Published in final edited form as:

Chemistry. 2020 April 16; 26(22): 4952–4957. doi:10.1002/chem.202000310.

Selenolysine: a new tool for traceless isopeptide bond formation

Rebecca Notis Dardashti^[a], Shailesh Kumar^[a], Shawn M. Sternisha^[b], Post Sai Reddy^[a], Brian G. Miller^[b], Norman Metanis^[a]

^[a]The Institute of Chemistry, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 91904, Israel

^[b]Department of Chemistry and Biochemistry, Florida State University, Tallahassee, FL, USA 32306-4390

Abstract

Despite their biological importance, post-translationally modified proteins are notoriously difficult to produce in a homogeneous fashion using conventional expression systems. Chemical protein synthesis or semi-synthesis offers a solution to this problem; however, traditional strategies often rely on sulfur-based chemistry that is incompatible with the presence of any cysteine residues in the target protein. To overcome these limitations, we present the design and synthesis of γ -selenolysine, a selenol-containing form of the commonly modified proteinogenic amino acid, lysine. The utility of γ -selenolysine is demonstrated with the traceless ligation of the small ubiquitin-like modifier protein, SUMO-1, to a peptide segment of human glucokinase. The resulting polypeptide is poised for native chemical ligation and chemoselective deselenization in the presence of unprotected cysteine residues. Selenolysine's straightforward synthesis and incorporation into synthetic peptides marks it as a universal handle for conjugating any ubiquitin-like modifying protein to its target.

Graphical Abstract

metanis@mail.huji.ac.il.

Publisher's Disclaimer: This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

The design and synthesis of γ -selenolysine, a selenol-containing form of the commonly modified proteinogenic amino acid, lysine is reported. The utility of γ -selenolysine is demonstrated with the traceless ligation of the small ubiquitin-like modifier protein, SUMO-1, to a peptide segment of human glucokinase.

Keywords

Chemical protein synthesis; post-translational modification; deselenization; expressed protein ligation; SUMOylation

Following ribosomal translation and folding, expressed proteins can undergo a great number of post-translational modifications (PTMs) before fulfilling their biological functions.^[1] These PTMs instruct the cell to carry out a variety of regulatory activities including transportation and degradation, to name a few.^[2] Additionally, Lys, which is one of the more reactive amino acids under biological conditions, can be modified with post-translationally modifying proteins such as ubiquitin and other ubiquitin-like modifiers (Ubl).^[3]

Techniques for controlling PTMs in cellular expression are not yet sufficiently well-developed to allow recombinant expression of homogeneously modified proteins.^[4] Alternatively, some methods for semi-selective modification of expressed proteins have been developed, most employing Cys chemistry. For example, dehydroalanine (Dha), a suitable electrophile for Michael addition reactions, can be generated from Cys,^[5] Ser,^[6] thioethers,^[7] and several selenium-containing unnatural amino acids.^[8] Michael addition to the resulting Dha has been used to introduce many small-molecule PTMs and entire PTM proteins, as in the case of ubiquitylated α -globin.^[9] However, a stereoselective method for addition to Dha has yet to be developed,^[9] a drawback in the chiral world of biology.

Chemical protein synthesis (CPS) and semi-synthesis, on the other hand, offers a direct route to selectively modified proteins.^[10] Based on solid-phase peptide synthesis (SPPS)^[11] and native chemical ligation (NCL)^[12], this approach allows the preparation of modified proteins, including proteins with multiple sites of different PTMs.^[4]

Additionally, Lys-mediated chemical ligation to form an isopeptide bond, typically referred to as isopeptide chemical ligation (ICL), has been used as a tool to conjugate post-translationally modifying proteins to their targets. Previously, Muir and coworkers reported ubiquitylated protein synthesis using a modified Lys with an auxiliary that was removed post-ligation;^[13] more recently, the Liu group employed an auxiliary to produce di- and tri-ubiquitin chains.^[14] Other studies utilized orthogonal protecting groups to incorporate isopeptide bonds on-resin,^[15] or installed dipeptides at desired modification sites to allow for downstream ligation.^[16] Brik and coworkers introduced δ -thiolysine in the conjugation of ubiquitin and synthesis of ubiquitylated proteins or ubiquitin chains.^[17] Another Lys derivative, γ -thiolysine, has been used as a handle in dual NCL reactions,^[18] in which the γ -thiolysine plays a role in both main chain and isopeptide ligation. Global desulfurization of the final products gives a native Lys at the ligation site (Scheme 1).

