Extremophiles (2012) 16:793–803
DOI 10.1007/s00792-012-0482-8

REVIEW

Translational recoding in archaea

Beatrice Cobucci-Ponzano · Mosè Rossi · Marco Moracci

Received: 21 June 2012 / Accepted: 9 September 2012 / Published online: 27 September 2012

Abstract Translational recoding includes a group of events occurring during gene translation, namely stop codon readthrough, programmed ±1 frameshifting, and ribosome bypassing, which have been found in organisms from all domains of life. They serve to regulate protein expression at translational level and represent a relatively less known exception to the traditional central 'dogma' of biology that information flows as DNA→RNA→protein and that it is stored in a co-linear way between the 5′→3′ of nucleic acids and N→C-terminal of polypeptides. In archaea, in which translational recoding regulates the decoding of the 21st and the 22nd amino acids selenocysteine and pyrrolysine, respectively, only one case of programmed −1 frameshifting has been reported so far and further examples, although promising, have not been confirmed yet. We here summarize the current state-of-the-art of this field that, especially in archaea, has relevant implications for the physiology of life in extreme environments and for the origin of life.

Keywords Programmed frameshifting · Stop codon readthrough · Gene expression · Hyperthermophiles · Pseudogenes

Abbreviations

Sec	Selenocysteine
SECIS	Selenocysteine inserting sequence
3′-UTR	3′ untranslated region
SBP2	SECIS-binding protein 2
Pyl	Pyrrolysine
MtmB, MtbB, and MttB, respectively	Mono-, di-, and trimethylamine methyltransferases
PYLIS	Pyrrolysine insertion element
Thg1	$tRNA^{His}$-guanylyltransferase
MALDIMS	Matrix-assisted laser desorption/ionization mass spectrometry
LCMSMS	Liquid chromatography online tandem mass spectrometry
fucA1	α-Fucosidase gene from *S. solfataricus*
RT-PCR	Reverse transcriptase
ICDS	Interrupted CoDing Sequence program

Introduction

The traditional central dogma of modern biology is that information flows as DNA→RNA→protein and that it is stored in a co-linear way between the 5′→3′ of nucleic acids and N→C-terminal of polypeptides. Nonetheless, the discovery of reverse transcriptase, non-standard genetic code in certain organelles and a small number of organisms, and RNA splicing and editing shattered these rules.

Another exception to the 'dogma' is translational recoding. This mechanism of regulation of protein expression at translational level includes several events occurring at the termination step (stop codon readthrough)

Communicated by S. Albers.

B. Cobucci-Ponzano · M. Rossi · M. Moracci (✉)
Institute of Protein Biochemistry, Consiglio Nazionale delle Ricerche, Via P. Castellino 111, 80131 Naples, Italy
e-mail: m.moracci@ibp.cnr.it

or during translation elongation (programmed ±1 frameshifting, ribosome hopping or bypassing). In stop codon readthrough a different meaning is temporarily assigned to a codon: usually, a stop codon becomes a sense codon for an amino acid different from the traditional 20 protein encoded. In programmed frameshifting, ribosomes switch to an alternative frame (±1) at a specific shift site, while in ribosome bypassing, ribosomes suspend translation at a certain site and then resume translation downstream (for reviews see Farabaugh 1996; Gesteland and Atkins 1996; Baranov et al. 2002; Namy et al. 2004; Dinman 2012).

Translational recoding has been found in organisms from all three domains, and it is particularly common in viruses (Farabaugh 1996; Baranov et al. 2002). In archaea, two different cases of stop codon readthrough are well documented with the 21st and the 22nd amino acids, selenocysteine and pyrrolysine, respectively, while only one case of programmed −1 frameshifting has been reported so far.

The reasons why translational recoding in archaea is not so commonly reported are many. Usually, bioinformatic inspection of data from the sequencing of novel genomes offers surveys of candidate genes that, being interrupted, can be potentially expressed by recoding (van Passel et al. 2007; Cobucci-Ponzano et al. 2010; Sharma et al. 2011; Guo et al. 2011). However, successively, the experimental demonstration that these events take place is less straightforward as it requires the isolation, characterization and sequencing of the protein products from cell extracts of the archaeon that, often, are available in tiny amounts. Finally, to understand the functional and physiological meaning of any new recoding event is even more complex, as it often requires molecular biology tools for recombinant expression that in several archaeal lineages are in their infancy and are not reliable as they are in bacteria and eukaryotes. Here, we summarize the state of the art on translational recoding in archaea in an effort to update the current knowledge of this interesting and relatively less famous mechanism of regulation of gene expression.

Selenocysteine

Since its discovery, selenium has been considered as a highly toxic element and only in 1954 its importance in the synthesis of enzymes involved in formate oxidation was recognized in *Escherichia coli* (Pinsent, 1954). Later, it was made clear that it is a trace element essential for many organisms, including humans (Kohrle et al. 2000; Arner and Holmgren 2006). The most important biological form of selenium is selenocysteine (Sec), an amino acid structurally identical to cysteine, but with the thiol group replaced by a selenol group. Sec was discovered as a unique amino acid in 1976 (Cone et al. 1976), and in 1986 it was found that it was co-translationally inserted into growing polypeptides (Chambers et al. 1986; Zinoni et al. 1986) and designated as the 21st genetically encoded amino acid (Bock et al. 1991). Incorporation of selenocysteine into proteins is specified in all three domains of life by dynamic translational redefinition of UGA codons. This occurs during translation elongation and is encoded by a combination of an internal UGA stop codon and a specific mRNA hairpin structure located further downstream, the selenocysteine inserting sequence (SECIS). The specificity of the stop codon is achieved by $tRNA^{Sec}$, whose UCA anticodon is complementary to the UGA stop codon.

In bacteria, selenocysteine synthesis occurs through the conversion of the seryl moiety on the seryl-$tRNA^{Sec}$ by a Sec synthase that employs a selenomonophosphate donor. In archaea, in which Sec synthesis and insertion follow similar strategies to eukaryotes, the conversion of Ser on the seryl-$tRNA^{Sec}$ molecule occurs by the concerted action of a phosphoseryl-$tRNA^{Sec}$ kinase and an *O*-phosphoseryl-$tRNA^{Sec}$-Sec synthase producing, respectively, a *O*-phosphoseryl- and a Sec-$tRNA^{Sec}$ (Fig. 1) (Carlson et al. 2004; Kaiser et al. 2005; Yuan et al. 2006).

In bacteria, the incorporation of Sec is mediated by a specialized translation elongation factor named SelB, which, carrying Sec-$tRNA^{Sec}$ and a GTP molecule, binds to the ribosome and SECIS, located in a position immediately adjacent to the UGA codon. A conformational change occurring in this complex promotes the insertion of Sec-$tRNA^{Sec}$ in the ribosome A-site and the translational recoding of the opal codon (Fig. 2a) (for a review see Stock and Rother 2009).

