Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Biotechnology

Recent advances in genetic code engineering in *Escherichia coli*

Michael Georg Hoesl and Nediljko Budisa

The expansion of the genetic code is gradually becoming a core discipline in Synthetic Biology. It offers the best possible platform for the transfer of numerous chemical reactions and processes from the chemical synthetic laboratory into the biochemistry of living cells. The incorporation of biologically occurring or chemically synthesized non-canonical amino acids into recombinant proteins and even proteomes via reprogrammed protein translation is in the heart of these efforts. Orthogonal pairs consisting of aminoacyl-tRNA synthetase and its cognate tRNA proved to be a general tool for the assignment of certain codons of the genetic code with a maximum degree of chemical liberty. Here, we highlight recent developments that should provide a solid basis for the development of generalist tools enabling a controlled variation of chemical composition in proteins and even proteomes. This will take place in the frame of a greatly expanded genetic code with emancipated codons liberated from the current function or with totally new coding units.

Address
Technische Universität Berlin (Berlin Institute of Technology), Department of Chemistry, Biocatalysis Group, Franklinstraße 29, 10587 Berlin, Germany

Corresponding author: Budisa, Nediljko (budisa@biocat.tu-berlin.de)

Current Opinion in Biotechnology 2012, **23**:751–757

This review comes from a themed issue on **Tissue, cell and pathway engineering**

Edited by **Hal Alper** and **Wilfried Weber**

For a complete overview see the Issue and the Editorial

Available online 9th January 2012

0958-1669/$ – see front matter,

http://dx.doi.org/10.1016/j.copbio.2011.12.027

Introduction

Nature builds up proteins with the 20 canonical amino acids (cAAs) encoded by the 61 sense codons. However, these 20 side-chain functionalities are obviously not sufficient for proteins to cover all the chemical diversity necessary to maintain many vital biological functions in both unicellular and multicellular organisms. Evolution invented two strategies to increase the side-chain inventory: a small fraction of proteins is co-translationally equipped with special proteinogenic amino acids such as selenocysteine (Sec) and pyrrolysine (Pyl) by reassignment of termination codons. However, the major classes of chemical modifications that contribute to the protein structure/function diversity are post-translational modifications (PTMs). These reactions are selectively and timely coordinated chemistries performed by dedicated enzymes and enzymatic complexes, usually in specialized cell compartments.

Certainly, one of the main goals of Synthetic Biology is to generate new and emergent biological functions in streamlined cells which are equipped with 'tailor-made biochemical production lines'. However, it is extremely difficult to mimic nature's complex machineries such as the PTM-apparatus. Thus, we usually highjack and/or divert cellular systems such as protein translation to gain additional chemical diversity. To achieve this goal, we need to find a way for efficient cellular uptake, metabolic stability and translational activity (i.e. incorporation) of useful non-canonical amino acids (ncAAs) which are usually chemically synthesized. Furthermore, we need to (re)assign coding units (i.e. codons) in the genetic code to accommodate ncAAs into target proteins.

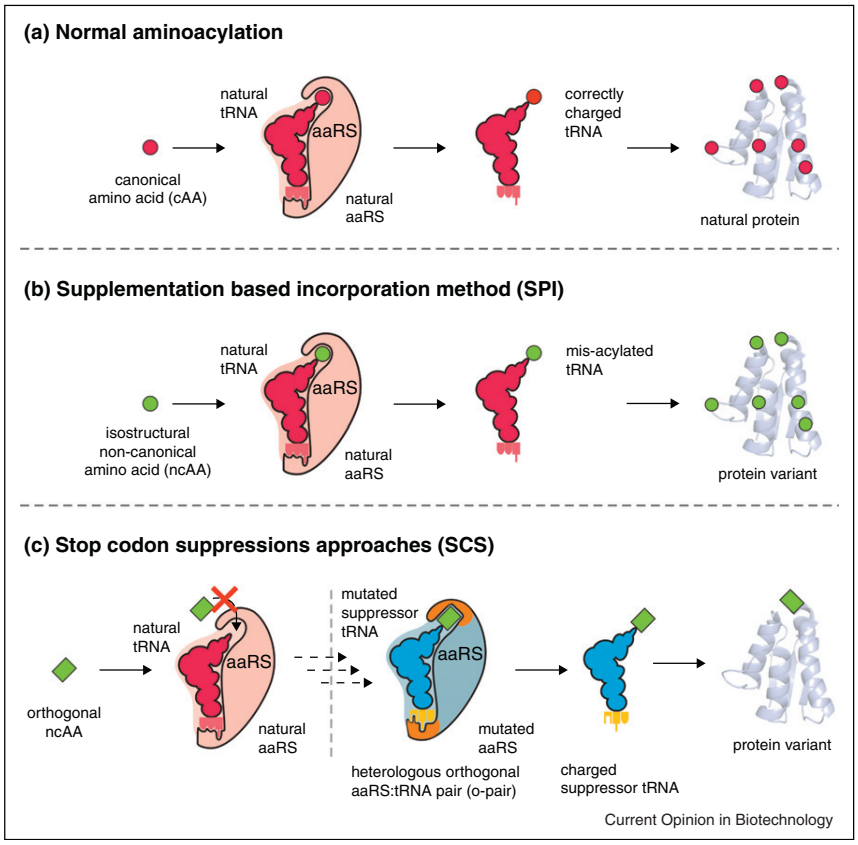
Here, we briefly sketch the most important developments in the field in the last three years. For further and more comprehensive information we refer to other recently published reviews which give an excellent overview of the methods, the incorporated ncAAs, and their diverse applications [1••,2••,3,4•,5•].

