

Evolving tRNA^{Sec} for Efficient Canonical Incorporation of Selenocysteine

Ross Thyer,* Scott A. Robotham, Jennifer S. Brodbelt, and Andrew D. Ellington

Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, Texas 78712, United States

S Supporting Information

ABSTRACT: Bacterial selenocysteine incorporation occurs in response to opal stop codons and is dependent on the presence of a selenocysteine insertion sequence (SECIS) element, which recruits the selenocysteine specific elongation factor and tRNA^{Sec} needed to reassign the UGA codon. The SECIS element is a stem-loop RNA structure immediately following the UGA codon and forms part of the coding sequence in bacterial selenoproteins. Although the site specific incorporation of selenocysteine is of great interest for protein engineering, the sequence constraints imposed by the adjoining SECIS element severely limit its use. We have evolved an *E. coli* tRNA^{Sec} that is compatible with the canonical translation machinery and can suppress amber stop codons to incorporate selenocysteine with high efficiency. This evolved tRNA^{Sec} allows the production of new recombinant selenoproteins containing structural motifs such as selenyl-sulfhydryl and diselenide bonds.

Incorporation of the rare amino acid selenocysteine into proteins confers unique biophysical properties and is essential for life in organisms spanning all three domains.¹ Unlike the 20 canonical amino acids, selenocysteine lacks an aminoacyl-tRNA synthetase, and is instead a modification of a precharged serine and is inserted into proteins in response to opal stop codons. The overall mechanism for cotranslational incorporation at particular stop codons requires several specific cis and trans acting protein and RNA factors,² including a dedicated selenocysteine tRNA (tRNA^{Sec}), a selenophosphate synthase (SelD), and selenocysteine synthase (SelA), which are required to form Sec-tRNA^{Sec},³ a selenocysteine-specific elongation factor (SelB), and a stop codon-adjacent selenocysteine insertion sequence (SECIS) element that forms a conserved stem-loop RNA structure.⁴ SelB, which is structurally related to EF-Tu, is capable of discriminating between serylated and selenylated tRNA^{Sec},^{5,6} and the SelB:Sec-tRNA^{Sec} complex is recruited by the SECIS element during translation to facilitate recoding of the UGA stop codon.^{5,7}

Selenocysteine has a significantly lower pK_a than cysteine (5.2 vs 8.5 for free amino acid) and much stronger nucleophilic properties, making it an attractive target for altering protein chemistry and function. Unfortunately, most selenoproteins have proven difficult to produce in *E. coli*, the standard host for recombinant protein production.² This is due to the inherently low efficiency of selenocysteine incorporation in bacteria (4–5% vs termination of protein synthesis). In addition, the requirement

that the SECIS element immediately follows the UGA codon, forming part of the coding sequence, greatly limits which proteins are amenable to selenocysteine insertion.⁸ Recently, Aldag et al.⁹ developed a hybrid tRNA (designated tRNA^{UTu}) in which the acceptor stem of *E. coli* tRNA^{Ser} was replaced with that of tRNA^{Sec}, and the anticodon was changed to CUA to enable recognition of amber stop codons. Unlike wild type tRNA^{Sec}, the hybrid tRNA was a substrate for EF-Tu (rather than SelB) and was shown to be compatible with canonical translation, greatly reducing the sequence constraints for selenocysteine incorporation.

Unfortunately, compared to wild-type tRNA^{Sec}, selenylation of Ser-tRNA^{UTu} was impaired, and in turn, serine was incorporated at a significant rate (35–45%).⁹ We hypothesized that the impairment was due to the loss of important contacts between SelA and the D- and T-loops of tRNA^{Sec}.¹⁰ To overcome this problem, we used *E. coli* tRNA^{Sec} as a scaffold for mutagenesis to identify tRNA^{Sec} variants capable of participating in canonical translation. While the unusual 8 bp acceptor stem in tRNA^{Sec} was thought to impair interactions with EF-Tu,¹¹ Rudinger and co-workers¹² showed that extended acceptor stems were compatible with EF-Tu binding and that a specific antideterminant sequence in tRNA^{Sec} was instead responsible for blocking the interaction. This sequence corresponded to the final base pair of the acceptor stem (C₇-G₆₆) and to the first two base pairs of the T-arm (G₄₉-U₆₅ and C₅₀-G₆₄) (Figure 1c). Sequence changes at either location abolished antideterminant activity. Replacement of this sequence in tRNA^{Sec} was shown to yield a tRNA capable of minimal interaction with EF-Tu.¹² Similarly, a partially overlapping region of the T-arm covering base pairs 49–65, 50–64, and 51–63 is known to modulate the affinity of canonical tRNAs for EF-Tu.¹³ On the basis of this information, we opted to randomize the antideterminant region of tRNA^{Sec} to identify sequences capable of interaction with EF-Tu and able to participate in canonical translation.

