Host-Dependent Producibility of Recombinant *Cypridina noctiluca* Luciferase With Glycosylation Defects

Yasuo Mitani1*, Rie Yasuno2, Kiyohito Kihira3, KwiMi Chung4, Nobutaka Mitsuda4, Shusei Kanie1, Azusa Tomioka2, Hiroyuki Kaji2 and Yoshihiro Ohmiya5,6

1Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Sapporo, Japan, 2Cellular and Molecular Biotechnology Research Institute, AIST, Tsukuba, Japan, 3Japan Aerospace Exploration Agency (JAXA), Tsukuba, Japan, 4Bioproduction Research Institute, AIST, Tsukuba, Japan, 5Biomedical Research Institute, AIST, Ikeda, Japan, 6Osaka Institute of Technology (OIT), Osaka, Japan

*Correspondence:
Yasuo Mitani
mitani-y@aist.go.jp

Specialty section:
This article was submitted to Biomaterials, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: 13 September 2021
Accepted: 17 January 2022
Published: 07 February 2022

Citation:
Mitani Y, Yasuno R, Kihira K, Chung K, Mitsuda N, Kanie S, Tomioka A, Kaji H and Ohmiya Y (2022) Host-Dependent Producibility of Recombinant *Cypridina noctiluca* Luciferase With Glycosylation Defects. Front. Bioeng. Biotechnol. 10:774786. doi: 10.3389/fbioe.2022.774786

Introduction

Luciferase is an enzyme that oxidizes the substrate luciferin; some are cellular proteins, while others are secreted. Secreted luciferases have been reported in a wide range of phyla including Arthropoda, Mollusca, and Annelida (Shimomura 2012). Among them, luciferase genes have been cloned from Copepoda *Gaussia princeps* (Bryan and Szent-Gyorgyi 2001) and *Metridia longa* (Markova et al., 2004), from Ostracoda *Vargula hilgendorfii* (Thompson et al., 1989), *Cypridina noctiluca* (Nakajima et al., 2004), and Caribbean species (Hensley et al., 2021); from the deep-sea shrimp *Oplophorus gracilirostris* (Inouye et al., 2000); and from syllid worms like *Odontosyllis undecimcostata* (Mitani et al., 2018; Schultz et al., 2018) and related species (Mitani et al., 2019). Secreted luciferases are useful for high-throughput analysis because the assay is made less difficult by using culture supernatants.
Thus far, luciferases from the genera *Oplophorus* and *Gaussia* have been widely used in application research (Hall et al., 2012; Suzuki and Inouye 2014; Markova et al., 2019). *Oplophorus* luciferase (OLuc) has some advantages, such as its catalytic domain with low molecular weight, excellent heat stability, and relatively easy expression in both prokaryotic and eukaryotic cells when some amino acid mutations are incorporated (Hall et al., 2012; Inouye et al., 2013). In the case of *Gaussia* luciferase (GLuc), a production method for low-temperature expression in the presence of chaperones has been established in an *Escherichia coli* expression system (Inouye and Sahara 2008). The three-dimensional (3D) structures of these luciferases have been elucidated by crystallization and X-ray diffraction of the partial domain exhibiting the catalytic function in OLuc (Tomabechi et al., 2016) and by heteronuclear NMR for GLuc (Wu et al., 2020).

*C. noctiluca* luciferase (CLuc) has been developed for applied research by modifying its surface with fluorescent molecules to emit infrared light (Wu et al., 2009) and by optimizing it for a high-throughput reporter assay (Tochigi et al., 2010). CLuc is also a stable protein with high thermal stability and has high potential for further applications. However, CLuc has a higher molecular weight (62–68 kDa) than OLuc and GLuc, and may have up to 17 potential disulfide bonds between the 34 cysteine residues in its 553 amino acids, while GLuc has only 5 disulfide bonds (Wu et al., 2020), thus making protein production and applied research with CLuc difficult (Inouye and Sahara 2008). There are no reports on the prokaryotic expression of full-length CLuc. Although CLuc does not show homology to other luciferases, its first- and second-half regions display weak homology to the von Willebrand factor type D (VWD) domain (Oakley 2005). These two VWD-like domains show approximately 20% homology to each other. The first half, containing 302 amino acids, has been expressed using an *E. coli* system, resulting in 45% activity of the intact protein (Hunt et al., 2017). However, there are no biochemical reports on the residues directly involved in CLuc activity, and there is no information on its 3D structure. The establishment of a system that can stably and massively produce CLuc is expected to lead to further progress in crystallographic analysis and future applications. Structural information will also help to modify CLuc using a simple approach to bioconjugation (Hu et al., 2018).

