DOI: 10.1002/anie.200905590

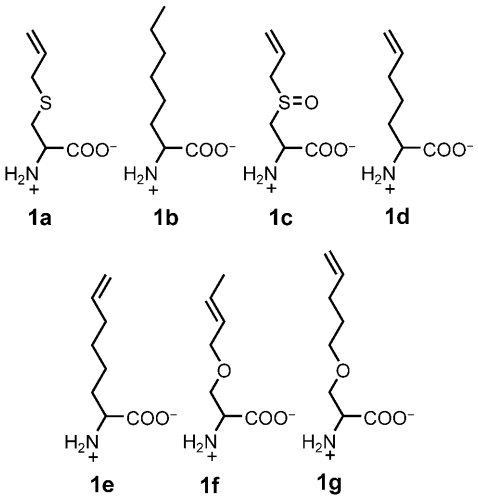
Expanded Genetic Code

Genetically Encoded Alkenes in Yeast**

*Hui-wang Ai, Weijun Shen, Eric Brustad, and Peter G. Schultz**

Several bioorthogonal chemical reactions have been explored for the selective modification of proteins,[1,2] including the coupling of alkoxyamines and hydrazides to ketones or aldehydes,[3] the Staudinger ligation of azides to modified phosphines,[4] and click reactions between azides and alkynes.[5] Recently, alkene moieties have also been exploited as uniquely reactive chemical handles. Examples include, the photoaddition of a diaryl tetrazole to alkenes;[6,7] a Diels–Alder reaction between tetrazines and *trans*-cyclooctenes;[8] the cross-metathesis of olefins with allyl thioether modified proteins;[9,10] the copolymerization of alkene-containing proteins and acrylamide;[11] and the coupling of two alkene-containing residues in peptides, resulting in improved stability and pharmacological properties.[12–15]

In addition to chemical semisynthesis, a number of in vitro and in vivo methods have been developed to incorporate the bioorthogonal alkene groups into proteins. For example, Davis and co-workers developed a variety of in vitro chemical methods to convert cysteine residues in proteins into *S*-allylcysteine (**1a**), a reactive cross-metathesis substrate;[10,16] methionyl-tRNA synthetase has been used to incorporate homoallylglycine in a methionine auxotroph *E. coli* strain;[17] and pyrrolysyl-tRNA synthetase has been used to charge the alkene-containing nonnatural amino acid (UAA) 6-*N*-allyloxycarbonyl-L-lysine onto its cognate tRNA.[18] In addition, we have genetically encoded *O*-allyltyrosine[19] and phenylselenocysteine (which can be converted to dehydroalanine by oxidative elimination)[20] in *E. coli* with engineered orthogonal tRNA/aminoacyl-tRNA synthetase (aaRS) pairs. Here we report the site-specific incorporation of several alkene-containing UAAs (Scheme 1) into proteins in eukaryotic cells with orthogonal tRNA/aaRS pairs evolved in *Saccharomyces cerevisiae*, and their subsequent application to protein modification.



Scheme 1. Structures of nonnatural amino acids described herein.

Our first attempt to genetically encode **1a** in *S. cerevisiae* made use of promiscuous aaRS mutants that incorporate amino acids with long aliphatic side chains, such as **1b**.[21] **1a** resembles **1b** in structure, so we directly tested these mutants for their ability to incorporate **1a**. The yeast strain MAV203:pGAD-Gal4(2TAG),[22] in which suppression of the amber codon (TAG) results in the expression of the reporter gene *ura3*, was transformed with plasmids encoding individual synthetases that aminoacylate **1b**. The resulting cells were then cultured on uracil-deficient (−Ura) agar plates in the presence of 1 mM **1a**. Since the gene *ura3* encodes an enzyme for uracil biosynthesis, TAG suppression is necessary for cell growth. Cells harbouring the plasmid encoding the aaRS Cap2X grew faster than cells containing other plasmids (Supporting Information, Figure S1-a);[21] therefore, Cap2X was further investigated. Additional experiments showed that Cap2X incorporated **1a** in response to the amber codon in human superoxide dismutase (hSOD-Trp33TAG) in SCY4 yeast (Figure S1-b). We then tested Cap2X in a yeast strain deficient in nonsense-mediated mRNA decay (Δ*upf1*) and expressing a significantly higher level of the amber suppressor tRNA.[23] This new system was reported to increase protein production by 300× relative to SCY4. However, the expression of GFP-Tyr39TAG in the presence of Cap2X and 2 mM **1a** in Δ*upf1* yeast resulted in heterogeneous GFP, which indicated that multiple cell-endogenous amino acids were incorporated at residue 39 of GFP (Figure S4-a).