While traditionally genetic code expansions have been evolved using reporter proteins containing amber stop codons,^{14,15} success is measured solely by the ability to make a full-length protein. Such selections are blind to the identity of the amino acid incorporated and rely on further negative selections to eliminate variants capable of nonspecific interactions. To prevent selection of tRNA^{Sec} variants that interact with EF-Tu but are poor substrates for SelA, we developed a novel genetic selection capable of discriminating different levels of selenocysteine incorporation. To specifically “addict” a reporter protein to

Received: October 17, 2014

Published: December 18, 2014

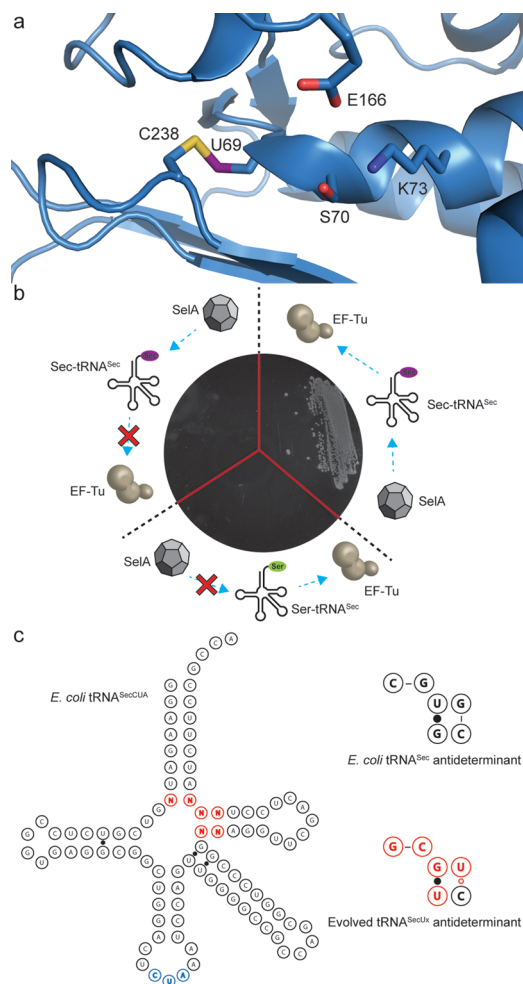


Figure 1. Selection of tRNAs capable of canonical incorporation of selenocysteine. (a) Representation of the NMC-A β -lactamase from *Enterobacter cloacae* (PDB: 1BUE) showing the engineered selenyl-sulfhydryl bond between residues 69 and 238 and its proximity to the catalytic site. (b) Validation of a β -lactamase reporter capable of discriminating serine from selenocysteine. Three NMC-A variants, C69X (TAG), C69, and C69S (clockwise from left), were constructed to determine the threshold for selection ($<50 \mu\text{g}\cdot\text{mL}^{-1}$). These approximate the following outcomes during selection: termination at the amber stop codon due to inefficient interaction with EF-Tu, impaired interaction with SelA promoting incorporation of serine, and efficient interactions with both SelA and EF-Tu resulting in incorporation of selenocysteine. (c) Representation of *E. coli* tRNA^{SecCUA} and the evolved anticodeterminant sequence. The anticodeterminant library is shown in red, and the CUA anticodon is shown in blue.

selenocysteine rather than serine we used the NMC-A β -lactamase from *Enterobacter cloacae*. This enzyme has high sequence similarity to the SME-1 β -lactamase from *Serratia marcescens*, an enzyme that has previously been shown to require a disulfide bond adjacent to the active site serine residue for activity but that confers a significant fitness cost on *E. coli*.¹⁶ We first constructed a C69S mutant of NMC-A, which failed to confer resistance to ampicillin ($\text{MIC} < 50 \mu\text{g}\cdot\text{mL}^{-1}$), indicating that the disulfide bond was essential for activity (Figure 1b). We then replaced cysteine 69 with an amber stop codon (X69; Figure 1a) for library selection, hypothesizing that the incorporation of selenocysteine and the formation of a selenyl-sulfhydryl bond would restore activity.¹⁷

