RESEARCH ARTICLE

Analysis of Selenoprotein F Binding to UDP-glucose:glycoprotein glucosyltransferase (UGGT) by a Photoreactive Crosslinker

Sayaka Higashi[a], Yuki Imamura[a], Takashi Kikuma[a], Takahiro Matoba[a], Saya Orita[a], Yoshiki Yamaguchi[b], Yukishige Ito*[c,d], and Yoichi Takeda*[a]

[a] S. Higashi, Y. Imamura, Dr. T. Kikuma, T. Matoba, S. Orita, Prof. Y. Takeda
Department of Biotechnology, College of Life Sciences, Ritsumeikan University, Kusatsu 525-8577, Japan
E-mail: yotakeda@fc.ritsumei.ac.jp
[b] Prof. Y. Yamaguchi
Faculty of Pharmaceutical Sciences, Tohoku Medical and Pharmaceutical University, Sendai 981-8558, Japan
[c] Prof. Y. Ito
Graduate School of Science, Osaka University, Toyonaka 560-0043 Japan
E-mail: yukito@chem.sci.osaka-u.ac.jp
[d] Prof. Y. Ito
RIKEN Cluster for Pioneering Research, Wako 351-0198 Japan

Supporting Information for this article is given via a link at the end of the document.

Abstract: In the endoplasmic reticulum glycoprotein quality control system, UDP-glucose:glycoprotein glucosyltransferase (UGGT) functions as a folding sensor. Although it is known to form a heterodimer with selenoprotein F (SelenoF), the details of the complex formation remain obscure. A pulldown assay using co-transfected SelenoF and truncated mutants of human UGGT1 (HUGT1) revealed that SelenoF binds to the TRXL2 domain of HUGT1. Additionally, a newly developed photoaffinity crosslinker was selectively introduced into cysteine residues of recombinant SelenoF to determine the spatial orientation of SelenoF to HUGT1. The crosslinking experiments showed that SelenoF formed a covalent bond with amino acids in the TRXL3 region and the interdomain between βS2 and GT24 of HUGT1 via the synthetic crosslinker. SelenoF might play a role in assessing and refining the disulfide bonds of misfolded glycoproteins in the hydrophobic cavity of HUGT1 as it binds to the highly flexible region of HUGT1 to reach its long hydrophobic cavity. Clarification of the SelenoF-binding domain of UGGT and its relative position will help predict and reveal the function of SelenoF from a structural perspective.

Introduction

In eukaryotes, the majority of polypeptides synthesized in the endoplasmic reticulum (ER) are N-glycosylated. The introduced glycans are involved in regulating various biological phenomena [1]. N-linked glycosylation occurs co-translationally on the Asn residue in an Asn-X-Ser/Thr sequence of nascent polypeptides, introducing a tetradecasaccharide composed of nine mannose, three glucose, and two N-acetylglucosamine (GlcNAc) residues [2]. This modification allows nascent glycoproteins to enter the glycan-dependent folding cycle that assists them in adopting the correct structure. In the first step of this process, the terminal and penultimate glucose residues are removed by glucosidases I and II, respectively, to generate the mono-glucosylated glycoform. The latter is recognized by the lectin chaperones calnexin (CNX) and calreticulin (CRT), which assist in the folding of client glycoproteins and prevent their premature transportation to the Golgi apparatus in unfolded states. Several lines of evidence indicate the close association of CNX/CRT with the function-specific adaptor proteins, ERp57, cyclophilin B, and Erp29, among which the role of Erp57 as a protein disulfide isomerase has been well-documented [3–6]. The remaining glucose residue is eventually removed by glucosidase II, and the generated non-glucosylated glycoproteins are scrutinized by the folding sensor UDP-glucose:glycoprotein glucosyltransferase (UGGT).

Whilst two UGGT paralogs, UGGT1 and UGGT2, have been reported [7,8], the former is considered to act primarily as the folding sensor [9]. When glycoproteins are correctly folded, they are transported to the Golgi apparatus for further modifications. On the other hand, if the folding is incorrect or incomplete, a glucose residue is re-added by UGGT1, providing the glycoproteins opportunities to iteratively interact with CNX/CRT [10–14]. Glycoproteins that have failed to attain correct folding within a certain timeframe are degraded via ER-associated degradation. In addition, we showed that UGGT1 facilitates the folding of denatured glycoproteins *in vitro*, suggesting that it might secondarily be a bonafide chaperone in the ER [15].

Selenoprotein F (SelenoF), which forms a heterodimer with UGGT1 and UGGT2, is deemed to act as a disulfide isomerase because the redox potential of *Drosophila melanogaster* SelenoF (*Dm*SelenoF) is within the range of that of oxidoreductases involved in cysteine thiol-disulfide exchange [16]. A TRX-like domain of human SelenoF containing Sec was recently synthesized and exhibited a redox potential of approximately −256 mV, although this partial structure did not affect the folding of two model proteins (bovine pancreatic trypsin inhibitor and hirudin) [17]. As SelenoF was revealed to enhance the glucose transfer activity of UGGT1 and 2 [18], the complex formation might cause their conformational change. However, although the binding was reported to be extremely tight (dissociation constant (Kd): 20 nM) [19], the SelenoF-binding domain of UGGT1 remains unknown.

SelenoF was reported to interact with UGGT1 and UGGT2 via its cysteine-rich domain [19]; however, the exact SelenoF-binding site of UGGTs remains unclear. To determine the binding

site, we first performed a pulldown assay using truncated mutants of UGGT1.

To gain deeper insights into the interaction between UGGT1 and SelenoF, we exploited the crosslinking technique, which has proven valuable in determining spatial proximity between proteins [20]. For our purpose, aryl azide was chosen as the photoreactive handle due to the following reasons: first, it is frequently used in PAL owing to its ease of synthesis and commercial availability [21]; second, aryl azides are valuable as bioorthogonal chemical handles [22] that can be directly used for phosphine-mediated Staudinger ligation.

