Expression of a Soluble Truncated *Vargula*

Luciferase in *Escherichia coli*

Eric A. Hunt ^{a,b†}, Angeliki Moutsopoulos ^{a,b†}, David Broyles ^a, Trajen Head ^a,
Emre Dikici ^a, Sylvia Daunert ^a, & Sapna K. Deo ^{a*}

^aUniversity of Miami, Leonard M. Miller School of Medicine, Department of Biochemistry & Molecular Biology, Miami, 33136, USA.

^bUniversity of Miami, Department of Chemistry, Coral Gables, 33146, USA.

[†]Co-first author

*Correspondence and requests for materials should be addressed to S.K.D.

email: sdeo@med.miami.edu,

Phone: 305-243-4421

Fax: 305-243-3955

Abstract

Marine luciferases are regularly employed as useful reporter molecules across a range of various applications. However, attempts to transition expression from their native eukaryotic environment into a more economical prokaryotic, i.e. bacterial, expression system often presents several challenges. Specifically, bacterial protein expression inherently lacks chaperone proteins to aid in the folding process, while *Escherichia coli* presents a reducing cytoplasmic environment in. These conditions contribute to the inhibition of proper folding of cysteine-rich proteins, leading to incorrect tertiary structure and ultimately inactive and potentially insoluble protein. *Vargula* luciferase (Vluc) is a cysteine-rich marine luciferase that exhibits glow-type bioluminescence through a reaction between its unique native substrate and molecular oxygen. Because most other commonly used bioluminescent proteins exhibit flash-type emission kinetics, this emission characteristic of Vluc is desirable for high-throughput applications where stability of emission is required for the duration of data collection. A truncated form of Vluc that retains considerable bioluminescence activity (55%) compared to the native full-length protein has been reported in the literature. However, expression and purification of this luciferase from bacterial systems has proven difficult. Herein, we demonstrate the expression and purification of a truncated form of Vluc from *E. coli*. This truncated Vluc (tVluc) was subsequently characterized in terms of both its biophysical and bioluminescence properties.

Keywords: *Vargula* luciferase, truncated protein, bioluminescence, bacterial expression

Abbreviations:

Vluc, *Vargula* luciferase; tVluc, truncated *Vargula* luciferase; Cluc, *Cypridina noctiluca* luciferase; LB, Luria Broth; OD, Optical density; IPTG, isopropyl- β -D-thiogalactopyranoside; Ni-NTA, Ni 2-nitrilotriacetic acid; NP-40, nonyl phenoxyethoxyethanol; β -ME, 10 mM 2-mercaptoethanol; SDS-PAGE, sodium dodecylsulfate-polyacrylamide gel electrophoresis; PBS, phosphate buffered saline; MWCO, molecular weight cutoff; CD, Circular Dichroism; IMAC, immobilized-metal affinity chromatography; trxB⁻, thioredoxin reductase; gor⁻, glutathione oxidoreductase; SRCD, synchrotron radiation CD;

Introduction

Numerous marine luciferases, such as those from *Renilla reniformis*, *Gaussia princeps*, and *Cypridina noctiluca* [1-9], are commonly used as bioluminescent reporters across a range of applications such as biosensing, reporter gene assays, and gene expression studies. Another such reporter is *Vargula* luciferase (Vluc), from the ostracod crustacean *Vargula hilgendorfii* also known as the “sea firefly.” In 1989, Thompson et al. successfully cloned the cDNA for Vluc and expressed the full-length protein in a mammalian cell system. The complete primary sequence of Vluc consists of 555 amino acids, with two unique potential sites of glycosylation in its native organism[10, 11]. The native substrate for Vluc, *Vargula* luciferin (vargulin), is also referred to as *Cypridina* luciferin due to the fact that the same substrate is utilized by the *Cypridina noctiluca* luciferase (Cluc) for bioluminescence emission. Although this substrate is similar to the more common luciferase substrate, coelenterazine, minor differences are observed in the substituents located around a conserved imidazopyrazine skeleton. However, the bioluminescent reaction for both substrates proceeds through a common dioxetanone intermediate that emits around 462 nm[12]. An important characteristic of Vluc bioluminescence is the extended, glow-type emission of light[13, 14]. This unique kinetic property makes Vluc a desirable reporter for imaging and various bioluminescent assays, specifically allowing for the time-resolved, multiplexed detection of multiple targets.

