tRNA

DOI: 10.1002/anie.200904035

Evolution of Amber Suppressor tRNAs for Efficient Bacterial Production of Proteins Containing Nonnatural Amino Acids**

*Jiantao Guo, Charles E. Melançon, III, Hyun Soo Lee, Dan Groff, and Peter G. Schultz**

tRNAs have evolved to act as highly efficient amino acid carriers and activators during each stage of protein synthesis. Each tRNA must be selectively charged by its cognate aminoacyl-tRNA synthetase (aaRS); the resulting aminoacyl-tRNA must be efficiently bound by elongation factor Tu (EF-Tu) for transport to the ribosome; after binding to the ribosomal A site, the aminoacyl-tRNA must function efficiently in translation as a substrate for peptidyl transferase; and finally the tRNA bearing the growing peptide chain must be translocated to the P site, undergo another acyl transfer reaction, and be released from the ribosome. We and others have used orthogonal tRNA/aaRS pairs for the site-specific incorporation of nearly 50 nonnatural amino acids in *E. coli*, *S. cerevisiae*, and mammalian cells in response to unique nonsense and frameshift codons.[1] An engineered *M. jannaschii* amber suppressor tyrosyl-tRNA/tRNA-synthetase ($MjtRNA^{Tyr}_{CUA}$/MjYRS) pair has been the most extensively used system for the evolution of aaRS variants that incorporate nonnatural amino acids in *E. coli*. Although nonnatural amino acids are typically incorporated into proteins with good efficiency and excellent fidelity, further system optimization resulting in increased protein yields is highly desirable.

Because $MjtRNA^{Tyr}_{CUA}$ is derived from an archaeal tRNA and therefore significantly differs in sequence from *E. coli* tRNAs, it may not function optimally with the *E. coli* translational machinery. Furthermore, in vitro binding studies have shown that while correctly acylated tRNAs bind EF-Tu with near uniform affinity, tRNAs bearing noncognate amino acids show a broad range of affinities for EF-Tu,[2] indicating that the tRNA body and the esterified amino acid make compensatory contributions to EF-Tu binding. A number of genetic, biochemical, and structural studies have implicated specific residues within the tRNA acceptor stem and T stem as being important for EF-Tu binding.[3] tRNA misacylation has also recently been shown to perturb binding to the ribosomal A site.[4] Thus, it is likely that tRNAs acylated with noncognate nonnatural amino acids result in a lower efficiency of protein synthesis owing to less than optimal interactions with EF-Tu and/or the ribosome. Therefore we have used in vitro evolution to optimize the sequence of $MjtRNA^{Tyr}_{CUA}$ with MjYRS and a panel of six evolved MjYRS variants incorporating nonnatural amino acids, and have identified several unique tRNA sequences that confer significantly improved expression of mutant proteins in *E. coli*. While most tRNAs identified show some improvement in protein yields for all aaRSs tested, the degree of improvement for each tRNA often varies depending on the identity of the esterified amino acid.

Examination of the X-ray crystal structure of the *T. aquaticus* EF-Tu·*E. coli* cysteinyl-$tRNA^{Cys}$·GDPNP ternary complex[6] reveals that residues 1–3, 50–54, 63–67, and 73–76, which reside in the acceptor stem and T stem of the tRNA, are in close proximity to EF-Tu (Figure 1). Interactions with EF-Tu almost exclusively involve the tRNA backbone, although contacts also exist between EF-Tu and the base of residue 63. Based on the notion that mutation of the corresponding positions in $MjtRNA^{Tyr}_{CUA}$ might modulate its binding affinity for EF-Tu, we created a tRNA library (Library I, theoretical diversity 1.05×10^6, >99% coverage) in which the five base pairs of the T stem (49–53, 61–65) were randomized. The tRNA library was subjected to a negative selection to remove tRNAs that are substrates for endogenous aaRSs using an amber mutant of the barnase gene as previously described;[6] positive selection of surviving clones to identify functional tRNA sequences was carried out in the presence of aaRSs specific for tyrosine (**1**) (MjYRS), L-3-(2)-naphthylalanine (**2**) (NapRS[7a]), *p*-azidophenylalanine (**3**) (pAzPheRS[7b]), *p*-iodophenylalanine (**4**) (pIPheRS[7c]), *p*-acetylphenylalanine (**5**) (pAcPheRS[7d]), and *p*-benzoylphenylalanine (**6**) (pBpaRS[7e]) (Scheme 1), chloramphenicol, and an amber mutant of chloramphenicol acetyltransferase (CAT) as previously described. The surviving tRNAs were then screened by cell fluorescence using an amber-suppressible GFPuv construct to identify the most active tRNAs (see the Supporting Information for experimental details). Nine unique tRNA sequences were identified with protein yields significantly improved compared with those of $MjtRNA^{Tyr}_{CUA}$ (see Table S3 in the Supporting Information). The G52–C62 and G53–C61 pairs, which are highly conserved among *E. coli* tRNAs, were invariant in all hits. The consensus sequence for the most active hits is $W^{49}NRGG^{53}$-$C^{61}CYNW^{65}$ (where W = A, T; N = A, T, G, C; R = A, G; and Y = C, T).