Secrated proteins in addition to luciferases need to undergo appropriate folding, including disulfide bond formation during the secretory process, and most secreted proteins undergo post-translational modification such as glycosylation (Helenius and Aebi 2001). Although glycosylation is generally considered to contribute to protein stability by increasing folding efficiency (Helenius and Aebi 2001; Wang et al., 2018), its effects on protein function itself are inconsistent. For example, it was reported that some hormones maintain their ability to bind to receptors when glycosylation is removed by hydrogen fluoride (HF) treatment, but the downstream signals are not transduced (Lapthorn et al., 1994). In another case, human follicle-stimulating hormone (hFSH) remains fully active even after glycan chains are removed via point mutation (Fox et al., 2001); the folding of these glycosylation-deficit mutants is unaffected (Wilbers et al., 2016). Since glycosylation can prevent crystallization due to its heterogeneity (Lapthorn et al., 1994; Columbus 2015), glycans can be removed by deglycosylation enzyme treatment or mutagenesis to the modification site (Woods 2018). CLuc is the only secreted luciferase whose glycan structure has been analyzed and reported to remain fully active even after enzymatic removal of the glycan from wild-type (Wt) recombinant protein produced in suspension-cultured tobacco Bright Yellow-2 (BY-2) cells (Mitani et al., 2017). In the case of mutant CLuc lacking glycan modification sites expressed in COS1 cells, its specific activity has been reported to be reduced to approximately 20% when compared to the Wt recombinant protein produced in COS1 cells. However, due to the low productivity, quantitative analysis using purified CLuc proteins remains to be conducted (Yasuno et al., 2018).

In this study, we tried to produce mutant recombinant CLuc lacking N-glycan modification in available hosts in our laboratory, including Expi293F cells, silkworms, and BY-2 cells. CLuc has two conserved N-glycan modification motifs, NIT and NTS, in its amino acid sequence, starting at residues N182 and N404, respectively. These two motifs were mutated to produce a double mutant recombinant protein (Dmt CLuc) containing T184A + N404D substitutions. In comparing the specific activities of the obtained recombinant CLuc by measuring the protein amount using highly purified CLuc or anti-CLuc antibody, we found that the mutant CLuc produced by Expi293F cells showed almost the same specific activity as the Wt. We also found important amino acid residues involved in the luciferase activity by expressing amino-terminal serial deletion mutants of Dmt CLuc.

**MATERIALS AND METHODS**

**Recombinant Proteins Used in This Study**

In this study, we produced Wt and Dmt CLuc using Expi293F cells, silkworms, and BY-2 cells, designating the recombinant proteins CLucEX, CLucSW, and CLucBY, respectively. For clarity, we designated Dmt CLuc previously expressed using COS1 cells (Yasuno et al., 2018) as Dmt CLucCO in this study.

