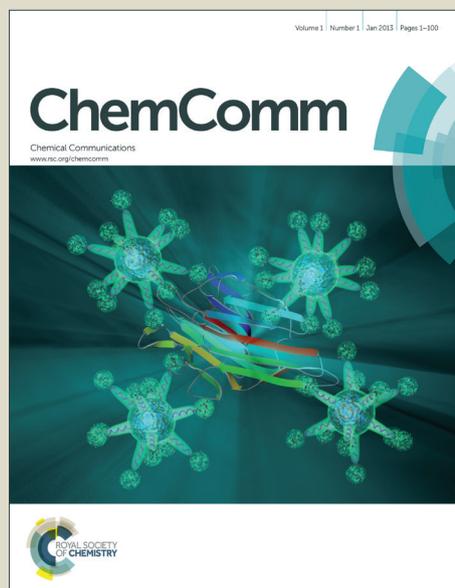


ChemComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Genetically Encoded Unstrained Olefins for Live Cell Labeling with Tetrazine Dyes

Yan-Jiun Lee,^a Yadagiri Kurra,^a Yanyan Yang,^a Jessica Torres-Kolbus,^b Alexander Deiters,^{b,c} and Wenshe R. Liu^{a*}⁵ Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

DOI: 10.1039/b000000x

A number of non-canonical amino acids (NCAAs) with unstrained olefins are genetically encoded using mutant pyrrolysyl-tRNA synthetase-tRNA^{Pyl}_{CUA} pairs. These NCAAs readily undergo inverse electron-demand Diels-Alder cycloadditions with tetrazine dyes, leading to selective labeling of proteins bearing these NCAAs in live cells.

Owing to its high selectivity, fast reaction kinetics, and non-catalytic nature, the inverse electron-demand Diels-Alder cycloaddition reaction between a tetrazine and an olefin has emerged as a state-of-the-art approach for selective bioconjugation in live cells.¹ Previous efforts focused on applications of strained cyclic olefins or alkynes that react rapidly with tetrazines.² However, similar applications with unstrained olefins have been largely overlooked except a recent report of metabolic glycan labeling,³ due to the common notion that the reaction between an unstrained olefin and a tetrazine is not kinetically favored. Here, we systematically analysed unstrained olefin-tetrazine reactions and showed that several unstrained olefins readily reacted with a tetrazine dye and a number of unstrained olefin-bearing NCAAs can be genetically incorporated into proteins in *Escherichia coli* for selective labeling with tetrazine dyes.

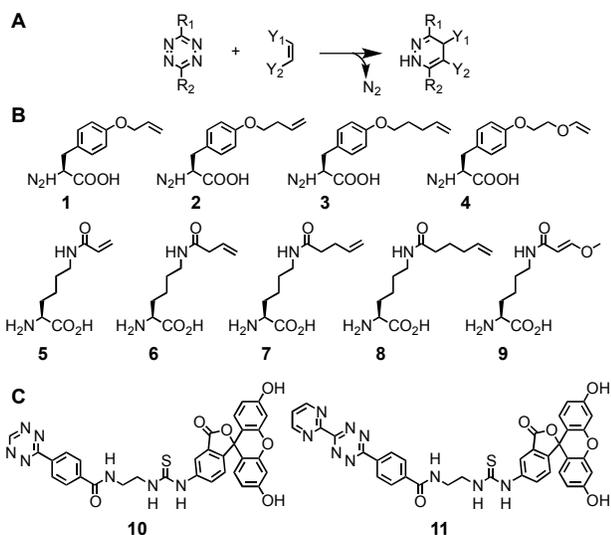


Figure 1. (A) The tetrazine-olefin reaction; (B) unstrained olefin-bearing NCAAs; (C) two fluorescein-tetrazine dyes used in this study.

