



# Tuning the Properties of Protein-Based Polymers Using High-Performance Orthogonal Translation Systems for the Incorporation of Aromatic Non-Canonical Amino Acids

## OPEN ACCESS

### Edited by:

Gustavo Fuertes,  
Institute of Biotechnology (ASCR),  
Czechia

### Reviewed by:

Huibin Zou,  
Qingdao University of Science and  
Technology, China  
Jeffery M. Tharp,  
Yale University, United States

### \*Correspondence:

Miriam Amiram  
mamiram@bgu.ac.il

<sup>†</sup>These authors have contributed  
equally to this work and share first  
authorship

### Specialty section:

This article was submitted to  
Nanobiotechnology,  
a section of the journal  
Frontiers in Bioengineering and  
Biotechnology

**Received:** 05 April 2022

**Accepted:** 27 April 2022

**Published:** 30 May 2022

### Citation:

Gueta O, Sheinenzon O, Azulay R,  
Shalit H, Strugach DS, Hadar D,  
Gelkop S, Milo A and Amiram M (2022)  
Tuning the Properties of Protein-Based  
Polymers Using High-Performance  
Orthogonal Translation Systems for the  
Incorporation of Aromatic Non-  
Canonical Amino Acids.  
Front. Bioeng. Biotechnol. 10:913057.  
doi: 10.3389/fbioe.2022.913057

Osher Gueta<sup>1†</sup>, Ortal Sheinenzon<sup>1†</sup>, Rotem Azulay<sup>1</sup>, Hadas Shalit<sup>2</sup>, Daniela S. Strugach<sup>1</sup>, Dagan Hadar<sup>1</sup>, Sigal Gelkop<sup>1</sup>, Anat Milo<sup>2</sup> and Miriam Amiram<sup>1\*</sup>

<sup>1</sup>Avram and Stella Goldstein-Goren Department of Biotechnology Engineering, Ben-Gurion University of the Negev, Beersheba, Israel, <sup>2</sup>Department of Chemistry, Ben-Gurion University of the Negev, Beersheba, Israel

The incorporation of non-canonical amino acids (ncAAs) using engineered aminoacyl-tRNA synthetases (aaRSs) has emerged as a powerful methodology to expand the chemical repertoire of proteins. However, the low efficiencies of typical aaRS variants limit the incorporation of ncAAs to only one or a few sites within a protein chain, hindering the design of protein-based polymers (PBPs) in which multi-site ncAA incorporation can be used to impart new properties and functions. Here, we determined the substrate specificities of 11 recently developed high-performance aaRS variants and identified those that enable an efficient multi-site incorporation of 15 different aromatic ncAAs. We used these aaRS variants to produce libraries of two temperature-responsive PBPs—elastin- and resilin-like polypeptides (ELPs and RLPs, respectively)—that bear multiple instances of each ncAA. We show that incorporating such aromatic ncAAs into the protein structure of ELPs and RLPs can affect their temperature responsiveness, secondary structure, and self-assembly propensity, yielding new and diverse families of ELPs and RLPs, each from a single DNA template. Finally, using a molecular model, we demonstrate that the temperature-responsive behavior of RLPs is strongly affected by both the hydrophobicity and the size of the unnatural aromatic side-chain. The ability to efficiently incorporate multiple instances of diverse ncAAs alongside the 20 natural amino acids can help to elucidate the effect of ncAA incorporation on these and many other PBPs, with the aim of designing additional precise and chemically diverse polymers with new or improved properties.

**Keywords:** non-canonical amino acids (ncAAs), elastin-like polypeptides (ELP), resilin-like polypeptides (RLP), smart biomaterials, genetic code expansion

## INTRODUCTION

Incorporating non-canonical amino acids (ncAAs) into proteins has emerged as a powerful methodology to improve, alter, or introduce new functions into proteins. The incorporation of ncAAs bearing a variety of chemical groups can facilitate the elucidation of protein structure–function relationship (Debelouchina and Muir, 2017; Chen et al., 2018) or protein–protein interactions (Nguyen et al., 2018), or the production of protein-based therapeutics (Huang and Liu, 2018) and biomaterials with novel functions (Connor and Tirrell, 2007; Israeli et al., 2019), among many other applications (Johnson et al., 2010; Chin, 2017; Voller and Budisa, 2017; Young and Schultz, 2018; Lee et al., 2019; Zhou and Deiters, 2021). There are two common strategies for expressing recombinant proteins containing ncAAs. The first entails substituting one of the natural AAs with an ncAA. This technique permits the incorporation of multiple instances of a ncAA into the target protein and in entire proteomes (Johnson et al., 2010). However, substitution allows the incorporation of only 19 natural AAs alongside the ncAA, and the ncAA must be a close analog of the natural AA that it replaces, therefore using this approach limits the flexibility of protein engineering. Alternatively, ncAAs can be incorporated by codon reassignment or frameshift codons using orthogonal translation systems consisting of an aminoacyl-tRNA synthetase (aaRS) and a tRNA pair that do not interact with the native aaRSs and tRNAs in the expression host. This technique has enabled the incorporation of >200 different ncAAs, using aaRS-tRNA pairs such as the tRNA<sup>Tyr</sup>·TyrRS pair from *Methanocaldococcus jannaschii* (Kobayashi et al., 2003) or the tRNA<sup>Pyl</sup>·PylRS pair from various methanogenic archaea and bacteria, such as from the *Methanosarcina* or *Methanomethylophilus* spp. (Wan et al., 2014; Willis and Chin, 2018). Typically, a TAG stop codon is assigned to the ncAA and the orthogonal tRNA anticodon is mutated to CUA (if needed). In addition, the evolution of the aaRS by mutagenesis of its AA-binding site is required to accommodate the ncAA. However, traditional evolution methodologies typically yield aaRS variants capable of introducing the ncAA at only one or a few instances within a polypeptide chain, and often with reduced yields compared to wild-type proteins (O'Donoghue et al., 2013; Vargas-Rodriguez et al., 2018).

We recently developed an aaRS evolution platform that utilizes multiplex automated genome engineering (MAGE) to create a library of chromosomal MjTyrRS variants. High-performance aaRS variants were selected from these libraries, enabling multi-site incorporation (10–30 instances per protein) of the aromatic ncAAs p-acetyl-phenylalanine (pAcF) (Amiram et al., 2015), p-azido-phenylalanine (pAzF) (Amiram et al., 2015), 4-propargyloxy-L-phenylalanine (pPR) (Hadar et al., 2021), and phenylalanine-4'-azobenzene (AzoPhe) (Israeli et al., 2021). Although these aaRS variants were selected based on their ability to incorporate the ncAAs mentioned above, it had been previously demonstrated by us and others that most aaRS variants exhibit some degree of poly-specificity, that is, they can recognize and charge their cognate tRNA with ncAAs bearing various chemical groups (Young

et al., 2011; Amiram et al., 2015). The elucidation of the substrate specificities of high-performance aaRS variants can be a facile means for generating highly efficient aaRS–ncAA pairs for the multi-site incorporation of additional ncAAs. Such an expanded set of aaRS–ncAA pairs for the efficient incorporation of multiple instances of ncAAs within a single protein chain can enable the engineering of proteins with the desired new or improved functions and their high-yield production (Ohtake et al., 2015; Rezhdo et al., 2019; Hayashi et al., 2021). In addition, analyzing the relationship between the mutations in the aaRS AA-binding site and the corresponding ncAA binding specificities can inform computational models and facilitate the selection of smaller, focused libraries for future aaRS evolution experiments (Hauf et al., 2017; Baumann et al., 2019).

Multi-site ncAA incorporation is particularly desirable for the design and production of protein-based polymers (PBPs), which consist of tandem repeats of either natural or artificial short peptide motifs, and whose properties are determined by the sequence of the AAs and ncAAs in these motifs (Connor and Tirrell, 2007; Israeli et al., 2019; Varanko et al., 2020; Chang et al., 2021). Although the incorporation of ncAAs in numerous PBPs can be used to expand their range of properties and functions (Connor and Tirrell, 2007; Israeli et al., 2019), we focus here on the effect of ncAA incorporation on the properties of two families of bio-inspired, thermo-responsive PBPs: elastin-like polypeptides (ELPs) and resilin-like polypeptides (RLPs). ELPs are arguably the most well-studied family of artificial PBPs that comprise multiple repeats (typically 5–200 repeats) of the VPGXG motif (variations of the tropoelastin-derived VPGVP motif), where X is permissive to any amino acid (AA) except proline (MacEwan and Chilkoti, 2010). ELPs undergo a reversible soluble-to-insoluble phase transition at their lower critical solution temperature (LCST), often termed the transition temperature, which depends on the ELP composition and is maintained in ELP fusion proteins and conjugates (MacEwan and Chilkoti, 2010; Amiram et al., 2013; MacEwan and Chilkoti, 2014; Despanie et al., 2016). Extensive characterization of the mechanism underlying this phase transition permits the prediction and tuning of the LCST of the ELP, primarily by varying the hydrophobicity of the X-guest residue and the molecular weight (MW) of the ELP (Meyer and Chilkoti, 2004; McDaniel et al., 2013). RLPs are artificial, repetitive PBPs, comprising sequences identical to, or inspired by, motifs found in natural resilin proteins (Balu et al., 2021). RLPs exhibit a temperature-responsive behavior characterized by an upper critical solution temperature (UCST), with some variants also showing an additional LCST-type phase transition. The role of aromatic AAs in the UCST- and LCST-type transition behaviors appears to have particular significance (Martin and Mittag, 2018). The LCST of ELPs is affected mainly by hydrophobic interactions (Pak et al., 2016; Ruff et al., 2018) but may also be affected by hydrogen bonding and  $\pi$ - $\pi$  stacking of the aromatic side-chains (Taylor et al., 2020). In contrast, several studies have demonstrated that the UCST transition appears to be strongly influenced by the fraction, position, and identity of aromatic AAs and their interactions with other AAs (e.g., cation- $\pi$  interactions), as well as by the overall hydrophobicity (Joseph et al., 2021).