Although the incorporation of amino acid analogs with photoreactive functional groups into proteins can be achieved via chemical [23] or genetic [24] methods, they are laborious. When a crosslinker is introduced randomly into a protein, site-specificity is impaired; therefore, various alternative methods are required. In these contexts, the intramolecular bridging technique developed by Brocchini et al. drew our attention, which was able to irreversibly modify proteins that have a pair of histidine or cysteine residues under mild conditions [25].

Results and Discussion

SelenoF binding with truncated mutants of human UGGT1 (HUGT1)

To determine the SelenoF-binding domain of HUGT1, a pulldown assay was performed. Specifically, 293T cells were transiently transfected with C-terminal myc-tagged SelenoF [18] and FLAG-tagged, intact, or C-terminal truncated HUGT1. The expression of intact HUGT1 and its truncates was confirmed by western blot (WB) analysis (Supporting Information, Figure S1). After removal of cell debris, the supernatants were immunoprecipitated with anti-FLAG antibody immobilized on agarose beads. Precipitates were subjected to WB analysis using anti-myc or anti-HA antibodies. Three variants of C-terminal-truncated HUGT1 proteins, HUGT1 (38–508), HUGT1 (38–744), and HUGT1 (38–980), and intact HUGT1 co-precipitated with SelenoF, whereas HUGT1 (38-271) did not (Figure 1a).

To determine the SelenoF-binding site among regions B to F, a co-IP assay was performed using myc-tagged SelenoF and FLAG-tagged N-terminal-truncated HUGT1 proteins (Figure 1b). The precipitated fractions were analyzed using WB analysis with an anti-myc antibody. Only HUGT1 (272–1555) and intact HUGT1 co-precipitated with SelenoF, indicating that the SelenoF-binding site is located in region B of HUGT1. This result was supported by the reverse pulldown assay performed with myc-tagged SelenoF and FLAG-tagged HUGT1 as bait and prey, respectively (Supporting Information, Figure S2). To determine the binding sites of SelenoF more precisely, a co-IP assay was carried out using more subtly prepared N-terminal truncates by removing 93, 157, and 199 amino acids in the region B, respectively (Figure 1c). WB analyses revealed the binding of HUGT1 (365–1555), HUGT1 (429–1555), and HUGT1 (466–1555), indicating that the SelenoF-binding site exists in the N-terminal region from residue 466 onwards. A co-IP assay was also performed using a combination of N-terminal HA-tagged SelenoF and FLAG-tagged C-terminal-truncated HUGT1 proteins. After precipitation with an anti-HA antibody immobilized on agarose beads, the precipitates were analyzed using WB with an anti-FLAG antibody, which did not detect any bands corresponding to HUGT1 or its truncated forms. Taken together with the protein sequence alignment of *Ct*UGGT and HUGT1 (Supporting Information, Figure S3) and the structure of HUGT1 predicted by AlphaFold [26,27], the binding domain of HUGT1 for SelenoF was concluded to exist in the TRXL2 domain. In contrast, based on the hydrogen-deuterium exchange mass spectroscopy analysis, Calles-Garcia et al. concluded that the binding domain of *Dm*UGGT toward SelenoF exists between the residues 267 and 307 in TRXL1 [28]. Whilst the origin of the discrepancy is not clear, the mode of binding might be different between two species.

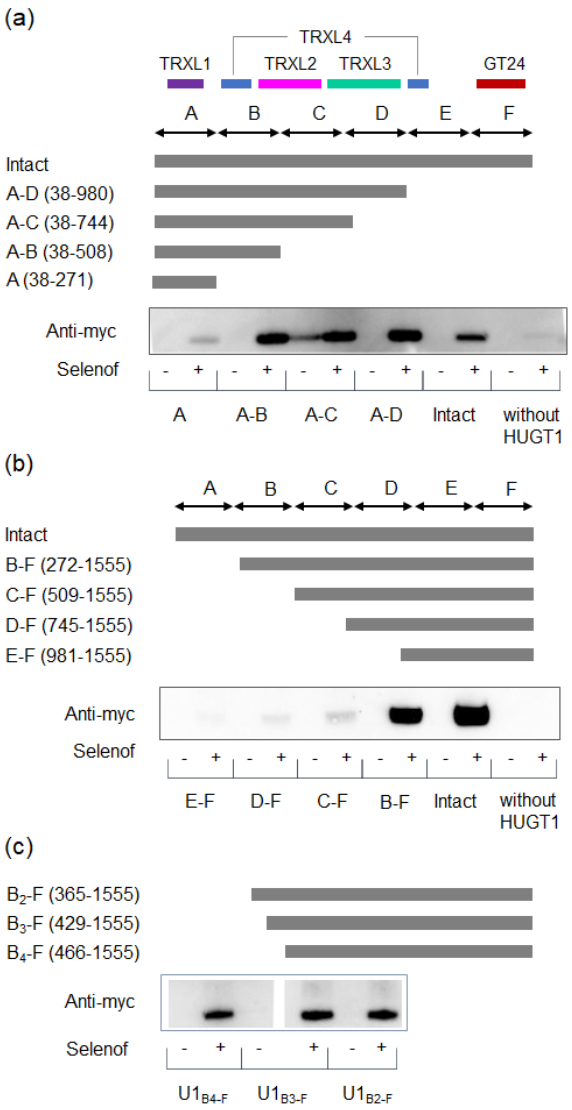
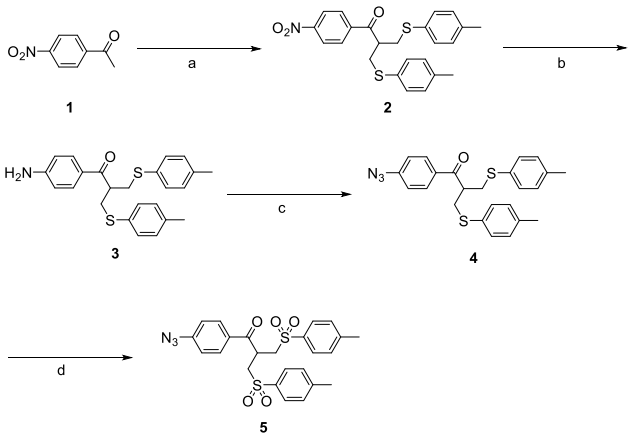