We first characterized the reaction kinetics between a tetrazine dye **10** (Figure 1) and a number of unstrained olefins in phosphate buffered saline (PBS) at ambient temperature. Three alkenyl

alcohols (allyl alcohol, 3-buten-1-ol, and 4-penten-1-ol) and NCAAs **5-8** (Figure 1) were chosen to undergo kinetic analysis. The alkenyl alcohols were used instead of their corresponding NCAAs **1-3** due to their high solubility in PBS buffer. Both **2** and **3** could hardly dissolve to more than 5 mM in PBS buffer. With this low solubility, their reactions with **10** were too long for reliable data collection. To test whether terminal and non-terminal olefins react differently with **10**, 2-buten-1-ol was also included in the analysis. Since the inverse electron-demand Diels-Alder reaction is highly influenced by the dienophile electron density, vinyl ethers with electron-rich olefins, 2-hydroxyethyl vinyl ether that corresponds to NCAA **4**, and **9** were also selected to undergo assays with **10**. Conjugation of fluorescein to a tetrazine moiety such as in **10** inherently quenches the fluorescence from the fluorophore.^{1d} Cycloaddition of the tetrazine moiety and an olefin abolishes this intrinsic quenching effect, leading to the rescue of fluorescence that can be readily detected with a fluorospectrometer. The reactions of all olefins with **10** were carried out in pseudo-first-order conditions in which the olefin concentration was at least 20-fold higher than the concentration of **10** (Supplementary Figures 1-10). The determined second-order reaction rate constants are presented in Table 1.

Table 1. Determined second-order reaction rate constants between various olefin dienophiles and **10**.

Entry	Olefin	k ($10^3 M^{-2} s^{-1}$)
a		9.4 ± 1.2
b		1.9 ± 0.1
c		26 ± 5
d		36 ± 4
e		81 ± 1
f		1.2 ± 0.1
g		1.7 ± 0.2
h		11 ± 2
i		16 ± 1
j		19 ± 1

As shown in Table 1, all tested unstrained olefins reacted

readily with **10** with second-order rate constants more than $0.001 \text{ M}^{-1}\text{s}^{-1}$. Both inductive and resonance effects of an olefin substituent significantly influence its reactivity with respect to **10**.⁴ Electron withdrawing effects of the hydroxyl group and the amide group gradually diminish, correlating well to gradually improved reactivities from allyl alcohol to 4-penten-1-ol and **6** to **8**, respectively. Two vinyl ethers, 2-hydroxyethyl vinyl ether and **9**, display relatively higher reactivities toward **10** in comparison to other tested olefins due to the strong electron donating resonance effect of the vinyl ether oxygen atom. The sluggish reaction between **5** and **10** is attributed to the electron withdrawing resonance effect of the amide group. In comparison to the allyl alcohol, 2-buten-1-ol reacts significantly slower with **10**, indicating that the methyl substituent adds steric hindrance. Although our kinetic analysis indicates much slower reactions with a tetrazine for unstrained olefins compared to strained ones,^{1e} the determined second-order rate constants are comparable to those of the Staudinger ligation ($k = 0.002 \text{ M}^{-1}\text{s}^{-1}$)⁵ and the copper-free dibenzocyclooctyne-azide cycloaddition ($k = 0.0565 \text{ M}^{-1}\text{s}^{-1}$),⁶ two reactions that have been well adopted for bioconjugation in live cells. In comparison to strained olefins, unstrained olefins can be more easily prepared and are generally more stable toward cellular nucleophiles (except **5**).⁷ Their genetic installation in proteins followed by selective bioconjugation with tetrazine dyes can potentially provide a simple and readily adoptable protein labeling approach in live cells.

