1-24-2017

MS-READ: Quantitative Measurement of Amino Acid Incorporation

Kyle Mohler
The Ohio State University

Hans-Rudolph Aerni
Yale University

Brandon Gassaway
Yale University

Jiqiang Ling
The Ohio State University

Michael Ibba
Chapman University, ibba@chapman.edu

See next page for additional authors

Recommended Citation

Mohler, K., Aerni, H.R., Gassaway, B., Ling, J., Ibba, M. and Rinehart, J. (2017) MS-READ: Quantitative measurement of amino acid incorporation. *BBA - Gen Subjects.* **1861**, 3081-3081. https://doi.org10.1016/j.bbagen.2017.01.025

MS-READ: Quantitative Measurement of Amino Acid Incorporation

Authors

Kyle Mohler, Hans-Rudolph Aerni, Brandon Gassaway, Jiqiang Ling, Michael Ibba, and Jesse Rinehart

Published in final edited form as:

Biochim Biophys Acta. 2017 November ; 1861(11 Pt B): 3081–3088. doi:10.1016/j.bbagen.2017.01.025.

MS-READ: Quantitative Measurement of Amino Acid Incorporation

Kyle Mohler[1,2], **Hans-Rudolph Aerni**[3,4], **Brandon Gassaway**[3,4], **Jiqiang Ling**[5], **Michael Ibba**[1,2], and **Jesse Rinehart**[3,4,†]

[1]Department of Microbiology, The Ohio State University, 318 West 12th Avenue, Columbus, Ohio 43210

[2]Center for RNA Biology, The Ohio State University, 484 West 12th Avenue, Columbus, Ohio 43210

[3]Systems Biology Institute, Yale University, West Haven, CT 06516

[4]Department of Cellular & Molecular Physiology, Yale, PO Box 27388, West Haven, CT 06516

[5]Department of Microbiology and Molecular Genetics, McGovern Medical School, University of Texas Health Science Center, Houston, TX 77030, USA

Abstract

Ribosomal protein synthesis results in the genetically programmed incorporation of amino acids into a growing polypeptide chain. Faithful amino acid incorporation that accurately reflects the genetic code is critical to the structure and function of proteins as well as overall proteome integrity. Errors in protein synthesis are generally detrimental to cellular processes yet emerging evidence suggest that proteome diversity generated through mistranslation may be beneficial under certain conditions. Cumulative translational error rates have been determined at the organismal level, however codon specific error rates and the spectrum of misincorporation errors from system to system remain largely unexplored. In particular, until recently technical challenges have limited the ability to detect and quantify comparatively rare amino acid misincorporation events, which occur orders of magnitude less frequently than canonical amino acid incorporation events. We now describe a technique for the quantitative analysis of amino acid incorporation that provides the sensitivity necessary to detect mistranslation events during translation of a single codon at frequencies as low as 1 in 10,000 for all 20 proteinogenic amino acids, as well as non-proteinogenic and modified amino acids.

Key terms

Misincorporation; Mistranslation; Mass Spectrometry Reporter

[†]Corresponding author and person to whom requests should be addressed: Jesse Rinehart, PhD, Department of Cellular & Molecular Physiology, Yale University, PO Box 27388, West Haven, CT 06516;, Phone: 203-737-3144, jesse.rinehart@yale.edu.

1. Background

1.1 Overview and definition of mistranslation

Protein synthesis is a complex and energy intensive cellular process. To achieve optimal growth rates while conserving cellular resources, cells have developed numerous mechanisms to ensure the accuracy of translation and maintenance of proteome fidelity. Translation fidelity is maintained in part through the accurate pairing of amino acids with cognate tRNAs and accurate selection of aminoacyl-tRNAs at the ribosome [1–3]. Despite these mechanisms, amino acid misincorporation occurs once in every 1,000 to 10,000 codons translated, resulting in around 15% of all proteins in the cell possessing at least one mistranslated amino acid [4, 5].

Errors in protein synthesis have traditionally been viewed as detrimental to cellular processes, yet emerging evidence suggests beneficial roles for limited mistranslation in certain biological contexts [6, 7]. In the yeast *Candida albicans*, for example, proteome diversity generated through mistranslation has been shown to increase phenotypic diversity and evasion of host immune response [8]. In contrast, misincorporation of non-proteinogenic amino acids often results in decreased cellular fitness [9]. These observations support the idea that protein synthesis error rates can vary substantially between different systems and that, beyond the level of error, amino acid specific chemicophysical properties play critical roles in the determination of cellular toxicity.

Cumulative translational error rates have been determined at the organismal level, yet the extent of codon specific error rates and the spectrum of misincorporation error from system to system remain less explored. Implementation of high resolution methods to determine global baselines for codon specific misincorporation events would provide an essential resource for the study of translation, and would prove particularly useful in the development of orthogonal translation systems where sensitive detection of off target incorporation is essential to the establishment of true orthogonality.

1.2 Uses and requirements within the field

Studies in protein science often depend on the reliable identification of amino acids at a specific position within the target protein. Additionally, the ability to accurately quantify site specific amino acid incorporation events is critical to our understanding of translation. A greater challenge, however, is the ability to measure amino acid misincorporation events which occur at frequencies orders of magnitude lower than canonical incorporation events. Furthermore, experimental approaches must also be robust in the detection of non-proteinogenic amino acids, central to the study of synthetic biology. A reliable method for the identification and quantification of site specific amino acid incorporation events must possess high sensitivity and low bias, be able to accommodate any amino acid moiety at all positions within a polypeptide, and must be robust in all manner of biologically relevant experimental contexts. This combination of requirements creates an ideal scenario for the application of mass spectrometry based approaches.

