JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Site-Specific Fluorescent Labeling of Nascent Proteins on the Translating Ribosome

Ishu Saraogi, Dawei Zhang, Sandhya Chandrasekaran, and Shu-ou Shan*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States

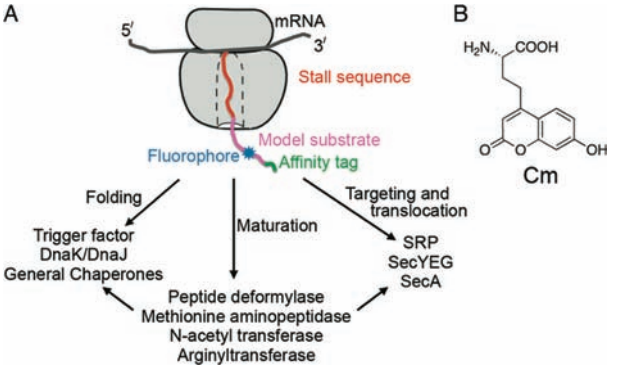
Supporting Information

ABSTRACT: As newly synthesized proteins emerge from the ribosome, they interact with a variety of cotranslational cellular machineries that facilitate their proper folding, maturation, and localization. These interactions are essential for proper function of the cell, and the ability to study these events is crucial to understanding cellular protein biogenesis. To this end, we have developed a highly efficient method to generate ribosome−nascent chain complexes (RNCs) site-specifically labeled with a fluorescent dye on the nascent polypeptide. The fluorescent RNC provides real-time, quantitative information on its cotranslational interaction with the signal recognition particle and will be a valuable tool in elucidating the role of the translating ribosome in numerous biochemical pathways.

The ribosome is the central molecular machine that translates genetic information into proteins. As newly synthesized proteins emerge from the ribosomal exit tunnel, they interact with a host of cellular factors that facilitate the folding, localization, maturation, and quality control of the nascent proteins. These include chaperones[1] such as trigger factor, Hsp70 (DnaK/J in bacteria), and the nascent chain associated complex; modification and processing enzymes[2] such as methionine aminopeptidase (or peptide deformylase in bacteria), *N*-acetyl transferase, and arginyl transferase; protein transport machineries[3] such as the signal recognition particle (SRP), the Sec61p (or secYEG) protein-conducting channel, and even possibly SecA[4] (Scheme 1A). These interactions, which often take place at the ribosome exit site, are essential for the proper function of the cell. A molecular understanding of the dynamic interactions of these machineries with the ribosome−nascent chain complex (RNC) is crucial to understanding cellular protein biogenesis, yet such information is challenging to obtain because of the lack of suitable tools. To monitor these interactions directly and quantitatively, we have developed an efficient method to generate synchronized RNCs site-specifically labeled with a fluorophore on the nascent protein. Large quantities of fluorescent RNC can be generated using this method, enabling kinetic and thermodynamic studies of early events during the biogenesis of the nascent chain as it emerges from the ribosome.

Several factors must be considered for efficient labeling and purification of RNCs. The ribosome comprises numerous proteins, which precludes labeling strategies based on protein chemistry; hence, selective labeling of the nascent chain necessitates the use of a bioorthogonal reactivity handle. Additionally, the fluorophore should be small to minimize perturbations due to probe

Scheme 1. (A) Schematic Depiction of the Fluorescent RNC and Some Cellular Factors That Interact with the Ribosome-Associated Nascent Protein; (B) Structure of Cm, the Fluorescent Amino Acid Used in This Study



incorporation. Finally, to ensure that an intact RNC is obtained, we need to minimize harsh and lengthy procedures during labeling and purification. An approach that satisfies these requirements is amber suppression, which incorporates a fluorophore into the nascent chain cotranslationally in response to the UAG (amber) stop codon.[5,6] Current methods for generating fluorescent RNCs[7,8] often use purified aminoacylated transfer RNA (tRNA); however, the yield is low, as the aminoacylated tRNA is chemically labile. Instead, we chose the *Methanococcus jannaschii* tyrosyl-tRNA synthetase (aaRS) developed by Schultz, which aminoacylates a cognate $tRNA_{CUA}$ with the fluorescent amino acid L-(7-hydroxycoumarin-4-yl)ethylglycine (Cm) (Scheme 1B).[9,10] The use of a cognate aaRS allows continuous regeneration of the aminoacylated $tRNA_{CUA}$, resulting in higher yield.