Encouraged by our kinetic analysis, we proceeded to recombinantly synthesize proteins carrying site-specifically incorporated unstrained olefins in *E. coli*. We previously reported that a rationally designed pyrrolysyl-tRNA synthetase (PylRS) mutant with two mutations, N346A and C348A, (PylRS(N346A/C348A)) together with $\text{tRNA}_{\text{CUA}}^{\text{Pyl}}$ is able to mediate the genetic incorporation of **1-4** into proteins at an amber codon in *E. coli*.⁸ This system was applied to synthesize four superfolder green fluorescent protein (sfGFP) variants with **1, 2, 3, or 4** incorporated at their S2 positions (sfGFP-1, sfGFP-2, sfGFP-3, and sfGFP-4). To synthesize the sfGFP variant with **5** incorporated at its S2 site (sfGFP-5), a previously evolved PylRS mutant PrKRS together with $\text{tRNA}_{\text{CUA}}^{\text{Pyl}}$ was used.⁹ Although systems for the genetic incorporation of **6-9** has not been described in the literature, our studies indicated that another previously evolved PylRS mutant BuKRS is able to recognize them as substrates.⁹ BuKRS together with $\text{tRNA}_{\text{CUA}}^{\text{Pyl}}$ was therefore applied for the synthesis of sfGFP variants with **6, 7, 8, and 9** at their S2 positions (sfGFP-6, sfGFP-7, sfGFP-8, and sfGFP-9) and the site-specific incorporation of **6-9** was verified by electrospray ionization mass spectrometry (ESI-MS) analysis (Supplementary Figures 11-14). The sfGFP variants were purified and then subjected to the fluorescent labeling with **10** at physiological pH in PBS buffer before they were denatured and analyzed by SDS-PAGE. Direct in-gel fluorescence detection showed strong fluorescent labeling of sfGFP-2, sfGFP-3, sfGFP-7, and sfGFP-8 (Figure 2). Labeling of sfGFP-1, sfGFP-5, and sfGFP-6 could be barely detected and, importantly, no labeling was observed for wild-type sfGFP. This labeling observation is consistent with the kinetic results in which 3-buten-1-ol, 4-penten-1-ol, **7**, and **8** react more favorably with **10** compared to allyl alcohol, **5**, and **6**. Although 2-hydroxyethyl vinyl ether displayed fast kinetics when reacted with **10**, sfGFP-4 did not show detectable labeling with **10** on the gel. This is likely due to the cleavage of the cycloaddition product from sfGFP-4 to generate an aromatized 1,2-diazine-fluorescein final product and a sfGFP-4 derivative with its vinyl group removed.⁴ Although a similar diazine formation reaction that followed the reaction between sfGFP-9 and **10** would retain the 1,2-diazine-fluorescein

moiety to sfGFP-9, we did not observed fluorescein-labeled sfGFP-9 on the gel. Since a fluorescent small molecule band was detected in the gel (Supplementary Figure 15), we suspect that the final fluorescein-labeled sfGFP-9 was not stable and hydrolysed during the SDS-PAGE analysis that required harsh treatments such as heating at 100 degree for 5 min. In gels, sfGFP-2 and sfGFP-7 displayed integrated fluorescence intensities that were about 40% and 80% in comparison to sfGFP-3 and sfGFP-8 after their labeling with **10**. The labeled sfGFP-3 was further subjected to the ESI-MS analysis that indicated close to 50% labeling efficiency (Supplementary Figure 16).

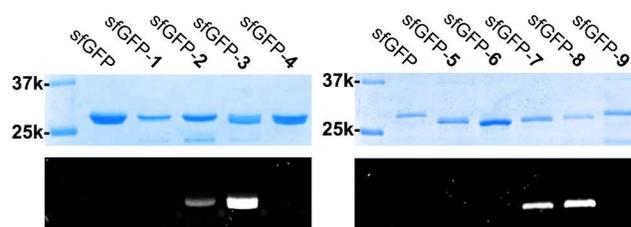


Figure 2. Labeling of sfGFP variants with **10**. The top panel shows Coomassie blue stained gels and the bottom panel presents the fluorescent imaging of the same gels before they were stained with Coomassie blue. Each protein (35 μM) was incubated with 0.5 mM of **10** in a 50 mM Tris-Cl buffer at pH 7.4 for 8 h before they were precipitated in 10% TFA, fully denatured with heating at 100 $^{\circ}\text{C}$ for 5 min, and then analyzed by SDS-PAGE. sfGFP with no NCAA incorporated was used as a control.

