University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Vadim Gladyshev Publications | Biochemistry, Department of

6-28-2006

Evolution of selenocysteine decoding and the key role of selenophosphate synthetase in the pathway of selenium utilization

Gustavo Salinas
Universidad de la República, Montevideo, Uruguay

Héctor Romero
Instituto de Biología, Montevideo, Uruguay

Xue-Ming Xu
National Cancer Institute, National Institutes of Health, Bethesda, MD

Bradley A. Carlson
National Cancer Institute, National Institutes of Health, Bethesda, MD

Dolph L. Hatfield
National Cancer Institute, National Institutes of Health, Bethesda, MD

See next page for additional authors

Follow this and additional works at: http://digitalcommons.unl.edu/biochemgladyshev

Part of the Biochemistry, Biophysics, and Structural Biology Commons

Salinas, Gustavo; Romero, Héctor; Xu, Xue-Ming; Carlson, Bradley A.; Hatfield, Dolph L.; and Gladyshev, Vadim N., "Evolution of selenocysteine decoding and the key role of selenophosphate synthetase in the pathway of selenium utilization" (2006). *Vadim Gladyshev Publications.* Paper 39.
http://digitalcommons.unl.edu/biochemgladyshev/39

This Article is brought to you for free and open access by the Biochemistry, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Vadim Gladyshev Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors
Gustavo Salinas, Héctor Romero, Xue-Ming Xu, Bradley A. Carlson, Dolph L. Hatfield, and Vadim N. Gladyshev

This article is available at DigitalCommons@University of Nebraska - Lincoln: http://digitalcommons.unl.edu/biochemgladyshev/39

SELENIUM
Its Molecular Biology and Role in Human Health, Second Edition

Edited by

Dolph L. Hatfield
National Cancer Institute, USA

Marla J. Berry
University of Hawaii, USA

and

Vadim N. Gladyshev
University of Nebraska, USA

Springer

2006

Chapter 4. Evolution of selenocysteine decoding and the key role of selenophosphate synthetase in the pathway of selenium utilization

Gustavo Salinas

Cátedra de Inmunología, Facultad de Química-Facultad de Ciencias, Universidad de la República. Instituto de Higiene, Avda. A. Navarro 3051, Montevideo, CP 11600, Uruguay

Héctor Romero

Laboratorio de Organización y Evolución del Genoma, Dpto. de Biología Celular y Molecular, Instituto de Biología, Facultad de Ciencias, Iguá 4225, Montevideo, CP 11400, Uruguay

Xue-Ming Xu, Bradley A. Carlson and Dolph L. Hatfield

Molecular Biology of Selenium Section, Laboratory of Cancer Prevention, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

Vadim N. Gladyshev

Department of Biochemistry, University of Nebraska, Lincoln, NE 68588, USA

Summary: The complete sequencing of genomes and the development of *in silico* methods for identification of genes encoding selenocysteine (Sec)-containing proteins have greatly contributed to shape our view on the evolution of selenium utilization in nature. Current evidence is consistent with the idea that Sec decoding is a late addition to the genetic code and it evolved once, before the separation of archaeal, bacterial and eukaryal domains. Many organisms have lost the Sec decoding trait, but recent evidence has shown that the loss is not irreversible. The distribution of organisms that use UGA as a Sec codon suggests that Sec decoding evolved as a result of speciation, differential gene loss and horizontal gene transfer. Selenium is also used in the synthesis 2-selenouridine, a modified base of unknown function located in the wobble position of certain tRNAs. It has been recently demonstrated that selenouridine and Sec-decoding traits can evolve independently of each other, but both require selenophosphate synthetase. This ATP-dependent enzyme emerged as a key feature of selenium utilization that allows separation of selenium from the pathways of

sulfur utilization and non-specific use of selenium. Some animals, including mammals, evolved two selenophosphate synthetases, highlighting an unknown complexity of selenium utilization in nature.