In this work, we have sought to characterize the substrate specificities of 11 of our previously described, MAGE-evolved aaRS variants using 38 different aromatic ncAAs. We identified efficient aaRS variants that are capable of incorporating 15 of these ncAAs in up to 30 instances in a single ELP chain. We applied this information to produce ELP and RLP libraries bearing multiple instances of each ncAA and characterized the effect of the newly introduced chemical group on the properties of these PBPs.

## MATERIALS AND METHODS

### Materials

ncAAs were purchased from various vendors, as indicated in **Supplementary Table S1**. Plasmid purification was conducted with Plasmid HiYield mini-prep (RBC Bioscience). SDS solution was purchased from Bio-Rad. Anhydrotetracycline hydrochloride was purchased from Sigma-Aldrich. Arabinose and 2xYT media were purchased from Tivan Biotech, and IPTG from Biolab-Biology. C321.  $\Delta A$  (Isaacs lab) was a gift from Farren Isaacs (Addgene plasmids # 73,581). Isomerization experiments were performed with a 365 nm UV lamp (VL-6.LC, 12W, Vilber Lourmat) and blue mounted LEDs: 405 nm, 1,000 mW, 800 mA (M405L4, Thor-Labs). Plasmids encoding the pAcFRS.1.t1, pAcFRS.2.t1, pAzFRS.1.t1, pAzFRS.2.t1, Mut1-RS, Mut2-RS, and AzORS-4 can be obtained from Addgene.

### Analysis of GFP Expression by Intact-Cell Fluorescence Measurements

For 96-well plate-based assays, strains harboring plasmids encoding the orthogonal translation systems and ELP-GFP reporter plasmids were inoculated with transformed cells from either a fresh agar plate or from stocks stored at  $-80^{\circ}\text{C}$ , and grown to confluence overnight. Cultures were then inoculated at a 1:30 dilution in 2xYT medium supplemented with kanamycin ( $30\ \mu\text{g ml}^{-1}$ ), chloramphenicol ( $25\ \mu\text{g ml}^{-1}$ ), and the ncAA (0.25 mM of ncAAs 1–5 or 1 mM of all other ncAAs). The expression of the aaRS was induced by adding arabinose (0.2%). Cells were allowed to grow at  $34^{\circ}\text{C}$  to an  $\text{OD}_{600}$  of 0.5–0.8 in a shaking plate incubator at 550 rpm ( $\sim 5$  h) and GFP expression was induced by adding anhydrotetracycline ( $60\ \text{ng ml}^{-1}$ ). Following overnight expression, the cells were centrifuged at 4,000 rpm for 3 min, the supernatant medium was removed, and the cells were resuspended in PBS. GFP fluorescence was measured on a Biotek spectrophotometric plate reader by using excitation and emission wavelengths of 485 nm and 528 nm, respectively. Fluorescence signals were normalized by dividing the fluorescence counts by the  $\text{OD}_{600}$  reading. Shown are representative results of at least three different assays conducted for each reporter protein and expression condition.

### ELP Expression and Purification

Before batch expression, starter cultures (1:40 v/v of final expression volume) of 2xYT media, supplemented with

kanamycin ( $30\ \mu\text{g ml}^{-1}$ ) and chloramphenicol ( $25\ \mu\text{g ml}^{-1}$ ), were inoculated with transformed cells from either a fresh agar plate or from stocks stored at  $-80^{\circ}\text{C}$ , incubated overnight at  $34^{\circ}\text{C}$  while shaking at 220 rpm, and transferred to expression flasks containing 2xYT media, antibiotics, arabinose (0.2%), and the ncAA (0.25 mM of ncAAs 5 or 1 mM of all other ncAAs). Cells were allowed to grow to an  $\text{OD}_{600}$  of 0.5–0.8 and protein expression was induced with isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG, 1 mM). The cells were harvested 24 h after inoculation by centrifugation at 4,000 g for 30 min at  $4^{\circ}\text{C}$ . The cell pellet was then resuspended by vortex in milli-Q water ( $\sim 4$  ml) and either stored at  $-80^{\circ}\text{C}$  or purified immediately. For purification, resuspended pellets were lysed by ultrasonic disruption (18 cycles of 10 s sonication, separated by 40 s intervals of rest). Poly (ethyleneimine) was added (0.2 ml of a 10% solution) to each lysed suspension before centrifugation at 4,000 rpm for 15 min at  $4^{\circ}\text{C}$  to separate cell debris from the soluble cell lysate. All ELP constructs were purified by a modified inverse transition cycling (ITC) protocol (Hassouneh et al., 2010) consisting of multiple “hot” and “cold” spins by using sodium chloride to trigger the phase transition. Before purification, the soluble cell lysate was incubated for up to 2 min at  $42$ – $65^{\circ}\text{C}$  to denature the native *E. coli* proteins. The cell lysate was then cooled on ice, centrifuged for 2 min at  $\sim 14,000$  rpm, and the pellet was discarded. For “hot” spins, the ELP phase transition was triggered by adding sodium chloride to the cell lysate or to the product of a previous cycle of ITC at a final concentration of up to  $\sim 5$  M. The solutions were then centrifuged at  $\sim 14,000$  rpm for 10 min and the pellets were resuspended in milli-Q water, after which a 2 min “cold” spin was performed without sodium chloride to remove denatured contaminants. Additional rounds of ITC were conducted as needed using a saturated solution of sodium chloride until sufficient purification was achieved. Purified proteins were visualized on SDS-PAGE.

Protein concentrations were calculated by measuring the  $\text{OD}_{280}$  of the purified protein according to the extinction coefficients in **Supplementary Table S2**. Predicted LogD values of the ncAAs were determined using ChemAxon logD predictor (Kubyskhin, 2021).

### RLP Expression and Purification

Batch expression was performed as described above for ELP production. The cells were harvested 24 h after inoculation by centrifugation at 4,000 g for 30 min at  $4^{\circ}\text{C}$ . Purification was performed according to a previously described protocol (Dzuricky et al., 2020). Purified proteins were visualized on SDS-PAGE. Protein concentrations were calculated by measuring the  $\text{OD}_{280}$  of the purified protein in 8M urea according to the extinction coefficients in **Supplementary Table S2**.

### Intact Mass Measurements

The intact mass of ELP and RLP variants was measured using MALDI-TOF/TOF autoflex speed at the Ilse Katz Institute for Nanoscale Science and Technology (Ben-Gurion University of the Negev). Spectrum analysis was performed by the Flexanalysis software.

**TABLE 1** | Annotations of specific mutations in evolved aaRS variants compared with the WT *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase (MjTyrRS) sequence. In addition to the indicated mutations, all mutants harbored the R257G and D286R mutations, which have been shown to improve tRNA binding.

Position	32	65	107	108	109	158	159	162	167
aaRS									
Tyr-RS	Y	L	E	F	Q	D	I	L	A
PacFRS.1.t1	L	L	E	F	Q	G	C	R	D
PacFRS.2.t1	L	V	E	F	Q	G	C	R	D
PazFRS.2.t1	L	L	T	Y	M	G	C	R	A
Mut1-RS	L	L	E	F	Q	S	M	K	H
Mut2-RS	L	V	E	F	Q	G	A	E	H
Mut3-RS	T	V	A	Y	M	G	C	R	D
Mut4-RS	L	V	E	F	Q	G	M	S	H
AzoRS-1	G	V	E	F	Q	G	Y	S	F
AzoRS-2	L	V	S	V	S	G	Y	S	F
AzoRS-3	L	V	N	V	L	G	Y	S	F
AzoRS-4	G	V	E	F	Q	G	Y	R	A

## Phase Transition Analysis

To characterize the inverse transition temperature of the ELP and RLP variants, the OD<sub>600</sub> of the ELP solution was monitored as a function of temperature, with heating and cooling performed at a rate of 1°C min<sup>-1</sup> on a UV-vis spectrophotometer equipped with a multicell thermoelectric temperature controller (Thermo Scientific).

## Dynamic Light Scattering (DLS) Analysis

ELP and RLP self-assembly was analyzed using a Zetasizer Nano ZS (Malvern Pananalytical). For each sample, 11–17 acquisitions (determined automatically by the instrument) were obtained at 5°C or 60°C for ELPs or RLPs, respectively. Populations comprising less than 1% of the total mass (by volume) were excluded from the analysis.

## Circular Dichroism (CD) Analysis

The secondary structure of ELPs was studied using an Jasco J-715 spectropolarimeter (Tokyo) equipped with a PTC-348WI temperature controller, using a 1-mm quartz cuvette by scanning from 280 to 180 nm at 5°C. Purified constructs were diluted to 5 μM in water. Data were considered for analysis whenever the Dynode voltage was below 800 V.

