ACS SyntheticBiology

Letter

pubs.acs.org/synthbio

Direct Polymerization of Proteins

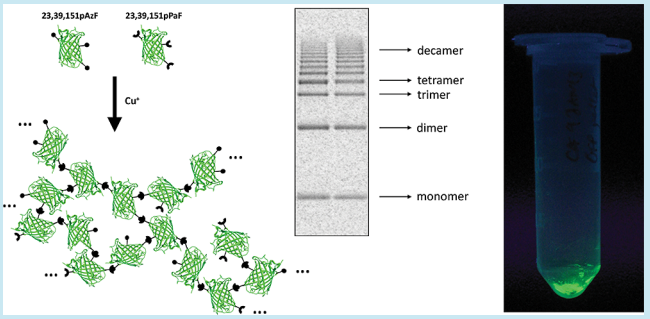
Cem Albayrak*,†,§ and James R. Swartz*,†,‡

†Department of Chemical Engineering, Stanford University, 381 North-South Mall, Stanford, California 94305, United States

‡Department of Bioengineering, Stanford University, 318 Campus Drive, Stanford, California 94305, United States

Supporting Information

ABSTRACT: We report the synthesis of active polymers of superfolder green fluorescent protein (sfGFP) in one step using Click chemistry. Up to six copies of the non-natural amino acids (nnAAs) *p*-azido-L-phenylalanine (pAzF) or *p*-propargyloxy-L-phenylalanine (pPaF) were site-specifically inserted into sfGFP by cell-free protein synthesis (CFPS). sfGFP containing two or three copies of these nnAAs were coupled by copper-catalyzed azide−alkyne cycloaddition to synthesize linear or branched protein polymers, respectively. The protein polymers retained ≥63% of their specific activity (i.e., fluorescence) after coupling. Polymerization of a concentrated solution of triply substituted sfGFP resulted in fluorescent macromolecular particles. Our method can be generalized to synthesize polymers of a protein or copolymers of any two or more proteins, and the conjugation sites can be determined exactly by standard genetic manipulation. Polymers of proteins and small molecules can also be created with this technology to make a new class of scaffolds or biomaterials.



KEYWORDS: *cell-free protein synthesis, non-natural amino acid, Escherichia coli, Click chemistry, protein polymerization, active biomaterials*

Current bioconjugation methods using natural amino acids lack precision because most proteins contain multiple surface-exposed copies of the modifiable amino acids such as lysines and cysteines. Covalent coupling of proteins via these residues would thus result in a heterogeneous mixture of products. Furthermore, it is likely that such random covalent coupling would reduce or eliminate the biological activity of the conjugated protein. Site-specific incorporation of non-natural amino acids (nnAAs) circumvents this problem by introducing unique chemical handles at specific positions on the protein to be conjugated.[1] Many nnAAs have been site-specifically incorporated for numerous applications ranging from improved biosimilars[2] and antibody-drug conjugates[3] and synthesis of fluorescent probes to studying cell biology in real time using photoactivatable proteins.[4,5]

Two different strategies exist for effective incorporation of nnAAs into proteins in *E. coli*-based protein synthesis platforms. In the first method, one of the aliphatic amino acids (i.e., isoleucines, leucines, or methionines) in all of the synthesized proteins is replaced by its analogue using strains that cannot produce the natural amino acid (i.e., auxotrophic strain). This residue-specific nnAA incorporation is achieved either by relying on the ability of the specific aminoacyl-tRNA synthetase (aaRS) to activate the analogue of its substrate, altering the amino acid binding site of the specific aaRS or by occluding the editing domain of the aaRS (e.g., the LeuRS T252Y mutant[6]). Different aliphatic nnAAs containing the alkyne[7,8] and azido[9,10] moieties, among others,[11,12] were effectively (≥85% when indicated) incorporated in proteins such as virus-like particles.[13] This method was also successfully adapted to the *E. coli*-based cell-free protein synthesis (CFPS) platform to produce modified proteins[14] and bioconjugates including surface-modified virus-like particles.[15−17] Despite the ability to effectively introduce multiple bioconjugation sites in each protein, this method suffers from the need to mutate most of the existing methionines (or isoleucines or leucines) to other amino acids and to introduce new ones on the surface of the protein. Such extensive mutagenesis is likely to adversely affect the folding and activity of the mutated protein.

In the second method, the nnAAs are introduced site-specifically at amber stop codons or more recently at quadruplet codons, using orthogonal tRNA (o-tRNA) and aaRS pairs. The orthogonal components are either derived from the tyrosyl tRNA/synthetase pair of the archaeon *Methanocaldococcus jannaschii*[18,19] or, more recently, from the pyrrolysyl tRNA/synthetase pair of the archaea *Methanosarcina barkeri* or *Methanosarcina mazei*.[20,21] More than 30 tyrosine analogues have been incorporated site-specifically into proteins using the *M. jannaschii* components, and lysine analogues containing azido[22] and alkyne functional groups,[23−25] among others, have been successfully inserted into proteins using the *Methanosarcinaceae* components. The incorporation of the lysine analogue that contains a strained alkyne is particularly exciting,

Special Issue: Cell-Free Synthetic Biology

Received: August 22, 2013
Published: November 7, 2013

   dx.doi.org/10.1021/sb400116x | *ACS Synth. Biol.* 2014, 3, 353−362

since it allows for facile bioconjugation using copper-free Click chemistry.[25] Orthogonal ribosomes that can effectively decode quadruplet and amber stop codons were evolved and used to incorporate two different nnAAs into the same protein.[26,27] Since quadruplet codons are not used by the genetic code, this powerful approach has the potential for incorporation of many different nnAAs into the same protein or polymer without interfering with the endogenous translational machinery.[28] However, despite optimization in orthogonal tRNA design[29] and expression of orthogonal components,[30,31] site-specific, *in vivo* nnAA incorporation currently suffers from low yields (2−100 mg/L) of singly substituted protein. When multiple copies of the same nnAA are incorporated, the modified protein yields are even lower (0.03−5 mg/L).[32,33]

In contrast, non-natural amino acids have been effectively incorporated into a variety of proteins in CFPS using the *M. jannaschii* orthogonal components, and moderate modified protein yields (200−660 μg/mL) were obtained by optimizing the supply of orthogonal components in the cell-free reaction.[34−36] A new method, where the o-tRNA is synthesized *in situ*, enabled the accumulation of 1.7 mg/mL of full-length and soluble, singly substituted superfolder green fluorescent protein (sfGFP).[37]

Here, we describe the synthesis of linear or branched protein polymers using this new method and the copper-catalyzed azido-alkyne cycloaddition reaction (CuAAC; i.e., Click chemistry). Multiple (up to six) copies of the nnAAs *p*-azido-L-phenylalanine (pAzF, Figure 1) or *p*-propargyloxy-L-phenylalanine (pPaF, Figure 1) were site-specifically incorporated into sfGFP. Doubly and triply substituted sfGFPs were then purified and directly coupled via CuAAC to synthesize linear or branched sfGFP polymers. Specific fluorescence measurements showed that most (≥63%) of the specific activity was retained in the sfGFP polymers. Thanks to its versatility and the absolute control over conjugation sites, we believe direct polymerization of active proteins using nnAAs can complement existing methods in biomaterial synthesis and cell-free metabolic engineering (CFME).[38,39]