Next, we constructed a second tRNA library (Library II, theoretical diversity 1.7×10^7, >99% coverage) in which the consensus sequences at positions 49–53 and 61–65 from the best hits from Library I were retained and the remaining

[*] J. Guo,[+] C. E. Melançon, III,[+] H. S. Lee, D. Groff, Prof. P. G. Schultz
Department of Chemistry and
The Skaggs Institute for Chemical Biology
The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-9440
E-mail: schultz@scripps.edu

[+] These authors contributed equally to this work.

[**] This work is supported by NIH grants R01 GM062159 (P.G.S.) and F32 GM080067 (C.E.M.), and the Skaggs Institute for Chemical Biology.

Supporting information for this article (including materials and methods) is available on the WWW under http://dx.doi.org/10.1002/anie.200904035.

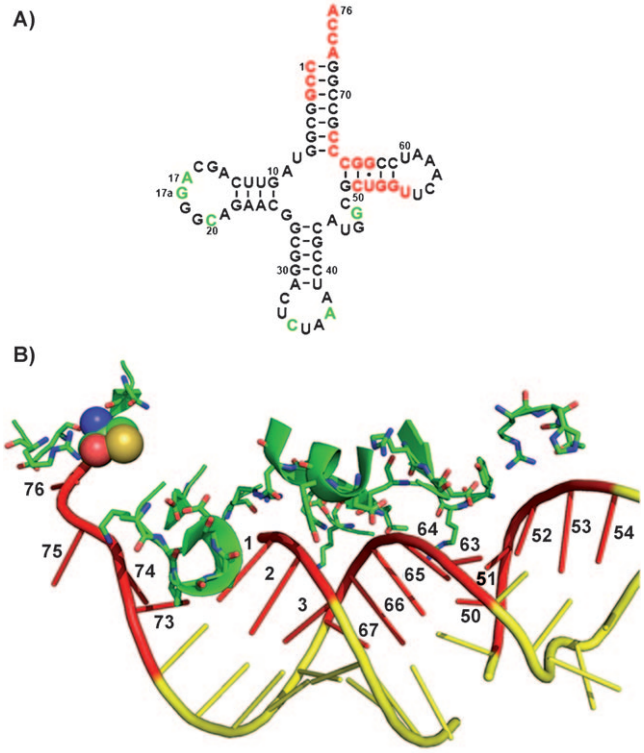
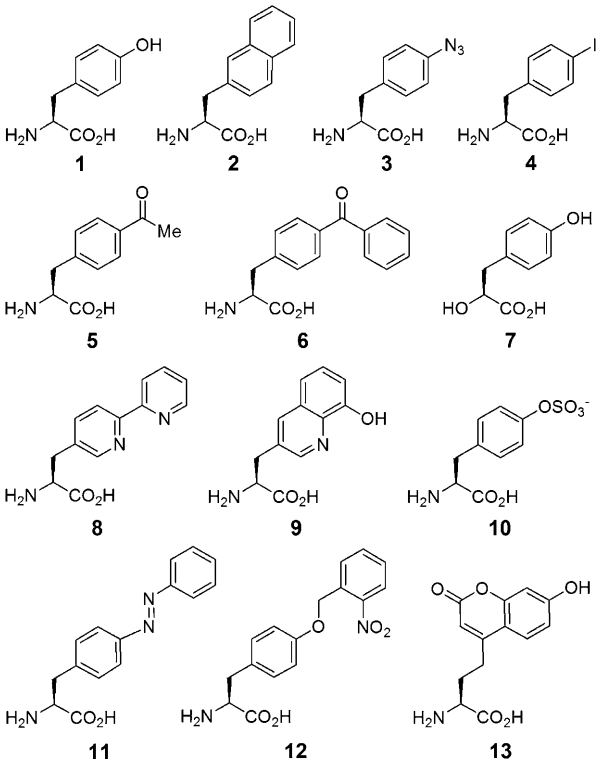