Introduction

Co-translational incorporation of selenocysteine (Sec) into nascent polypeptides is neither canonical nor universal. A Sec-decoding apparatus is needed to reprogram specific UGA codons [1-3]. The Sec-decoding apparatus and selenoprotein genes are present in the three domains of life; yet, many taxa lack them. In Sec decoding species, the selenoproteome consists of a restricted number of proteins [4,5]. All these observations have raised important questions regarding the evolution of Sec utilization in nature. For example, how and when did the translation machinery to decode Sec evolve? If it evolved once, has it been perpetuated solely by vertical descent? Has the UGA codon evolved from nonsense to sense or vice versa? Have extant selenoproteins evolved from Cys-containing proteins or vice versa? What are the selective forces that result in maintenance, loss and acquisition of the Sec-decoding trait and selenoproteins? In a broader scenario, studies on the evolution of Sec invite more in-depth questions regarding the evolution of the genetic code and the translation machinery. Recent work allowed some of these questions to be answered providing a provisional evolutionary scenario [5-7]. At the same time, some unknowns remain. In this chapter we review the current knowledge regarding Sec and selenium utilization in nature, their evolution, and highlight a key role of selenophosphate synthetase in these processes.

Sec decoding: common origin before the division of the three domains?

Current evidence strongly suggests that the Sec decoding trait evolved once, before the division of bacterial, archaeal and eukaryal domains. For example, there are fundamental similarities in the three domains: i) Sec is decoded by UGA-matching $tRNA^{Sec}$ (also known as *selC*) and a dedicated elongation factor (EFsec, also known as *selB*); ii) the translational reprogramming is fulfilled by the SECIS element present in selenoprotein mRNAs; and iii) Sec synthesis occurs on a tRNA scaffold as reviewed in [2,3,8,9]. A common origin is further supported by the recent phylogenetic analysis of the genes involved in Sec decoding [7], which indicates that the trait is monophyletic in the bacterial domain and that eukaryal and archaeal Sec-decoding genes have a common ancestor. These observations suggest that the most parsimonious and likely evolutionary scenario for the trait is a common origin between the three domains, and not independent origins. The greater similarity between archaeal and eukaryal domains may reflect the fact that the transcription and translation machinery in archaea and eukarya is thought to be of common origin. Further studies should be carried out to identify and

date the time of divergence of the different genes involved in Sec decoding and compare these patterns with divergence of the three domains. Syvanen [10] has proposed that the unity of the genetic code is the product of an evolutionary process that has continued since the diversification of the major domains and specifically suggested that the last common ancestor (which defines the origin of the three domains) did not use arginine and tryptophan. In this alternative scenario, horizontal gene transfer (HGT) would have played a critical role in homogenization of the code. If this proposition is correct, it is then possible that Sec might not have been the 21st amino acid added to the genetic code.

Despite the remarkable similarities, differences do exist in the Sec-decoding traits in the three domains of life, including an increased complexity of the pathway in eukaryotes [11-14]. It has been argued that the differences between bacterial and eukaryal Sec incorporation are due to a refinement of the mechanism to provide an increased efficiency in Sec incorporation [15]. It is also likely that some changes favored a greater flexibility in reprogramming. Indeed, the location of the SECIS element within the untranslated region in archaea and eukarya released the constraints imposed by the location of SECIS immediately downstream of UGA^{Sec} within the coding region of bacterial messages.

The Sec decoding trait can be lost, but not irreversibly

Recently, the distribution of the Sec-decoding trait was analyzed systematically by searching complete genomes for the presence of genes involved in Sec decoding and selenoprotein(s) and using this information to construct a provisional "Sec decoding map" within the "tree of life" [7]. This study revealed that the trait is present in most phyla, but absent in many species, and provided clues regarding the evolution of Sec. Then, the phylogenies of the Sec-decoding genes were inferred and compared to organismal phylogenies. This approach explained the spread and "holed" pattern of Sec-decoding species within the tree of life as the result of speciation, differential gene loss and horizontal gene transfer (HGT). It also revealed that the loss of the trait is a phenomenon that takes place within clades at different evolutionary levels, implying that the loss occurred not only at rather basal evolutionary levels (*e.g.*, phylum and class), but also in recent lineages (*e.g.*, genus and species). A stunning example of the latter is the case of the CO92 and mediaevalis strains of *Yersinia pestis* that have lost the ability to decode Sec (possess functional tRNA^{Sec} and Sec synthase but an EFsec pseudogene) while the KIM strain retains this ability (unpublished). Yet, the main disclosure of the study was that it clearly demonstrated that the loss of the trait was not irreversible, indicating that the genetic code can be "rewired" by HGT, a possibility previously thought as highly unlikely [16]. This phenomenon was patently observed in the case of