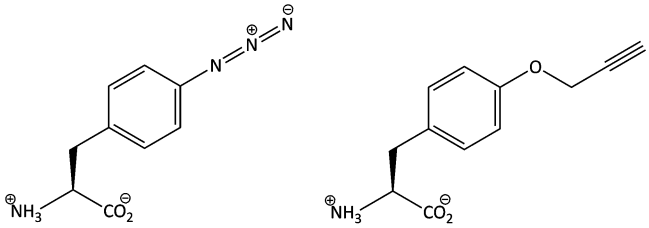


Figure 1. Structures of the non-natural amino acids (nnAAs) used in this work. (1) *p*-azido-L-phenylalanine (pAzF) and (2) *p*-propargyloxy-L-phenylalanine (pPaF).

In our new platform, the o-tRNA required for nnAA incorporation is produced simultaneously with the modified protein in the CFPS reaction. 0.9 to 1.7 mg/mL of singly substituted (i.e., modified with one pAzF or pPaF) sfGFP accumulated in the CFPS solutions, and quantitative incorporation of the nnAAs at desired positions was validated by LC/MS.[37] We next explored if this method could be used to site-specifically incorporate multiple copies of these nnAAs. Two to six amber stop codons were inserted into the sfGFP sequence, and the full-length and soluble yields of modified radiolabeled sfGFP were measured using liquid scintillation counting (LSC) and densitometry (Figure 2).[37] 290 to 740 μg/mL of full-length and soluble, doubly substituted sfGFP accumulated in small-scale CFPS reactions. These yields were reduced to 250−390 μg/mL when three pPaF's were site-specifically inserted into sfGFP; similar yields were obtained with pAzF (data not shown). About 70 μg/mL of sfGFP containing six pPaF's still accumulated in small-scale CFPS reactions. In general, the modified protein yields varied inversely with the number of nnAAs, regardless of the location of the incorporation sites along the polypeptide backbone (i.e., the position effect). This is probably due to the presence of the endogenous Release Factor 1 (RF1), which competes with the amber suppressor orthogonal tRNA bearing the nnAA; the more amber stop codons in the sfGFP mRNA, the more chances RF1 has to prematurely terminate translation.

Several of these modified and radiolabeled sfGFP proteins were then synthesized in milliliter-scale CFPS reactions, purified using StrepTactin affinity chromatography, and coupled by copper-catalyzed azide−alkyne cycloaddition (CuAAC) in three sets of analytical reactions (25 or 50 μL). The first set was conducted to test the accessibility of the nnAA incorporation sites for post-translational modification. The second and third sets were conducted to assess polymerization of these modified sfGFP proteins.

In the first set of Click reactions, the modified sfGFP with three pAzF's at positions 23, 39, and 151 (sfGFP23,39,151pAzF) was coupled to sfGFP carrying one pPaF at position 23 (sfGFP23pPaF) at two different molar ratios (Supporting Information (SI) Figure S1A). The reactions yielded only dimers, trimers, and tetramers of sfGFP as expected. Dimers and trimers with different linkages migrated differently resulting in multiple bands on the polyacrylamide gel (SI Figure S1B). The complex banding pattern also demonstrated that the CuAAC reaction was not complete, despite up to a 10-fold molar excess of sfGFP23pPaF. All of the nnAA incorporation sites were nonetheless accessible for covalent coupling with another macromolecule, and conjugation was restricted to the desired sites. If the coupling reaction was not specific to the nnAAs, then sfGFP polymers larger than the tetramer would have appeared on the polyacrylamide gel and/or the autoradiogram.

In the second set of CuAAC reactions, doubly substituted sfGFP proteins were coupled to form linear sfGFP polymers (Figure 3A). Since each protein harbored two conjugation sites, there was no intrinsic limit to product size. The doubly substituted proteins sfGFP23,39pAzF and sfGFP23,39pPaF were covalently coupled in equimolar mixtures catalyzed by either 0.25 mM or 0.5 mM each of the tetrakis Cu(I) catalyst and the TTMA enhancer (SI Figure S2). A protein ladder was obtained, containing a mixture of unreacted monomers, dimers, and polymers ranging to more than 15 conjugated sfGFP molecules (Figure 3B).

We then explored if larger sfGFP polymers could be obtained by increasing the number of conjugation sites per sfGFP monomer. In the third set of analytical CuAAC reactions, triply substituted sfGFP polymers were coupled to form branched protein polymers (Figure 4A). Larger polymers were obtained and less unreacted monomeric sfGFP remained in these reactions, despite lower initial protein concentrations than those in the second set. In fact, some of the protein polymers were so large that they did not enter the 10% polyacrylamide gel (Figure 4B). Two dimeric sfGFP bands were observed,

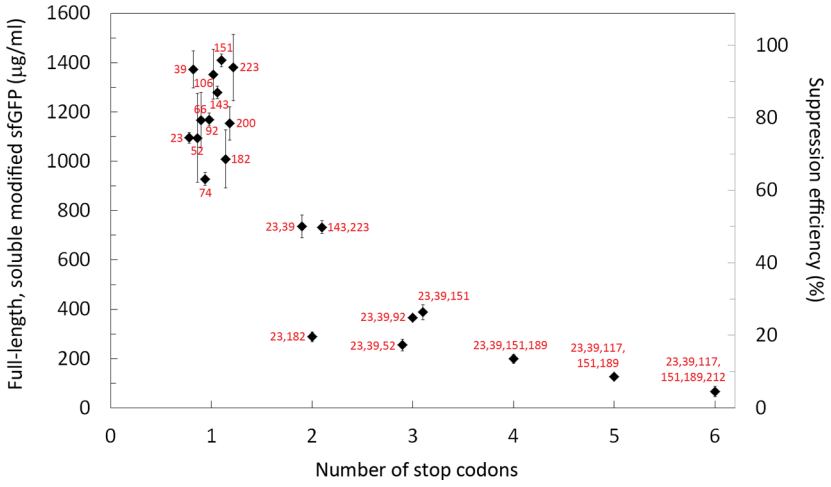


Figure 2. Full-length and soluble modified sfGFP yields containing one to six copies of the tyrosine analogue pPaF. The numbers indicate the site of nnAA substitution along the primary sequence. Suppression efficiency is defined as the ratio of modified protein yields to the natural sfGFP yield. Error bars indicate ±1 standard deviation for 3 reactions. The singly substituted modified sfGFP yields were reported previously[37] and are included here for comparison. Reprinted with permission from reference 37. Copyright 2013 Oxford University Press.