To eliminate any crosstalk between the tRNA^{Sec} library and the endogenous selenocysteine incorporation machinery, the *selA*, *selB*, and *selC* genes (encoding SelA, SelB, and tRNA^{Sec}, respectively) were deleted from *E. coli* DH10B (designated DH Δ abc). Cells containing the reporter plasmid pNMC-A C69X and the accessory plasmid pRSF-eSelA (expressing SelA) were transformed with plasmid pMB1-ZU containing the tRNA^{Sec} anticodeterminant library. Transformants were plated on media containing a gradient of ampicillin concentrations for selection of mutants capable of selenocysteine-specific suppression. The single colonies that arose covered a range of ampicillin concentrations. Some 12 colonies from each plate were sequenced and revealed three distinct tRNA^{Sec} mutants: G₇-C₆₆:U₄₉-G₆₅:C₅₀-U₆₄, C₇-G₆₆:U₄₉-G₆₅:C₅₀-U₆₄, and C₇-U₆₆:U₄₉-A₆₅:A₅₀-A₆₄ (where underlined, bases represent changes from the parental anticodeterminant sequence). Of these tRNA^{Sec} variants, only G₇-C₆₆:U₄₉-G₆₅:C₅₀-U₆₄ was detected at the two highest ampicillin concentrations (200 and 250 $\mu\text{g}\cdot\text{mL}^{-1}$).

The tRNA^{Sec} variant containing the G₇-C₆₆:U₄₉-G₆₅:C₅₀-U₆₄ anticodeterminant sequence was designated tRNA^{SecUx} and was compared with the previously designed chimera (tRNA^{UTu}) and with a tRNA^{Sec} derivative designed to have an anticodeterminant region that should tightly bind EF-Tu (tRNA^{UG}; Figure 2a). The parental tRNA^{Sec} containing a CUA anticodon and tRNA^{UG} failed to produce active β -lactamase. Cells containing the hybrid tRNA^{UTu} incorporated selenocysteine and could grow on 75 $\mu\text{g}\cdot\text{mL}^{-1}$. In contrast, expression of tRNA^{SecUx} resulted in significantly higher β -lactamase activity (up to 400 $\mu\text{g}\cdot\text{mL}^{-1}$), but only when coexpressed with SelA, confirming activity was selenocysteine dependent. To further confirm tRNA^{SecUx} incorporated selenocysteine in response to amber stop codons we employed a standard colorimetric assay based on the activity of the endogenous *E. coli* selenoprotein formate dehydrogenase H (FdhH) (Figure 2b). FdhH is expressed under anaerobic conditions and catalyzes the oxidation of formate to produce CO₂ with the concomitant reduction of the electron acceptor benzyl viologen resulting in the development of a deep purple color.¹⁸ Formate oxidation by FdhH is strictly dependent on the selenocysteine residue at position 140; the mutant FdhH U140S was completely inactive. Only tRNA^{SecUx} and tRNA^{UTu} when coexpressed with SelA produced active FdhH.

The selected tRNA contained a nonstandard sequence in the junction that normally interacts with EF-Tu. Given that neither the base of the acceptor stem nor the adjoining T-arm base pairs are believed to play a role in the interaction between tRNA^{Sec} and SelA, our results suggest that the selected C:U leads to stronger binding to EF-Tu than the wild-type tRNA^{Sec} sequence.¹⁰ That said, the unusual C₅₀-U₆₄ base pair is not predicted to bind strongly to EF-Tu based on models developed for canonical tRNAs,¹³ and expression of a hybrid tRNA^{UG} containing the strong EF-Tu binding region from the major *E. coli* tRNA^{Gly} did not lead to the production of active β -lactamase, suggesting that the nonstandard sequence was functionally important. Thus, it is possible that portions of the engineered tRNA^{Sec} bind to EF-Tu differently than do canonical tRNAs, which would not necessarily be surprising given that tRNA^{Sec} normally interacts with SelB.²³