## RESULTS AND DISCUSSION

### Profiling the Substrate Specificity of High-Performance aaRSs

We first profiled the substrate specificities of our previously selected aaRS variants: pAcFRS.1.t1, pAcFRS.2.t1, and pAzFRS.2.t1 were originally selected for pAcF and pAzF incorporation (Amiram et al., 2015), MutRS-1–MutRS-4 were selected for pPR incorporation (Hadar et al., 2021), and AzoRS-1–AzoRS-4 were selected for AzoPhe incorporation (Israeli et al., 2021). A summary of the mutations in the AA binding site of these aaRS variants, as compared with the parent MjTyrRS, which

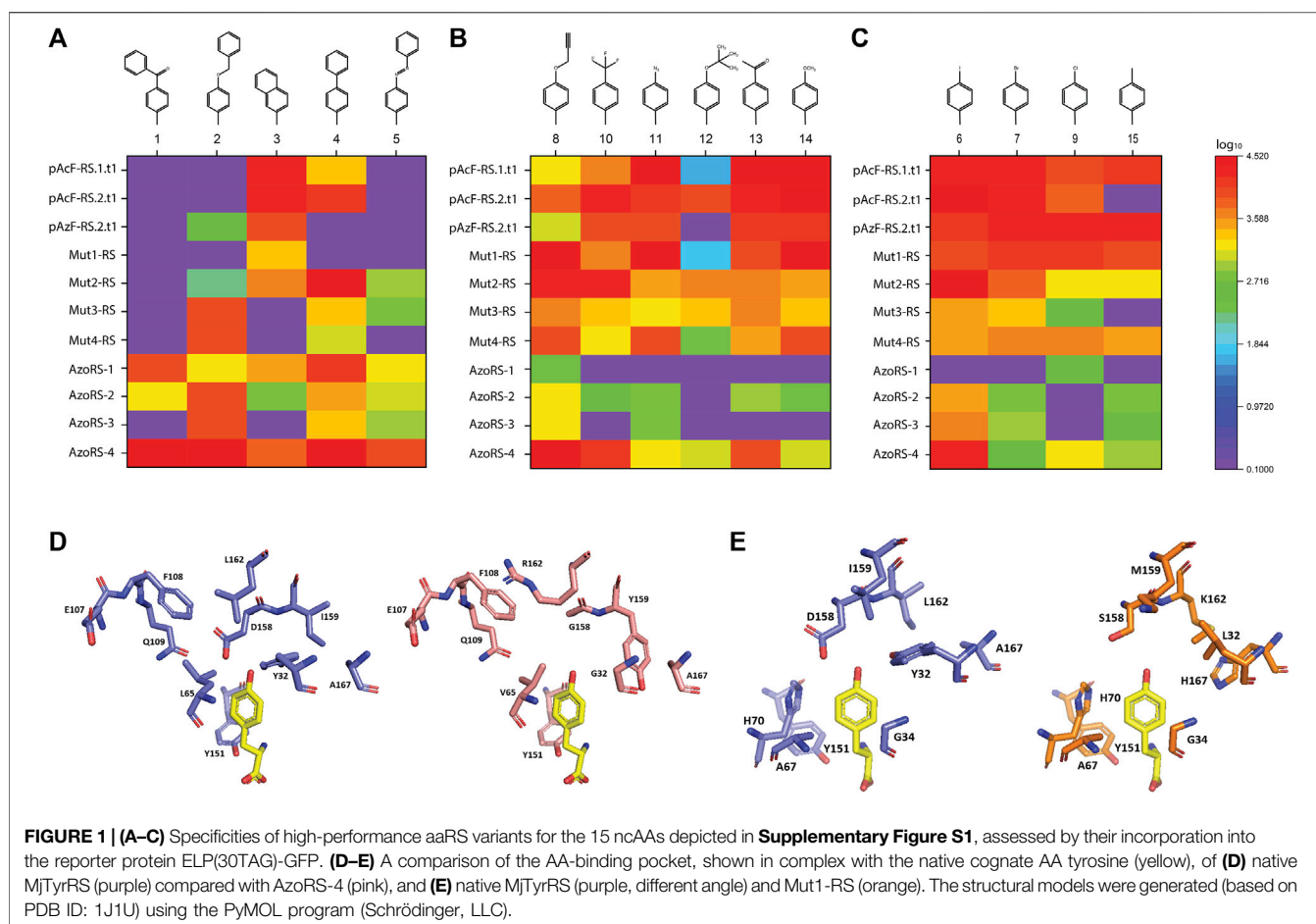
natively recognizes tyrosine, is provided in **Table 1**. To identify high-performance aaRSs capable of efficient incorporation of multiple instances of the ncAAs per protein, we assessed the ability of the aaRSs to support the incorporation of either 10 or 30 instances of each ncAA per protein via the expression of two reporter proteins: ELP(10TAG)-GFP and ELP(30TAG)-GFP (**Supplementary Table S3**). These reporter proteins were produced in the presence of each aaRS-tRNA pair using the *Escherichia coli* strain C321.ΔRF1, which lacks all the native TAG codons and the associated release factors (RF-1) (Lajoie et al., 2013).

Notably, some high-performance aaRS variants, particularly pAcFRS.1.t1 and pAzFRS.2.t1, also exhibit a relatively high incorporation of natural AAs in the absence of their cognate ncAA. This property results in high levels of “background” expression of the ELP-GFP reporters in the absence of a ncAA, although, when present, they accurately incorporate their cognate ncAA (Amiram et al., 2015; Hadar et al., 2021; Israeli et al., 2021). Such background expression can be reduced in conditions that highlight the differences in aaRS efficiencies, such as by increasing the number of TAG codons in the reporter protein, since the aaRSs incorporate the ncAAs more efficiently than the natural AAs. Specifically, relative protein production in the presence of the cognate ncAA, as compared with the absence of any ncAA, is lower for ELP(30TAG)-GFP than for ELP(10TAG)-GFP (Amiram et al., 2015; Hadar et al., 2021). Consequently, we conducted our initial evaluation of multi-site ncAA incorporation using the ELP(30TAG)-GFP reporter to eliminate the possibility of masking the production of ncAA-containing reporter proteins by background expression in the absence of the ncAA.

The expression of the ELP(30TAG)-GFP reporter protein in the presence of 38 different ncAAs revealed that 15 of the ncAAs (including four of the cognate ncAAs for which the aaRS variants were initially selected; **Supplementary Figure S1**) are efficiently incorporated by one or more of the aaRS variants (**Figures 1A–C**, **Supplementary Figure S2**). Notably, all but one (ncAA 3) of the 15 ncAAs that were successfully incorporated are phenylalanine derivatives that harbor 4' (para) substituents. The expression of the ELP(10TAG)-GFP reporter protein in the presence of the same 15 ncAAs confirms the high background incorporation previously observed for some of the variants and reveals that some aaRS variants appear to have increased efficiency in the expression of this reporter, as compared with ELP(30TAG)-GFP (**Supplementary Figure S3**). This finding is in agreement with previous studies demonstrating that differences in aaRS efficiencies are revealed by increasing the number of TAG codons per protein (Amiram et al., 2015; Hadar et al., 2021; Israeli et al., 2021).

An examination of the specificities of the aaRS variants (indicated as heat maps in **Figures 1A–C**) reveals several distinctions in the substrate profiles of some of the aaRSs. First, generally, the larger ncAAs (1–5) were suitable substrates for AzoRS-1-4, which were selected for the incorporation of 5 (AzoPhe). In contrast, smaller ncAAs were more efficiently recognized by the remaining aaRSs, which were selected for the incorporation of pAcF (13), pAzF (11), and pPR





(8). In addition, while some aaRSs, such as pAcFRS.1.t1, pAcFRS.2.t1, and AzoRS-4, exhibited a broad substrate spectrum and could efficiently accommodate >9 different ncAAs, other aaRSs, such as Mut3-RS, Mut4-RS, AzoRS-1, and AzoRS-3, were more selective and can efficiently aminoacylate only 1–3 ncAAs. We also profiled the substrate specificities of an additional aaRS, namely pAzFRS.1.t1, using our panel of 38 ncAAs. We found that it is exceptionally specific for pAzF and excluded all other ncAAs in our panel (**Supplementary Figure S4**), in accordance with our previous analysis of this aaRS using another panel of ncAAs (Amiram et al., 2015). Finally, mutually orthogonal aaRS–ncAA pairs, such as AzoRS-1:ncAA1 and Mut2-RS:ncAA6, could be identified from these analyses. Such orthogonal pairs may assist in constructing multiple orthogonal translation systems for the incorporation of two or more ncAAs within a single protein.

A comparison of the mutations in each of the aaRS variants (**Table 1**) provides information that may aid in the construction of future libraries for aaRS evolution. First, the differences and similarities in the mutations found in the aaRS variants revealed patterns in the types of AAs substituted at certain positions (**Figures 1D,E**). For example, Y32 and D158, which originally form hydrogen bonds with the tyrosine hydroxyl group, were mutated in all variants to a smaller

AA—leucine, threonine, or glycine (for Y32) or glycine and serine (for D158), enlarging the ncAA binding pocket. Similarly, I159 was mutated in all variants to either smaller AAs, such as cysteine or alanine in aaRSs selected for the incorporation of 8, 11, or 13, or to tyrosine in variants selected for the incorporation of 5, where  $\pi$ - $\pi$  interactions between the side chain of tyrosine 159 and the additional phenyl ring in ncAAs 1–5 may facilitate their binding. Finally, A167 was frequently replaced with a charged AA or with phenylalanine in aaRSs that are compatible with ncAAs bearing the relatively small, or phenyl-based substituents, respectively. Second, although 12 AAs in the AA-binding pocket were targeted for diversification in the selections of all aaRS variants, mutations in only nine of these AAs were evident in the selected aaRSs. It is possible that any changes in the remaining three AAs—G34, A67 and Y151—which remained unaffected in all of our variants, are not beneficial to this group of ncAAs or deleterious to the binding of any AA. A mutagenesis of G34 and A67, which are relatively small AAs, may increase the size of the AA-binding pocket and, thereby, interfere with the binding of aromatic AAs and ncAAs. The mutagenesis of Y151, which forms a hydrogen bond with the AA/ncAA backbone amino group (Kobayashi et al., 2003), may prevent AA binding. Finally, it may be that other AAs in

the AA-binding pocket, such as H70, must be mutagenized to enable the efficient incorporation of aromatic ncAAs bearing meta or ortho substituents (Neumann et al., 2008; Sakamoto et al., 2009; Hauf et al., 2017).

