

# Reprogramming the Genetic Code: From Triplet to Quadruplet Codes

Kaihang Wang, Wolfgang H. Schmied, and Jason W. Chin\*

amino acids · codons · gene expression ·  
protein design · ribosome

*The genetic code of cells is near-universally triplet, and since many ribosomal mutations are lethal, changing the cellular ribosome to read nontriplet codes is challenging. Herein we review work on the incorporation of unnatural amino acids into proteins in response to quadruplet codons, and the creation of an orthogonal translation system in the cell that uses an evolved orthogonal ribosome to efficiently direct the incorporation of unnatural amino acids in response to quadruplet codons. Using this system multiple distinct unnatural amino acids have been incorporated and used to genetically program emergent properties into recombinant proteins. Extension of approaches to incorporate multiple unnatural amino acids may allow the combinatorial biosynthesis of materials and therapeutics, and drive investigations into whether life with additional genetically encoded polymers can evolve to perform functions that natural biological systems cannot.*

response to quadruplet codons using extended anticodon tRNAs. We describe the creation of an orthogonal translation pathway that allows evolution of an orthogonal ribosome in the laboratory, and the evolution of the orthogonal ribosome to efficiently decode quadruplet codons using extended anticodon tRNAs. We describe the incorporation of multiple unnatural amino acids into recombinant proteins through an orthogonal translation pathway assembled from the modular combination of evolved orthogonal ribosomes, orthogonal mRNA, orthogonal aminoacyl-tRNA synthetases, and tRNAs, and the use of this pathway to direct emergent properties in recombinant proteins. Finally we describe experiments that address the limitations of orthogonal synthetase/tRNA pair discovery, and suggest future directions.

## 1. Introduction

The genetic code describes the correspondence between triplet codons, composed of a combination of the four bases (A,C,G,T), and the 20 common amino acids.<sup>[1–6]</sup> The code is enforced by the translational machinery of cells, which decodes the sequence of non-overlapping triplet codons in an mRNA to produce proteins of defined length, amino acid composition, and sequence.

The translational machinery has been modified for the site-specific incorporation of unnatural amino acids into proteins in vitro,<sup>[7–10]</sup> in bacterial and eukaryotic cells,<sup>[11,12]</sup> and in an animal.<sup>[13]</sup> It has also been possible to incorporate multiple unnatural amino acids, using pre-aminoacylated tRNAs, in in vitro translation reactions.<sup>[14–16]</sup>

In this Minireview we describe alterations in the translational machinery that allow quadruplet codons to be decoded and used as unique insertion signals for encoding the incorporation of unnatural amino acids into proteins. We describe the incorporation of unnatural amino acids in

response to quadruplet codons using extended anticodon tRNAs. We describe the creation of an orthogonal translation pathway that allows evolution of an orthogonal ribosome in the laboratory, and the evolution of the orthogonal ribosome to efficiently decode quadruplet codons using extended anticodon tRNAs. We describe the incorporation of multiple unnatural amino acids into recombinant proteins through an orthogonal translation pathway assembled from the modular combination of evolved orthogonal ribosomes, orthogonal mRNA, orthogonal aminoacyl-tRNA synthetases, and tRNAs, and the use of this pathway to direct emergent properties in recombinant proteins. Finally we describe experiments that address the limitations of orthogonal synthetase/tRNA pair discovery, and suggest future directions.

## 2. Protein Translation

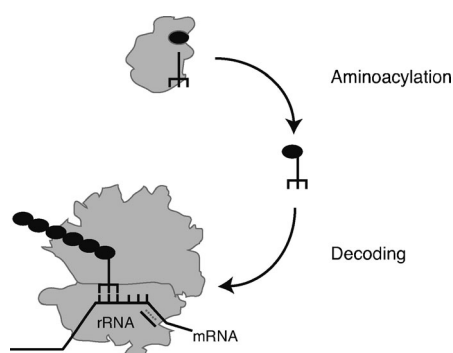
Protein translation is directed on the ribosome: a 2.5 MDa molecular machine composed of three large ribosomal (r) RNAs and more than 50 proteins partitioned between a large and a small subunit. Translation has four steps: initiation, elongation, termination, and recycling.<sup>[17,18]</sup>

In *E. coli*, protein translation is initiated when a complex between initiation factors IF1/3 and the small subunit of the ribosome bind to cellular mRNAs.<sup>[19–21]</sup> This interaction is rate determining for protein synthesis, and is mediated by base pairing between the 16S rRNA in the small subunit and the Shine–Dalgarno sequence (AGGAGG), which is 7 to 12 bases

[\*] Dr. K. Wang, Mag. W. H. Schmied, Dr. J. W. Chin  
Medical Research Council Laboratory of Molecular Biology  
Hills Rd, Cambridge, CB20QH (UK)  
E-mail: chin@mrc-lmb.cam.ac.uk

upstream of an AUG initiation codon on the mRNA.<sup>[22,23]</sup> This interaction positions the AUG codon in the P site of the ribosome, which defines, in most genes, the sequence of non-overlapping triplet codons that will be read. The complex between the IF1/3/ribosomal small subunit and mRNA recruits the aminoacylated initiator tRNA and the large subunit of the ribosome in sequential steps, leading to a ribosome poised for translational elongation.

The fidelity of protein translation in the elongation phase of protein synthesis is primarily defined by two processes: 1) the aminoacylation of a tRNA with the correct amino acid by an aminoacyl-tRNA synthetase<sup>[24,25]</sup> and the delivery of the aminoacylated tRNA to the ribosome,<sup>[26]</sup> and 2) the ribosomal decoding and subsequent translocation of tRNA/mRNA complexes with cognate codon–anticodon interactions,<sup>[27–30]</sup> to direct the correct amino acid into the nascent polypeptide in response to a specific codon (Figure 1).



**Figure 1.** Aminoacylation and decoding are key steps in the fidelity of translation. Amino acids (black ovals), tRNAs (black tridents). mRNA = messenger RNA, rRNA = ribosomal RNA.

In the elongation phase of translation an aminoacylated tRNA is delivered into the decoding center, within the A site of the ribosome, by EF-Tu/GTP. The tRNA (anticodon)/mRNA (codon) helix formed in the A site is actively monitored by the ribosome to exclude noncognate interactions and maintain the fidelity of the genetic code.<sup>[27]</sup> Several nucleotides (1492, 1493, 530) in the 16S ribosomal RNA form key interactions with the first two nucleotides of the codon–anticodon interaction and ensure strict Watson–Crick base pairing by excluding noncognate and wobble pairs. In contrast the ribosome packs less tightly against the third base of the codon–anticodon interaction, thus providing a molecular explanation for the fact that wobble pairing is excluded at the first two positions but allowed at the third position of the codon.<sup>[27]</sup> Cognate tRNA/mRNA interactions facilitate GTP hydrolysis in EF-Tu and the release of the aminoacylated tRNA acceptor stem into the A site of the peptidyl transferase center in the large subunit of the ribosome for peptide bond formation. Following peptide bond formation the peptidyl tRNA/mRNA complex is translocated by three bases by EF-G/GTP, thus positioning a new codon in the A site of the ribosome for the next round of elongation.<sup>[17,18,31]</sup>

Protein translation terminates at the end of the elongation phase when one of the three stop codons is reached.



**Kaihang Wang** was an undergraduate at Peking University and University College London. He obtained his PhD from Cambridge University under the supervision of Prof. Jason Chin at the Medical Research Council Laboratory of Molecular Biology (MRC-LMB). He was a Junior Research Fellow at Trinity College, Cambridge and is currently a Career Development Fellow at the MRC-LMB.



**Wolfgang Schmied** studied molecular biology at the University of Vienna. He received his diploma in 2010, for work on the type three secretion system under the supervision of Dr. Thomas Marlovits at IMBA, Vienna. He is currently a graduate student with Prof. Jason Chin and works on engineering the protein translation machinery.