1.3 Problems with current approaches

Current approaches for detecting amino acid misincorporation include the use of gain-of-function reporters, recombinant expressed proteins, and shotgun proteomics. In each case, technical limitations impose restrictions on their effective use as reporters of amino acid misincorporation. Gain-of-function reporters rely on the site specific misincorporation of a target amino acid to restore enzymatic or fluorogenic properties of the reporter. In this case, restoration of protein activity is used as a proxy for misincorporation rate. Protein reporters for this experimental approach have included β-lactamase, GFP, luciferase, and β-Galactosidase [10–13]. However, these reporters often lack the sensitivity necessary to accurately detect low level misincorporation events. Additionally, functionality is often restored by substitution of more than one amino acid at the target position, complicating interpretation of results.

The use of recombinantly expressed proteins as reporters of protein mistranslation has been explored in a variety of contexts. This approach relies on the measurement of naturally occurring misincorporation events within the entire polypeptide, rather than focusing on a site specific misincorporation event [14, 15]. Affinity purification of the reporter protein from the cytoplasmic pool of proteins is often used to isolate the protein of interest prior to analysis. As a result, misincorporation events that lead to misfolded proteins are often underrepresented due to degradation or reduction in solubility [16]. Post isolation, the protein is analyzed by mass spectrometry, where, depending upon the impact of misincorporation events on protein structure, the ionization of differentially mistranslated proteins may create biases in peptide detection [17]. Additionally, traditional analysis of peptides is confounded by the occurrence of multiple misincorporation events within the same peptide, further reducing overall sensitivity of the approach. An ideal reporter of mistranslation must faithfully report all translation events regardless of their impact on structure or function of the protein, creating a context independent parallel of the biological question.

Global analysis of proteomic misincorporation events has been explored using recent developments in “shotgun” proteomics techniques. The shotgun approach, in theory, allows for a comprehensive analysis of the entire cellular complement of proteins and has been used to detect misincorporation events across the proteome [18]. In practice, however, the predicted misincorporation event must be anticipated due to limitations in current data analysis strategies based on data-base searching algorithms and unbiased sequence surveys which lack the ability to accurately detect all possible mistranslation events and have been associated with extremely high false discovery rates [19, 20]. As with most mass spectrometry based proteomic analyses, coverage is often incomplete due to variations in the primary sequence and overall structure of a target protein affect fragmentation and ionization of target peptides, decreasing overall abundance below the limit of detection and reducing the likelihood of observing rare misincorporation events [21].

1.4 Overview of how MS-READ overcomes these issues

As outlined above, the use of a proteins as reporters of mistranslation has been explored previously. However, these techniques are prone to peptide ionization biases due, in part, to

the highly structured conformation of most proteins. Misincorporation events may exacerbate this effect by altering the structure of the protein and potentially the charge state of the target peptide. As a result, the peptide of interest may vary in abundance, confounding detection and analysis of misincorporation events. The **M**ass **S**pectrometry **R**eporter for **E**xact **A**mino acid **D**ecoding (**MS-READ**) approach largely overcomes ionization issues through the use of an unstructured elastin-derived polypeptide as the reporter component of the protein construct. Elastin is an amorphous protein composed primarily of hydrophobic amino acids which allow for predictable and consistent ionization of peptide constituents. Our approach offers a low-cost, label free quantitation method that is easy to implement and allows for the sensitive and accurate detection of amino acids at a flexible codon position.

1.5 Design and composition of MS-READ reporter

MS-READ centers on the monitoring of site-specific amino acid incorporations within a reporter peptide. The reporter peptide was constructed based on an elastin-like polypeptide sequence containing a VPGXG repeat where X has been shown to be highly permissive to the incorporation of any amino acid expect proline [22]. The elastin-like polypeptide domain was fused to green fluorescent protein (GFP) to aid the stable expression of the reporter peptide while providing a simple means to monitor reporter expression via fluorescence spectroscopy. GFP was selected based on its high tolerance for amino acid misincorporation events and ease of detection based on its fluorescent properties. The GFP fusion was tagged at the C-terminus with six His residues to facilitate reporter isolation using standard metal affinity techniques (Figure 1A).

2. Methods

2.1 MS-READ Plasmid Construction and Reporter Expression

2.1.1 EcMS-READ Plasmid Construction and Reporter Expression—The MS-READ plasmid was constructed by direct gene synthesis (IDT gBlocks) of the N-terminal extension (MSKGPGKVPGAGVPGXGVPGVGKGGGT, see Fig S1) and by sub-cloning synthetic DNA fragments into a PCRT7 NT Topo tetR pLtetO plasmid containing GFP (5′-KpnI/3′HindIII) as described in Pirman *et al.* 2015. Specific gBlock variants were made to add different codons (ACA-Thr, TAG-Amber, and an NNK random library) at position X in the N-terminal extension. All plasmid variants were validated by sequencing. NNK library members were sub-cloned out and individual codons representing all 20 amino acids were confirmed by sequencing (Figure 1 and S1). Protein expression was carried out as in Pirman *et al.* 2015. For experiments in *E. coli,* culture growth in LB or minimal medium (M9) with or without Ser supplementation was monitored by UV/Vis spectroscopy at 600nm (OD_{600}). *E. coli* MG1655 harboring EcMS-READ plasmid was maintained in the presence of ampicillin (50 μg/ml). Protein expression was induced with 100 ng/mL anhydotetracycline at an OD_{600} of 0.5. Cultures were grown an additional 4 hours, quenched on ice and pelleted at 2,000 × g (15 min at 4°C). The supernatant was discarded and the cell pellets were frozen at −80°C to assist with subsequent protein extraction.

