Cell-Free Synthesis of Selenoproteins in High Yield and Purity for Selective Protein Tagging

Adarshi P. Welegedara,[a, d] Ansis Maleckis,[b] Ruchira Bandara,[a] Mithun C. Mahawaththa,[a] Iresha Dilhani Herath,[a] Yi Jiun Tan,[a] Angeliki Giannoulis,[c] Daniella Goldfarb,[c] Gottfried Otting,*[a] and Thomas Huber*[a]

The selenol group of selenocysteine is much more nucleophilic than the thiol group of cysteine. Selenocysteine residues in proteins thus offer reactive points for rapid post-translational modification. Herein, we show that selenoproteins can be expressed in high yield and purity by cell-free protein synthesis by global substitution of cysteine by selenocysteine. Complete alkylation of solvent-exposed selenocysteine residues was achieved in 10 minutes with 4-chloromethylene dipicolinic acid (4Cl-MDPA) under conditions that left cysteine residues unchanged even after overnight incubation. Gd^{III}–Gd^{III} distances measured by double electron–electron resonance (DEER) experiments of maltose binding protein (MBP) containing two selenocysteine residues tagged with 4Cl-MDPA-Gd^{III} were indistinguishable from Gd^{III}–Gd^{III} distances measured of MBP containing cysteine reacted with 4Br-MDPA tags.

Introduction

Site-specific chemical modifications of proteins are fundamental to protein research and required in many applications, such as immobilization of proteins on chips or the posttranslational attachment of solubility tags or spectroscopic probes for studies by fluorescence, NMR spectroscopy, or electron paramagnetic resonance (EPR).[1] EPR spectroscopy with paramagnetic tags has recently emerged as an important tool in structural biology, where double electron-electron resonance (DEER) experiments yield the distance distribution between two site-specifically attached paramagnetic tags.[2] DEER experiments are attractive, because they give access to accurate distances ranging between about 2–8 nm, which are difficult to measure by any other spectroscopic technique and particularly suited to study large conformational changes of proteins.[3] Furthermore, DEER experiments yield not only distances, but entire distance distributions.

Most often, the site-specific attachment of the paramagnetic tags is achieved by reaction with the thiol group of a cysteine (Cys) residue in the target protein, with formation of either a covalent disulfide or a thioether bond. For example, the most commonly used EPR tag, *S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1*H*-pyrrol-3-yl)methyl methanesulfonothioate (MTSL), spontaneously reacts with Cys to form a disulfide bond.[4] Cys stands out as the most nucleophilic and oxidation sensitive amino acid among the 20 canonical amino acids, enabling site-specific tagging. Tagging of site-specifically introduced Cys residues, however, is incompatible with proteins that depend on Cys in their wild-type amino acid sequence to maintain structural and functional integrity. It is therefore of interest to develop chemistries that allow site-specific tagging in a manner that leaves native Cys residues intact.

An ideal scenario would involve an unnatural amino acid that contains a paramagnetic center and can be incorporated into proteins in response to a stop codon with the help of a suppressor tRNA and a suitable aminoacyl-tRNA synthetase. Such a system has been established, but the amino acid carries the paramagnetic center at the end of a long side chain, limiting its use for accurate distance measurements.[5] Alternatively, *p*-acetylphenylalanine or *p*-azido-phenylalanine can be incorporated into proteins by similar systems and their terminal groups reacted with EPR tags.[6] This also allows the attachment of tags containing Gd^{III} ions, which are particularly suited for sensitive DEER measurements with W-band EPR instruments.[6,7] Again, these schemes generate tethers between protein and EPR-active paramagnetic center that are long compared with tags attached to Cys residues. Even though the tethers can be relatively rigid,[6] the ultimate tether length limits the utility of these amino acids in structural biology. In the context of DEER experiments, the paramagnetic centers of the tags are best positioned close to the protein backbone, so that the width of the DEER distance distribution reflects the structure and conformational variability of the protein rather than the tags.

A possible solution of these limitations could be provided by a scheme, where photocaged selenocysteine is incorporated into the target protein as an unnatural amino acid and

[a] Dr. A. P. Welegedara, R. Bandara, Dr. M. C. Mahawaththa, I. Dilhani Herath, Y. Jiun Tan, Prof. G. Otting, Prof. T. Huber
Australian National University, Research School of Chemistry
Canberra, ACT 2601 (Australia)
E-mail: gottfried.otting@anu.edu.au
t.huber@anu.edu.au
[b] Dr. A. Maleckis
Latvian Institute of Organic Synthesis
1006 Riga (Latvia)
[c] Dr. A. Giannoulis, Prof. D. Goldfarb
Department of Chemical and Biological Physics
Weizmann Institute of Science
Rehovot 76100 (Israel)
[d] Dr. A. P. Welegedara
Department of Chemistry, University of Peradeniya
Peradeniya 20400 (Sri Lanka)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202000785

subsequently deprotected to yield selenocysteine (Sec), which can be tagged like Cys while showing greater reactivity at neutral pH owing to the greater acidity of the selenol group ($pK_a = 5.2$) versus the thiol group ($pK_a = 8.5$).[8] This scheme would produce a short tether between the protein and the label, while the increased reactivity of Sec versus Cys provides selectivity. The present work explores the suitability of Sec versus Cys for the selective attachment of Gd^{III} tags.