**Cypridina noctiluca Luciferase Expression Using Expi293 Expression System**

Expression vectors for recombinant Wt CLucEX and Dmt CLucEX pcDNA-CL (Nakajima et al., 2004), and pcDNA-CL (T184A + N404D) (Yasuno et al., 2018) were prepared as previously described. The pcDNA-CL (T184A + N404D) was modified by the insertion of the sequence for His-tag and TEV protease (HHHHHHHENLYFQG) following the predicted signal sequence (1–18 a.a.). Expression vectors for deletion mutants were designed with deleted amino-terminal or carboxyl-terminal residues of the mature region by modification of the His-TEV-pcDNA-CL (T184A + N404D) plasmid. These insertion and deletion modifications were outsourced to Genscript Japan. Each expression vector was transfected into Expi293F mammalian cells (Thermo Fisher Scientific, MA,
United States) according to the manufacturer's instructions. On day 5 post-transfection, the culture medium was centrifuged at 350 x g for 10 min, and the resulting supernatant was used for the following analysis, hereafter called the medium fraction. Precipitated cells were lysed with 10 mM Tris-HCl (pH 7.4) by sonication, the cell lysate was centrifuged at 20,000 x g for 5 min, and the resulting supernatant was used as the cell extract fraction. Immunoblot analysis for CLuc proteins was conducted as described previously (Yasuno et al., 2018), except that 10 μL of each medium fraction sample was prepared under reducing conditions with dithiothreitol before electrophoresis and that the anti-CLuc antibody was raised using purified recombinant CLuc in rabbit. The activities of the recombinant CLuc proteins were measured as described previously (Yasuno et al., 2018), and the intensities were obtained as relative light units (RLUs).

**LC-MS/MS Analyses of *Cypridina noctiluca* Luciferase Mutated at N-Glycosylation Sites**

Dmt CLuc\textsubscript{EX} was S-reduced and alkylated as described previously (Mitani et al., 2017). The protein was digested with a mixture of trypsin and Lys-C endopeptidase and/or chymotrypsin (TL/C/CTL). An aliquot of each digest was heated at 80°C for 2 h in 0.1% TFA to remove sialic acid and to increase the detection sensitivity of glycopeptides. Each digest was analyzed by an LC-MS system using a nanoflow LC (Ultimate 3000, Thermo Fisher Scientific) and Orbitrap Fusion Trividil mass spectrometer (Thermo Fisher Scientific). MS and MS/MS spectra were obtained by the Orbitrap analyzer with a resolution of 120,000 at 200 m/z. Data dependently selected precursor ions were fragmented by high-energy collision-induced dissociation (HCD) and HCD-fragment-triggered EThcD. Glycan peptide cluster analysis was performed by glycan-heterogeneity-based relational identification of glycopeptide signals on elution profile (Glyco-RIDGE; Togayachi et al., 2018). MS/MS spectra were searched using Mascot (ver. 2.5.1, Matrix Science, MA, United States) with the UniProt protein sequence database for Homo sapiens including sequences of CLuc (gi|41152712) with mutations (T184A + N404D). Search parameters were set as missed cleavage for chymotrypsin; 6.
transferred to a pBCKK T-DNA vector (Mitsuda et al., 2006) and transformed into *Rhizobium radiobacter* GV3101.

**RESULTS AND DISCUSSION**

**Recombinant *Cypridina noctiluca* Luciferase Expression Using Expi293 Expression System**

In a previous work, we reported the expression of recombinant CLuc deficient in N-glycosylation binding sites (Dmt CLu**c**CO) in a mammalian adherent cell, COS1. In the present study, we first focused on Expi293F cells, which have been adapted to grow in high density and to give higher recombinant protein yields, and tried a suspension culture of the Expi293 expression system, with the expectation of high expression of CLuc. Each expression vector for Wt or Dmt CLucEX was transfected into Expi293F cells. After 5 days of incubation, the accumulation of CLucEX in the medium was confirmed by immunoblot analysis using an anti-CLuc antibody (**Figure 1A**). Mobility shift differences between Wt and Dmt CLucEX were presumably due to deficiencies in N-glycans. Over 90% of the luciferase activities were observed in the medium in both Wt and Dmt CLucEX, indicating that the majority of CLuc produced in Expi293F cells was secreted into the medium efficiently. The activity of CLucEX was determined by measuring the luminescence intensity using synthetic cypridinid luciferin. The specific activities of Wt and Dmt CLucEX in the medium were calculated by dividing the luminescence intensities by the estimated intensities of their immunoblot CLucEX bands (**Figure 1B**). The visibility of the immunoblot band differed among multiple experiments, and the estimation range was broad. Thus, we performed 3 independent experiments and calculated the average of the specific activity ratio for Dmt CLucEX relative to that for Wt CLuc EX. A typical result is shown in **Figure 1B**, where the ratio was approximately 120%, but the average ratio was 90.7 ± 21.8%. These data showed that the specific activities of Wt and Dmt CLucEX were almost the same, unlike the previous result of approximately 20% for Dmt CLu**c**CO expressed in COS1 cells.