2.1.2 ScMS-READ Plasmid Construction and Reporter Expression—For construction of the ScMS-READ plasmid, the ELP reporter sequence was fused to the N-

terminus of a yeGFP and placed under the control of a constitutive promoter (TEF1). The coding sequences for TEF1, the ELP reporter fragment, and yeGFP were amplified from the plasmids pYM-N18, EcMS-READ, and pYM25, respectively, by PCR. Fragments were assembled into the MCS of plasmid pXRH3 using Gibson Assembly per manufacturer recommendations (NEB, Ipswich, MA). For experiments in *S. cerevisiae*, culture growth in minimal medium without His (SDMM-His) was monitored by UV/Vis spectroscopy at 600nm (OD_{600}). WT and PheRS proofreading deficient (*frs1-1*) *S. cerevisiae* strains were transformed with ScMS-READ plasmid with a codon for Phe (UUU) at the variable position and maintained in the absence of His. At an OD_{600} of 0.8–1.0, cultures were quenched on ice and pelleted at 2,000 × g (15 min at 4°C). The supernatant was discarded and the cell pellets were frozen at −80°C.

2.2 Reporter Protein Purification

2.2.1 Purification from *E. coli*—Frozen *E. coli* cell pellets were thawed on ice and pellets were lysed by sonication with lysis buffer consisting of 50 mM Tris-HCl (pH 7.4, 23°C), 500 mM NaCl, 0.5 mM EGTA, 1mM DTT, 10 % glycerol, 50 mM NaF, and 1 mM Na_3O_4V. The extract was clarified with two rounds of centrifugation performed for 20 minutes at 4 °C and 14,000 × g. Cell free extracts were applied to Ni-NTA metal affinity resin and purified according to the manufacturer's instructions. Wash buffers contained 50 mM Tris pH 7.5, 500 mM NaCl, 0.5 mM EGTA, 1mM DTT, 50 mM NaF, 1 mM Na_3VO_4 and increasing concentrations of imidazole 20 mM, 40mM, and 60mM, sequentially. Proteins were eluted with wash buffer containing 250 mM imidazole. Eluted protein was subjected to 4 rounds of buffer exchange (20mM Tris pH 8.0 and 100mM NaCl) and concentrated using a 30 kDa molecular weight cutoff spin filter (Amicon).

2.2.2 Purification from *S. cerevisiae*—Frozen yeast cell pellets were thawed and incubated with 300U/ml of zymolyase for 20 minutes at RT. Lysis buffer was added to the digested cells, along with sterile 0.1 mm acid washed glass beads. Cell suspensions were vortexed 6 times, 30 seconds per round. Samples were briefly sonicated and the supernatant was removed after two rounds of centrifugation at 14,000 × g for 20 minutes. The remaining pellet was re-extracted and resulting fractions were combined. Cell free extracts were applied to Ni-NTA metal affinity resin and purified as above. SDS–PAGE electrophoresis followed by staining with Coomassie blue revealed greater than 90% purity. Protein concentration was determined by UV/Vis spectroscopy at 280nm using a nanodrop spectrophotometer. Samples were frozen and stored at −80°C.

2.3 Protein Digestion and mass spectrometry

2.3.1 Sample Preparation—Affinity purified, buffer exchanged protein was digested using a 2-step digestion protocol with Lysyl Endopeptidase (LysC) and trypsin and analyzed by mass spectrometry. A detailed description of the digestion protocol has been provided earlier [23]. Briefly, the concentration of protein was determined by UV280 spectroscopy on a NanoDrop (Thermo Fisher) and aliquots corresponding to 40 μg or 20 μg GFP reporter from *E. coli* or *S. cerevisiae* respectively were dissolved in 50 μl solubilization buffer consisting of 10 mM Tris-HCl pH=8.5 (23°C), 10 mM DTT, 1 mM EDTA and 0.5 % acid labile surfactant (ALS-101, Protea). Samples were heat denatured for 6 min at 95 °C in a

heat block. Alkylation of cysteines was performed with iodoacetamide (IAA) using a final IAA concentration of 24 mM. The alkylation reaction proceeded for 30 min at room temperature and in the dark. Excess IAA was quenched with DTT and the buffer concentration was adjusted using a 1 M Tris-HCl stock solution (pH=8.5) resulting in a final Tris-HCl concentration of 150 mM. Digestion with LysC (Wako) was performed for 4 h at 37 °C using $3.5 \cdot 10^{-5}$ units LysC per 1 μg of GFP. The reaction was then diluted with water and 1 M $CaCl_2$ solution to obtain a ALS-101 concentration of 0.045 % and 2 mM $CaCl_2$ respectively. Finally, sequencing grade porcine trypsin (Promega) was added to obtain an enyzme/protein ratio of 1/5.3 and the digest was incubated for 15 h at 37 °C without shaking. The digest was quenched with 20% TFA solution resulting in a sample pH of 2. Cleavage of the acid cleavable detergent proceeded for 15 min at room temperature. Digests were frozen at −80 °C until further processing.