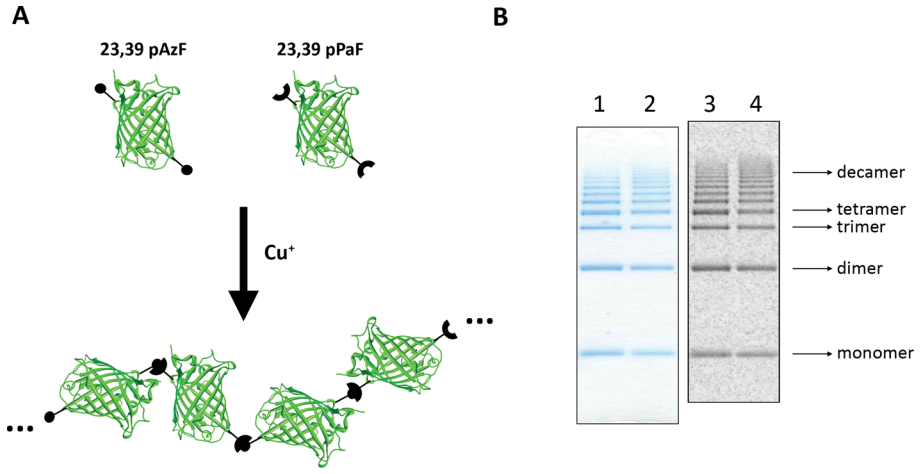


Figure 3. Synthesis of linear sfGFP polymers. (A) Cartoon showing the generation of a linear sfGFP polymer from directly coupling doubly substituted sfGFPs by CuAAC. (B) Generation of a sfGFP ladder by direct polymerization of doubly substituted sfGFP. Equimolar amounts of sfGFP containing two copies of pAzF or two copies of pPaF were reacted in an anaerobic glovebox for 9 h at room temperature. The numbers indicate the nnAA incorporation sites. The first reaction contained 6.5 μM sfGFP23,39pAzF, 6.5 μM sfGFP23,39pPaF, 0.25 mM Cu(I) catalyst, and 0.25 mM TTMA (Lanes 1 and 3). The second reaction contained 6.0 μM sfGFP23,39pAzF, 6.0 μM sfGFP23,39pPaF, 0.5 mM Cu(I) catalyst, and 0.5 mM TTMA (Lanes 2 and 4). The left panel (Lanes 1 and 2) shows the Coomassie-stained protein bands, and the right panel (Lanes 3 and 4) shows the autoradiogram of the SDS-PAGE gel.

presumably because dimers with different connectivities migrated differently across the polyacrylamide gel, a feature not possible with dimers synthesized from doubly substituted sfGFP (Figures 3B and 4B). More of the larger polymers were obtained when the Cu(I) catalyst and TTMA enhancer concentrations were increased from 0.25 to 0.5 mM (left vs right lanes in Figure 4B). We were thus able to produce linear and branched sfGFP polymers in one step by directly coupling proteins containing nnAAs at multiple sites using Click chemistry.

We next assessed activity retention after polymerization. The modified sfGFP proteins were concentrated using centrifugal filter devices prior to CuAAC. Research from our laboratory showed that sfGFP fluorescence can be recovered after CuAAC by removing the small molecules (i.e., Cu(I) catalyst and TTMA) from the reaction mixture.[35] Therefore, the reaction products from separate conjugations of doubly and triply substituted sfGFP proteins were fractionated by size exclusion chromatography (SEC) to separate the newly created polymers from the unreacted monomer as well as the catalysts. Finally, the fluorescence and protein concentration of each SEC fraction was measured to determine the specific activity of sfGFP polymers.

In the first CuAAC reaction with concentrated reactants, 22.6 μM each of sfGFP23,39pAzF and sfGFP23,39pPaF were coupled in an anerobic glovebox (to retain the copper catalyst in its active, +1 oxidation state[40,41]) in a 110 μL solution containing 1 mM Cu(I) catalyst and 0.5 mM TTMA to form linear polymers (Figure 5). The proteins had been labeled with ^{14}C-leucine to facilitate quantification. After desalting the reaction solution in the anaerobic glovebox (to remove the TTMA and the Cu(I) catalyst in its reduced state), the products were separated by analytical SEC (Figure 5A). The SEC fractions were first analyzed by SDS-PAGE (Figure 5B)

dx.doi.org/10.1021/sb400116x | ACS Synth. Biol. 2014, 3, 353−362

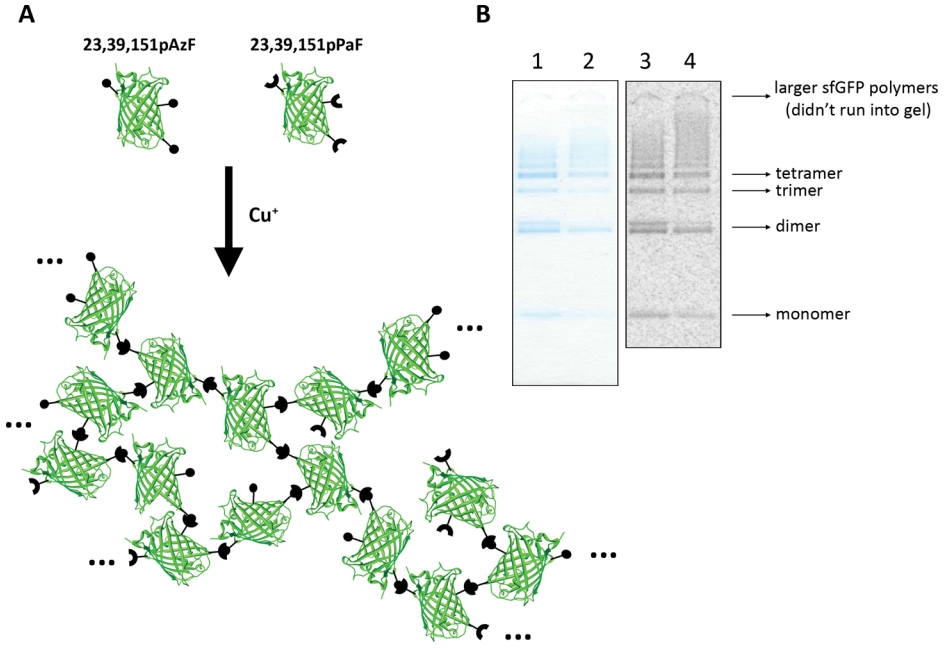


Figure 4. Synthesis of branched sfGFP polymers. (A) Cartoon showing the generation of a branched sfGFP polymer from directly coupling triply substituted sfGFPs using Click chemistry. (B) Larger branched sfGFP polymers were obtained when triply substituted sfGFP proteins were covalently coupled. As seen more clearly on the right panel, some of the branched polymers were so large that they did not enter the gel. Equimolar amounts of sfGFP containing three copies of pAzF or three copies of pPaF were reacted in an anaerobic glovebox for 9 h at room temperature. The numbers indicate the nnAA incorporation sites. The first reaction contained 2.9 μM sfGFP23,39,151pAzF, 2.9 μM sfGFP23,39,151pPaF, 0.25 mM Cu(I), and 0.25 mM TTMA (Lanes 1 and 3). The second reaction contained 2.7 μM sfGFP23,39,151pAzF, 2.7 μM sfGFP23,39,151pPaF, 0.5 mM Cu(I), and 0.5 mM TTMA (Lanes 2 and 4). The left panel (Lanes 1 and 2) shows the Coomassie-stained protein bands, and the right panel (Lanes 3 and 4) shows the autoradiogram of the SDS-PAGE gel.