Figure 1. The EF-Tu/tRNA interface. a) Putative *M. jannaschii* $tRNA^{Tyr}_{CUA}$ nucleotides that interact with EF-Tu are shown in red. Previously mutated positions[6] are shown in green. b) A diagram derived from the X-ray crystal structure of *T. aquaticus* EF-Tu/*E. coli* cysteinyl-$tRNA^{Cys}$ [5] showing EF-Tu/tRNA interactions. The tRNA residues that interact with EF-Tu are numbered and shown in red. Interacting residues of EF-Tu are shown in green. Note that all interactions between tRNA and EF-Tu involve the tRNA backbone except in the case of residue 63.



Scheme 1. Structures of amino acids used in this study.

positions in contact with EF-Tu (2, 3, 6, 7, 66, 67, 70, and 71) were randomized. Positions 1 and 72–76 were not mutated because of their high degree of conservation and their importance in recognition by MjYRS. Library II was subjected to the same selection/screening procedure as was used for Library I. In addition to the six previously used aaRSs, the aaRS specific for *p*-hydroxy-L-phenyllactic acid (**7**) (PlaRS)[1c], was also used for positive selection. Fourteen unique tRNA sequences were identified that lead to significant improvements in protein yields relative to that of $MjtRNA^{Tyr}_{CUA}$ (see Table S4 in the Supporting Information). Of these, ten unique sequences were found to be among the top three tRNAs in terms of fluorescence intensity for at least one aaRS, and resulted in mutant protein yields 200–2200% of those obtained with $MjtRNA^{Tyr}_{CUA}$ (Table 1). Yield improvements were far greater with **6** (18 to 21-fold better than $MjtRNA^{Tyr}_{CUA}$) than with the other amino acids (100–350% better than $MjtRNA^{Tyr}_{CUA}$). Several of the best hits contained base deletions in the D loop which arose as artifacts of library construction.

These ten tRNAs were isolated, crossed against thirteen aaRSs (the seven used in this study and six additional aaRSs previously evolved for incorporation of bipyridylalanine (**8**) (BbyAla),[8a] hydroxyquinolinylalanine (**9**) (HQAla),[8b] sulfotyrosine (**10**) (SfY),[8c] *p*-azobenzylphenylalanine (**11**) (pABPhe),[8d] *o*-nitrobenzyltyrosine (**12**) (ONBY),[8e] and 7-hydroxycoumarinylethylglycine (**13**) (Cou)[8f] (Scheme 1) and assayed for GFP production in order to identify the tRNA sequence that confers the largest improvement in expressed protein yield for each amino acid. Improvements in yield of 175–320% were seen with the best tRNAs charged with compounds **1**–**5**, **8**, **10**, and **11**. However, significantly greater improvements were observed with **9** (405%), **7** (420%), **13** (790%), **12** (2070%), and **6** (2520%) (Figure 2). Interestingly, the compounds with the most significant improvements in yield have bulky side chains. Notable exceptions include **7**, which differs from tyrosine only by the replacement of the α-amino group with an α-hydroxy group; and **11**, which gave only a 255% improvement, yet has a bulky side chain. This data suggests that the structure of the nonnatural amino acid can affect the ability of the corresponding aminoacyl-tRNA to function productively in translation. This may be a result of interactions between the amino acid and the EF-Tu binding pocket. Consistent with this hypothesis, it was recently shown that mutation of the EF-Tu amino acid binding pocket resulted in more efficient binding of tRNAs bearing bulky nonnatural amino acids.[9] The 195% yield increase observed with the natural amino acid **1** represents the baseline improvement of $MjtRNA^{Tyr}_{CUA}$ arising from optimization of this archaeal tRNA for the *E. coli* translational machinery. Five tRNAs (Tyr1, pAzPhe1, Nap1, Nap3, and pIPhe1) consistently give the largest increases in yield. Of these, Nap1 is the best "general" hit (the tRNA that gives the best overall yield improvements with all amino acids). Yields of purified GFP expressed using Nap1 and a GFPN149TAG