General *in vivo* incorporation strategies

ncAAs for protein engineering can be divided into two groups: first, amino acids which are isostructural to cAAs and therefore recognized by the endogenous host cell machinery, and second, amino acids which are orthogonal to the host cell system because they do not participate in conventional translation. To exploit the beneficial features of both ncAA groups for protein engineering, two distinct *in vivo* approaches are available for their co-translational incorporation (see Figure 1).

For isostructural ncAAs, residue-specific replacement of cAAs is performed with the supplementation-based incorporation method (SPI) using auxotrophic host strains [6–8]. In contrast, orthogonal ncAAs are added to the amino acid repertoire by site-specific incorporation in response to stop or quadruplet codons (stop codon suppression, SCS) using orthogonal aminoacyl-tRNA synthetase:tRNA pairs (o-pairs) [9,10]. Both approaches have particular advantages and pitfalls, depending on the specific biological questions and problems being addressed. Biological functions based on collective effects of many residues (e.g. activity [11], stability [12,13], or conformational preferences [14]) are particularly well studied and tailored by

Figure 1



Aminoacylation with canonical and non-canonical amino acids for protein translation. **(a)** In the natural scenario, tRNA aminoacylation is catalyzed by the corresponding aminoacyl-tRNA synthetase (aaRS) responsible for charging the tRNA with the cognate amino acids. **(b)** The supplementation-based incorporation method (SPI) exploits the natural substrate tolerance of the endogenous host aaRSs by using auxotrophic host strains. This allows the simultaneous exchange of many residues in a target protein by sense-codon reassignment [57]. The substrates for this procedure are non-canonical amino acids (ncAAs) which are isostructural to their canonical counterparts (i.e. atomic mutation concept [58]). **(c)** Stop codon suppression methodologies (SCS) are nominally site-specific and make use of a heterologous orthogonal aaRS:tRNA pair (o-pair) to incorporate an orthogonal amino acid in response to a stop or quadruplet codon. At this level, orthogonality is defined by a lack of cross-reactivity between the o-pair (including the ncAA) and the endogenous host synthetases, amino acids and tRNAs. Orthogonality of the o-pair is based on species-specific differences in tRNA recognition by the aaRS. The o-pair has to be evolved to enable the participation of a desired orthogonal ncAA in ribosome mediated protein synthesis (reviewed in [1••]).

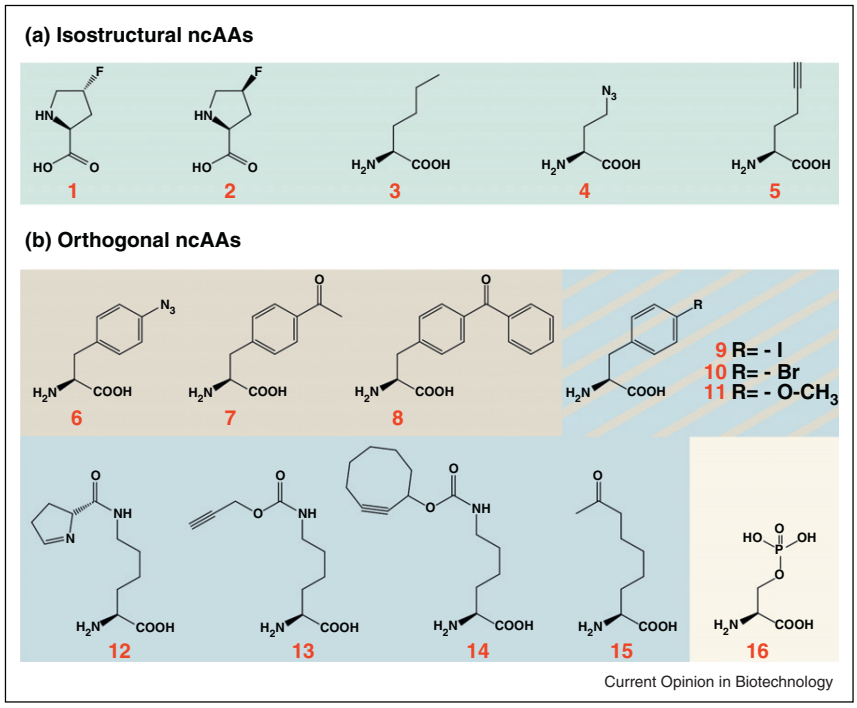
SPI. In contrast, the site-specific SCS approach enables the punctual dissection and modification of proteins (e.g. [15,16]).