and autoradiography (Figure 5C). The concentration and the fluorescence of the resulting fractions were then measured by LSC and fluorimetry, respectively (Figure 5D). The doubly substituted sfGFP (i.e., sfGFP23,39pPaF) has 83% of the specific activity of natural sfGFP. The polymerized sfGFP proteins in fractions 1−13 retained most (≥78%) of their specific activity after the Click reaction; this corresponds to ≥65% of the specific activity of natural sfGFP (Figure 5D). A single SEC step was not sufficient to isolate individual sfGFP polymers, but did separate the polymers from unreacted sfGFP monomer. Fractions 1−9 contained the different sfGFP polymers, fractions 10−13 contained a mixture of polymer and unreacted monomer, while fractions 14−17 contained only the unreacted monomer (Figure 5A−C). The amount of each polymer in each fraction was estimated by SDS polyacrylamide gel densitometry, and these estimates were used to calculate the distribution of polymerization products: ∼40% of the monomeric sfGFP was polymerized ($N \geq 3$), 40% assembled into dimers, 15% did not react, and the remaining 5% was degraded. 9% of the monomeric sfGFP assembled into homopolymers with 5 or more monomers (SI Table 1).

The same protocol was repeated for a second CuAAC reaction with concentrated triply substituted sfGFP proteins. In this reaction, 105 μM each of sfGFP23,39,151pAzF and sfGFP23,39,151pPaF were coupled in a 150 μL solution containing 1 mM Cu(I) catalyst and 0.5 mM TTMA (Figure 6). (An additional CuAAC reaction was incubated with all components except the Cu(I) catalyst to serve as negative control.) Fluorescent macromolecular particles were formed by this Click reaction (Figure 6A). After centrifugation at 14000 g for 10 min, the supernatant was recovered, desalted, and fractionated by SEC, as explained previously (Figure 6B). The fluorescent particles were washed five times with deoxygenated buffer, and the supernatants from each wash were also analyzed by SDS-PAGE (Figure 6C) and autoradiography (Figure 6D). The specific activity of the sfGFP polymers in each wash fraction was then determined using LSC and fluorescence measurements (Figure 6E). Most of the sfGFP polymers (0.59 mg, ∼70% of total protein) precipitated to form the fluorescent pellet at the bottom of the reaction tube, thereby demonstrating that a biomaterial composed entirely of covalently coupled active proteins can be formed in one step using nnAAs and Click chemistry (Figure 6A). The largest of the pelleted particles was approximately 2 mm in diameter; the weight of the pelleted particles was 12 mg, hence the particles consist of ∼5% w/v coupled sfGFP. For a significant fraction (14%, 4% of total protein) of the supernatant sfGFP, the polymers were so large that they did not run into the SDS-PAGE gel (Figures 6C and 6D). SEC fractions 1−28 contained the different soluble sfGFP polymers in the supernatant, while fractions 29−33 contained the unreacted sfGFP monomer, which comprised only a tiny fraction (0.7%) of the proteins after the CuAAC reaction. The triply substituted sfGFP proteins had higher specific activities (22% and 32% higher for the azide- and alkyne-containing sfGFPs, respectively) than the natural sfGFP. As for the doubly substituted sfGFP, the polymerized triply substituted sfGFP proteins retained most (≥63%) of the specific activity, which corresponds to ≥83% of the specific activity of natural sfGFP (Figure 6E).

To the best of our knowledge, this is the first report both of a biomaterial composed entirely of folded and active proteins and of the synthesis of active protein polymers in one step using Click chemistry. The nnAA incorporation (i.e., the conjugation) sites can be precisely determined using standard genetic