Based on the determined sequence of the Vluc cDNA, it was established that the

full-length protein contained two homologous domains, each with notable similarity to the photoprotein aequorin from the jellyfish *Aequorea victoria*[15]. This is a feature shared by many of marine luciferases including *Gaussia*, *Metridia*, and *Cypridina*. In terms of emission intensity, *Vargula* luciferase is significantly brighter (10- to 20-fold) as a gene reporter than the commonly utilized firefly luciferase. The unique substrate utilized by Vluc allows for its use as a multiplex reporter in conjunction with coelenterazine-dependent luciferases[16]. Additionally, the glow-type bioluminescence of Vluc provides a means to develop multiplexed systems based on time resolution of different luciferase signals. In 1996, Maeda et al. demonstrated that a fusion consisting of protein A with the N-terminal homologous domain of Vluc (P28-C312) could be expressed in a mammalian system, and retained ~40% of the full-length wild-type Vluc bioluminescent activity[10]. This fusion containing truncated Vluc (tVluc) enabled the bioluminescence-based detection of the anti-protein A antibody, opening up a range of novel applications for tVluc. However, the production of Vluc in bacterial systems has thus far remained elusive. Inouye and Sahara provide the only example of soluble Vluc production in *E. coli*, achieved through cold shock expression of a fusion consisting of the luciferase and a synthetic IgG-binding domain as a solubilizing partner[17]. However, the resulting fusion showed no bioluminescence activity, likely a result of improper protein folding. Moreover, inclusion of a fused solubility partner can lead to undesirable effects in downstream applications, such as binding interference or reduction of reporter activity. Although such additions can be subsequently removed via proteolytic

cleavage, this process requires additional reagents and the necessity of a secondary purification step, reducing overall protein yield.

Thus, a major consideration for the successful soluble expression of *Vargula* luciferase in a bacterial system involves the proper folding of the protein – specifically the precise formation of cysteine-cysteine bonds. The truncated form of Vluc (tVluc), representing the N-terminal homologous domain of the full-length luciferase, contains only 16 of the 34 cysteine residues found in the wild-type protein. This reduction in possible disulfide bond formation requirements could make production in a bacterial system without a solubilizing partner more feasible, while simultaneously facilitating downstream applications that may require bioconjugation or cellular delivery. In this study, we demonstrate the successful bacterial expression of tVluc in a soluble and active form from *E. coli* and provide a complete characterization of its bioluminescent properties. It is believed that this work will help guide the further development of Vluc and Cluc luciferase variants for more efficient expression and purification from bacterial systems.

Materials and methods

Molecular Cloning

The 555 amino acid sequence for Vluc was obtained from the NCBI GenBank (accession number AAA30332, “luciferase [*Vargula hilgendorffii*]”) and was used

to create the tVluc sequence for expression in *E. coli*. The signal peptide, which was identified using the online prediction tool provided by the Center for Biological Sequencing Analysis at the Technical University of Denmark[18], was predicted to be the first 15 amino acids of the native sequence and was not present in the truncated sequence.

The tVluc sequence used in this study consists of 302 amino acids with a theoretical molecular weight of 33.3 kDa and includes the Pro28–Cys312 segment of the native sequence with an N-terminal 6xHis tag and factor Xa cleavage site. The sequence was synthesized commercially (GeneCopoeia Inc., Rockville, Maryland) following codon optimization for expression in an *E. coli* host and inserted into the pCold-I Cold Shock Expression System vector (Takara Bio. Inc., Japan) using the NdeI (CA^{*}TATG) and XhoI (C^{*}TCGAG) restriction sites of the multiple cloning site (Figure 1). The resulting pCold-I::tVluc (ptVluc) plasmid was then transformed into the cloning strain NEB5- α (New England Biolabs, Ipswich, Massachusetts) for propagation and storage.

Figure 1. The pCold-I Cold Shock Expression System introduces an N-terminal 6xHis tag followed by a factor Xa cleavage site for tag removal following purification. The gene of interest inserted into the multiple cloning site is under the control of the cspA promoter and lac operon for expression control. The vector imparts ampicillin resistance through the amp^r gene encoding β -lactamase.