**Jason Chin** was an undergraduate at Oxford, obtained his PhD with Prof. Alanna Schepartz at Yale, and was a postdoctoral fellow with Prof. Peter Schultz at Scripps. Currently, he is Head of the Centre for Chemical & Synthetic Biology at the Medical Research Council Laboratory of Molecular Biology, where he is also a Programme Leader. He has a joint appointment at the University of Cambridge Department of Chemistry and is a fellow in Natural Sciences at Trinity College, Cambridge.

Termination of protein synthesis is mediated by protein release factors that bind to the ribosome when a stop codon (UAA, UAG, UGA) on the mRNA is presented at the A site of the ribosome. The bound release factor directs the hydrolysis of the nascent polypeptide chain from the ribosome P-site-bound tRNA, and the ribosome is subsequently recycled in preparation for additional rounds of translation. Detailed accounts of the steps in protein translation have been published elsewhere.<sup>[17,18,28]</sup>

### 3. Challenges and Opportunities in Reprogramming Protein Translation

While cellular translation on the ribosome normally makes proteins composed of the natural 20 amino acids, it provides the ultimate paradigm for the encoded, and evolvable, synthesis of proteins containing additional amino acids, and ultimately the synthesis and evolution of completely unnatural polymers. Since the ribosome uses tRNA adapter molecules it makes polymers that are chemically independent of the template, this independence should allow distinct polymers to be made without interfering with their encoding. Since the ribosome uses a single set of active sites for bond

formation, and couples these to a translocation activity, it can make very long polymers with very high fidelity.

There are several major challenges in genetically encoding multiple unnatural amino acids into proteins in living cells. First, unique new codons are required to encode the incorporation of unnatural amino acids at specific sites in proteins. Since the 64 triplet codons are already used in the genome of most organisms for encoding natural proteins, additional codons (such as quadruplet codons) might be used to encode the incorporation of unnatural amino acids. Second, new aminoacyl-tRNA synthetase/tRNA pairs that are orthogonal to the aminoacyl-tRNA synthetase/tRNA pairs in the host organism and uniquely direct the incorporation of an unnatural amino acid in response to a unique codon, are required. Finally the scope of cellular protein translation is limited to  $\alpha$ -L amino acids and their close analogues, and alteration of the ribosome and potentially other components of the translational machinery are required to increase the chemical scope of translation.

## 4. Noncanonical Decoding of Cellular mRNAs

### 4.1. Frameshifting and Hopping

While the triplet genetic code is near universally conserved there is a growing body of work on polypeptide sequences that do not directly correlate with the product that would be expected from the iterative decoding of non-overlapping triplets between the start codon and the termination codon in an mRNA.<sup>[32–34]</sup> Such mRNAs may contain additional bases or have bases deleted with respect to a hypothetical “canonical gene” that would be expected to encode the protein they produce. Sequences that contain an additional base in the canonical gene are said to contain a +1 frameshift and create a “quadruplet codon” in the reading frame. Sequences containing a deletion with respect to the canonical gene are said to contain a –1 frameshift. Longer insertions of up to 50 bases or more with respect to the canonical gene can be bypassed in translation by a process described as hopping.<sup>[35–38]</sup>

Frameshifting (the process of decoding a quadruplet codon to restore the reading frame found in the canonical gene) and hopping using the common set of tRNAs usually reach appreciable levels only in the presence of elements in the mRNA that cause translational pausing.<sup>[33]</sup> Elements known to cause pausing and facilitate noncanonical decoding include unusual mRNA structures,<sup>[39–41]</sup> upstream Shine–Dalgarno-like sequences,<sup>[35,42]</sup> or upstream underused codons for which the cognate tRNAs are in low abundance.<sup>[43]</sup> Frameshifting and hopping can be used to regulate the amino acid sequence of a protein produced,<sup>[44]</sup> to control the ratio of two proteins produced in alternate reading frames, or to control the abundance of a protein through feedback regulation.<sup>[45,46]</sup>

### 4.2. Decoding Quadruplet Codons using Extended Anticodon tRNAs

Known suppressors of quadruplet codons include natural tRNAs bearing extended anticodon loops. Prolyl and glycyl tRNAs bearing extended anticodons, as a result of insertions in the anticodon, were first identified as suppressors of +1 frameshifts in *Salmonella typhimurium*.<sup>[47–49]</sup> Subsequently, extended anticodon tRNAs that suppress quadruplet codons have been identified in *E. coli* and *S. cerevisiae*.<sup>[50–52]</sup> Moore and co-workers examined a limited range of extended anticodon tRNAs derived from *E. coli* tRNA<sup>Leu</sup><sub>CAA</sub> and discovered that *E. coli* tRNA<sup>Leu</sup><sub>UCUA</sub> efficiently decodes UAGA codons.<sup>[50]</sup> Magliery et al. set out to discover a set of functional extended anticodon/quadruplet codon pairs in *E. coli*.<sup>[51,52]</sup> They started from an *E. coli* tRNA<sub>2<sup>Ser</sup></sub> and mutated the anticodon loop, which normally contains seven nucleotides, to create a library of all possible eight and nine nucleotide anticodon loops. They combined the extended anticodon tRNA libraries with a library of all possible quadruplet codons at a single site in an antibiotic resistance gene in *E. coli*. Finally they selected extended anticodon tRNA library members that mediate frameshift suppression and restore the antibiotic resistance. Using this approach they discovered extended anticodon tRNA variants of tRNA<sub>2<sup>Ser</sup></sub> that decode AGGA, UAGA, CCCU, and CUAG codons. These codons correspond to some of the least used codons in *E. coli* with one additional base. In mammalian cells variants of human tRNA<sup>Tyr</sup> bearing extended NCUA anticodons were reported to decode extended UAGN codons in green fluorescent protein (GFP), thus leading to fluorescence that was characterized by fluorescence microscopy.<sup>[53]</sup> Similarly a tRNA<sup>Tyr</sup> variant bearing an AUAG anticodon was reported to decode a CUAU codon.

### 4.3. Encoding Unnatural Amino Acids at Quadruplet Codons

There have been numerous efforts to use quadruplet codons to encode unnatural amino acids. These approaches, which are based on prior work with amber suppressors,<sup>[8,54]</sup> have primarily used pre-aminoacylated, extended anticodon tRNAs in in vitro translation reactions.<sup>[55–57]</sup> These approaches have been extended through microinjection of pre-aminoacylated, extended anticodon tRNAs in *Xenopus* oocytes.<sup>[58–60]</sup> In these approaches, aminoacylation of the tRNA is performed prior to its addition to the translation mix and, to avoid the deacylated tRNA that accumulates over the course of the translation reaction from being re-acylated with natural amino acids by aminoacyl-tRNA synthetases in the extract or oocyte, the tRNA must be orthogonal.

Hohsaka, Sisido, and co-workers used variants of yeast tRNA<sup>Phe</sup> with altered anticodons, tRNAs<sub>NCCU</sub> (tRNA<sub>ACCUC</sub>, tRNA<sub>UCCUC</sub>, tRNA<sub>CCCUC</sub>, and tRNA<sub>GCCUC</sub>), to incorporate unnatural amino acids into proteins produced in an *E. coli* S30 extract. They replaced the Tyr83 UAU codon in the Streptavidin mRNA with AGGN and demonstrated unnatural amino acid incorporation in response to quadruplet codons in vitro.<sup>[57]</sup> Several unnatural amino acids have been incorporated

in response to distinct quadruplet codons, including AGGU, CGGU, CGCU, CGAU, CCCU, CUCU, CUAU, GGGU, and their fourth base variants.<sup>[61–63]</sup> In vitro translation using quadruplet codons has been extended from *E. coli* extracts to insect<sup>[64,65]</sup> and rabbit reticulocyte cell extracts,<sup>[66]</sup> and model in vitro selection experiments using four-base codons as part of an mRNA display based strategy have been performed.<sup>[67]</sup> Two<sup>[68]</sup> or three<sup>[69]</sup> distinct unnatural amino acids have been incorporated using distinct quadruplet codons and their cognate pre-aminoacylated tRNAs, and this approach has been used to site-specifically incorporate two BODIPY fluorophores into calmodulin at distinct sites to make distance measurements using fluorescence resonance energy transfer (FRET).<sup>[70]</sup>