A unique translation mechanism exists in nature, by which Sec can be incorporated in response to a UGA stop codon. It involves a series of specific enzymes, a suppressor tRNA ($tRNA^{Sec}$), and a conserved mRNA sequence (Sec insertion sequence or SECIS) in the vicinity of the target UGA codon.[8a,c] The efficiency and applicability of this mechanism has been greatly broadened by enzyme and RNA engineering, but it is still not as efficient as the incorporation of Cys.[9] Söll and co-workers also demonstrated a cell-free system to incorporate Sec into proteins genetically with yields of less than 100 ng of protein per cell-free reaction.[9f] In a much simpler approach, proteins with Sec can be made in bacteria simply by replacing Cys in the feedstock by Sec, substituting all Cys residues by Sec by misloading cysteinyl-tRNA with Sec.[9b,j] Notably, however, Sec is highly toxic and Cys released by protein turnover will compete with Sec incorporation. A 90% replacement yield of single Cys residues by Sec has been reported for a Cys-auxotroph *Escherichia coli* strain in a defined growth medium containing Sec instead of Cys.[9j]

Here we introduce a cell-free protein synthesis (CFPS) platform to synthesize proteins with Sec in high yield and purity. CFPS is ideally suited for synthesis under cytotoxic conditions and, as many biosynthetic enzymes are inactivated in cell-free extracts, we hypothesized that global substitution of Cys for Sec would be straightforward. We show for three proteins, the NS2B-NS3 protease from Zika virus (ZiPro), GB1, and *E. coli* maltose binding protein (MBP), that milligram quantities of protein containing Sec can be produced from a single milliliter of reaction mixture. Furthermore, we demonstrate that Sec can be tagged with alkylating reagents under conditions that barely affect Cys residues. The specificity of the tagging reaction is illustrated by DEER distance distributions that are indistinguishable between tags attached to the double-Sec mutant of MBP and the corresponding double-Cys mutant.

Results

Production of selenoproteins by cell-free protein synthesis

Sec versions of four protein mutants, ZiPro C80U, ZiPro V36U, GB1 Q32U and MBP T237U/T345U, where U stands for Sec, were produced in good yield by misloading cysteinyl tRNA with Sec during CFPS. As Sec is highly sensitive to oxidation and not commercially available, our experiments started from selenocystine, which had to be reduced in the reaction mixture. We used dithiothreitol (DTT) to maintain reducing conditions, as DTT has been reported to be more effective than sodium borohydrate or β-mercaptoethanol for the reduction of diselenide bonds.[10] Initial CFPS experiments using our standard amount of DTT (1 mM), however, failed to produce significant amounts of selenoproteins. A series of CFPS experiments with increasing concentrations of DTT showed that ZiPro C80U was produced in much greater yields simply by raising the concentration of DTT to 10 mM (Figure 1A). Therefore, we conducted all subsequent CFPS reactions of selenoproteins in the presence of 10 mM DTT. At this DTT concentration, 1 mL of CFPS reaction mixture yielded about 1.2 mg of ZiPro C80U with Sec. LC-MS/MS of a tryptic digest of the ZiPro C80U isolated by SDS-PAGE confirmed the presence of a Sec residue in position 80 as expected (Figure 1B). The same CFPS reaction conditions produced 1.5 mg of the double-Sec mutant of MBP (MBP

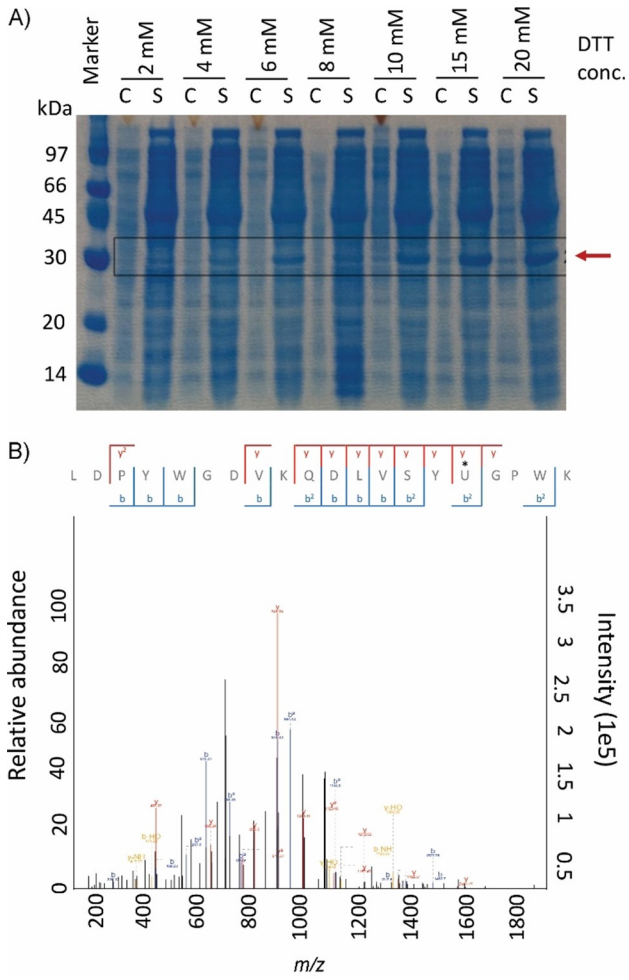


Figure 1. Optimization of DTT concentrations in CFPS of the C80U mutant of ZiPro and confirmation of the incorporation of Sec at position 80. A) SDS-PAGE of seven CFPS reactions carried out in the presence of different concentrations of DTT (2 to 20 mM). The arrow identifies the band of ZiPro C80U. Crude samples (C) are from the inner reaction mixture after overnight CFPS, and the corresponding supernatants (S) were obtained after spinning down the precipitates of the inner reaction mixtures. B) LC-MS/MS analysis of ZiPro C80U, confirming Sec incorporation in response to the Cys codon at position 80. Annotated peaks in the LC-MS/MS spectrum correspond to the series of b ions (blue) and y ions (red) that resulted from fragmentation of the carbamidomethylated peptide from ZiPro C80U shown at the top. b- and y-ion fragmentation is also indicated in the amino acid sequence, where the position of the Sec residue is marked by a star. The 3-charged (+3) ion with *m/z* of 825.706 was fragmented to produce the MS/MS spectrum.

