Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

Synthesis and semisynthesis of selenopeptides and selenoproteins

Jun Liu[1], Rujin Cheng[2] and Sharon Rozovsky[2]

The versatile chemistry of the genetically encoded amino acid selenocysteine (Sec) is employed in Nature to expand the reactivity of enzymes. In addition to, its role in biology, Sec is used in protein engineering to modify folding, stability, and reactivity of proteins, to introduce conjugations and to facilitate reactions. However, due to limitations related to Sec's insertion mechanism in Nature, much of the production of Sec containing peptides and proteins relies on synthesis and semisynthesis. Here, we review recent advances that have enabled the assembly of complicated selenoproteins, including novel uses of protecting groups for solid phase peptide synthesis, rapid selenoester driven chemical ligations and versatile expressed protein ligations.

Addresses
[1] University of California, San Francisco, Department of Pharmaceutical Chemistry, 555 Mission Bay Blvd. South, San Francisco, CA 94158, United States
[2] University of Delaware, Department of Chemistry and Biochemistry, Newark, DE 19716, United States

Corresponding authors: Liu, Jun (jun.liu@ucsf.edu), Rozovsky, Sharon (rozovsky@udel.edu)

Current Opinion in Chemical Biology 2018, **46**:41–47

This review comes from a themed issue on **Synthetic biomolecules**

Edited by **Richard J Payne** and **Nicolas Winssinger**

https://doi.org.10.1016/j.cbpa.2018.04.008

Introduction

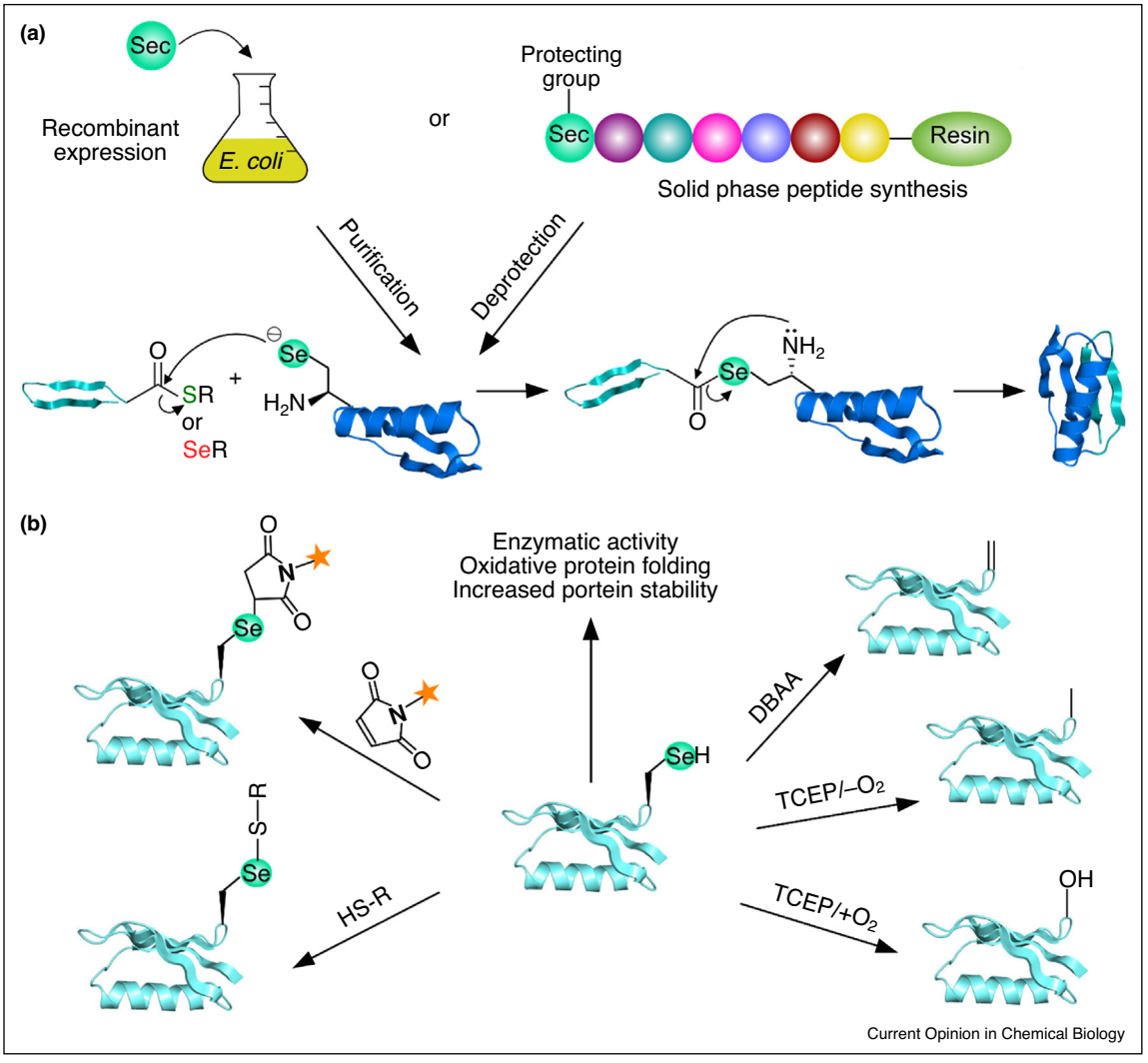
Selenium is used in Nature in nucleic acids, sugars, small compounds, and proteins, but its most significant utilization is in the form of the genetically encoded amino acid selenocysteine (Sec). Sec and Cys share many physiochemical properties, but Sec is a better nucleophile and electrophile because of selenium's electronic structure [1,2]. Another marked difference to their sulfur containing counterparts is the much lower redox potential of selenocompounds [1,3]. Thus Sec and Cys can not only play similar roles in proteins, but because Sec is more reactive it broadens the range of chemical reactions that are facilitated by enzymes [1,2]. Accordingly, the rapidly growing interest in proteins and peptides containing selenium originates from two overlapping areas: Firstly, the study of selenoproteins in their own right, as proteins with unique and specialized cellular roles [4], and secondly from protein engineering for which Sec's particular properties offer unique chemical approaches toward the production, manipulation, and augmentation of macromolecules.