Following work on microinjection of pre-aminoacylated amber suppressor tRNAs in *Xenopus* oocytes,<sup>[58]</sup> investigators have explored the use of quadruplet codons and extended anticodon tRNAs in *Xenopus* oocytes.<sup>[71]</sup> Dougherty and co-workers injected pre-aminoacylated tRNA<sub>CCCG</sub> and tRNA<sub>ACCC</sub> (derived from yeast tRNA<sup>Phe</sup>) into *Xenopus* oocytes together with muscle nicotinic acetylcholine receptor (nAChR) mRNA having the quadruplet codon CGGG and GGGU. Using this approach they demonstrated the site-specific quadruplet incorporation of the two distinct unnatural amino acids through electrophysiological recordings on the receptor.<sup>[72]</sup> Schultz and co-workers demonstrated the use of a quadruplet codon to encode homoglutamine by using a variant *Pyrococcus horikoshii* lysyl-tRNA synthetase/tRNA pair in *E. coli*.<sup>[73]</sup> This variant was used in combination with a variant *Methanococcus jannaschii* tyrosyl-tRNA synthetase (MjTyrRS)/tRNA<sub>CUA</sub> pair that directs the incorporation of O-methyl-L-tyrosine in response to the amber codon.<sup>[74]</sup>

#### 4.4. Mechanisms of Quadruplet Decoding

While the decoding of quadruplet codons by extended anticodon tRNAs is well established, and in most cases appears to require a potential Watson–Crick or wobble base pair between the fourth base in the codon and the anticodon, the molecular mechanism by which quadruplet decoding occurs is not well understood. Indeed it may be that different mechanisms operate in different mRNA contexts and with different tRNAs.

While the conceptually simplest mechanism of quadruplet decoding by tRNAs with extended anticodons arguably involves the binding of all four bases of a quadruplet codon to the anticodon in the A site of the ribosome and quadruplet translocation, the overall process of quadruplet decoding could be theoretically explained by other combinations of noncanonical codon/anticodon interactions, noncanonical translocation step size, and tRNA slipping.<sup>[34]</sup> The incoming aminoacyl tRNA may bind to a codon larger or smaller than the normal triplet, and/or bind out of frame. The step size for translocation may differ from the normal three bases, and the translocated peptidyl tRNA may slip, or re-pair, on the mRNA out of frame.<sup>[75]</sup> Two limiting models that are consistent with the experimental data for particular tRNA codon pairs have been described: the yardstick model

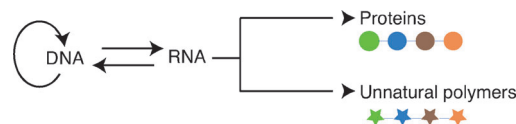
proposes triplet or quadruplet interactions between the codon and anticodon in the A site of the ribosome with subsequent quadruplet translocation,<sup>[76–78]</sup> while the slippery model proposes a triplet interaction in the A site and triplet translocation followed by a slip of the mRNA by one base.<sup>[79]</sup>

## 5. Incorporating Multiple Unnatural Amino Acids into Proteins through an Orthogonal Translation Pathway

### 5.1. Orthogonal Translation

While a host of quadruplet codons can be read on the natural ribosome using extended anticodon tRNAs, the decoding of these tRNAs is inefficient. Extended anticodons may be poorly accommodated in the ribosome and efforts to increase the efficiency of quadruplet decoding in natural translation by extended anticodon tRNAs may lead to mis-synthesis of the proteome, toxicity, and cell death. A solution to the efficient decoding of quadruplet codons comes from our work on the creation and evolution of an orthogonal translation pathway in the cell.<sup>[80–86]</sup>

We realized that since the genetic code is a correspondence between amino acids and codons, which is set by the translational machinery, we could create a parallel genetic code (Figure 2) in the cell in two steps. In the first step we created an orthogonal ribosome that is selectively directed to a new orthogonal mRNA that is not read by endogenous ribosomes.<sup>[80]</sup> This creates a ribosome that, unlike the natural ribosome (which is required to synthesize the proteome and keep the cell alive and is therefore refractory to alteration of its sequence and structure), can be altered by synthetic evolution in the laboratory. The new orthogonal ribosome provides a platform for engineering the functional centers of the ribosome, that have been revealed by years of biochemistry as well as recent stunning advances in structural biology, but which it has not been possible to alter to perform new functions. In a second step we altered the orthogonal ribosome to efficiently decode tRNAs, which are not efficiently decoded on cellular mRNA, on the orthogonal mRNA.<sup>[84,86]</sup> This is achieved by creating a “bump” on the tRNA (using an extended anticodon tRNA) and a corresponding “hole” in the orthogonal ribosome to accommodate the expanded tRNA. This approach allows the efficiency of the quadruplet decoding to be selectively enhanced on the orthogonal mRNA, without affecting the decoding of cellular mRNAs by natural ribosomes, and therefore enhances the



**Figure 2.** Orthogonal translation. Natural translation makes proteins composed of natural amino acids (circles). An orthogonal translation pathway can make proteins containing unnatural amino acids (stars), and ultimately may be used to make completely unnatural polymers.



efficiency of quadruplet decoding without creating toxic misreading of the proteome.

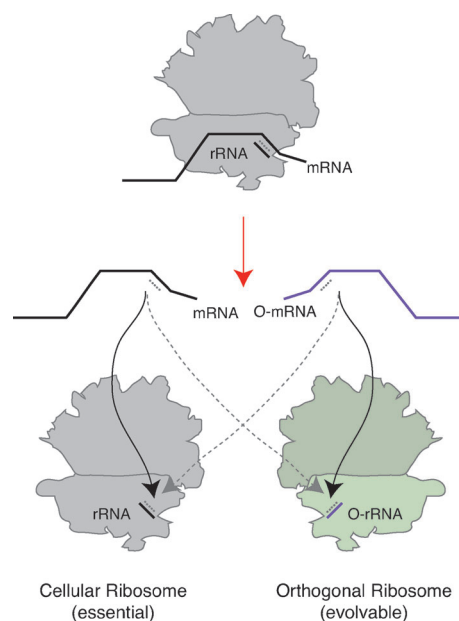
### 5.2. Creating Orthogonal Ribosome/mRNA Pairs

To create an orthogonal ribosome/mRNA pair in *E. coli* that operates in parallel with, but independent of, the natural ribosome we took advantage of the fact that the *E. coli* ribosome selects its cognate mRNA in the rate-determining, initiation stage of protein translation through a Shine–Dalgarno sequence 7 to 12 bases 5' of the AUG initiation codon.<sup>[80,81]</sup> Genomic Shine–Dalgarno sequences show substantial sequence variation, but the consensus sequence (AGGAGG) is complementary to the 3'-end of the 16S rRNA in the small subunit of the ribosome (CCUCCU). We created a large library of alternative Shine–Dalgarno sequences that cover all combinations of nucleotides from 7 to 12 bases upstream to the AUG start codon in the mRNA. This allowed the alternative Shine–Dalgarno sequences to vary in their sequence and spacing with respect to the start codon. We placed these sequences upstream of a gene encoding a fusion between *cat* (chloramphenicol acetyl transferase) and *upp* (uracil phosphoribosyl transferase), which allows either positive or negative selection on the expression of the gene fusion upon the addition of distinct small molecules. In the presence of chloramphenicol the *cat* gene allows the selection of functional sequences on chloramphenicol and in the presence of 5-fluorouracil (5FU) the *upp* gene allows selection of nonfunctional sequences.<sup>[87]</sup>

We first selected for 5FU for sequences upstream of the *cat-upp* fusion that do not allow the gene to be translated by the cell's endogenous ribosome, thus creating a pool of potentially orthogonal ribosome binding sites. Next we created a library, based on the gene for the *rrnB* ribosomal RNA operon. This library allows cells to produce ribosomes that contain all possible combinations of nucleotides at the 3'-end of 16S rRNA. The growth of cells in the presence of chloramphenicol allows selection for orthogonal ribosomes that specifically translate the orthogonal mRNAs. From the 10<sup>9</sup> combinations of potential ribosome mRNA pairs interrogated in this two-step selection we found three classes of orthogonal ribosome/orthogonal mRNA pairs. These pairs differ in the base paired sequence by which the ribosome binds to its cognate mRNA (Figure 3).