Expression and Purification from *Escherichia coli*

After propagating the ptVluc via bacterial growth at 37 °C, the plasmid was purified using a QIAprep Spin Miniprep Kit (Qiagen, Valencia, California) and transformed into three different expression strains: NEB Express (New England

Biolabs, Ipswich, Massachusetts), Origami™ 2 (Novagen - EMD Millipore, Billerica, Massachusetts) and SHuffle® Express (New England Biolabs, Ipswich, Massachusetts). As a preliminary step to culture expansion, small cultures were prepared for overnight growth at 37 °C by inoculating 5 mL of Miller LB supplemented with 0.1 mg/mL ampicillin from glycerol stocks of tVluc for each of the three expression strains. Following overnight growth, the small cultures were centrifuged at 4,000×g, resuspended in fresh media supplemented with 0.1 mg/mL ampicillin, and used to inoculate large 250 mL cultures.

Protocol for NEB Express. Large 250 mL cultures were grown to an OD₆₀₀ of 0.6 in Miller LB supplemented with 0.1 mg/mL ampicillin before initiating cold-shock on ice for 1 hour. Cultures were then induced with a final concentration of 1 mM IPTG, followed by growth at 15 °C for 48 hours. The cells were collected by centrifugation at 4,000×g and 4 °C for 20 minutes and resuspended in a lysis buffer of 50 mM sodium phosphate pH 8.0, 300 mM sodium chloride, 10 mM imidazole, and 0.05% (vol.) polyoxyethylene (20) sorbitan monolaurate (Tween-20). The cells were lysed, following which the insoluble material was removed by centrifugation at 17,000×g and 4 °C for 20 minutes, and the supernatant was filtered by syringe through a 0.22 µm filter. The filtered crude protein was then incubated with 500 µL of Ni-NTA agarose (Qiagen, Valencia, California) per culture at 4 °C for 2 hours, collected on a Pierce™ Centrifuge Column (Life Technologies, Grand Island, New York) by gravity flow, and washed with 2 column volumes of a wash buffer of 50 mM sodium phosphate pH 8.0, 300 mM

sodium chloride, and 20 mM imidazole. The protein was then eluted in 500 μ L increments with a stepwise gradient of imidazole in 50 mM sodium phosphate pH 8.0, 300 mM sodium chloride (four column volumes were collected at each stepped concentration).

Protocol for Origami™ 2 Expression. Large 250 mL cultures were grown to an OD₆₀₀ of 1.6 in Terrific broth supplemented with 0.1 mg/mL ampicillin prior to induction of cold-shock on ice for 1 hour. The media was then replenished, and overnight induction at 15 °C was initiated using a final concentration of 0.1 mM IPTG. The cells were collected by centrifugation at 4000 $\times$ g and 4 °C for 20 minutes and resuspended in a lysis buffer of 50 mM Tris HCl pH 8.0, 150 mM sodium chloride, 10 mM imidazole, 1% (vol.) nonyl phenoxy polyethoxy ethanol (NP-40), 0.2% (vol.) Tween-20, and 10 mM 2-mercaptoethanol (β -ME). The cell suspension was supplemented with 1X ProBlock™ Gold Bacterial Protease Inhibitor Cocktail (Gold Biotechnology Inc., St. Louis, Missouri) to prevent protein degradation following cell lysis. The insoluble material was removed by centrifugation at 10,000 $\times$ g and 4 °C for 30 minutes, and the supernatant was filtered by syringe through a 0.22 μ m filter.

The filtered crude protein was then incubated with 250 μ L of Ni-NTA agarose per 250 mL culture at 4 °C for 1–2 hours, collected on a centrifuge column by gravity flow, washed with 10 column volumes of lysis buffer followed by 20 column volumes of wash buffer containing 50 mM Tris HCl pH 8.0, 150 mM sodium chloride, and 20 mM imidazole. The protein was then eluted with elution buffer of

50 mM Tris-HCl pH 8.0, 150 mM sodium chloride, and 150 mM imidazole in 1-column volume increments.

Protocol for SHuffle[®] Express Expression. SHuffle[®] Express proceeded similarly to the protocol for Origami[™] 2, except that the large cultures were grown to an OD₆₀₀ of 1.2 and were induced with a final concentration of 0.5 mM IPTG. The protein was bound to the Ni-NTA agarose for 30–45 minutes and subsequently purified in the same manner.