T237U/T345U), 1.8 mg of ZiPro V36U, and 5.0 mg of GB1 Q32U. Red precipitate formed in the course of the CFPS reaction, indicating the generation of a limited amount of elemental selenium. Attempts to purify ZiPro C80U by chromatography over a Ni-NTA column resulted in a product of molecular weight 296 Dalton greater than expected, suggesting reaction of the Sec residue with an unknown compound in the CFPS reaction.

ESI mass spectrometry of MBP T237U/T345U and of ZiPro V36U indicated that the proteins were monomers, and there was no indication of misincorporation of any of the canonical amino acids in response to the Cys codon (Figure 2A and C). SDS-PAGE analysis of purified MBP T237U/T345U and MBP T237C/T345C show dimers (Figure S1 in the Supporting Information). GB1 Q32U was also a dimer (Figures 2D and S1). Sec is prone to oxidation to selenocystine[11] and ready formation of a selenium-selenium bond would be expected for the solvent-exposed Sec residue of MBP T237U/T345U and GB1 Q32U.

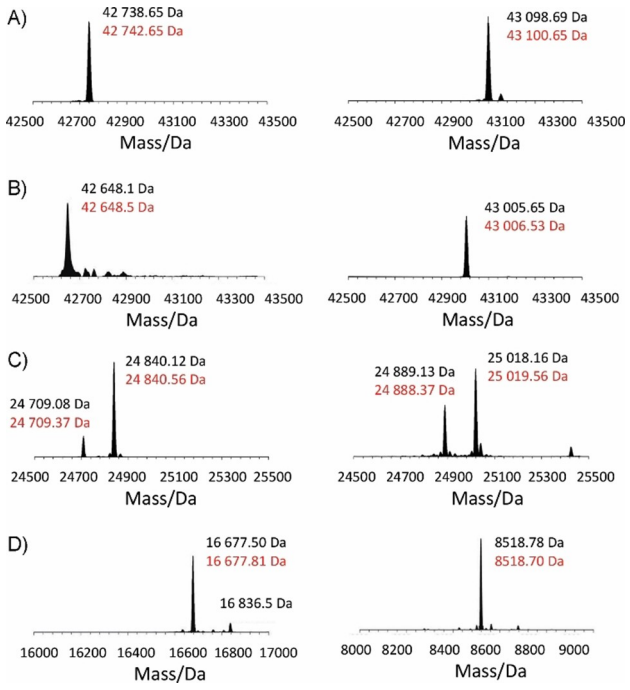


Figure 2. Labeling Sec or Cys with alkylating DPA tags (as shown in Figure 3). Observed and expected masses are indicated in black and red, respectively. The expected mass increase after tagging with two MDPA tags is 358 Da. The left and right panels show the mass spectra before and after tagging, respectively. A) Mass spectra of MBP T237U/T345U before and after tagging with Cl-MDPA. The absence of untagged protein suggests complete tagging yields. B) Mass spectra of MBP T237C/T345C before and after tagging with Br-MDPA. C) Mass spectra of ZiPro V36U before and after tagging with Cl-MDPA. The peak at 24 709.08 corresponds to ZiPro V36U without the first methionine residue. The peak at 24 889.13 Da corresponds to ZiPro V36U without the first methionine residue tagged with Cl-MDPA. D) Mass spectra of unlabeled GB1 Q32U and after labeling with Cl-MDPA.

Relative reactivities of 4F-MDPA, 4Cl-MDPA, and 4Br-MDPA with cysteine and selenocysteine

To compare the reactivity of Sec and Cys in nucleophilic tagging reactions, we tested reactions with the three 4-halogen-methylene-dipicolinic acid (4halogen-MDPA) tags 4Br-MDPA, 4Cl-MDPA and 4F-MDPA (Figure 3). In particular, high tagging yields were sought with the aim of performing DEER experiments with Gd^{III} ions on tagged double-Sec and double-Cys mutants of MBP (MBP T237U/T345U and MBP T237C/T345C, respectively). 4Cl-MDPA did not react with the double-Cys mutant even after 16 hours at room temperature and pH 7.5. In contrast, tagging of the double-Sec mutant with 4Cl-MDPA was complete within 10 minutes at pH 7.5 and room temperature (Figure 2A). To prevent oxidation of the selenol group, the tagging reaction was conducted in the presence of 5 mM DTT. In contrast, tagging of the double-Cys mutant required 4Br-MDPA and was conducted in the absence of DTT to avoid reaction of the tag with DTT. Complete ligation of the double-Cys mutant was achieved by incubation with 4Br-MDPA for 16 hours at room temperature and pH 7.5 (Figure 2B). These results highlight the much greater reactivity of selenol groups in nucleophilic substitution reactions compared with thiol groups. 4F-MDPA proved completely inert, failing to react with the Sec mutant ZiPro V36U to any significant extent even after 32 h incubation at pH 7.5 in the presence of TCEP, whereas 4Cl-MDPA reacted within 10 minutes (Figure 2C).

Presence of a Se–Se bond in the dimer of GB1 Q32U did not prevent tagging, as the addition of 5 mM DTT proved sufficient to tag GB1 Q32U with 4Cl-MDPA within 10 minutes (Figure 2D).

DEER experiment of double-tagged MBP

To assess the utility of the 4Cl-MDPA and 4Br-MDPA tags for distance measurements by DEER experiments, the double-Cys and double-Sec mutants of MBP tagged with MDPA were titrated with Gd^{III} in 1:1 ratio of Gd^{III}/tag and measured on a W-band EPR spectrometer. The echo-detected EPR spectra (Figure S2) revealed full widths at half height of the central transition of, respectively, 5.2 and 6.6 mT for the double-Sec