Photobacterium profundum, which did not acquire its extant trait and the selenoproteins involved in the glycine reductase complex by vertical descent from a proteobacterial ancestor, but rather from a different lineage. Incongruences possibly attributable to HGT, between gene and species trees were also observed in the case of *Pseudomomas* spp. (γ-proteobacteria), *Sinorhizobium meliloti* (α-proteobacterium) and *Burkholderia* spp (β-proteobacteria). In these species, the only selenoprotein is the α-subunit of formate dehydrogenase. In this regard, it should be noted that *Pseudomonas* spp, *S. meliloti* and *Burkholderia pseudomallei* are soil colonizing bacteria [17,18]; whereas *S. meliloti* and *Pseudomonas* spp. even compete for nodulation on some plants. It was speculated that a putative vector for acquisition of the trait may exist, on the basis of the observation that *S. meliloti* contains the genetic information for selenoproteins and the Sec-decoding trait within a megaplasmid (pSymA) with a high number of transposons [15].

Selenophosphate synthetase: an essential enzyme for selenium utilization

Selenophosphate synthetase (SPS) is an essential enzyme for selenium utilization: it catalyzes the synthesis of monoselenophosphate [19], a reduced and reactive form of selenium, which provides the selenium atom for synthesis of Sec and 2-selenouridine, another biologically relevant form of selenium in nature (see below and [20]). There appears to be two groups of SPS enzymes. One group contains Sec or Cys at the active site and the corresponding *E. coli* enzyme can catalyze, *in vitro*, the synthesis of selenophosphate from selenide and ATP [21] as follows: $ATP + HSe^- + H_2O \rightarrow H_2SePO_3^- + AMP + P_i$. It should be emphasized that the *Km* value for selenide is 20 μM, a concentration that would be noxious for the cell *in vivo*, suggesting that selenide would not be the physiological selenium donor for SPS2/SelD. The other group (designated SPS1) is present exclusively in some eukaryotic organisms that also possess SPS2 (*e.g.*, mammals); SPS1 neither contains Sec or Cys at the predicted active site position nor appears to catalyze the *in vitro* reaction depicted above [22]. *In vivo*, human SPS2 complemented an *E. coli selD* mutant strain when the medium was supplemented with selenite or Sec, restoring the activity of the selenoprotein formate dehydrogenase to *E. coli* wild type levels. In contrast, complementation with SPS1 was ineffective when selenite was used as a selenium supply, although it improved when Sec was used in the medium [22]. These results led Tamura et al [22] to propose that human SPS2 functions in the pathway of *de novo* synthesis of selenophosphate from selenite, after reduction of the latter, presumably, by intracellular thiols; the Sec residue of SPS2 active site would bind this reduced selenium to form an enzyme substrate complex. Alternatively, for SPS1 catalysis, mammalian

cells would supply an atom of Se derived from a Sec salvage pathway that may recycle Sec derived from selenoproteins or from the promiscuous incorporation of selenium instead of sulfur in the Cys metabolic pathway. Sec β-lyases, enzymes that catalyze the conversion of Sec to Ala and a selenium transfer form (Se*, since the redox state of Se has not been determined), and NifS and NifS-like proteins (enzymes that provide a sulfur atom to iron-sulfur clusters by catalyzing Cys desulfuration, and also convert Sec to Ala and Se*) are candidate enzymes to participate in selenium mobilization from Sec [23,24], and have been proposed as key players for the Sec-salvage pathway.