All purifications were analyzed by SDS-PAGE using 4–20% gradient Mini-PROTEAN[®] TGX[™] Precast Gels (Bio-Rad Laboratories Inc., Hercules, California) under denaturing conditions with running buffers containing sodium dodecyl sulfate (SDS).

The purified tVluc was then dialyzed into either phosphate buffered saline (PBS) consisting of 10 mM sodium phosphate pH 7.2 and 150 mM sodium chloride or a CD buffer consisting of 5 mM potassium phosphate and 25 mM ammonium sulfate pH 7.8, depending on the intended application following purification. Slide-A-Lyzer[™] Dialysis Cassettes (Life Technologies, Grand Island, New York) with a molecular weight cutoff (MWCO) of 10 kDa were used.

Characterization of Bioluminescence Spectra

Bioluminescence measurements were carried out in a 96-well black polystyrene non-binding microplate (Greiner Bio-One Inc., Monroe, North Carolina). Cypridina luciferin (vargulin) was purchased from NanoLight[™] Technology (Prolume Ltd.,

Pinetop, Arizona).

Briefly, 100 μL of 600 μM vargulin was injected into 100 μL of purified luciferase at a concentration of ~ 3.5 μM , and the spectrum was recorded. Using a CLARIOstar[®] microplate reader (BMG LABTECH GmbH, Ortenberg, Germany), the spectrum was scanned from 380-650 nm using 1 s/nm integration. Five separate spectra were recorded and averaged using GraphPad Prism 6 to produce the final spectrum.

Kinetic studies

Decay kinetics were obtained using a POLARstar[®] Optima (BMG LABTECH GmbH, Ortenberg, Germany). The emission filter was set to “lens” (i.e. no filter) and total light was collected for all measurements. All kinetic measurements were carried out in 96-well black polystyrene non-binding microplates (Greiner Bio-One Inc., Monroe, North Carolina). Cypridina luciferin (vargulin) was purchased from Nanolight[™] Technology (Prolume Ltd., Pinetop, Arizona). Briefly, 100 μL of luciferase was added to a 96-well microplate at a concentration of 1.4 μM . The bioluminescent emission was quantified by integrating the signal in 10-second intervals for 20 minutes. At 30 seconds, 100 μL of 100 μM vargulin was injected to begin the reaction.

Circular Dichroism Spectroscopy

Circular dichroism (CD) spectroscopy was performed on a Jasco J-815 Circular Dichroism Spectropolarimeter using a 0.1 cm path length quartz cuvette. tVluc

was dialyzed into CD buffer and diluted to a concentration of 0.05 mg/mL. The spectrum was acquired from 260-185 nm using the parameter set described in Table 1.

Parameter	Value
Scan Mode	<i>Continuous</i>
Scan Speed	<i>50 nm/min</i>
Data Pitch	<i>0.5 nm</i>
Bandwidth	<i>2 nm</i>
D.I.T	<i>2 s</i>
Accumulations	<i>5</i>

Table 1: Parameter used for the circular dichroism spectroscopy

Computational Analysis of Structure

All analysis of the CD spectrum was performed using DichroWeb (Department of Crystallography, Institute of Structural and Molecular Biology, Birkbeck College, University of London, United Kingdom)[19]. The CDSSTR analysis program[20] was used to generate secondary structure assignments for tVluc using reference sets 3, 4, 6, 7, SP175, and SMP180. The CDSSTR analysis program was chosen as it provided the best fit to experimental data; however, all analysis programs available through DichroWeb including SELCON3, CONTIN, VARSLC, and K2D were evaluated. Data input for the analysis was limited to the wavelength range of 240-185 nm.

Vluc Expression from mammalian cells: HEK293T cells were transfected with the plasmid, pCMV-VargLuc vector (Targeting Systems, El Cajo, CA) using the Xfect™ (Takara Bio Inc., Mountain View, CA) transfection reagent protocol. The cells were cultured using Dulbecco's Modified Eagle Medium (DMEM) (ThermoFischer Scientific, Waltham, MA) containing 10% fetal bovine serum (FBS) (Sigma, St. Louis, MO), antibiotic and antimycotic in a humidified 5% CO₂ incubator at 37 °C. After 2 days of incubation, the culture supernatant was collected for the bioluminescence characterization of the expressed protein. The protein was analyzed by SDS-PAGE using 4–20% gradient Mini- PROTEAN® TGX™ Precast Gels (Bio-Rad Laboratories Inc., Hercules, California) under denaturing conditions with running buffers containing sodium dodecyl sulfate (SDS).