In eukaryotes and archaea (Fig. 2b, c, respectively), SECIS is located in the 3′ untranslated region (3′-UTR) of selenoprotein mRNA. As in bacteria, SelB (Rother et al. 2000) mediates the incorporation of Sec. However, while the bacterial SECIS is specifically bound by SelB, the archaeal and eukaryotic counterparts are not; in fact, the respective SelBs do not contain a C-terminal extension shown to be responsible for SECIS binding. Eukaryotic SECIS is bound by a SECIS-binding protein 2 (SBP2) and the ribosomal protein L30 (Copeland et al. 2000; Chavatte et al. 2005), but other factors have also been described (Fig. 2b) (Squires and Berry 2008; Allmang et al. 2009). So far, in archaea no SECIS-binding factors have been identified. Homologs of SBP2 are not encoded in any available archaeal genome and L30 homologs identified in *Methanococcus maripaludis* and *Methanocaldococcus jannaschii* appear not to function in selenoprotein synthesis (Rother and Krzycki 2010). Thus, it remains to be established how the SECIS element communicates with SelB in archaea (Fig. 2c).

Despite the increasing sequence data, comparative genome analyses, and studies on archaeal selenoprotein

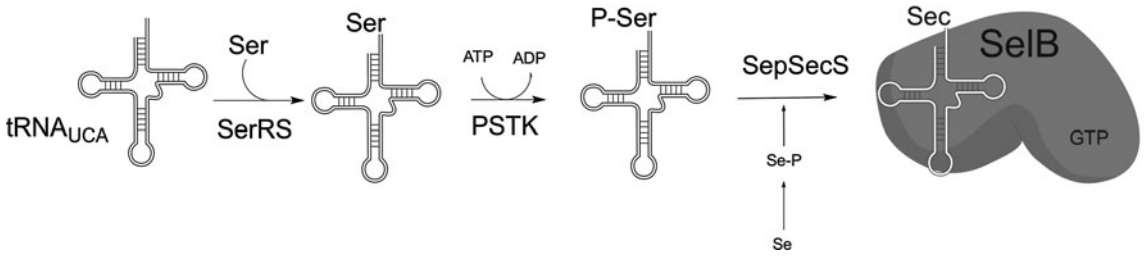

Fig. 1 Mechanism of selenocysteine synthesis in archaea. *Ser* serine, *SerRS* seryl-tRNA synthetase, *PSTK* seryl-tRNAsec kinase, *Se* reduced Se-species, *Se-P* seleno(mono)phosphate, *SepSecS* *O*-phosphoseryl-tRNAsec-Sec synthase, *SelB* Sec-specific elongation factor. See the text for details

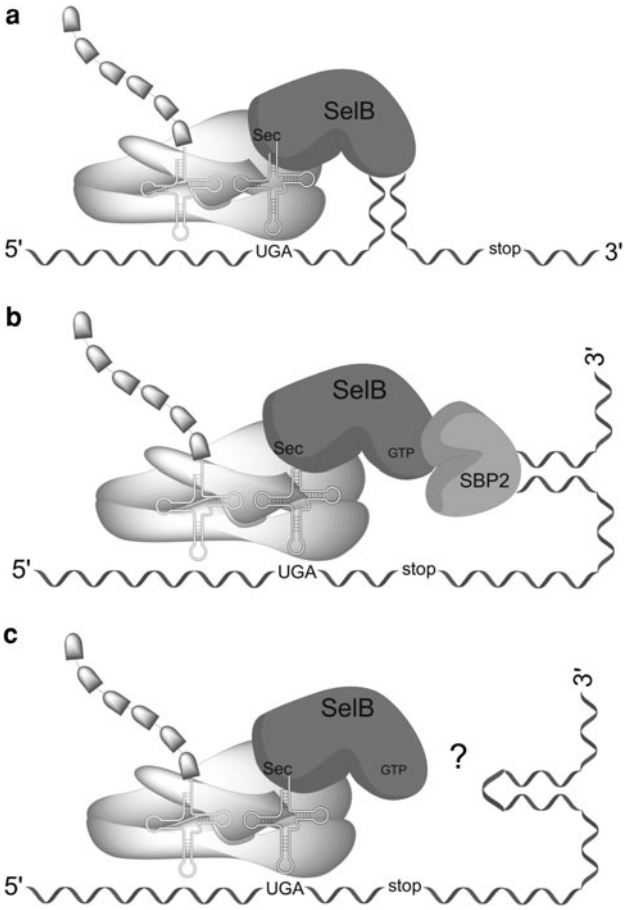

Fig. 2 Selenocysteine decoding in the three living domains. Comparison of Sec mechanism of insertion in bacteria (**a**), eukaryotes (**b**), and archaea (**c**). See the text for details

synthesis, there is still considerable missing information on the biochemical and physiological function of selenoproteins and the evolution of their synthesis. Most organisms for which genome information is available do not employ Sec and why some organisms use Sec is still not understood, mostly because the specific function of Sec is still unknown. The use of Sec might partly be explained by its high nucleophilicity and the fact that the selenol group is mostly deprotonated at physiological pH making it more reactive than Cys (Zinoni et al. 1987; Axley et al. 1991). Therefore, it is almost exclusively found in the catalytic site of numerous redox-active enzymes. In archaea, the presence of selenoproteins appears to be restricted to two genera, *Methanococcus* and *Methanopyrus* methanogens (Kryukov and Gladyshev 2004; Rother et al. 2001), obligatory dependent on the hydrogenotropic pathway of methanogenesis. Interestingly, the verified selenoenzymes of archaea, with the exception of selenophosphate synthetase, are all involved in methanogenesis (Table 1).

Selenoproteins are present in members of all three domains of life with some fundamental features conserved suggesting that the utilization of Sec emerged before the division of the three domains and must have been present in the last universal common ancestor (Yuan et al. 2006). Subsequently, the ability to use Sec was possibly lost in many lineages during evolution and some lineages later regained it through lateral gene transfer (Zhang et al. 2006). It is suggested that in archaea, the utilization of Sec is a trait that was lost and those still synthesizing selenoproteins might just thrive under conditions like permanent absence of oxygen and low reductant concentration, which are not selected against this trait, explaining why some methanogens use Sec and others do not (Rother and Krzycki 2010).

Pyrrolysine

The first hint of the presence of pyrrolysine (Pyl) in archaea has been reported about a decade ago by the group of Krycki who identified in several *Methanosarcina* species a total of 21 genes of mono-, di-, and trimethylamine methyltransferases (MtmB, MtbB, and MttB, respectively) showing an in-frame amber TAG codon (James et al. 2001). The amino acid decoding the stop codon was initially identified as Lys, but, successively, the 3D-structure of the MtmB enzyme and the finding that a novel tRNAPyl decoded the UAG codon allowed the unequivocal identification of Pyl and demonstrated that the process occurred by translational recoding (Srinivasan et al. 2002; Hao et al. 2002). Since then, important progresses have been made in elucidating the biosynthetic pathway of Pyl, its mechanism

Table 1 Selenoproteins of archaea

Selenoprotein	Annotated gene	Organism	Reference
Formate dehydrogenase	Mevan_0608 (FdhA)	*Methanococcus vannielii*	Jones and Stedman (1981); Jones et al. (1979)
Formyl-methanofuran dehydrogenase	MK1527 (FwuB)	*Methanopyrus kandleri*	Vorholt et al. (1997)
F420-reducing hydrogenase	Mvol_0991 (FruA)	*Methanococcus voltae*	Halboth and Klein (1992)
F420-nonreducing hydrogenase	Mvol_1264 (VhuD), Mvol_1267 (VhuU)	*Methanococcus voltae*	Halboth and Klein (1992); Sorgenfrei et al. (1993)
Heterodisulfide reductase	MJ1190m (HdrA)	*Methanocaldococcus jannaschii*	Wilting et al. (1997)
Selenophosphate synthetase	MMP0904	*Methanococcus maripaludis*	Wilting et al. (1997); Stock et al. (2010)
HesB-like protein	MMP0252	*Methanococcus maripaludis*	Stock et al. (2010); Kryukov and Gladyshev (2004)

Fig. 3 Genomic environment. *M. acetivorans* pyl *genes*

of translational insertion in the nascent protein, the function of this amino acid in the proteins translationally recoded, and the distribution of Pyl in other organisms. These results substantially modified the conclusions of early works and our knowledge of this process (Namy et al. 2004; Cobucci-Ponzano et al. 2005b).

