Chapter 3
Molecular Mechanism of Eukaryotic Selenocysteine Incorporation

Michael T. Howard, Jonathan N. Gonzales-Flores, and Paul R. Copeland

Abstract Although substantial progress has been made in determining which factors are required for eukaryotic Sec incorporation, the mechanism by which the factors are able to alter the coding potential of an mRNA at specific UGA codons is still not known. What is clear is that a complex interplay between *cis*- and *trans*-acting factors regulates the selenocysteine (Sec) incorporation event both at the basal level and in determining the efficiency of the process. In this chapter, we dissect the current state of knowledge regarding this interplay and delve into the increasingly important role that in vitro systems will play in determining the precise mechanism by which Sec is incorporated into selenoproteins.

3.1 Introduction

The a priori assumption upon discovery of the genetic code, over 4 decades ago, was that it would be universal in nature [1]. Any change in the code would result in global changes in the proteome and likely be cataclysmic to the fitness of the organism; consequently, the code once established would be immutable and fixed. Subsequent discoveries of organisms and organelles that reassign codons on a genome-wide level, as well as a growing number of examples of *cis*- and *trans*-acting signals that alter decoding of select codons in specific mRNAs, and the addition of selenocysteine (Sec) and pyrrolysine to the list of 20 cotranslationally

M.T. Howard (✉)
Department of Human Genetics, University of Utah, Salt Lake, UT 84112, USA
e-mail: mhoward@genetics.utah.edu

J.N. Gonzales-Flores • P.R. Copeland
Department of Molecular Genetics, Microbiology and Immunology,
UMDNJ-Robert Wood Johnson Medical School, Piscataway, NJ 08854, USA

D.L. Hatfield et al. (eds.), *Selenium: Its Molecular Biology and Role in Human Health*,
DOI 10.1007/978-1-4614-1025-6_3,

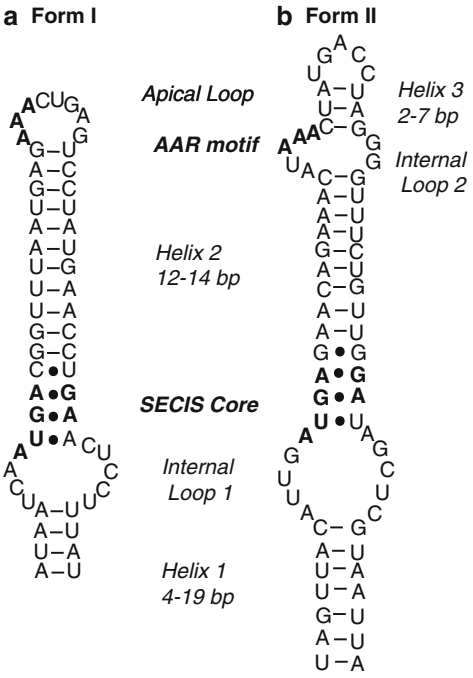

Fig. 3.1 SECIS element diagram. The two types of SECIS elements are exemplified here by the first (**a**) and second (**b**) SECIS elements in the selenoprotein P 3′ UTR. The Form II SECIS element is characterized by an additional short helix (helix 3) above the conserved AAR motif

inserted amino-acids dispelled the notion of a fixed universal genetic code (reviewed in [2–4]). The genetic code is evolving. The standard rules of decoding in most organisms dictate that the ribosome will terminate translation upon encountering any one of the three stop codons, UAA, UAG, or UGA. It may be no coincidence that gene-specific redefinition of codons occurs most often with stop codons. Termination codons occur only once per gene; consequently, assigning a dual meaning to these codons in an organism would minimize the impact on existing coding sequences and the resulting proteome. Further, impact on global protein expression is reduced by the evolution of *cis*-acting sequence elements and *trans*-acting factors that direct stop codon redefinition to select mRNAs in an organism. To date there are two *trans*-acting factors and one *cis*-acting RNA sequence that are required for converting specific UGA codons from signaling termination to one that signals Sec incorporation. In 1991 it was reported that type 1 iodothyronine deiodinase (DIO1) 3′ UTR contained a sparsely conserved sequence, termed the Sec insertion sequence (SECIS) element, that was required for the translation of full-length DIO1 both in vitro and in injected *Xenopus* oocytes [5]. Subsequent studies clarified that the SECIS element consisted of three small regions of conservation surrounded by an overall similar topology as shown in Fig. 3.1 [6–10]. In 1997 a protein was shown to specifically bind to the conserved AUGA motif, and in 2000 this protein was identified as a novel factor that was required for the Sec incorporation reaction in vitro. Because a previous study had reported a nonspecific SECIS-binding activity that they termed SECIS-binding protein (SBP) [11], the SBP

identified in 2000 was termed SBP2. That same year saw the identification of the Sec-specific elongation factor, eEFSec [12, 13], and there has been no further identification of factors that are required for Sec incorporation. Layered over the core of required components are other factors, both *cis* and *trans*-acting, which have been reported to modulate the Sec incorporation reaction. These include the novel SBPs discussed in Chap. 4 of this book as well as mRNA sequences in the coding regions of some selenoprotein genes that have dramatic effects on the efficiency of Sec incorporation both in vitro and in vivo, which are discussed in detail below.