In *Ec*LRS (the aminoacyl-tRNA synthetase from which the promiscuous synthetases were derived), the additional CP1 editing domain corrects mischarged amino acids,[24,25] leading to the incorporation of leucine with high fidelity. Therefore, we hypothesized that the fidelity of these synthe-

[*] Dr. H. W. Ai, Dr. W. Shen, Dr. E. Brustad,[+] Prof. Dr. P. G. Schultz
Department of Chemistry, The Scripps Research Institute
10550 N. Torrey Pines Rd., La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-9440
E-mail: schultz@scripps.edu

[+] Present address: Division of Chemistry and Chemical Engineering
California Institute of Technology, Pasadena, CA 91125 (USA)

[**] We thank Prof. Benjamin Davis, Justin Chalker, and Yuya Lin (University of Oxford) for helpful discussion; Prof. Karol Grela and Lukasz Gulajski (University of Warsaw), and Prof. Lei Wang (Salk Institute) for providing reagents; and Dr. Chang Liu and Emily Remba for manuscript preparation. This work is supported by the US Department of Energy, Division of Materials Sciences, under Award No. DE-FG03-00ER46051 and the Skaggs Institute for Chemical Biology.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200905590.

tases might be improved by mutations in the CP1 domain. To test this notion, we generated a new library in which the active site of the synthetase CP1 editing domain (residues Thr247, Thr248, and Thr252) was fully randomized. This library also contains some recombined sequences of the aforementioned promiscuous synthetases, since a mixture of plasmids encoding these aaRSs was used as the template for library construction and recombination can occur in the overlap polymerase chain reaction (PCR) (Figure S2). We then subjected the new aaRS library to four positive and three negative rounds of selection as previously described.[22] We observed convergence to two independent sequences, designated AK-1 and AK-2, respectively (Table S1; the sequences of synthetase AK-1 and AK-2 can be found under GenBank accession numbers GU059871 and GU059870). Both were tested for expression of the model protein hSOD using the plasmid hSOD-Trp33TAG. As expected, we observed high hSOD production in the presence of 1 mM **1a**, and little hSOD in the absence of **1a**. (Figure S1-b). We next attempted to express GFP-Tyr39TAG in Δ*upf1* yeast in the presence of 2 mM **1a**; approximately 5 mg protein could be purified from a 1 L culture using either the AK-1 or AK-2 synthetase. The resulting proteins were then subjected to ESI-MS characterization (Figure S4-b). Although a single protein peak was detected, the molecular weight did not match the calculated theoretical number (Table 1). We reasoned that **1a** is unstable in *S. cerevisiae*, and is oxidized to alliin (**1c**). In fact, when **1a** was replaced with 2 mM **1c** during protein expression, we detected GFP with the same molecular weight.

The **1a** analogues, **1d–1g**, are chemically resistant to oxidation, so we next tested whether AK-1 or AK-2 could incorporate **1d–1g**. Δ*upf1* yeast cells harbouring the GFP-Tyr39TAG gene, the corresponding tRNA gene, and synthetase AK-1 or AK-2 genes were cultured with 2 mM of each UAA. Cells cultured with **1d–1g** were all highly fluorescent. Suppression with AK-1 showed better contrast between the presence and absence of UAAs (Figure S3); thus, AK-1 was chosen for further characterization. Yields of alkene-containing GFPs produced with AK-1 were between 3.3 mg L^{-1} and 6.3 mg L^{-1} in the presence of 2 mM **1d–1g**, and molecular weights of purified GFPs agreed well with the corresponding calculated numbers (Table 1 and Figures S4–S5). The yield of protein production by AK-1 is comparable to other previously reported tRNA/aaRS pairs in yeast.[23] In addition, although there was noticeable basal amber suppression in the absence of UAAs, there was no apparent incorporation of endogenous amino acids detectable when 2 mM UAAs were present to compete with cell-endogenous amino acids as determined by ESI-MS (Figures S4–S5).

Next, we investigated the ability of proteins containing *O*-crotylserine (**1f**) to be selectively modified by olefin metathesis.

Table 1: Observed molecular weights (by ESI-MS) and calculated molecular weights of GFP and cpVenus (expressed with AK-1 in the presence of 2 mM UAAs).