Table 1: Evolved tRNA variants with improved amber suppression activity. The sequence of each tRNA at randomized positions and the expressed protein yield compared to MjtRNA$^{Tyr}_{CUA}$ are listed. The three most active tRNAs obtained from each selection are shown. S = G, C; W = A, T; B = G, T, C; V = A, G, C.

aaRS	tRNA	Position						Yield [% of MjtRNA$^{Tyr}_{CUA}$ activity]
		2,3	6,7	16–18	49–51	63–67	70,71	
–	MjtRNA$^{Tyr}_{CUA}$	CG	GG	CAG	GCT	GGCCC	CG	–
Nap	Nap1	CG	CG	CAG	ATG	CATCGT	CG	260 ± 7
	Nap2	CC	CC	–A–	TCG	CGAGG	GG	224 ± 16
	Nap3	CC	CT	CAG	ACG	CGTAG	GG	206 ± 9
pAzPhe	pAzPhe1	CC	CT	–AG	ACG	CGTAG	GG	365 ± 16
	pAzPhe2	same as Nap1						281 ± 17
	pAzPhe3	same as Nap2						219 ± 4
Tyr	Tyr1	same as Nap2						281 ± 11
	Tyr2	CG	CT	CAG	ATG	CATAG	CG	221 ± 8
	Tyr3	same as Nap3						216 ± 3
plPhe	plPhe1	CG	CG	CAG	TCG	CGACG	CG	345 ± 17
	plPhe2	same as Nap1						332 ± 30
	plPhe3	same as pAzPhe1						290 ± 16
pAcPhe	pAcPhe1	same as Nap2						274 ± 9
	pAcPhe2	same as Nap1						254 ± 5
	pAcPhe3	same as Nap3						211 ± 12
pBpa	pBpa1	same as pAzPhe1						2266 ± 12
	pBpa2	same as Nap1						2182 ± 26
	pBpa3	CG	CT	CAG	ACG	CGTAG	CG	1977 ± 8
Pla	Pla1	CC	CT	CAG	ATG	CATAG	GG	446 ± 15
	Pla2	CC	CA	CAG	TGG	CCATG	GG	400 ± 12
	Pla3	CC	CA	CAG	ATG	CATTG	GG	314 ± 14
–	consensus	CS	CN	CAG	WBG	CVWNG	SG	–

construct with **1–7** range from 2.7–16.3 mg L^{-1} (380–1175 % of yields obtained with MjtRNA$^{Tyr}_{CUA}$) in rich media (see Figure S9 in the Supporting Information). The fidelity of non-natural amino acid incorporation with Nap1 and **1–7** was determined by LC-ESI-MS of purified GFP proteins produced using evolved tRNAs and is comparable to the high fidelity obtained with MjtRNA$^{Tyr}_{CUA}$ (see Table S2 and Figures S11–S17 in the Supporting Information).

Growth rates were determined for *E. coli* TOP10 cells expressing the evolved tRNAs (with GFP N149TAG reporter, aaRS, IPTG, and **1–7**) (see Figure S10 and Table S1 in the Supporting Information). In all cell lines except those expressing PlaRS, the exponential growth rates for cells expressing the evolved tRNAs were significantly faster [$\mu = (0.47 \pm 0.05)\ h^{-1}$] than for those expressing MjtRNA$^{Tyr}_{CUA}$ [$\mu = (0.30 \pm 0.05)\ h^{-1}$], indicating that the evolved tRNAs are in general less toxic than MjtRNA$^{Tyr}_{CUA}$. Because all tRNAs were subjected to negative selection to remove those that are substrates for endogenous aaRSs, it is unlikely that altered toxicity results from differences in orthogonality of the tRNAs. The decreased toxicity of the evolved tRNAs could result from decreased readthrough of natural stop codons, or altered interactions of the tRNA with EF-Tu and/or the ribosome which result in enhanced translation of heterologously expressed amber mutant genes. Several examples of suppressor tRNA-dependent cellular toxicity, in which toxicity was suggested to be the result of readthrough of endogenous stop codons, have been reported.[10] The fact that cells expressing all evolved tRNAs charged with **1–6** have nearly the same growth rates, yet display a broad range of protein yields, suggests that other factors in addition to ameliorated toxicity contribute to the improvements in protein yield.