In this chapter, we review current developments in our understanding of the Sec incorporation mechanism with an emphasis on the known roles of each of the required factors, the *cis*-acting elements that facilitate Sec incorporation, and describe potential in vitro systems that could prove to be essential in further deciphering the molecular mechanism(s) controlling UGA redefinition and Sec incorporation.

3.2 The Core Factors

3.2.1 SECIS Elements

The basic structure and function of SECIS elements has been extensively reviewed in the previous editions of this book [14, 15] and elsewhere [16–18]. The importance of SECIS function in vivo was underscored recently when it was found that a mutation in the Selenoprotein N gene (*SEPN1*) that caused SEPN1-related myopathy was mapped to the SECIS element. The AUGA → ACGA mutation was severe enough to completely eliminate detectable SEPN1 expression as determined by immunoblot of patient samples [19]. Recent investigations into SECIS function have shown dramatic variability in the efficiency with which different SECIS elements function in the context of a reporter gene where the difference in Sec insertion efficiency between high- and low-efficiency SECIS elements spanned more than three orders of magnitude [20]. This effect was shown to be primarily dictated by the sequence of the SECIS core, helix 2, and the apical loop (see Fig. 3.1) Although there was a generally similar trend across SECIS elements, the cell type in which the experiment was done, or the use of a cell free in vitro translation system (rabbit reticulocyte lysate), had a significant effect on which SECIS elements were strongest and the total difference between high- and low-efficiency sequences. Interestingly, the differences in SECIS efficiency could not be attributed to differential SBP2 binding. Together these findings strongly suggest the existence of other cell-specific factors that regulate the efficiency of Sec incorporation in vivo. One caveat to this study, however, is that these SECIS elements were taken out of their natural contexts, so any stabilizing or modulating effects mediated by the surrounding 3′ UTR sequence or more distant *cis*-acting elements were not considered. Indeed, the idea that a so-called “efficiency factor” could be a *cis*-acting RNA sequence is a possibility discussed in detail below.

The SECIS element has also recently been shown to be a platform for complex formation as it was found to be required for the association of two independently expressed SBP2 domains (see below) and for recruitment of eEFSec to form a stable SBP2/SECIS/eEFSec complex [21]. This finding provides valuable mechanistic insight into the process of Sec incorporation and implicates SBP2 and the SECIS in driving an eEFSec conformation that can be recognized by the ribosome. Future work designed to decipher the conformational changes that accompany complex formation will be required to build a molecular model of the events required for Sec incorporation.

In the 20 years that have passed since the discovery of the SECIS element, one of its features remains a mystery. The conserved sequence in the apical loop, the AAR motif, is absolutely required for Sec incorporation, but its function remains elusive. To date, no AAR sequence-specific binding proteins have been identified, and this is not for a lack of attempts (P.R. Copeland, unpublished observation). The fact that two SECIS elements (SelM and SelO) have C residues in place of the AAR motif reduces the likelihood that a sequence-specific RNA-binding protein is interacting with the terminal loop [22, 23]. This gives support to an as-yet unsubstantiated model where the terminal loop may be playing a role directly on the ribosome, likely assisting with eEFSec binding to the ribosomal A site and/or with Sec-tRNA^{Sec} accommodation.

3.2.2 SECIS-Binding Protein 2

SBP2 possesses three biochemically distinct domains (Fig. 3.2), the C-terminal half of the protein comprised of a Sec incorporation domain (SID), which is evolutionarily unique, and an RNA-binding domain (RBD), which is a member of the L7Ae RNA-binding family of proteins that interact with a variety of RNAs, specifically at kink turn motifs such as those found in rRNA, snRNA, and SECIS elements. These two domains, the SID and RBD, are sufficient for all three of the known functions of SBP2: SECIS element binding, ribosome binding, and Sec incorporation. The N-terminal half of the protein is also evolutionarily unique and has no known function, but presumably serves a regulatory role since SBP2 in many organisms (e.g., insects, protists, and worms) lacks the N-terminal domain entirely [24]. Although much is known about the RNA-binding properties of SBP2, little progress has been made in determining how it works to promote Sec incorporation. One model proposes that SBP2 stably binds ribosomes and upon SECIS element binding promotes a conformational change in the ribosome that allows eEFSec binding at the expense of the translation termination factor eRF1 [25]. Two findings have recently posed significant challenges to this model. First, it was found that the SID and RBD domains, when expressed as separate proteins, are fully active in Sec incorporation in vitro, but they do not stably interact with ribosomes, providing fairly clear evidence that the stable ribosome-binding activity is not required for Sec incorporation [21]. Second, it was found that SBP2 forms a stable SECIS-dependent, Sec-tRNA^{Sec}-independent complex