### 5.3. Evolving an Orthogonal Ribosome for Quadruplet Decoding

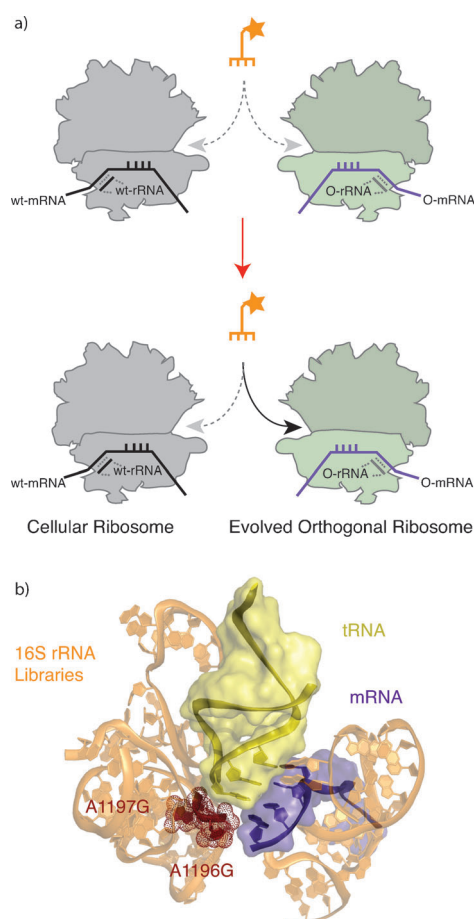
With orthogonal ribosome/mRNA pairs in hand we set out to evolve an orthogonal ribosome to efficiently decode quadruplet codons, which are poorly decoded on the natural ribosome.<sup>[86]</sup> To create a ribosome that decodes quadruplet codons we first designed and created 13 structurally guided<sup>[31,88–90]</sup> libraries in the decoding center within the A site of the ribosome, which is responsible for maintaining the fidelity of triplet decoding. Each library contains approximately 10<sup>8</sup> members, and together the libraries cover a molecular surface defined by 127 nucleotides of 16S rRNA. To select for



**Figure 3.** Creation of orthogonal ribosome/mRNA pairs by gene duplication and specialization. The cellular ribosome (grey) translates natural mRNA (black). The orthogonal ribosome (green) translates an orthogonal mRNA (O-mRNA, blue).

orthogonal ribosomes that efficiently decode extended anticodon tRNAs we combined the library of orthogonal ribosomes with an extended anticodon tRNA that is aminoacylated by the *E. coli* seryl-tRNA synthetase, and incorporates serine in response to the AGA quadruplet codon. Cells also contained an orthogonal mRNA bearing a *cat* gene with an in-frame AAGA codon. Reading of the AAGA codon on mutant orthogonal ribosomes, using the extended anticodon tRNA, led to synthesis of the full-length chloramphenicol acetyl transferase protein. This allowed the selection on chloramphenicol, of cells bearing orthogonal ribosomes that more efficiently decode extended anticodon tRNAs (Figure 4).

We selected a new orthogonal ribosome, Ribo-Q1, that was able to efficiently decode a range of different quadruplet codons using extended anticodon tRNAs.<sup>[86]</sup> Indeed, the levels of chloramphenicol resistance in cells bearing Ribo-Q1, a gene encoding *cat* with a AGGA codon, and the corresponding extended anticodon tRNA that is aminoacylated by *E. coli* seryl-tRNA synthetase was comparable to the levels of resistance in cells bearing a wild-type *cat* gene. This observation suggests that the level of quadruplet decoding on the evolved ribosome, in the presence of an efficiently aminoacylated tRNA, can approach the level of triplet decoding on the natural ribosome. Ribo-Q1 was derived from Ribo-X (an evolved orthogonal ribosome selected for enhanced amber suppression which allowed the first efficient and quantitative incorporation of multiple identical unnatural amino acids at specific sites<sup>[84]</sup>) and was additionally able to efficiently decode the amber codon on an orthogonal mRNA using amber suppressor tRNAs. Although we mutated 127 nucleotides in the decoding center of Ribo-X to select for Ribo-Q1, Ribo-Q1 contains only two mutations with respect to Ribo-X.



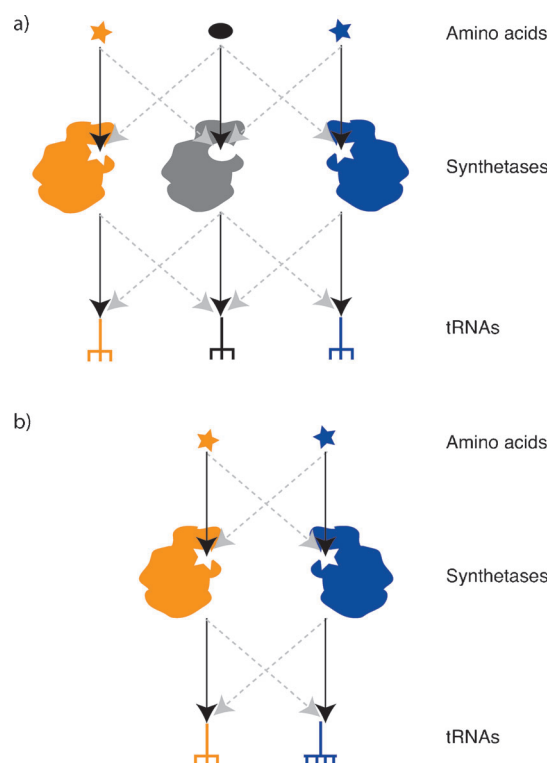
**Figure 4.** Creation of a ribosome for efficiently decoding quadruplet codons using extended anticodon tRNAs. a) Quadruplet codons are inefficiently decoded on the cellular and orthogonal ribosome by aminoacylated, extended anticodon tRNAs (orange). The orthogonal ribosome can be evolved in the laboratory (red arrow) to efficiently decode quadruplet codons. b) The mutations (in red) selected in the 16S ribosomal RNA (rRNA) of the orthogonal ribosome. Image created using Pymol (<http://www.pymol.org>) and Protein Data Bank (PDB) accession 2J00.

Mapping the selected mutations onto a crystal structure of the ribosome bound to a tRNA (Figure 4b) reveals that the selected mutations cluster where the fourth base of the extended anticodon might be expected to reside. This observation suggests that the selected mutations may result in a local rearrangement of ribosomal RNA structure. We demonstrated that Ribo-Q1 had excellent translational fidelity. Moreover we were able to efficiently decode several distinct quadruplet codons using tRNAs bearing distinct extended anticodons and Ribo-Q1. This work therefore provides several blank codons that can be specifically decoded on the orthogonal mRNA by Ribo-Q1, and potentially assigned to new amino acids.