Peptides were desalted on C_{18} UltraMicroSpin columns (The Nest Group Inc.) essentially following the instructions provided by the manufacturer but using 300 μl elution solvent consisting of 80% ACN, 0.1% TFA for peptide elution. Peptides were dried in a vacuum centrifuge at room temperature. Dried peptides were reconstituted by vortex by sequentially adding 2.5 μl 70 % formic acid, 3.75 μl 1-propanol and 8.75 μl 0.5 % acetic acid and the peptide concentration was determined by UV280 using a NanoDrop. Peptide stock solution were diluted to a concentration of 50ng/ul in glass HPLC vials (Agilent) using LC-MS solvent. The composition of the LC-MS solvent was 100 μg/ul sodium deoxycholate prepared in a solvent mixture of 5 % DMSO and 95 % of a solvent mixture consisting of 19 % formic acid/0.07 % TFA in water.

2.3.1 Liquid chromatograph mass spectrometry—LC-MS/MS was performed on an Orbitrap Velos using a Top10 HCD method as described previously [23] but using an optimized solvent gradient for peptide separation. The trapping column consisted of a 3 cm × 150 μm Kasil frit terminated fused silica capillary packed with 3 μm particle size Reprosil-Pur 120 C18-AQ (Dr Maisch GmBH). The analytical column was a 20 cm × 75 μm ID picofrit (New Objective) column packed with 1.9 μm particle size Reprosil-Pur 120 C18-AQ (Dr Maisch GmBH) to a length of 20 cm. Trapping column and analytical column were connected using a vented split setup using a low dead volume T metal connector where the spray voltage was applied. Eluent A was 0.1 % formic acid in water and eluent B was 0.1 % formic acid in acetonitrile. Trapping of peptides was performed for 5 min at a flow rate of 2.5 μl/min with an eluent composition of 2 % B. Gradient separation of peptides proceeded at a flow rate of 300 nl/min using the following linear gradient program (min/%B): 0.0/2.0, 0.1/2.0, 60.00/25.0, 70.0/40, 72.0/95.0, 78.0/95.0 80.0/2.0, 90.0/2.0. An estimated 150 ng of the digest was injected for each experiment.

2.4.2 Data Analysis—A detailed description of the bioinformatics strategy for the identification of amino acids at a target site has been described previously [24, 25]. Briefly, raw spectra were processed and matched using MaxQuant [26] v. 1.5.1.2 or 1.5.3.30 software. Unless noted otherwise, the default processing parameters for MaxQuant were used. Searches were performed with a custom database of the MS-READ reporter N-terminal extension containing all potential canonical AA at position X, and an *E. coli*

(UniProt strain K12, 4313 sequences, release version 2014_08) or *S. cerevisiae* protein database (Uniprot, strain AWRI1631, 5450 sequences, downloaded 27 August 2015) respectively. The enzyme specificity was trypsin\P allowing up to 3 missed cleavages. The search tolerances were 20 ppm for precursor and 20 ppm for fragment ions respectively. Carbamidomethylation of cysteines was specified as a fixed modification and deamidation (N/Q) and oxidation (M) were variable modifications. Peptides and proteins were reported with an estimated false discovery rate of 1 % and all peptides and proteins matching common sample contaminants or the reverse database were removed in Perseus [27] software v. 1.4.0.20.

3. Results

3.1 Comprehensive resolution of amino acid incorporation based on retention time

To demonstrate the permissibility of the reporter to amino acid incorporation, we constructed a library of reporters using a degenerate NNK codon at position 16 of the polypeptide which yielded 30 out of 32 potential codons (Figure 1B). The entire population of reporter constructs was expressed recombinantly *in E. coli* and tryptic reporter peptides were analyzed according to the protocol outlined above (Figure 2A). Reporter peptides (with the exception of Ile and Leu containing peptides) displayed unique retention signatures based on the amino acid incorporated at the variable codon position (Figure 2B). The ability to chromatographically resolve unique peptides within a mixed population streamlines deconvolution of mass spectra and highlights the ability of the reporter polypeptide to accurately identify amino acid incorporation events (Figure S2).

3.2 Measurement of aaRS-mediated mistranslation in E. coli and S. cerevisiae

To explore the use of MS-READ as a means to measure aaRS-mediated mistranslation, the incorporation of amino acids was monitored at a Thr (ACA) codon within the reporter construct transformed into a strain of *E. coli* harboring an proofreading deficient threonyl-tRNA synthetase (ThrRS). The proofreading defective ThrRS has previously been shown to misactivate Ser and misacylate $tRNA^{Thr}$ to form stable Ser-$tRNA^{Thr}$ *in vivo* [10]. As expected, only Thr was detected at the Thr codon in reporter peptides isolated from the WT ThrRS *E. coli* strain. In accordance with previous observations, Ser incorporation at the Thr codon was also detected in the ThrRS editing deficient strain of *E. coli* (Figure 3A). The cellular effects of Ser misincorporation were explored through determination of temperature dependent growth rates. WT and proofreading deficient cells were grown with and without Ser supplementation. The growth rate of WT and proofreading deficient cells was similar across all growth temperatures in minimal media, yet the growth rate of proofreading deficient strain was reduced during Ser supplementation compared to WT (Figure 3B). By demonstrating Ser misincorporation, MS-READ analysis contributes data, previously assumed, linking misincorporation events to phenotypic observations [28].

To expand the functionality of MS-READ, we constructed an MS-READ reporter for use in *Saccharomyces cerevisiae*. Similar to the experiments in *E. coli* above, we sought to demonstrate the use of MS-READ in the measurement of aaRS-mediated mistranslation in yeast. A reporter construct with a Phe codon at the flexible position was expressed in a strain

of *S. cerevisiae* with an aminoacyl-tRNA proofreading deficient allele of phenylalanyl-tRNA synthetase (PheRS). aa-tRNA proofreading deficient PheRS allows for the production and accumulation of misacylated Tyr-tRNA^{Phe} [29]. As expected, incorporation of both the amino acids Phe and Tyr was observed at the Phe codon in reporter peptides analyzed from the proofreading deficient PheRS strain, whereas, only the incorporation of Phe was observed from the WT PheRS background (Figure 4). Taken together, these observations provide support for the application of MS-READ analysis across all domains of life.