After the early identification of pyrrolysine in *Methanosarcina* species and in the bacterium *Desulfitobacterium hafniense* (Srinivasan et al. 2002), the annotation of other genomes revealed the presence of the *pyl* genes in archaeal and bacterial species. Among archaea, *pyl* genes were identified in anaerobic methanogenes living in environments where methylamines are available, namely, several Methanosarcinales (*M. barkeri*, *M. thermophila*, *M. acetivorans*, *M. mazei*) (Galagan et al. 2002; Deppenmeier et al. 2002; Maeder et al. 2006), in *Methanococcus burtonii* (psychrophile) (Goodchild et al. 2004), and in *Methanoalophilus mahii* and *Methanohalobium evestigatum* (halophiles) (Rother and Krzycki 2010; Gaston et al. 2011a). The description of the bacterial groups containing *pyl* genes goes beyond the aims of this review and details can be found in (Gaston et al. 2011a) and reference therein; however, it is worth mentioning that these organisms include also pathogens of the human gut.

Pyrrolysine is synthesized and incorporated into methylamine methyltransferases through the combined action of the product of the *pyl* genes, which, as shown in Fig.-Ã3 in *M. acetivorans*, are usually clustered near the genes encoding the methylamine methyltransferases and other genes involved in methylamine metabolism (for a more detailed description of the genomic contexts of Pyl-related genes see also below and (Gaston et al. 2011a)). The *pylT* and *pylS* genes encode, respectively, the $tRNA^{Pyl}$ and the Pyl-tRNA synthetase (Blight et al. 2004; Polycarpo et al. 2004), while the *pylBCD* gene products are involved in the synthesis of this amino acid (Longstaff et al. 2007b) (Fig.-Ã3). This was demonstrated by transforming *E. coli* with the archaeal *pylTSBCD* gene cluster, which promoted the recoding of the UAG in the *mtmB1* gene of *M. acetivorans* and in the reporter *uidA* gene from *E. coli*, encoding a +Î-glucuronidase engineered with an in-frame amber codon (Longstaff et al. 2007b). More recently, the same authors unequivocally demonstrated by stable isotope labeling experiments that Pyl is synthesized from two lysines, ruling out that ornithine, Glu, Pro, or Ile were the precursors of the methylated pyrroline ring as previously believed (Krzycki 2004; Longstaff et al. 2007b; Ambrogelly et al. 2007). In addition, using the *E. coli* recombinant *pyl* system, they also dissected elegantly the function of the PylB, C and D proteins (Gaston et al. 2011b). In particular, the addition of D-ornithine to cells transformed with *pylCD* produced D-ornithyl-$^{+\hat{I}}$*N*-Lys that was subsequently oxidized by PylD and spontaneously eliminated water leading to desmethylpyrrolysine. Interestingly, under these conditions in which *pylB* was lacking, full-length MtmB could be synthesized only if D-ornithine was externally added, but desmethylpyrrolysine and not pyrrolysine was found in MtmB (in Fig.-Ã4 desmethylpyrrolysine resembles compound ***3***, but lacks the methyl group on the pyrroline ring), indicating that PylB produced a derivative of D-ornithine (Gaston et al. 2011b). The authors have proposed that PylB is a lysine mutase, which, exploiting *S*-adenosyl-L-methionine and lysine, catalyzes the synthesis of (*R*,*R*)-3-methylornithine (compound ***2*** in Fig.-Ã4) as the first step of Pyl biosynthesis.

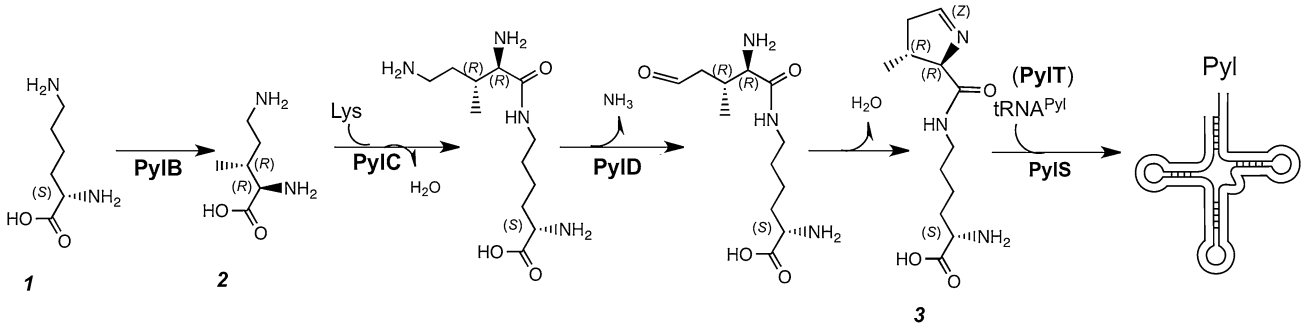

Fig. 4 Mechanism of pyrrolysine synthesis. Lysine (*1*), (2*R*,3*R*)-3-methylornithine (*2*), pyrrolysine (*3*)

These data demonstrated that pyrrolysine, unlike selenocysteine resulting from the modification of a serine already bound to the $tRNA^{Sec}$ (see above), is charged onto a $tRNA^{Pyl}$ after its biosynthesis. The pyrrolysyl-tRNA synthetase catalyzing this reaction, PylS, is a homodimer, showing N- and C-terminal domains with different functions. The C-terminal domain is homolog to class II aminoacyl-tRNA synthetases, while the N-terminal shows no homology with proteins of known function (Srinivasan et al. 2002). It is worth mentioning that in bacteria these two domains are encoded by two independent genes, named *pylSn* and *pylSc* (Krzycki 2005; Gaston et al. 2011a). The description of the mechanism of recognition between PylS and $tRNA^{Pyl}$ (Yanagisawa et al. 2008a; Nozawa et al. 2009) is well described in a recent excellent review (Gaston et al. 2011a). Regarding the mechanism of Pyl insertion in the translationally recoded protein, it is worth mentioning here that convincing evidences of the existence of a pyrrolysine insertion element (PYLIS, a sequence forming a secondary structure adjacent to the recoded UAG) and of specific translation factors previously suggested (Namy et al. 2004; Cobucci-Ponzano et al. 2005b; Longstaff et al. 2007a; Alkalaeva et al. 2009; Atkinson et al. 2011) have not been confirmed. In fact, a detailed bioinformatic analysis showed that PYLIS is often absent in organisms using pyrrolysine (Zhang et al. 2005) and, as mentioned above, the simple insertion of an in-frame amber codon in the *E. coli uidA* gene allowed the translation of a full-length product with 20–30 % efficiency in *M. acetivorans* (Longstaff et al. 2007a). This is confirmed to be true also with other foreign genes in heterologous cells carrying the *pyl* system (see below) and, in addition, substitution of sequences downstream of the UAG codon, although dramatically increased the UAG-termination, produced still high levels of Pyl incorporation in full-length products (Gaston et al. 2011a). Therefore, again unlike Sec, signaling sequences that might be involved in the efficient translation of the amber codon in Pyl have not been identified.