Methanocaldococcus jannaschii TyrRS versus *Methanosarcinaceae* PylRSs

Since the introduction of orthogonality as a tool for protein engineering by Furter [9], various o-pairs were developed (see [1••] and [17] for comprehensive lists). Nevertheless, until recently the most frequently used o-pair was based on the TyrRS from *Methanocaldococcus jannaschii* (*mj*TyrRS). Altogether, the incorporation of around 40 different ncAAs, mainly aromatic analogs of Phe and Tyr, was reported [1••]. Among these, 4-benzoylphenylalanine (Bpa) for crosslinking [18] or 4-acetylphenylalanine (ActF) [19] and 4-azidophenylalanine (AzF) [20] for site-specific coupling reactions are certainly the most popular ones.

Now we are witnessing the spreading utilization of pyrrolysyl-tRNA synthetase (PylRS) and its cognate suppressor tRNA *pylT* from *Methanosarcinaceae* species in the whole field. In the native context, this 'natural o-pair' enables cells to incorporate the mostly aliphatic Pyl into target proteins in response to the amber (UAG) stop codon. In contrast to *mj*TyrRS, however, PylRSs already naturally display a broad substrate tolerance towards orthogonal ncAAs [21]. Within a very short time, a large number of different PylRS variants for incorporation of highly valuable aliphatic ncAAs were developed (see Figure 2 and [4•]). Remarkably, the evolved Pyl o-pair systems show a very high flexibility with respect to amino

Figure 2



Survey of some important ncAAs incorporated by SPI or SCS methods. **(a)** Isostructural ncAAs incorporated by SPI: (1) (4*s*)-fluoroproline ((4*s*)-FPro) [12], (2) (4*R*)-fluoroproline ((4*R*)-FPro) [13], (3) norleucine (Nle) [59], (4) azidohomoalanine (Aha) [60], (5) homopropargylglycine (Hpg) [61]. 1 and 2 are incorporated in response to Pro codons while 3–5 are substrates for the endogenous MetRS. **(b)** Orthogonal ncAAs incorporated by SCS: (6) 4-azidophenylalanine (AzF) [62], (7) 4-acetylphenylalanine (ActF) [63], (8) 4-benzoylphenylalanine (Bpa) [64], (9) 4-iodophenylalanine (IF) [22,65], (10) 4-bromophenylalanine (BrF) [22,66], (11) *O*-methyltyrosine (OmeY) [10,23], (12) pyrroline-carboxy-lysine (Pcl) [56••], (13) *N6*-[(2-propynyloxy)-carbonyl]-lysine (PoxK) [67], (14) *N6*-[(cyclooct-2-yn-1-yloxy)carbonyl]-lysine (CoK) [68], (15) 2-amino-8-oxononanoic acid (KetoK) [69], (16) *O*-phosphoserine (Sep) [54••]. 6–8 are incorporated by variants of *mj*TyrRS, 12-15 by variants of PylRS, and 9–11 were incorporated by both systems. Currently available PylRS-based o-pairs are derived from *Methanosarcina barkeri* [70], *Methanosarcina mazei* [71], and *Desulfitobacterium hafniense* [72]. *O*-phosphoserine was genetically encoded by a *Methanococcus maripaludis* SepRS:*mj*tRNA$^{Cys}_{CUA}$ o-pair and a mutated EF-Tu. If available, particular applications were cited for the presented ncAAs. For a more comprehensive view we refer to reviews [1••,2••,3,4•,5•].

acid charging. For example, the groups of Wang and Liu evolved PylRSs for aromatic Phe and Tyr analogs (see Figure 2 and [22,23]). Most importantly, the system is very tolerant in regard to different tRNA anticodons. UAG, UAA, UGA, UAGA were read without further adaptations in the tRNA^{Pyl} [24]. Doubtlessly, the PylRS systems are not only highly valuable tools for ncAA incorporation but are also superior to *mj*TyrRS systems. However, it should always be kept in mind that selection of orthogonal aaRS variants for new ncAAs is still a laborious and time consuming task.

Another important issue of o-pairs tackled only recently is the degree of substrate tolerance in these systems. Like endogenous aaRSs, orthogonal aaRSs can also display significant substrate tolerance. In contrast to SPI, the whole system maintains orthogonality towards all cAAs. Remarkably, Young *et al.* found that *mj*TyrRS evolved for 4-cyanophenylalanine can incorporate 18 different Phe and Tyr analogs [25]. Similarly, Miyake-Stoner *et al.* generated *mj*TyrRSs capable of incorporating a variety of differently fluorinated 4-methylphenylalanines which are useful in ^{19}F NMR studies [26]. These studies nicely demonstrate that the re-screening of previously available o-pairs can also be a strategy to identify an aaRS capable of charging a desired ncAA.

Parallel incorporation of multiple ncAAs in a single expression experiment

The incorporation of a single ncAA at a single position in a protein is certainly useful as tool to study particular biological problems (e.g. receptor–ligand interactions [27]). However, it is often not sufficient for many academic and biotechnological applications. This will leverage the whole field towards the introduction of two or more reactive handles or probes into recombinant proteins. The first experiment was reported by Schultz and coworkers [28] followed by recent works of Chin and Liu (reviewed in [29•]). The combination of *mj*TyrRS and *Methanosarcina mazei* PylRS derived o-pairs enabled the simultaneous reassignment of an amber stop and an ochre (UAA) [24] or quadruplet [30] codon. However, the

reassignment of further codons using SCS methodologies is currently impeded by two crucial limitations. First, suppressor tRNAs have to compete with the endogenous release factors RF1 and RF2 during translation at the ribosome. Second, frame-shift suppression is still very inefficient [31].