We sought to develop selenium-containing analogs of modified Lys residues in order to facilitate traceless isopeptide bond formation through ICL. A subsequent chemoselective deselenization reaction in the presence of unprotected Cys residues (Scheme 1) would give the native Lys.

The aforementioned γ/δ -thiolysines used in ICL/desulfurization approaches are not compatible in the presence of proteins with natural Cys residues in their sequences, unless orthogonal Cys protecting groups are installed. This presents a true challenge in expressed protein ligation (EPL),¹⁵ where installation of protected Cys residues in Ubl proteins would require specialized expression methods. In contrast, thanks to selenium's unique chemical properties, γ/δ -selenolysine (Se-Lys, Sek) could be used in an ICL/deselenization approach in the target protein, without the risk of desulfurizing natural Cys residues.^[19]

Here we report the synthesis and use of γ -Se-Lys as a new tool for traceless isopeptide bond formation. The design and synthesis of δ -Se-Lys, which could not be completed due to spontaneous Se-elimination, is described in the SI (Scheme S1).

To obtain a protected form of γ -Se-Lys suitable for SPPS, we turned to work by Merckx *et al.*, where an Fmoc-Lys[S γ (*S*tBu)-N^e(Boc)]-OH was successfully synthesized from γ -chlorolysine.^[20] We opted to follow the previous protocol until selenium incorporation, protecting the α -amino and carboxylate groups of γ -chlorolysine with 9-BBN and then the ϵ -NH₂ group with Boc to give **2** (Scheme 2).

Nucleophilic substitution of Cl to Se in DMF was performed using Na₂Se₂, generated *in situ* from Se powder, hydrazine monohydrate, and NaOH.^[21] The reaction gave the γ -diselenide product **3** in 42% yield. A reduction with NaBH₄ and protection with *p*-methoxybenzyl chloride (Mob-Cl or PMB-Cl) gave **4** in 60% yield. Finally, removal of 9-BBN and subsequent Fmoc-protection gave the desired Fmoc-Lys[Se γ (Mob)-N^e(Boc)]-OH, **6**, in 77% yield (Scheme 2).

Following our successful synthesis of a protected Se-Lys building block for SPPS, we set out to design an analog with an orthogonal protecting group at the ϵ -NH₂ site to allow for dual chemical ligation. In the proposed application, Se-Lys with an orthogonal protection at ϵ -NH₂ would first be used in NCL along the main peptide chain. After NCL, the protecting group of ϵ -NH₂ would be removed to allow for a second ligation, forming an isopeptide bond with the desired PTM protein-thioester. Final deselenization would give native Lys at the PTM site (Scheme 3). Disappointingly, using γ -chlorolysine, we were unsuccessful in incorporating an orthogonal protecting group on the ϵ -NH₂. Instead, an alternative route was pursued from known aldehyde **10**, which can be obtained from *L*-aspartic acid precursor in 3 steps (Scheme 4).^[22] In this synthesis, the aldehyde **10** was treated with methyl (triphenylphosphoranylidene)acetate to afford the α,β -unsaturated ester **11** (85% for 2 steps).^[23] Michael addition on **11** with Mob-selenolate (generated *in situ* from (Mob-se)-2 and NaBH₄)^[24] gave compound **12** in 90% yield.

