Copper-mediated selenazolidine deprotection enables one-pot chemical synthesis of challenging proteins

Zhenguang Zhao[a] and Norman Metanis*[^a]

Abstract: While chemical protein synthesis (CPS) has granted access to challenging proteins, synthesis of longer proteins is often limited by low abundance or non-strategic placement of cysteine (Cys) residues, essential for native chemical ligations (NCL), as well as multiple purification and isolation steps. Selective deselenization and one-pot CPS serve as key technologies to circumvent these issues. Herein, we describe the one-pot total synthesis of human thiocysteine: glutathione sulfurtransferase (TSTD1), a 115-residue protein with a single Cys residue at its active site, and its seleno-analogue. WT-TSTD1 was synthesized in a C-to-N synthetic approach employing multiple NCL reactions, Cu(II)-mediated deprotection of selenazolidine (Sez), and chemoselective deselenization, all in one-pot. In addition, the protein’s seleno-analogue (Se-TSTD1), in which the active site Cys is replaced with selencysteine, was synthesized with a kinetically controlled ligation in a one-pot, N-to-C synthetic approach. TSTD1’s one-pot synthesis was made possible by the newly reported, rapid, and facile copper-mediated selenazolidine deprotection that can be accomplished in one minute. Finally, catalytic activity of the two proteins indicated that Se-TSTD1 possessed only four-fold lower activity than WT-TSTD1 as a thiosulfate: glutathione sulfurtransferase, suggesting that selenoproteins can have physiologically comparable sulfutransferase activity as their cysteine counterparts.

Due to their critical role in life function, proteins have long been subject of intense study in all fields of the natural sciences. Of special interest is proteins’ remarkable ability to achieve a wide range of function using only a few, simple functional groups.[1] Protein study demands the preparation of sufficient amounts of pure and homogeneous samples for downstream investigation. Although many proteins are accessed using recombinant expression in host cells, commonly used methods are not adequate for modified or toxic protein production. Likewise, expression of proteins containing unnatural amino acids or the rare amino acid selencysteine (Sec, U) requires specialized methods and suffer low yields and homogeneity.[2] In such cases, chemical protein synthesis (CPS) offers an alternative route to control protein structure and function at atomic precision.[3] The mainstream method is based on solid phase peptide synthesis (SPPS)[4] combined with selective chemical ligation reactions of unprotected peptide segments in aqueous solutions.[5–8] The most widely used method, native chemical ligation (NCL),[9] typically proceeds through the chemoselective condensation of two unprotected polypeptide chains, one equipped with a C-terminal thioester and the other with an N-terminal cysteine (Cys, C) residue, at neutral pH and in aqueous solution.[8–7] Because of length limitations of SPPS (usually less than ~50 amino acids) of synthetic polypeptide chains, total synthesis of proteins with more than 100 amino acids usually requires more than one NCL reaction.[9] Thus, the Cys residue on the N-terminus of central peptide segment(s) has to be protected to prevent cyclization side reactions, and several different protecting groups for the N-terminus Cys have been employed.[9] The most popular of these, thiazolidine (Thz), is valued for its commercial availability, low price, and efficient deprotection under mild conditions.[9c, 10] Though NCL can be performed at either Cys or Sec,[5–11] many proteins lack these required residues or have them at non-strategic locations. These limitations are nowadays circumvented using NCL/desulfurization and NCL/deselenization approaches,[9a, 12] whose chemistries have been applied to enable ligation at a wide variety of amino acids.[13] Notably, NCL/deselenization is chemoselective for seleno-containing amino acids,[12c–14] even in the presence of unprotected thiols elsewhere in the protein sequence.