The studies reported suggest that $tRNA^{Pyl}$ acts as a suppressor tRNA on UAG codon. The ability of the *pyl* system to promote the efficient incorporation of pyrrolysine in reporter genes interrupted by an amber codon has been observed in *E. coli*, yeasts (Hancock et al. 2010) and mammalian cells (Gautier et al. 2011). Remarkably, the orthogonal pair PylS/$tRNA^{Pyl}$ and the site-directed engineering of the amino acid binding site of PylS have been exploited as a general method to insert novel and unnatural amino acids (Neumann et al. 2008) that found application for fluorescent tagging (Yanagisawa et al. 2008b), FRET studies (Fekner et al. 2009), ubiquitination (Li et al. 2009), and control of transport into organelles (Gautier et al. 2011).

The function of Pyl in the methylamine methyltransferase reaction has not been unraveled yet. It is usually ascribed to the peculiar electrophilic nature of this amino acid, which has unique characteristics compared to the other genetically encoded amino acids, suggesting that it might participate in corrinoid-dependent methylamine methyltransferase reactions by interacting with either the methylamine substrate or product (Hao et al. 2002, 2004). If confirmed, this very specific function would explain why *pyl* system and methylamine methyltransferases are strictly linked. In fact, noticeably, each organism possessing the five *pyl* genes also has one or more methylamine methyltransferases interrupted by UAG and other genes involved in methylamine metabolism. Conversely, other genomes, including some from *Methanosarcinaceae,* which lack the *pylTSBCD* genes, did not contain interrupted methylamine methyltransferases (Heinemann et al. 2009; Gaston et al. 2011a). The function of these enzymes as true methylamine methyltransferases has not been proved yet and the substitution of pyrrolysine with other amino acids containing small or bulky side chains suggests for this residue a non-catalytic role (Rother and Krzycki 2010). These observations would suggest that the *pyl* system evolved under a selective pressure to include an amino acid with a novel function giving metabolic advantages; however, the

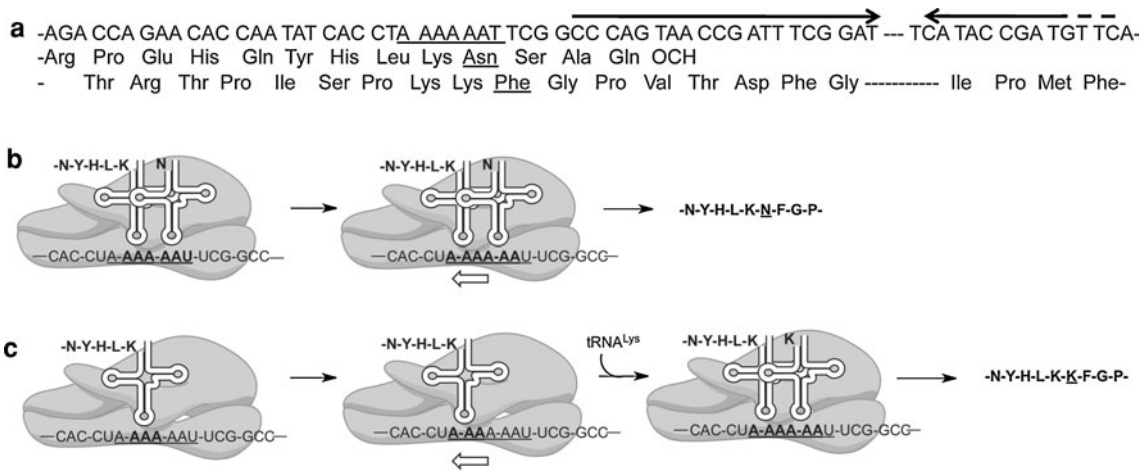

Fig. 5 Programmed −1 frameshifting of the *fucA1* gene in *S. solfataricus*. Closeup of the zone of overlapping between the two ORFs of *fucA1*: the slippery sequence and the amino acids in which the frameshifting occurs are *underlined*; the putative stem loop is indicated by *arrows* (**a**). Comparison of the experimentally demonstrated mechanisms of programmed −1 frameshifting (see the text for details) (**b**, **c**)

identification of pyrrolysine in a $tRNA^{His}$-guanylyltransferase (Thg1) in *M. acetivorans* contradicted this hypothesis. This gene, whose product is unrelated to methylamine metabolism as it post-transcriptionally adds a guanosine to the 5′-terminus of $tRNA^{His}$, contains an in-frame TAG codon and leads to a full-length enzyme containing pyrrolysine. Interestingly, UAG is not critical, in fact Pyl or Trp insertions produce full-length fully functional Thg1 in equal amounts (Heinemann et al. 2009). From these studies, the authors conclude that Pyl appeared in Tgh1 from *M. acetivorans* as the result of neutral evolution and predict that many other Pyl-proteins may be present in species displaying the *pyl* genes.

Programmed frameshifting

The only other kind of translational recoding identified in archaea is programmed frameshifting. Preliminary studies in the thermoacidophilic archaeon *Sulfolobus solfataricus* strain P2 demonstrated that the *fucA1* gene was formed by the ORFs SSO11867 and SSO3060 encoding for the N- and C-termini, respectively, of a α-fucosidase. These ORFs, actively transcribed in vivo, were separated by a −1 frameshifting and showed the typical sequences regulating programmed −1 frameshifting in *cis*, namely a *slippery* sequence (A-AAA-AAT) flanked by rare codons and a putative stem loop at the 5′ and 3′, respectively (Fig. 5a) (Cobucci-Ponzano et al. 2003a).