The less sterically hindered ester in **12** was selectively converted to an aldehyde with DIBAL-H in dry ether at -78 °C, which was treated with NaBH₄ in EtOH to give alcohol **13**

(68% for 2 steps). The hydroxyl group in **13** was converted to mesylate using MsCl and Et₃N in CH₂Cl₂ followed by treatment with NaN₃ in DMF at 80 °C, which afforded the azide compound **14** in 88% yield. Reduction of azide **14** with Ph₃P in MeOH, followed by treatment with Alloc-Cl and Et₃N, gave an Alloc-protected amine **15** (71%). Conversion of *di*-Boc (**15**) to *mono*-Boc (**16**) with LiBr^[25] was carried out in CH₃CN at 65 °C for 24 h in 98% yield. Finally, hydrolysis of methyl ester with LiOH^[17a] in aqueous THF gave the carboxylic acid **17** in 97% yield (**17** can be used in Boc-SPPS). Compound **17** was treated with TFA in CH₂Cl₂ to give an amine-salt, which was finally converted to Fmoc-protected amino acid **18** using Fmoc-OSu and aqueous Na₂CO₃ in 90% yield. The final product, Fmoc-Lys[Se^γ(Mob)-N^ε(Alloc)]-OH, was achieved in 17% overall yield for 13 steps of synthesis (Scheme 4).

Following successful main-chain NCL, the Alloc protecting group can be removed off-resin with a ruthenium catalyst^[26] and the resulting peptide can be subjected to isopeptide ligation. Studies showing the utility of Fmoc-Lys[N^ε(Alloc)-Se^γ(Mob)]-OH in dual chemical ligation are currently underway.

To assess the utility of γ -Se-Lys in protein semi-synthesis, we incorporated our first-generation Se-Lys **6** (Scheme 2) into post-translationally modified human glucokinase (GCK). Prior studies demonstrated that the pancreatic isoform of this key glucose homeostatic enzyme is modified by the small ubiquitin-like modifier protein, SUMO-1, at one or more lysine residues near the N-terminus.^[27] Unfortunately, the structural and functional impact of site-specific GCK SUMOylation has been difficult to assess due to an inability to produce homogeneously SUMOylated GCK. We viewed the SUMO-1-GCK conjugate as an attractive target, since both proteins contain unprotected cysteine residues, and thus require the unique chemical properties of γ -Se-Lys.

We devised a three-segment strategy to obtain semisynthetic SUMOylated human GCK (Scheme 5). First, SPPS was used to incorporate our Fmoc-protected γ -Se-Lys **6** at residue 15*,^[28] within a peptide consisting of the first 18 residues of GCK, GCK(2–19)-NHNH₂ (Scheme 5, Figure S1). This peptide also contained a C-terminal masked thioester for downstream use in NCL.^[29] Next, SUMO-1 was produced as a C-terminal fusion with the MxeGyrA C-terminal intein (Figures S2 and S3),^[30] which was subsequently converted to its 2-mercaptoethane sulfonic acid thioester (SUMO-1-Mes) for ICL (Figure S4, details and method in SI). Following ICL, the product underwent a one-pot deselenization/thioesterification (Figure S5), giving an overall yield of 12.3% of the GCK(2–19) segment linked to SUMO-1 via an isopeptide bond at Lys15 (Scheme 5, Figure 1).

*Using online analysis programs such as SUMOplot™ or SUMOsp 2.0, which predict possible SUMOylation sites, K12 and K15 are predicted as likely SUMOylation sites in the N-terminus of pancreatic GCK.^[27b, 28a] The first site, K12, is part of SUMO consensus motif Ψ -K-x-D/E, where Ψ is a large hydrophobic amino acid, K is the lysine at which the SUMO is linked, x is any amino acid and D/E is an acidic amino acid (Asp or Glu, respectively).^[28c] The second SUMOylation site is K15, which is not part of a typical SUMO consensus motif. Both analysis programs did not predict K13 as a possible SUMOylation site. In early crystal structures of GCK (PDB entries 1V4S, 1V4T)^[28b] the N-terminus was truncated to facilitate crystallization. In these structures the N terminal region of GCK, which begins with Lys12, adopts an α -helix.^[28d] Although K12, K13 and K15 are adjacent residues, each residue would face a different environment as part of the α -helix, with some inaccessible to the SUMOylation system. Indeed, based on a more recent crystal structure (PDB entry 3IDH),^[28d] only K15 is solvent exposed, while K12 and K13 are on the opposite face of the helix and make electrostatic and hydrogen bonding interactions with the rest of the protein. Based on these observations, we initially focused on Lys15 as the SUMOylation site in pancreatic GCK; however, our method should be applicable to Lys12 or any Lys residue in a protein sequence.