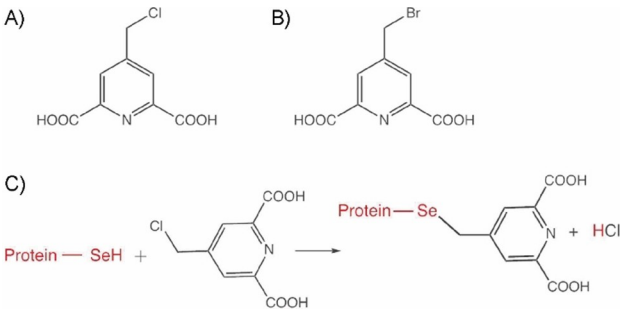


Figure 3. Tags and tagging reaction used. Chemical structures of A) 4-chloromethylene-dipicolinic acid (4Cl-MDPA) and B) 4-bromomethylene-dipicolinic acid (4Br-MDPA). C) Reaction scheme between 4Cl-MDPA tag and a Sec residue of a protein.

and double-Cys mutants of MBP. As the selenium or sulfur atoms are too far from the Gd^{III} ion to influence its zero-field splitting, we attribute the difference to the presence of some free Gd^{III} in the sample of the double-Sec mutant, which would also explain the longer Gd^{III} phase memory time observed for the double-Sec mutant. The DEER data yielded very similar distance distributions, indicating equivalent performance of the double-Sec and double-Cys mutants (Figure 4). A somewhat lesser modulation depth observed for the double-Sec mutant again is consistent with the presence of some free Gd^{III} in this sample. The maxima of the experimental distance distributions, observed at about 4.2 and 4.3 nm, respectively, were in reasonable agreement with the distance distribution modeled on the crystal structure 1OMP,[12] which suggested 3.9 and 4.0 nm as the most frequent Gd–Gd distance (Figure 4).

Discussion

Cell-free production of selenoproteins

The present work shows that, with modified conditions, CFPS provides exceptionally facile access to pure selenoproteins by global substitution of Cys. Compared to *in vivo* expression systems, there is no need to take recourse to Cys auxotrophs or to guard for the presence of Cys produced by biosynthesis or released by protein turnover. Furthermore, CFPS is ideally suited for the production of proteins with amino acids that are toxic *in vivo*, such as Sec, and the proteins produced in this work were obtained in yields comparable to proteins produced from the 20 canonical amino acids. While *in vivo* protocols recommend washing of the cells to flush out Cys prior to the addition of sSec,[9j] the CFPS protocol is much more straightforward. Not having to lyse cells in order to release the selenoproteins also minimizes the chances for selenium oxidation and elimination that have been reported previously.

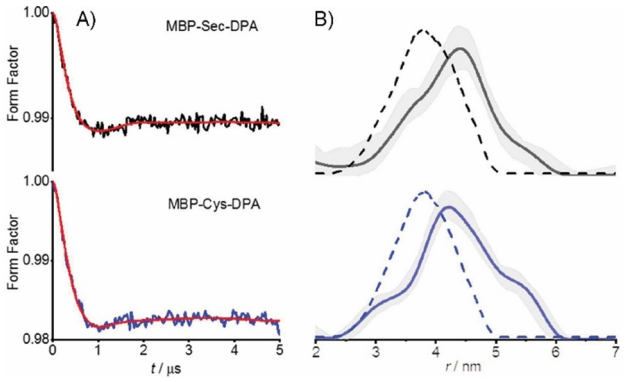


Figure 4. DEER results obtained with different tags on the MBP Sec mutant (black) and the MBP Cys mutant (blue). a) Background-corrected DEER traces along with the fit obtained with the distance distributions shown on the right. The primary DEER traces are shown in Figure S3. b) Analyzed distance distributions including confidence ranges (in shaded gray) and comparison with modeled distances (in dashed lines).

To misload cysteinyl-tRNA with Sec requires the provision of Sec rather than its oxidized form selenocystine, in which it is commercially available. The diselenide bond in selenocystine can be reduced with water soluble reducing agents such as DTT, TCEP, or $NaBH_4$, but too strong reducing conditions are detrimental as the selenium of selenol groups is easily reduced to elemental selenium. On the other hand, oxygen readily re-oxidizes Sec to selenocystine. In our hands, increased concentrations of DTT proved to provide the right balance and only a small amount of elemental selenium precipitation was observed. Furthermore, analysis by mass spectrometry provided no evidence of selenium elimination to dehydroalanine or serine, which can arise as a consequence of selenoprotein oxidation followed by *syn*-β-elimination of selenenic acid.[9a,13]

Reactivities of 4F-MDPA, 4Cl-MDPA, and 4Br-MDPA towards Sec and Cys

Sec residues in proteins present sites of greater chemical reactivity than Cys residues, allowing enhanced rates of chemical labeling. In the case of 4Cl-MDPA and at pH 7.5, labeling of Sec residues was complete within 10 minutes, whereas Cys residues were not noticeably modified even after 16 h incubation at room temperature. Tagging of Cys residues was achieved with 4Br-MDPA but, in the example of MBP T237C/T345C, still required overnight incubation to go to completion, as 60% was still untagged after 3 hours incubation. 4F-MDPA was expected to be the least reactive tag, as the relative performance of halogen ions as leaving group in nucleophilic substitution reactions decreases in the sequence $Br^- > Cl^- > F^-$. Our failure to tag Sec with 4F-MDPA thus was not unexpected.

The 4Cl-MDPA tag thus emerged as ideally tuned to discriminate between Sec and Cys. The very large difference in reactivity of 4Cl-MDPA towards Sec as opposed to Cys at near-neutral pH revealed in the present work suggests that, quite generally, Sec residues can be selectively and quantitatively tagged in the presence of solvent-exposed Cys residues without any significant conversion of the Cys thiol groups. Furthermore, the reaction rate of 4Cl-MDPA with Sec was fast, which bodes well for labeling of proteins that are sensitive to precipitation or other pathways of degradation. This opens the door to a bio-orthogonal labeling scheme, whereby the installation of a single chemically caged Sec residue in the target protein provides a selective site for tag attachment. Systems for genetic encoding of photocaged Sec have been described, but unfortunately are not very efficient.[9a,k] Enhancing the yields of proteins containing masked Sec will be a worthwhile project.