In contrast, a multiple incorporation of three and even more ncAAs into a single target protein is possible by SPI without substantial loss of protein yield. For example, the parallel incorporation of homopropargylglycine (Hpg, click handle), 4-azatryptophan (fluorescence tag), and (4*s*)-fluoroproline ((4*s*)-FPro) (stabilizing amino acid) yielded active tailor-made protein [32•]. Similarly, a highly fluorinated active lipase with amino acid exchanges at as many as 24 positions was reported [33]. We also explored the potential of multi-labeling by SPI in combination with site-specific SCS. In particular, Bpa along with norleucine or (4*s*)-FPro incorporation provided a useful combination of desired features in the recombinant protein [34]. At the same time, Yun and coworkers expressed a green fluorescent protein (GFP) variant harboring 3,4-dihydroxy-phenylalanine and Hpg or azidohomoalanine (Aha) using the same strategy [35].

Engineering of bacterial strains free of release factors or particular termination codons

The major challenge in the further development of all suppression-based methodologies is to achieve unlimited reassignments of the codons of interest, for example the amber stop codon. The main obstacles include release factor competition, catalytic performance of o-pairs and poorly understood mRNA context effects. In this regard, systems based on amber suppression have been considerably improved in the last years. These improvements include first, the use of more suppressor tRNA copies [36], second, optimized suppressor tRNAs for enhanced ncAA-tRNA binding to elongation factor Tu (EF-Tu) [37], third, enhanced co-expression plasmids for higher aaRS expression [38,39], and finally, the co-expression of the C-terminal domain of the ribosomal protein L11 [40]. Furthermore, the decreased affinity of orthogonal ribosomes towards RF1 led to higher amber suppression [41]. Despite these important improvements, the expression of target proteins with more than three in-frame amber stop codons had yet to be reported.

Doubtlessly, the most appealing idea to efficiently eliminate this drawback for the amber stop codon was the removal of RF1 from *Escherichia coli* since it would eliminate the competition reaction with the amber suppressor tRNA during translation. However, the *prfA* gene coding for RF1 is essential in *E. coli* [42]. Therefore, Church and coworkers set out to exchange all 314 TAG codons in the *E. coli* genome by TAA and remove RF1, subsequently [43]. In their study, they succeeded in generating four strains with around 80 complementary TAG to TAA mutations each. Unfortunately, the final step to unify these four subsets to one TAG free *E. coli* strain and the RF1 knockout was not performed in the study.

In contrast to this genome remodeling approach, the teams of Yokoyama/Sakamoto [44•] and Lei Wang [45•] came up with two different genetic strategies to solve the 'RF1 problem' without the removal of all TAG codons from the genome. In the frame of the first strategy it was speculated that the degradation of essential proteins due to ribosome stalling at unassigned UAG codons [46] may be a possible reason for RF1 essentiality. Thus, the *E. coli* strain was *trans* complemented with TAA ending versions of seven essential genes originally terminated with TAG. In the second step, an engineered o-pair reading UAG was introduced. Host cells configured in this way, allowed successful knockout of *prfA*. Not surprisingly, an enhanced version of this Δ*prfA* strain [47] enabled glutathione S-transferase expression with different ncAAs at up to seven amber positions in parallel. In contrast, the second strategy assumes that insufficient termination at UAA instead of UAG codons (RF1 and RF2 are both reading UAA) upon *prfA* knockout causes RF1 essentiality. Indeed, by enhancing RF2 expression and activity, Wang and coworkers could successfully knockout *prfA* [45•] from a genome reduced *E. coli* strain [48]. The approach enabled the successful read-though of 10 UAG codons in GFP.

It is still difficult to rationalize how these two different strategies can be reconciled. However, the Δ*prfA* strain designed by Wang and coworkers has the advantage of being independent of *trans* complementation. This permits the facile use of already available expression systems with o-pairs virtually without limitations.

New orthogonal pairs and mutually orthogonal *mj*TyrRSs

The successful knockout of RF1 leads to the liberation of the amber codon for complete reassignment. However, multiple *in vivo* incorporations with distinct ncAAs will need a further extended number of blank codons. A quite popular opinion in the academic community is that this 'lack of blank codons' could be solved by introducing quadruplet codons [49]. In addition, we are convinced that the liberation of some (rare) codons from their natural assignments (i.e. codon emancipation) might also be very powerful to obtain new blank codons for reassignment to ncAAs [50]. In any case, many different (and mutually orthogonal) o-pairs from various sources will be needed for this task. Therefore, a lot of effort should be made in the further development of existing o-pairs and the search for natural ones. For example, Hughes *et al.* rationally designed an o-pair based on *Saccharomyces cerevisiae* TrpRS:$tRNA^{Trp}$ for use in *E. coli* [17]. Furthermore, the *Caulobacter crescentus* HisRS:$tRNA^{His}$ pair was identified as a natural o-pair in *E. coli* [51]. Finally, starting from

the parent *mj*TyrRS:tRNA$^{Tyr}_{CUA}$ o-pair, two mutually orthogonal systems were developed recently [52•].