To further improve the efficiency of CPS, there is considerable interest in the preparation of proteins in a one-pot manner, which could minimize intermediate isolation and purification steps, and hence improve protein recovery yields.[9e, 15] Recently, we developed selenazolidine (Sez), an easily accessible, N-terminal protected form of Sec, which grants synthetic access to proteins with non-strategically placed Cys residues.[16] For Sez to Sec conversion, previously used methoxylamine (MeONH₂) treatment required hours-long reaction times similar to thiazolidine (Thz) to Cys conversion.[9e, 13c, 16] Moreover, the large excess of MeONH₂ typically used in these reactions gave unwanted side-products after reacting with the C-terminal thioester.[17] Thus, we sought to develop an alternative approach for Sez deprotection under mild conditions that would not affect sequential ligation or deselenization reactions, in the hope of achieving one-pot protein synthesis. We now report the one-pot chemical synthesis of peptides and proteins using Cu(II)-mediated Sez deprotection, NCL, and subsequent deselenization reactions (Fig. 1a). The method is demonstrated in high-yield, one-pot chemical syntheses of both human thiosulfate:glutathione sulfurtransferase (TSTD1) and its seleno-analogue (Se-TSTD1). Enzymatic characterization indicates that Se-TSTD1 possesses only four-fold lower sulfurtransferase activity than the WT-TSTD1.

We began by exploring milder conditions for efficient deprotection of Sez that would also be compatible under NCL and deselenization reactions conditions. Though Brik and coworkers’ Pd/thiol mediated deprotection of Thz was incompatible with Sez deprotection,[10] we were inspired by their work to seek other metal ions for the deprotection of Sez, with an eye towards lower concentrations and greater affordability. Thus, the deprotection of Sez in peptide 1, Sez-LKFAG-NH₂, was screened with a series of transition metal ions.

Confirming Brik’s report,[10] peptide 1 was completely consumed in a reaction with 1.5 equiv allylpalladium(II) chloride dimer [Pd(allyl)Cl₂], within 0.5 h (Table 1, entry 1), yet the desired product was not observed, perhaps due to Pd-Se complex formation (Fig. S1 in the SI). As Liu and coworkers had

[a] Zhenguang Zhao[^a] and Prof. Norman Metanis[^a]
Institute of Chemistry, The Hebrew University of Jerusalem, Edmond J. Safra, Givat Ram, Jerusalem 91904 (Israel)
E-mail: Metanis@mail.huji.ac.il
previously demonstrated that silver ions are able to deprotect Thz,\textsuperscript{[18]} we showed that under similar conditions, 1.5 equiv AgCl deprotected Sez to give only 58% of the desired product \(1'\) after an extended reaction time (32 h, entry 2). On the other hand, quantitative yields of Sez deprotection were observed when 1.5 equiv Fe(III), Cu(I) or Cu(II) were used (entries 3-5). Full deprotection of Sez in the presence of FeCl\(_2\)-6H\(_2\)O was completed in 6 h (entry 3), while the same reaction took only 1 h with CuCl (entry 4). The most efficient metal ion additive for Sez unmasking was CuCl\(_2\), which could be completed within 0.5 h (entry 5), with equal efficiency observed at both 37 °C and 25 °C (entry 6). Notably, Sez deprotection could be completed within 1 min when 5 equiv CuCl\(_2\) was used (Table 1, entry 7, and Fig. S2). Further, we found that Cu(II)-mediated deprotection of Sez is mild and epimerization-free (Fig. S3 in the SI). This makes CuCl\(_2\) an excellent alternative to MeONH\(_2\) in the deprotection of Sez, and its faster reaction times should be beneficial for one-pot chemical protein synthesis.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Metal ion additives</th>
<th>Time (h)</th>
<th>Yield (%)([b])</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[Pd(allyl)Cl](_2)</td>
<td>0.5</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>AgCl</td>
<td>32([c])</td>
<td>58</td>
</tr>
<tr>
<td>3</td>
<td>FeCl(_2)-6H(_2)O</td>
<td>6</td>
<td>quantitative</td>
</tr>
<tr>
<td>4</td>
<td>CuCl</td>
<td>1</td>
<td>quantitative</td>
</tr>
<tr>
<td>5</td>
<td>CuCl(_2)</td>
<td>0.5</td>
<td>quantitative</td>
</tr>
<tr>
<td>6</td>
<td>CuCl(_2)</td>
<td>0.5([d])</td>
<td>quantitative</td>
</tr>
<tr>
<td>7</td>
<td>CuCl(_2)</td>
<td>1 min([e])</td>
<td>quantitative</td>
</tr>
</tbody>
</table>