Figure 1. SelenoF binding was analyzed by performing a pulldown assay using truncated mutants of HUGT1. The whole sequence was divided into six domains, namely A (from 38 to 271), B (from 272 to 508), C (from 509 to 744), D (from 745 to 980), E (from 981 to 1244), and F (from 1245 to 1555). Each domain (TRXL1, TRXL2, TRXL3, TRXL4, and GT24) was deduced from amino acid sequence alignment between *Ct*UGGT and HUGT1 (Supporting Information, Figure S3).

Synthesis of the crosslinker diTAAP (**5**)

The newly prepared photoreactive crosslinker (**5**) comprises a photoreactive phenyl azide group along with a bidentate electrophilic component, a functional group that specifically reacts with a pair of histidine or cysteine residues. The synthesis was

carried out based on the method reported for the site-specific PEGylation of proteins [29]. As shown in Scheme 1, the aldol condensation and Michael addition reactions of formaldehyde, 4'-nitroacetophenone, and 4-methylbenzenethiol occurred in ethanol at room temperature in the presence of piperidine as a catalyst and resulted in the formation of the condensation product (**2**). Reduction of the nitro group of **2** by iron in the presence of acetic acid, following azidation with sodium azide and sodium nitrite, led to the generation of the azide derivative **4**. The oxidation of sulfides with oxone in EtOH/H_2O generated compound **5** (1,1-ditosylmethyl-*p*-azidoacetophenone: diTAAP).



Scheme 1. Synthesis of diTAAP (**5**). (a) formaldehyde, 4-methylbenzenethiol, piperidine, EtOH, reflux, 6 h, 58%; (b) AcOH, Fe, EtOH, reflux (under N_2), 6 h, 96%; (c) $NaNO_2$, NaN_3, 20% HCl, 0 °C, 4 h, 60%; (d) Oxone, MeOH:H_2O (1:1), r.t., 1 d.

Optimization of reaction conditions for modification of SelenoF with diTAAP

In order to optimize the conditions for the SelenoF modification, *Escherichia coli*-expressed SelenoF was prepared as a recombinant fused to a His-tagged small ubiquitin-like modifier (SUMO) at its N-terminus. Monitoring the reaction with diTAAP was carried out through the fluorescence labeling of the phenyl azide component via Staudinger ligation [22,30] using a triphenylphosphine derivative with a fluorophore (Figure 2a). The recombinant SelenoF was incubated with 400 μM diTAAP for 24 h at 4 °C, 24 °C, and 37 °C in a phosphate buffer (pH 8.0). After being incubated with 47 μM DyLight 488-Phosphine for 24 h at room temperature, the mixture was analyzed by SDS-PAGE and fluorescence imaging. As shown in Figure 2b, a major Coomassie brilliant blue (CBB)-stained band corresponding to SelenoF (approximately 30 kDa) was clearly visible by fluorescence, indicating that diTAAP was incorporated into SelenoF as expected. Next, we examined the optimal amount of diTAAP for protein modification. As shown in Figure 2c, 200 μM diTAAP was sufficient for protein modification. The recombinant SelenoF harbored a His-tag at the N-terminus and eight cysteines in the region spanning 50% of the C-terminus. Since α,β-unsaturated β′-mono-sulfone was reported to bridge both contiguous cysteine and histidine residues [31], the selectivity of modification by diTAAP was not immediately clear. In order to determine the modification site, the His-SUMO-SelenoF was subjected to cleavage using recombinant SUMO protease 1, which can recognize the tertiary structure of SUMO and remove SUMO from the SUMO fusion protein. The two fragments generated by SUMO protease 1 were labeled with a fluorophore, which indicated that diTAAP was introduced both into a cysteine diad in the SelenoF portion and into the His-tag in the SUMO portion.

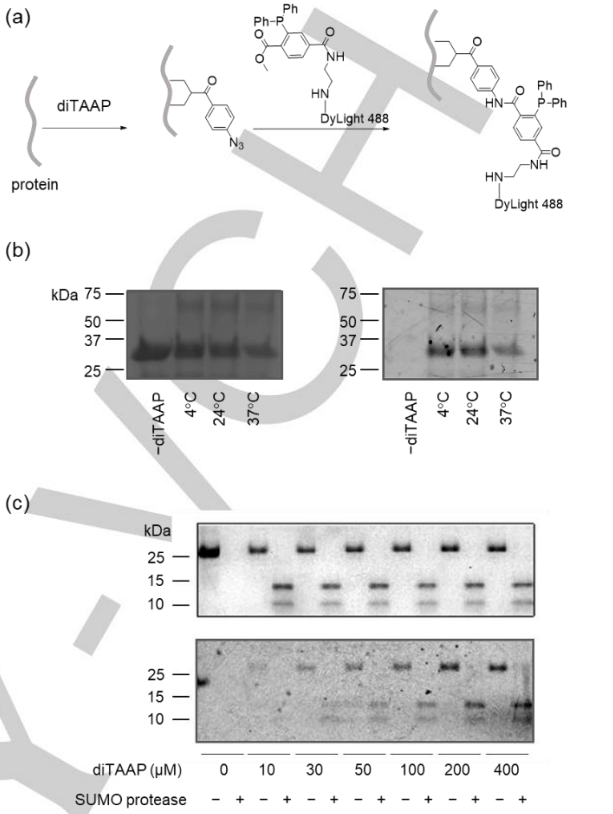


Figure 2. (a) Schematic illustration of modification of the recombinant SelenoF with diTAAP and subsequent fluorescence labeling with DyLight 488-phosphine. (b) Temperature dependence of the reaction. The left image represents CBB staining, and the right image shows immunofluorescence detection. (c) diTAAP concentration dependence of the reaction, and SUMO protease digestion of fluorophore-labeled SelenoF. The upper image represents CBB staining, and the lower image shows immunofluorescence detection.