Interestingly, merging by site-directed mutagenesis, the two ORFs of *fucA1* gene in a novel full-length gene mimicking the frameshifting event produced a full-length α-fucosidase that was fully functional and has been characterized in detail (Cobucci-Ponzano et al. 2003a; Cobucci-Ponzano et al. 2003b; Rosano et al. 2004; Cobucci-Ponzano et al. 2005a; Cobucci-Ponzano et al. 2009). These studies did not say anything on whether this event occurred in vivo. More recently, the observation that the *fucA1* frameshifted gene directed the expression of low α-fucosidase activity in *E. coli*, led to the isolation and characterization of the polypeptides expressed in recombinant form. Experiments of matrix-assisted laser desorption/ionization mass spectrometry (MALDI–MS) and liquid chromatography online tandem mass spectrometry (LC–MS–MS) demonstrated that the wild-type *fucA1* in *E. coli* produced a mixture of two full-length polypeptides, both functional, with a total

Table 2 Programmed −1 frameshifting of *fucA1* alleles in *E. coli* and *S. solfataricus*

fucA1 alleles		Programmed −1 frameshifting products[a] and efficiency	
fucA1 nomenclature	Slippery heptamer (0 frame)	*E. coli*	*S. solfataricus*
Wild type (*fucA1*)	A-AAA-AAT	-Leu-Lys-Asn-Phe-Gly-[b] -Leu-Lys-Lys-Phe-Gly- 5 %	Unknown 10 %
Single mutant (*fucA1*^{sm})	A-AAG-AAT	-Leu-Lys-Glu-Phe-Gly- 10 %	–
Triple mutant (*fucA1*^{tm})	C-AAG-AAC	–	–

Data from Cobucci-Ponzano et al. (2006)

[a] Referred to the peptides identified by MALDI, LC–MS–MS, and in vitro translation

[b] The sequence reported is that of the region in which the programmed −1 frameshifting occurred (see also Fig. 5)

efficiency of about 5 %, a value even higher than that previously found in other genes expressed by programmed −1 frameshifting in heterologous expression systems (Table 2) (Xu et al. 2004; Cobucci-Ponzano et al. 2006). In addition, the introduction of a single mutation (*fucA1*^{sm}), which, although maintaining the frameshift, modified the heptameric sequence and surprisingly increased the frameshifting efficiency to about 10 %. Instead, a triple mutant allele (*fucA1*^{tm}), in which the slippery sequence was completely disrupted, did not produce any full-length α-fucosidase (Table 2) (Cobucci-Ponzano et al. 2006). These data demonstrated that in *E. coli*, the integrity of the A-AAA-AAT heptamer of *fucA1* was essential for programmed −1 frameshifting expression and that the efficiency of this process was even increased with the *fucA1*^{sm} allele because it contained the sequence A-AAG that is identical to that previously found for the programmed −1 frameshifting of a transposase gene (Sekine et al. 1994).

The analysis of recoding of *fucA1* in *S. solfataricus* was performed by in vitro translation. In this case, only the wild-type allele led to a full-length product by programmed −1 frameshifting with good efficiency (about 10 % in Table 2), demonstrating that this process occurred in archaea. The mutated alleles did not produce any detectable translation, indicating that any modification of the slippery sequence was detrimental to the translational recoding event (Cobucci-Ponzano et al. 2006). Full-length polypeptides from *fucA1* in *S. solfataricus* extracts were identified by Western blots with antibodies raised against recombinant α-fucosidase, while specific assays demonstrated that this enzyme was expressed at very low level. This was confirmed by reverse transcriptase (RT-PCR) and real-time PCR experiments demonstrating that *fucA1* was expressed at very low levels at the conditions tested (Cobucci-Ponzano et al. 2006).

The sequence of the recoded polypeptide in *S. solfataricus* could not be determined; however, the sum of these experiments allowed us to propose the models summarized in Fig. 5b and c. The polypeptides identified by LC–MS–MS in the *E. coli* expressing wild-type *fucA1* indicated that the translational recoding of the interrupted gene containing the A-AAA-AAT sequence might occur in two ways. The sequence -Leu-Lys-Asn-Phe-Gly- in Table 2 results from a simultaneous backward slippage of the ribosome when both the P- and the A-site tRNAs are occupied (Fig. 5b). Instead, the other polypeptide, in which a Lys replaced Asn, can result only from the repositioning of the ribosome in the −1 frame when only the P-site tRNA is bound (Fig. 5c). In the former mechanism, as a result of the backward shifting, the A-site tRNA decoding AAU→Asn recognizes the codon with a mismatch on the third base. Instead, the shifting of the ribosome carrying only the P-site tRNA (Fig. 5c) allows access to the A-site of a tRNA decoding AAA→Lys codon. It is worth mentioning that MALDI and LC–MS–MS experiments demonstrated that the *fucA1*^{sm} allele followed only the mechanism described in Fig. 5c. If the ribosome would perform frameshifting as described in Fig. 5b, Lys would be found in any case, since both AAA and AAG encode for this amino acid. Instead, no such sequence was identified, but only Glu resulting from the AAA→GAA mutation in the −1 frame. The lack of expression of the *fucA1*^{sm} allele by translation in vitro in *S. solfataricus* (Table 2) contrasts with its efficient decoding in *E. coli*, suggesting that the two organisms recognize different slippery sequences.

These studies, although demonstrating for the first time that programmed −1 frameshifting is present in archaea, leaves several questions unanswered: to date, we still do not know whether, how, and why the translational recoding of *fucA1* in *S. solfataricus* is a regulated process and if other genes are expressed by translational recoding in this or other archaea. The regulation of the expression of the α-fucosidase gene is currently the object of further studies; however, recent genomic sequencing showed that the *fucA1* gene was present also in other archaea. Interestingly, all the species showing this gene belong to Crenarchaeota, including *Ignisphaera aggregans*, *Caldivirga maquilingensis*, eight strains of *Sulfolobus islandicus*, and *S. solfataricus* strain 98/2 (for the compilation of these genes, see the Carbohydrate Active Enzyme database http://www.cazy.org/). Sulfolobales α-fucosidases, which showed >96 % amino acid sequence identity, were all full length with the exception of the *S. solfataricus*, strain 98/2 presenting the frameshifting in the same position of the gene from strain P2. However, all Sulfolobales genes presented 100 % DNA sequence identity in the region of frameshifting, maintaining the rare codon, the slippery sequence (in which the stretch of A is shortened by one nucleotide in full-length genes) and the putative stem loop. Instead, α-fucosidases from Sulfolobales showed identities >60 % and between 45 and 60 % at DNA and protein level to their homologs from *I. aggregans* and *C. maquilingensis*, in which the slippery sequence is not conserved. These observations indicate that α-fucosidase is present in different strains of archaea and that in phylogenetically close species its gene is either full length or frameshifted, suggesting that the latter status might be maintained for functional/physiological reasons. Remarkably, full-length α-fucosidases, in the corresponding region, show the same Lys or Asn amino acids observed in the full-length product of the wild-type interrupted *fucA1* (Cobucci-Ponzano et al. 2006).

Other examples of translational recoding?

The identification of novel cases of genes expressed by translational recoding is not easy, either because of the

technical reasons explained in "Introduction", or because disrupted genes are commonly considered nonfunctional pseudogenes, i.e., disabled copies or decayed remnants of genes that do not merit further consideration. Nonfunctional pseudogenes certainly are present in organisms from all the living domains; however, in some cases they have been demonstrated to be functional or at least useful for an organism's survival and adaptation to particular environmental changes (Harrison and Gerstein 2002; Balakirev and Ayala 2003; Hirotsune et al. 2003). In archaea, a bioinformatic analysis on 15 different species revealed a surprisingly high number of predicted pseudogenes present in the highest fraction in *S. solfataricus* as 8.6 % of the total of annotated protein coding sequences (van Passel et al. 2007). Tests on whether or not these predicted pseudogenes were functional were not reported in that study; however, remarkably, in archaea all the frameshifts occurred in A/T rich tracts, sequences that resemble the slippery sequences regulating programmed −1 frameshifting in *cis* (van Passel et al. 2007).