The development of engineered *E. coli* strains lacking the *prfA* gene encoding release factor 1 (RF1) has allowed efficient incorporation of a range of unnatural amino acids,^{19,20} and the development of the genome-engineered “Amberless” *E. coli* C321. $\Delta\Delta$ ²⁰ provided an excellent opportunity to determine whether we could express proteins that efficiently incorporated selenocysteine. The *selA*, *selB*, and *selC* genes were deleted in

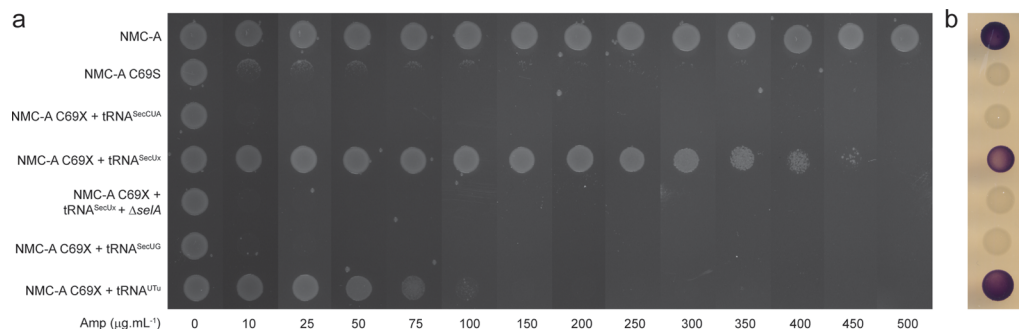


Figure 2. Activity of tRNA^{Sec} variants in canonical translation. (a) NMC-A β-lactamase reporter assay. From top row: Wild type NMC-A β-lactamase, NMC-A C69S, NMC-A C69X with tRNA^{SecCUA}, NMC-A C69X with tRNA^{SecUx}, NMC-A C69X with tRNA^{SecUx} and inactive SelA, NMC-A C69X with tRNA^{SecUG} containing a strong EF-Tu binding sequence, and NMC-A C69X with tRNA^{UTu}. The observed dynamic range of the selection exceeds 20-fold. (b) Reduction of benzyl viologen by *E. coli* formate dehydrogenase H confirms selenocysteine is incorporated by the tRNA^{SecUx}. Plate layout is identical to panel a with FdhH, FdhH U140S, and FdhH U140TAG reporters replacing the NMC-A reporter.

C321.ΔA (designated strain RTΔA), and cells were transformed with the amber-containing NMC-A reporter and accessory plasmids (Figure S1, Supporting Information (SI)). β-lactamase activity was dramatically increased in RF1-deficient cells compared to *prfA*⁺ DHΔabc cells that still contain RF1. In addition, in a RF1-deficient background tRNA^{SecUx} could now support the formation of a functional diselenide bond (via amber-mediated incorporation of two selenocysteine residues, U69 and U238; Figure S1, SI).

To further enhance the efficiency of selenocysteine incorporation, a number of steps were taken to improve the levels of Sec-tRNA^{Sec} relative to Ser-tRNA^{Sec}, including increasing the level of SelA, decreasing the gene dose of tRNA^{SecUx}, and coexpressing a phosphoseryl-tRNA^{Sec} kinase (Methods, SI). To monitor the efficiency of selenocysteine incorporation and demonstrate the possibilities for protein engineering, *E. coli* dihydrofolate reductase (DHFR) was produced containing an engineered nonessential selenyl-sulfhydryl bond.²¹ Top down mass spectrometry showed close to 100% selenocysteine incorporation with no detectable background corresponding to DHFR containing serine (Figure 3). The rationally designed tRNA^{UTu} chimera was also observed to incorporate selenocysteine in DHFR containing a P39X substitution, but resulted in a much lower level of selenocysteine incorporation (38%) and significant serine incorporation (62%) (Figure S1, SI). No masses corresponding to the incorporation of other canonical amino acids were observed in the mass spectra (Figure S4, SI). In order to further validate selenocysteine incorporation, we also expressed the *Pseudomonas aeruginosa* metalloprotein azurin with its essential cysteine (C112) replaced by selenocysteine and the human selenoprotein cellular glutathione peroxidase (GPx-1) (Figures S3 and S5, SI). For azurin, this chemical change had previously been proven possible only through expressed protein ligation;²² the essential cysteine could now be biologically replaced with selenocysteine with good efficiency as measured by mass spectrometry of the intact protein.