Detection of SelenoF-HUGT1 crosslinking by diTAAP

To extract peptides photo-crosslinked by diTAAP, we constructed a plasmid vector for SelenoF having an HA-tag immediately next to the His-tag, as shown in Figure 3a. The recombinant SelenoF (U96C) and HUGT1 carrying a FLAG-tag were co-expressed in 293T cells. After purification with anti-FLAG antibody-conjugated affinity beads, the protein complex was removed by the addition of the 3XFLAG peptide. diTAAP in dimethyl sulfoxide (DMSO) was added to the protein mixture and incubated for 18 h at 4 °C. Subsequently, the mixture was analyzed by performing SDS-PAGE after UV irradiation (254 nm). Proteins crosslinked to SelenoF were visualized using WB with HA-tag specific antibody. As shown in Figure 3b, the new products in lane 4 had an apparent molecular mass of approximately 200 kDa, reflecting the molecular mass increase corresponding to SelenoF. Further analysis using WB revealed that the band was responsive to an anti-HA antibody, while no such band was observed in the no-UV control (Lane 3). The results of the experiment showed that diTAAP was photo-crosslinked between SelenoF and HUGT1, although its efficiency is unknown, because it has not been compared to other crosslinking methods.

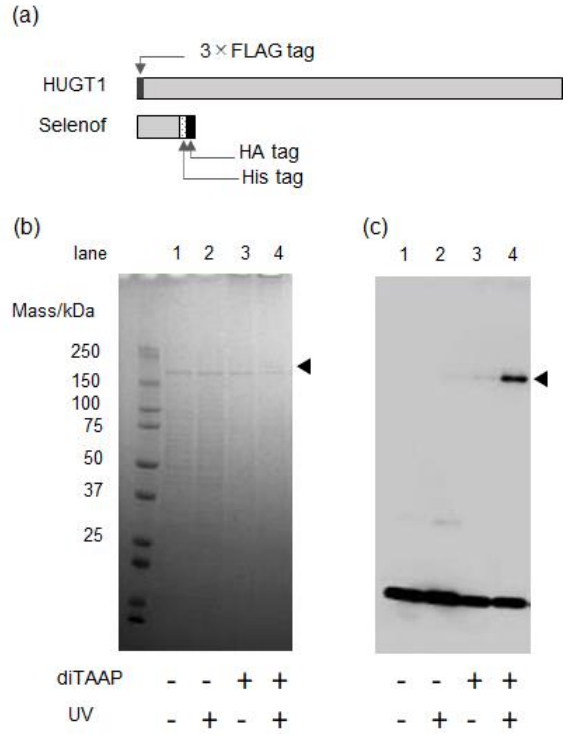


Figure 3. Detection of SelenoF-HUGT1 crosslinking by diTAAP (a) Illustration of N-terminal-FLAG-tagged HUGT1 and C-terminal-His/HA-tagged recombinant SelenoF. (b) SDS-PAGE analysis of recombinant HUGT1/SelenoF purified with anti-FLAG antibody-conjugated agarose beads. The black arrow represents HUGT1, which formed a covalent bond with SelenoF. (c) Immunoblotting analysis of the same gel as in (b) using an HRP-conjugated-anti-HA antibody.

Determination of the SelenoF-interacting sites of HUGT1 crosslink

In the 293T cell expression system, SelenoF was only obtained by co-expression with HUGT1; hence diTAAP was added to the HUGT1-SelenoF complex, and photo-crosslinking was subsequently performed. To determine the sites covalently modified by diTAAP, labeled HUGT1 was subjected to in-gel digestion with trypsin and proteinase K, and the resulting fragments were extracted. The eluent was analyzed using a liquid reverse-phase C18 chromatography (LC) system equipped with a UV detector connected to an electrospray tandem mass spectrometry (ESI-MS^2) system (Supporting Information, Figure S4-S6). The MS data that inferred the introduction of diTAAP into the cysteine residues of SelenoF were limited; however, a doubly charged ion was detected at m/z 1418.6642 and 1397.0941, which corresponded to the doubly protonated ion of peptide A and B (calculated for $[M + 2H]^+$ m/z 2835.3129 and 2792.1846), respectively (Figure 4a). This peptide was assigned to derive from the diTAAP-mediated covalent bond formation between V938-L953 of HUGT1 and C41-L50 of SelenoF (peptide A), and between S1222-K1237 of HUGT1 and C96-F101 of SelenoF (peptide B). Peptide B contains an N-glycosylation consensus sequence, which, however, not likely to be glycosylated [32]. This result indicated that diTAAP introduced into the cysteine residues of SelenoF covalently binds to peptides existing in the predicted TRXL3 and interdomain between βS2 and GT24 in HUGT1 (Figure 4b). In contrast, no crosslinking with diTAAP introduced into the His-tag of recombinant SelenoF was detected. Notably, in peptide B, the cysteine residue into which diTAAP was introduced is part of the redox-active triad CGU, although U is replaced by C in the present study. This suggests that the redox-active center is accessible to various client glycoproteins across the long hydrophobic cavity of HUGT1. A recent study using crystal structure analysis and molecular dynamic simulations of *Ct*UGGT proposed that twists and bends in the TRXL2 domain allow UGGT to accept various substrates in the saddle-like groove [33]. Furthermore, the maximum distance between the misfolded region of client glycoprotein and the N-glycan that could be re-glucosylated by UGGT was defined as the Parodi limit [33]. Taken together with previous studies showing that SelenoF functions as a thioredoxin-like oxidoreductase in the ER [34–36], our results provide evidence that SelenoF can access misfolding sites of client glycoproteins existing within the Parodi limit and might function as an oxidoreductase for misfolded client proteins (Figure 4c).