3.4 MS-READ in synthetic biology: monitoring the efficacy of orthogonal translation systems

Orthogonal translation systems (OTS) operate in parallel with existing translation machinery but provide substrates which expand the repertoire of programmed protein synthesis beyond the 20 canonical amino acids [30]. These systems have been used to explore the impact of non-proteinogenic amino acids on cellular function and, more recently, to provide a method for the site specific incorporation of phosphoserine into a recombinantly expressed protein (Figure 5A) [31–33]. Extension of OTS into site specific applications necessitates the ability to monitor amino acid incorporation at the site of interest. To demonstrate the application of MS-READ towards the assessment of OTS, we expressed and purified reporter protein containing a TAG codon at the flexible position which is decoded as phosphoserine when used in parallel with the SepOTS. Phsophoserine incorporation was monitored by Phos-Tag mediated electrophoretic shift and confirmed by mass spectrometry (Figure 5B and C). The MS-READ approach provides a highly sensitive means to monitor the efficacy of OTSs, providing a critical tool for both development and implementation.

4. Discussion and Conclusions

4.1 Additional Applications

In addition to the applications explored above, MS-READ may be applied to any number of academic or industrial problems. In an academic context, the approach may be used in assay development to monitor the bioavailability and utilization of non-proteinogenic amino acids *in vivo*. MS-READ also provides a means to conduct investigations of both global and site specific translation events across all major domains of life with previously unobtainable sensitivity and resolution. Applied to industry, MS-READ may be utilized to access the accuracy of production scale protein synthesis in systems which necessitate high translational fidelity. To match the diversity of potential applications, MS-READ workflow may be adapted to accommodate any method of mass tagging and quantification (e.g. SILAC, SRM/MRM, iTRAQ, etc.) and provides both the versatility to utilize labeling strategies *in vivo* as well as spike in standards for *ex vivo* quantification.

4.2 Overall Impact

MS-READ combines the best features of previously devised amino acid incorporation techniques by allowing the site specific quantification of amino acids within a mass spectrometry optimized recombinant reporter protein. Applications of this technology, thus far, highlight the global utility of the reporter across multiple domains of life in various biologically relevant scenarios. Further, we describe the use of the reporter to monitor the

performance of orthogonal translation systems, thus providing additional functionality within the field of synthetic biology. Overall, MS-READ is a powerful and versatile approach which promises to yield amino acid incorporation data of unprecedented resolution, with potential applications limited only by the ability to construct organism specific expression platforms.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Karl Barber, Miriam Amiram, Adrian Haimovich, and Farren Isaacs for helpful input and advice. This work was supported by grants from the National Science Foundation (MCB 1412611 (M.I.)), the National Institutes of Health (P01DK01743341 (J.R.), R01GM117230 (J.R.)), the Army Research Office (W911NF1510105 (M.I.)), NIH Training Grant T32 GM086252 (K.M.) and a fellowship from the OSU Center for RNA Biology (K.M.).

References

1. Marc D. Aminoacyl-tRNA synthetases. Current opinion in structural biology. 1995; 5:48. [PubMed: 7773747]
2. Jakubowski H. Quality control in tRNA charging, Wiley interdisciplinary reviews. RNA. 2012; 3
3. Reynolds NM, Lazazzera BA, Ibba M. Cellular mechanisms that control mistranslation. Nature reviews Microbiology. 2010; 8:849–856. [PubMed: 21079633]
4. Moghal A, Mohler K, Ibba M. Mistranslation of the genetic code. FEBS letters. 2014; 588:4305–4310. [PubMed: 25220850]
5. Schimmel P. Mistranslation and its control by tRNA synthetases. Philosophical transactions of the Royal Society of London. Series B Biological sciences. 2011; 366:2965–2971. [PubMed: 21930589]
6. Ribas de Pouplana L, Santos MA, Zhu JH, Farabaugh PJ, Javid B. Protein mistranslation: friend or foe? Trends in biochemical sciences. 2014; 39:355–362. [PubMed: 25023410]
7. Li L, Boniecki MT, Jaffe JD, Imai BS, Yau PM, Luthey-Schulten ZA, Martinis SA. Naturally occurring aminoacyl-tRNA synthetases editing-domain mutations that cause mistranslation in Mycoplasma parasites. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108:9378–9383. [PubMed: 21606343]
8. Santos MA, Cheesman C, Costa V, Moradas-Ferreira P, Tuite MF. Selective advantages created by codon ambiguity allowed for the evolution of an alternative genetic code in Candida spp. Molecular microbiology. 1999; 31:937–947. [PubMed: 10048036]
9. Bullwinkle T, Reynolds NM, Raina M, Moghal AB, Matsa E, Rajkovic A, Kayadibi H, Fazlollahi F, Ryan C, Howitz N, Faull KF, Lazazzera B, Ibba M. Oxidation of cellular amino acid pools leads to cytotoxic mistranslation of the genetic code. Place Published. 2014
10. Ling J, Soll D. Severe oxidative stress induces protein mistranslation through impairment of an aminoacyl-tRNA synthetase editing site. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:4028–4033. [PubMed: 20160114]
11. Javid B, Sorrentino F, Toosky M, Zheng W, Pinkham JT, Jain N, Pan M, Deighan P, Rubin EJ. Mycobacterial mistranslation is necessary and sufficient for rifampicin phenotypic resistance. Proceedings of the National Academy of Sciences of the United States of America. 2014; 111:1132–1137. [PubMed: 24395793]
12. Gomes AC, Kordala AJ, Strack R, Wang X, Geslain R, Delaney K, Clark WC, Keenan R, Pan T. A dual fluorescent reporter for the investigation of methionine mistranslation in live cells. Rna. 2016; 22:467–476. [PubMed: 26729921]