Tagging of Sec and Cys residues for DEER experiments

Dipicolinic acid is a small and rigid chemical moiety that binds lanthanide ions with nanomolar affinities.[14] Lanthanide tags based on DPA thus offer straightforward ways to attach paramagnetic Gd^{III} ions to proteins for DEER measurements.[15]

The first tag was 4-mercaptomethyl-dipicolinic acid (4MMDPA), which required activation of the Cys residues prior to tagging and produced a relatively flexible tether with a disulfide bond.[15a,b] In subsequent tags, the methylene group in the tether was eliminated by the use of 3-mercaptodipicolinic acid (3MDPA) or 4-mercaptodipicolinic acid (4MDPA), which reduces the flexibility of the tether.[15c,16] To obtain a ligation product that is stable towards reducing agents, 4-vinyl-DPA was introduced, which leads to a thioether but also two methylene groups between the sulfur and DPA moiety.[17] The same number of heavy atoms between the sulfur atom (or selenium atom in the case of Sec) and the DPA moiety are obtained with the new 4Cl-MDPA tag. 4-(phenylsulfonyl)-pyridine-2,6-dicarboxylic acid (4PS-DPA) is the only DPA tag that features fewer atoms between the DPA moiety and the protein backbone by direct attachment of the pyridine ring to the sulfur of the Cys side chain.[18]

The width of the Gd^{III}–Gd^{III} distance distribution observed with the 4Cl-MDPA tag was 0.3 nm wider at half height than the width obtained previously with C9-Gd^{III} tags attached to the same double-Cys mutant of MBP.[19] In contrast to the DPA tags, the C9 tag is a cyclen tag with bulky phenylethyl-amide pendants that reacts with Cys residues by formation of a disulfide bridge, and at least one of the methylene groups in the tether is flexible.[20] The narrow width of the distance distributions obtained with the C9-Gd^{III} tag may be attributed to the greater steric restrictions associated with a bulky tag compared to those associated with a smaller tag such as DPA. The maxima of the DEER distance distributions indicated distances that were about 0.3 nm longer than modeled by a rotamer library that ignores potential preferential interactions with other residues (Figure S4). In the case of the T237/T345 mutants of MBP, the site of T237 is highly solvent exposed, whereas the site of T345 is in the middle of a helix with other amino acid side chains nearby. As the modeled distances ranged from 2.4 to 5.1 nm, preferential interactions could well cause a bias in the distance distribution. Nonetheless, the maximum of the DEER distance distribution obtained with the 4Br-MDPA tag on MBP T237C/T345C was within about 0.2 nm of that obtained with the C9-Gd^{III} tag. Narrower and more defined distance distributions might be obtained for sites where the lanthanide ion can also coordinate to a negatively charged carboxyl group.[21]

Conclusion

Sec has long been recognized to offer an exquisitely reactive chemical handle for chemical modifications of proteins, but the complexity of selenoprotein synthesis restricted its use in protein chemistry. The CFPS protocol established in the present work provides easy access to sSec in high yield and exceptional purity by global substitution of Cys. The very fast and selective alkyation of Sec versus Cys is highly attractive for bioconjugations. Our results encourage the search for methods, by which Sec can be incorporated into proteins site-selectively in high yield and in the presence of Cys.

Experimental Section

Synthesis of 4Cl-MDPA, 4Br-MDPA, and 4F-MDPA: The 4Cl-MDPA and 4Br-MDPA tags were synthesized from 4-(hydroxymethyl) pyridine-2,6-dicarboxylate. The protocols and analytical data are provided in the Supporting Information.

Plasmid constructs: In the following, we denote Sec as "U" for brevity in mutant names, even though Sec was incorporated in response to a Cys codon. The ZiPro C80U and ZiPro V36U constructs contained a C-terminal His_6 tag and were cloned into the NdeI and EcoRI sites of the T7 expression vector pETMSCI.[22] The wild-type ZiPro construct was described previously.[23] It contained residues 48*–95* of NS2B (where the star indicates residues in NS2B) linked to the NS3 protease domain (residues 1–170) by a $Gly_4SerGly_4$ linker. It also contained the mutations R95*A, K15N, and R29G to increase stability toward autocleavage and the mutations C80S and C143S to eliminate all Cys residues from the wild-type protein. In the ZiPro V36U mutant, the codon of Val36 was replaced by a Cys codon, while in the ZiPro C80U mutant, the natural Cys codon was retained for residue 80.

The double-mutant T237C/T345C of MBP was cloned into the NdeI and EcoRI sites of the T7 expression pETMCSIII vector[22] with a N-terminal His_6 tag followed by a tobacco etch virus (TEV) protease recognition site.

The GB1 Q32U construct was cloned into the NdeI and EcoRI sites of the T7 expression vector pETMSCI.[22] It contained a MASMTG expression tag at the N terminus and a TEV protease site followed by a His_6 tag at the C terminus.

Protein production: All protein samples were produced by continuous exchange cell-free protein synthesis conducted at 30 °C for 16 h, using 1 mL inner reaction mixture of S30 cell extract from *E. coli* BL21 and 10 mL outer buffer. The genes of the ZiPro, MBP and GB1 mutants were PCR-amplified with eight-nucleotide single-stranded overhangs to generate DNA suitable for use in CFPS.[24] The cell-free reactions were set up according to an established protocol[25] modified to promote the incorporation of Sec. Cys was omitted and L-selenocystine (Sigma-Aldrich #545996) was added in the inner and outer buffers at a final concentration of 1 mM. DTT was used to reduce selenocystine to Sec. No further special precautions were taken to prevent oxidation. MBP T237U/T345U was produced using the same vector as for the production of MBP T237C/T345C. The DTT concentration was optimized in CFPS experiments of ZiPro C80 U, using DTT concentrations in the inner and outer buffers ranging from 2 to 20 mM. 10 mM DTT was used for all subsequent protein expressions.