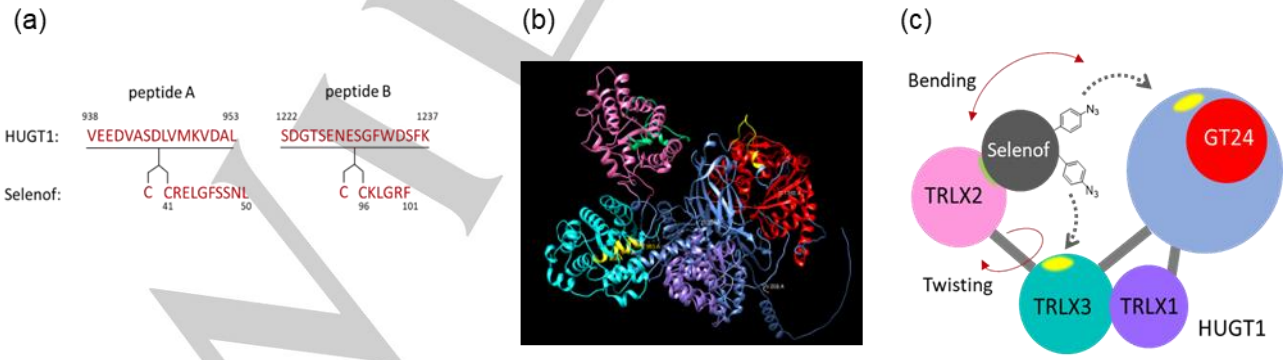


Figure 4. SelenoF-interacting sites of HUGT1 (a) Structure of crosslinked peptides detected using the LC-MS/MS system. (b) Predicted structure of HUGT1 calculated using AlphaFold 2. Purple, TRXL1; cyan, TRXL3; pink, TRXL2; red, GT24; green, predicted SelenoF-binding region; yellow, peptides crosslinked with SelenoF. (c) Schematic representation of the HUGT1–SelenoF structure based on the structure of *Ct*UGGT revealed by Modenutti et al. 2021.

Conclusion

In summary, this work reports evidence that will help clarify the SelenoF-binding site of UGGTs as well as the relative orientation of these proteins. A series of pulldown assays with truncated mutants of HUGT1 and SelenoF revealed that SelenoF binds to the TRXL2 region of UGGT. Moreover, photo-crosslinking experiments using a novel crosslinker (**5**) that was selectively introduced into the cysteine residues of recombinant SelenoF revealed that SelenoF formed a covalent bond with amino acids in the TRXL3 region and the interdomain between βS2 and GT24 of HUGT1, suggesting the cooperative action of SelenoF with these sites. These results are intriguing as the presence of folding sensors in TRXL3 [37] and GT24 [38] domains of UGGT has been proposed. We predict that the photoreactive crosslinking technique will find further use in analyzing the interaction between cysteine-rich proteins and their binding partners.

Experimental Section

Construction of expression vectors and mutagenesis

N-terminal- and C-terminal-truncated HUGT1 constructs, which were inserted between the NotI and XbaI sites of p3XFLAG-CMV-9 (Sigma-Aldrich), were constructed as previously reported [38]. The sequences of truncated HUGT1 proteins were also inserted between the SalI and BglII sites of pCMV-HA-N using the In-Fusion HD cloning kit (Takara Bio, Kusatsu, Japan).

Cell culture, transfection, and protein purification

293T human embryonic kidney cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum and penicillin (50 μg/mL) in a 100 mm dish. Cell transfection was performed using PEI MAX (Polysciences, Inc. Warrington, PA). After a 6 h incubation at 37 °C with 5% CO_2, cells were washed once with the medium and then incubated further for 42 h. Cultured cells were harvested and collected by centrifugation at 4000 × *g* for 5 min at 4 °C. The cells were then lysed and centrifuged at 8000 × *g* for 15 min at 4 °C. Next, the supernatants were incubated with anti-FLAG M2 agarose beads (Sigma-Aldrich) for 2 h at 4 °C. The suspensions were placed on disposable empty columns and washed with a wash buffer containing 50 mM tris-HCl (pH 7.4), 150 mM NaCl, 1 mM EDTA, and 1% Triton X-100. Proteins on anti-FLAG M2 agarose beads were eluted with 3XFlag-peptide.

Pulldown assay

293T cells co-transfected with pDNA coding myc-tagged SelenoF (6 μg) and pDNA coding FLAG-tagged HUGT1 mutants (6 μg) were cultured in 10-mm dish as mentioned above. After 48 h, they were harvested and collected by centrifugation at 4000 × g for 5 min at 4 °C. The cells were then lysed and centrifuged at 8000 × g for 15 min at 4 °C. The supernatants were incubated with anti-FLAG M2 agarose beads (15 μL) for 2 h at 4 °C. The suspensions were placed on Pierce disposable empty columns (Thermo Fisher Scientific, Waltham, MA) and washed thrice with a buffer containing 50 mM tris-HCl (pH 7.4), 150 mM NaCl, 1 mM EDTA, and 1% Triton X-100. The suspension of agarose beads was analyzed with SDS-PAGE and WB with an Anti-Myc-tag mAb-HRP-DirecT (MBL, Tokyo, Japan).

Construction of a vector for expression in E. coli

His_6-SUMO-tagged SelenoF (29-165, U96C) cDNA (for the gene sequence, see Supporting Information, Figure S7) was synthesized and subcloned into the BamHI/PmlI sites of the pET302NT-His vector using GeneArt Gene Synthesis (Thermo Fisher Scientific, Rockford, IL, USA).