Our results suggest that the aaRS variants analyzed in this study may also efficiently incorporate other phenylalanine derivatives that harbor 4' (para) substituents, which were not included in the current panel. It is also possible that some of these aaRS variants can efficiently incorporate one or a few instances of other aromatic ncAAs, including the remaining 23 ncAAs that we examined in this study and for which we did not identify an efficient aaRS for the incorporation of 10 or 30 instances per protein. However, the experimental conditions for the identification of such low-to-medium efficiency aaRS:ncAA pairs must be carefully selected, since the background expression of reporter protein in the absence of the ncAA can be high, and it increases with decreasing number of TAG codons per protein. Such analyses may be facilitated by decreasing aaRS production, for example by expression of a chromosomally integrated aaRS, or by reducing the copy number of the plasmid encoding the aaRS, or the amount of the respective inducer of aaRS expression (Amiram et al., 2015; Hadar et al., 2021). Finally, the high-performance aaRS:ncAA pairs identified in this study may also be utilized to improve the expression of ncAA-containing proteins and PBPs in other *E. coli* strains, such as derivatives of the commonly utilized BL21 strain, although competition with RF-1 is expected to reduce the yield of the full-length protein in these strains (Hadar et al., 2021).

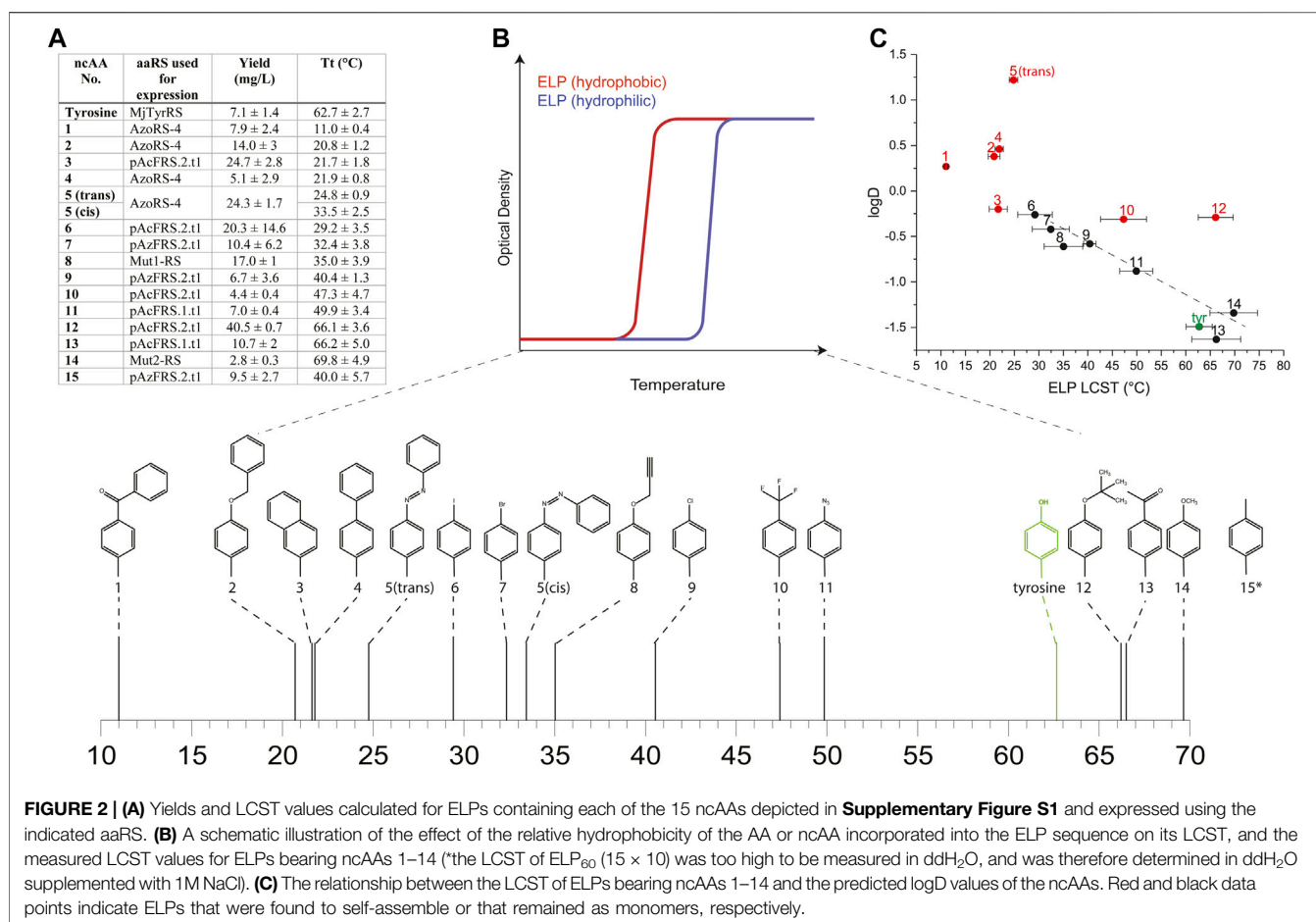
## Multi-Site ncAA Incorporation in ELPs

The characterization of ncAA-containing ELPs requires a suitable ELP variant, such that ncAA incorporation will result in ELPs with measurable LCSTs (i.e., ~10–90°C). However, the LCST of the above-mentioned ELP-GFP reporters is predicted to be either above (in GFP-fused proteins) or below (in unfused proteins) the measurable temperature range used for LCST analysis and characterization. Therefore, we selected a gene that encodes another ELP variant, termed ELP<sub>60</sub>(10TAG), which we previously utilized to determine the effect of the incorporation and isomerization of 5 on the ELP LCST (Israeli et al., 2021). The protein encoded by the ELP<sub>60</sub>(10TAG) gene is based on previously described hydrophilic ELPs, which have a high LCST (>90°C for a 25 μM solution (Meyer and Chilkoti, 2002)) and are composed of glycine and alanine, alternating in the X-guest residue position (Supplementary Table S3). We chose this variant since we hypothesized that the multi-site incorporation of aromatic ncAAs is expected to dramatically reduce the LCST of the ELP, based on the observation that the incorporation of aromatic residues, such as tyrosine, phenylalanine, tryptophan (Urry et al., 1992; Seifried et al., 2018), and 5 (Israeli et al., 2021), reduce the LCST, as compared with the incorporation of more hydrophilic AAs. Below, we name proteins according to the identity of the ncAA incorporated in the TAG codons. For example, ELP<sub>60</sub>(1 × 10) is the protein product of the ELP<sub>60</sub>(10TAG) gene, wherein 1 is incorporated in 10 encoded TAG codons.

We expressed proteins from the ELP<sub>60</sub>(10TAG) gene in the C321. ΔRF1 *E. coli* strain by using either aaRSs suitable for each of the 15 ncAAs, or the native MjTyrRS (for the incorporation of tyrosine), as designated in Figure 2A. We quantified protein yields by using small-batch expression and found that the yields of ncAA-containing ELP<sub>60</sub>(10TAG) proteins ranged from ~3 to 40 mg/L, as compared with the yield of tyrosine-containing ELP<sub>60</sub>(tyrosine×10), which was ~7 mg/L (Figure 2A). The yield of purified proteins is likely affected both by the purification process, in which more hydrophobic ELPs generally tend to precipitate more easily, and by the efficiency of the aaRS. Protein purity and MW were determined by SDS-PAGE and intact mass-spectrometry (Supplementary Table S4 and Supplementary Figure S5).

To determine the effect of ncAA incorporation on ELP properties, we first examined the effect of ncAA identity on the LCST of the ELPs in ddH<sub>2</sub>O at a fixed concentration of 25 μM (Figures 2A,B, Supplementary Figure S6). The incorporation of 10 instances of each ncAA in the ELP<sub>60</sub>(10TAG) protein profoundly affected the LCST, which ranged from ~11 to 70°C for ELP<sub>60</sub>(1 × 10)–ELP<sub>60</sub>(14 × 10), while the LCST of ELP<sub>60</sub>(15 × 10) was above 90°C in these conditions. We next examined the effect of ncAA incorporation on the self-assembly of this ELP family, motivated by an earlier finding that the incorporation of 5 engendered the self-assembly of ELP<sub>60</sub>(5 × 10) into thin 2-dimensional sheets. An assessment of self-assembly was conducted by a dynamic light scattering (DLS) analysis of a solution of each protein (25 μM in ddH<sub>2</sub>O) at a temperature below the LCST of all ELPs (5°C). The DLS analysis indicated that ncAAs 1–5, 10, 12, and 15 appeared to engender ELP self-assembly (Supplementary Figure S7). We note that previously reported ELP<sub>60</sub>(1 × 10) produced using a first-generation aaRS did not self-assemble, perhaps due to differences in mis-incorporation tendencies or in the purification protocol (Israeli et al., 2021). The effect of ncAA incorporation on the LCST of the ELPs can be viewed as an extension to Urry's AA hydrophobicity scale, which is based on the effect of natural AAs on the ELP LCST (Urry et al., 1992). Therefore, we examined the relationship between the ELP LCST and predicted logD values of the ncAAs at pH 7 (computed using ChemAxon, which was previously shown to generate values that correlated well with the experimental logD measurements of AAs and similar ncAAs (Kubyshkin, 2021)). While a poor correlation was observed overall between the LCST of the ELPs and the computed logD values, a strong correlation ( $R^2 = 0.92$ ) was observed for those ELPs that did not self-assemble in the solution (Figure 2C). This finding could be explained by the fact that self-assembly effectively alters the local concentration of the ELPs—and, therefore, its apparent LCST—as was suggested in other self-assembled ELP systems (Dreher et al., 2008; Prhashanna et al., 2019).