It is relevant to emphasize that the bacterial NifS-like proteins CsdB, CSD, and IscS also have an important role in selenium mobilization from Sec and selenophosphate synthesis. Indeed, Sec and bacterial NifS-like proteins can effectively replace the high level of free selenide in the *in vitro* SPS assay [25,26]. Furthermore, the *E. coli* SPS (C17S) mutant, which is inactive in the standard *in vitro* assay with selenide as substrate, was found to be active in the presence of Sec and NifS proteins, suggesting a selenium delivery function for these proteins [26]. Evidence that selenium is mobilized *in vivo* from free Sec has also been obtained by Lacourciere [27]: growth of *E. coli* in the presence 0.1 μM $^{75}Se\ SeO_3^{2-}$ and increasing amounts of Sec resulted in a concomitant decrease in ^{75}Se incorporation in formate dehydrogenase and bulk tRNA. This led Lacourciere to propose that NifS-like proteins are components of a selenium delivery system for the biosynthesis of selenophosphate.

Recently, other potential selenium-binding and delivery proteins for SPS have been characterized. Human 3-mercaptopyruvate sulfur transferase (MST) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH), bound selenium supplied as selenodiglutathione formed from SeO_3^{2-} and glutathione; the bound selenium was readily released from MST and GAPDH and available as a substrate for bacterial SPS [28]. Thus, *in vitro*, these selenium-binding proteins and a low level of selenodigultathione (formed from selenite and glutathione) could effectively replace the high concentrations of selenide used in SPS assays.

Despite this considerable progress, the metabolic pathway(s) of selenium assimilation and the physiological system that donates selenium to SPS remain to be completely characterized and represents one of the challenges ahead in the selenium field.

Selenophosphate synthetase: one enzyme, two selenium utilization traits

2-Selenouridine is the second major biological form of Se in nature [29]. It is a modified base so far identified exclusively in the wobble position of the anticodon of $tRNA^{Lys}$, $tRNA^{Glu}$ and $tRNA^{Gln}$ in some bacteria [30]. 2-Selenouridine is synthesized by the protein designated YbbB from

thiouridine tRNA and selenophosphate (the latter as the selenium donor). Although it was originally thought that organisms able to synthesize 2-selenouridine were the same as those able to decode Sec, it has recently been shown that the set of organisms that synthesize 2-selenouridine overlaps with, yet is distinct from, the set of organisms able to decode Sec (Figure 1).

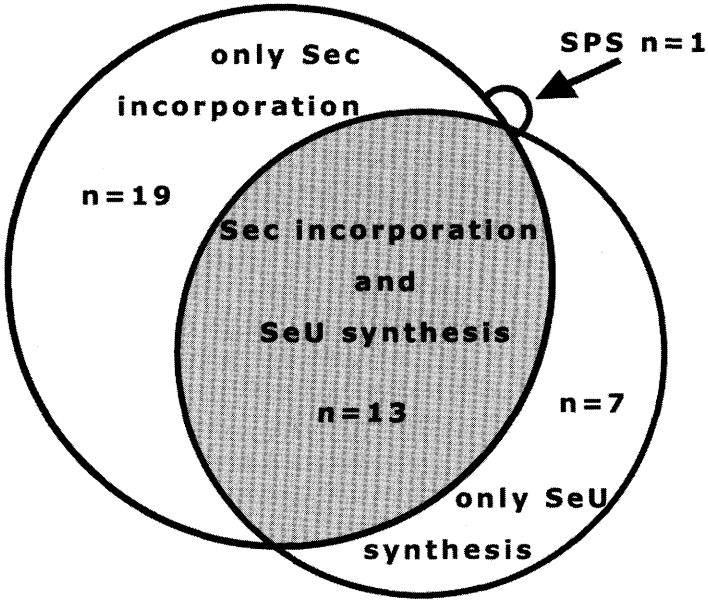

Figure 1. Sec-decoding and selenouridine (SeU) synthesis traits both require selenophosphate synthetase (SPS). However, the set of species that decode Sec overlaps with, but is different from, the set of species that synthesize SeU. The representation excludes species of the eukaryal domain. Note that there is one species possessing SPS, but neither trait (see text). n= numbers of completed prokaryotic genomes. (Total number of complete genomes analyzed was 153.)