Protein purification: ZiPro V36U and C80U were purified using a 1 mL His GraviTrap column (GE Healthcare Life Sciences) equilibrated with buffer A (50 mM Tris·HCl, 300 mM NaCl, pH 7.5), washed with buffer B (same as buffer A but with 10 mM imidazole), and eluted with buffer C (same as buffer A but with 300 mM imidazole, 100 mM EDTA). The purified protein samples were dialyzed overnight against storage buffer D (same as buffer A but with 5 mM DTT and 10% glycerol) and concentrated using an Amicon centrifugal ultrafiltration tube with a molecular weight cutoff (MWCO) of 10 kDa.

MBP T237U/T345U and MBP T237C/T345C were purified using a 1 mL His GraviTrap column (GE Healthcare Life Sciences) equilibrated with buffer A, washed with buffer E (same as buffer A but with 15 mM imidazole), and eluted with buffer F (same as buffer A but with 300 mM imidazole). Finally, the protein samples were dialyzed against buffer A and concentrated using an Amicon centrifugal ultrafiltration tube (MWCO 10 kDa). GB1 Q32U was

purified in the same way, except that the final concentration step used an ultrafiltration tube with MWCO 3 kDa.

Protein tagging with 4Cl-MDPA and 4Br-MDPA: To maintain reducing conditions, the double-Sec and double-Cys mutants of MBP (MBP T237U/T345U and MBP T237C/T345C) were kept in buffer A with 5 mM DTT. MBP T237U/T345U was incubated with the tags at room temperature for 1 h. 4Cl-MDPA dissolved in DMSO was added in fivefold molar excess. The samples were shaken at room temperature for 10 min and the reactions were quenched by putting the samples into liquid nitrogen. The completeness of the ligation reactions was assessed by mass spectrometry. ZiPro V36U and GB1 Q32U were tagged in the same way. The same tagging protocol was also applied to MBP T237C/T345C, except that, prior to the addition of tag, the protein was incubated with 5 mM DTT for 1 h followed by removal of the DTT using an Amicon centrifugal ultrafiltration tube (MWCO 10 kDa). For complete tagging of MBP T237C/T345C with 4Br-MDPA, a 50 μM solution of the protein was incubated overnight with tenfold molar excess of tag at room temperature.

Mass spectrometry: The proteins were analyzed using an Orbitrap Elite Hybrid Ion Trap-Orbitrap mass spectrometer coupled with an UltiMate 3000 UHPLC (Thermo Scientific). Samples were injected into the mass analyzer via an Agilent ZORBAX SB-C3 Rapid Resolution HT Threaded Column, using an acetonitrile gradient (10–85%) and 0.1% formic acid. In-gel digestion for mass spectrometric characterization of ZiPro C80U was performed according to a published protocol.[26] In-gel tryptic digestion was performed as described[27] and the peptide containing the Sec identified by LC-MS/MS.

Sample preparation for EPR spectroscopy: The tagged MBP mutants were concentrated and exchanged into EPR buffer (50 mM MES·KOH in D_2O, pH 6.7, uncorrected pH meter reading) to a final concentration of 100 μM using an Amicon ultracentrifugation tube (MWCO 10 kDa). Perdeuterated glycerol was added to reach a 20% (*v*/*v*) final composition. $GdCl_3$ dissolved in D_2O was added in 2:1 molar ratio of $GdCl_3$ to conjugated Cl/Br-MDPA tag.

DEER experiments: All data were recorded on a home-built W-band (94.9 GHz) spectrometer.[26] The data were recorded at 10 K and the temperature was stabilized with a cryo-free cooling system from ColdEdge Technologies. Echo-detected EPR (ED-EPR) spectra (Figure S2) were recorded with a Hahn echo sequence ($\pi/2$-τ–π-echo) and sweeping the magnetic field. The microwave pulse lengths were $\pi/2 = 15$ ns, $\pi = 30$ ns, the inter-pulse delay $\tau = 550$ ns and the repetition rate 1 ms. Echo decay traces (Figure S3) were recorded at the maximum of the Gd^{III} EPR spectrum monitoring the echo intensity of a Hahn echo sequence by increasing the τ interval in steps of 50 ns and using pulse lengths and repetition times as above.

Gd^{III}–Gd^{III} DEER data were recorded with the four-pulse "reversed" ($\pi/2_{vobs}$-τ_1-π_{vobs}-(τ_1-t)-π_{vpump}-(τ_2+t)-π_{vobs}-τ_2-echo) DEER sequence (rDEER)[28] using chirp pump pulses[29] and monitoring the echo intensity with increasing time *t*. An eight-step phase cycle was applied. The microwave pulse lengths at the observe frequency were $\pi/2_{vobs} = 15$ ns, $\pi_{vobs} = 30$ ns, the inter-pulse delays $\tau_1 = 5.7$ μs and $\tau_2 = 5$ μs, and the repetition rate 0.2 ms. The DEER data were recorded by observing at the maximum of the EPR line at 94.9 GHz and by pumping spins with two chirp pulses set at +100 MHz and −100 MHz from the observed position with a width of 300 MHz (i.e., at 95.0–95.3 and 94.5–94.8 GHz, respectively) using pulse lengths of 96 ns. Due to limits set by the cavity bandwidth, the actual excitation bandwidth of the chirp pulses is estimated to be about 100 MHz.

The primary DEER data (Figure S4) were transformed into distance distributions using the DeerAnalysis2018[30] software and Tikhonov regularization. The background contributions to the primary DEER data were removed by fitting the dimensionality of the background function to obtain reasonable fits in the frequency and time domain data, as judged by visual inspection. The regularization parameter was 100. The removal of the background contributions was evaluated within the validation tool of the DeerAnalysis2018 program performed from 5 to 80% of the DEER time trace in 16 trials, 50% random noise was added in 10 trials and datasets within 15% of the best root mean square deviation were retained (i.e., default prune level 1.15).