Next, we examined the secondary structure of ncAA-containing ELPs by using circular dichroism (CD) spectroscopy at a temperature below the LCST of all ELPs (5°C). All ELPs, including the control ELP<sub>60</sub>(Tyrosine×10), showed negative peaks at around ~190 nm, and most ELPs exhibited an additional negative peak at ~220 nm; these peaks

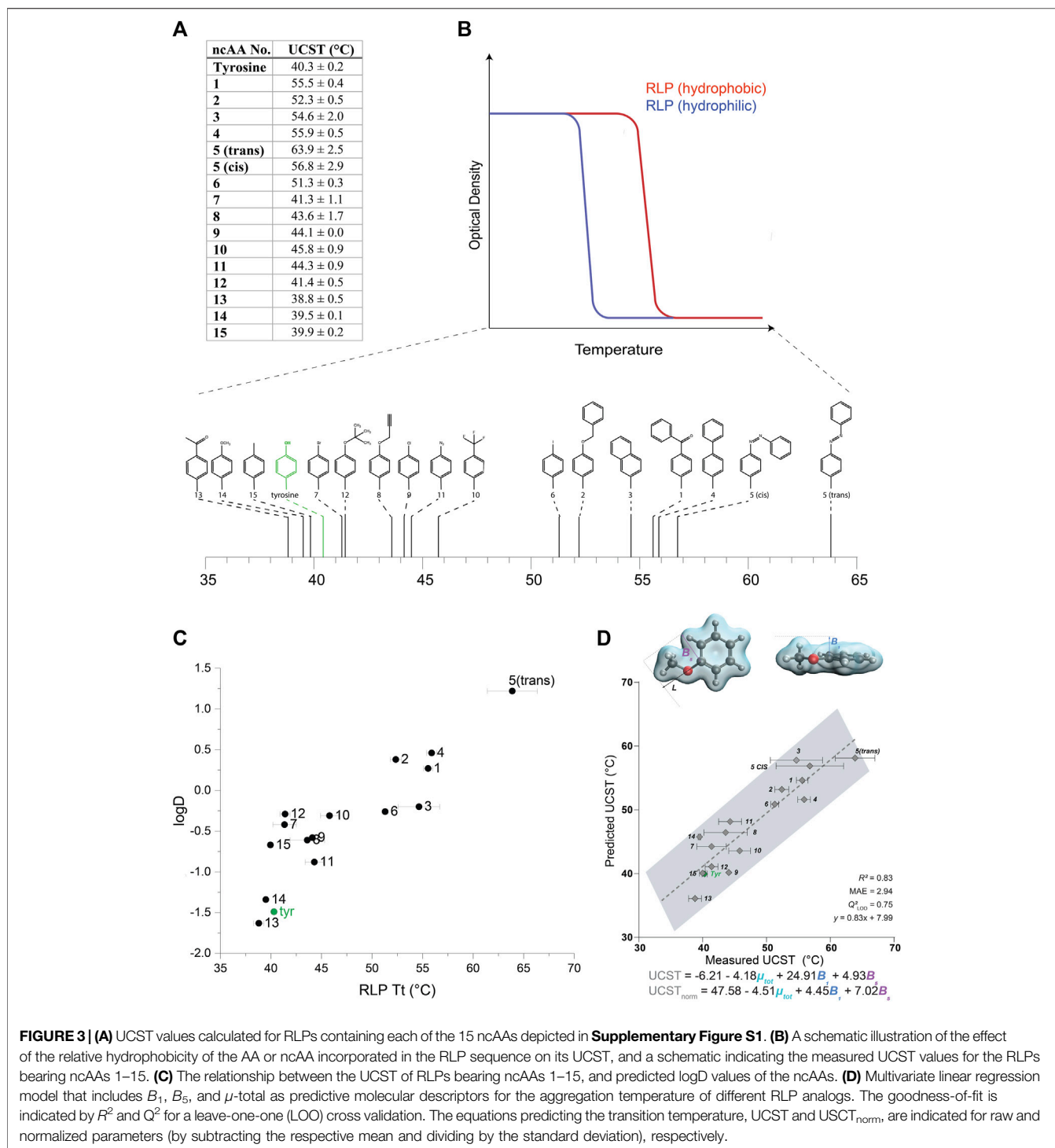


are characteristic of random-coil and  $\beta$ -turn structures, respectively (**Supplementary Figure S8**) (Janib et al., 2014). The magnitude and ratio of these peaks varied for each ELP, indicating that ncAA incorporation can also affect the secondary structure of these proteins. However, no correlation was found between these structural properties, the LCST, or the self-assembly propensity of the ELPs, as determined in these conditions.

Precise tuning of the properties, such as the LCST, of ELPs (and other PBPs), either as single- or multi-block polymers or as fusions with other proteins and PBPs (e.g., ELP–RFP fusions), is required in many applications (McDaniel et al., 2013), such as for drug delivery (Bhattacharyya et al., 2016; Varanko et al., 2020; Balu et al., 2021), tissue engineering (Bhattacharyya et al., 2016; Balu et al., 2021), sensing (Dhandhukia et al., 2013; Dhandhukia et al., 2017), metabolic engineering (Dzuricky et al., 2020), and protein purification (Hassouneh et al., 2010; Bhattacharyya et al., 2016). Although factors such as salinity and MW can also be used to tune the LCST or UCST of PBPs, such adjustments are often limited by the specific application (McDaniel et al., 2013). Previous studies have demonstrated that either installing or incorporating unnatural chemical groups can modulate ELP properties, such as LCST. For example, a global replacement

of proline residues with various proline analogs in ELPs resulted in changes to the secondary structure of the ELPs and altered their LCST (Kim et al., 2005; Kim et al., 2006; Catherine et al., 2015). Similarly, the chemical modification of methionine residues encoded in the guest-residue position in ELPs also enabled the tuning of the LCST (Kramer et al., 2015; Petitdemange et al., 2017; Rosselin et al., 2019) and triggered the self-assembly of di-block ELPs (Dai et al., 2021). However, until recently, the moderate efficiency of orthogonal translation systems has limited the production of PBPs that contain multiple instances of ncAAs by genetic code expansion, and thereby hindered the analysis of the resulting properties (Wang et al., 2021).

Here, we demonstrate that high-performance aaRSs enable the incorporation of 15 different aromatic ncAAs in ELPs—and, in turn, the production of a family of ELPs with LCSTs that vary within a wide temperature range—from a single DNA template. In addition, the analysis of these ELPs revealed that ncAA incorporation affects the secondary structure and self-assembly of the ELPs. Importantly, as compared with global replacement or chemical modification, which modify every instance of a particular AA, incorporating ncAAs using high-performance orthogonal translation components enables the multi-site



**FIGURE 3 | (A)** UCST values calculated for RLPs containing each of the 15 ncAAs depicted in **Supplementary Figure S1**. **(B)** A schematic illustration of the effect of the relative hydrophobicity of the AA or ncAA incorporated in the RLP sequence on its UCST, and a schematic indicating the measured UCST values for the RLPs bearing ncAAs 1–15. **(C)** The relationship between the UCST of RLPs bearing ncAAs 1–15, and predicted logD values of the ncAAs. **(D)** Multivariate linear regression model that includes  $B_1$ ,  $B_2$ , and  $\mu$ -total as predictive molecular descriptors for the aggregation temperature of different RLP analogs. The goodness-of-fit is indicated by  $R^2$  and  $Q^2$  for a leave-one-one (LOO) cross validation. The equations predicting the transition temperature, UCST and  $UCST_{norm}$ , are indicated for raw and normalized parameters (by subtracting the respective mean and dividing by the standard deviation), respectively.

incorporation of the selected unnatural side-chain alongside the entire set of the 20 natural AAs. The distinct advantage of this approach is particularly valuable in the design of ELP-fusion proteins wherein unintended AA modifications can result in suboptimal PBP properties or reduce or abolish the activity of the fused protein (Costa et al., 2019; Vanderschuren et al., 2022).

### Multi-Site ncAA Incorporation in RLPs

While a few previous studies have reported the effects of ncAA incorporation or side-chain modifications on the properties and behaviors of ELPs (Kim et al., 2005; Kim et al., 2006; Catherine et al., 2015; Kramer et al., 2015; Petitdemange et al., 2017; Rosselin et al., 2019), we are not aware of any similar analyses



of such alterations in other PBPs. To examine the effect of ncAA incorporation on the UCST-type phase transition behavior, we turned our attention to a recently described family of RLPs constructed from tandem repeats of the GRGDSPYS peptide (Dzuricky et al., 2020). We produced a series of proteins from the GRGDSPYS<sub>40</sub>(6TAG) gene (**Supplementary Table S3**), with either tyrosine (which results in the original “wild type” sequence) or ncAAs 1–15 encoded in the TAG codons. Protein purity and MW were determined by SDS-PAGE and intact mass-spectrometry (**Supplementary Table S5** and **Supplementary Figure S9**).