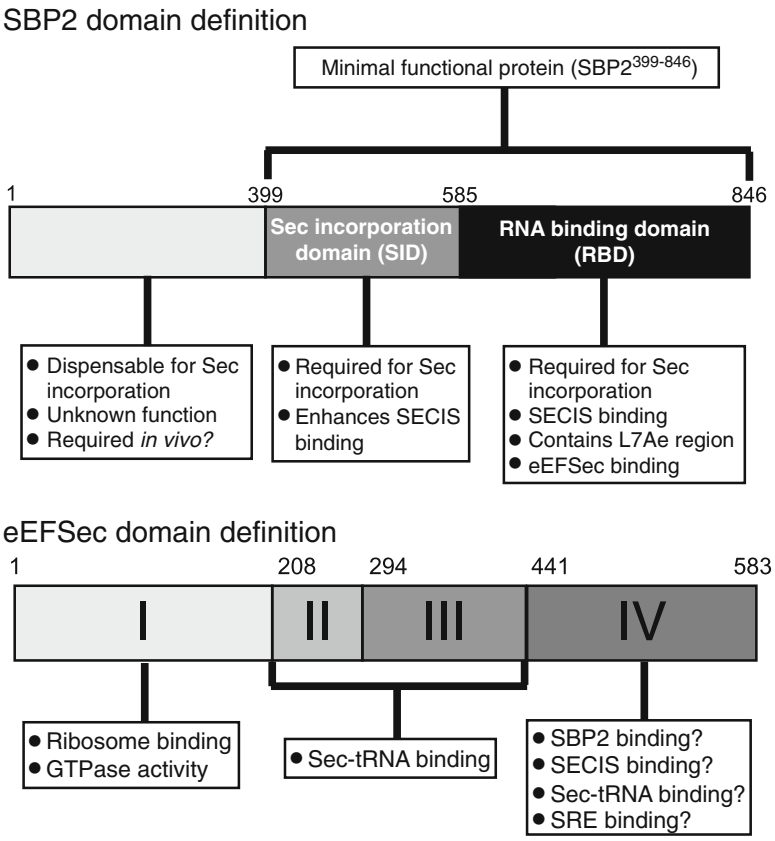

Fig. 3.2 SBP2 and eEFSec domain structures. Known or putative functions for each domain are indicated

with eEFSec, suggesting that SBP2 may act directly on eEFSec conformation upstream of the actual incorporation event, perhaps prior to ribosome binding [21].

Progress has been made in determining how the SID in SBP2 may function with respect to the RBD. Using the SID and RBD domains as separate proteins, it was shown that they form a SECIS-dependent complex and that the SID is involved in enhancing the SECIS-binding activity of the RBD. Since the SID does not exhibit a stable binding activity toward the SECIS element, it follows that the SID and RBD make direct contacts that are conformationally driven by SECIS binding. Interestingly when the conserved SID residues IILKE$^{526–530}$ (rat numbering -NP_076492.1) are mutated to alanine, the enhancement of RBD SECIS-binding activity is preserved, but the stable interaction between the SID and RBD is lost [21]. Surprisingly, the IILKE$^{526–530}$ mutant protein that contains both the SID and RBD domains completely lacks SECIS-binding activity, indicating that the presence of the mutated sequence is "blocking" access to the SECIS element. This finding aligns nicely with a separate study showing that the IILKE$^{526–530}$ residues are a determinant for SECIS specificity.

In that case it was found that the corresponding residues in the *Drosophila* version of SBP2 (SVRVY), which are not highly conserved, are necessary for binding to the form 2 SECIS elements found in *Drosophila* [26]. Based on these findings, this region of the SBP2 SID has clearly been identified as an indirect regulator of SECIS element affinity, making it a likely focus for determining the molecular basis for selective SECIS binding.

Beyond the SECIS and ribosome-binding activities of SBP2, two regions have been identified as being critical for an as-yet unidentified function that is likely proximal to the Sec incorporation event (i.e., driving a conformational change in eEFSec and/or the ribosome). One of these lies just upstream of the IILKE$^{526–530}$ sequence and was identified when the PLMKK$^{504–508}$ sequence was mutated to alanines. This version of SBP2 has neither SECIS nor ribosome-binding defects and yet is completely unable to support Sec incorporation. An identical phenotype was found when a sequence at the N-terminus of the RBD (FQ$^{648–649}$) was mutated, suggesting that these two regions form a single functional interface or that they perform two separate but essential functions. The latter scenario would fit with a model where SBP2 is required to promote conformational changes in both the ribosome and eEFSec either simultaneously or even sequentially.