Indeed, SPS is required for both Sec and selenouridine synthesis. This study allowed us to define SPS as the gene signature of selenium utilization, YbbB as the gene signature for 2-selenouridine synthesis and EFsec and $tRNA^{Sec}$ as the gene signature of the Sec-decoding trait. Thus, the likely evolutionary scenario is that SPS is required for both traits, but Sec-decoding and 2-selenouridine synthesis traits can evolve independently of each other. Analyses of genomic organization of Se utilization genes in the bacterial domain revealed that SPS is more often arranged in an operon with *ybbB* than with EFsec (*selB*) and Sec synthase (*selA*). HGT events were also identified for the selenouridine synthesis trait; thus, the pattern of species

distribution for both Se traits is the result of speciation, differential gene loss and HGT.

An interesting parallel that might be sketched for both traits is the restricted use of selenium: it is certainly peculiar that the Sec decoding apparatus has been maintained in some species for insertion of a single amino acid into a
single protein; equally unusual is the finding that 2-selenouridine is used in only three bases in the entire transcriptomes of bacteria possessing *ybbB*. Although several functions have been postulated for selenouridine, its function is not known. Based on the facts that 2-selenouridine is found exclusively at the wobble position of codons ending in a purine, and that these codons pose a problem for the translation machinery, it has been postulated that 2-selenouridine would have been an adjustment of the decoding apparatus to increase translational fidelity [7]. Whether this base modification occurs outside the bacterial domain is not known; a low identity homolog to bacterial *ybbB* is present in *Methanococcus jannaschii* and *Methanopyrus kandleri* [7,30].

Finally, the presence of a SPS homolog with high identity in *Enterococcus faecalis*, a species that neither decodes Sec nor possesses *ybbB* is interesting. Furthermore, a Sec lyase homolog and a protein involved in sulfur reduction flank this gene in the genome. This observation suggests that there may be an additional Se utilization trait that occurs in *E. faecalis* and perhaps other organisms (unpublished).

Looking at the phenotype: evolution of selenoproteins is a highly dynamic process

Evolution of selenoproteins is a highly dynamic process linked to the evolution of the Sec decoding trait itself. From the mechanistic point of view, this process involves: i) evolution of selenoproteins from Cys-containing homologs concomitant with the acquisition of SECIS elements [5], ii) evolution of Cys-containing proteins from selenoproteins (fossil SECIS elements have been well documented even in Sec decoding species) [4,31], iii) gene duplication [32,33], and iv) conceivably "invention" of entirely new selenoprotein families for which no Cys homologs have been detected (possibly the glycine reductase component, grdA, in bacteria [5,15], and SelJ in eukarya [5,34]). This highly dynamic process is at least in part due to the following processes: i) Sec and Cys can be replaced by each other concomitant with appropriate adjustments in the protein context [35] and/or protein concentration, and ii) the SECIS element appears to evolve easily, especially in archaea and eukarya, in which the 3'-UTR location of this structure is not constrained by the coding region.

Yet, the evolutionary forces that shape Sec utilization, in particular the delicate balance between processes that maintain, acquire or lose this trait

are not clear. It has been previously postulated that the selective advantage provided by selenoenzymes over Cys-homologs (due to the better nucleophilicity and lower pKa of SeH in Sec over SH in Cys), might become a disadvantage if selenium supply becomes limiting [7,15,36]. In these situations, enzymes containing Sec as catalytic residues could have evolved into Cys-containing proteins, or alternatively, both Sec-containing and Cys-containing forms could have been maintained allowing organisms to use Se in a facultative manner [36]. An illustrative example of a Se facultative organism is *Methanococcus maripaludis*, which represses the synthesis of the Cys homologs when grown in a medium that contains adequate amounts of Se, but this repression is not observed in a mutant with disrupted *selB* [36]. Nevertheless, the occurrence of organisms carrying only one of the selenium utilization traits indicates that Se availability might not be the sole factor involved in the loss of either trait.

It is also of interest that selenoproteomes, while differing in size in composition between various organisms, always represent a very small subset of the total protein complement. One may consider Sec as a two-edged weapon – the advantage of having a highly reactive selenol group might become a disadvantage if used indiscriminately to replace Cys. This may restrict the pervasive use of Sec in the environments where Se supply may be adequate or excessive [7]. Another conspicuous feature of Sec utilization in nature is its idiosyncratic use by different taxa: different sets of selenoproteins have evolved in different lineages [6,34,37]. This finding indicates adaptations in the use of Sec, presumably to fulfill particular needs. It has been suggested that both lineage-specific expansion (presumably of recent origin), and the presence of core selenoproteins (ancient origin) appear to contribute to extant selenoproteomes [5].