Confirmation of N-glycosylation site mutations and search for O-glycosylation site(s) of Dmt CLucEX expressed in Expi293F cells by LC-MS.

Dmt CLucEX protein samples digested with chymotrypsin and/or a mixture of trypsin and Lys-C endopeptidase were analyzed by LC-MS/MS, and the obtained MS/MS spectra (HCD) were searched using Mascot to identify Dmt CLucEX-derived peptides (**Supplementary Tables S1–S6**). For the trypsin + Lys-C (TL) digest, 81% of peptides were identified, and a peptide containing mutation T184A was included. In addition,
for the chymotrypsin digest, 87% of peptides were identified, including a peptide with an N404D mutation (Supplementary Tables S1, S3). These results indicate that both N-glycosylation sites were mutated and lost.

To search for other glycosylated site(s), EThcD MS/MS spectra were obtained by triggering with HCD fragments of m/z = 138, 204, or 366, that is, diagnostic fragment ions of glycan corresponding to [HexNAc-CH6O3]+, [HexNAc]+, or [HexHexNAc]+, respectively. Several O-glycosylated peptides, which were modified with glycan at a serine or threonine residue, were found by the Mascot search, peptide sequences were identified from HCD spectra, and glycosylated sites were assigned by EThcD spectra (Supplementary Tables S2, S6, Supplementary Figures S1–S5). In a TL digest, the HCD MS/MS spectrum of peptide sequence PPNTVPTSCEAK (residues 27–38) having a Hex(1) HexNAc(1) modification was found, presenting a series of diagnostic ions of glycan including a NeuAc-H2O fragment (m/z = 274) and a ladder-like signal of glycopeptides suggesting a NeuAc-Hex-HexNAc-peptide (Supplementary Figure S1). A partial sequence of the core peptide was assigned. In an acid-treated digest with chymotrypsin + trypsin + Lys-C (CTL), glycopeptides having the same core sequence, PPNTVPTSCEAK, carrying 1 and 2 Hex(1) HexNAc(1) modification(s), were identified. A partial peptide sequence could be assigned from their HCD spectra, but the modification sites were not identified because there are 3 potential Ser/Thr in the peptide (Supplementary Figure S2).

Glyco-RIDGE analysis of the acid-treated TL digest revealed a cluster of the core peptide having different glycan compositions, that is, members carrying glycans of Hex:HexNAc:dHex = 0:0:0 (no glycosylation), 0:1:0, 1:1:0, 1:2:0, 2:2:0, and 3:3:0 (data not shown). Extracted ion chromatograms are shown in Supplementary Figure S3 for the signals of 0:0:0, 0:1:0, 1:1:0, and 2:2:0. Signals of 1:2:0 and 3:3:0 were too weak to be seen in the figure. The signal of non-glycosylated peptide was highest even when the scale was reduced to 1/100, suggesting that the glycan occupancy of the peptide was quite low, about 2% of the area. The glycopeptide carrying Hex(1) HexNAc(1) showed twin broad peaks. The EThcD spectra of glycopeptide ions at 3 different retention times were compared (Supplementary Figure S4), revealing multiple O-glycosylated sites in the peptide sequence: Thr-30 and Thr-33 or Ser-34. MS/MS acquired at time (2) in Supplementary Figure S3 suggested that both Thr-30 and Thr-33 were glycosylated (Figure 2). The EThcD spectrum of the

FIGURE 3 | Luminescence activities of Dmt CLucEX deletion mutants. (A) Schematic representation of full-length Dmt CLucEX and its deletion mutants. Each construct was designed to have a natural signal peptide (SP 1–18) and His tag (His x 6)–TEV sequence at the amino terminus. The O-glycosylation sites are at 30 and 33. VWD-like domains are indicated by black bars at the top of the panel. (B) Immunoblot analysis of Dmt CLucEX deletion mutants. Expi293F cells were transfected with the indicated constructs, and the produced protein secreted into the medium was analyzed by Western blotting using anti-CLuc antibody. (C) Relative specific activity of each Dmt CLucEX deletion mutant. Specific CLucEX activities were determined as shown in Figure 1. Note that the band in lane 3 is very faint, probably due to the low expression level of this construct.
Deletion Mutants of O-Glycosylation Residues