A recent study on the EF-Tu binding affinities of a collection of acceptor stem and T-stem mutations allows prediction of the strength of the EF-Tu interaction based on sequence.[3a] This model predicts that our evolved tRNAs bind EF-Tu 0.3–0.9 kcal mol^{-1} more tightly than MjtRNA$^{Tyr}_{CUA}$. The 51–63 base pair, which was changed from a U–G pair in

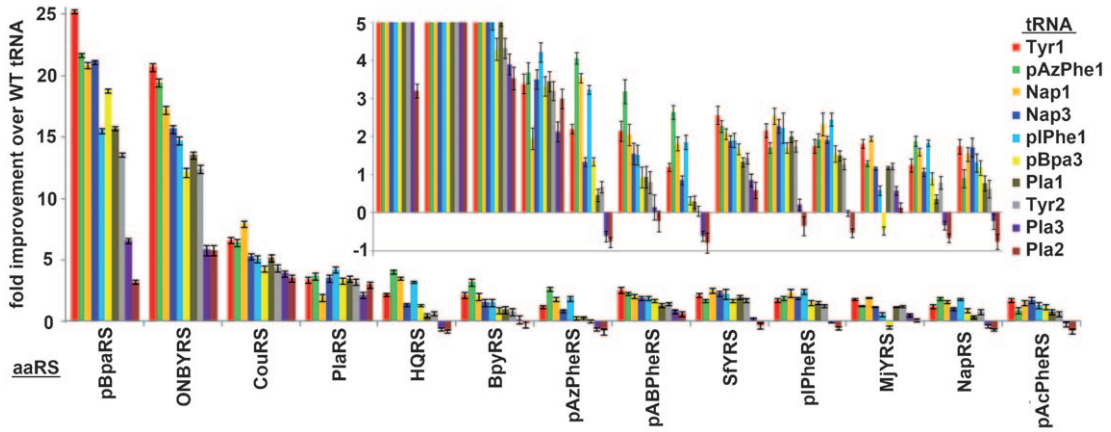


Figure 2. Activity of the ten best evolved tRNAs with each aaRS shown as fold improvement over MjtRNA$^{Tyr}_{CUA}$. tRNAs and aaRSs are arranged from right to left in order of increasing yield improvement. Inset shows the region of the main graph between −1- and 5-fold improvement.

$MjtRNA^{Tyr}_{CUA}$ to a G–C pair in all evolved tRNAs, is predicted to strengthen the EF-Tu binding affinity by 1.0 kcal mol^{-1}. The gain in function observed for the evolved tRNAs may be due in large part to this mutation, although convergence to non-wild-type sequences at other positions suggests that these mutations may act to fine tune tRNA affinity for EF-Tu. It is possible that the predicted increase in binding energy between EF-Tu and the evolved tRNAs compensates for weaker binding of nonnatural amino acids in the EF-Tu binding pocket, and that this compensation is the reason for the large improvements in yield observed with some amino acids. Detailed analysis of the interaction of the acylated tRNAs with EF-Tu will provide additional insights into the mechanistic basis for the improved activities of these tRNAs.

In conclusion, we have utilized an in vitro evolution approach to identify $MjtRNA^{Tyr}_{CUA}$ variants with significantly enhanced activity for the incorporation of nonnatural amino acids into proteins, especially the photocrosslinking amino acid **6**, which is widely used to map biomolecular interactions. We also found that the degree of yield improvement is in some cases dependent on the nonnatural amino acid. These tRNAs will facilitate the creation of an optimized and standardized system for the genetic incorporation of nonnatural amino acids into proteins in *E. coli*.[11]

Received: July 22, 2009
Published online: October 23, 2009

Keywords: amber suppression · evolution · nonnatural amino acids · protein modifications · tRNA