	Amino acid	Obsd [Da]	Calcd[a] [Da]	Yield [mg L^{-1}]
GFP-Y39TAG	**1a**	28752	28734	5.0
	1d	28719	28716	3.3
	1e	28731	28730	4.8
	1f	28733	28732	6.3
	1g	28745	28746	5.5
cpVenus-2TAG	**1g** (before reaction)	29446	29451	0.3
	1g (after reaction)	29391	29395	

[a] GFP and cpVenus were N-acetylated in yeast. Listed are average molecular weights of acetylated mature peptide (GFP loses 20 Da upon chromophore maturation).[26]

Previous studies have suggested that heteroatoms (e.g., oxygen) at the allylic positions of alkenes can facilitate olefin metathesis reactions; and that the substituted alkylidene, which is the propagating catalytic species resulting from **1f**, is more stable in water than the methylidene resulting from terminal olefins.[10,27] Circularly permutated yellow fluorescent protein (cpVenus-2TAG) with **1f** incorporated at two spatially adjacent residues (Arg168 and Leu178 in Venus[28]) was expressed and purified in Δ*upf1* yeast. The resulting protein was subjected to olefin metathesis catalyzed by the 2nd generation Hoveyda–Grubbs catalyst in aqueous solution containing 30% *tert*-butanol (Supporting Information). The reaction mixture was directly analyzed by LC-ESI-MS, and near-complete conversion was observed in 5 h. The observed mass loss was consistent with the formation of a new

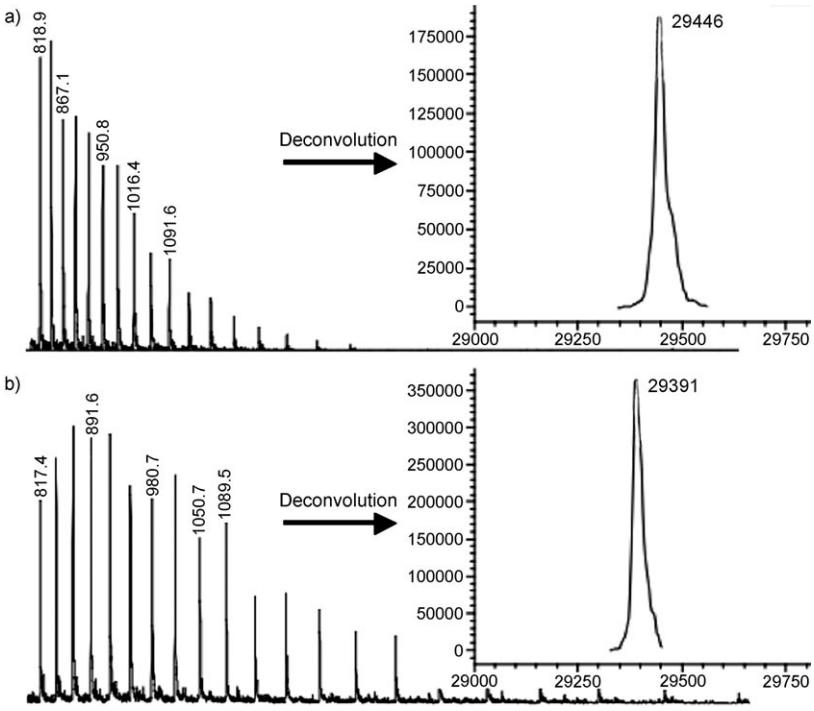


Figure 1. ESI-MS of cpVenus double *O*-crotylserine mutant, a) before and b) after olefin metathesis.

olefin bridge on the side chains of the two adjacent residues through an intramolecular metathesis reaction (Table 1 and Figure 1). This result shows that the genetically encoded alkenes are reactive in proteins and can be used for selective coupling chemistry.

In summary, we have developed new tRNA/aaRS pairs that make possible the genetic incorporation of a set of alkenes (**1c**–**1g**) into proteins in *S. cerevisiae*. These alkene moieties have flexible and relatively long side chains, and are useful bioorthogonal handles for various protein modification reactions. In one example, we demonstrated that protein side-chain bridges could be formed from intramolecular olefin metathesis of two alkene-containing residues. Similar reactions have been used to produce stable and protease-resistant peptides, such as the stapled BH3 helix for apoptosis activation,[12] and the stapled p53 peptide for the activation of the p53 tumour suppressor pathway.[13] In addition, the development of more active water-soluble catalysts may allow intermolecular reactions with olefin-containing biophysical probes, toxins, PEGs and the like;[29–31] the methodology described here may also allow the evolution of conformationally constrained peptides by yeast surface display. In conclusion, this work makes it possible to directly encode alkene functionality in eukaryotic cells, and thus, should greatly facilitate the further exploration of alkene functionality for protein modification both in vitro and in vivo.

