J.* **477**, 1651–1668
49. Mekhattu Tharayil, S., Mahawaththa, M. C., Loh, C.-T., Adekoya, I., and Otting, G. (2021) Phosphoserine for the generation of lanthanide-binding sites on proteins for paramagnetic nuclear magnetic resonance spectroscopy. *Magn. Reson.* **2**, 1–13
50. Stuber, K., Schneider, T., Werner, J., Kovermann, M., Marx, A., and Scheffner, M. (2021) Structural and functional consequences of NEDD8 phosphorylation. *Nat. Commun.* **12**, 5939
51. Frey, S., and Görlich, D. (2014) A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins. *J. Chromatogr. A* **1337**, 95–105
52. Zhang, Y., Werling, U., and Edelman, W. (2012) SLICE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res.* **40**, e55
53. Forsythe, H. M., Rodriguez Galvan, J., Yu, Z., Pinckney, S., Reardon, P., Cooley, R. B., *et al.* (2021) Multivalent binding of the partially disordered SARS-CoV-2 nucleocapsid phosphoprotein dimer to RNA. *Biophys. J.* **120**, 2890–2901
54. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293
55. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.* **278**, 313–352
56. Ying, J., Delaglio, F., Torchia, D. A., and Bax, A. (2017) Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. *J. Biomol. NMR* **68**, 101–118