This method, originally developed for in vivo production of proteins, is not optimal for the generation of fluorescent RNC for several reasons. First, during the generation of a stalled RNC, the ribosome remains stably bound to the mRNA and the nascent chain and hence cannot be reused for multiple rounds of protein synthesis, resulting in low yields. Second, during amber suppression, release factor-1 (RF-1) competes with the suppressor $tRNA_{CUA}$ for decoding the amber codon, which terminates protein synthesis and further reduces the yield. Finally, a high concentration of Cm (typically 1 mM) is required in in vivo experiments. This not only limits the scale of RNC production but also complicates subsequent purification of the RNC, as

Received: July 15, 2011
Published: August 26, 2011

hydrophobic fluorophores can bind to the ribosome nonspecifically at high concentrations. To address these problems, we generated labeled RNC in a highly efficient in vitro transcription−translation system.[11−13] Cell-free translation permits the use of an RF-1 inhibitor to improve the suppression yield, a technique that might be toxic in vivo. Furthermore, the concentration of Cm required for efficient amber suppression is much lower (see below), which facilitates both labeling and purification. Finally, in vitro protein synthesis offers the flexibility to generate multiple RNCs with different nascent chain compositions, lengths, and locations of the fluorescent dye using the same cell lysate. This greatly reduces the time and effort required to explore different nascent chain constructs relative to those in vivo.

To test and optimize the efficiency of amber suppression in vitro, we started with a model protein chloramphenicol acetyltransferase (CAT). We generated S30 cell lysates from an *Escherichia coli* KC6 strain[14] that constitutively expressed $tRNA_{CUA}$.[15,16] An amber suppression reaction [see the Supporting Information for experimental details] using an amber codon engineered within the coding sequence of CAT produced the full-length protein in 45−80% yield relative to wild-type CAT (SI Figure 1). In the absence of aaRS or Cm, <2% of the full-length product was obtained, indicating highly specific incorporation of Cm into the protein. Addition of an RNA aptamer[17] that inhibits RF-1 (RF-apt) increased the yield of the full-length product 2-fold (SI Figure 2). Interestingly, in a less efficient suppression system using an earlier-generation $tRNA_{CUA}$,[9] addition of RF-apt provided greater enhancement of the suppression yield (SI Figure 3).

We next used this method to generate a synchronized RNC site-specifically labeled with Cm (Scheme 1A, blue) on the nascent chain. To ensure a homogeneous population of nascent chains bound to the ribosome, a SecM stall sequence that arrests translation elongation at a specific point[19] was appended to the C-terminus of the protein (Scheme 1A, red). The nascent chain also contained a triple strep tag at the N-terminus for affinity purification (Scheme 1A, green).[18] As a model protein substrate (Scheme 1A, magenta), we used the N-terminal region of 1A9L, a variant of alkaline phosphatase (phoA).[20] In vitro transcription−translation of a 1A9L template with an amber codon at position 56 generated a polypeptide that matched the size expected for the full-length translation product. The suppression yield was close to 100% relative to wild-type 1A9L (Figure 1A, cf. lanes 6 and 7 vs 1). About 5−6% background suppression was observed in the absence of aaRS or Cm (Figure 1A, lanes 3 and 4); this background is likely lower in the presence of the orthogonal aminoacylation machinery.