It should be kept in mind that only the design of o-pairs and their use in suitable host cells will not be sufficient for incorporation of many desired ncAAs. For example, the issue of ncAA-tRNA interaction with EF-Tu was not considered for a long time in the whole field. Sisido and co-workers were the first ones to discover that it is possible to improve EF-Tu interaction with various ncAA-tRNAs *in vitro* [53]. The most striking example for EF-Tu interaction engineering was recently provided by Söll and coworkers [54••]. In particular, they developed an amber suppression based co-translational *O*-phosphoserine (Sep) incorporation system starting from a natural pathway in *Methanococcus maripaludis*. The system is based on a naturally orthogonal SepRS and its establishment included first, engineering of a cognate amber suppressor tRNA^{Sep}, second, screening for EF-Tu mutants capable of binding Sep-tRNA^{Sep}, and finally, host cell engineering to block endogenous phosphoserine phosphatase activity. Interestingly, the strategy for designing the Sep encoding system was borrowed from a natural example, since natural selenocysteine incorporation includes a specialized elongation factor as well (reviewed in [55]).

Outlook — codon emancipated cells with maximum chemical liberty

The rapid development of new orthogonal pairs and new aaRS specificities for various ncAAs discussed in this review will allow the assignment of novel functionalities of the genetic code with a maximum degree of chemical liberty. Orthogonal pairs should be designed to serve as generalist tools so that ncAAs-mediated protein engineering will not only be relevant for single recombinant proteins, but also feasible throughout the entire *E. coli* proteome. The ground-breaking works of Dieter Söll on GluRS systems clearly highlight how codons can be emancipated and liberated from the current function [50]. In addition, the use of genome remodeling [43] will enable stable and valuable ncAA additions to the entire proteome of a cell. As the whole research area moves towards maturity, more and more approaches will contribute to solve industrially relevant bio-production problems, including advanced peptide and protein production.

Acknowledgements

We thank Dr Lars Merkel and Dr Patrick Durkin for critically reading this manuscript.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Liu CC, Schultz PG: **Adding new chemistries to the genetic code**. *Annu Rev Biochem* 2010, **79**:413-444.
•• Excellent, comprehensive review focusing on SCS methodology in both prokaryotic and eukaryotic cells. Lists all important non-canonical amino acids incorporated with *mj*TyrRS o-pairs and their applications.

2. Johnson JA, Lu YY, Van Deventer JA, Tirrell DA: **Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications**. *Curr Opin Chem Biol* 2010, **14**:774-780.
•• Excellently written review focusing on SPI methodology and its application for proteins and proteomes.

3. Antonczak AK, Morris J, Tippmann EM: **Advances in the mechanism and understanding of site-selective noncanonical amino acid incorporation**. *Curr Opin Struct Biol* 2011, **21**:481-487.

4. Fekner T, Li X, Chan MK: **Pyrrolysine analogs for translational incorporation into proteins**. *Eur J Org Chem* 2010:4171-4179.
• Review about ncAAs incorporated with the PylRS systems with clear focus on chemical coupling reactions.

5. Liu WSR, Wang YS, Wan W: **Synthesis of proteins with defined posttranslational modifications using the genetic noncanonical amino acid incorporation approach**. *Mol Biosyst* 2011, **7**:38-47.
• Good review summarizing all current possibilities of direct co-translational incorporation of post-translational modifications into recombinant proteins using SCS methodologies.

6. Cowie DB, Cohen GN: **Biosynthesis by *Escherichia coli* of active altered proteins containing selenium instead of sulfur**. *Biochim Biophys Acta* 1957, **26**:252-261.

7. Yoshikawa E, Fournier MJ, Mason TL, Tirrell DA: **Genetically engineered fluoropolymers. synthesis of repetitive polypeptides containing *p*-fluorophenylalanine residues**. *Macromolecules* 1994, **27**:5471-5475.

8. Budisa N, Steipe B, Demange P, Eckerskorn C, Kellermann J, Huber R: **High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli***. *Eur J Biochem* 1995, **230**:788-796.

9. Furter R: **Expansion of the genetic code: site-directed *p*-fluorophenylalanine incorporation in *Escherichia coli***. *Protein Sci* 1998, **7**:419-426.

10. Wang L, Brock A, Herberich B, Schultz PG: **Expanding the genetic code of *Escherichia coli***. *Science* 2001, **292**:498-500.

11. Hoesl MG, Acevedo-Rocha CG, Nehring S, Royter M, Wolschner C, Wiltschi B, Budisa N, Antranikian G: **Lipase congeners designed by genetic code engineering**. *ChemCatChem* 2011, **3**:213-221.