dx.doi.org/10.1021/sb400116x | ACS Synth. Biol. 2014, 3, 353−362

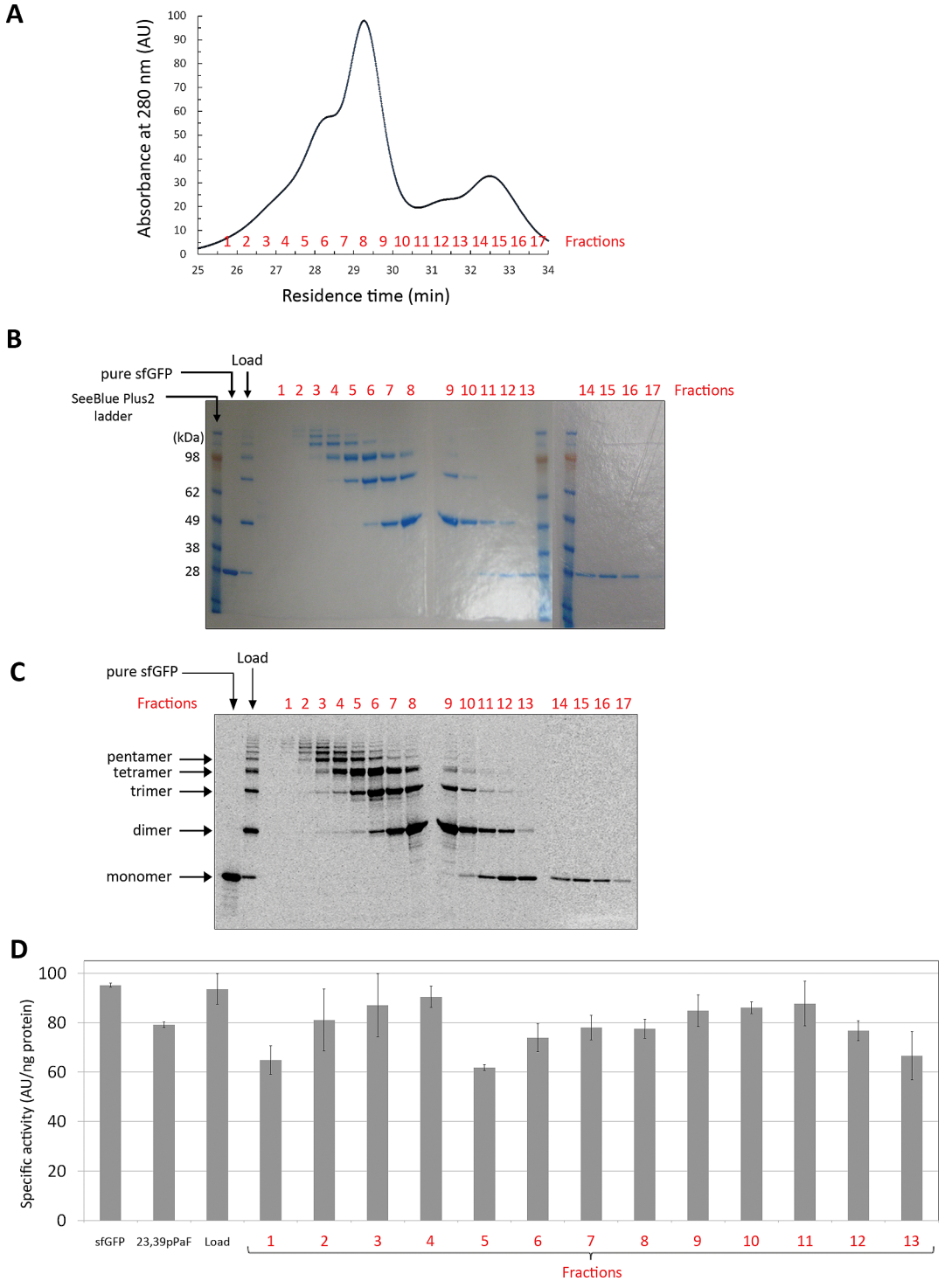


Figure 5. Separation of linear sfGFP polymers by size exclusion chromatography (SEC). (A) Absorbance at 280 nm shows the presence of different polymers. After 7.5 mL, each fraction was manually collected for 0.5 min and contained 150 μL of protein solution. (B) Coomassie-stained SDS-PAGE gel showing the contents of the collected fractions. Purified natural (monomeric) sfGFP ("pure sfGFP") and the protein solution that was loaded onto the SEC column ("Load") were included for comparison. The protein ladder was also included to determine the polymerization state of each protein observed on the gel. (C) Autoradiogram of the SDS-PAGE gels containing the different SEC fractions. The polymerization state of each band is indicated on the left. (D) Specific activity of fractions 1−14 that contained either sfGFP polymers or a mixture of polymers and unreacted monomer. The first three bars indicate the specific activities of the purified natural sfGFP, the doubly substituted sfGFP that contains pPaF at positions 23 and 39, and the protein mixture that was loaded onto the SEC column. Error bars indicate ±1 standard deviation for 6 specific activity values that were calculated from 2 independent LSC measurements and 3 independent fluorescence measurements.

manipulation, and our methods can be applied to virtually any protein. Proteins that are sensitive to the Cu(I) catalyst[35] can be coupled using nnAAs with strained alkyne moieties,[25] which enable copper-free Click chemistry. Furthermore, the azido and the alkyne moieties on each of the nnAAs react specifically with each other, thereby minimizing the possibility of unwanted side