The practical large-scale generation of fluorescent RNC is limited by the high concentration of Cm typically required for in vivo amber suppression. Thus, we examined the dependence of the suppression yield on the concentration of Cm in the in vitro reaction. A suppression yield close to saturation was obtained at ∼50 μM (Figure 1A, lanes 6 and 7, and Figure 1B), ∼20-fold lower than that in vivo. This, combined with the smaller volume of in vitro translation reactions, reduced the consumption of Cm by over 2 orders of magnitude. The use of a lower concentration of Cm also minimized nonspecific binding of the fluorescent dye to the ribosome and simplified subsequent purification of the fluorescently labeled RNC.

To evaluate the generality of the method, we tested amber suppression in RNCs carrying another nascent protein, the

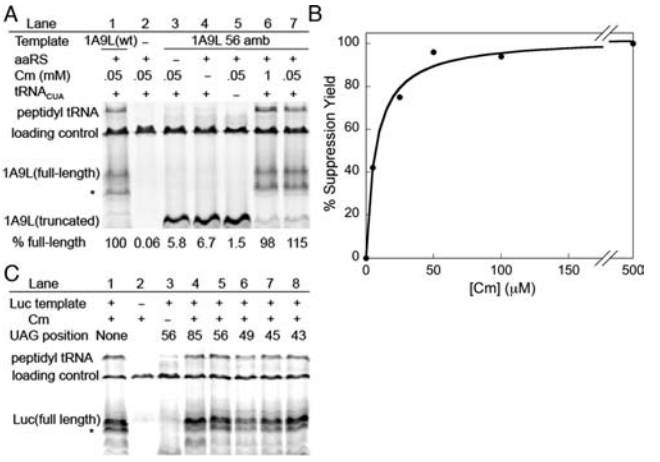


Figure 1. (A) Translation and amber suppression of a model substrate, 1A9L, without (lane 1) or with (lanes 3−7) an amber codon at residue 56 from the N-terminus. Translation products were visualized by labeling with ^{35}S-methionine and SDS-PAGE. The full-length band was quantified relative to that of wild-type 1A9L (lane 1) by autoradiography. The lower-molecular-weight band (marked with an asterisk) was from the polysome fraction.[18] (B) Suppression yield of full-length RNC_{1A9L} at different concentrations of Cm, quantified as in part A. (C) Translation of a 125 amino acid-long nascent chain of firefly luciferase (Luc) without (lane 1) and with (lanes 3−8) the amber codon placed at various locations in the nascent chain. When the amber codon was placed at position 85 (lane 4), the fluorophore was estimated to be near the peptide exit site.

N-terminal region of firefly luciferase (RNC_{Luc}).[20] Furthermore, to test the dependence of the suppression yield on the sequence context, we varied the position of the amber codon in the open reading frame. Highly efficient incorporation of Cm to generate the full-length nascent chain was observed for RNC_{Luc} at virtually all of the positions tested (Figure 1C), even when the amber codon was placed near the C-terminus of the nascent polypeptide (Figure 1C, lane 4). Thus, this method is highly robust and versatile, allowing incorporation of the fluorescent probe at any desired location. In comparison, an established method that uses chemically labeled Met-$tRNA^{fMet}$ for fluorescent labeling of the RNCs is restricted to labeling the N-terminus of proteins.[8]

To obtain large quantities of RNC for biophysical measurements, we scaled up the translation reaction and purified fluorescently labeled RNC in a single step using a Strep-Tactin column.[18,21] The yields (Table 1), which were measured using the optical density at 260 nm, varied from one substrate protein to another and mirrored the translation efficiency of the corresponding wild-type construct. Notably, 200−500 μg of purified, fluorescently labeled RNC was obtained per milliliter of translation (equivalent to 2−4 nmol from 1 L of cell culture), comparable to the yields of unlabeled RNC.[18] The purified RNC generated the fluorescence emission profile expected for 7-hydroxycoumarin (SI Figure 4 and Figure 2B). To test nonspecific binding of Cm to the ribosome, RNC was generated from a wild-type 1A9L template, translated in the presence of Cm, and purified in parallel. This RNC showed ≤15% fluorescence background (SI Figure 4), indicating specific incorporation of Cm into the nascent chain in response to the amber codon.