Analyses of prokaryotic selenoproteomes revealed that formate dehydrogenase is present in most Sec decoding organisms [5], suggesting that under anaerobic respiration Sec-containing formate dehydrogenase confers a specific advantage [7]. Indeed, most of selenoprotein formate dehydrogenase-containing species are obligatory anaerobes or facultative aerobes; the sole exception appears to be *S. melilloti*, a symbiotic nitrogen-fixing obligatory aerobe that lives in the oxygen-limited environment of the nodule. On the other hand, glycine reductases, present in *T. denticola*, *P. profundum*, *T. tengcongensis* and several species of the genus Clostridium might have conferred a selective advantage allowing certain anaerobic bacteria to conserve energy via a soluble substrate level phosphorylation system [38].

Synthesis of Sec-tRNA^{Sec}: a non-canonical mechanism

A conspicuous feature of Sec synthesis is that it occurs on its tRNA, from Ser-tRNA^{Sec}, which is then converted into Sec-tRNA^{Sec} using

monoselenophosphate as the selenium donor [39,40]. This latter reaction is catalyzed by Sec synthase (SelA) in the bacterial domain. The equivalent enzyme(s) in archaea and eukarya is (are) not known, although the mechanism is thought to involve a Sep-tRNA^{Sec} intermediate (Sep=phosphoserine) [41].

Table 1. Non-cognate charging of amino acids into tRNAs

Final aa attached to tRNA	Initial aa charged to tRNA	Charging aminoacyl-tRNA synthetase	Phylogenetic Distribution
Asn	Asp, Glu	aspRS gluRS	archaea bacteria
Cys	Sep	sepRS	archaea
fMet	Met	metRS	bacteria, organelles
Gln	Glu, Asp	gluRS aspRS	archaea bacteria
Pyl	Lys	lysRS1+lysRS2	archaea
Sec	**Ser**	**serRS**	**archaea bacteria eukarya**

Nature has evolved two different mechanisms for synthesis of aminoacyl-tRNAs: i) the canonical one (*i.e.*, a specific amino acyl tRNA synthetase recognizes an amino acid and its cognate tRNA), and ii) the non-canonical mechanism wherein the tRNA is loaded first with a "non-cognate" amino acid that is then modified to the amino acid to be incorporated into protein. Although the amino acid biosynthesis on a tRNA scaffold is not unique to Sec (Table 1, after [42] and [43]), it is interesting to point out that Sec is the sole example in which the second mechanism appears to occur in the three domains of life, and might be the predominant or even exclusive biosynthetic mechanism. It is also important to note that Sec synthesis appears to resemble Cys synthesis on tRNA in two aspects: i) in archaea and eukarya, Ser-tRNA^{Sec} is first converted into Sep-tRNA^{Sec}, whereas Sep is the intermediate in Cys synthesis on tRNA in some archaea [44], and ii) synthesis of Cys on tRNA is the sole mechanism by which Cys is synthesized in a subset of archaea that use this strategy. It has been speculated that SepCysS (the enzyme that charges Sep directly to tRNA^{Cys}) provided a means by which both Cys and Sec may have been originally added to the genetic code [44]. More generally, it could be speculated that *in*

situ synthesis of an amino acid provides a strategy by which the amino acid repertoire (and the genetic code) could have been, and conceivably can be expanded.

Lessons from the genetic code

Nearly four decades ago the genetic code was deciphered [45]. Yet, the signals involved in translation of information are neither universal nor completely known, and several mechanisms of reprogramming have been documented [46]. Sec decoding provides clues to how, with very few genes, a codon can be reprogrammed in specific messages, increasing the amino acid repertoire. In addition, it illustrates how flexible and dynamic is the evolution of the genetic code: the ability to decode Sec can be lost, but later reacquired by HGT of a handful of genes. The case of Sec is also interesting in that it is a "non-standard" amino acid that is co-translationally incorporated, but has not been fixed as the only (or major) function of UGA in the genetic code. In turn, this raises the question of why Sec has not been "hardwired" (i.e., reassigned a codon), and what are the differences between increasing the amino acid repertoire post-translationally or co-translationally.