We then examined the effect of the different ncAAs on the UCST of this set of RLPs (**Figures 3A,B**, **Supplementary Figure S10**). The incorporation of more hydrophilic ncAAs, as estimated by predicted logD values, typically resulted in a reduction of the UCST, such that the solubilization of the protein occurred at lower temperatures (**Figures 3A,B**) (we note that all the RLP variants appeared to self-assemble above their UCST, **Supplementary Figure S11**). As expected, this effect was generally opposite to that displayed by the ELPs, wherein the more hydrophilic ncAAs increased the LCST, such that the aggregation of the protein occurred at higher temperatures. Indeed, the correlation ( $R^2 = 0.77$ ) between the predicted logD values and the UCST of the RLPs indicated that increasing ncAA hydrophobicity generally increases the UCST of the RLP (**Figure 3C**).

Although hydrophobic interactions broadly explain the aggregation of different RLP analogs at various temperatures, it is well established that other interactions (e.g.,  $\pi$ - $\pi$ , cation- $\pi$ , hydrogen bonds) also strongly affect the UCST. Therefore, we were interested in utilizing the data collected on this large set of aromatic substitutions to gain a more in-depth understanding of their effect on the phase-transition behavior at the molecular level. To this end, a multivariate model was identified to elucidate the specific molecular features of the RLP derivatives that may influence their UCST. Based on the versatile modifications on the aromatic ring of the tyrosine side chain, we evaluated the steric, electronic, and stereo-electronic molecular descriptors, such as Sterimol parameters (Verloop et al., 1976), vibration frequencies and intensities (Milo et al., 2014), dipole moment and NBO charges (Glendening et al., 2012) (**Supplementary Appendix 1**). The best-fitted model included the total dipole moment of the side-chain and two steric Sterimol parameters, B1 and B5, representing the minimal and maximal width of the substituent, respectively (**Figure 3D**). Overall, wider analogs with decreased polarity, reflected as larger B1 and B5 values and a smaller total dipole variable, led to an increased transition temperature. We speculate that the decreased polarity could reflect hydrophobicity, whereas the width parameters could represent steric hindrance. These results suggest a decreased transition temperature for both more hydrophilic residues and for less sterically hindered substituents that do not impede hydrogen bonding with the aqueous solution.

Several studies demonstrated that the number, position, and identity of aromatic AAs have a significant effect on the UCST of RLPs and other natural and artificial intrinsically disordered proteins/regions (IDP/Rs) known to exhibit UCST-type behavior. Specifically, an extensive characterization of RLPs composed of the “GRGDSPYS” motif, and the effect of substitutions of tyrosine residues

to other aromatic or non-aromatic AAs, revealed that the UCST was substantially affected by the identity of the aromatic AA in the repeat sequence, and by an arginine-to-lysine substitution, which affects cation- $\pi$  interactions with the aromatic AAs (Dzuricky et al., 2020). Here, we show that substituting only six of the 40 tyrosine residues in the repeating GRGDSPYS<sub>40</sub> protein sequence can generate a difference of  $\sim 25^\circ\text{C}$  in the UCST of these PBPs. Likewise, it is expected that ncAA incorporation will enable the tuning of the UCST of other PBPs and IDP/Rs whose properties can be altered by aromatic AAs. For example, mutations of aromatic and positively charged AAs have been shown to modulate the UCST of several natural IDP/Rs, such as proteins derived from the Fused in Sarcoma (FUS) family (Lin et al., 2017; Wang et al., 2018), IDR regions of Ddx4 (Nott et al., 2015; Brady et al., 2017), and the intrinsically disordered RGG domain of the LAF-1 protein (a component of P granules) (Schuster et al., 2020). We expect that the ability to incorporate multiple instances of various aromatic ncAAs in these and many other PBPs and IDPs will continue to inform the molecular-level behavior of their phase transition and enable the rational tuning of their UCSTs and, thereby, also of their self-assembly propensities.

## CONCLUSION

We report a suit of high-performance aaRSs for the multi-site incorporation of 15 different aromatic ncAAs. Furthermore, our comparative analysis of 11 aaRS variants that evolved from a single ancestor (MjTyRS), using an identical mutagenesis strategy, enabled the analysis of the mutational landscape at the ncAA-binding site which can inform future aaRS evolution efforts. This analysis indicates a distinct pattern of mutations found in the ncAA-binding site and shows that the mutagenesis of only nine of the 12 residues targeted in our library designs may suffice for the selection of efficient aaRS variants for the incorporation of para-substituted aromatic ncAAs. The application of this knowledge toward the creation of more focused libraries will enable the exploration of a larger subset of the relevant theoretical library space, and thus the selection of more efficient variants (Hauf et al., 2017; Baumann et al., 2019). Employing these capabilities, we were able to provide the first investigation of the effect of various unnatural aromatic groups on the ELP LCST- and RLP UCST-type phase-transition behaviors. We show that the incorporation of aromatic ncAAs can generate, from a single DNA template, a family of ELP proteins that have a wide range (spanning  $>60^\circ\text{C}$ ) of LCSTs. Further analysis of the ELP properties reveals that ncAA incorporation affects both the secondary structure of the ELPs and, for some ncAAs, the propensity to self-assemble. Moreover, we show that aromatic ncAA incorporation can also be utilized to tune the phase-transition behavior of RLPs, and that this behavior is strongly affected by both the hydrophobicity and the size of the ncAA side-chain. Given the recognized importance of aromatic residues for the phase transition behavior—primarily of the UCST-type, but also of the LCST-type—the ability to efficiently encode aromatic ncAAs alongside the entire set of 20 natural AAs will allow access to new families of precise and chemically diverse PBPs and PBP-protein fusions. Expanding the chemical diversity of aromatic side chains incorporated in ELP- and RLP-based PBPs will permit further investigation of the effect of aromatic character on the different

phase-transition behaviors and advance our understanding of the functions and sequence determinants of LCST- and UCST-type phase transitions in natural and artificial IDPs and PBBs (Ruff et al., 2018). Beyond the specific families explored in this work, these capabilities will empower future studies of the effect of ncAA incorporation on the self-assembly propensities and morphologies of multi-block polymers composed of domains derived from ELPs, RLPs, or other PBBs. We expect that further elucidation of the sequence and molecular determinants of these and many other PBBs, IDP/Rs, and ncAAs will enable the rational design of semi-synthetic polymers with bespoke new or improved properties.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

OG and OS contributed equally. OG, OS, RA, DSS, DH, and SG carried out the experiments, HS and AM conducted the molecular modeling analysis, MA supervised the study and wrote the manuscript with feedback from all authors.

## REFERENCES

- Amiram, M., Haimovich, A. D., Fan, C., Wang, Y.-S., Aerni, H.-R., Ntai, I., et al. (2015). Evolution of Translation Machinery in Recoded Bacteria Enables Multi-Site Incorporation of Nonstandard Amino Acids. *Nat. Biotechnol.* 33 (12), 1272–1279. doi:10.1038/nbt.3372
- Amiram, M., Luginbuhl, K. M., Li, X., Feinglos, M. N., and Chilkoti, A. (2013). A Depot-Forming Glucagon-like Peptide-1 Fusion Protein Reduces Blood Glucose for Five Days with a Single Injection. *J. Control. Release* 172 (1), 144–151. doi:10.1016/j.jconrel.2013.07.021
- Balu, R., Dutta, N. K., Dutta, A. K., and Choudhury, N. R. (2021). Resilin-mimetics as a Smart Biomaterial Platform for Biomedical Applications. *Nat. Commun.* 12 (1), 149. doi:10.1038/s41467-020-20375-x
- Baumann, T., Hauf, M., Richter, F., Albers, S., Möglich, A., Ignatova, Z., et al. (2019). Computational Aminoacyl-tRNA Synthetase Library Design for Photocaged Tyrosine. *Ijms* 20 (9), 2343. doi:10.3390/ijms20092343
- Bhattacharyya, J., Bellucci, J. J., and Chilkoti, A. (2016). “Elastin-Like Polypeptides: Bio-Inspired Smart Polymers for Protein Purification, Drug Delivery and Tissue Engineering,” in *Biomaterials from Nature for Advanced Devices and Therapies* (Hoboken, NJ: John Wiley & Sons, Inc.), 106–126. doi:10.1002/9781119126218.ch7
- Brady, J. P., Farber, P. J., Sekhar, A., Lin, Y.-H., Huang, R., Bah, A., et al. (2017). Structural and Hydrodynamic Properties of an Intrinsically Disordered Region of a Germ Cell-specific Protein on Phase Separation. *Proc. Natl. Acad. Sci. U.S.A.* 114 (39), E8194–E8203. doi:10.1073/pnas.1706197114
- Catherine, C., Oh, S. J., Lee, K.-H., Min, S.-E., Won, J.-I., Yun, H., et al. (2015). Engineering Thermal Properties of Elastin-like Polypeptides by Incorporation of Unnatural Amino Acids in a Cell-free Protein Synthesis System. *Biotechnol. Bioproc E* 20 (3), 417–422. doi:10.1007/s12257-015-0190-1
- Chang, M. P., Huang, W., and Mai, D. J. (2021). Monomer-scale Design of Functional Protein Polymers Using Consensus Repeat Sequences. *J. Polym. Sci.* 59 (22), 2644–2664. doi:10.1002/pol.20210506

## FUNDING

This work was funded by the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation program, grant agreement 756996 and by the Israel Science Foundation (grant number 939\21) (to MA). MA gratefully acknowledges support from the Elaine S. and Alvin W. Wene Career Development Chair in Biotechnology Engineering. HS acknowledges support from the Emergency Postdoctoral Fellowships for Israeli Researchers in Israel.