Acknowledgements

Financial support by the Australian Research Council (DP200100348, DP210100088 and FL170100019), and through a Centre of Excellence (CE200100012) is gratefully acknowledged.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: electron paramagnetic resonance spectroscopy · paramagnetic tags · protein labeling · selenocysteine · selenoproteins

[1] a) J. M. Chalker, G. J. L. Bernardes, Y. A. Lin, B. G. Davis, *Chem. Asian J.* **2009**, *4*, 630–640; b) E. Baslé, N. Joubert, M. Pucheault, *Chem. Biol.* **2010**, *17*, 213–227; c) X.-C. Su, G. Otting, *J. Biomol. NMR* **2010**, *46*, 101–112; d) C. D. Spicer, B. G. Davis, *Nat. Commun.* **2014**, *5*, 4740; e) O. Boutureira, G. J. L. Bernardes, *Chem. Rev.* **2015**, *115*, 2174–2195; f) N. Krall, F. P. da Cruz, O. Boutureira, G. J. L. Bernardes, *Nat. Chem.* **2016**, *8*, 103–113; g) P. Roser, M. J. Schmidt, M. Drescher, D. Summerer, *Org. Biomol. Chem.* **2016**, *14*, 5468–5476; h) C. Nitsche, G. Otting, *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *98–99*, 20–49.
[2] M. M. Haugland, J. E. Lovett, E. A. Anderson, *Chem. Soc. Rev.* **2018**, *47*, 668–680.
[3] G. Jeschke, *Annu. Rev. Phys. Chem.* **2012**, *63*, 419–446.
[4] W. L. Hubbell, A. Gross, R. Langen, M. A. Lietzow, *Curr. Opin. Struct. Biol.* **1998**, *8*, 649–656.
[5] M. J. Schmidt, J. Borbas, M. Drescher, D. Summerer, *J. Am. Chem. Soc.* **2014**, *136*, 1238–1241.
[6] a) E. H. Abdelkader, A. Feintuch, X. Yao, L. A. Adams, L. Aurelio, B. Graham, D. Goldfarb, G. Otting, *Chem. Commun.* **2015**, *51*, 15898–15901; b) M. C. Mahawaththa, M. D. Lee, A. Giannoulis, L. A. Adams, A. Feintuch, J. D. Swarbrick, B. Graham, C. Nitsche, D. Goldfarb, G. Otting, *Phys. Chem. Chem. Phys.* **2018**, *20*, 23535–23545; c) E. G. B. Evans, G. L. Millhauser, *Methods in Enzymology, Vol. 563* (Eds.: P. Z. Qin, K. Warncke), Academic Press, **2015**, pp. 503–527.
[7] a) A. Feintuch, G. Otting, D. Goldfarb, *Methods in Enzymology, Vol. 563* (Eds.: P. Z. Qin, K. Warncke), Academic Press, **2015**, pp. 415–457; b) M. R. Fleissner, E. M. Brustad, T. Kálai, C. Altenbach, D. Cascio, F. B. Peters, K. Hideg, S. Peuker, P. G. Schultz, W. L. Hubbell, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 21637–21642.
[8] a) L. Johansson, G. Gafvelin, E. S. J. Arnér, *Biochim. Biophys. Acta Gen. Subj.* **2005**, *1726*, 1–13; b) E. S. J. Arnér, *Exp. Cell Res.* **2010**, *316*, 1296–1303; c) R. Mousa, R. Notis Dardashti, N. Metanis, *Angew. Chem. Int. Ed.* **2017**, *56*, 15818–15827; *Angew. Chem.* **2017**, *129*, 16027–16037.
[9] a) R. Rakauskaitė, G. Urbanavičiūtė, A. Rukšėnaitė, Z. Liutkevičiūtė, R. Juškėnas, V. Masevičius, S. Klimašauskas, *Chem. Commun.* **2015**, *51*, 8245–8248; b) J. Liu, Q. Chen, S. Rozovsky, *J. Am. Chem. Soc.* **2017**, *139*, 3430–3437; c) Q. Cheng, E. S. J. Arnér, *J. Biol. Chem.* **2017**, *292*, 5476–

5487; d) C. Aldag, M. J. Bröcker, M. J. Hohn, L. Prat, G. Hammond, A. Plummer, D. Söll, *Angew. Chem. Int. Ed.* **2013**, *52*, 1441–1445; *Angew. Chem.* **2013**, *125*, 1481–1485; e) R. Thyer, S. A. Robotham, J. S. Brodbelt, A. D. Ellington, *J. Am. Chem. Soc.* **2015**, *137*, 46–49; f) C. Miller, M. J. Bröcker, L. Prat, K. Ip, N. Chirathivat, A. Feiock, M. Veszprémi, D. Söll, *FEBS Lett.* **2015**, *589*, 2194–2199; g) Z. Fan, J. Song, T. Guan, X. Lv, J. Wei, *ACS Synth. Biol.* **2018**, *7*, 249–257; h) X. Fu, D. Söll, A. Sevostyanova, *RNA Biol.* **2018**, *15*, 461–470; i) T. Mukai, A. Sevostyanova, T. Suzuki, X. Fu, D. Söll, *Angew. Chem. Int. Ed.* **2018**, *57*, 7215–7219; *Angew. Chem.* **2018**, *130*, 7333–7337; j) S. Mueller, H. Senn, B. Gsell, W. Vetter, C. Baron, A. Boeck, *Biochemistry* **1994**, *33*, 3404–3412; k) A. P. Welegedara, L. A. Adams, T. Huber, B. Graham, G. Otting, *Bioconjugate Chem.* **2018**, *29*, 2257–2264; l) J. Liu, F. Zheng, R. Cheng, S. Li, S. Rozovsky, Q. Wang, L. Wang, *J. Am. Chem. Soc.* **2018**, *140*, 8807–8816; m) Q. Cheng, S. Stone-Elander, E. S. J. Arnér, *Nat. Protoc.* **2006**, *1*, 604–613; n) M. H. Pinkerton, P. R. Copeland, *Selenoproteins: Methods and Protocols* (Ed.: L. Chavatte), Springer, New York, **2018**, pp. 93–101.