#### 5.4. Synthetases and tRNAs for an Orthogonal Genetic Code

To take advantage of Ribo-Q1 for incorporating multiple distinct unnatural amino acids we require an orthogonal

synthetase/tRNA pair for each new codon. Two useful aminoacyl-tRNA synthetase (RS) pairs that are orthogonal in *E. coli*, namely the *Mj*TyrRS/tRNA<sub>CUA</sub> pair<sup>[11]</sup> and pyrrolysyl (Pyl)RS/tRNA<sub>CUA</sub> pair from *Methanosarcina* species,<sup>[91,92]</sup> have been discovered by import of the pairs from archaea (Figure 5). This approach takes advantage of the fact that, while the genetic code is near universally conserved between organisms, the identity elements by which these archaeal synthetases recognize their cognate tRNA have fortuitously diverged over the timescale of natural evolution to create orthogonal pairs.<sup>[93–98]</sup> The active site of *Mj*TyrRS has been evolved in the laboratory to allow the incorporation of a range of aromatic unnatural amino acid variants.<sup>[11]</sup> The PylRS/tRNA<sub>CUA</sub> pair is a recent discovery in certain methanogens that naturally recognizes pyrrolysine and incorporates this amino acid in response to the amber stop codon.<sup>[91]</sup> Unlike the pathway for incorporating selenocysteine, a synthetase/tRNA and amino acid are sufficient to direct the incorporation of pyrrolysine in response to the amber stop codon.<sup>[91]</sup> Because the PylRS/tRNA<sub>CUA</sub> pair is orthogonal with respect to *E. coli* synthetases and tRNAs and does not use the natural 20 amino acids, it has been possible to use this pair and



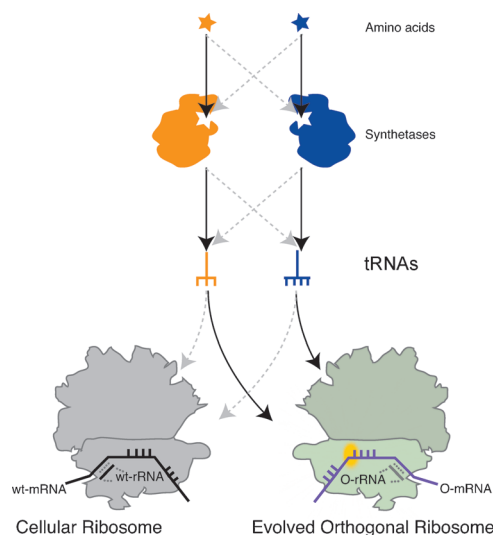
**Figure 5.** Orthogonal aminoacyl-tRNA synthetases for genetic code expansion. a) The pyrrolysyl-tRNA synthetase/tRNA<sub>CUA</sub> pair from *Methanosarcina* species (orange) can use, or be evolved to use, unnatural amino acids (orange star). The *Mj*Tyrosyl-tRNA synthetase/tRNA<sub>CUA</sub> pair (blue) can use, or be evolved to use, unnatural amino acids (blue star). Both synthetase tRNA pairs are orthogonal with respect to the *E. coli* synthetases and tRNAs (grey/black), that use natural amino acids (black oval). b) Pyrrolysyl-tRNA synthetase/tRNA<sub>CUA</sub>-derived pairs and *Mj*Tyrosyl-tRNA synthetase/tRNA<sub>UCCU</sub>-derived pairs are mutually orthogonal, and orthogonal with respect to endogenous synthetases and tRNAs.

its synthetically evolved variants to incorporate diverse unnatural amino acids into proteins.<sup>[91,92,99–109]</sup>

We demonstrated that the noncognate *Mj*TyrRS/*Methanoscarcina barkeri* (*Mb*)tRNA<sub>CUA</sub> and *Mb*PylRS/*Mj* tRNA<sub>CUA</sub> pairs do not function as amber suppressors, suggesting that the *Mj*TyrRS/tRNA<sub>CUA</sub> pair and *Mb*PylRS/tRNA<sub>CUA</sub> pairs are mutually orthogonal in their aminoacylation specificities.<sup>[86]</sup> To use these synthetases together to direct the incorporation of distinct amino acids in response to distinct codons we needed to differentiate the codons they decode. To alter the *Mj*TyrRS/tRNA<sub>CUA</sub> pair so that it incorporates an unnatural amino acid in response to a unique codon we first mutated the anticodon of the *Mj*tRNA<sub>CUA</sub> to UCCU, thus creating *Mj*tRNA<sub>UCCU</sub>. Since *Mj*TyrRS recognizes the anticodon, this alteration produced a nonfunctional synthetase/tRNA pair. To create a synthetase that functions with *Mj*tRNA<sub>UCCU</sub> to incorporate a useful unnatural amino acid in response to a quadruplet codon, we swapped *Mj*TyrRS for *Mj*AzPheRS (a synthetase variant that recognizes *p*-azido-*L*-phenylalanine<sup>[110]</sup>), and created a library of mutants in the region of the synthetase that recognize the anticodon. We then selected variant *Mj*AzPheRS/tRNA<sub>UCCU</sub> pairs that can incorporate *p*-azido-*L*-phenylalanine in response to the AGGA codon (Figure 5 b).

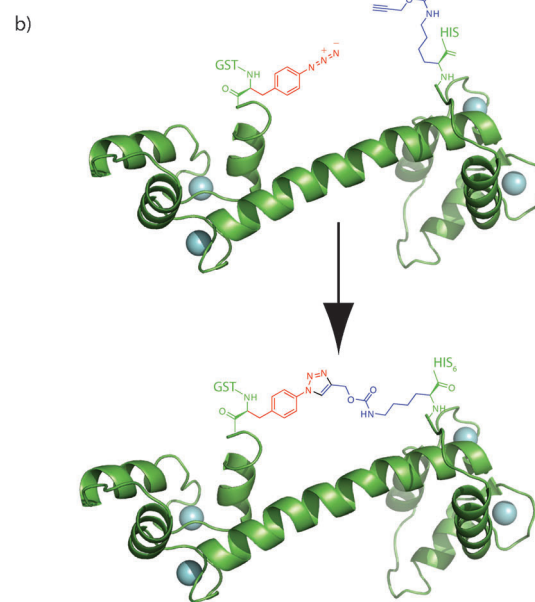
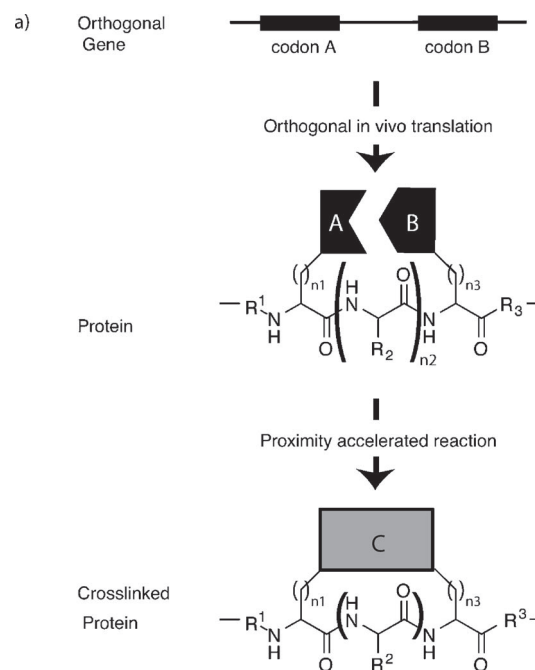
### 5.5. Genetically Encoding Emergent Properties in Proteins through Orthogonal Translation

To put together an orthogonal translation pathway for site-specifically incorporating two unnatural amino acids we combined Ribo-Q with an mRNA containing both a UAG and an AGGA codon, *Mj*AzPheRS/tRNA<sub>UCCU</sub>, and *Mb*PylRS/tRNA<sub>CUA</sub>. We demonstrated that full-length protein was generated from the mRNA only in the presence of *p*-azido-*L*-phenylalanine and an aliphatic alkyne, which we had



**Figure 6.** An orthogonal translation pathway for incorporating multiple unnatural amino acids. The modular combination of unnatural amino acids, mutually orthogonal evolved synthetases, tRNAs, and an evolved orthogonal ribosome/mRNA pair allows unnatural amino acids to be encoded on the orthogonal mRNA.

previously shown to be a good substrate for PylRS,<sup>[103]</sup> and that both amino acids were incorporated at the genetically encoded sites as judged by mass spectrometry. This experiment demonstrated the creation of a parallel pathway in the cells for the incorporation of unnatural amino acids in response to amber and quadruplet codons on the orthogonal mRNA (Figure 6).<sup>[86]</sup>



**Figure 7.** Programming properties that emerge from combinations of amino acids into recombinant proteins. a) Genetically programming protein cyclization. b) Structure of calmodulin indicating the sites of two genetically encoded unnatural amino acids specifically cyclize the protein in a proximity accelerated reaction (with Cu<sup>+</sup> catalyst) to provide a redox insensitive nanoscale crosslink.