3.2.3 Sec-Specific Elongation Factor

The Sec-specific elongation factor (eEFSec) in eukaryotes is a GTPase that is the exclusive carrier of the Sec-tRNA^{Sec} [12, 13]. The binding affinity of eEFSec to GTP is approximately three times higher than to GDP, thus it may not require a guanine nucleotide exchange factor (GEF). As for its tRNA-binding properties, eEFSec can only interact with Sec-tRNA^{Sec}, but not the serylated-tRNA^{Sec} precursor or the canonical aminoacyl-tRNAs [12, 13].

eEFSec consists of four domains (Fig. 3.2). Leibundgut et al. reported the complete crystal structure of archaeal EFSec [27]. This revealed a “chalice-like” structure consisting of Domains I, II, and III forming the cup of the chalice, whereas Domain IV is separated from the first three domains and forms the base of the chalice. The function of each eEFSec domain remains untested, but based on sequence conservation, the first three domains in eEFSec may have similar properties to the eukaryotic translation elongation factor, eEF1A. The elongation factor eEF1A is the main protein carrier that delivers all canonical aminoacyl-tRNAs to the ribosomal A-site during protein synthesis and is composed of three domains [28]. Domain I is required for GTPase activity and ribosomal factor-binding site interaction. Domain II is mainly involved in aminoacyl-tRNA binding, and Domain III is proposed to be involved in interactions with the T arm of aminoacyl-tRNAs (reviewed in [29]). Recently, it was shown that Sec-tRNA^{Sec} contains an anti-determinant for eEF1A binding at the base of the T arm, thus providing evidence of Domain III importance in tRNA recognition [30]. Domain IV in eEFSec, which is not present in eEF1A,

was proposed to be involved in interactions with SBP2 [31] and the "extra arm" of the Sec-tRNA^{Sec} [32].

eEFSec and SBP2 were first demonstrated to interact by a co-immunoprecipitation experiment in mammalian cells, forming a complex that was RNase sensitive [13]. Further studies in mammalian cells showed that eEFSec and SBP2 interactions were further enhanced by overexpression of the tRNA^{Sec} gene [31]. However, it was later demonstrated in a pure component system that SBP2 could not form a complex with eEFSec/GTP/Sec-tRNA^{Sec} and instead it caused Sec-tRNA^{Sec} release from eEFSec, although, the addition of a SECIS element was not tested [33]. Recently, Donovan et al. demonstrated in a native gel shift assay that eEFSec can form a complex with SBP2 in the presence of a SECIS element without the requirement of Sec-tRNA^{Sec} and/or GTP nucleotide [21]. This report also showed that the RBD of SBP2 and the SECIS element are sufficient for eEFSec recruitment. This suggests that the RBD of SBP2 and the SECIS element could form together a binding interface that is favorable for eEFSec interaction. Thus, taking together the in vivo and in vitro data so far, complex formation between eEFSec and SBP2 is driven by the SECIS element and possibly stabilized in the presence of Sec-tRNA^{Sec}. Putting together what is known about the function of SBP2, eEFSec, and the SECIS element, it seems likely that they form a stable and "active" Sec-tRNA^{Sec} delivery complex that can bind the ribosomal factor-binding site. Some elements in the complex (e.g., the SECIS loop) may play roles downstream, e.g., in Sec-tRNA^{Sec} accommodation into the A-site. This, however, is likely only half the story as the ribosome may also need to be primed to accept this active complex as discussed in Chap. 5.

eEFSec must be denied general access to the ribosomal A-site to prevent the Sec-tRNA^{Sec} from acting as a suppressor tRNA. Unlike eEFSec, eEF1A can obtain ribosomal A-site access without the requirement of additional factors. Near the ribosomal A-site is found the elongation factor-binding site that is composed of two elements: (1) the GTPase associating center (GAC) and (2) the sarcin-ricin loop (SRL). The GAC and SRL main function is to activate GTP hydrolysis in elongation factors, such as eEF1A and eEF2 [34]. Hüttenhofer et al. reported that the bacterial version of eEFSec, SelB, obtains ribosome-dependent GTP hydrolysis only when a bacterial SECIS element is added [35]. Their conclusion was that the bacterial SECIS element, which resides immediately downstream of the UGA-Sec codon rather than in the 3′ UTR, induces a conformational change within SelB to promote functional interactions with the ribosome. Indeed, eEFSec could be using a similar mechanism where SBP2 and the SECIS element act in concert to directly modify the eEFSec/GTP/Sec-tRNA^{Sec} ternary complex. In addition to the basal activity of the core complex and by analogy to the bacterial system, *cis*-acting RNA elements located immediately downstream of eukaryotic UGA-Sec codons could be interacting with the ternary complex to induce conformational changes that enhance the accommodation of Sec-tRNA^{Sec} (see Sect. 3.3). Further investigations are required to clarify the activating mechanism(s) that promotes functional interactions between eEFSec and the ribosome to allow Sec incorporation into nascent peptides during decoding of in-frame UGA-Sec codons.

3.3 *Cis*-Acting Elements Affecting Sec Incorporation

Although the efficiency of Sec incorporation is not known in vivo, the observation that termination appears to be the predominant event for selenoprotein genes with a single Sec-encoding UGA codon [5, 13, 36–39] has widely been interpreted as evidence for competition between termination of translation by eRF1/3 and Sec-tRNA^{Sec} decoding of the stop codon. Although the SECIS element itself can have a large impact on Sec incorporation efficiency (as discussed above in Sect. 3.2.1 and [40]), UGA-Sec sequence contexts that either favor eEFSec delivery of the Sec-tRNA^{Sec} to the ribosome and Sec incorporation or antagonize release factor catalyzed termination of translation are expected to increase redefinition efficiency.