The pure, full-length Vluc was then dialyzed into phosphate buffered saline (PBS) consisting of 10 mM sodium phosphate pH 7.2 and 150 mM sodium chloride. Slide-A-Lyzer™ Dialysis Cassettes (Life Technologies, Grand Island, New York) with a molecular weight cutoff (MWCO) of 10 kDa was used.

Results

Molecular Cloning of Truncated *Vargula* Luciferase

The pColdI::tVluc vector (ptVluc, Figure 1b) was created by inserting a codon-optimized gene for tVluc into the pCold-I cold shock expression vector (Figure 1a).

Expression and Purification of Truncated *Vargula* Luciferase from *E. coli*

We evaluated three expression hosts of *E. coli*; NEB Express, Origami™ 2 and SHuffle® Express. Utilizing the NEB Express system, protein bands corresponding to tVluc were observed during SDS-PAGE analysis, but were extremely faint with protein elution beginning at 40 mM imidazole and completing before the 60 mM imidazole wash aliquot (data not shown). This result indicated that the N-terminal 6xHis tag may have been sequestered internally, while the observed lack of bioluminescent activity in the presence of vargulin suggested overall protein misfolding.

In contrast, the tVluc product obtained from Origami™ 2 expression showed a band observed in SDS-PAGE analysis (Figure 2) corresponding to overexpressed and successfully isolated tVluc. Although bioluminescent activity of the protein-containing fraction was high (as verified from bioluminescence emission upon substrate addition), these elution fractions contained numerous contamination bands indicating overall poor purity of the recovered tVluc.

Figure 2. Molecular weight analysis of tVluc expressed in Origami™ 2 *E. coli*. Analysis of band migration was done using Adobe Photoshop CS6 and GraphPad Prism 6. The main over-expressed band was calculated to be 34 kDa. The theoretical molecular weight of tVluc is 33.3 kDa. The 66 kDa band may be evidence of dimerization. Lanes 1-2) flowthrough; 3) lysis buffer wash; 4) Precision Plus Protein™ Dual Color Standard; 4) wash buffer wash; 6-9) elution fractions.

Subsequently, a third strain of *E. coli*, SHuffle® Express was attempted for expression of tVluc. Expression from this strain produced active tVluc in an exceptionally pure form from the soluble fraction (Figure 3), yielding

approximately 1.5 mg of pure protein per 1 L culture. Because of its purity and yield, subsequent characterization of the bacterially-expressed tVluc was performed on protein obtained using this expression system.

Figure 3. Purification of tVluc from SHuffle® Express *E. coli*. Lanes 1) Precision Plus Protein™ Dual Color Standard; 2) flowthrough; 3) lysis buffer wash; 4) wash buffer wash; 5–7) elution fractions. The purified band was calculated to have a molecular weight of 35 kDa, while the theoretical molecular weight of tVluc is 33.3 kDa.

The SDS-PAGE analysis results from expression in both Origami™ 2 and SHuffle were further utilized to determine the molecular weight of the purified tVluc. From the expression performed in Origami™ 2 cells, four prominent bands were chosen from the elution fractions of the tVluc purification. The migration of each band in the Precision Plus Protein™ Dual Color Standard (Bio-Rad Laboratories Inc., Hercules, California) was measured using Photoshop CS6 and plotted with automated curve fitting in Prism 6 to extrapolate a precise molecular weight value for the over-expressed bands (Figure 2). The most prominent over-expressed band was calculated to be 34 kDa, approximately the theoretical molecular weight of tVluc. Additionally, a 66 kDa band was observed, potentially representing the dimerization between individual tVluc monomers. A similar calculation was performed using SDS-PAGE analysis of purified tVluc from expression in SHuffle® Express cells (Figure 3). As before, the migration of the bands of the protein standard were measured, and a molecular weight of approximately 35 kDa for purified tVluc was extrapolated from the resulting curve fit, corroborating the results demonstrated in the Origami™ 2 expression and

purification.