The ability to direct two unnatural amino acids into recombinant proteins allows us to program properties into proteins that are not a property of either amino acid individually but emerge from the interaction between the two amino acids (Figure 7a). We demonstrated that by encoding an aliphatic alkyne- and an azide-containing amino acid it is possible to genetically program a rapid, proximity accelerated cycloaddition between these bioorthogonal groups to form a nanoscale, redox insensitive, triazole crosslink in a protein (Figure 7b).<sup>[86]</sup> Extensions of this approach may allow us to rapidly explore all possible crosslinks in proteins, and find utility in trapping particular functional states of proteins or in stabilizing protein therapeutics.

Since the synthetases derived from *MjTyrRS*<sup>[11]</sup> and *PyIRS*<sup>[91,92,99–109]</sup> have each been used to encode numerous unnatural amino acids, it is now possible to encode several hundred pairwise combinations of unnatural amino acids into proteins by simple extension of our approach. By encoding new combinations of unnatural amino acids new properties, such as fluorescence, may be programmed into proteins, and this may facilitate the labeling of specific proteins in vivo.

## 6. De novo Evolution of Orthogonal Aminoacyl-tRNA Synthetase/tRNA Pairs

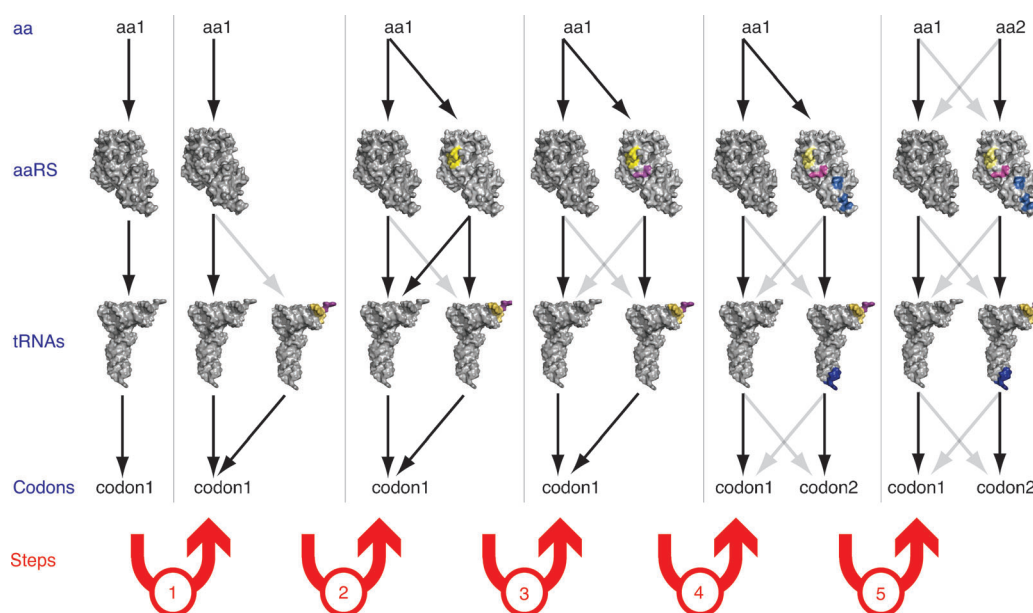
Ribo-Q1 provides numerous additional codons on the orthogonal mRNA. However, since only two orthogonal synthetase/tRNA pairs exist that can be used to incorporate distinct amino acids, only two distinct unnatural amino acids can be incorporated into a protein in the cell. A clear challenge in going from incorporating two unnatural amino acids to the synthesis of completely unnatural polymers is therefore to discover or invent strategies for generating new

orthogonal aminoacyl-tRNA synthetase/tRNA pairs that can be used to encode additional unnatural amino acids into a single polypeptide.

The two existing orthogonal synthetase tRNA pairs in *E. coli* were derived by import from heterologous organisms, thus taking advantage of the fact that while the genetic code is near universally conserved between known organisms the sequences and structures of synthetases and tRNAs have diverged through evolution. However, it is unclear how many mutually orthogonal synthetase/tRNA combinations can be discovered by taking advantage of natural evolutionary divergence. Moreover since the evolutionary record suggests that the current set of synthetases and tRNAs arose by gene duplication and specialization from a simpler basis set (for example, tyrosyl- and tryptophanyl-tRNA synthetase appear to be derived from a common ancestor),<sup>[111]</sup> we realized that it might be possible to extend this evolutionary process in the laboratory to generate new orthogonal synthetases and tRNAs de novo. We demonstrated that by a series of genetic selections for structurally targeted libraries in a synthetase/tRNA pair it is possible to evolve a new synthetase/tRNA pair that is orthogonal to both the synthetase/tRNA pair from which it was evolved and every other synthetase and tRNA in the cell (Figure 8).<sup>[112]</sup> This work demonstrates that the small number of orthogonal synthetase/tRNA pairs that have been discovered in nature do not intrinsically limit the potential of genetic code expansion for encoding additional unnatural amino acids.

## 7. Conclusions and Future Directions

Future work will aim to extend the strategies we have described, for providing new codons, and to pair new codons



**Figure 8.** De novo generation of orthogonal synthetase/tRNA pairs. Each step is a directed evolution or screening experiment. These experiments were performed with the *MjTyrosyl*-tRNA synthetase/tRNA<sub>CUA</sub> pair. codon1 is UAG, codon2 is AGGA. aa = amino acid, aaRS = aminoacyl-tRNA synthetase.



with additional orthogonal synthetases and tRNAs to extend the orthogonal genetic code to the cellular synthesis of completely unnatural polymers. We will also investigate further evolving the orthogonal ribosome to allow the in vivo biosynthesis of unnatural polymers composed on non- $\alpha$ -L amino acids. This biosynthesis will likely require evolution of the peptidyl-transferase center and possibly other parts of the ribosome, however, the demonstrated evolvability of the orthogonal ribosome provides a starting point for this approach. By creating cells endowed with genetically encoded heritable polymers it may be possible to explore the combinatorial biosynthesis of materials and therapeutics. It may also be possible to investigate whether life with additional genetically encoded polymers can evolve functions that natural biology cannot.

Work in our laboratory is supported by The Medical Research Council (U105181009, UD99999908), The European Research Council, The Human Frontiers of Science Program, and the Louis Jenatet Foundation (LJ-YICA). K.W. was funded by a Research Fellowship from Trinity College, Cambridge.