dx.doi.org/10.1021/sb400116x | ACS Synth. Biol. 2014, 3, 353−362

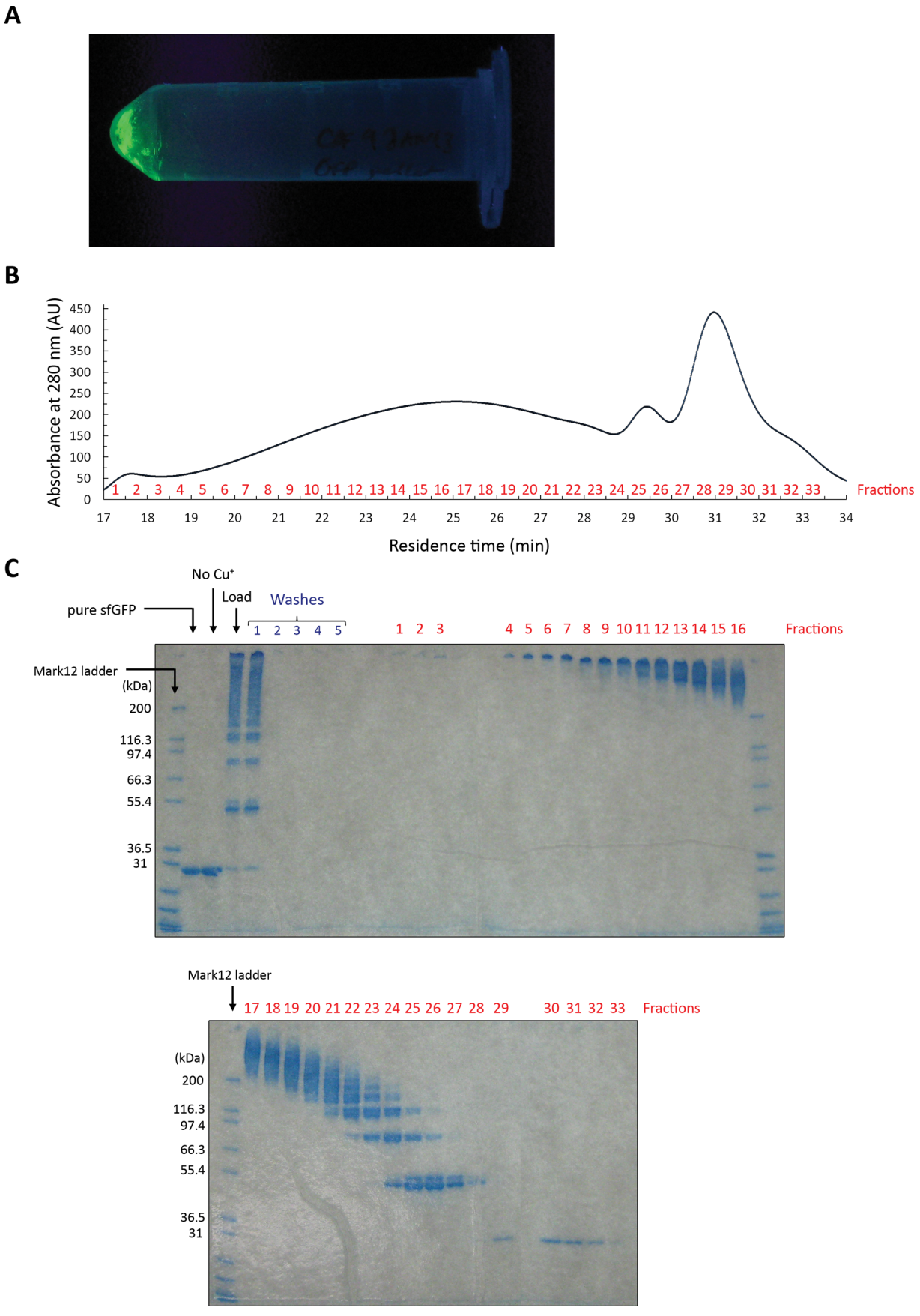


Figure 6. continued

dx.doi.org/10.1021/sb400116x | *ACS Synth. Biol.* 2014, 3, 353−362

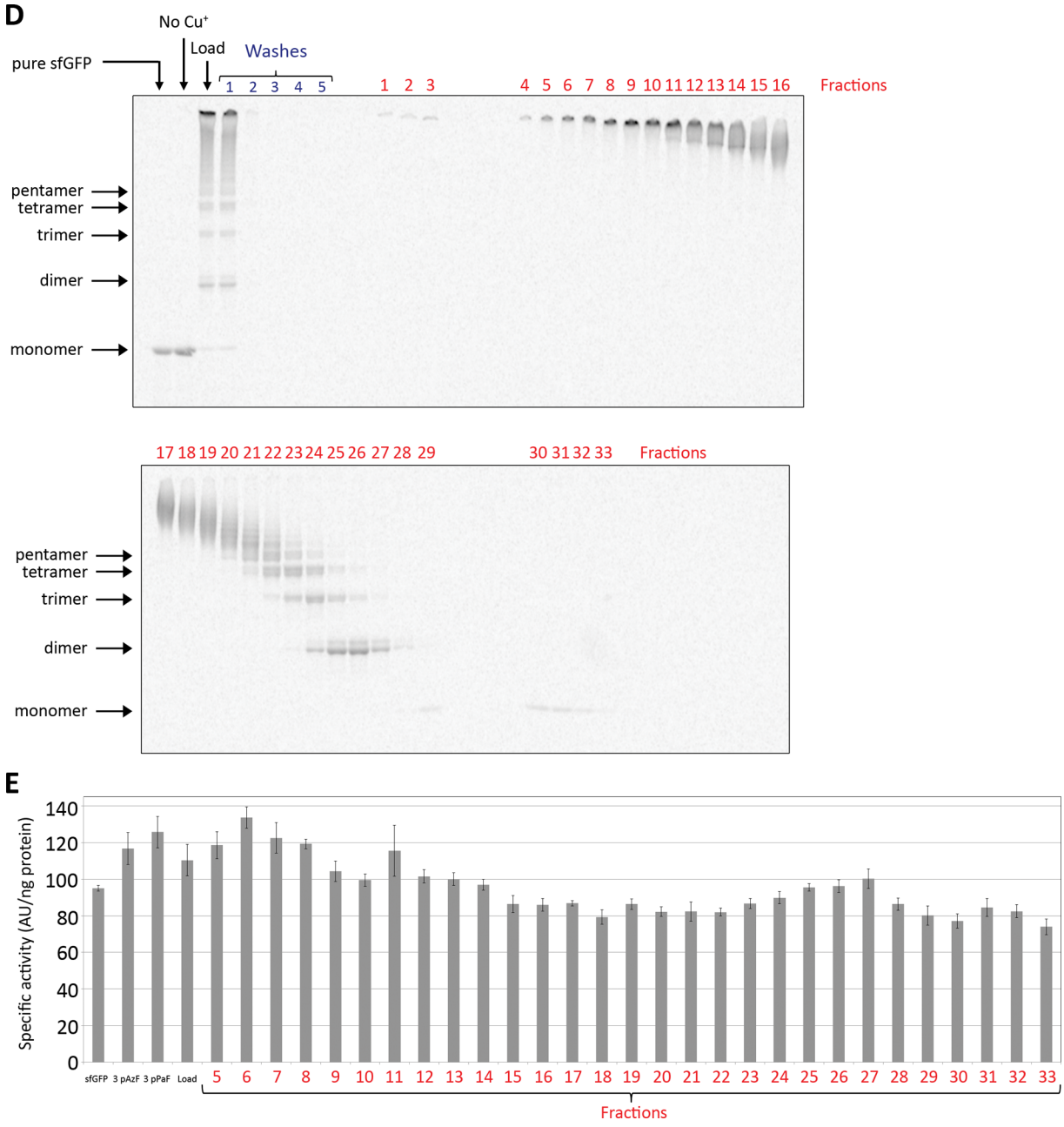


Figure 6. Separation of branched sfGFP polymers by size exclusion chromatography (SEC). (A) Photograph of the fluorescent sfGFP pellet that formed after the Click reaction in the microcentrifuge tube. The pellet was illuminated using a 302 nm UV transilluminator. (B) Absorbance at 280 nm shows the presence of different polymers. After 5.1 mL, each fraction was manually collected for 0.5 min and contained 150 μL of protein solution. (C) SDS-PAGE gels showing the contents of the collected fractions. Purified monomeric sfGFP ("pure sfGFP"), the reaction solution with no Cu(I) catalyst (negative control, "No Cu^{+}") and the protein solution that was loaded onto the SEC column ("Load") were included for comparison. The solutions obtained by washing the sfGFP pellet ("Washes 1−5") were also included. The Mark12 MW ladder was included to indicate the polymerization state of the observed bands. (D) Autoradiogram of the SDS-PAGE gels. The polymerization state of each band is indicated on the left. (E) Specific activity of fractions 5−33 that contained either sfGFP polymers or the unreacted sfGFP monomer. ≥63% of the specific sfGFP activity (i.e., fluorescence) was retained after polymerization. The first four bars indicate the specific activities of the purified natural sfGFP ("sfGFP"), the triply substituted sfGFP that contains pAzF at positions 23,39 and 151 ("3 pAzF"), the triply substituted sfGFP that contains pPaF at the same positions ("3 pPaF"), and the protein mixture that was loaded onto the SEC column ("Load"). Error bars indicate ±1 standard deviation for 9 specific activity values that were calculated from 3 independent LSC measurements and three independent fluorescence measurements.

reactions and the formation of polymers with undesired configurations. Despite its demonstration with only one protein, this method can also be generalized to form copolymers of two or more proteins or hybrid polymers composed of active proteins and small molecules.

Our protein cross-linking experiments included modified proteins in equimolar amounts and resulted in heterogeneous protein mixtures containing unreacted monomers and different multimers. However, multistep polymerization of proteins would be desirable so that specific multimeric species can be

dx.doi.org/10.1021/sb400116x | ACS Synth. Biol. 2014, 3, 353−362

enriched in the assembly reaction. Monomer concentration, distribution (stoichiometric vs nonstoichiometric), reaction time, time and order of additions, and addition of quenchers (e.g., singly substituted modified protein) can be adjusted to alter the extent of reaction and cross-linked protein distribution. Our data indicate that larger polymers are obtained and the polymer distribution is shifted toward higher molecular weight species at higher reactant concentrations (Figures 4 and 6). When synthesizing copolymers, one of the monomers can be added at limiting amounts so that polymerization would be terminated when the limiting monomer is depleted in the reaction mixture. Finally, a quencher can be added to terminate polymerization after a certain amount of time, and the quencher concentration and timing of quencher addition can be adjusted to achieve the desired polymer distribution. Our results demonstrate the feasibility of this polymerization strategy; additional process development research will be needed to develop methods for producing polymerized protein biomaterials with uniform and consistent functionality. Improved conjugation reagents and methods will also help to reduce production costs in the future.

Specific activity data show no correlation with polymer size or the nature of nnAA substitutions, but a fraction (up to 37% for certain oligomers) of the specific activity is lost after the CuAAC reaction. The nnAA incorporation sites are located on the surface of the protein, while the chromophore is at the center of the sfGFP barrel and hence relatively far from the conjugation sites. As a result, we did not expect nnAA substitution or conjugation to significantly alter specific fluorescence. Some of the protein may be degrading during CuAAC incubation, thereby reducing the specific activity of the protein polymers.