To test whether the fluorescent RNC was functional, we developed a fluorescence resonance energy transfer (FRET) assay to monitor the interaction of RNC_{1A9L} with SRP. SRP is

Table 1. Yields of Fluorescent RNCs after Affinity Purification[a]

RNC	RNC yield (mg/5 mL of translation)	
	$-tRNA_{CUA}$	$+tRNA_{CUA}$
1A9L (wt)	2	1.4
1A9L 56 amb	NA	1.4
3A7L 56 amb	NA	0.86
Luc 56 amb	NA	2.1
Luc 65 amb	NA	2.6
Luc 75 amb	NA	2.8
Luc 85 amb	NA	2.1

[a] 3A7L is an SRP substrate similar to 1A9L.[20] The number after the protein name indicates the position of the amber codon relative to the N-terminus. Abbreviations: Luc, luciferase; NA, not applicable; wt, wild-type.

a well-characterized, cotranslational protein-targeting machinery that recognizes a hydrophobic signal sequence on its substrate proteins. The protein component of *E. coli* SRP, Ffh, binds the signal sequence via a hydrophobic groove on its methionine-rich M-domain (Figure 2A, blue).[22,23] 1A9L, in which the natural signal sequence of phoA was replaced with a more hydrophobic core, serves as a model substrate for SRP.[20] RNC_{1A9L} with Cm placed two residues downstream of the signal sequence served as the donor in the FRET assay. As the acceptor, we introduced BODIPY-Fl at a nonconserved residue (Figure 2A, gold) on the helix lining the signal-sequence-binding groove of Ffh using cysteine−maleimide chemistry.

Efficient energy transfer was observed when BODIPY-Fl-labeled SRP was added to Cm-labeled RNC_{1A9L} (Figure 2B); this FRET signal could be competed off by excess unlabeled SRP (SI Figure 5). At saturating concentrations of labeled SRP, the FRET efficiency with RNC_{1A9L} was ~0.8 (Figures 2B and 3A). Equilibrium titrations based on the FRET signal gave a K_d value of 0.73 ± 0.43 nM (Figure 3A), in good agreement with reported values.[7,20,24] The FRET assay further allowed us to measure the kinetics of RNC·SRP complex assembly on the millisecond time scale. The assembly of the complex fit well to single-exponential kinetics (SI Figure 6). The observed rate constant increased linearly with increasing SRP concentration, giving a k_{on} value of $2.2 \times 10^6\ M^{-1}\ s^{-1}$ as measured by the slope of the concentration dependence (Figure 3B).

The specificity of the SRP−nascent chain interaction was tested by comparison with RNC_{Luc}, whose nascent chain lacks a hydrophobic signal sequence. Addition of BODIPY-Fl-labeled SRP to RNC_{Luc} with Cm incorporated at several different positions resulted in no significant change in the fluorescence signal (SI Figure 7), indicating the lack of interaction between the SRP M-domain and the nascent chain of an SRP-independent substrate.

The generation of large quantities of fluorescently labeled RNC has allowed for the first time the direct measurement of the dynamic interactions of the nascent polypeptide with the SRP. This method can be applied to study a variety of cotranslational biochemical pathways by engineering the fluorophore at a designated site in the nascent protein of interest. Since the fluorescent dye is on the protein, this method can also be used to study the dynamics of the nascent chain on the ribosome[25,26] or to follow the fate of newly synthesized proteins in post-translational pathways. In comparison with in vivo amber