Received: July 18, 2011

Published online: January 19, 2012

- [1] A. Tsugita, H. Fraenkelconrat, *Proc. Natl. Acad. Sci. USA* **1960**, 46, 636.
- [2] F. H. Crick, S. Brenner, R. Watstobi, L. Barnett, *Nature* **1961**, 192, 1227.
- [3] M. W. Nirenberg, J. H. Matthae, O. W. Jones, *Proc. Natl. Acad. Sci. USA* **1962**, 48, 104.
- [4] H. G. Khorana, *Harvey Lect.* **1966**, 62, 79.
- [5] H. G. Khorana, H. Buchi, H. Ghosh, N. Gupta, T. M. Jacob, H. Kossel, R. Morgan, S. A. Narang, E. Ohtsuka, R. D. Wells, *Cold Spring Harbor Symp. Quant. Biol.* **1966**, 31, 39.
- [6] F. H. C. Crick, *J. Mol. Biol.* **1968**, 38, 367.
- [7] G. Baldini, B. Martoglio, A. Schachenmann, C. Zugliani, J. Brunner, *Biochemistry* **1988**, 27, 7951.
- [8] C. J. Noren, S. J. Anthonycahill, M. C. Griffith, P. G. Schultz, *Science* **1989**, 244, 182.
- [9] J. D. Bain, C. G. Glabe, T. A. Dix, A. R. Chamberlin, E. S. Diala, *J. Am. Chem. Soc.* **1989**, 111, 8013.
- [10] D. Mendel, V. W. Cornish, P. G. Schultz, *Annu. Rev. Biophys. Biomol. Struct.* **1995**, 24, 435.
- [11] J. M. Xie, P. G. Schultz, *Nat. Rev. Mol. Cell Biol.* **2006**, 7, 775.
- [12] J. W. Chin, *EMBO J.* **2011**, 30, 2312.
- [13] S. Greiss, J. W. Chin, *J. Am. Chem. Soc.* **2011**, 133, 14196.
- [14] K. Josephson, M. C. T. Hartman, J. W. Szostak, *J. Am. Chem. Soc.* **2005**, 127, 11727.
- [15] T. Kawakami, H. Murakami, H. Suga, *Chem. Biol.* **2008**, 15, 32.
- [16] A. C. Forster, Z. P. Tan, M. N. L. Nalam, H. N. Lin, H. Qu, V. W. Cornish, S. C. Blacklow, *Proc. Natl. Acad. Sci. USA* **2003**, 100, 6353.
- [17] T. M. Schmeing, V. Ramakrishnan, *Nature* **2009**, 461, 1234.
- [18] V. Ramakrishnan, *Cell* **2002**, 108, 557.
- [19] A. Simonetti, S. Marzi, L. Jenner, A. Myasnikov, P. Romby, G. Yusupova, B. P. Klaholz, M. Yusupov, *Cell. Mol. Life Sci.* **2009**, 66, 423.
- [20] B. S. Laursen, H. P. Sorensen, K. K. Mortensen, H. U. Sperling-Petersen, *Microbiol. Mol. Biol. Rev.* **2005**, 69, 101.
- [21] M. L. Sprengart, A. G. Porter, *Mol. Microbiol.* **1997**, 24, 19.
- [22] J. Shine, L. Dalgarno, *Biochem. J.* **1974**, 141, 609.
- [23] J. A. Steitz, K. Jakes, *Proc. Natl. Acad. Sci. USA* **1975**, 72, 4734.
- [24] C. R. Woese, G. J. Olsen, M. Ibba, D. Soll, *Microbiol. Mol. Biol. Rev.* **2000**, 64, 202.
- [25] P. O'Donoghue, Z. Luthey-Schulten, *Microbiol. Mol. Biol. Rev.* **2003**, 67, 550.
- [26] F. J. LaRiviere, A. D. Wolfson, O. C. Uhlenbeck, *Science* **2001**, 294, 165.
- [27] J. M. Ogle, F. V. Murphy, M. J. Tarry, V. Ramakrishnan, *Cell* **2002**, 111, 721.
- [28] J. M. Ogle, A. P. Carter, V. Ramakrishnan, *Trends Biochem. Sci.* **2003**, 28, 259.
- [29] M. V. Rodnina, K. B. Gromadski, U. Kothe, H. J. Wieden, *FEBS Lett.* **2005**, 579, 938.
- [30] M. V. Rodnina, W. Wintermeyer, *Biochem. Soc. Trans.* **2011**, 39, 658.
- [31] A. Korostelev, S. Trakhanov, M. Laurberg, H. F. Noller, *Cell* **2006**, 126, 1065.
- [32] J. F. Atkins, R. B. Weiss, S. Thompson, R. F. Gesteland, *Annu. Rev. Genet.* **1991**, 25, 201.
- [33] P. J. Farabaugh, *Microbiol. Rev.* **1996**, 60, 103.
- [34] J. F. Atkins, G. R. Bjork, *Microbiol. Mol. Biol. Rev.* **2009**, 73, 178.
- [35] R. B. Weiss, D. M. Dunn, J. F. Atkins, R. F. Gesteland, *Cold Spring Harbor Symp. Quant. Biol.* **1987**, 52, 687.
- [36] W. M. Huang, S. Z. Ao, S. Casjens, R. Orlandi, R. Zeikus, R. Weiss, D. Winge, M. Fang, *Science* **1988**, 239, 1005.
- [37] R. B. Weiss, D. M. Dunn, J. F. Atkins, R. F. Gesteland, *Prog. Nucleic Acid Res. Mol. Biol.* **1990**, 39, 159.
- [38] I. Benhar, H. Engelbergkulka, *Cell* **1993**, 72, 121.
- [39] E. Tendam, I. Brierley, S. Inglis, C. Pleij, *Nucleic Acids Res.* **1994**, 22, 2304.
- [40] E. B. Tendam, P. W. G. Verlaan, C. W. A. Pleij, *RNA* **1995**, 1, 146.
- [41] R. C. L. Olsthoorn, R. Reumerman, C. W. Hilbers, C. W. A. Pleij, H. A. Heus, *Nucleic Acids Res.* **2010**, 38, 7665.
- [42] R. B. Weiss, D. M. Dunn, A. E. Dahlberg, J. F. Atkins, R. F. Gesteland, *EMBO J.* **1988**, 7, 1503.
- [43] M. O'Connor, *Nucleic Acids Res.* **2002**, 30, 759.
- [44] T. Jacks, H. D. Madhani, F. R. Masiarz, H. E. Varmus, *Cell* **1988**, 55, 447.
- [45] R. B. Weiss, D. M. Dunn, M. Shuh, J. F. Atkins, R. F. Gesteland, *New Biol.* **1989**, 1, 159.
- [46] P. J. Farabaugh, *J. Biol. Chem.* **1995**, 270, 10361.
- [47] D. L. Riddle, J. Carbon, *Nature New Biol.* **1973**, 242, 230.
- [48] D. L. Riddle, J. R. Roth, *J. Mol. Biol.* **1970**, 54, 131.
- [49] J. Yourno, T. Kohno, *Science* **1972**, 175, 650.
- [50] B. Moore, B. C. Persson, C. C. Nelson, R. F. Gesteland, J. F. Atkins, *J. Mol. Biol.* **2000**, 298, 195.
- [51] T. J. Magliery, J. C. Anderson, P. G. Schultz, *J. Mol. Biol.* **2001**, 307, 755.
- [52] J. C. Anderson, T. J. Magliery, P. G. Schultz, *Chem. Biol.* **2002**, 9, 237.
- [53] M. Taki, J. Matsushita, M. Sisido, *ChemBioChem* **2006**, 7, 425.
- [54] J. D. Bain, E. S. Diala, C. G. Glabe, D. A. Wacker, M. H. Lyttle, T. A. Dix, A. R. Chamberlin, *Biochemistry* **1991**, 30, 5411.
- [55] J. R. Roesser, M. S. Chorghade, S. M. Hecht, *Biochemistry* **1986**, 25, 6361.
- [56] C. H. Ma, W. Kudlicki, O. W. Odom, G. Kramer, B. Hardesty, *Biochemistry* **1993**, 32, 7939.
- [57] T. Hohsaka, Y. Ashizuka, H. Murakami, M. Sisido, *J. Am. Chem. Soc.* **1996**, 118, 9778.
- [58] M. W. Nowak, J. P. Gallivan, S. K. Silverman, C. G. Labarca, D. A. Dougherty, H. A. Lester, *Ion Channels, Pt. B, Vol. 293*, Academic Press, San Diego, **1998**, p. 504.
- [59] D. L. Beene, D. A. Dougherty, H. A. Lester, *Curr. Opin. Neurobiol.* **2003**, 13, 264.
- [60] D. A. Dougherty, *Curr. Opin. Chem. Biol.* **2000**, 4, 645.