Stop codons have repeatedly evolved particular meanings (Trp, Cys, Sec, Pyl) either by codon reassignment or by reprogramming, suggesting that a change of meaning could be less deleterious for a nonsense codon. In particular, whether the UGA codon originally specified Sec or stop has been a matter of debate. It was speculated [16] that UGA evolved from sense to nonsense, and postulated a possible scenario where Sec was one of the earliest amino acids in evolutionary history. And once oxygen evolved in the atmosphere, the susceptibility to oxidation could have counterselected Sec resulting in the stop function of UGA and maintaining the Sec function by means of SECIS element and EF-sec. On the other hand, it was also proposed [47] that UGA evolved from nonsense to sense, and that Sec was added to the already fixed code. Although the question of the ancestral meaning of UGA is difficult to solve, the very need of reprogramming specific messages suggests that the Sec insertion has evolved on top of an existing translational machinery to acquire a new meaning for an already assigned codon. In other words, "loose" programming (for specific messages) must have been a novelty added to the "hardwired" code (i.e., universal, for all messages).

Further clues to the idiosyncrasies of the genetic code and its dynamic evolution have arisen from pyrrolysine (Pyl), a recently discovered "non-standard" amino acid [48]. The mechanism of Pyl incorporation, in particular whether it requires recoding (and if it does, what are the signals) is not clear [49]. Although Sec and Pyl use dissimilar decoding strategies, there are clear parallels: Sec and Pyl are amino acids used in a small set of

proteins and both are expansions of the amino acid repertoire by redefining the meaning of stop codons. The diversity of mechanisms involved in translation, together with the simple machinery needed to expand the genetic code, opens the door to the possibility that 22 genetically encoded amino acids may not be the final number and that additional amino acids may exist in the genetic code.

References

1. A Bock 2000 *Biofactors* 11:77
2. MJ Berry, RM Tujebajeva, PR Copeland, XM Xu, BA Carlson, GW Martin, 3rd, SC Low, JB Mansell, E Grundner-Culemann, JW Harney et al 2001 *Biofactors* 14:17
3. DL Hatfield, VN Gladyshev 2002 *Mol Cell Biol* 22:3565
4. GV Kryukov, S Castellano, SV Novoselov, AV Lobanov, O Zehtab, R Guigo, VN Gladyshev 2003 *Science* 300:1439
5. GV Kryukov, VN Gladyshev 2004 *EMBO Rep* 5:538
6. Y Zhang, DE Fomenko, VN Gladyshev 2005 *Genome Biol* 6:R37
7. H Romero, Y Zhang, VN Gladyshev, G Salinas 2005 *Genome Biol* 6:R66
8. M Rother, A Resch, R Wilting, A Bock 2001 *Biofactors* 14:75
9. A Lescure, D Fagegaltier, P Carbon, A Krol 2002 *Curr Protein Pept Sci* 3:143
10. M Syvanen 2002 *Trends Genet* 18:245
11. SA Kinzy, K Caban, PR Copeland 2005 *Nucleic Acids Res* 33:5172
12. XM Xu, H Mix, BA Carlson, PJ Grabowski, VN Gladyshev, MJ Berry, DL Hatfield 2005 *J Biol Chem* 280:41568
13. L Chavatte, BA Brown, DM Driscoll 2005 *Nat Struct Mol Biol* 12:408
14. AL Small-Howard, MJ Berry 2005 *Biochem Soc Trans* 33:1493
15. PR Copeland 2005 *Genome Biol* 6:221
16. A Bock, K Forchhammer, J Heider, C Baron 1991 *Trends Biochem Sci* 16:463
17. MT Holden, RW Titball, SJ Peacock, AM Cerdeno-Tarraga, T Atkins, LC Crossman, T Pitt, C Churcher, K Mungall, SD Bentley et al 2004 *Proc Natl Acad Sci U S A* 101:14240
18. PM Riccillo, CI Muglia, FJ de Bruijn, AJ Roe, IR Booth, OM Aguilar 2000 *J Bacteriol* 182:1748
19. W Leinfelder, K Forchhammer, B Veprek, E Zehelein, A Bock 1990 *Proc Natl Acad Sci U S A* 87:543
20. TC Stadtman, JN Davis, E Zehelein, A Bock 1989 *Biofactors* 2:35
21. IY Kim, Z Veres, TC Stadtman 1992 *J Biol Chem* 267:19650
22. T Tamura, S Yamamoto, M Takahata, H Sakaguchi, H Tanaka, TC Stadtman, K Inagaki 2004 *Proc Natl Acad Sci U S A* 101:16162
23. A Lescure, D Gautheret, P Carbon, A Krol 1999 *J Biol Chem* 274:38147
24. H Mihara, T Kurihara, T Watanabe, T Yoshimura, N Esaki 2000 *J Biol Chem* 275:6195
25. GM Lacourciere, TC Stadtman 1998 *J Biol Chem* 273:30921
26. GM Lacourciere, H Mihara, T Kurihara, N Esaki, TC Stadtman 2000 *J Biol Chem* 275:23769
27. GM Lacourciere 2002 *J Bacteriol* 184:1940
28. Y Ogasawara, GM Lacourciere, K Ishii, TC Stadtman 2005 *Proc Natl Acad Sci U S A* 102:1012
29. Z Veres, TC Stadtman 1994 *Proc Natl Acad Sci U S A* 91:8092
30. MD Wolfe, F Ahmed, GM Lacourciere, CT Lauhon, TC Stadtman, TJ Larson 2003 *J Biol Chem* 279:1801
31. A Bock, M Rother 2004 *Arch Microbiol* 183:148
32. JA Vorholt, M Vaupel, RK Thauer 1997 *Mol Microbiol* 23:1033