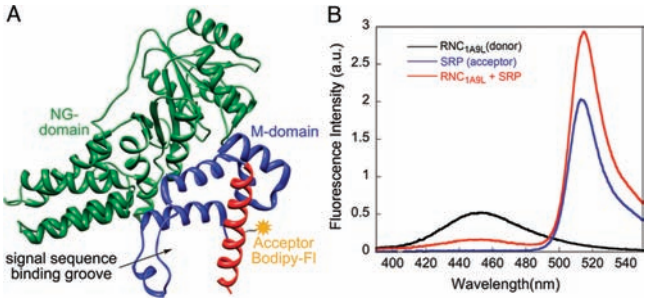


Figure 2. FRET assay for measuring the binding parameters of the RNC−SRP interaction. (A) Crystal structure of SRP (PDB ID 2ffh) with its signal-sequence-binding M-domain in blue and the helix containing the acceptor fluorophore in red. The arrow marks the hydrophobic groove that binds the signal sequence on the nascent chain. (B) Fluorescence emission spectra of 20 nM RNC_{1A9L} labeled with Cm (black), 30 nM SRP labeled with BODIPY-Fl (blue), and their complex (red).

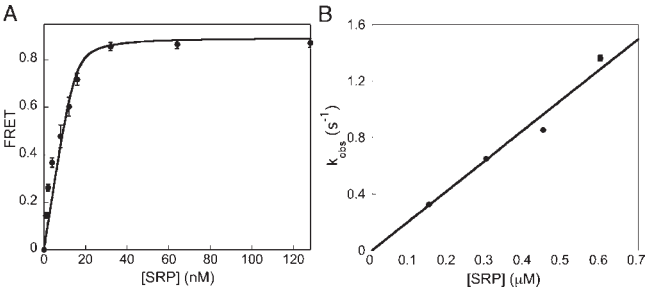


Figure 3. (A) Equilibrium titration for binding of 20 nM fluorescently labeled RNC_{1A9L} to BODIPY-Fl-labeled SRP. The data were fit to eq 2 in the SI and gave a dissociation constant of 0.73 ± 0.43 nM. (B) Observed rate constant of association of RNC_{1A9L} with SRP in the presence of a nonhydrolyzable GTP analogue, GppNHp (SI Figure 6) plotted against the SRP concentration. A linear fit of the data to eq 3 in the SI gave a k_{on} value of $2.2 \times 10^6\ M^{-1}\ s^{-1}$. Measurements were performed in triplicate and are representative of several experiments. Error bars are depicted but are not visible in B.

suppression, the in vitro setting offers higher flexibility in generating RNCs with different nascent chain composition and fluorophore position and eliminates problems with cellular uptake and toxicity of the unnatural amino acid. In principle, other unnatural amino acids could also be inserted into RNCs using this method (e.g., the photo-cross-linker *p*-benzoylphenylalanine[27] for cross-linking experiments). Finally, this method can also be used to generate site-specifically labeled fluorescent proteins that are difficult to produce in vivo because of their toxicity, protease susceptibility, or low solubility.

■ ASSOCIATED CONTENT

Supporting Information. Experimental protocols and supplementary figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
sshan@caltech.edu

■ ACKNOWLEDGMENT

We thank Prof. Swartz (Stanford University) for help with the generation of S30 extracts; Prof. Schultz (TSRI) for providing the pEVOL plasmid containing the $tRNA_{CUA}$/aaRS pair; Prof. Chapman (TSRI) for a sample of Cm; Profs. Tirrell, Barton, and Dougherty (Caltech) for use of their facilities; Prof. Sando (Kyushu University) for the RF-1 inhibitor aptamer sequence; and D. Dougherty and members of the Shan group for helpful comments on the manuscript. This work was supported by NIH Grant GM078024 to S.S. S.S. was supported by the Henry Dreyfus Teacher−Scholar Award, the Beckman Young Investigator Award, and the Packard and Lucile Award in Science and Engineering.