For such protein engineering applications, it is particularly beneficial that Sec, due to its specific chemistry, can be manipulated in ways that leave any of the 20 canonical amino acids unaffected. In that sense Sec is often considered a 'hetero amino acid' in the protein and its high reactivity can be exploited to diversify and control peptide and protein chemistry (Figure 1) [5]. For example, Sec can be employed directly for site-specific protein labeling and conjugation [6] or it can be selectively converted into dehydroalanine [7] — an electrophilic handle for natural posttranslational modification or other bioconjugations [8,9]. In addition, selenoproteins and selenopeptides have been used to probe the kinetics and dynamics of protein's oxidative folding [10], as well as to alter chemical properties of proteins [11]. Finally, ^{77}Se-Sec is utilized to study proteins by NMR spectroscopy [12]. The interested reader can find more detailed discussions of selenoproteins and selenopeptides and their fascinating role in biology, biochemistry and protein engineering in Refs. [3,4,13–15].

Despite the great promise that Sec offers for protein engineering, its full potential has yet to be unlocked. One of the major challenges is the peculiar bioincorporation of Sec in Nature. Unlike the 20 canonical amino acids, which are genetically encoded by unique codons, Sec relies on dedicated ancillary proteins that enable the alternative use of the stop codon UGA for translational incorporation into the polypeptide chain [16,17]. Unfortunately, termination of transcription is thus a common side reaction when overexpressing an organism's native Sec-incorporation machinery. To overcome this limitation, novel approaches have been recently developed in which Sec-specific tRNA was engineered and the native selenium incorporation path was altered [18–25]. Despite the promise of such approaches, obtaining high yield of protein can be challenging due to low efficiency and fidelity of Sec loading. Thus this technique is not yet routinely employed for preparation of selenopeptides or selenoproteins [26].

Nevertheless, in recent years the reliable production of complex selenopeptides and selenoproteins has been

Figure 1



Selenopeptides and Selenoproteins: their synthesis, semisynthesis and applications. **(a)** Fragments with N-terminal Sec are most commonly prepared either by solid phase peptide synthesis (right) or by growing bacteria with selenocystine in the growth medium (left). In standard native chemical ligation reactions, an amide bond is formed between fragments containing a selenolate and a thioester, respectively. **(b)** Selenopeptides and selenoproteins can be utilized for a wide and diverse range of applications.

achieved by several methods, that combine automated solid phase peptide synthesis (SPPS), recombinant expression of fragments, and native chemical ligation (NCL). The latter, is a widely used synthetic method that relies on an amide-forming reaction to generate proteins from their respective fragments. Sec-mediated NCL reactions take place between a N-terminal peptide fragment with a C-terminal reactive thioester (or, as discussed below, now selenoester), and a C-terminal peptide fragment containing Sec at the N-terminus. The nucleophilic Sec attacks the N-terminal fragment and forms a transient thioester between the two fragments, which then rapidly forms the peptide bond through an intramolecular acyl Se→N shift [27–30]. Relative to Cys-mediated NCL, Sec-mediated ligations exhibit faster kinetics and are compatible with a wider pH range. Sec's high nucleophilicity facilitates fast reaction rates and its low pK_a allows for ligations at low pH to minimize undesired side reactions such as thioester hydrolysis [29]. Furthermore, following the reaction Sec can be selectively converted into the abundant amino acids Ala or Ser under mild conditions. Thus, Sec offers more choices for the ligation site [31•,32•].

Numerous articles and book chapters have surveyed the different approaches for preparing selenoproteins [5,15,26,33,34]. Here, we review recent advances in the field of synthetic and semisynthetic preparation of