## ACKNOWLEDGMENTS

We thank Dr. Mark Karpasas from the Ilse Katz Institute for Nanoscale Science and Technology for the professional help with the mass spectrometry experiments.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fbioe.2022.913057/full#supplementary-material>

- Chen, H., Venkat, S., McGuire, P., Gan, Q., and Fan, C. (2018). Recent Development of Genetic Code Expansion for Posttranslational Modification Studies. *Molecules* 23 (7), 1662. doi:10.3390/molecules23071662
- Chin, J. W. (2017). Expanding and Reprogramming the Genetic Code. *Nature* 550 (7674), 53–60. doi:10.1038/nature24031
- Connor, R. E., and Tirrell, D. A. (2007). Non-Canonical Amino Acids in Protein Polymer Design. *Polym. Rev.* 47 (1), 9–28. doi:10.1080/15583720601109552
- Costa, S. A., Mozhdehi, D., Dzuricky, M. J., Isaacs, F. J., Brustad, E. M., and Chilkoti, A. (2019). Active Targeting of Cancer Cells by Nanobody Decorated Polypeptide Micelle with Bio-Orthogonally Conjugated Drug. *Nano Lett.* 19 (1), 247–254. doi:10.1021/acs.nanolett.8b03837
- Dai, M., Georgilis, E., Goudounet, G., Garbay, B., Pille, J., van Hest, J. C. M., et al. (2021). Refining the Design of Diblock Elastin-Like Polypeptides for Self-Assembly into Nanoparticles. *Polymers* 13 (9), 1470. doi:10.3390/polym13091470
- Debelouchina, G. T., and Muir, T. W. (2017). A Molecular Engineering Toolbox for the Structural Biologist. *Q. Rev. Biophysics* 50, e7. doi:10.1017/s0033583517000051
- Despanie, J., Dhandhukia, J. P., Hamm-Alvarez, S. F., and MacKay, J. A. (2016). Elastin-like Polypeptides: Therapeutic Applications for an Emerging Class of Nanomedicines. *J. Control. Release* 240, 93–108. doi:10.1016/j.jconrel.2015.11.010
- Dhandhukia, J. P., Brill, D. A., Kouhi, A., Pastuszka, M. K., and MacKay, J. A. (2017). Elastin-like Polypeptide Switches: A Design Strategy to Detect Multimeric Proteins. *Protein Sci.* 26 (9), 1785–1795. doi:10.1002/pro.3215
- Dhandhukia, J., Weitzhandler, I., Wang, W., and MacKay, J. A. (2013). Switchable Elastin-Like Polypeptides that Respond to Chemical Inducers of Dimerization. *Biomacromolecules* 14 (4), 976–985. doi:10.1021/bm301558q
- Dreher, M. R., Simnick, A. J., Fischer, K., Smith, R. J., Patel, A., Schmidt, M., and Chilkoti, A. (2008). Temperature Triggered Self-Assembly of Polypeptides into Multivalent Spherical Micelles. *J. Am. Chem. Soc.* 130 (2), 687–694. doi:10.1021/ja0764862
- Dzuricky, M., Rogers, B. A., Shahid, A., Cremer, P. S., and Chilkoti, A. (2020). De Novo engineering of Intracellular Condensates Using Artificial Disordered Proteins. *Nat. Chem.* 12 (9), 814–825. doi:10.1038/s41557-020-0511-7

- Glendening, E. D., Landis, C. R., and Weinhold, F. (2012). Natural Bond Orbital Methods. *WIREs Comput. Mol. Sci.* 2 (1), 1–42. doi:10.1002/wcms.51
- Hadar, D., Gekkop, S., Vaserman, L., and Amiram, M. (2021). Efficient Incorporation of Clickable Unnatural Amino Acids Enables Rapid and Biocompatible Labeling of Proteins *In Vitro* and in Bacteria. *ChemBiochem* 22, 1379–1384. doi:10.1002/cbic.202000663
- Hassouneh, W., Christensen, T., and Chilkoti, A. (2010). Elastin-Like Polypeptides as a Purification Tag for Recombinant Proteins. *Curr. Protoc. Protein Sci.* 6 (61), 1–20. Unit 6 11. doi:10.1002/0471140864.ps0611s61
- Hauf, M., Richter, F., Schneider, T., Faidt, T., Martins, B. M., Baumann, T., et al. (2017). Photoactivatable Mussel-Based Underwater Adhesive Proteins by an Expanded Genetic Code. *ChemBiochem* 18 (18), 1819–1823. doi:10.1002/cbic.201700327
- Hayashi, A., Haruna, K. i., Sato, H., Ito, K., Makino, C., Ito, T., et al. (2021). Incorporation of Halogenated Amino Acids into Antibody Fragments at Multiple Specific Sites Enhances Antigen Binding. *ChemBiochem* 22 (1), 120–123. doi:10.1002/cbic.202000429
- Huang, Y., and Liu, T. (2018). Therapeutic Applications of Genetic Code Expansion. *Synthetic Syst. Biotechnol.* 3 (3), 150–158. doi:10.1016/j.synbio.2018.09.003
- Israeli, B., Strugach, D. S., Gekkop, S., Weber, S., Gozlan, D. S., and Amiram, M. (2021). Genetically Encoding Light-Responsive Protein-Polymers Using Translation Machinery for the Multi-Site Incorporation of Photo-Switchable Unnatural Amino Acids. *Adv. Funct. Mater.* 31, 1–14. doi:10.1002/adfm.202011276
- Israeli, B., Vaserman, L., and Amiram, M. (2019). Multi-Site Incorporation of Nonstandard Amino Acids into Protein-Based Biomaterials. *Isr. J. Chem.* 60, 1118–1128. doi:10.1002/ijch.201900043
- Janib, S. M., Pastuszka, M. F., Aluri, S., Folchman-Wagner, Z., Hsueh, P. Y., Shi, P., et al. (2014). A Quantitative Recipe for Engineering Protein Polymer Nanoparticles. *Polym. Chem.* 5 (5), 1614–1625. doi:10.1039/c3py00537b
- Johnson, J. A., Lu, Y. Y., Van Deventer, J. A., and Tirrell, D. A. (2010). Residue-specific Incorporation of Non-canonical Amino Acids into Proteins: Recent Developments and Applications. *Curr. Opin. Chem. Biol.* 14 (6), 774–780. doi:10.1016/j.cbpa.2010.09.013
- Joseph, J. A., Reinhardt, A., Aguirre, A., Chew, P. Y., Russell, K. O., Espinosa, J. R., et al. (2021). Physics-driven Coarse-Grained Model for Biomolecular Phase Separation with Near-Quantitative Accuracy. *Nat. Comput. Sci.* 1 (11), 732–743. doi:10.1038/s43588-021-00155-3
- Kim, W., Hardcastle, K. I., and Conticello, V. P. (2006). Fluoroproline Flip-Flop: Regiochemical Reversal of a Stereoelectronic Effect on Peptide and Protein Structures. *Angew. Chem. Int. Ed.* 45 (48), 8141–8145. doi:10.1002/anie.200603227
- Kim, W., McMillan, R. A., Snyder, J. P., and Conticello, V. P. (2005). A Stereoelectronic Effect on Turn Formation Due to Proline Substitution in Elastin-Mimetic Polypeptides. *J. Am. Chem. Soc.* 127 (51), 18121–18132. doi:10.1021/ja054105j
- Kobayashi, T., Nureki, O., Ishitani, R., Yaremchuk, A., Tukalo, M., Cusack, S., et al. (2003). Structural Basis for Orthogonal tRNA Specificities of Tyrosyl-tRNA Synthetases for Genetic Code Expansion. *Nat. Struct. Mol. Biol.* 10 (6), 425–432. doi:10.1038/nsb934
- Kramer, J. R., Petitdemange, R., Bataille, L., Bathany, K., Wirotius, A.-L., Garbay, B., et al. (2015). Quantitative Side-Chain Modifications of Methionine-Containing Elastin-Like Polypeptides as a Versatile Tool to Tune Their Properties. *ACS Macro Lett.* 4 (11), 1283–1286. doi:10.1021/acsmacrolett.5b00651
- Kubyskhin, V. (2021). Experimental Lipophilicity Scale for Coded and Noncoded Amino Acid Residues. *Org. Biomol. Chem.* 19 (32), 7031–7040. doi:10.1039/d1ob01213d
- Lajoie, M. J., Rovner, A. J., Goodman, D. B., Aerni, H.-R., Haimovich, A. D., Kuznetsov, G., et al. (2013). Genomically Recoded Organisms Expand Biological Functions. *Science* 342 (6156), 357–360. doi:10.1126/science.1241459
- Lee, K. J., Kang, D., and Park, H. S. (2019). Site-Specific Labeling of Proteins Using Unnatural Amino Acids. *Mol. Cells* 42 (5), 386–396. doi:10.14348/molcells.2019.0078
- Lin, Y., Currie, S. L., and Rosen, M. K. (2017). Intrinsically Disordered Sequences Enable Modulation of Protein Phase Separation through Distributed Tyrosine Motifs. *J. Biol. Chem.* 292 (46), 19110–19120. doi:10.1074/jbc.M117.800466
- MacEwan, S. R., and Chilkoti, A. (2014). Applications of Elastin-like Polypeptides in Drug Delivery. *J. Control. Release* 190, 314–330. doi:10.1016/j.jconrel.2014.06.028
- MacEwan, S. R., and Chilkoti, A. (2010). Elastin-like Polypeptides: Biomedical Applications of Tunable Biopolymers. *Biopolymers* 94 (1), 60–77. doi:10.1002/bip.21327
- Martin, E. W., and Mittag, T. (2018). Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions. *Biochemistry* 57 (17), 2478–2487. doi:10.1021/acs.biochem.8b00008
- McDaniel, J. R., Radford, D. C., and Chilkoti, A. (2013). A Unified Model for De Novo Design of Elastin-like Polypeptides with Tunable Inverse Transition Temperatures. *Biomacromolecules* 14 (8), 2866–2872. doi:10.1021/bm4007166
- Meyer, D. E., and Chilkoti, A. (2002). Genetically Encoded Synthesis of Protein-Based Polymers with Precisely Specified Molecular Weight and Sequence by Recursive Directional Ligation: Examples from the Elastin-like Polypeptide System. *Biomacromolecules* 3 (2), 357–367. doi:10.1021/bm015630n
- Meyer, D. E., and Chilkoti, A. (2004). Quantification of the Effects of Chain Length and Concentration on the Thermal Behavior of Elastin-like Polypeptides. *Biomacromolecules* 5 (3), 846–851. doi:10.1021/bm034215n
- Milo, A., Bess, E. N., and Sigman, M. S. (2014). Interrogating Selectivity in Catalysis Using Molecular Vibrations. *Nature* 507 (7491), 210, 214. doi:10.1038/nature13019
- Neumann, H., Hazen, J. L., Weinstein, J., Mehl, R. A., and Chin, J. W. (2008). Genetically Encoding Protein Oxidative Damage. *J. Am. Chem. Soc.* 130 (12), 4028–4033. doi:10.1021/ja710100d
- Nguyen, T.-A., Cigler, M., and Lang, K. (2018). Expanding the Genetic Code to Study Protein-Protein Interactions. *Angew. Chem. Int. Ed.* 57 (44), 14350–14361. doi:10.1002/anie.201805869
- Nott, T. J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., et al. (2015). Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Mol. Cell* 57 (5), 936–947. doi:10.1016/j.molcel.2015.01.013
- O'Donoghue, P., Ling, J., Wang, Y.-S., and Söll, D. (2013). Upgrading Protein Synthesis for Synthetic Biology. *Nat. Chem. Biol.* 9 (10), 594–598. doi:10.1038/nchembio.1339
- Ohtake, K., Yamaguchi, A., Mukai, T., Kashimura, H., Hirano, N., Haruki, M., et al. (2015). Protein Stabilization Utilizing a Redefined Codon. *Sci. Rep.* 5, 9762. doi:10.1038/srep09762
- Pak, C. W., Kosno, M., Holehouse, A. S., Padrick, S. B., Mittal, A., Ali, R., et al. (2016). Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Mol. Cell* 63 (1), 72–85. doi:10.1016/j.molcel.2016.05.042
- Petitdemange, R., Garanger, E., Bataille, L., Dierck, W., Bathany, K., Garbay, B., et al. (2017). Selective Tuning of Elastin-like Polypeptide Properties via Methionine Oxidation. *Biomacromolecules* 18 (2), 544–550. doi:10.1021/acs.biomac.6b01696
- Prashanna, A., Taylor, P. A., Qin, J., Kiick, K. L., and Jayaraman, A. (2019). Effect of Peptide Sequence on the LCST-Like Transition of Elastin-Like Peptides and Elastin-Like Peptide-collagen-Like Peptide Conjugates: Simulations and Experiments. *Biomacromolecules* 20 (3), 1178–1189. doi:10.1021/acs.biomac.8b01503
- Rezhdo, A., Islam, M., Huang, M., and Van Deventer, J. A. (2019). Future Prospects for Noncanonical Amino Acids in Biological Therapeutics. *Curr. Opin. Biotechnol.* 60, 168–178. doi:10.1016/j.copbio.2019.02.020
- Rosselin, M., Xiao, Y., Belhomme, L., Lecommandoux, S., and Garanger, E. (2019). Expanding the Toolbox of Chemospecific Modifications of Protein-Like Polymers at Methionine Residues. *ACS Macro Lett.* 8 (12), 1648–1653. doi:10.1021/acsmacrolett.9b00862
- Ruff, K. M., Roberts, S., Chilkoti, A., and Pappu, R. V. (2018). Advances in Understanding Stimulus-Responsive Phase Behavior of Intrinsically Disordered Protein Polymers. *J. Mol. Biol.* 430 (23), 4619–4635. doi:10.1016/j.jmb.2018.06.031
- Sakamoto, K., Murayama, K., Oki, K., Iraha, F., Kato-Murayama, M., Takahashi, M., et al. (2009). Genetic Encoding of 3-Iodo-L-Tyrosine in *Escherichia coli* for Single-Wavelength Anomalous Dispersion Phasing in Protein Crystallography. *Structure* 17 (3), 335–344. doi:10.1016/j.str.2009.01.008
- Schuster, B. S., Dignon, G. L., Tang, W. S., Kelley, F. M., Ranganath, A. K., Jahnke, C. N., et al. (2020). Identifying Sequence Perturbations to an