With the demonstration of the tetrazine reaction of unstrained alkenes for protein labeling *in vitro*, we then proceeded to test this reaction to label proteins bearing unstrained alkenes in live cells. Two previously constructed plasmids pEVOL-pylT-PylRS(N346A/C348A)^{8b} and pETDuet-OmpXTAG⁹ were used to transform *E. coli* BL21 cells to express an *E. coli* outer membrane protein OmpX with **3** incorporated at its extracellular domain (OmpX-3). Plasmid pEVOL-pylT-PylRS(N346A/C348A) carried genes for PylRS(N346A/C348A) and $\text{tRNA}_{\text{CUA}}^{\text{Pyl}}$ and plasmid pETDuet-OmpXTAG contained a gene coding OmpX with an AAAXAA (A denotes alanine and X denotes an amber mutation) insertion between two extracellular residues 53 and 54. The transformed cells were grown in M9 minimal medium supplemented with 1% glycerol, 1 mM IPTG, 0.2% arabinose, and 5 mM **3** to express OmpX with **3** incorporated. After the expression of OmpX-3, cells were washed with PBS buffer and incubated with **10** (0.5 mM) in PBS buffer at room temperature for 8 h. The residual dye was removed by washing cells with isotonic saline. Finally the labeled cells were subjected to fluorescent microscope imaging. *E. coli* cells expressing OmpX-3 showed detectable fluorescence (Figure 3). In contrast, control cells that were grown in the absence of **3** did not display any fluorescence. A similar labeling reaction was carried out to label *E. coli* cells that expressed OmpX with **8** incorporated (OmpX-8). To express OmpX-8, plasmids pEVOL-pylT-BuKRS and pETDuet-OmpXTAG were used to transform *E. coli* BL21 cells. Plasmid pEVOL-pylT-BuKRS carried genes coding BuKRS and $\text{tRNA}_{\text{CUA}}^{\text{Pyl}}$. OmpX-8 was expressed similarly as OmpX-3 except that the expression was done in LB medium. Since we noticed some background reaction between **10** and bovine serum albumin, we employed another tetrazine-fluorescein dye **11** that displayed a much lower reactivity toward cysteine than **10** for the following cell labeling analysis (Supplementary Figure 17). Cells were incubated with **11** (0.5 mM) for 8 h, washed with isotonic saline, and then subjected to fluorescent microscope imaging. Cells expressing OmpX-8 displayed strong green fluorescence. However, cells grown in the absence of **8** were not visibly labeled

with **11**. These results indicate that the unstrained alkene labeling reaction with a tetrazine is biocompatible and highly selective.

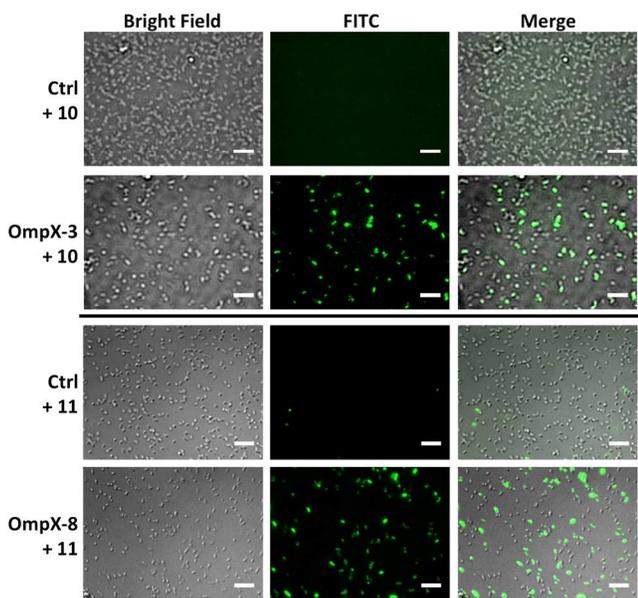


Figure 3. Selective labeling of *E. coli* cells expressing OmpX-3 and OmpX-8. Labeling reactions were carried out in the presence of 0.5 mM **10** or **11** in the PBS buffer at pH 7.4 for 8 h. *E. coli* cells grown in corresponding media without NCAA supplemented served as controls (Ctrls). The left panels show bright-field images of *E. coli* cells, the middle panels show green fluorescent images of the same cells, and the right panels shows composite images of bright-field and fluorescent images. Scale bars represent 10 μ m.