33. S Castellano, SV Novoselov, GV Kryukov, A Lescure, E Blanco, A Krol, VN Gladyshev, R Guigo 2004 *EMBO Rep* 5:71
34. S Castellano, AV Lobanov, C Chapple, SV Novoselov, M Albrecht, D Hua, A Lescure, T Lengauer, A Krol, VN Gladyshev et al 2005 *Proc Natl Acad Sci U S A* 102:16188
35. HY Kim, VN Gladyshev 2005 *PLoS Biol* 3:e375
36. M Rother, I Mathes, F Lottspeich, A Bock 2003 *J Bacteriol* 185:107
37. RJ Stillwell, MJ Berry 2005 *Proc Natl Acad Sci U S A* 102:16123
38. JR Andreesen 2004 *Curr Opin Chem Biol* 8:454
39. BJ Lee, PJ Worland, JN Davis, TC Stadtman, DL Hatfield 1989 *J Biol Chem* 264:9724
40. BJ Lee, M Rajagopalan, YS Kim, KH You, KB Jacobson, D Hatfield 1990 *Mol Cell Biol* 10:1940
41. BA Carlson, XM Xu, GV Kryukov, M Rao, MJ Berry, VN Gladyshev, DL Hatfield 2004 *Proc Natl Acad Sci U S A* 101:12848
42. M Di Giulio 2005 *Biosystems* 80:175
43. M Ibba, D Soll 2004 *Genes Dev* 18:731
44. A Sauerwald, W Zhu, TA Major, H Roy, S Palioura, D Jahn, WB Whitman, JR Yates, 3rd, M Ibba, D Soll 2005 *Science* 307:1969
45. FH Crick 1966 *Cold Spring Harbor Symp Quant Biol* 31:1
46. O Namy, JP Rousset, S Napthine, I Brierley 2004 *Mol Cell* 13:157
47. VN Gladyshev, GV Kryukov 2001 *Biofactors* 14:87
48. B Hao, W Gong, TK Ferguson, CM James, JA Krzycki, MK Chan 2002 *Science* 296:1462
49. Y Zhang, PV Baranov, JF Atkins, VN Gladyshev 2005 *J Biol Chem* 280:20740