13. Lu J, Bergert M, Walther A, Suter B. Double-sieving-defective aminoacyl-tRNA synthetase causes protein mistranslation and affects cellular physiology and development. Nature communications. 2014; 5
14. Cvetesic N, Semanjski M, Soufi B, Krug K, Gruic-Sovulj I, Macek B. Proteome-wide measurement of non-canonical bacterial mistranslation by quantitative mass spectrometry of protein modifications. Scientific reports. 2016; 6:28631. [PubMed: 27377007]
15. Ruan B, Palioura S, Sabina J, Marvin-Guy L, Kochhar S, Larossa RA, Soll D. Quality control despite mistranslation caused by an ambiguous genetic code. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105:16502. [PubMed: 18946032]
16. Baneyx F, Mujacic M. Recombinant protein folding and misfolding in Escherichia coli. Nature biotechnology. 2004; 22:1399–1408.
17. Konermann L, Tong X, Pan Y. Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. Journal of mass spectrometry : JMS. 2008; 43:1021–1036. [PubMed: 18523973]
18. Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR 3rd. Protein analysis by shotgun/bottom-up proteomics. Chem Rev. 2013; 113:2343–2394. [PubMed: 23438204]
19. Reiter L, Claassen M, Schrimpf SP, Jovanovic M, Schmidt A, Buhmann JM, Hengartner MO, Aebersold R. Protein identification false discovery rates for very large proteomics data sets generated by tandem mass spectrometry. Molecular & cellular proteomics : MCP. 2009; 8:2405–2417. [PubMed: 19608599]
20. Colaert N, Degroeve S, Helsens K, Martens L. Analysis of the resolution limitations of peptide identification algorithms. Journal of proteome research. 2011; 10:5555–5561. [PubMed: 21995378]
21. Domon B, Aebersold R. Options and considerations when selecting a quantitative proteomics strategy. Nature biotechnology. 2010; 28:710–721.
22. Amiram M, Quiroz FG, Callahan DJ, Chilkoti A. A highly parallel method for synthesizing DNA repeats enables the discovery of 'smart' protein polymers. Nature materials. 2011; 10:141–148. [PubMed: 21258353]
23. Pirman NL, Barber KW, Aerni HR, Ma NJ, Haimovich AD, Rogulina S, Isaacs FJ, Rinehart J. A flexible codon in genomically recoded Escherichia coli permits programmable protein phosphorylation. Nature Communications. 2015; 6:8130.
24. 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 Res. 2015; 43:e8. [PubMed: 25378305]
25. Amiram M, Haimovich AD, Fan C, Wang YS, Aerni HR, Ntai I, Moonan DW, Ma NJ, Rovner AJ, Hong SH, Kelleher NL, Goodman AL, Jewett MC, Soll D, Rinehart J, Isaacs FJ. Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. Nat Biotechnol. 2015; 33:1272–1279. [PubMed: 26571098]
26. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008; 26:1367–1372. [PubMed: 19029910]
27. Cox J, Mann M. 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. BMC Bioinformatics. 2012; 13(Suppl 16):S12.
28. Wu J, Fan Y, Ling J. Mechanism of oxidant-induced mistranslation by threonyl-tRNA synthetase. Nucleic acids research. 2014; 42:6523–6531. [PubMed: 24744241]
29. Reynolds NM, Ling J, Roy H, Banerjee R, Repasky SE, Hamel P, Ibba M. Cell-specific differences in the requirements for translation quality control. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:4063–4068. [PubMed: 20160120]
30. Des Soye BJ, Patel JR, Isaacs FJ, Jewett MC. Repurposing the translation apparatus for synthetic biology. Current opinion in chemical biology. 2015; 28:83–90. [PubMed: 26186264]
31. Wang K, Sachdeva A, Cox DJ, Wilf NM, Lang K, Wallace S, Mehl RA, Chin JW. Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. Nature chemistry. 2014; 6:393–403.

32. Pirman NL, Barber KW, Aerni HR, Ma NJ, Haimovich AD, Rogulina S, Isaacs FJ, Rinehart J. A flexible codon in genomically recoded Escherichia coli permits programmable protein phosphorylation. Nature communications. 2015; 6:8130.