- Intrinsically Disordered Protein that Determine its Phase-Separation Behavior. *Proc. Natl. Acad. Sci. U.S.A.* 117 (21), 11421–11431. doi:10.1073/pnas.2000223117
- Seifried, B. M., Cao, J., and Olsen, B. D. (2018). Multifunctional, High Molecular Weight, Post-Translationally Modified Proteins through Oxidative Cysteine Coupling and Tyrosine Modification. *Bioconjugate Chem.* 29 (6), 1876–1884. doi:10.1021/acs.bioconjchem.7b00834
- Taylor, P. A., Huang, H., Kiick, K. L., and Jayaraman, A. (2020). Placement of Tyrosine Residues as a Design Element for Tuning the Phase Transition of Elastin-Peptide-Containing Conjugates: Experiments and Simulations. *Mol. Syst. Des. Eng.* 5 (7), 1239–1254. doi:10.1039/d0me00051e
- Urry, D. W., Gowda, D. C., Parker, T. M., Luan, C.-H., Reid, M. C., Harris, C. M., et al. (1992). Hydrophobicity Scale for Proteins Based on Inverse Temperature Transitions. *Biopolymers* 32 (9), 1243–1250. doi:10.1002/bip.360320913
- Vanderschuren, K., Arranz-Gibert, P., Khang, M., Hadar, D., Gaudin, A., Yang, F., et al. (2022). Tuning Protein Half-Life in Mouse Using Sequence-Defined Biopolymers Functionalized with Lipids. *Proc. Natl. Acad. Sci. U.S.A.* 119 (4). doi:10.1073/pnas.2103099119
- Varanko, A., Saha, S., and Chilkoti, A. (2020). Recent Trends in Protein and Peptide-Based Biomaterials for Advanced Drug Delivery. *Adv. Drug Deliv. Rev.* 156, 133–187. doi:10.1016/j.addr.2020.08.008
- Vargas-Rodriguez, O., Sevostyanova, A., Söll, D., and Crnković, A. (2018). Upgrading Aminoacyl-tRNA Synthetases for Genetic Code Expansion. *Curr. Opin. Chem. Biol.* 46, 115–122. doi:10.1016/j.cbpa.2018.07.014
- Verloop, A., Hoogenstraaten, W., and Tipker, J. (1976). “Chapter 4 - Development and Application of New Steric Substituent Parameters in Drug Design,” in *Drug Design*. Editor E. J. Ariens (Amsterdam: Academic Press), 165–207. doi:10.1016/b978-0-12-060307-7.50010-9
- Völler, J.-S., and Budisa, N. (2017). Coupling Genetic Code Expansion and Metabolic Engineering for Synthetic Cells. *Curr. Opin. Biotechnol.* 48, 1–7. doi:10.1016/j.copbio.2017.02.002
- Wan, W., Tharp, J. M., and Liu, W. R. (2014). Pyrrolysyl-tRNA Synthetase: An Ordinary Enzyme but an Outstanding Genetic Code Expansion Tool. *Biochimica Biophysica Acta (BBA) - Proteins Proteomics* 1844 (6), 1059–1070. doi:10.1016/j.bbapap.2014.03.002
- Wang, B., Patkar, S. S., and Kiick, K. L. (2021). Application of Thermoresponsive Intrinsically Disordered Protein Polymers in Nanostructured and Microstructured Materials. *Macromol. Biosci.* 21 (9), 1–20. doi:10.1002/mabi.202100129
- Wang, J., Choi, J.-M., Holehouse, A. S., Lee, H. O., Zhang, X., Jahnel, M., et al. (2018). A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* 174 (3), 688–699. e616. doi:10.1016/j.cell.2018.06.006
- Willis, J. C. W., and Chin, J. W. (2018). Mutually Orthogonal Pyrrolysyl-tRNA Synthetase/tRNA Pairs. *Nat. Chem.* 10 (8), 831–837. doi:10.1038/s41557-018-0052-5
- Young, D. D., and Schultz, P. G. (2018). Playing with the Molecules of Life. *ACS Chem. Biol.* 13 (4), 854–870. doi:10.1021/acscchembio.7b00974
- Young, D. D., Young, T. S., Jahnz, M., Ahmad, I., Spraggon, G., and Schultz, P. G. (2011). An Evolved Aminoacyl-tRNA Synthetase with Atypical Polysubstrate Specificity. *Biochemistry* 50 (11), 1894–1900. doi:10.1021/bi101929e
- Zhou, W., and Deiters, A. (2021). Chemogenetic and Optogenetic Control of Post-translational Modifications through Genetic Code Expansion. *Curr. Opin. Chem. Biol.* 63, 123–131. doi:10.1016/j.cbpa.2021.02.016

**Conflict of Interest:** MA, DH, DSS, and SG filed a patent related to this technology.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Gueta, Sheinenzon, Azulay, Shalit, Strugach, Hadar, Gelkop, Milo and Amiram. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.