\([a]\) 2 mM of 1 was dissolved into the phosphate buffer (0.2 M, 6 M Gn•HCl, pH 6) and 1.5 equiv metal additives were added to the reaction. \([b]\) The yield is calculated according to HPLC integration. \([c]\) The reaction was not completed within 32 h. \([d]\) The reaction temperature was 25 °C. \([e]\) 5 equiv CuCl\(_2\) was used.

With the optimized conditions for Sez deprotection in hand, we next explored its application in multi-step CPS. To develop an efficient one-pot methodology, we explored the performance of a peptide-arylthioester (peptide-MPAA thioester, MPAA is 4-mercaptophenylacetic acid\textsuperscript{[19]} ) in deprotection, ligation, and deselenization (Fig. 1a). We reasoned that the ligation step should proceed smoothly with the activated arylthioester, while avoiding traditional excess concentrations of MPAA\textsuperscript{[19]} which inhibit subsequent, desired deselenization reactions.\textsuperscript{[12c, 13b]} On the other hand, to avoid unwanted deselenization during NCL\textsuperscript{[11b]} but promote reduction of inactive diselenides, the use of TCEP should be present, but in minimized amounts. Although ascorbate could be used to inhibit the deselenization side-reaction during NCL\textsuperscript{[13b, 16, 20]} purification would be required before the final deselenization step, which is inconsistent with our target one-pot strategy.

To test this approach, we prepared three peptides: 4a (LKFAG-MPAA), 4b (Sez-LKFAG-MPAA) and 4c (URAFS-NH\(_2\) isolated as a diselenide dimer), which were used for the one-pot ligation, Cu(II)-mediated deprotection of Sez, second ligation, and final deselenization to give 4g (Fig. 1). The first ligation between peptides 4b and 4c proceeded smoothly with 0.5 equiv TCEP, and was followed by a 30 min deprotection of ligated product 4d with 1.5 equiv CuCl\(_2\) at pH 6. Then 1.0 equiv 4a was added for the second ligation in the presence of 2.5 equiv TCEP at pH 7, and the reaction was completed after 10 h. Lastly, deselenization of 4f could be achieved within 10 h by adding 50 equiv TCEP at pH 5 (Fig. 1b). It is worth noting that 2 equiv MPAA were released over both NCL steps, causing the observed slower deselenization step.\textsuperscript{[12c, 13b]} The final target peptide 4g was obtained in 39% isolated yield over four steps.

Next, we used this approach for the one-pot synthesis of proteins with a non-strategically placed Cys residue. To test the applicability of our strategy, we chose the human thiosulfate: glutathione sulfotransferase (TSTD1), which is thought to play a role in protein persulfidation\textsuperscript{[21]} and H\(_2\)S-based metabolism, although its exact physiological role is still not fully understood.\textsuperscript{[22]} TSTD1 is a 115-residue protein with a single Cys residue located at its active site (Cys79),\textsuperscript{[23]} which makes it a challenging target for chemical synthesis using a typical NCL/desulfurization approach. Because of the similarity between sulfur and selenium in natural (seleno)proteins, we wondered how replacing Cys at the active site with Sec would impact the biological activity of TSTD1. The possible existence of natural selenoproteins with sulfotransferase activity has been reported by Gladyshev and coworkers based on genetic analysis of microbial selenoproteomes, but have not yet been isolated or characterized.\textsuperscript{[23]} We used an NCL/deselenization approach in both the synthesis of TSTD1 and its seleno-analogue, Se-TSTD1, choosing ligation sites at Gly39–Ala40 and Phe78–Cys79 (Fig. 2).