33. Park HS, Hohn MJ, Umehara T, Guo LT, Osborne EM, Benner J, Noren CJ, Rinehart J, Söll D. Expanding the genetic code of Escherichia coli with phosphoserine. Science. 2011; 333:1151–1154. [PubMed: 21868676]

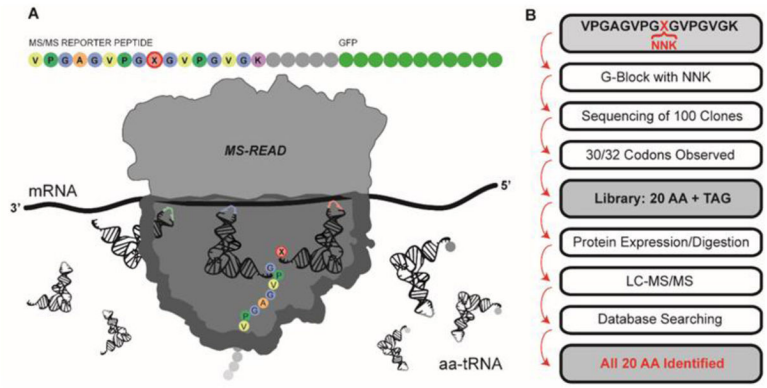

Figure 1. Overview or MS-READ method

A: Diagram of MS-READ features and components of reporter which include flexible reporter peptide fused to GFP. Amino acid incorporation is monitored at the variable X position within the reporter peptide sequence, where X may be any codon. **B:** Design and construction workflow for MS-READ reporter and validation studies.

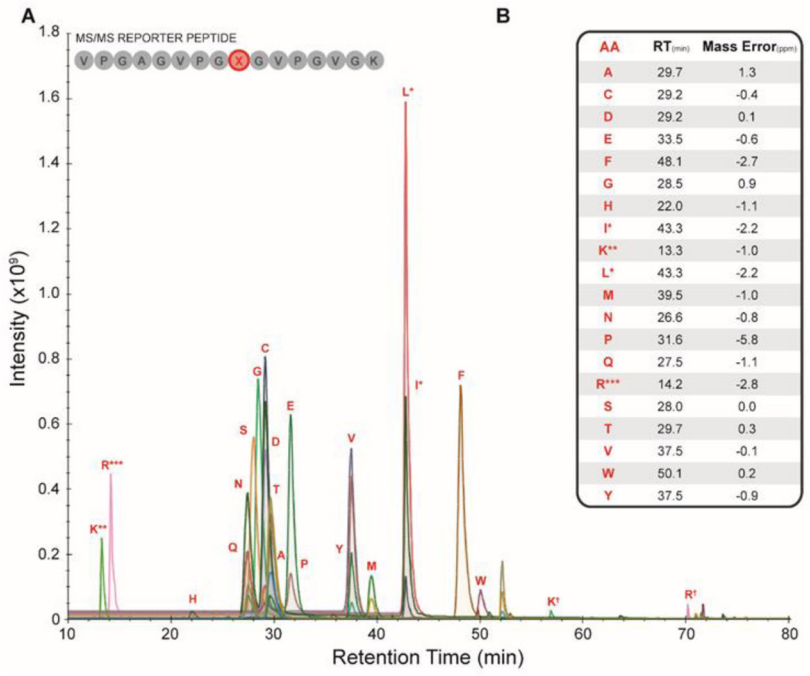

Figure 2. Analysis of amino acid incorporation and reporter peptide resolution

A: MS-READ reporter constructs with 30/32 codons were expressed and processed in parallel. Peptides containing each of the 20 proteinogenic amino acids were chromatographically resolved and observed within the population. **B:** Each amino acid specific peptide (with the exception of L and I, denoted by *) displayed a unique retention time signature. Peptides containing K and R were subject to additional cleavage events during tryptic digest, denoted by ** and ***, respectively. † denotes miss-cleaved K and R peptides that were observed but not reported in (B).

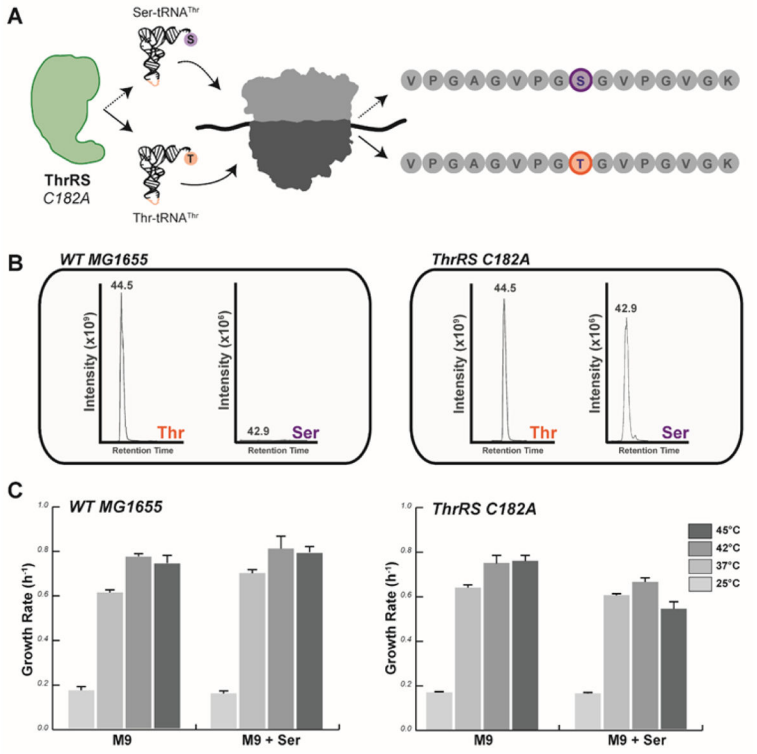

Figure 3. Application of MS-READ to monitor aaRS-mediated misincorporation in *E. coli*

A: WT and ThrRS proofreading deficient strains of *E. coli* harboring MS-READ plasmids with a Thr codon at the flexible position were examined for Thr → Ser misincorporation events. **B:** Extracted ion chromatograms display unique peptide signatures for both Thr and Ser containing peptides. **C:** Phenotypic effects of Ser misincorporation were examined by monitoring growth of both WT and ThrRS proofreading deficient strains of *E. coli* in response to challenge with Ser at increasing growth temperatures. Growth data was obtained by monitoring absorbance at 600nm and the growth rate was determined as a function of time. Error bars are representative of 1 SD from experiments conducted in triplicate.