[10] J. N. Burnell, J. A. Karle, A. Shrift, *J. Inorg. Biochem.* **1980**, *12*, 343–351.

[11] L. Moroder, *J. Pept. Res.* **2005**, *11*, 187–214.

[12] A. J. Sharff, L. E. Rodseth, J. C. Spurlino, F. A. Quiocho, *Biochemistry* **1992**, *31*, 10657–10663.

[13] S. Ma, R. M. Caprioli, K. E. Hill, R. F. Burk, *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 593–600.

[14] I. Grenthe, *J. Am. Chem. Soc.* **1961**, *83*, 360–364.

[15] a) X.-C. Su, B. Man, S. Beeren, H. Liang, S. Simonsen, C. Schmitz, T. Huber, B. A. Messerle, G. Otting, *J. Am. Chem. Soc.* **2008**, *130*, 10486–10487; b) A. Potapov, H. Yagi, T. Huber, S. Jergic, N. E. Dixon, G. Otting, D. Goldfarb, *J. Am. Chem. Soc.* **2010**, *132*, 9040–9048; c) B. Man, X.-C. Su, H. Liang, S. Simonsen, T. Huber, B. A. Messerle, G. Otting, *Chem. Eur. J.* **2010**, *16*, 3827–3832.

[16] X. Jia, A. Maleckis, T. Huber, G. Otting, *Chem. Eur. J.* **2011**, *17*, 6830–6836.

[17] Q.-F. Li, Y. Yang, A. Maleckis, G. Otting, X.-C. Su, *Chem. Commun.* **2012**, *48*, 2704–2706.

[18] Y. Yang, J.-T. Wang, Y.-Y. Pei, X.-C. Su, *Chem. Commun.* **2015**, *51*, 2824–2827.

[19] A. P. Welegedara, Y. Yang, M. D. Lee, J. D. Swarbrick, T. Huber, B. Graham, D. Goldfarb, G. Otting, *Chem. Eur. J.* **2017**, *23*, 11694–11702.

[20] E. H. Abdelkader, M. D. Lee, A. Feintuch, M. R. Cohen, J. D. Swarbrick, G. Otting, B. Graham, D. Goldfarb, *J. Phys. Chem. Lett.* **2015**, *6*, 5016–5021.

[21] C. Nitsche, M. C. Mahawaththa, W. Becker, T. Huber, G. Otting, *Chem. Commun.* **2017**, *53*, 10894–10897.

[22] C. Neylon, S. E. Brown, A. V. Kralicek, C. S. Miles, C. A. Love, N. E. Dixon, *Biochemistry* **2000**, *39*, 11989–11999.

[23] M. C. Mahawaththa, B. J. G. Pearce, M. Szabo, B. Graham, C. D. Klein, C. Nitsche, G. Otting, *Antiviral Res.* **2017**, *142*, 141–147.

[24] P. S. C. Wu, K. Ozawa, S. P. Lim, S. G. Vasudevan, N. E. Dixon, G. Otting, *Angew. Chem. Int. Ed.* **2007**, *46*, 3356–3358; *Angew. Chem.* **2007**, *119*, 3420–3422.

[25] a) M. A. Apponyi, K. Ozawa, N. E. Dixon, G. Otting, *Methods Mol. Biol.* **2008**, *426*, 257–268; b) K. Ozawa, K. V. Loscha, K. V. Kuppan, C. T. Loh, N. E. Dixon, G. Otting, *Biochem. Biophys. Res. Commun.* **2012**, *418*, 652–656.

[26] D. Goldfarb, Y. Lipkin, A. Potapov, Y. Gorodetsky, B. Epel, A. M. Raitsimring, M. Radoul, I. Kaminker, *J. Magn. Reson.* **2008**, *194*, 8–15.

[27] A. Shevchenko, H. Tomas, J. Havli, J. V. Olsen, M. Mann, *Nat. Protoc.* **2006**, *1*, 2856–2860.

[28] T. Bahrenberg, Y. Yang, D. Goldfarb, A. Feintuch, *Magnetochemistry* **2019**, *5*, 20.

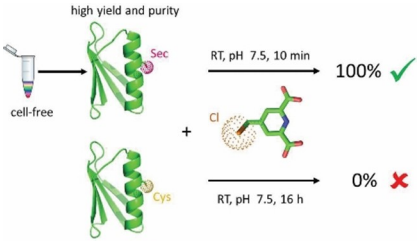
[29] a) T. Bahrenberg, Y. Rosenski, R. Carmieli, K. Zibzener, M. Qi, V. Frydman, A. Godt, D. Goldfarb, A. Feintuch, *J. Magn. Reson.* **2017**, *283*, 1–13; b) A. Doll, M. Qi, N. Wili, S. Pribitzer, A. Godt, G. Jeschke, *J. Magn. Reson.* **2015**, *259*, 153–162.

[30] G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger, H. Jung, *Appl. Magn. Reson.* **2006**, *30*, 473–498.

Manuscript received: November 19, 2020
Revised manuscript received: December 10, 2020
Accepted manuscript online: December 14, 2020
Version of record online: ■■, ■■■

FULL PAPERS

[34]Se-lective protein tagging: Cell-free protein synthesis is used to produce selenoproteins in high yield and purity. Selenocysteine has long been recognized as offering an exquisitely reactive chemical handle for chemical modifications of proteins, and we show that the alkylation of solvent-exposed selenol groups with a chloro-alkyl group proceeds quickly under conditions that leaves cysteine residues unchanged.



Dr. A. P. Welegedara, Dr. A. Maleckis, R. Bandara, Dr. M. C. Mahawaththa, I. Dilhani Herath, Y. Jiun Tan, Dr. A. Giannoulis, Prof. D. Goldfarb, Prof. G. Otting, Prof. T. Huber**