Purification of His6-SUMO-tagged SelenoF expressed in E. coli

Cells were dissolved in 30 mL of lysis buffer (20 mM tris-HCl, 300 mM NaCl, 1 mM $CaCl_2$, pH 8.0) containing a protease inhibitor cocktail (cOmplete ™, EDTA-free, Roche Diagnostics). The suspension was sonicated (10 times) and centrifuged (19,000 × g, 60 min, 4 °C). Protein purification was performed using Ni-NTA agarose (Qiagen, Venlo, Netherlands) according to the manufacturer's instructions.

Modification of His-SUMO-SelenoF with diTAAP and DyLight 488

Protein solution eluted with 200 mM imidazole was replaced with 50 mM sodium phosphate buffer (pH 7.9) by dialysis (MEMBRA-CEL DIALYSIS MEMBRANES, cut-off 14000, Viskase, Lombard, IL) and concentrated with ultrafiltration (Vivaspin, Sartorius Stedim Biotech, Göttingen, Germany). diTAAP was dissolved in DMSO to 4.1 mg/mL. In 1.5-mL microtubes, 50 μL of His-SUMO-SelenoF solution was added at a concentration of 0.305 mg/mL, mixed with 40 equivalents of diTAAP, and incubated at 4˚C, 24˚C, and 37˚C for 24 h. To the reaction solution, 1 μL of 2.5 mM DyLight 488-phosphine (Thermo Fisher Scientific) in DMSO was added, followed by incubation at 24˚C for 24 h. Fluorescence was detected using a UV transilluminator (wavelength 312 nm) on an ImageQuant LAS 4000 (GE Healthcare). Gels were continuously stained using Quick CBB Plus (Fujifilm Wako Pure Chemicals, Osaka, Japan) and captured with ImageQuant LAS 4000.

SUMO protease digestion

To 10 μL of Staudinger reaction mixture, 10 μL of 0.263 mg/mL SUMO protease 1 (50 mM sodium phosphate buffer, pH 8.0) was added, followed by incubation at 25˚C for 1.5 h. The reaction was analyzed with a UV transilluminator and CBB staining.

Modification of SelenoF-HUGT1 complex with diTAAP and Photo-crosslinking

293T cells co-transfected with pDNA coding HA-tagged SelenoF (6 μg) and pDNA coding FLAG-tagged HUGT1 mutants (6 μg) were cultured in 100-mm dish as mentioned above. After 48 h, they were harvested and collected by centrifugation at 4000 × g for 5 min at 4 °C. The cells were then lysed and centrifuged at 8000 × g for 15 min at 4 °C. The supernatants were incubated with anti-FLAG M2 agarose beads (15 μL) for 2 h at 4 °C. The suspensions were placed on Pierce disposable empty columns (Thermo Fisher Scientific, Waltham, MA) and washed thrice with a buffer containing 50 mM tris-HCl (pH 7.4), 150 mM NaCl, 1 mM EDTA, and 1% Triton X-100. The suspension of agarose beads was incubated with a buffer containing 3×FLAG peptide (Protein Ark, Rotherham, UK) for 3 h at 4 °C. For this, 15 mg/mL diTAAP in DMSO was added to the protein solution to a final concentration of 0.15 mg/mL. The samples were then incubated at 37 °C for 18 h in the dark. After incubation, the sample was transferred to 96-well flat-bottom polystyrene plates (AGC Techno Glass, Yoshida, Japan) and UV-irradiated at 254 nm at room temperature using a CL-1000 UV Crosslinker (Analytik Jena, Jena, Germany). This was confirmed by SDS-PAGE and WB using anti-HA-tag mAb-HRP-DirecT (Medical and Biological Laboratories, Nagoya, Japan)

Proteinase digestion of crosslinked SelenoF-HUGT1

The gel corresponding to a molecular weight of approximately 200 kDa was cut for in-gel digestion. The excised piece was subjected to in-gel digestion. Briefly, destaining solution (10 mM tris-HCl/MeCN=1:1) was applied at room temperature for 10 min. Further, 10 mM DTT was applied

at 56 °C for the reduction of disulfide bonds and 55 mM iodoacetamide was added at room temperature in the dark for the alkylation of cysteine residues. Trypsin digestion was performed at 37 °C for 3 h followed by proteinase K digestion at 37 °C for 5 h. After digestion, peptides were extracted from the gel pieces by adding 400 µL 5% formic acid/ MeCN (v/v) and incubating the sample for 15 min at 37 °C. Subsequently, the supernatants were collected and dried using a Centrifugal concentrator CC-105 (TOMY Seiko Co. Ltd., Japan). The dried extracts were dissolved in 100 µL 5% MeCN in H_2O containing 0.5% TFA. Finally, tryptic peptides were eluted using 20 µL 50% ACN in H_2O with 0.1% formic acid for MS analysis.

Reversed-phase HPLC fractionation and LC−MS analysis of the digested peptide mixture

The peptide mixture was separated using a Nano Frontier nLC nanoflow HPLC system (EASY-nLC 1200 System; Thermo Fisher Scientific Inc., USA). A tapered capillary column was used as a sprayer tip and filled with C18 silica particles (particle size of 3 µm, 75 µm (inside diameter) × 150 mm (length); Thermo Fisher Scientific Inc., USA). Mobile phase A comprised 0.1% formic acid in water, and mobile phase B comprised 0.1% formic acid in 80 % MeCN (LC−MS grade; Merck, Darmstadt, Germany). The gradient conditions were as follows: 0 to 10 min, linear gradient from 5 to 35% B; 10 to 13 min, linear gradient from 35 to 100% B; 13 to 20 min, isocratic 100% B. The samples were analyzed using a quadrupole ion trap mass spectrometer (QIT-MS) coupled with a nanoelectrospray interface (Q Exactive Plus, Thermo Fisher Scientific Inc., USA) attached to a UV detector. The parameters for analysis were as follows: (1) dry temperature, 120 °C; (2) dry gas (N_2), 3.0 L/min; (3) scan range, m/z 600−2250; (4) compound stability, 100%; (5) target mass, m/z 1750; (6) ion charge control, on, target, 400000; (7) maximal accumulation time, 200 ms; (8) average, five spectra; and (9) polarity, positive. In the MS/MS experiments, the end-cap radiofrequency amplitude was 1.2 V, and the isolation width was m/z 4.0.