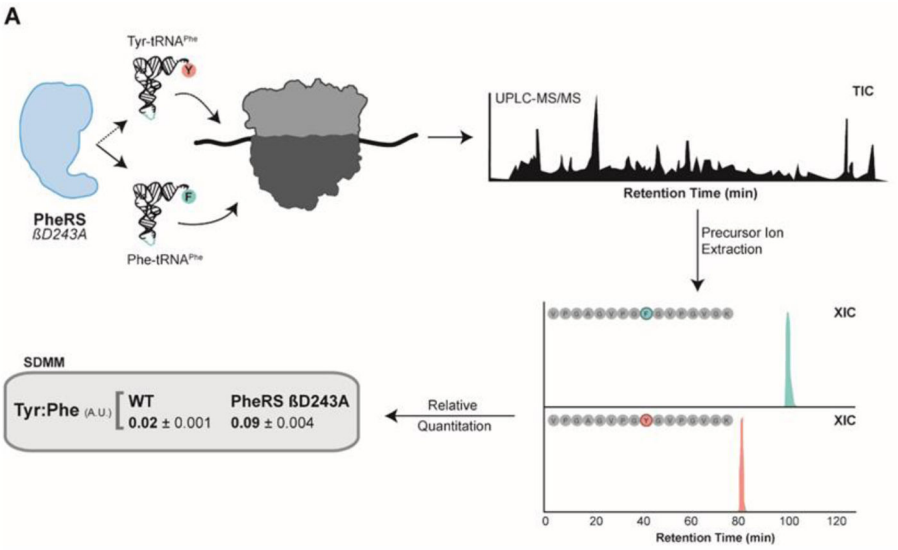

Figure 4. Application of MS-READ to monitor aaRS-mediated misincorporation in yeast

A: WT and PheRS proofreading deficient strains of *S. cerevisiae* harboring MS-READ plasmids with a Phe codon at the flexible position were examined for Phe → Tyr misincorporation events. Extracted ion chromatograms display unique peptide signatures for both Phe and Tyr- containing peptides. Relative misincorporation rates were determined from extracted ion chromatograms and reported as the fraction of Tyr incorporation relative to Phe. The standard deviation was determined from duplicate biological replicates.

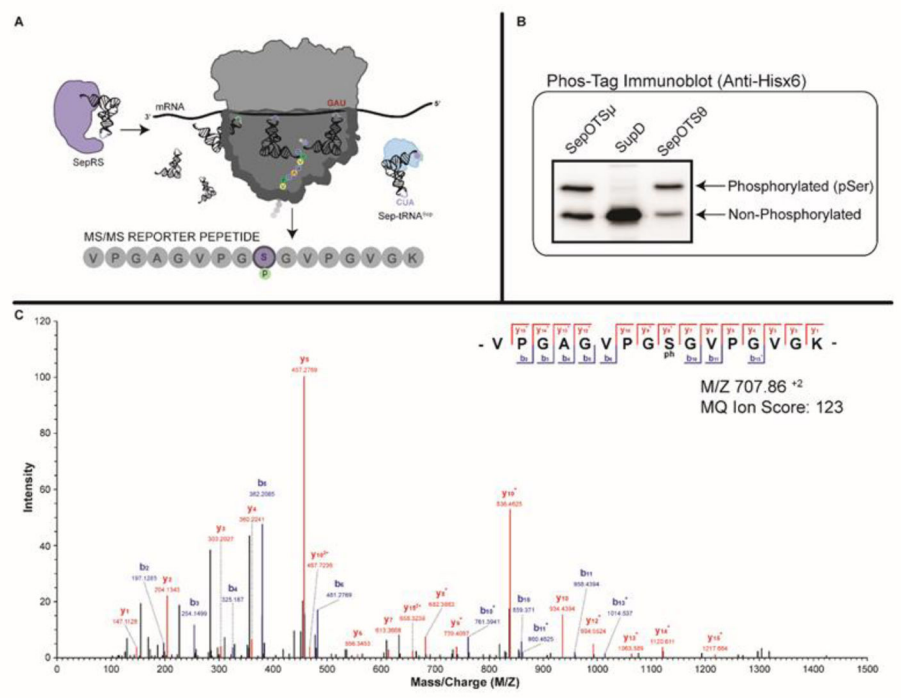

Figure 5. Application of MS-READ to monitor orthogonal translation systems

A: Diagram of the application of MS-READ to monitor phosphoserine incorporation through the use of orthogonal phosphoserine translation systems *in vivo*. Site specific incorporation of phosphoserine may be achieved using a suppressor tRNA which recognizes an amber stop codon. **B:** EcMS-READ reporter protein was expressed in the presence of SepOTS. Phosphoserine containing protein was separated from non-phsophoserine containing protein electrophoreticlaly using SDA-PAGE gel containing Phos-Tag acrylamide. Protein was detected by immunoblot using antibody specific for Hisx6. **C:** Annotated tandem mass spectrum from EcMS-READ reporter protein confirming the site-specific incorporation of phosphoserine (ph). The doubly charged parent ion mass to charge ratio (M/Z) is shown with the corresponding MaxQuant (MQ) ion score. Fragments ions are annotated (y and b) with neutral loss, phosphorylation signature fragments identified (*).