In this study we developed a new and robust genetic selection that relies on an engineered, essential selenyl-sulfhydryl bond in NMC-A β-lactamase. The selection system has a wide dynamic range and can function independently of the host genetic background. A new tRNA^{Sec} variant with a nonstandard T-arm sequence and dramatically improved activity in canonical translation was evolved. A simple two-plasmid system carrying the new tRNA enabled highly efficient, site-specific incorporation of selenocysteine under culture conditions optimal for *E. coli* with no observable toxicity. Using this system, we have

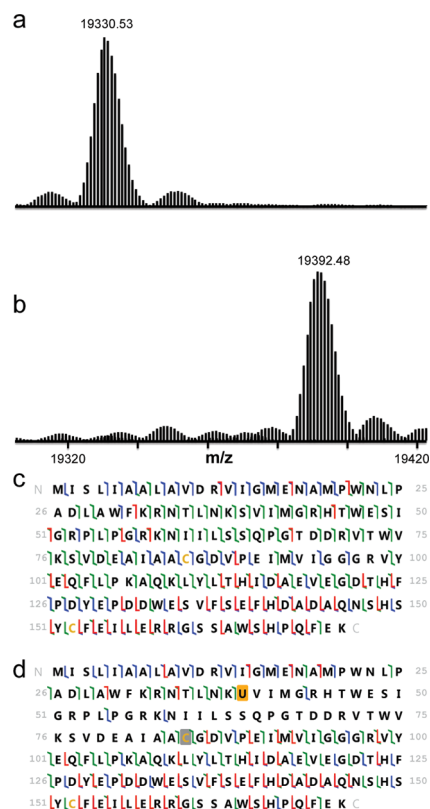


Figure 3. Intact mass and 193 nm UVPD results are shown for DHFR and seleno-DHFR produced using tRNA^{SecUx}. For the sequence information shown, disulfide-bound cysteines are shaded in gray, and selenocysteine is shaded in orange. When selenocysteine is located at position 39, it forms a selenyl-sulfhydryl bond with the cysteine at position 85. The different colors of slash marks represent the cleavage sites that lead to different N-terminal and C-terminal ions (green is used for a and x ions, blue is for b and y ions, and red is for c and z ions). (a) Expanded region of the deconvoluted mass spectrum of DHFR with serine at position 39 (expected average mass *m/z* 19331.72). (b) Expanded region of the deconvoluted mass spectrum of DHFR with selenocysteine incorporated at position 39 using tRNA^{SecUx} (expected average mass *m/z* 19392.66). (c) UVPD fragmentation map of DHFR P39S obtained for the 20+ charge state. (d) UVPD fragmentation map of DHFR with selenocysteine at position 39 (18+ charge state).

demonstrated the production of recombinant selenoproteins containing structural motifs such as selenyl-sulfhydryl and diselenide bonds at yields comparable to those obtained using

expressed protein ligation. Such motifs have a number of protein engineering applications including bond formation in reducing intracellular environments and directing ordered folding of peptide scaffolds.^{24,25} Furthermore, this system will complement the development of new RF1-deficient strains derived from genetic backgrounds optimized for protein production,²⁶ forming a powerful platform for numerous synthetic biology and biotechnology applications as well as a useful tool for potentially exploring the evolutionary expansion of the bacterial selenoproteome.

■ ASSOCIATED CONTENT

📄 Supporting Information

Experimental details and figures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*ross.thyer@utexas.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Funding from the NSF (CHE-1402753) and the Welch Foundation (F-1155) is acknowledged.