Acknowledgments

We would like to thank Dr. Chiaki Yoshioka (APRO Science) for performing LC-MS/MS analysis and Ms. Fumie Tanaka and Mr. Ryoma Aoki for their technical assistance. This work was supported by the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (JSPS KAKENHI), grant numbers JP16H06290 (YI), and JP21K05282 (YT).

Keywords: Glycoproteins • Photoaffinity labeling • Protein-protein interactions • Selenoprotein F • UGGT

References

[1] K. Olden, B.A. Bernard, M.J. Humphries, T.-K.K. Yeo, K.-T.T. Yeo, S.L. White, S.A. Newton, H.C. Bauer, J.B.B. Parent, *Trends Biochem. Sci.* **1985**, *10*, 78–82. DOI: 10.1016/0968-0004(85)90238-5.

[2] E. Bause, *Biochem. J.* **1983**, *209*, 331–336. DOI: 10.1042/bj2090331.

[3] A. Helenius, M. Aebi, *Annu. Rev. Biochem*. **2004**, *73*, 1019–1049. DOI: 10.1146/annurev.biochem.73.011303.073752.

[4] L.W. Ruddock, M. Molinari, *J. Cell Sci*. **2006**, *119*, 4373–4380. DOI: 10.1242/jcs.03225.

[5] H. Nakao, A. Seko, Y. Ito, M. Sakono, *Biochem. Biophys. Res. Commun*. **2017**, *487*, 763–767. DOI: 10.1016/j.bbrc.2017.04.139.

[6] G. Kozlov, K. Gehring, *FEBS J.* **2020**, *287*, 4322–4340. DOI: 10.1111/febs.15330.

[7] S.M. Arnold, R.J. Kaufman, *J. Biol. Chem*. **2003**, *278*, 43320–43328. DOI: 10.1074/jbc.M305800200.

[8] M.B. Prados, J.J. Caramelo, S.E. Miranda, *Biochim. Biophys. Acta Mol. Cell Res.* **2013**, *1833*, 3368–3374. DOI: 10.1016/j.bbamcr.2013.09.022.

[9] B.M. Adams, N.P. Canniff, K.P. Guay, I.S.B. Larsen, D.N. Hebert, *eLife* **2020**, *9*, 1–27. DOI: 10.7554/eLife.63997.

[10] J.J. Caramelo, A.J. Parodi, *J. Biol. Chem*. **2008**, *283,* 10221–10225. DOI: 10.1074/jbc.R700048200.

[11] C. D'Alessio, J.J. Caramelo, A.J. Parodi, *Semin. Cell Dev. Biol*. **2010**, *21*, 491–499. DOI: 10.1016/j.semcdb.2009.12.014.

[12] A. Tannous, G.B. Pisoni, D.N. Hebert, M. Molinari, *Cell Dev. Biol*. **2015**, *41*, 79–89. DOI: 10.1016/j.semcdb.2014.12.001.

[13] A. Zapun, S.M. Petrescu, P.M. Rudd, R.A. Dwek, D.Y. Thomas, J.J.M. Bergeron, *Cell* **1997**, *88*, 29–38. DOI: 10.1016/S0092-8674(00)81855-3.

[14] L.A. Rutkevich, D.B. Williams, *Curr. Opin. Cell Biol*. **2011**, *23*, 157–166. DOI: 10.1016/j.ceb.2010.10.011.

[15] N. Wang, A. Seko, Y. Takeda, Y. Ito, *Biochim. Biophys. Acta Gen. Subj*. **2020**, *1864*, 129709. DOI: 10.1016/j.bbagen.2020.129709.

[16] A.D. Ferguson, V.M. Labunskyy, D.E. Fomenko, D. Araç, Y. Chelliah, C.A. Amezcua, J. Rizo, V.N. Gladyshev, J. Deisenhofer, *J. Biol. Chem*. **2006**, *281*, 3536–3543. DOI: 10.1074/jbc.M511386200.

[17] Z. Zhao, R. Mousa, N. Metanis, *Chem. Eur. J*. **2022**, *28*, e202200279. DOI: 10.1002/chem.202200279.

[18] Y. Takeda, A. Seko, M. Hachisu, S. Daikoku, M. Izumi, A. Koizumi, K. Fujikawa, Y. Kajihara, Y. Ito, *Glycobiology* **2014**, *24*, 344–350. DOI: 10.1093/glycob/cwt163.

[19] V.M. Labunskyy, A.D. Ferguson, D.E. Fomenko, Y. Chelliah, D.L. Hatfield, V.N. Gladyshev, *J. Biol. Chem*. **2005**, *280*, 37839–37845. DOI: M508685200 [pii] 10.1074/jbc.M508685200.

[20] K.L. Bennett, M. Kussmann, P. Björk, M. Godzwon, M. Mikkelsen, P. Sørensen, P. Roepstorff, *Protein Sci*. **2000**, *9*, 1503–1518. DOI: 10.1110/ps.9.8.1503.

[21] E. Scriven, Academic Press Inc., New York, **1984**.

[22] F.L. Lin, H.M. Hoyt, H. van Halbeek, R.G. Bergman, C.R. Bertozzi, J. *Am. Chem. Soc*. **2005**, *127*, 2686–2695. DOI: 10.1021/ja044461m.

[23] J. Liang, L. Zhang, X.-L. Tan, Y.-K. Qi, S. Feng, H. Deng, Y. Yan, J.-S. Zheng, L. Liu, C.-L. Tian, *Angew. Chem. Int. Ed*. **2017**, *56*, 2744–2748. DOI: 10.1002/anie.201611659.