3.3.1 Sequence Context Effects on Termination Efficiency

One factor known to effect termination efficiency is the identity of the stop codon, where in mammals the order is: [UAA > UAG > UGA]. The differences likely stem from the nature of direct contacts between the eukaryotic termination factor (eRF1) and the stop codon that induce conformational changes in the ribosome complex required to trigger peptide hydrolysis [41, 42]. In addition, multiple studies have highlighted the importance of local stop codon sequence context, especially the two codons preceding and the base following the stop codon, in determining termination efficiency [43–46]. However, these *cis*-acting sequences alone are not sufficient to predict termination efficiency [47], indicating that a larger sequence context is involved. The nature of this effect is likely to be complex and may include RNA secondary structure as well as the primary sequence, and even the composition of the nascent peptide chain in the exit tunnel of the ribosome.

3.3.2 Sequence Context Effects on Sec Incorporation

In contrast to stop codon readthrough due to near-cognate tRNA decoding, redefinition of UGA-Sec codons to encode Sec requires recruiting the eEFSec ternary complex for cognate decoding by Sec-tRNA^{Sec}. Evidence for an effect of UGA-sequence context on the efficiency of Sec insertion efficiency initially came from studies where nucleotides adjacent to the UGA-Sec codons for the iodothyronine deiodinase [48, 49] and PHGPx [39] were varied. In most cases, readthrough efficiency was found to be increased in contexts that resulted in inefficient termination. A thorough analysis of readthrough efficiency of the 10 UGA-Sec codons encoded by the rat Selenoprotein P (*SEPP1*) gene (with each UGA-Sec codon containing the surrounding native 24 nucleotide sequence context) in rabbit reticulocyte lysate revealed a lack of correlation between Sec incorporation efficiency and the nucleotide immediately following the UGA-Sec codon [50]. This study illustrated that

the sequence context effect is complex and supports a model in which a larger *cis*-acting sequence context determines Sec insertion efficiency. Here, it was proposed that this larger context may work together with Sec incorporation factors to determine readthrough efficiency.

Further evidence for an extended context effect on Sec incorporation efficiency comes from phylogenetic analysis demonstrating the potential for stable and conserved RNA structures located downstream of the UGA-Sec codons in a subset of selenoprotein genes [51]. In support of the importance of these elements, the same RNA secondary structures were independently identified in two selenoprotein genes, *SEPN1* and *SELT*, in a genome-wide search for deeply conserved functional RNA structures [52]. Detailed experimental analysis of the larger sequence context surrounding the *SEPN1* UGA-Sec codon demonstrated an effect on Sec incorporation efficiency in vitro [51, 53] and in vivo [54] (also see Chap. 22). This *cis*-acting element, designated the Sec codon redefinition element (SRE), consists of upstream sequences and a highly conserved stem-loop structure that starts six nucleotides downstream of the UGA-Sec codon. Using reporter assays in cultured mammalian cells, the *SEPN1* SRE is sufficient to cause high level (4–6%) readthrough of UGA and UAG codons in mammalian cells, which was not dependent on the presence of the 3′ UTR SECIS element [51, 53]. When the *SEPN1* SECIS element was included in the 3′ UTR, the SRE was not required for readthrough but had a significant stimulatory effect. Experiments in rabbit reticulocyte lysates provided direct evidence that the SRE stimulates Sec incorporation, rather than near-cognate tRNA decoding of the UGA codon [53] suggesting it may play a direct role in recruiting or ribosomal accommodation of the Sec-tRNA^{Sec}. In contrast to the cell-based model originally used, SRE stimulation of readthrough was specific for UGA codons and required both SBP2 supplementation and the 3′ UTR SECIS element. Importantly, by supplementing the rabbit reticulocyte lysate with 75-Se labeled Sec-tRNA^{Sec}, it was also shown that the SRE increased incorporation of Sec into the full-length product.

3.3.3 Sequence Context Effects from a Distance

Evidence that distant (non-SECIS) *cis*-acting elements can alter UGA redefinition was recently discovered in selenoprotein mRNAs from the ciliate *Euplotes crassus* [55]. Several genes were identified that contain multiple UGA codons. *E. crassus* has the requisite Sec incorporation machinery as well as tRNAs capable of decoding UGA codons as cysteine. The *Euplotes* thioredoxin reductase 1 (eTxnrd1) gene contains seven UGA codons. Transfection experiments in HEK293 cells and mass spectrometry analysis of the native protein purified from *E. crassus* revealed that the first six UGA codons are decoded as cysteine and only the final UGA codon in the penultimate codon position was decoded as Sec. Replacing the eTxnrd1 3′ UTR with the 3′ UTR of *SELT* from *Toxoplasma* relaxed the positional requirement allowing insertion of Sec at upstream UGA codons. The authors propose a model in which the 3′ UTR from eTxnrd1 contains a *cis*-acting RNA structure that prevents

the SECIS element from accessing the ribosome during decoding of upstream UGA codons. The contribution of the local sequence context at each UGA codon was not examined in this study.