Received: October 7, 2009
Published online: December 18, 2009

Keywords: nonnatural amino acids · olefin metathesis · protein modification

[1] C. P. Hackenberger, D. Schwarzer, *Angew. Chem.* **2008**, *120*, 10182; *Angew. Chem. Int. Ed.* **2008**, *47*, 10030.
[2] E. M. Sletten, C. R. Bertozzi, *Angew. Chem.* **2009**, *121*, 7108; *Angew. Chem. Int. Ed.* **2009**, *48*, 6974.
[3] A. Dirksen, T. M. Hackeng, P. E. Dawson, *Angew. Chem.* **2006**, *118*, 7743; *Angew. Chem. Int. Ed.* **2006**, *45*, 7581.
[4] K. L. Kiick, E. Saxon, D. A. Tirrell, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 19.
[5] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708; *Angew. Chem. Int. Ed.* **2002**, *41*, 2596.
[6] Y. Wang, W. Song, W. J. Hu, Q. Lin, *Angew. Chem.* **2009**, *121*, 5434; *Angew. Chem. Int. Ed.* **2009**, *48*, 5330.
[7] W. Song, Y. Wang, J. Qu, Q. Lin, *J. Am. Chem. Soc.* **2008**, *130*, 9654.
[8] M. L. Blackman, M. Royzen, J. M. Fox, *J. Am. Chem. Soc.* **2008**, *130*, 13518.
[9] G. J. Bernardes, J. M. Chalker, J. C. Errey, B. G. Davis, *J. Am. Chem. Soc.* **2008**, *130*, 5052.
[10] Y. A. Lin, J. M. Chalker, N. Floyd, G. J. Bernardes, B. G. Davis, *J. Am. Chem. Soc.* **2008**, *130*, 9642.
[11] J. Xiao, T. J. Tolbert, *Biomacromolecules* **2009**, *10*, 1939.
[12] L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, G. L. Verdine, S. J. Korsmeyer, *Science* **2004**, *305*, 1466.
[13] F. Bernal, A. F. Tyler, S. J. Korsmeyer, L. D. Walensky, G. L. Verdine, *J. Am. Chem. Soc.* **2007**, *129*, 2456.
[14] N. E. Shepherd, H. N. Hoang, G. Abbenante, D. P. Fairlie, *J. Am. Chem. Soc.* **2005**, *127*, 2974.
[15] A. Brik, *Adv. Synth. Catal.* **2008**, *350*, 1661.
[16] J. M. Chalker, Y. A. Lin, O. Boutureira, B. G. Davis, *Chem. Commun.* **2009**, 3714.
[17] J. C. M. van Hest, K. L. Kiick, D. A. Tirrell, *J. Am. Chem. Soc.* **2000**, *122*, 1282.
[18] T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto, S. Yokoyama, *Chem. Biol.* **2008**, *15*, 1187.
[19] Z. Zhang, L. Wang, A. Brock, P. G. Schultz, *Angew. Chem.* **2002**, *114*, 2964; *Angew. Chem. Int. Ed.* **2002**, *41*, 2840.
[20] J. Wang, S. M. Schiller, P. G. Schultz, *Angew. Chem. Int. Ed.* **2007**, *119*, 6973; *Angew. Chem. Int. Ed.* **2007**, *46*, 6849–6851.
[21] E. Brustad, M. L. Bushey, A. Brock, J. Chittuluru, P. G. Schultz, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6004.
[22] J. W. Chin, T. A. Cropp, J. C. Anderson, M. Mukherji, Z. Zhang, P. G. Schultz, *Science* **2003**, *301*, 964.
[23] Q. Wang, L. Wang, *J. Am. Chem. Soc.* **2008**, *130*, 6066.
[24] J. F. Chen, N. N. Guo, T. Li, E. D. Wang, Y. L. Wang, *Biochemistry* **2000**, *39*, 6726.
[25] Y. Liu, J. Liao, B. Zhu, E. D. Wang, J. Ding, *Biochem. J.* **2006**, *394*, 399.
[26] H.-w. Ai, R. E. Campbell, *Biotechnol. Focus* **2007**, *10*, 16.
[27] D. M. Lynn, B. Mohr, R. H. Grubbs, L. M. Henling, M. W. Day, *J. Am. Chem. Soc.* **2000**, *122*, 6601.
[28] T. Nagai, K. Ibata, E. S. Park, M. Kubota, K. Mikoshiba, A. Miyawaki, *Nat. Biotechnol.* **2002**, *20*, 87.
[29] Y. A. Lin, J. M. Chalker, B. G. Davis, *ChemBioChem* **2009**, *10*, 959.
[30] J. B. Binder, R. T. Raines, *Curr. Opin. Chem. Biol.* **2008**, *12*, 767.
[31] K. Kirshenbaum, P. S. Arora, *Nat. Chem. Biol.* **2008**, *4*, 527.