■ REFERENCES

(1) Fedyukina, D. V.; Cavagnero, S. *Annu. Rev. Biophys.* **2011**, *40*, 337–359.
(2) Kramer, G.; Boehringer, D.; Ban, N.; Bukau, B. *Nat. Struct. Mol. Biol.* **2009**, *16*, 589–597.
(3) Cross, B. C. S.; Sinning, I.; Luirink, J.; High, S. *Nat. Rev. Mol. Cell. Biol.* **2009**, *10*, 255–264.
(4) Huber, D.; Rajagopalan, N.; Preissler, S.; Rocco, M. A.; Merz, F.; Kramer, G.; Bukau, B. *Mol. Cell* **2011**, *41*, 343–353.
(5) Wang, L.; Xie, J.; Schultz, P. G. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225–249.
(6) Crowley, K. S.; Reinhart, G. D.; Johnson, A. E. *Cell* **1993**, *73*, 1101–1115.
(7) Flanagan, J. J.; Chen, J. C.; Miao, Y. W.; Shao, Y. L.; Lin, J. L.; Bock, P. E.; Johnson, A. E. *J. Biol. Chem.* **2003**, *278*, 18628–18637.
(8) Ellis, J. P.; Bakke, C. K.; Kirchdoerfer, R. N.; Jungbauer, L. M.; Cavagnero, S. *ACS Chem. Biol.* **2008**, *3*, 555–566.
(9) Wang, J. Y.; Xie, J. M.; Schultz, P. G. *J. Am. Chem. Soc.* **2006**, *128*, 8738–8739.
(10) Charbon, G.; Brustad, E.; Scott, K. A.; Wang, J.; Lobner-Olesen, A.; Schultz, P. G.; Jacobs-Wagner, C.; Chapman, E. *ChemBioChem* **2011**, *12*, 1818–1821.
(11) Goerke, A. R.; Swartz, J. R. *Biotechnol. Bioeng.* **2009**, *102*, 400–416.
(12) Jewett, M. C.; Swartz, J. R. *Biotechnol. Bioeng.* **2004**, *86*, 19–26.
(13) Goerke, A. R.; Swartz, J. R. *Biotechnol. Bioeng.* **2008**, *99*, 351–367.
(14) Calhoun, K. A.; Swartz, J. R. *J. Biotechnol.* **2006**, *123*, 193–203.
(15) Young, T. S.; Ahmad, I.; Yin, J. A.; Schultz, P. G. *J. Mol. Biol.* **2010**, *395*, 361–374.
(16) Guo, J. T.; Melancon, C. E.; Lee, H. S.; Groff, D.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2009**, *48*, 9148–9151.
(17) Sando, S.; Ogawa, A.; Nishi, T.; Hayami, M.; Aoyama, Y. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1216–1220.
(18) Schaffitzel, C.; Ban, N. *J. Struct. Biol.* **2007**, *158*, 463–471.
(19) Evans, M. S.; Ugrinov, K. G.; Frese, M. A.; Clark, P. L. *Nat. Methods* **2005**, *2*, 757–762.
(20) Zhang, X.; Rashid, R.; Wang, K.; Shan, S. O. *Science* **2010**, *328*, 757–760.
(21) Zhang, X.; Schaffitzel, C.; Ban, N.; Shan, S. O. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 1754–1759.
(22) Janda, C. Y.; Li, J.; Oubridge, C.; Hernandez, H.; Robinson, C. V.; Nagai, K. *Nature* **2010**, *465*, 507–510.
(23) Hainzl, T.; Huang, S. H.; Merilainen, G.; Brannstrom, K.; Sauer-Eriksson, A. E. *Nat. Struct. Mol. Biol.* **2011**, *18*, 389–391.
(24) Bornemann, T.; Jockel, J.; Rodnina, M. V.; Wintermeyer, W. *Nat. Struct. Mol. Biol.* **2008**, *15*, 494–499.
(25) Hsu, S. T. D.; Cabrita, L. D.; Fucini, P.; Christodoulou, J.; Dobson, C. M. *J. Am. Chem. Soc.* **2009**, *131*, 8366–8367.
(26) O'Brien, E. P.; Christodoulou, J.; Vendruscolo, M.; Dobson, C. M. *J. Am. Chem. Soc.* **2011**, *133*, 513–526.
(27) Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11020–11024.