In summary, we demonstrate that an unstrained olefin can react with a tetrazine dye with an appreciable reaction rate that allows selective labeling of proteins on the surface of live bacterial cells. Given its small size, easy preparation, and high stability, an unstrained olefin provides a significant advantage over a strained olefin in promoting the general adoption of the tetrazine labeling approach for cell biology studies. With a size comparable to alkyne and azide, an unstrained olefin can be similarly integrated into metabolic surrogates for cell labeling.^{3, 5, 10} Its biocompatible reaction with a tetrazine probe potentially provides an alternative tool for metabolic labeling research. That can be applied to investigate the temporal and spatial synthesis of DNA, RNA, proteins, and carbohydrates in live cells.

We acknowledge financial support from the National Institute of Health (grant 1R01CA161158 to W.R.L.), the National Science Foundation (grants CHE-1148684 to W.R.L. and CHE-0848398 to A.D.), and the Welch Foundation (grant A-1715 to W.R.L.). J.T.K. acknowledges support from the National Science Foundation Graduate Research Fellowship (grant NSF-0750733). We also thank Dr. Yohannes H. Rezenom from the Laboratory for Biological Mass Spectrometry at Texas A&M University for characterizing our proteins by electrospray ionization mass spectrometry.

Notes and references

^a Department of Chemistry, Texas A&M University, College Station, TX 77843 (USA); wliu@chem.tamu.edu.

^b Department of Chemistry, North Carolina State University, Raleigh, NC 27695 (USA).

^c Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260 (USA).

Electronic Supplementary Information (ESI) available: Synthesis, protein expression, protein labeling, and mass spectrometry analysis. See DOI: 10.1039/c000000x/