- [61] T. Hohsaka, R. Abe, K. Shiraga, M. Sisido, *Nucleic Acids Res. Suppl.* **2003**, 271.
- [62] T. Hohsaka, N. Muranaka, C. Komiyama, K. Matsui, S. Takaura, R. Abe, H. Murakami, M. Sisido, *FEBS Lett.* **2004**, 560, 173.
- [63] T. Hohsaka, M. Sisido, *Curr. Opin. Chem. Biol.* **2002**, 6, 809.
- [64] M. Taki, Y. Tokuda, T. Ohtsuki, M. Sisido, *J. Biosci. Bioeng.* **2006**, 102, 511.
- [65] Y. Tokuda, M. Taki, M. Sisido, *Nucleic Acids Symp. Ser.* **2006**, 50, 277.
- [66] H. Taira, M. Fukushima, T. Hohsaka, M. Sisido, *J. Biosci. Bioeng.* **2005**, 99, 473.
- [67] N. Muranaka, T. Hohsaka, M. Sisido, *Nucleic Acids Res.* **2006**, 34, e7.
- [68] T. Hohsaka, Y. Ashizuka, H. Taira, H. Murakami, M. Sisido, *Biochemistry* **2001**, 40, 11060.
- [69] T. Ohtsuki, T. Manabe, M. Sisido, *FEBS Lett.* **2005**, 579, 6769.
- [70] D. Kajihara, R. Abe, I. Iijima, C. Komiyama, M. Sisido, T. Hohsaka, *Nat. Methods* **2006**, 3, 923.
- [71] A. M. Shafer, T. Kalai, S. Q. B. Liu, K. Hideg, J. C. Voss, *Biochemistry* **2004**, 43, 8470.
- [72] E. A. Rodriguez, H. A. Lester, D. A. Dougherty, *Proc. Natl. Acad. Sci. USA* **2006**, 103, 8650.
- [73] S. W. Santoro, J. C. Anderson, V. Lakshman, P. G. Schultz, *Nucleic Acids Res.* **2003**, 31, 6700.
- [74] J. C. Anderson, N. Wu, S. W. Santoro, V. Lakshman, D. S. King, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **2004**, 101, 7566.
- [75] C. M. Dunham, M. Selmer, S. S. Phelps, A. C. Kelley, T. Suzuki, S. Joseph, V. Ramakrishnan, *RNA* **2007**, 13, 817.
- [76] J. F. Curran, M. Yarus, *Science* **1987**, 238, 1545.
- [77] S. S. Phelps, C. Gaudin, S. Yoshizawa, C. Benitez, D. Fourmy, S. Joseph, *J. Mol. Biol.* **2006**, 360, 610.
- [78] S. E. Walker, K. Fredrick, *J. Mol. Biol.* **2006**, 360, 599.
- [79] Q. Qian, J. N. Li, H. Zhao, T. G. Hagervall, P. J. Farabaugh, G. R. Bjork, *Mol. Cell* **1998**, 1, 471.
- [80] O. Rackham, J. W. Chin, *Nat. Chem. Biol.* **2005**, 1, 159.
- [81] O. Rackham, J. W. Chin, *J. Am. Chem. Soc.* **2005**, 127, 17584.
- [82] O. Rackham, K. Wang, J. W. Chin, *Nat. Chem. Biol.* **2006**, 2, 254.
- [83] W. L. An, J. W. Chin, *Proc. Natl. Acad. Sci. USA* **2009**, 106, 8477.
- [84] K. H. Wang, H. Neumann, S. Y. Peak-Chew, J. W. Chin, *Nat. Biotechnol.* **2007**, 25, 770.
- [85] O. P. T. Barrett, J. W. Chin, *Nucleic Acids Res.* **2010**, 38, 2682.
- [86] H. Neumann, K. H. Wang, L. Davis, M. Garcia-Alai, J. W. Chin, *Nature* **2010**, 464, 441.
- [87] C. Fabret, S. D. Ehrlich, P. Noirot, *Mol. Microbiol.* **2002**, 46, 25.
- [88] M. Selmer, C. M. Dunham, F. V. Murphy, A. Weixlbaumer, S. Petry, A. C. Kelley, J. R. Weir, V. Ramakrishnan, *Science* **2006**, 313, 1935.
- [89] H. F. Noller, M. M. Yusupov, G. Z. Yusupova, A. Baucom, K. Lieberman, L. Lancaster, A. Dallas, K. Fredrick, T. N. Earnest, J. H. D. Cate, *Cold Spring Harbor Symp. Quant. Biol.* **2001**, 66, 57.
- [90] M. M. Yusupov, G. Z. Yusupova, A. Baucom, K. Lieberman, T. N. Earnest, J. H. D. Cate, H. F. Noller, *Science* **2001**, 292, 883.
- [91] A. Ambrogelly, S. Palioura, D. Soll, *Nat. Chem. Biol.* **2007**, 3, 29.
- [92] H. Neumann, S. Y. Peak-Chew, J. W. Chin, *Nat. Chem. Biol.* **2008**, 4, 232.
- [93] D. Soll, *Experientia* **1990**, 46, 1089.
- [94] R. Gieger, M. Sissler, C. Florentz, *Nucleic Acids Res.* **1998**, 26, 5017.
- [95] P. J. Beuning, K. Musier-Forsyth, *Biopolymers* **1999**, 52, 1.
- [96] P. Fechter, J. Rudinger-Thirion, M. Tukalo, R. Gieger, *Eur. J. Biochem.* **2001**, 268, 761.
- [97] T. Kobayashi, O. Nureki, R. Ishitani, A. Yaremchuk, M. Tukalo, S. Cusack, K. Sakamoto, S. Yokoyama, *Nat. Struct. Biol.* **2003**, 10, 425.
- [98] K. Nozawa, P. O'Donoghue, S. Gundllapalli, Y. Arais, R. Ishitani, T. Umehara, D. Soll, O. Nureki, *Nature* **2009**, 457, 1163.
- [99] T. Mukai, T. Kobayashi, N. Hino, T. Yanagisawa, K. Sakamoto, S. Yokoyama, *Biochem. Biophys. Res. Commun.* **2008**, 371, 818.
- [100] T. Fekner, X. Li, M. M. Lee, M. K. Chan, *Angew. Chem.* **2009**, 121, 1661; *Angew. Chem. Int. Ed.* **2009**, 48, 1633.
- [101] A. Gautier, D. P. Nguyen, H. Lusic, W. A. An, A. Deiters, J. W. Chin, *J. Am. Chem. Soc.* **2010**, 132, 4086.
- [102] X. Li, T. Fekner, J. J. Ottesen, M. K. Chan, *Angew. Chem.* **2009**, 121, 9348; *Angew. Chem. Int. Ed.* **2009**, 48, 9184.
- [103] D. P. Nguyen, M. M. G. Alai, P. B. Kapadnis, H. Neumann, J. W. Chin, *J. Am. Chem. Soc.* **2009**, 131, 14194.
- [104] D. P. Nguyen, M. M. G. Alai, S. Virdee, J. W. Chin, *Chem. Biol.* **2010**, 17, 1072.
- [105] D. P. Nguyen, H. Lusic, H. Neumann, P. B. Kapadnis, A. Deiters, J. W. Chin, *J. Am. Chem. Soc.* **2009**, 131, 8720.
- [106] C. R. Polycarpo, S. Herring, A. Bérubé, J. L. Wood, D. Söll, A. Ambrogelly, *FEBS Lett.* **2006**, 580, 6695.
- [107] S. Virdee, Y. Ye, D. P. Nguyen, D. Komander, J. W. Chin, *Nat. Chem. Biol.* **2010**, 6, 750.
- [108] S. Virdee, P. B. Kapadnis, T. Elliott, K. Lang, J. Madrzak, D. P. Nguyen, L. Riechmann, J. W. Chin, *J. Am. Chem. Soc.* **2011**, 133, 10708.
- [109] D. P. Nguyen, T. Elliott, M. Holt, T. W. Muir, J. W. Chin, *J. Am. Chem. Soc.* **2011**, 133, 11418.
- [110] J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, *J. Am. Chem. Soc.* **2002**, 124, 9026.
- [111] L. Ribas de Pouplana, P. Schimmel, *Trends Biochem. Sci.* **2001**, 26, 591.
- [112] H. Neumann, A. L. Slusarczyk, J. W. Chin, *J. Am. Chem. Soc.* **2010**, 132, 2142.