12. Steiner T, Hess P, Bae JH, Wiltschi B, Moroder L, Budisa N: **Synthetic Biology of proteins: tuning GFPs folding and stability with fluoroproline**. *PLoS ONE* 2008, **3**: Article no. e1680.

13. Crespo MD, Rubini M: **Rational design of protein stability: effect of (2*s*,4*R*)-4-fluoroproline on the stability and folding pathway of ubiquitin**. *PLoS ONE* 2011, **6**:e19425.

14. Wolschner C, Giese A, Kretzschmar HA, Huber R, Moroder L, Budisa N: **Design of anti- and pro-aggregation variants to assess the effects of methionine oxidation in human prion protein**. *Proc Natl Acad Sci U S A* 2009, **106**:7756-7761.

15. Brustad EM, Lemke EA, Schultz PG, Deniz AA: **A general and efficient method for the site-specific dual-labeling of proteins for single molecule fluorescence resonance energy transfer**. *J Am Chem Soc* 2008, **130**:17664-17665.

16. Hutchins BM, Kazane SA, Staflin K, Forsyth JS, Felding-Habermann B, Schultz PG, Smider VV: **Site-specific coupling and sterically controlled formation of multimeric antibody Fab fragments with unnatural amino acids**. *J Mol Biol* 2011, **406**:595-603.

17. Hughes RA, Ellington AD: **Rational design of an orthogonal tryptophanyl nonsense suppressor tRNA**. *Nucleic Acids Res* 2010, **38**:6813-6830.

18. Chin JW, Martin AB, King DS, Wang L, Schultz PG: **Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli***. *Proc Natl Acad Sci U S A* 2002, **99**:11020-11024.

19. Wang L, Zhang Z, Brock A, Schultz PG: **Addition of the keto functional group to the genetic code of *Escherichia coli***. *Proc Natl Acad Sci U S A* 2003, **100**:56-61.

20. Chin JW, Santoro SW, Martin AB, King DS, Wang L, Schultz PG: **Addition of *p*-azido-L-phenylaianine to the genetic code of *Escherichia coli***. *J Am Chem Soc* 2002, **124**:9026-9027.

21. Polycarpo CR, Herring S, Berube A, Wood JL, Söll D, Ambrogelly A: **Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase**. *FEBS Lett* 2006, **580**:6695-6700.

22. Wang Y-S, Russell WK, Wang Z, Wan W, Dodd LE, Pai P-J, Russell DH, Liu WR: **The de novo engineering of pyrrolysyl-tRNA synthetase for genetic incorporation of L-phenylalanine and its derivatives**. *Mol Biosyst* 2011, **7**:714-717.

23. Takimoto JK, Dellas N, Noel JP, Wang L: **Stereochemical basis for engineered pyrrolysyl-tRNA synthetase and the efficient in vivo incorporation of structurally divergent non-native amino acids**. *ACS Chem Biol* 2011, **6**:733-743.

24. Wan W, Huang Y, Wang Z, Russell William K, Pai P-J, Russell David H, Liu WR: **A facile system for genetic incorporation of two different noncanonical amino acids into one protein in *Escherichia coli***. *Angew Chem Int Ed Engl* 2010, **49**:3211-3214.

25. Young DD, Young TS, Jahnz M, Ahmad I, Spraggon G, Schultz PG: **An evolved aminoacyl-tRNA synthetase with atypical polysubstrate specificity**. *Biochemistry* 2011, **50**:1894-1900.

26. Miyake-Stoner SJ, Refakis CA, Hammill JT, Lusic H, Hazen JL, Deiters A, Mehl RA: **Generating permissive site-specific unnatural aminoacyl-tRNA synthetases**. *Biochemistry* 2010, **49**:1667-1677.

27. Beck-Sickinger AG, Budisa N: **Genetically encoded photocrosslinkers as molecular probes to study g-protein-coupled receptors (GPCRs)**. *Angew Chem Int Ed Engl* 2011 doi: 10.1002/anie.201107211. [epub].

28. Anderson JC, Wu N, Santoro SW, Lakshman V, King DS, Schultz PG: **An expanded genetic code with a functional quadruplet codon**. *Proc Natl Acad Sci U S A* 2004, **101**:7566-7571.

29. • Hoesl MG, Budisa N: **In vivo incorporation of multiple noncanonical amino acids into proteins**. *Angew Chem Int Ed Engl* 2011, **50**:2896-2902.
A detailed, critical review discussing the current status of multiple parallel *in vivo* incorporations of noncanonical amino acids only briefly touched in the present review.

30. Neumann H, Wang K, Davis L, Garcia-Alai M, Chin JW: **Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome**. *Nature* 2010, **464**:441-444.

31. Hayashi G, Goto Y, Suga H: **Ribosome evolution for two artificial amino acids in *E. coli***. *Chem Biol* 2010, **17**:320-321.

32. • Lepthien S, Merkel L, Budisa N: **In vivo double and triple labeling of proteins using synthetic amino acids**. *Angew Chem Int Ed Engl* 2010, **49**:5446-5450.
This study reports for the first time the parallel incorporation of three different non-canonical amino acids and clearly points out the benefits of this approach for protein engineering.