- 45
1 (a) F. Thalhammer, U. Wallfahner and J. Sauer, *Tetrahedron Lett*, 1990, **31**, 6851; (b) A. C. Knall and C. Slugovc, *Chem Soc Rev*, 2013, **42**, 5131; (c) M. F. Debets, S. S. van Berkel, J. Dommerholt, A. T. Dirks, F. P. Rutjes and F. L. van Delft, *Acc Chem Res*, 2011, **44**, 805;
50 (d) N. K. Devaraj and R. Weissleder, *Acc Chem Res*, 2011, **44**, 816; (e) M. R. Karver, R. Weissleder and S. A. Hilderbrand, *Bioconjugate Chem*, 2011, **22**, 2263; (f) J. C. Carlson, L. G. Meimetis, S. A. Hilderbrand and R. Weissleder, *Angew Chem Int Ed Engl*, 2013, **52**, 6917.
55 2 (a) M. L. Blackman, M. Royzen and J. M. Fox, *J Am Chem Soc*, 2008, **130**, 13518; (b) N. K. Devaraj, R. Weissleder and S. A. Hilderbrand, *Bioconjugate Chem*, 2008, **19**, 2297; (c) M. Royzen, G. P. Yap and J. M. Fox, *J Am Chem Soc*, 2008, **130**, 3760; (d) N. K. Devaraj, S. Hilderbrand, R. Upadhyay, R. Mazitschek and R. Weissleder, *Angew Chem Int Ed Engl*, 2010, **49**, 2869; (e) M. T. Taylor, M. L. Blackman, O. Dmitrenko and J. M. Fox, *J Am Chem Soc*, 2011, **133**, 9646; (f) K. Lang, L. Davis, S. Wallace, M. Mahesh, D. J. Cox, M. L. Blackman, J. M. Fox and J. W. Chin, *J Am Chem Soc*, 2012, **134**, 10317; (g) Y. Liang, J. L. Mackey, S. A. Lopez, F. Liu and K. N. Houk, *J Am Chem Soc*, 2012, **134**, 17904; (h) D. S. Liu, A. Tangpeerachaikul, R. Selvaraj, M. T. Taylor, J. M. Fox and A. Y. Ting, *J Am Chem Soc*, 2012, **134**, 792; (i) J. L. Seitchik, J. C. Peeler, M. T. Taylor, M. L. Blackman, T. W. Rhoads, R. B. Cooley, C. Refakis, J. M. Fox and R. A. Mehl, *J Am Chem Soc*, 2012, **134**, 2898; (j) K. Lang, L. Davis, J. Torres-Kolbus, C. Chou, A. Deiters and J. W. Chin, *Nat Chem*, 2012, **4**, 298; (k) T. Plass, S. Milles, C. Koehler, J. Szymanski, R. Mueller, M. Wiessler, C. Schultz and E. A. Lemke, *Angew Chem Int Ed Engl*, 2012, **51**, 4166; (l) A. Borrmann, S. Milles, T. Plass, J. Dommerholt, J. M. Verkade, M. Wiessler, C. Schultz, J. C. van Hest, F. L. van Delft and E. A. Lemke, *Chembiochem*, 2012, **13**, 2094; (m) W. Wan, J. M. Tharp and W. R. Liu, *Biochim Biophys Acta*, 2014, **1844**, 1059.
60 3 A. Niederwieser, A. K. Spate, L. D. Nguyen, C. Jungst, W. Reutter and V. Wittmann, *Angew Chem Int Ed Engl*, 2013, **52**, 4265.
80 4 (a) A. Meier and J. Sauer, *Tetrahedron Lett*, 1990, **31**, 6855; (b) E. G. Kovalev, I. Y. Postovskii, G. L. Rusinov and I. L. Shegal, *Chem Heterocycl Compd*, 1981, **17**, 1063.
5 E. Saxon and C. R. Bertozzi, *Science*, 2000, **287**, 2007.
6 (a) M. F. Debets, S. S. van Berkel, S. Schoffelen, F. P. Rutjes, J. C. van Hest and F. L. van Delft, *Chem Commun*, 2010, **46**, 97; (b) N. J. Agard, J. A. Prescher and C. R. Bertozzi, *J Am Chem Soc*, 2004, **126**, 15046; (c) J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli and C. R. Bertozzi, *Proc Natl Acad Sci U S A*, 2007, **104**, 16793; (d) C. G. Gordon, J. L. Mackey, J. C. Jewett, E. M. Sletten, K. N. Houk and C. R. Bertozzi, *J Am Chem Soc*, 2012, **134**, 9199; (e) J. C. Jewett, E. M. Sletten and C. R. Bertozzi, *J Am Chem Soc*, 2010, **132**, 3688.
7 R. van Geel, G. J. Pruijn, F. L. van Delft and W. C. Boelens, *Bioconjugate Chem*, 2012, **23**, 392.
95 8 (a) Y. S. Wang, X. Fang, A. L. Wallace, B. Wu and W. R. Liu, *J Am Chem Soc*, 2012, **134**, 2950; (b) Y. S. Wang, X. Fang, H. Y. Chen, B. Wu, Z. U. Wang, C. Hilty and W. R. Liu, *ACS Chem Biol*, 2013, **8**, 405; (c) A. Tuley, Y. S. Wang, X. Fang, Y. Kurra, Y. H. Rezenom and W. R. Liu, *Chem Commun*, 2014, **50**, 2673; (d) J. M. Tharp, Y. S. Wang, Y. J. Lee, Y. Yang and W. R. Liu, *ACS Chem Biol*, 2014, **9**, 884.
9 Y. J. Lee, B. Wu, J. E. Raymond, Y. Zeng, X. Fang, K. L. Wooley and W. R. Liu, *ACS Chem Biol*, 2013, **8**, 1664.
10 (a) R. Breinbauer and M. Kohn, *Chembiochem*, 2003, **4**, 1147; (b) Q. Wang, T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless and M. G. Finn, *J Am Chem Soc*, 2003, **125**, 3192; (c) P. M. Clark, J. F. Dweck, D. E. Mason, C. R. Hart, S. B. Buck, E. C. Peters, B. J. Agnew and L. C. Hsieh-Wilson, *J Am Chem Soc*, 2008, **130**, 11576; (d) D. C. Dieterich, A. J. Link, J. Graumann, D. A. Tirrell and E. M. Schuman, *Proc Natl Acad Sci U S A*, 2006, **103**, 9482.
110