[24] J.W. Chin, A.B. Martin, D.S. King, L. Wang, P.G. Schultz, *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11020–11024. DOI: 10.1073/pnas.172226299.

[25] Y. Cong, E. Pawlisz, P. Bryant, S. Balan, E. Laurine, R. Tommasi, R. Singh, S. Dubey, K. Peciak, M. Bird, A. Sivasankar, J. Swierkosz, M. Muroni, S. Heidelberger, M. Farys, F. Khayrzad, J. Edwards, G. Badescu, I. Hodgson, C. Heise, S. Somavarapu, J. Liddell, K. Powell, M. Zloh, J.W. Choi, A. Godwin, S. Brocchini, *Bioconjug. Chem*. **2012**, *23*, 248–263. DOI: 10.1021/bc200530x.

[26] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Žídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, S. Velankar, *Nucleic Acids Res*. **2022**, *50*, D439–D444. DOI: 10.1093/nar/gkab1061.

[27] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **2021**, *596*, 583–589. DOI: 10.1038/s41586-021-03819-2.

[28] D. Calles-Garcia, M. Yang, N. Soya, R. Melero, M. Ménade, Y. Ito, J. Vargas, G.L. Lukacs, J.M. Kollman, G. Kozlov, K. Gehring, *J. Biol. Chem*. **2017**, *292*, 11499–11507. DOI: 10.1074/jbc.M117.789495.

[29] S. Balan, J.W. Choi, A. Godwin, I. Teo, C.M. Laborde, S. Heidelberger, M. Zloh, S. Shaunak, S. Brocchini, *Bioconjug. Chem*. **2007**, *18*, 61–76. DOI: 10.1021/bc0601471.

[30] J.A. Restituyo, L.R. Comstock, S.G. Petersen, T. Stringfellow, S.R. Rajski, *Org. Lett*. **2003**, *5*, 4357–4360. DOI: 10.1021/ol035635s.

[31] K. Peciak, E. Laurine, R. Tommasi, J.W. Choi, S. Brocchini, *Chem. Sci*. **2019**, *10*, 427–439. DOI: 10.1039/C8SC03355B.

[32] S. Daikoku, A. Seko, Y. Ito, O. Kanie, *Biochem. Biophys. Res. Commun*. **2014**, *451*, 356–360. DOI: 10.1016/j.bbrc.2014.07.095.

[33] C.P. Modenutti, J.I. Blanco Capurro, R. Ibba, D.S. Alonzi, M.N. Song, S. Vasiljević, A. Kumar, A. V. Chandran, G. Tax, L. Marti, J.C. Hill, A. Lia, M. Hensen, T. Waksman, J. Rushton, S. Rubichi, A. Santino, M.A. Martí, N. Zitzmann, P. Roversi, *Structure* **2021**, *29*, 357–370.e9. DOI: 10.1016/j.str.2020.11.017.

[34] V.M. Labunskyy, D.L. Hatfield, V.N. Gladyshev, *IUBMB Life* **2007**, *59*, 1–5. DOI: 10.1080/15216540601126694.

[35] S.H. Yim, R.A. Everley, F.A. Schildberg, S.G. Lee, A. Orsi, Z.R. Barbati, K. Karatepe, D.E. Fomenko, P.A. Tsuji, H.R. Luo, S.P. Gygi, R. Sitia, A.H. Sharpe, D.L. Hatfield, V.N. Gladyshev, *Cell Rep*. **2018**, *23*, 1387–1398. DOI: 10.1016/j.celrep.2018.04.009.

[36] M. V. Kasaikina, D.E. Fomenko, V.M. Labunskyy, S.A. Lachke, W. Qiu, J.A. Moncaster, J. Zhang, M.W. Wojnarowicz, S.K. Natarajan, M. Malinouski, U. Schweizer, P.A. Tsuji, B.A. Carlson, R.L. Maas, M.F. Lou, L.E. Goldstein, D.L. Hatfield, V.N. Gladyshev, *J. Biol. Chem*. **2011**, *286*, 33203–33212. DOI: 10.1074/jbc.M111.259218.

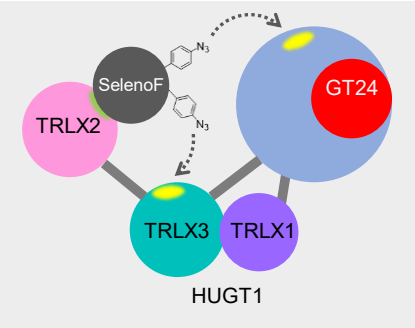
[37] T. Zhu, T. Satoh, K. Kato, *Sci. Rep*. **2015**, *4*, 7322. DOI: 10.1038/srep07322.

[38] K. Ohara, Y. Takeda, S. Daikoku, M. Hachisu, A. Seko, Y. Ito, *Biochemistry* **2015**, *54*, 4909–4917. DOI: 10.1021/acs.biochem.5b00785.

[39] Y. Takeda, A. Seko, K. Fujikawa, M. Izumi, Y. Kajihara, Y. Ito, *Glycobiology* **2016**, *26*, 999–1006. DOI: 10.1093/glycob/cww069.

[40] F.A. Liberatore, R.D. Comeau, J.M. McKearin, D.A. Pearson, B.Q. Belonga, S.J. Brocchini, J. Kath, T. Phillips, K. Oswell, R.G. Lawton, *Bioconjug. Chem*. **1990**, *1*, 36–50. DOI: 10.1021/bc00001a005.

Entry for the Table of Contents



Selenoprotein F binds to the TRXL2 domain of human UGGT1, and it forms a covalent bond with amino acids in the TRXL3 region and the interdomain between βS2 and GT24 of human UGGT1 via a synthetic crosslinker.