[1] Selected references include: a) J. Xie, P. G. Schultz, *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 775; b) E. A. Lemke, D. Summerer, B. H. Geierstanger, S. M. Brittain, P. G. Schultz, *Nat. Chem. Biol.* **2007**, *3*, 769; c) J. Guo, J. Wang, J. C. Anderson, P. G. Schultz, *Angew. Chem.* **2008**, *120*, 734; *Angew. Chem. Int. Ed.* **2008**, *47*, 722; d) E. Brustad, M. L. Bushey, J. W. Lee, D. Groff, W. Liu, P. G. Schultz, *Angew. Chem.* **2008**, *120*, 8344; *Angew. Chem. Int. Ed.* **2008**, *47*, 8220; e) H. Neumann, J. L. Hazen, J. Weinstein, R. A. Mehl, J. W. Chin, *J. Am. Chem. Soc.* **2008**, *130*, 4028; f) H. Neumann, S. Y. Peak-Chew, J. W. Chin, *Nat. Chem. Biol.* **2008**, *4*, 232; g) T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto, S. Yokoyama, *Chem. Biol.* **2008**, *15*, 1187; h) F. B. Peters, A. Brock, J. Wang, P. G. Schultz, *Chem. Biol.* **2009**, *16*, 148; i) W.-T. Li, A. Mahapatra, D. G. Longstaff, J. Bechtel, G. Zhao, P. T. Kang, M. K. Chan, J. A. Krzycki, *J. Mol. Biol.* **2009**, *385*, 1156; j) T. Kobayashi, T. Yanagisawa, K. Sakamoto, S. Yokoyama, *J. Mol. Biol.* **2009**, *385*, 1352.

[2] a) F. J. LaRiviere, A. D. Wolfson, O. C. Uhlenbeck, *Science* **2001**, *294*, 165; b) H. Asahara, O. C. Uhlenbeck, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 3499; c) T. Dale, L. E. Sanderson, O. C. Uhlenbeck, *Biochemistry* **2004**, *43*, 6159.

[3] a) J. M. Schrader, S. J. Chapman, O. C. Uhlenbeck, *J. Mol. Biol.* **2009**, *386*, 1255; b) J. Eargle, A. A. Black, A. Sethi, L. G. Trabuco, Z. Luthey-Schulten, *J. Mol. Biol.* **2008**, *377*, 1382; c) C. Forster, K. Chakraburtty, M. Sprinzl, *Nucleic Acids Res.* **1993**, *21*, 5679–5683; d) J. Rudinger, R. Hillenbrandt, M. Sprinzl, R. Giege, *EMBO J.* **1996**, *15*, 650–657.

[4] T. Dale, R. P. Fahlman, M. Olejniczak, O. C. Uhlenbeck, *Nucleic Acids Res.* **2009**, *37*, 1202.

[5] P. Nissen, S. Thirup, M. Kjeldgaard, J. Nyborg, *Structure* **1999**, *7*, 143.

[6] L. Wang, P. G. Schultz, *Chem. Biol.* **2001**, *8*, 883.

[7] a) L. Wang, A. Brock, P. G. Schultz, *J. Am. Chem. Soc.* **2002**, *124*, 1836; b) J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, *J. Am. Chem. Soc.* **2002**, *124*, 9026; c) J. Xie, L. Wang, N. Wu, A. Brock, G. Spraggon, P. G. Schultz, *Nat. Biotechnol.* **2004**, *22*, 1297; d) L. Wang, Z. Zhang, A. Brock, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 56; e) J. W. Chin, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11020.

[8] a) J. Xie, W. Liu, P. G. Schultz, *Angew. Chem.* **2007**, *119*, 9399; *Angew. Chem. Int. Ed.* **2007**, *46*, 9239; b) H. S. Lee, G. Spraggon, P. G. Schultz, F. Wang, *J. Am. Chem. Soc.* **2009**, *131*, 2481; c) C. Liu, P. G. Schultz, *Nat. Biotechnol.* **2006**, *24*, 1436; d) M. Bose, D. Groff, J. Xie, E. Brustad, P. G. Schultz, *J. Am. Chem. Soc.* **2006**, *128*, 388; e) A. Deiters, D. Groff, Y. Ryu, J. Xie, P. G. Schultz, *Angew. Chem.* **2006**, *118*, 2794; *Angew. Chem. Int. Ed.* **2006**, *45*, 2728; f) J. Wang, J. Xie, P. G. Schultz, *J. Am. Chem. Soc.* **2006**, *128*, 8738.

[9] Y. Doi, T. Ohtsuki, Y. Shimizu, T. Ueda, M. Sisido, *J. Am. Chem. Soc.* **2007**, *129*, 14458–14462.

[10] a) R. E. Doerig, B. Suter, M. Gray, E. Kubli, *EMBO J.* **1988**, *7*, 2579–2584; b) C. E. Leuker, J. F. Ernst, *Mol. Gen. Genet.* **1994**, *245*, 212–217; c) V. T. C. Carneiro, G. Pelletier, I. Small, *Plant Mol. Biol.* **1993**, *22*, 681–690.

[11] T. S. Young, I. Ahmad, P. G. Schultz, *J. Mol. Biol.*, in press.