We expected the ligation step to proceed efficiently, as the size of the amino acid adjacent to the thioester typically dictates the rate of the ligation reaction.^[31] SUMO-1 contains a desirable Gly-Gly sequence adjacent to the thioester. However, the ligation step was surprisingly slow. SUMO-1-Mes was converted to the more reactive SUMO-1-MPAA thioester within 1 hour. Yet, ligation of this adduct with the small GCK peptide took up to 22 hours to complete despite a large excess of the peptide and the presence of TCEP and sodium ascorbate, which are known to accelerate the selenium-mediated NCL reaction.^[19b, 32] It is unlikely that the γ -position of the selenol contributes to a slower rate of ligation, based on a previous report indicating that this has no impact on ligation rate.^[33] Instead, a direct comparison of isopeptide ligation in the synthesis of diubiquitin with γ - and δ -thiolysine also showed a slower-than-expected rate of ligation.^[20] In fact, the pK_a of the peptide's amino groups is likely to blame, as the N-terminal α -amino group of a peptide has an average pK_a of 7.7^[34], while the ϵ -amino group of Lys has a pK_a of 10.5^[35], making it significantly less reactive at pH 7, the optimized pH for NCL. Together, these observations suggest that the inefficiency lies in the nature of isopeptide bond formation, as the first, typically rate-determining nucleophilic attack in ICL is performed by the less reactive ϵ -amine.

Due to the slow rate of ligation, a significant portion of SUMO-1-MPAA underwent hydrolysis to inactive SUMO-1-OH rather than to the desired ligation. As a result, the maximum yield we observed was 22% for the ligation step, which severely limited the expected yield of the overall strategy. Notably, for the subsequent deselenization reaction, the unprotected, native Cys52 of SUMO-1 was unaffected by this procedure^[16a] and unmasking of the peptide's cryptic thioester provided a segment poised for NCL. As a final step in the preparation of SUMOylated human GCK, we recombinantly produced a human GCK variant containing a Factor Xa recognition sequence directly upstream of a Cys at position 20 (Scheme 5, and Figure S6). Following purification and Factor Xa digestion, this procedure yielded a truncated, N-terminal Cys containing GCK variant, 20CysGCK, that displayed wild-type kinetic characteristics (Figure S7). Unfortunately, efforts to perform NCL between 20CysGCK and the SUMO-1-GCK(1–19)-COSR segment proved unsuccessful. We observed no ligation after prolonged incubation of the reaction components in the absence of denaturants, whereas inclusion of guanidium hydrochloride in the ligation buffer caused extensive precipitation of the protein, as has been reported earlier.^[36] Nonetheless, we expect that the methods presented here will provide a route towards Cys-rich Ubl-protein conjugates. Further exploration of this topic could include an investigation of bacterial expression using the orthogonal tRNA synthetase/tRNA pairs, as has been done for thiol-containing lysine analogs.^[37]

In summary, lysine, often a site of ubiquitylation, SUMOylation, and other Ubl protein PTM conjugations, was designed with an artificial γ -selenol moiety. The resulting amino acid was successfully used for traceless isopeptide ligation of SUMO-1 on a peptide of its natural target protein, human glucokinase. Additionally, a selenolysine analog, possessing orthogonal protecting groups to be used in traceless dual chemical ligation, was designed and synthesized. Selenolysine presents an attractive building block in the synthesis of post-translationally modified proteins.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the Dr. Tsafi Danieli, Dr. Mario Lebendiker, and Dr. Yael Keren for their assistance in SUMO-1-COSR expression and purification in the Wolfson Centre for Applied Structural Biology at HUJI. Thanks to Dr. Bill Breuer of the Mass Spectrometry Unit at the Institute of Life Sciences, HUJI for assistance with HR-MS. We wish to thank Prof. Dr. Huib Ovaa (Leiden University Medical Center, The Netherlands) for the generous gift of chlorolysine. RND was supported by the Kaete Klausner Ph.D. scholarship, SK was supported by PBC fellowship for postdocs, NM acknowledges the support of ISF (783/18) and the US-Israel Binational Science Foundation (BSF) (2014167), and BGM acknowledges the support of the NIH (GM115388).