More recently, a bioinformatic analysis of the genomes of 16 archaea, supported by a high throughput proteomic analysis and functional characterization specifically in *S. solfataricus* strain P2, allowed us to identify a large stock of disrupted genes (Cobucci-Ponzano et al. 2010). They included genes containing insertion/deletion of single/multiple base pairs, truncated ORFs, and genes separated by ±1 frameshifts, stop codons, and fragments of different length. In *S. solfataricus*, a manual search supported by the Interrupted CoDing Sequence (ICDS) program (Perrodou et al. 2006) led to 34 entries in which 10 were protein-encoding genes interrupted by ±1 frameshifts and stop codons. Interestingly, the position of the disruption was conserved in homologs from different archaea, indicating that the interruptions did not occur randomly as it would be expected for inactive pseudogenes, but, rather, that these genes were selectively maintained in their interrupted status for functional reasons (Cobucci-Ponzano et al. 2010).

Experimental proof that the interrupted genes in *S. solfataricus* were expressed in vivo was found by RT-PCR and shotgun proteomic analysis. Among the genes tested, a putative universal translation initiation factor SUI-1 was particularly intriguing. This protein is essential in yeast and its gene is present in all sequenced genomes of archaea. It forms the translation initiation complex, monitors the maintenance of the correct translational reading frame in eukaryotes, and it was proposed that it might govern programmed −1 frameshifting as a *trans*-acting factor (Kyrpides and Woese 1998; Cui et al. 1998).

In the genome of *S. solfataricus*, *sui-1* is interrupted by a −1 frameshift (She et al. 2001), but the re-sequencing of the gene led to a full-length ORF, indicating that the observed interruption was the consequence of a sequencing error (Cobucci-Ponzano et al. 2010). However, surprisingly, the high-throughput proteomic and LC–MS–MS analyses in protein extracts from this archaeon allowed the unambiguous identification of two polypeptides resulting either from a full-length *sui-1* allele or from an interrupted version, which might be expressed by programmed −1 frameshifting. In fact, *sui-1* gene showed also the sequences putatively regulating the process in *cis* (Cobucci-Ponzano et al. 2010). The s*ui-1* gene is present in a single copy in the genome, and the DNA sequencing of different clones of the gene taken randomly demonstrated that *sui-1* tends to accumulate mutations. Therefore, the conundrum of the two polypeptides found was explained speculating that the sample used for the shotgun proteomic analysis contained a mixed population of *S. solfataricus* cells showing the two alleles of *sui-1*. Possibly, since the gene accumulated mutations for unknown reasons, the programmed −1 frameshifting was used to rescue full-length copies of SUI-1 protein to sustain the survival of cell population until functional copies were available again. More experiments are certainly needed to test if *sui-1* is a second example of programmed −1 frameshifting in archaea.

Conclusions

Translational recoding is far less documented in archaea than in the other domains of life, as a consequence of the relatively limited knowledge we have on the molecular biology of these organisms. A typical example of this disparity can be seen in viruses where translational recoding is well known in bacteria and eukaryotes (Dinman 2012), while in those from archaea, for which candidate genes putatively expressed by recoding have been identified (Guo et al. 2011), no experimental data are yet available. Despite these limitations, the study of translational recoding in archaea would be particularly interesting for its implications on the evolution of the genetic code and the origin of life; therefore, possible new achievements in this field would justify the great expectations.

Acknowledgments This work was supported by the project MoMa n. 1/014/06/0 of the Agenzia Spaziale Italiana.

References

Alkalaeva E, Eliseev B, Ambrogelly A, Vlasov P, Kondrashov FA, Gundllapalli S, Frolova L, Soll D, Kisselev L (2009) Translation termination in pyrrolysine-utilizing archaea. FEBS Lett 583: 3455–3460

Allmang C, Wurth L, Krol A (2009) The selenium to selenoprotein pathway in eukaryotes: more molecular partners than anticipated. Biochim Biophys Acta 1790:1415–1423

Ambrogelly A, Palioura S, Soll D (2007) Natural expansion of the genetic code. Nat Chem Biol 3:29–35

Arner ESJ, Holmgren A (2006) The thioredoxin system in cancer. Semin Cancer Biol 16:420–426

Atkinson GC, Hauryliuk V, Tenson T (2011) An ancient family of SelB elongation factor-like proteins with a broad but disjunct distribution across archaea. BMC Evol Biol 11:22–31

Axley MJ, Bock A, Stadtman TC (1991) Catalytic properties of an *Escherichia coli* formate dehydrogenase mutant in which sulfur replaces selenium. Proc Natl Acad Sci USA 88:8450–8454

Balakirev ES, Ayala FJ (2003) Pseudogenes: are they "Junk" or functional DNA? Annu Rev Genet 37:123–151

Baranov PV, Gesteland RF, Atkins JF (2002) Recoding: translational bifurcations in gene expression. Gene 286:187–201

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

Bock A, Forchhammer K, Heider J, Leinfelder W, Sawers G, Veprek B, Zinoni F (1991) Selenocysteine: the 21st amino acid. Mol Microbiol 5:515–520

Carlson BA, Xu XM, Kryukov GV, Rao M, Berry MJ, Gladyshev VN, Hatfield DL (2004) Identification and characterization of phosphoseryl-tRNA([Ser]Sec) kinase. Proc Natl Acad Sci USA 101:12848–12853

Chambers I, Frampton J, Goldfarb P, Affara N, McBain W, Harrison PR (1986) The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the 'termination' codon, TGA. EMBO J 5:1221–1227

Chavatte L, Brown BA, Driscoll DM (2005) Ribosomal protein L30 is a component of the UGA-selenocysteine recoding machinery in eukaryotes. Nat Struct Mol Biol 12:408–416

Cobucci-Ponzano B, Trincone A, Giordano A, Rossi M, Moracci M (2003a) Identification of an archaeal alpha-L-fucosidase encoded by an interrupted gene—production of a functional enzyme by mutations mimicking programmed-1 frameshifting. J Biol Chem 278:14622–14631

Cobucci-Ponzano B, Trincone A, Giordano A, Rossi M, Moracci M (2003b) Identification of the catalytic nucleophile of the family 29 alpha-L-fucosidase from *Sulfolobus solfataricus* via chemical rescue of an inactive mutant. Biochemistry 42:9525–9531

Cobucci-Ponzano B, Mazzone M, Rossi M, Moracci M (2005a) Probing the catalytically essential residues of the alpha-L-fucosidase from the hyperthermophilic Archaeon *Sulfolobus solfataricus*. Biochemistry 44:6331–6342

Cobucci-Ponzano B, Rossi M, Moracci M (2005b) Recoding in archaea. Mol Microbiol 55:339–348

Cobucci-Ponzano B, Conte F, Benelli D, Londei P, Flagiello A, Monti M, Pucci P, Rossi M, Moracci M (2006) The gene of an archaeal alpha-L-fucosidase is expressed by translational frameshifting. Nucleic Acids Res 34:4258–4268

Cobucci-Ponzano B, Conte F, Bedini E, Corsaro MM, Parrilli M, Sulzenbacher G, Lipski A, Dal Piaz F, Lepore L, Rossi M, Moracci M (2009) beta-Glycosyl azides as substrates for alpha-glycosynthases: preparation of efficient alpha-L-fucosynthases. Chem Biol 16:1097–1108