Deletion mutants were constructed to remove the region around O-glycosylation residues at Thr-30 and Thr-33 (Figure 3A). Protein expression using the Exp293 expression system was confirmed by immunoblot analysis (Figure 3B), and specific activities were determined based on the immunoblot band intensity (Figure 3C). Specific activities were halved in mutants with deletions at residues 19–24 and 19–30 when compared to the full length of Dmt CLucEX. However, no activities were exhibited by the mutants with further deletions at residues 19–36, 19–42, or 19–48, even when the protein amounts were almost equal to those of the full-length protein. These results suggested that the amino acid residues at 31–36 (VPTSCE) were important for CLuc function. This region was highly conserved with the luciferase of the related cypridinid species, including cysteine residues (Supplementary Figure S6).

Moreover, the removal effects of carboxyl-terminal regions of Dmt CLucEX were investigated. A deletion mutant (19–314) designed to remove the second VWD-like domain exhibited no luciferase activity even though protein production was confirmed at the predicted size of 36 kDa (Figure 3B), suggesting that the residues at 315–553 were indispensable regions for CLuc function. Deletion of the residues at 453–553, which was designed to remove the carboxyl region behind the second VWD-like domain, resulted in a significant reduction of protein production, observed as a slight band around 50 kDa (Figure 3B). Deletion of the region at 453–553 might lead to the instability of the protein or to a reduction of secretion efficiency in this expression system.

Recombinant Cypridina noctiluca Luciferase Expression Using Silkworm

The recombinant proteins of Dmt and Wt CLucSW were produced using the silkworm expression system. First, each recombinant protein was simply purified using anti-FLAG affinity gel, and a single band was observed for each case in the SDS-PAGE analysis (Figure 4). Depending on the protein, treatment with a reducing agent differently affected the mobility in the gel, as shown previously (Mitani et al., 2017). The Wt CLucSW band looked larger than the Dmt CLucSW band, and this difference was more obvious under the non-reducing conditions (Figure 4). This size difference would be due to the glycan modification, suggesting that an N-glycan chain would be added to CLuc when produced in silkworm, similar to the case with COS1 and BY-2 cell expressions, as previously reported (Mitani et al., 2017; Yasuno et al., 2018). The luciferase activities of Dmt and Wt CLucSW were $4.24 \times 10^7$ and $1.67 \times 10^8$ RLU ng$^{-1}$, respectively. These data indicated that almost 5% activity was retained in the Dmt CLucSW even without glycan modification. These data were similar to those obtained using a mammalian expression system, which resulted in approximately 20% activity for Wt CLucCO (Yasuno et al., 2018). We also produced another mutant, N182D + N404D, and expected an improvement in its specific activity. This mutant recombinant protein was successfully produced (Figure 4), but its activity was almost the same as that of Dmt CLucSW ($4.18 \times 10^8$ RLU ng$^{-1}$). Thus, we decided to use Dmt CLucSW for further analysis.