33. Merkel L, Schauer M, Antranikian G, Budisa N: **Parallel incorporation of different fluorinated amino acids: on the way to teflon proteins**. *ChemBioChem* 2010, **11**:1505-1507.

34. Hoesl MG, Budisa N: **Expanding and engineering the genetic code in a single expression experiment**. *ChemBioChem* 2011, **12**:552-555.

35. Ayyadurai N, Deepankumar K, Prabhu NS, Lee S, Yun H: **A facile and efficient method for the incorporation of multiple unnatural amino acids into a single protein**. *Chem Commun* 2011, **47**:3430-3432.

36. Ryu YH, Schultz PG: **Efficient incorporation of unnatural amino acids into proteins in *Escherichia coli***. *Nat Methods* 2006, **3**:263-265.

37. Guo J, Charles EM III, Lee HS, Groff D, Schultz Peter G: **Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing nonnatural amino acids**. *Angew Chem Int Ed Engl* 2009, **48**:9148-9151.

38. Young TS, Ahmad I, Yin JA, Schultz PG: **An enhanced system for unnatural amino acid mutagenesis in *E. coli***. *J Mol Biol* 2010, **395**:361-374.

39. Cellitti SE, Jones DH, Lagpacan L, Hao XS, Zhang Q, Hu HY, Brittain SM, Brinker A, Caldwell J, Bursulaya B *et al.*: **In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy**. *J Am Chem Soc* 2008, **130**:9268-9281.

40. Huang Y, Russell WK, Wan W, Pai PJ, Russell DH, Liu WS: **A convenient method for genetic incorporation of multiple noncanonical amino acids into one protein in *Escherichia coli***. *Mol Biosyst* 2010, **6**:683-686.

41. Wang KH, Neumann H, Peak-Chew SY, Chin JW: **Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion**. *Nat Biotechnol* 2007, **25**:770-777.

42. Ryden SM, Isaksson LA: **A temperature-sensitive mutant of *Escherichia coli* that shows enhanced misreading of UAG/A and increased efficiency for some tRNA nonsense suppressors**. *Mol Gen Genet* 1984, **193**:38-45.

43. Isaacs FJ, Carr PA, Wang HH, Lajoie MJ, Sterling B, Kraal L, Tolonen AC, Gianoulis TA, Goodman DB, Reppas NB *et al.*: **Precise manipulation of chromosomes in vivo enables genome-wide codon replacement**. *Science* 2011, **333**:348-353.

44. • Mukai T, Hayashi A, Iraha F, Sato A, Ohtake K, Yokoyama S, Sakamoto K: **Codon reassignment in the *Escherichia coli* genetic code**. *Nucleic Acids Res* 2010, **38**:8188-8195.
The first report of a strategy to knock out release factor 1 for amber stop codon emancipitation in *E. coli*.

45. • Johnson DBF, Xu J, Shen Z, Takimoto JK, Schultz MD, Schmitz RJ, Xiang Z, Ecker JR, Briggs SP, Wang L: **RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites**. *Nat Chem Biol* 2011, **7**:779-786.
The second, very elegant strategy for release factor 1 knock out by 'fixing' release factor 2 in *E. coli*.

46. Moore SD, Sauer RT: **The tmRNA system for translational surveillance and ribosome rescue**. *Annu Rev Biochem* 2007, **76**:101-124.

47. Mukai T, Yanagisawa T, Ohtake K, Wakamori M, Adachi J, Hino N, Sato A, Kobayashi T, Hayashi A, Shirouzu M *et al.*: **Genetic-code evolution for protein synthesis with non-natural amino acids**. *Biochem Biophys Res Commun* 2011, **411**:757-761.

48. Kolisnychenko V, Plunkett G, Herring CD, Feher T, Posfai J, Blattner FR, Posfai G: **Engineering a reduced *Escherichia coli* genome**. *Genome Res* 2002, **12**:640-647.

49. Chen IA, Schindlinger M: **Quadruplet codons: one small step for a ribosome, one giant leap for proteins**. *Bioessays* 2010, **32**:650-654.

50. O'Donoghue P, Sheppard K, Nureki O, Söll D: **Rational design of an evolutionary precursor of glutaminyl-tRNA synthetase**. *Proc Natl Acad Sci U S A* 2011.

51. Yuan J, Gogakos T, Babina AM, Söll D, Randau L: **Change of tRNA identity leads to a divergent orthogonal histidyl-tRNA synthetase/tRNAHis pair**. *Nucleic Acids Res* 2011, **39**:2286-2293.

52. • Neumann H, Slusarczyk AL, Chin JW: **De novo generation of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs**. *J Am Chem Soc* 2010, **132**:2142-2144.
The paper describes the successful implementation of an engineering and selection approach to design two mutually orthogonal aaRS:tRNA pairs from one parent o-pair.

53. Doi Y, Ohtsuki T, Shimizu Y, Ueda T, Sisido M: **Elongation factor tu mutants expand amino acid tolerance of protein biosynthesis system**. *J Am Chem Soc* 2007, **129**:14458-14462.