Collectively, these results clearly demonstrate that an extended UGA-Sec sequence context and even distant *cis*-acting elements can affect both termination and Sec incorporation efficiency. It is unclear if these *cis*-acting elements share common mechanisms with *cis*-acting elements stimulating near-cognate tRNA decoding of stop codons as the mechanism(s) by which these elements act remains uncertain. The intriguing possibility that the *cis*-acting elements in selenoprotein mRNAs interact directly with components of the Sec insertion machinery or the ribosome to facilitate decoding by Sec-tRNA^{Sec} is suggested by several lines of evidence but requires further study.

3.3.4 Mechanism of Sec Incorporation in Transcripts with Multiple UGA Codons

Several selenoprotein genes have now been identified with the potential to encode more than one Sec residue. These include an alternatively spliced isoform of *SEPN1* (alternate transcript 1) [56], selenoprotein L [57], and *SEPP1*. Surprisingly, the total number of UGA-Sec codons in *SEPP1* ranges among species from 10 in humans to as many as 28 in sea urchin [58]. As demonstrated in rabbit reticulocyte lysate, the efficiency of Sec incorporation at each UGA-Sec codon when examined alone and in its native context varies between ~5 and 25% [50]. Yet, purification of selenoprotein P from plasma reveals the majority of protein to be full-length with several prematurely UGA-terminated species having been identified [59, 60]. In contrast to the model whereby Sec incorporation competes inefficiently with termination, the production of full-length protein from these messages would seem to demand highly efficient Sec incorporation due to the compounding effect of termination at each UGA-Sec codon.

It has been suggested that *SEPP1* may utilize a special mechanism for Sec incorporation due to the exceptional number of UGA-Sec codons and the observation that *SEPP1* mRNAs are unique in having two conserved 3′ UTR SECIS elements. In one model [61], it is proposed that each of the two SECIS elements in the *SEPP1* RNA has different functions with the distal SECIS element serving to incorporate Sec at the first UGA inefficiently, acting as a checkpoint for Sec incorporation factors, and the proximal SECIS dedicated to redefinition of the remaining UGA codons with high efficiency. It was concluded that the *SEPP1* gene has evolved unique properties to accommodate the incorporation of multiple Secs into one polypeptide.

A recent study examining Sec incorporation efficiency in messages containing a subset of the *SEPP1* UGA-Sec codons suggests that the ability to incorporate Sec with high efficiency may not be a unique feature of the *SEPP1* mRNA but rather an intrinsic property of Sec incorporation [62]. In this study, it was found that

incorporation of Sec was inefficient at a first UGA-Sec codon but increased by roughly one order of magnitude at downstream UGA-Sec codons. The observed increase in Sec incorporation "processivity" was not unique to *SEPP1* SECIS, as replacing the two *SEPP1* SECIS elements with single SECIS elements derived from other selenoprotein genes revealed similar results. A modified model was proposed in which the SECIS complex is loaded onto the ribosome prior to, or during, decoding of the first UGA-Sec codon. Once assembled, the ribosome proceeds to the next UGA-Sec codon reprogrammed for highly efficient Sec incorporation. A key observation in this study was that the efficiency of Sec incorporation at each UGA-Sec codon in the message was dependent on the identity and relative strength of the SECIS in the 3′ UTR. This was interpreted as evidence for a continued interaction of the SECIS element with the ribosome at each UGA-Sec codon. Consequently, it was proposed that the SECIS element and associated factors track with the ribosome following Sec insertion at the first UGA-Sec codon. The high level of termination at the first UGA-Sec codon could then be explained by the preceding ribosomes encountering the first UGA-Sec codon without having access to the SECIS element and the *trans*-acting factors associated with the ribosome engaged in decoding the remainder of the open reading frame. Under this model, ribosomes that have recruited the Sec incorporation machinery decode UGA-Sec codons as Sec with high efficiency and termination at UGA codons is inefficient.

3.4 Putting It All Together: In Vitro Reconstitution

In this chapter, we have provided a "bottom up" perspective on Sec incorporation, describing the factors, *cis*-sequences, and events most proximal to the actual Sec-tRNA^{Sec} delivery event (see Fig. 3.3). One of the major hurdles in definitively determining the core mechanism of Sec incorporation is creating a system in which Sec incorporation can be reconstituted from purified components. The use of rabbit reticulocyte lysates has been a valuable intermediate in this endeavor as they are naturally devoid of endogenous SBP2. Two recent studies highlight the utility of rabbit reticulocyte lysates in not only helping to decipher the core mechanism but also reproducing results obtained in living cells [20, 63]. Our efforts to find or create a similar lysate that is devoid of both SBP2 and eEFSec have thus far been unsuccessful, but current efforts to make home-made lysates from eEFSec-null *Drosophila* embryos will likely yield favorable results and provide an alternative intermediate system that will allow significant progress to be made with regard to eEFSec function (P.R. Copeland, unpublished results). The major hurdle in building a completely reconstituted system is that in vitro reconstitution of eukaryotic translation initiation is inefficient and technically challenging. One potential way to circumvent this is to use a translation elongation system that bypasses the initiation phase by the use of an internal ribosome entry site (IRES) from the cricket paralysis virus [64]. This *cis*-acting sequence allows for a complete bypass of translation initiation as it