Cobucci-Ponzano B, Guzzini L, Benelli D, Londei P, Perrodou E, Lecompte O, Tran D, Sun J, Wei J, Mathur EJ, Rossi M, Moracci M (2010) Functional characterization and high-throughput proteomic analysis of interrupted genes in the archaeon *Sulfolobus solfataricus*. J Proteome Res 9:2496–2507

Cone JE, Del Rio RM, Davis JN, Stadtman TC (1976) Chemical characterization of the selenoprotein component of clostridial glycine reductase: identification of selenocysteine as the organoselenium moiety. Proc Natl Acad Sci USA 73:2659–2663

Copeland PR, Fletcher JE, Carlson BA, Hatfield DL, Driscoll DM (2000) A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. EMBO J 19:306–314

Cui Y, Dinman JD, Kinzy TG, Peltz SW (1998) The Mof2/Sui1 protein is a general monitor of translational accuracy. Mol Cell Biol 18:1506–1516

Deppenmeier U, Johann A, Hartsch T, Merkl R, Schmitz RA, Martinez-Arias R, Henne A, Wiezer A, Baumer S, Jacobi C, Bruggemann H, Lienard T, Christmann A, Bomeke M, Steckel S, Bhattacharyya A, Lykidis A, Overbeek R, Klenk HP, Gunsalus RP, Fritz HJ, Gottschalk G (2002) The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between bacteria and archaea. J Mol Microb Biotech 4:453–461

Dinman JD (2012) Control of gene expression by translational recoding. Adv Protein Chem Struct Biol 86:129–149

Farabaugh PJ (1996) Programmed translational frameshifting. Microbiol Rev 60:103

Fekner T, Li X, Lee MM, Chan MK (2009) A pyrrolysine analogue for protein click chemistry. Angew Chem Int Ed 48:1633–1635

Galagan JE, Nusbaum C, Roy A, Endrizzi MG, Macdonald P, FitzHugh W, Calvo S, Engels R, Smirnov S, Atnoor D, Brown A, Allen N, Naylor J, Stange-Thomann N, DeArellano K, Johnson R, Linton L, McEwan P, McKernan K, Talamas J, Tirrell A, Ye WJ, Zimmer A, Barber RD, Cann I, Graham DE, Grahame DA, Guss AM, Hedderich R, Ingram-Smith C, Kuettner HC, Krzycki JA, Leigh JA, Li WX, Liu JF, Mukhopadhyay B, Reeve JN, Smith K, Springer TA, Umayam LA, White O, White RH, de Macario EC, Ferry JG, Jarrell KF, Jing H, Macario AJL, Paulsen I, Pritchett M, Sowers KR, Swanson RV, Zinder SH, Lander E, Metcalf WW, Birren B (2002) The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. Genome Res 12:532–542

Gaston MA, Jiang RS, Krzycki JA (2011a) Functional context, biosynthesis, and genetic encoding of pyrrolysine. Curr Opin Microbiol 14:342–349

Gaston MA, Zhang LW, Green-Church KB, Krzycki JA (2011b) The complete biosynthesis of the genetically encoded amino acid pyrrolysine from lysine. Nature 471:647–650

Gautier A, Deiters A, Chin JW (2011) Light-activated kinases enable temporal dissection of signaling networks in living cells. J Am Chem Soc 133:2124–2127

Gesteland RF, Atkins JF (1996) Recoding: dynamic reprogramming of translation. Annu Rev Biochem 65:741–768

Goodchild A, Saunders NFW, Ertan H, Raftery M, Guilhaus M, Curmi PMG, Cavicchioli R (2004) A proteomic determination of cold adaptation in the Antarctic archaeon, *Methanococcoides burtonii*. Mol Microbiol 53:309–321

Guo L, Brugger K, Liu C, Shah SA, Zheng HJ, Zhu YQ, Wang SY, Lillestol RK, Chen LM, Frank J, Prangishvili D, Paulin L, She QX, Huang L, Garrett RA (2011) Genome analyses of Icelandic strains of *Sulfolobus islandicus*, model organisms for genetic and virus–host interaction studies. J Bacteriol 193:1672–1680

Halboth S, Klein A (1992) *Methanococcus voltae* harbors four gene clusters potentially encoding two [NiFe] and two [NiFeSe] hydrogenases, each of the cofactor F420-reducing or F420-non-reducing types. Mol Gen Genet 233:217–224

Hancock SM, Uprety R, Deiters A, Chin JW (2010) Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. J Am Chem Soc 132:14819–14824

Hao B, Gong WM, Ferguson TK, James CM, Krzycki JA, Chan MK (2002) A new UAG-encoded residue in the structure of a methanogen methyltransferase. Science 296:1462–1466

Hao B, Zhao G, Kang PT, Soares JA, Ferguson TK, Gallucci J, Krzycki JA, Chan MK (2004) Reactivity and chemical synthesis