54. •• Park H-S, Hohn MJ, Umehara T, Guo L-T, Osborne EM, Benner J, Noren CJ, Rinehart J, Soell D: **Expanding the genetic code of *Escherichia coli* with phosphoserine**. *Science* 2011, **333**:1151-1154.
One of the most important recent studies in the field. It beautifully applies a very holistic approach for the incorporation of the industrially and academically highly relevant *O*-phosphoserine. Moreover, the paper

highlights the importance of a system perspective in reprogramming protein translation.

55. Yoshizawa S, Böck A: **The many levels of control on bacterial selenoprotein synthesis**. *Biochim Biophys Acta* 2009, **1790**:1404-1414.

56. •• Ou W, Uno T, Chiu H-P, Grünewald J, Cellitti SE, Crossgrove T, Hao X, Fan Q, Quinn LL, Patterson P *et al.*: **Site-specific protein modifications through pyrroline-carboxy-lysine residues**. *Proc Natl Acad Sci U S A* 2011, **108**:10437-10442.
The paper nicely combines metabolic engineering for *in situ* production of pyrroline-carboxy-lysine from D-ornithine and the later exploitation of its side chain for site-specific coupling reactions. This forward-looking study clearly demonstrates that metabolic engineering must be an inevitable part of genetic code engineering.

57. Link AJ, Mock ML, Tirrell DA: **Non-canonical amino acids in protein engineering**. *Curr Opin Biotechnol* 2003, **14**:603-609.

58. Minks C, Huber R, Moroder L, Budisa N: **Atomic mutations at the single tryptophan residue of human recombinant annexin V: effects on structure, stability, and activity**. *Biochemistry* 1999, **38**:10649-10659.

59. Cirino PC, Tang Y, Takahashi K, Tirrell DA, Arnold FH: **Global incorporation of norleucine in place of methionine in cytochrome P450 BM-3 heme domain increases peroxygenase activity**. *Biotechnol Bioeng* 2003, **83**:729-734.

60. Merkel L, Beckmann HSG, Wittmann V, Budisa N: **Efficient N-terminal glycoconjugation of proteins by the N-end rule**. *ChemBioChem* 2008, **9**:1220-1224.

61. Dieterich DC, Hodas JJL, Gouzer G, Shadrin IY, Ngo JT, Triller A, Tirrell DA, Schuman EM: **In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons**. *Nat Neurosci* 2010, **13**: 897-U149.

62. Seo MH, Han J, Jin Z, Lee DW, Park HS, Kim HS: **Controlled and oriented immobilization of protein by site-specific incorporation of unnatural amino acid**. *Anal Chem* 2011, **83**:2841-2845.

63. Cho H, Daniel T, Buechler YJ, Litzinger DC, Maio ZW, Putnam AMH, Kraynov VS, Sim BC, Bussell S, Javahishvili T *et al.*: **Optimized clinical performance of growth hormone with an expanded genetic code**. *Proc Natl Acad Sci U S A* 2011, **108**:9060-9065.

64. Ye ZX, Bair M, Desai H, Williams GJ: **A photocrosslinking assay for reporting protein interactions in polyketide and fatty acid synthases**. *Mol Biosyst* 2011, **7**:3152-3156.

65. Xie J, Wang L, Wu N, Brock A, Spraggon G, Schultz PG: **The site-specific incorporation of *p*-iodo-L-phenylalanine into proteins for structure determination**. *Nat Biotechnol* 2004, **22**:1297-1301.

66. Wang L, Xie J, Deniz AA, Schultz PG: **Unnatural amino acid mutagenesis of green fluorescent protein**. *J Org Chem* 2002, **68**:174-176.

67. Kaya E, Gutsmiedl K, Vrabel M, Muller M, Thumbs P, Carell T: **Synthesis of threefold glycosylated proteins using click chemistry and genetically encoded unnatural amino acids**. *ChemBioChem* 2009, **10**:2858-2861.

68. Plass T, Milles S, Koehler C, Schultz C, Lemke EA: **Genetically encoded copper-free click chemistry**. *Angew Chem Int Ed Engl* 2011, **50**:3878-3881.

69. Huang Y, Wan W, Russell WK, Pai P-J, Wang Z, Russell DH, Liu W: **Genetic incorporation of an aliphatic keto-containing amino acid into proteins for their site-specific modifications**. *Bioorg Med Chem Lett* 2010, **20**:878-880.

70. Blight SK, Larue RC, Mahapatra A, Longstaff DG, Chang E, Zhao G, Kang PT, Church-Church KB, Chan MK, Krzycki JA: **Direct charging of tRNA(CUA) with pyrrolysine in vitro and in vivo**. *Nature* 2004, **431**:333-335.

71. Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama S: **Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode N-epsilon-(o-azidobenzyloxycarbonyl) lysine for site-specific protein modification**. *Chem Biol* 2008, **15**:1187-1197.

72. Nozawa K, O'Donoghue P, Gundllapalli S, Araiso Y, Ishitani R, Umehara T, Söll D, Nureki O: **Pyrrolysyl-tRNA synthetase-tRNA(Pyl) structure reveals the molecular basis of orthogonality**. *Nature* 2009, **457**: 1163-U1127.