The versatility and the precision of this coupling strategy can complement existing technologies in biomaterials synthesis or metabolic engineering. For example, a recent report described the manufacture of a patterned hydrogel using photoinduced Click chemistry.[42] Proteins containing the azide or alkyne moieties can be integrated into these biomaterials, and such biomaterials containing covalently coupled active proteins (e.g., growth factors) can be used to engineer tissues.[43] Alternatively, enzymes comprising a pathway can be linked using this method to increase the substrate flux by increasing local enzyme concentrations, thereby reducing diffusion distances between enzymes. If the enzymes catalyze the maturation or regeneration of a cofactor, for example, then they may be coupled to the enzyme requiring the cofactor for maximal holoenzyme activity. In this manner, "clicking" different enzymes of a pathway post-translationally would complement the protein scaffolding approach, which was shown to increase metabolic flux and product titers.[44]

Finally, further improvements in modified protein yields (especially of proteins with multiple nnAAs) and conjugation efficiency would enhance the usefulness of our methods; our protein yields dropped from ~1000 to ~70 μg/mL when the number of nnAAs in sfGFP was increased from one to six (Figure 2). Our methods currently enable the synthesis of biopolymers and biomaterials composed of active proteins using a single conjugation reaction, but we are limited to single conjugation chemistry. With the advent of more bio-orthogonal moieties (in addition to the azido−alkyne pair) that can be directly incorporated into proteins, better methods to site-specifically incorporate many different nnAAs into the same protein, and improved bioconjugation methods, we anticipate that scientists and engineers will be able to synthesize complex biomaterials and other macromolecular structures with precise configurations for a broad variety of applications.

METHODS

Detailed protocols on plasmid construction, CFPS, sfGFP purification and quantification, CuAAC reactions, size exclusion chromatography, and specific activity measurements are available in the Supporting Information.

ASSOCIATED CONTENT

Supporting Information

Experimental methods, Figures S1 and S2, supplementary Table 1, and supplementary references. This material is available free of charge *via* the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*Tel.: +41 61 387 3387. Email: cemalb@gmail.com.
*Tel.: 650-723-5398. Fax: 650-725-0555. E-mail: jswartz@stanford.edu.

Present Address

[§]Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26, 4058 Basel, Switzerland

Notes

The authors declare the following competing financial interest(s): The authors declare competing financial interests as could potentially share in royalty payments received by Stanford University resulting from issuance and licensing of a filed provisional patent.

ACKNOWLEDGMENTS

We thank Dr. Randall Lowe for providing us with the TTMA ligand, Drs. Kedar Patel and William Yang for T7 RNA polymerase and the orthogonal synthetase specific to pPaF, Prae Numswangneth for her assistance with the production and purification of the modified proteins, and Dr. Yuan Lu and Wei Chan for their advice on size exclusion chromatography. We are also grateful to the Global Climate and Energy Project (GCEP) at Stanford University for providing partial funding for this work (http://gcep.stanford.edu).

ABBREVIATIONS

CFME: cell-free metabolic engineering
CFPS: cell-free protein synthesis
CuAAC: copper-catalyzed azide−alkyne cycloaddition
LSC: liquid scintillation counting
nnAA: non-natural amino acid
PAGE: polyacrylamide gel electrophoresis
pAzF: *p*-azido-L-phenylalanine
pPaF: *p*-propargyloxy-L-phenylalanine
SDS: sodium dodecyl sulfate
SEC: size exclusion chromatography
sfGFP: superfolder green fluorescent protein
TTMA: tris(triazolylmethyl)amine

REFERENCES

(1) Stephanopoulos, N., and Francis, M. B. (2011) Choosing an effective protein bioconjugation strategy. *Nat. Chem. Biol.* 7, 876−884.
(2) Cho, H., Daniel, T., Buechler, Y. J., Litzinger, D. C., Maio, Z., Putnam, A.-M. H., Kraynov, V. S., Sim, B.-C., Bussell, S., Javahishvili, T., Kaphle, S., Viramontes, G., Ong, M., Chu, S., Becky, G. C., Lieu, R.,