For WT-TSTD1 synthesis, we used three peptide segments with two NCLs and a final deselenization step, in a one-pot, C-to-N synthetic approach, where Ala40 was temporarily substituted with Sez (Fig. 2a). All peptide segments were synthesized by standard Fmoc-SPPS (Fig. S9, S10 and S12 in the SI), in which

![Figure 1](https://example.com/image1.png)

**Figure 1.** a. The synthesis of a sample peptide in one-pot using Cu(II)-mediated deprotection of Sez. b. HPLC traces for beginning and end of each reaction step; for full discussion of conditions see SI.
TSTD1(2-39)-MPAA and TSTD1(40-78)(A40Sez)-MPAA were first synthesized using an acyl pyrazoles precursor and subsequently converted into aryl thioesters,[24] The ligation between TSTD1(40-78)(A40Sez)-MPAA and TSTD1(79-115) was performed in the presence of 0.5 equiv TCEP at pH 7 and was completed in 5 h (Fig. 3a). The crude mixture was treated with 1.5 equiv CuCl₂ at pH 6, affording TSTD1(40-115)(A40U) (Fig. 3a). Then, TSTD1(2-39)-MPAA was added to the ligation buffer with 2.5 equiv TCEP, and ligation was completed in 10 h. The deselenization reaction was performed using 100 equiv TCEP to give the WT-TSTD1 in 28% isolated yield over four steps.

Subsequently, we turned to the chemical synthesis of the seleno-analogue, Se-TSTD1, which presented an additional challenge over the wildtype, as we wanted to preserve the Sec79 in the active site while allowing deselenization of Sec40 into the natural Ala in the final product. Hence, we decided to prepare Se-TSTD1 by kinetically controlled ligation (KCL)[25] using one-pot N-to-C synthetic approach (Fig. 2b), in which a peptide aryl thioester could react with a Sec-peptide with a C-terminal alkyl thioester. For this, only a single additional peptide, TSTD1(79-115)(C79U), was synthesized while the other peptides had already been prepared by modifications from the previous syntheses (see SI). The ligation between TSTD1(2-39)-MPAA and TSTD1(40-78)(A40U)-MMP (MMP is methyl 3-mercaptopyrroline) was performed in the presence of 0.5 equiv TCEP at pH 7 and was completed in 8 h to give TSTD1(2-78)(A40U)-MMP. Then, 100 equiv TCEP were added to accelerate the deselenization reaction at pH 5 under argon atmosphere, which was accomplished in 12 h to yield TSTD1(2-78)-MMP (Fig. 3b), with a minor by-product, TSTD1(2-78)-OH, observed from hydrolysis of TSTD1(2-78)-MMP. Here, for the first time, we show that the deselenization reaction is compatible in the presence of a C-terminal thioester.[26] Afterwards, the second ligation between TSTD1(2-78)-MMP and TSTD1(79-115)(C79U) was performed in ligation buffer with 0.2 M MPAA, 50 mM TCEP and 0.1 M sodium ascorbate at pH 7 and was completed in 8 h. The target protein Se-TSTD1 was obtained in 25% isolated yield over three steps (Fig. 3b).

The two proteins, WT-TSTD1 and Se-TSTD1 (characterized by HPLC and HRMS, Fig. S14 and S15 in the SI), were separately folded in phosphate buffer, and their circular dichroism (CD) spectra were recorded. CD analysis indicated that the two proteins are composed of mainly α-helical secondary structures, with only minor differences observed (Fig. S28 in the SI).