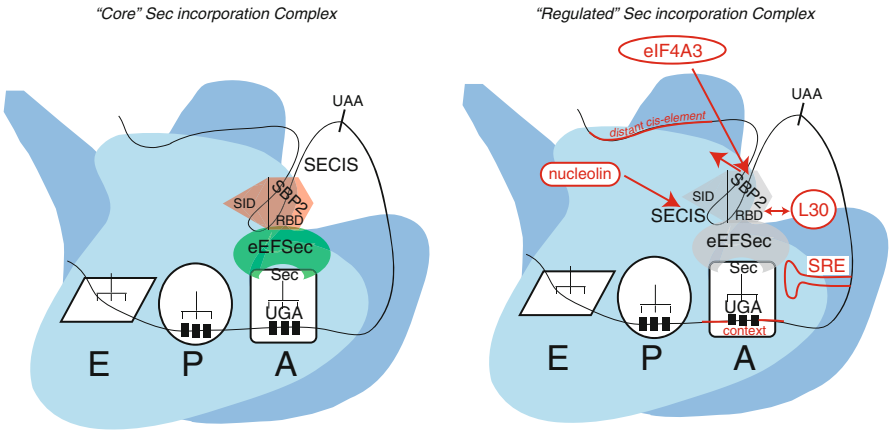

Fig. 3.3 Models of Sec incorporation. The *left panel* shows a Sec incorporation event mediated by the factors known to be required, while the *right panel* shows the factors (*red*) that may impact the efficiency and/or processivity of the Sec incorporation reaction. These factors include the *cis*-elements discussed in the text such as the SRE, codon context and distant 3′ UTR-based elements as well as *trans*-factors discussed in Chap. 4 such as nucleolin, ribosomal protein L30, and eukaryotic initiation factor A3 (eIF4A3)

directly recruits 80 S ribosomes. One potential caveat of such a system is that there may be a role for translation initiation factors in supporting Sec incorporation. A recent study has addressed this by determining that CrPV IRES-driven Sec incorporation is possible in vitro, albeit at a slight but consistently lower efficiency [65]. Ultimately, such a system will allow detailed molecular interaction studies based on fluorescence resonance energy transfer (FRET) as well as the ability to create intermediate Sec incorporation complexes that will reveal the steps required for the Sec incorporation event and the role that newly identified *cis*- and *trans*-acting factors play in modulating this process. One clear example of how this system may be utilized is in carefully dissecting the interplay between translation termination and Sec incorporation. By altering the ratios of Sec incorporation factors and termination factors and assessing the ability of an active Sec incorporation complex to form on purified ribosomes in the presence or absence of eRF1, one could gain clear insight into how these two competing processes work in concert to provide regulated synthesis of selenoproteins. This is just one of countless experiments that will reveal the inner workings of the Sec incorporation machinery, ultimately shedding light on how one might regulate this process in vivo.

Acknowledgements This work was supported by National Institutes of Health Grants to MTH and PRC.