Next, to improve the recombinant protein purity, each sample was successively purified by FLAG-tag affinity chromatography and anion exchange chromatography. The results showed that Dmt CLucSW was separated into two major peaks (Supplementary Figure S5A). Considering the peak area ratio of the chromatograph, the existence ratio of these fractions was estimated to be about 1:3. The two separated peaks of Dmt CLucSW were collected, and each sample was subjected to gel filtration chromatography. The higher molecular weight fraction was eluted in the void volume, and the lower one was eluted in the volume corresponding to the molecular weight of the CLucSW monomer. Next, electrophoresis under native conditions was performed on each of these fractions. As shown in Supplementary Figure S5B, the Wt CLucSW and the Dmt CLucSW, which were derived from the lower molecular weight fraction in the
gel filtration chromatography, showed a single band near the calculated molecular weight in each case, while the Dmt CLucSW, which was obtained as the void fraction in the gel filtration, showed a much higher molecular weight than the monomer. This aggregation is probably due to misfolding, since glycosylation is known to be involved in the correct folding of secreted proteins during translation (Wang et al., 2018). These aggregated Dmt CLucSW did not show any luciferase activity. These data seemed reasonable because only 25% of Dmt CLucSW was produced as a monomer, probably with normal activity, while the remaining 75% was aggregated during protein production, resulting in no luciferase activity. Thus, our first estimates of the specific activity of the Dmt CLucSW when purified only with anti-FLAG affinity were artificially low due to the presence of a high percentage of aggregated, non-functional luciferase. In the case of Dmt CLucCO expressed using COS1 cells, the specific activity was 20% of that of the Wt CLucCO (Yasuno et al., 2018). Although we did not perform further purification for this protein due to the limited amount of protein produced, 80% of the recombinant protein may have aggregated during protein production, as in the case of the silkworm expression system.

Recombinant Cypridina noctiluca Luciferase Expression Using Suspension-Cultured Tobacco BY-2 Cells

Wt CLucBY was efficiently expressed using cultured tobacco BY-2 cells (Mitani et al., 2017), and thus, we tried to produce Dmt CLucBY. However, none of the clones produced recombinant Dmt CLucBY even though we tested 4 independent BY-2 callus clones. This result suggested that the N-glycosylation of CLucBY was essential for its recombinant protein production in BY-2 cells.

CONCLUSION

In this study, we produced recombinant mutant CLuc lacking N-glycosylation motifs using Expi293F cells, silkworms, and tobacco BY-2 cells. The relative activity of the produced Dmt CLucEX was almost the same as that of the Wt CLucEX, whereas Dmt CLucCO was approximately 20% of that of Wt CLucCO in the previous study (Yasuno et al., 2018). In the case of the silkworm system, the relative activity of mutant CLuc purified using only anti-FLAG affinity gel was 25% of that of Wt CLucSW. However, further purification revealed that approximately 75% of the FLAG-purified recombinant protein was aggregated and its enzyme activity was lost. In the BY-2 cells, our attempt to express Dmt CLucBY resulted in no recombinant protein production despite the use of 4 independent transgenic callus clones. Taken together, our results suggested that the productivity of CLuc mutated in the N-glycosylation sites differed depending on the expression host. Therefore, we should choose the recombinant protein expression system carefully in order to minimize the effect of glycosylation, especially for cases such as protein crystallization, which requires the removal of glycan. In our case, the Expi293 expression system was the best choice for Dmt CLuc production. To understand the reason for these differences in protein production efficiency, we need to study the protein production pathway in detail and to investigate which point is critical for the efficient production of recombinant protein. Finally, we identified important residues involved in CLuc activity. Crystallographic analysis will be necessary to clarify the functions of these residues in the oxidation process of cypridinid luciferin.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://repository.jpostdb.org/ and JPST001397, PXD029916.

AUTHOR CONTRIBUTIONS

YM, RY, KK, and YO designed the research. RY and KK performed mammalian cell expression, purification, and activity analysis. KC and NM performed Dmt CLuc construction and transformation, and supervised BY-2 transfection. SK performed BY-2 transfection, purification of recombinant protein, and activity assay. AT and HK performed glycan modification analysis. All authors read and approved the final manuscript.

FUNDING

This work was partially supported by JSPS KAKENHI (grant number JP18KK0199) (to YM).

ACKNOWLEDGMENTS

We are grateful to Mami Komatsu (AIST) for her assistance in experiments using BY-2. We thank Hensley M. Nicholai (Cornell University) for his useful comments to the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2022.774786/full#supplementary-material

REFERENCES

Bryan, B. I., and Szent-Gyorgyi, C. (2001). Luciferases, Fluorescent Proteins, Nucleic Acids Encoding the Luciferases and Fluorescent Proteins and the Use Thereof in Membrane Proteins. Curr. Opin. Struct. Biol. 32, 131–138. doi:10.1016/j.sbi.2015.04.005