References

- [1]. Wold F, Annu. Rev. Biochem 1981, 50, 783–814. [PubMed: 6791580]
- [2]. Yang XJ, Oncogene 2005, 24, 1653–1662. [PubMed: 15744326]
- [3]. Welchman RL, Gordon C, Mayer RJ, Nat. Rev. Mol. Cell Biol 2005, 6, 599–609. [PubMed: 16064136]
- [4]. Siman P, Brik A, Org. Biomol. Chem 2012, 10, 5684–5697. [PubMed: 22527305]
- [5]. Chalker JM, Gunnoo SB, Boutureira O, Gerstberger SC, Fernández-González M, Bernardes GJL, Griffin L, Hailu H, Schofield CJ, Davis BG, Chem. Sci 2011, 2, 1666–1676.
- [6]. Strumeyer DH, White WN, Koshland DE Jr., Proc. Natl. Acad. Sci. USA 1963, 50, 931–935. [PubMed: 14082360]
- [7]. aHolmes TJJ, Lawton RG, J. Am. Chem. Soc 1977, 99, 1984–1986; [PubMed: 839018] bBernardes GJL, Chalker JM, Errey JC, Davis BG, J. Am. Chem. Soc 2008, 130, 5052–5053. [PubMed: 18357986]
- [8]. Mousa R, Dardashti RN, Metanis N, Angew. Chem. Int. Ed 2017, 56, 15818–15827.
- [9]. Meledin R, Mali SM, Singh SK, Brik A, Org. Biomol. Chem 2016, 14, 4817–4823. [PubMed: 27143624]
- [10]. Kochendoerfer GG, Kent SBH, Curr. Opin. Chem. Biol 1999, 3, 665–671. [PubMed: 10600727]
- [11]. Merrifield RB, J Am Chem Soc 1963, 85, 2149–2154.
- [12]. Dawson PE, Muir TW, Clark-Lewis I, Kent SBH, Science 1994, 266, 776–779. [PubMed: 7973629]
- [13]. aChatterjee C, McGinty RK, Pellois JP, Muir TW, Angew. Chem. Int. Ed 2007, 46, 2814–2818; bMcGinty RK, Kim J, Chatterjee C, Roeder RG, Muir TW, Nature 2008, 453, 812–816. [PubMed: 18449190]
- [14]. aPan M, Gao S, Zheng Y, Tan X, Lan H, Tan X, Sun D, Lu L, Wang T, Zheng Q, Huang Y, Wang J, Liu L, J Am Chem Soc 2016, 138, 7429–7435; [PubMed: 27268299] bPan M, Zheng Q, Ding S, Zhang L, Qu Q, Wang T, Hong D, Ren Y, Liang L, Chen C, Mei Z, Liu L, Angewandte Chemie 2019, 58, 2627–2631. [PubMed: 30589182]
- [15]. Tang S, Liang LJ, Si YY, Gao S, Wang JX, Liang J, Mei ZQ, Zheng JS, Liu L, Angew Chem Int Edit 2017, 56, 13333–13337.
- [16]. aBouchenna J, Senechal M, Drobecq H, Stankovic-Valentin N, Vicogne J, Melnyk O, Bioconjugate Chem 2019, 30, 2684–2696; bBouchenna J, Senechal M, Drobecq H, Vicogne J, Melnyk O, Bioconjugate Chem. 2019, 30, 2967–2973.
- [17]. aKumar KSA, Haj-Yahya M, Olschewski D, Lashuel HA, Brik A, Angew. Chem. Int. Ed 2009, 48, 8090–8094; bHemantha HP, Bavikar SN, Herman-Bachinsky Y, Haj-Yahya N, Bondalapati S, Ciechanover A, Brik A, J. Am. Chem. Soc 2014, 136, 2665–2673; [PubMed: 24437386] cSeenaiah M, Jbara M, Mali SM, Brik A, Angew. Chem. Int. Ed 2015, 54, 12374–12378.
- [18]. aYang R, Pasunooti KK, Li F, Liu XW, Liu CF, J. Am. Chem. Soc 2009, 131, 13592–13593; [PubMed: 19728708] bYang R, Pasunooti KK, Li F, Liu XW, Liu CF, Chem Commun 2010, 46, 7199–7201.