of L-pyrrolysine—the 22(nd) genetically encoded amino acid. Chem Biol 11:1317–1324
Harrison PM, Gerstein M (2002) Studying Genomes through the aeons: protein families, pseudogenes and proteome evolution. J Mol Biol 318:1155–1174
Heinemann IU, O'Donoghue P, Madinger C, Benner J, Randau L, Noren CJ, Soll D (2009) The appearance of pyrrolysine in tRNA(His) guanylyltransferase by neutral evolution. Proc Natl Acad Sci USA 106:21103–21108
Hirotsune S, Yoshida N, Chen A, Garrett L, Sugiyama F, Takahashi S, Yagami K, Wynshaw-Boris A, Yoshiki A (2003) An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. Nature 423:91–96
James CM, Ferguson TK, Leykam JF, Krzycki JA (2001) The amber codon in the gene encoding the monomethylamine methyltransferase isolated from *Methanosarcina barkeri* is translated as a sense codon. J Biol Chem 276:34252–34258
Jones JB, Stadtman TC (1981) Selenium-dependent and selenium-independent formate dehydrogenases of *Methanococcus vannielii*. Separation of the two forms and characterization of the purified selenium-independent form. J Biol Chem 256:656–663
Jones JB, Dilworth GL, Stadtman TC (1979) Occurrence of selenocysteine in the selenium-dependent formate dehydrogenase of *Methanococcus vannielii*. Arch Biochem Biophys 195:255–260
Kaiser JT, Gromadski K, Rother M, Engelhardt H, Rodnina MV, Wahl MC (2005) Structural and functional investigation of a putative archaeal selenocysteine synthase. Biochemistry 44:13315–13327
Kohrle J, Brigelius-Flohe R, Bock A, Gartner R, Meyer O, Flohe L (2000) Selenium in biology: facts and medical perspectives. Biol Chem 381:849–864
Kryukov GV, Gladyshev VN (2004) The prokaryotic selenoproteome. EMBO Rep 5:538–543
Krzycki JA (2004) Function of genetically encoded pyrrolysine in corrinoid-dependent methylamine methyltransferases. Curr Opin Chem Biol 8:484–491
Krzycki JA (2005) The direct genetic encoding of pyrrolysine. Curr Opin Microbiol 8:706–712
Kyrpides NC, Woese CR (1998) Universally conserved translation initiation factors. Proc Natl Acad Sci USA 95:224–228
Li X, Fekner T, Ottesen JJ, Chan MK (2009) A pyrrolysine analogue for site-specific protein ubiquitination. Angew Chem Int Ed 48:9184–9187
Longstaff DG, Blight SK, Zhang LW, Green-Church KB, Krzycki JA (2007a) In vivo contextual requirements for UAG translation as pyrrolysine. Mol Microbiol 63:229–241
Longstaff DG, Larue RC, Faust JE, Mahapatra A, Zhang LW, Green-Church KB, Krzycki JA (2007b) A natural genetic code expansion cassette enables transmissible biosynthesis and genetic encoding of pyrrolysine. Proc Natl Acad Sci USA 104:1021–1026
Maeder DL, Anderson I, Brettin TS, Bruce DC, Gilna P, Han CS, Lapidus A, Metcalf WW, Saunders E, Tapia R, Sowers KR (2006) The *Methanosarcina barkeri* genome: comparative analysis with *Methanosarcina acetivorans* and *Methanosarcina mazei* reveals extensive rearrangement within methanosarcinal genomes. J Bacteriol 188:7922–7931
Namy O, Rousset JP, Napthine S, Brierley I (2004) Reprogrammed genetic decoding in cellular gene expression. Mol Cell 13: 157–168
Neumann H, Peak-Chew SY, Chin JW (2008) Genetically encoding N-epsilon-acetyllysine in recombinant proteins. Nat Chem Biol 4:232–234
Nozawa K, O'Donoghue P, Gundllapalli S, Araiso Y, Ishitani R, Umehara T, Soll D, Nureki O (2009) Pyrrolysyl-tRNA synthetase-tRNA(Pyl) structure reveals the molecular basis of orthogonality. Nature 457:1163–1167
Perrodou E, Deshayes C, Muller J, Schaeffer C, Van Dorsselaer A, Ripp R, Poch O, Reyrat JM, Lecompte O (2006) ICDS database: interrupted CoDing sequences in prokaryotic genomes. Nucl Acids Res 34:D338–D343
Pinsent J (1954) The need for selenite and molybdate in the formation of formic dehydrogenase by members of the coli-aerogenes group of bacteria. Biochem J 57:10–16
Polycarpo C, Ambrogelly A, Berube A, Winbush SAM, McCloskey JA, Crain PF, Wood JL, Soll D (2004) An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. Proc Natl Acad Sci USA 101:12450–12454
Rosano C, Zuccotti S, Cobucci-Ponzano B, Mazzone M, Rossi M, Moracci M, Petoukhov MV, Svergun DI, Bolognesi M (2004) Structural characterization of the nonameric assembly of an Archaeal alpha-L-fucosidase by synchrotron small angle X-ray scattering. Biochem Biophys Res Commun 320:176–182
Rother M, Krzycki JA (2010) Selenocysteine, pyrrolysine, and the unique energy metabolism of methanogenic archaea. Archaea Int Microbiol J
Rother M, Wilting R, Commans S, Bock A (2000) Identification and characterisation of the selenocysteine-specific translation factor SelB from the archaeon *Methanococcus jannaschii*. J Mol Biol 299:351–358
Rother M, Resch A, Wilting R, Bock A (2001) Selenoprotein synthesis in archaea. BioFactors 14:75–83
Sekine Y, Eisaki N, Ohtsubo E (1994) Translational control in production of transposase and in transposition of insertion sequence-Is3. J Mol Biol 235:1406–1420
Sharma V, Firth AE, Antonov I, Fayet O, Atkins JF, Borodovsky M, Baranov PV (2011) A pilot study of bacterial genes with disrupted ORFs reveals a surprising profusion of protein sequence recoding mediated by ribosomal frameshifting and transcriptional realignment. Mol Biol Evol 28:3195–3211
She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, Chan-Weiher CCY, Clausen IG, Curtis BA, De Moors A, Erauso G, Fletcher C, Gordon PMK, Heikamp-de Jong I, Jeffries AC, Kozera CJ, Medina N, Peng X, Thi-Ngoc HP, Redder P, Schenk ME, Theriault C, Tolstrup N, Charlebois RL, Doolittle WF, Duguet M, Gaasterland T, Garrett RA, Ragan MA, Sensen CW, Van der Oost J (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc Natl Acad Sci USA 98:7835–7840
Sorgenfrei O, Linder D, Karas M, Klein A (1993) A novel very small subunit of a selenium containing [NiFe] hydrogenase of *Methanococcus voltae* is post-translationally processed by cleavage at a defined position. Eur J Biochem 213:1355–1358
Squires JE, Berry MJ (2008) Eukaryotic selenoprotein synthesis: mechanistic insight incorporating new factors and new functions for old factors. IUBMB Life 60:232–235
Srinivasan G, James CM, Krzycki JA (2002) Pyrrolysine encoded by UAG in archaea: charging of a UAG-decoding specialized tRNA. Science 296:1459–1462
Stock T, Rother M (2009) Selenoproteins in archaea and gram-positive bacteria. Biochim Biophys Acta 1790:1520–1532
Stock T, Selzer M, Rother M (2010) In vivo requirement of selenophosphate for selenoprotein synthesis in archaea. Mol Microbiol 75:149–160
van Passel MW, Smillie CS, Ochman H (2007) Gene decay in archaea. Archaea 2:137–143
Vorholt JA, Vaupel M, Thauer RK (1997) A selenium-dependent and a selenium-independent formylmethanofuran dehydrogenase and their transcriptional regulation in the hyperthermophilic *Methanopyrus kandleri*. Mol Microbiol 23:1033–1042

Wilting R, Schorling S, Persson BC, Bock A (1997) Selenoprotein synthesis in archaea: identification of an mRNA element of *Methanococcus jannaschii* probably directing selenocysteine insertion. J Mol Biol 266:637–641

Xu J, Hendrix RW, Duda RL (2004) Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. Mol Cell 16:11–21

Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama S (2008a) Crystallographic studies on multiple conformational states of active-site loops in pyrrolysyl-tRNA synthetase. J Mol Biol 378:634–652

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

Yuan J, Palioura S, Salazar JC, Su D, O'Donoghue P, Hohn MJ, Cardoso AM, Whitman WB, Soll D (2006) RNA-dependent conversion of phosphoserine forms selenocysteine in eukaryotes and archaea. Proc Natl Acad Sci USA 103:18923–18927

Zhang Y, Baranov PV, Atkins JF, Gladyshev VN (2005) Pyrrolysine and selenocysteine use dissimilar decoding strategies. J Biol Chem 280:20740–20751

Zhang Y, Romero H, Salinas G, Gladyshev VN (2006) Dynamic evolution of selenocysteine utilization in bacteria: a balance between selenoprotein loss and evolution of selenocysteine from redox active cysteine residues. Genome Biol 7:R94

Zinoni F, Birkmann A, Stadtman TC, Bock A (1986) Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from *Escherichia coli*. Proc Natl Acad Sci USA 83: 4650–4654

Zinoni F, Birkmann A, Leinfelder W, Bock A (1987) Cotranslational insertion of selenocysteine into formate dehydrogenase from *Escherichia coli* directed by a UGA codon. Proc Natl Acad Sci USA 84:3156–3160