References

1. Crick FH (1968) J Mol Biol 38:367
2. Ambrogelly A, Palioura S, Söll D (2007) Nat Chem Biol 3:29
3. Namy O, Rousset JP, Napthine S et al (2004) Mol Cell 13:157
4. Atkins JF, Baranov PV, Fayet O et al (2001) Cold Spring Harb Symp Quant Biol 66:217
5. Berry MJ, Banu L, Chen YY et al (1991) Nature 353:273
6. Hubert N, Walczak R, Sturchler C et al (1996) Biochimie 78:590
7. Martin GW, Harney JW, Berry MJ (1996) RNA 2:171
8. Walczak R, Westhof E, Carbon P et al (1996) RNA 2:367
9. Berry MJ, Banu L, Harney JW et al (1993) EMBO J 12:3315
10. Hill KE, Lloyd RS, Burk RF (1993) Proc Natl Acad Sci USA 90:537
11. Hubert N, Walczak R, Carbon P et al (1996) Nucleic Acids Res 24:464
12. Fagegaltier D, Hubert N, Yamada K et al (2000) EMBO J 19:4796
13. Tujebajeva RM, Copeland PR, Xu XM et al (2000) EMBO Rep 1:158
14. GW Martin, MJ Berry 2001 Selenium: its molecular biology and role in human health DL Hatfield (ed) Kluwer Academic Publishers Boston pp 45
15. C Allmang, A Krol 2006 Selenium: its molecular biology and role in human health DL Hatfield, MJ Berry, VN Gladyshev (eds) Kluwer Academic Publishers Boston pp 51
16. Papp LV, Lu J, Holmgren A et al (2007) Antioxid Redox Signal 9:775
17. Squires JE, Berry MJ (2008) IUBMB Life 60:232
18. Donovan J, Copeland PR (2010) Antioxid Redox Signal 12:881
19. Allamand V, Richard P, Lescure A et al (2006) EMBO Rep 7:450
20. Latrèche L, Jean-Jean O, Driscoll DM et al (2009) Nucleic Acids Res 37:5868
21. Donovan J, Caban K, Ranaweera R et al (2008) J Biol Chem 283:35129
22. Korotkov KV, Novoselov SV, Hatfield DL et al (2002) Mol Cell Biol 22:1402
23. Kryukov GV, Castellano S, Novoselov SV et al (2003) Science 300:1439
24. Donovan J, Copeland PR (2009) BMC Evol Biol 9:229
25. Caban K, Copeland PR (2006) Cell Mol Life Sci 63:73
26. Takeuchi A, Schmitt D, Chapple C et al (2009) Nucleic Acids Res 37:2126
27. Leibundgut M, Frick C, Thanbichler M et al (2005) EMBO J 24:11
28. Andersen GR, Valente L, Pedersen L et al (2001) Nat Struct Biol 8:531
29. Noble CG, Song H (2008) Cell Mol Life Sci 65:1335
30. Bouzaidi-Tiali N, Aeby E, Charrière F et al (2007) EMBO J 26:4302
31. Zavacki AM, Mansell JB, Chung M et al (2003) Mol Cell 11:773
32. Itoh Y, Chiba S, Sekine SI et al (2009) Nucleic Acids Res 37:6259
33. Small-Howard A, Morozova N, Stoytcheva Z et al (2006) Mol Cell Biol 26:2337
34. Sergiev PV, Kiparisov SV, Burakovsky DE et al (2005) J Mol Biol 353:116
35. Hüttenhofer A, Böck A (1998) Biochemistry 37:885
36. Suppmann S, Persson BC, Böck A (1999) EMBO J 18:2284
37. Jung JE, Karoor V, Sandbaken MG et al (1994) J Biol Chem 269:29739
38. Mehta A, Rebsch CM, Kinzy SA et al (2004) J Biol Chem 279:37852
39. Nasim MT, Jaenecke S, Belduz A et al (2000) J Biol Chem 275:14846
40. Novoselov SV, Lobanov AV, Hua D et al (2007) Proc Natl Acad Sci USA 104:7857
41. Freistroffer DV, Kwiatkowski M, Buckingham RH et al (2000) Proc Natl Acad Sci USA 97:2046
42. Youngman EM, McDonald ME, Green R (2008) Annu Rev Microbiol 62:353
43. Howard MT, Shirts BH, Petros LM et al (2000) Ann Neurol 48:164
44. Manuvakhova M, Keeling K, Bedwell DM (2000) RNA 6:1044
45. Mottagui-Tabar S, Björnsson A, Isaksson LA (1994) EMBO J 13:249
46. Namy O, Hatin I, Rousset JP (2001) EMBO Rep 2:787
47. Bidou L, Hatin I, Perez N et al (2004) Gene Ther 11:619
48. Grundner-Culemann E, Martin GW, Tujebajeva R et al (2001) J Mol Biol 310:699

49. McCaughan KK, Brown CM, Dalphin ME et al (1995) Proc Natl Acad Sci USA 92:5431
50. Gupta M, Copeland PR (2007) J Biol Chem 282:36797
51. Howard MT, Aggarwal G, Anderson CB et al (2005) EMBO J 24:1596
52. Pedersen JS, Bejerano G, Siepel A et al (2006) PLoS Comput Biol 2:e33
53. Howard MT, Moyle MW, Aggarwal G et al (2007) RNA 13:912
54. Maiti B, Arbogast S, Allamand V et al (2008) Hum Mutat 30:411
55. Turanov AA, Lobanov AV, Fomenko DE et al (2009) Science 323:259
56. Moghadaszadeh B, Petit N, Jaillard C et al (2001) Nat Genet 29:17
57. Shchedrina VA, Novoselov SV, Malinouski MY et al (2007) Proc Natl Acad Sci USA 104:13919
58. Lobanov AV, Hatfield DL, Gladyshev VN (2008) Genome Biol 9:R62
59. Himeno S, Chittum HS, Burk RF (1996) J Biol Chem 271:15769
60. Ma S, Hill KE, Caprioli RM et al (2002) J Biol Chem 277:12749
61. Stoytcheva Z, Tujebajeva RM, Harney JW et al (2006) Mol Cell Biol 26:9177
62. Fixsen SM, Howard MT (2010) J Mol Biol 399:385
63. Bubenik JL, Driscoll DM (2007) J Biol Chem 282:34653
64. Pestova TV, Hellen CU (2005) Methods 36:261
65. Donovan J, Copeland PR (2010) J Mol Biol 400:659