Knudsen, N., Castiglioni, P., Norman, T. C., Axelrod, D. W., Hoffman, A. R., Schultz, P. G., DiMarchi, R. D., and Kimmel, B. E. (2011) Optimized clinical performance of growth hormone with an expanded genetic code. *Proc. Natl. Acad. Sci. U.S.A. 108*, 9060−9065.
(3) Axup, J. Y., Bajjuri, K. M., Ritland, M., Hutchins, B. M., Kim, C. H., Kazane, S. A., Halder, R., Forsyth, J. S., Santidrian, A. F., Stafin, K., Lu, Y., Tran, H., Seller, A. J., Biroc, S. L., Szydlik, A., Pinkstaff, J. K., Tian, F., Sinha, S. C., Felding-Habermann, B., Smider, V. V, and Schultz, P. G. (2012) Synthesis of site-specific antibody-drug conjugates using unnatural amino acids. *Proc. Natl. Acad. Sci. U.S.A. 109*, 16101−16106.
(4) Davis, L., and Chin, J. W. (2012) Designer proteins: applications of genetic code expansion in cell biology. *Nat. Rev. Mol. Cell Biol. 13*, 168−182.
(5) Neumann, H. (2012) Rewiring translation—Genetic code expansion and its applications. *FEBS Lett. 586*, 2057−2064.
(6) Mursinna, R. S., Lee, K. W., Briggs, J. M., and Martinis, S. A. (2004) Molecular dissection of a critical specificity determinant within the amino acid editing domain of leucyl-tRNA synthetase. *Biochemistry 43*, 155−165.
(7) Van Hest, J. C. M., Kiick, K. L., and Tirrell, D. A. (2000) Efficient incorporation of unsaturated methionine analogues into proteins *in vivo*. *J. Am. Chem. Soc. 122*, 1282−1288.
(8) Tang, Y., and Tirrell, D. A. (2002) Attenuation of the editing activity of the *Escherichia coli* leucyl-tRNA synthetase allows incorporation of novel amino acids into proteins *in vivo*. *Biochemistry 41*, 10635−10645.
(9) Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U.S.A. 99*, 19−24.
(10) Link, A. J., Vink, M. K. S., Agard, N. J., Prescher, J. A., Bertozzi, C. R., and Tirrell, D. A. (2006) Discovery of aminoacyl-tRNA synthetase activity through cell-surface display of noncanonical amino acids. *Proc. Natl. Acad. Sci. U.S.A. 103*, 10180−10185.
(11) Tang, Y., Wang, P., Van Deventer, J. A., Link, A. J., and Tirrell, D. A. (2009) Introduction of an aliphatic ketone into recombinant proteins in a bacterial strain that overexpresses an editing-impaired leucyl-tRNA synthetase. *ChemBioChem 10*, 2188−2190.
(12) Wang, P., Tang, Y., and Tirrell, D. A. (2003) Incorporation of trifluoroisoleucine into proteins *in vivo*. *J. Am. Chem. Soc. 125*, 6900−6906.
(13) Strable, E., Prasuhn, D. E., Udit, A. K., Brown, S., Link, A. J., Ngo, J. T., Lander, G., Quispe, J., Potter, C. S., Carragher, B., Tirrell, D. A., and Finn, M. G. (2008) Unnatural amino acid incorporation into virus-like particles. *Bioconjug. Chem. 19*, 866−875.
(14) Welsh, J. P., Patel, K. G., Manthiram, K., and Swartz, J. R. (2009) Multiply mutated *Gaussia luciferases* provide prolonged and intense bioluminescence. *Biochem. Biophys. Res. Commun. 389*, 563−568.
(15) Lu, Y., Welsh, J. P., Chan, W., and Swartz, J. R. (2013) *Escherichia coli*-based cell free production of flagellin and ordered flagellin display on virus-like particles. *Biotechnol. Bioeng. 110*, 2073−2085.
(16) Patel, K. G., Ng, P. P., Kuo, C.-C., Levy, S., Levy, R., and Swartz, J. R. (2009) Cell-free production of *Gaussia princeps* luciferase−antibody fragment bioconjugates for *ex vivo* detection of tumor cells. *Biochem. Biophys. Res. Commun. 390*, 971−976.
(17) Patel, K. G., and Swartz, J. R. (2011) Surface functionalization of virus-like particles by direct conjugation using azide−alkyne click chemistry. *Bioconjug. Chem. 22*, 376−387.
(18) Wang, L., and Schultz, P. G. (2001) A general approach for the generation of orthogonal tRNAs. *Chem. Biol. 8*, 883−890.
(19) Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of *Escherichia coli*. *Science 292*, 498−500.
(20) Polycarpo, C. R., Herring, S., Bérubé, A., Wood, J. L., Söll, D., and Ambrogelly, A. (2006) Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase. *FEBS Lett. 580*, 6695−6700.
(21) Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem. 79*, 413−444.
(22) Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., and Yokoyama, S. (2008) Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode N(epsilon)-(o-azidobenzyloxycarbonyl) lysine for site-specific protein modification. *Chem. Biol. 15*, 1187−1197.
(23) Nguyen, D. P., Lusic, H., Neumann, H., Kapadnis, P. B., Deiters, A., and Chin, J. W. (2009) Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA Synthetase/tRNA(CUA) pair and click chemistry. *J. Am. Chem. Soc. 131*, 8720−8721.
(24) Wan, W., Huang, Y., Wang, Z., Russell, W. K., Pai, P.-J., Russell, D. H., and Liu, W. R. (2010) A facile system for genetic incorporation of two different noncanonical amino acids into one protein in *Escherichia coli*. *Angew. Chem., Int. Ed. Engl. 49*, 3211−3214.
(25) Plass, T., Milles, S., Koehler, C., Schultz, C., and Lemke, E. A. (2011) Genetically encoded copper-free click chemistry. *Angew. Chem., Int. Ed. Engl. 50*, 3878−3881.
(26) Neumann, H., Slusarczyk, A. L., and Chin, J. W. (2010) De novo generation of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs. *J. Am. Chem. Soc. 132*, 2142−2144.
(27) Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., and Chin, J. W. (2010) Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature 464*, 441−444.
(28) Wang, K., Schmied, W. H., and Chin, J. W. (2012) Reprogramming the genetic code: From triplet to quadruplet codes. *Angew. Chem., Int. Ed. Engl. 51*, 2288−2297.
(29) Guo, J., Melançon, C. E., Lee, H. S., Groff, D., and Schultz, P. G. (2009) Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing nonnatural amino acids. *Angew. Chem., Int. Ed. Engl. 48*, 9148−9151.
(30) Ryu, Y., and Schultz, P. G. (2006) Efficient incorporation of unnatural amino acids into proteins in *Escherichia coli*. *Nat. Methods 3*, 263−265.
(31) Young, T. S., Ahmad, I., Yin, J. A., and Schultz, P. G. (2010) An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J. Mol. Biol. 395*, 361−374.
(32) Kaya, E., Gutsmiedl, K., Vrabel, M., Müller, M., Thumbs, P., and Carell, T. (2009) Synthesis of threefold glycosylated proteins using click chemistry and genetically encoded unnatural amino acids. *ChemBioChem 10*, 2858−2861.
(33) Johnson, D. B. F., Xu, J., Shen, Z., Takimoto, J. K., Schultz, M. D., Schmitz, R. J., Xiang, Z., Ecker, J. R., Briggs, S. P., and Wang, L. (2011) RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat. Chem. Biol. 7*, 779−786.
(34) Albayrak, C., and Swartz, J. R. (2013) Using *E. coli*-based cell-free protein synthesis to evaluate the kinetic performance of an orthogonal tRNA and aminoacyl-tRNA synthetase pair. *Biochem. Biophys. Res. Commun. 431*, 291−295.
(35) Bundy, B. C., and Swartz, J. R. (2010) Site-specific incorporation of *p*-propargyloxyphenylalanine in a cell-free environment for direct protein-protein click conjugation. *Bioconjug. Chem. 21*, 255−263.
(36) Goerke, A. R., and Swartz, J. R. (2009) High-level cell-free synthesis yields of proteins containing site-specific non-natural amino acids. *Biotechnol. Bioeng. 102*, 400−416.
(37) Albayrak, C., and Swartz, J. R. (2013) Cell-free co-production of an orthogonal transfer RNA activates efficient site-specific non-natural amino acid incorporation. *Nucleic Acids Res. 41*, 5949−5963.
(38) Hodgman, C. E., and Jewett, M. C. (2012) Cell-free synthetic biology: Thinking outside the cell. *Metab. Eng. 14*, 261−269.
(39) Swartz, J. R. (2012) Transforming biochemical engineering with cell-free biology. *AIChE J. 58*, 5−13.
(40) Rostovtsev, V. V, Green, L. G., Fokin, V. V, and Sharpless, K. B. (2002) A stepwise huisgen cycloaddition process: Copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem., Int. Ed. Engl. 41*, 2596−2599.

(41) Tornøe, C. W., Christensen, C., and Meldal, M. (2002) Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(i)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem. 67*, 3057−3064.
(42) Adzima, B. J., Tao, Y., Kloxin, C. J., DeForest, C. A., Anseth, K. S., and Bowman, C. N. (2011) Spatial and temporal control of the alkyne−azide cycloaddition by photoinitiated Cu(II) reduction. *Nat. Chem. 3*, 258−261.
(43) Hudalla, G. A., and Murphy, W. L. (2011) Biomaterials that regulate growth factor activity via bioinspired interactions. *Adv. Funct. Mater. 21*, 1754−1768.
(44) Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V, Prather, K. L. J., and Keasling, J. D. (2009) Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol. 27*, 753−759.