- [19]. aMetanis N, Keinan E, Dawson PE, Angew. Chem. Int. Ed 2010, 49, 7049–7053; bDery S, Reddy PS, Dery L, Mousa R, Notis Dardashti R, Metanis N, Chem. Sci 2015, 6, 6207–6212. [PubMed: 30090236]
- [20]. Merckx R, de Bruin G, Kruithof A, van den Bergh T, Snip E, Lutz M, El Oualid F, Ovaa H, Chem. Sci 2013, 4, 4494–4498.
- [21]. Siebum AHG, Woo WS, Raap J, Lugtenburg J, Eur. J. Org. Chem 2004, 2905–2913.
- [22]. Padron JM, Kokotos G, Martin T, Markidis T, Gibbons WA, Martin VS, Tetrahedron-Asymmetr 1998, 9, 3381–3394.
- [23]. Padrón JM, Kokotos G, Martín T, Markidis T, Gibbons WA, Martin VS, Tetrahedron-Asymmetr 1998, 9, 3381–3394.
- [24]. Gieselman MD, Xie L, van der Donk WA, Org. Lett 2001, 3, 1331–1334. [PubMed: 11348227]
- [25]. Hernández JN, Ramírez MA, Martin VS, J Org Chem 2003, 68, 743–746. [PubMed: 12558394]
- [26]. Castaneda C, Liu J, Chaturvedi A, Nowicka U, Cropp TA, Fushman D, J. Am. Chem. Soc 2011, 133, 17855–17868. [PubMed: 21962295]
- [27]. aAukrust I, Bjørkhaug L, Negahdar M, Molnes J, Johansson BB, Müller Y, Haas W, Gygi SP, Søvik O, Flatmark T, Kulkarni RN, Njølstad PR, J. Biol. Chem 2013, 288, 5951–5962; [PubMed: 23297408] bJohansson BB, Fjeld K, Solheim MH, Shirakawa J, Zhang E, Keindl M, Hu J, Lindqvist A, Døskeland A, Mellgren G, Flatmark T, Njølstad PR, Kulkarni RN, Wierup N, Aukrust I, Bjørkhaug L, Mol. Cell. Endocrinol 2017, 454, 146–157. [PubMed: 28648619]
- [28]. aXu J, He Y, Qiang B, Yuan J, Peng X, Pan XM, BMC Bioinformatics 2008, 9; bRen J, Gao XJ, Jin CJ, Zhu M, Wang XW, Shaw A, Wen LP, Yao XB, Xue Y, Proteomics 2009, 9, 3409–3412; [PubMed: 29658196] cKamata K, Mitsuya M, Nishimura T, Eiki JI, Nagata Y, Structure 2004, 12, 429–438; [PubMed: 15016359] dRodriguez MS, Dargemont C, Hay RT, J. Biol. Chem 2001, 276, 12654–12659; [PubMed: 11124955] ePetit P, Antoine M, Ferry G, Boutin JA, Lagarde A, Gluais L, Vincentelli R, Vuillard L, Acta Crystallogr D Biol Crystallogr 2011, 67, 929–935. [PubMed: 22101819]
- [29]. aZheng JS, Tang S, Qi YK, Wang ZP, Liu L, Nat. Protoc 2013, 8, 2483–2495; [PubMed: 24232250] bFlood DT, Hintzen JCJ, Bird MJ, Cistrone PA, Chen JS, Dawson PE, Angew Chem Int Edit 2018, 57, 11634–11639; cBouchenna J, Senechal M, Drobecq H, Vicogne J, Melnyk O, Bioconjugate Chem 2019, 30, 2967–2973; dFang GM, Li YM, Shen F, Huang YC, Li JB, Lin Y, Cui HK, Liu L, Angewandte Chemie 2011, 50, 7645–7649. [PubMed: 21648030]
- [30]. Shah NH, Dann GP, Vila-Perelló M, Liu Z, Muir TW, J. Am. Chem. Soc 2012, 134, 11338–11341. [PubMed: 22734434]
- [31]. Hackeng TM, Griffin JH, Dawson PE, Proc. Natl. Acad. Sci. USA 1999, 96, 10068–10073. [PubMed: 10468563]
- [32]. aReddy PS, Dery S, Metanis N, Angew. Chem. Int. Ed 2016, 55, 992–995; bRohde H, Schmalisch J, Harpaz Z, Diezmann F, Seitz O, Chembiochem : a European journal of chemical biology 2011, 12, 1396–1400. [PubMed: 21557429]
- [33]. Dardashti RN, Metanis N, Bioorg. Med. Chem 2017, 25, 4983–4989. [PubMed: 28526476]
- [34]. Grimsley GR, Scholtz JM, Pace CN, Protein Sci 2009, 18, 247–251. [PubMed: 19177368]
- [35]. Oregioni A, Stieglitz B, Kelly G, Ritinger K, Frenkiel T, Sci Rep 2017, 7, 43748. [PubMed: 28252051]
- [36]. Molnes J, Teigen K, Aukrust I, Bjørkhaug L, Søvik O, Flatmark T, Njølstad PR, FEBS J 2011, 278, 2372–2386. [PubMed: 21569204]
- [37]. Nguyen DP, Elliott T, Holt M, Muir TW, Chin JW, J Am Chem Soc 2011, 133, 11418–11421. [PubMed: 21736333]