■ REFERENCES

- (1) Copeland, P. R. *Genome Biol.* **2005**, *6*, 27.
- (2) Arner, E. S.; Sarioglu, H.; Lottspeich, F.; Holmgren, A.; Bock, A. *J. Mol. Biol.* **1999**, *292*, 1003.
- (3) Yuan, J.; O'Donoghue, P.; Ambrogelly, A.; Gundllapalli, S.; Sherrer, R. L.; Palioura, S.; Simonovic, M.; Soll, D. *FEBS Lett.* **2010**, *584*, 342.
- (4) Mansell, J. B.; Guevremont, D.; Poole, E. S.; Tate, W. P. *EMBO J.* **2001**, *20*, 7284.
- (5) Leibundgut, M.; Frick, C.; Thanbichler, M.; Bock, A.; Ban, N. *EMBO J.* **2005**, *24*, 11.
- (6) Paleskava, A.; Konevega, A. L.; Rodnina, M. V. *J. Biol. Chem.* **2010**, *285*, 3014.
- (7) Yoshizawa, S.; Rasubala, L.; Ose, T.; Kohda, D.; Fourmy, D.; Maenaka, K. *Nat. Struct. Mol. Biol.* **2005**, *12*, 198.
- (8) Suppmann, S.; Persson, B. C.; Bock, A. *EMBO J.* **1999**, *18*, 2284.
- (9) Aldag, C.; Brocker, M. J.; Hohn, M. J.; Prat, L.; Hammond, G.; Plummer, A.; Soll, D. *Angew. Chem., Int. Ed.* **2013**, *52*, 1441.
- (10) Itoh, Y.; Brocker, M. J.; Sekine, S.; Hammond, G.; Suetsugu, S.; Soll, D.; Yokoyama, S. *Science* **2013**, *340*, 75.
- (11) Baron, C.; Bock, A. *J. Biol. Chem.* **1991**, *266*, 20375.
- (12) Rudinger, J.; Hillenbrandt, R.; Sprinzl, M.; Giege, R. *EMBO J.* **1996**, *15*, 650.
- (13) Schrader, J. M.; Uhlenbeck, O. C. *Nucleic Acids Res.* **2011**, *39*, 9746.
- (14) Wang, L.; Schultz, P. G. *Chem. Biol.* **2001**, *8*, 883.
- (15) Ellefson, J. W.; Meyer, A. J.; Hughes, R. A.; Cannon, J. R.; Brodbelt, J. S.; Ellington, A. D. *Nat. Biotechnol.* **2014**, *32*, 97.
- (16) Majiduddin, F. K.; Palzkill, T. *Antimicrob. Agents Chemother.* **2003**, *47*, 1062.
- (17) Swarén, P.; Maveyraud, L.; Raquet, X.; Cabantous, S.; Duez, C.; Pédelaçq, J.-D.; Mariotte-Boyer, S.; Mourey, L.; Labia, R.; Nicolas-Chanoine, M.-H.; Nordmann, P.; Frère, J.-M.; Samama, J.-P. *J. Biol. Chem.* **1998**, *273*, 26714.
- (18) Zinoni, F.; Birkmann, A.; Leinfelder, W.; Bock, A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3156.
- (19) Mukai, T.; Hayashi, A.; Iraha, F.; Sato, A.; Ohtake, K.; Yokoyama, S.; Sakamoto, K. *Nucleic Acids Res.* **2010**, *38*, 8188.
- (20) Lajoie, M. J.; Rovner, A. J.; Goodman, D. B.; Aerni, H. R.; Haimovich, A. D.; Kuznetsov, G.; Mercer, J. A.; Wang, H. H.; Carr, P. A.; Mosberg, J. A.; Rohland, N.; Schultz, P. G.; Jacobson, J. M.; Rinehart, J.; Church, G. M.; Isaacs, F. J. *Science* **2013**, *342*, 357.

(21) Villafranca, J. E.; Howell, E. E.; Oatley, S. J.; Xuong, N. H.; Kraut, J. *Biochemistry* **1987**, *26*, 2182.

(22) Clark, K. M.; van der Donk, W. A.; Lu, Y. *Methods Enzymol.* **2009**, *462*, 97.

(23) Li, W. Q.; Yarus, M. *J. Mol. Biol.* **1992**, *223*, 9.

(24) Shchedrina, V. A.; Novoselov, S. V.; Malinouski, M. Y.; Gladyshev, V. N. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 13919.

(25) Armishaw, C. J.; Daly, N. L.; Nevin, S. T.; Adams, D. J.; Craik, D. J.; Alewood, P. F. *J. Biol. Chem.* **2006**, *281*, 14136.

(26) Isaacs, F. J.; Carr, P. A.; Wang, H. H.; Lajoie, M. J.; Sterling, B.; Kraal, L.; Tolonen, A. C.; Gianoulis, T. A.; Goodman, D. B.; Reppas, N. B.; Emig, C. J.; Bang, D.; Hwang, S. J.; Jewett, M. C.; Jacobson, J. M.; Church, G. M. *Science* **2011**, *333*, 348.