Next, the catalytic sulfurtransferase activities of the two proteins were tested using p-toluenethiosulfonate (p-Tol-SO₂S⁻) as a substrate and glutathione (GSH) as the acceptor co-substrate (Eq. 1 and 2, Fig. 4a).[21] The p-Tol-SO₂S⁻ exhibits a moderate absorption band in the UV region (ε₂₅₀ = 6760 M⁻¹·cm⁻¹), which is lost upon transfer of the sulfane sulfur to the acceptor and formation of p-Tol-SO₂-. The Se-TSTD1 exhibited p-Tol-SO₂S⁻ : glutathione sulfurtransferase activity (Fig. 4) that was four-fold lower than the WT-TSTD1 (apparent turnover rates k₅cat (app) = 3.8 and 16.8 s⁻¹, respectively).

Figure 2. CPS of WT-TSTD1 and Se-TSTD1. a. The chemical synthesis of WT-TSTD1 in one-pot by ligation of three unprotected peptide segments using C-to-N approach. b. CPS of Se-TSTD1 in one-pot by ligation of three unprotected peptide segments using N-to-C approach.

Figure 3. a. Analytical HPLC traces for NCL, Sez-deprotection, and deselenization reactions for WT-TSTD1 synthesis. b. Analytical HPLC traces of NCL, and deselenization reactions for Se-TSTD1 synthesis.

Figure 4. Side-by-side comparison of WT-TSTD1 and Se-TSTD1 with p-Tol-SO₂S⁻ and GSH as the sulfane sulfur donor and acceptor, respectively. Reactions were conducted at 37 °C in 50 mM phosphate buffer (pH 8.0) containing 0.15 mM p-Tol-SO₂S⁻ and 0.58 mM GSH, with the indicated enzyme(s). (a) The proposed catalytic mechanism of WT-TSTD1[21]; (b) Se-TSTD1 takes significantly longer time to catalyze the reaction compared to WT-TSTD1. The apparent gap of first point between blue and red curves is due to the enzyme reaction that occurs during mixing (~5 sec).[22]; (d) Varying protein concentrations shows that WT-TSTD1 is four-fold more active than Se-TSTD1; (d) Steady-state kinetic analysis of WT-TSTD1 and Se-TSTD1 with the respective k₅cat and k₉ indicated.
Furthermore, steady-state kinetic analysis for WT-TSTD1 and Se-TSTD1 were performed by varying the concentration of the substrate p-Tol-SO\textsubscript{3}H with the acceptor GSH at saturation (Fig. 4d). We found that the $k_{\text{cat}}$ for Se-TSTD1 was only 4-fold lower than WT-TSTD1, with identical $K_m$ observed (Fig. 4d). These results validated the crucial catalytic role of Cys79 in WT-TSTD1; even Sec substitution lowers the catalytic activity of the enzyme to a certain extent, in agreement with previous studies.[23] These results also suggest that natural selenoproteins may exhibit comparable sulfurtransferase activity similar to their cysteine containing counterparts.[23]

In summary, we developed a Cu(II)-mediated deprotection of Sez to facilitate the chemical synthesis of proteins. This new deprotection method is not only faster than previously reported methods, it avoids unwanted side reactions at vulnerable thioesters. It is this development which allowed us to pursue the one-pot synthesis of both small peptides and wildtype TSTD1 in high yields. Additionally, TSTD1’s seleno-analogue, Se-TSTD1, was prepared in one-pot using KCl in the less-utilized, N-to-C sequential assembly. The CD spectra and enzymatic activity of the two proteins indicated that the Cys to Sec substitution had a minimal effect on structure and function of TSTD1. The synthetic approach reported here will surely prove a useful tool for the synthesis of more complex proteins. Such studies are currently being explored in our laboratory.

Acknowledgements

We would like to thank Dr. Shailesh Kumar and Ms. Reem Mousa for technical assistance and helpful discussions. We also thank Mrs. Ricki Notis Dardashti for input on the manuscript. Z.Z. is grateful for a CSC fellowship. N.M. acknowledges the financial support of Israel Science Foundation (783/18) and ICRF Acceleration Grant.

Keywords: selenocysteine • human thiocysteine: glutathione sulfurtransferase • one-pot chemical protein synthesis• native chemical ligation • selenoproteins