Figure 1.

A. HPLC traces of one-pot SUMO-1-GCK deselenization and thioesterification over time. **1** is SUMO-1-GCK(2-19)(K15Sek)-NHNH₂, **2** is SUMO-1-GCK(2-19)-NHNH₂, **3** shows two peaks, both with the mass of the first intermediate in thioesterification (Schiff base),^[29b] **4** is the final SUMO-GCK(2-19)-MPAA, and **5** is the second intermediate in thioesterification (Knorr pyrazole).^[29b] * shows a mass that most likely is truncated protein from expression of the intein, which coeluted with all intermediate products. The slow rate of thioesterification is attributed to a C-terminal Ile^[31]; **B.** HPLC trace for the final product SUMO-GCK(2-19)-MPAA, and **C.** Overlay of the simulated HR-MS (above) and deconvoluted HR-MS (below).

Scheme 1.

Previously developed δ -thiolysine, by Brik,^{16, 17} and γ -thiolysine, by Yang,^{19, 20} used for NCL and ICL followed by desulfurization for the conjugation of Ubl proteins to Lys residues. Our designed γ/δ -selenolysines for NCL and/or ICL followed by chemoselective deselenization for traceless conjugation of Ubl proteins to Lys residues.

Scheme 2.
 Synthesis of Fmoc-Lys[Se^y(Mob)-N^ε(Boc)]-OH

Scheme 3.

γ -thiolysine^[18] and γ -selenolysine as tools for dual NCL and ICL at Lys. Selenolysine has the advantage of allowing chemoselective deselenization in the presence of unprotected Cys residues.

Scheme 4.
Synthesis of Fmoc-Lys[Se^y(Mob)-N^e(Alloc)]-OH.

Scheme 5.

Proposed synthesis of the semisynthetic SUMOylated GCK.