As further confirmation of the successful purification of tVluc from Origami™ 2, Western blot analyses were performed using an anti-6xHis epitope tag antibody (Life Technologies, Grand Island, New York) and an anti-Cypridina luciferase epitope antibody (antibodies-online Inc., Atlanta, Georgia). Because both *Vargula* and *Cypridina* luciferases share a high degree of homology (85 % identity and 94 % similarity in two sequences), we speculated that the polyclonal anti-cypridina antibody would be capable of binding tVluc. The anti-6xHis epitope tag antibody blot confirmed that largest over-expression band that was observed at ~34 kDa (Figure 3) represented tVluc produced from the pCold-I cold shock expression system (Figure 4). Although the anti-Cypridina blot did not generate any readable signal, the purified tVluc includes only the N-terminal domain of the full-length Vluc protein. Thus the epitope targeted by this antibody may have potentially been removed with truncation.

Figure 4. Western blots of tVluc expressed in Origami™ 2 *E. coli* using anti-6xHis tag (left) and anti-Cluc (right) primary antibodies. No binding was observed for the anti-Cluc antibody, potentially due to epitope removal following truncation.

Expression and characterization of *Vargula* Luciferase from HEK293T mammalian cells

Full-length Vluc was expressed using mammalian HEK293T cells. We observed only a single band at the expected size for full-length Vluc after loading the cell supernatant to the SDS-PAGE gel (Figure 5a). It wasn't required to do any

purification since the protein was excreted selectively into the culture media of the mammalian HEK293T cells. Although the yield of the protein was low, the observed protein purity indicated a lack of need for further purification. This was sufficient for use as a control in our study. However, we would like to point out this method may not be appropriate for practical applications. We measured the bioluminescence activity of the full-length Vluc, and compared these values with the purified tVluc from SHuffle® Express *E. coli* (Figure 5b). The bioluminescence activity of the purified tVluc was 45% of the full-length Vluc.

Figure 5: a) SDS-PAGE gel of expression of Vluc from HEK293T mammalian cells. Lanes 1: Precision Plus Protein™ Dual Color Standard and Lane 2 the expressed Vluc. The band was calculated to have approximately a molecular weight 66 kDa. b) The bioluminescence activity of the matched concentration of full-length Vluc from HEK293T compared with the purified tVluc expressed from SHuffle® Express *E. coli*.

Secondary Structure Analysis of Truncated *Vargula* Luciferase by Circular Dichroism Spectroscopy

After isolating purified tVluc expressed from SHuffle® Express, a CD spectrum was obtained from 260–180 nm (Figure 6a) and analyzed by the CDSSTR analysis package to generate secondary structure assignments for the luciferase (Figure 6b). The online secondary structure prediction server Phyre2 [21] was able to model ~65% of residues with a >90% confidence level (data not shown). While these types of *ab initio* structure analyses can vary widely in their accuracy, it was reassuring to note that the predicted model exhibited ~8% α_r character and ~20% β_r character. These values correspond well with the

CDSSTR secondary structure assignments calculated from the obtained CD data (Figure 6b).

Figure 6. (a) CD spectrum for tVluc and (b) secondary structure assignments made using the CDSSTR analysis program from DichroWeb.

Bioluminescent Characteristics of Truncated *Vargula* Luciferase

Characterization of the bioluminescence properties of tVluc began with the acquisition of a complete bioluminescence emission spectrum (Figure 7). Immediately following injection of vargulin into a sample of purified tVluc, emission intensity was measured from 380-650 nm. The emission profile of the purified tVluc exhibited a characteristic broad emission, with a maximum bioluminescence emission intensity occurring at approximately 510 nm.

Figure 7. Bioluminescence emission spectra of tVluc expressed using SHuffle[®] Express *E. coli*.

Examining the kinetic parameters of the tVluc bioluminescence, we determined that the purified protein exhibited a glow-type bioluminescent emission with a half-life of over two hours (Figure 8b). The bioluminescence reaction occurs rapidly upon injection of the vargulin substrate and is well suited for high-throughput analytical applications where extended signal lifetime is necessary to compensate for delays in signal acquisition following simultaneous substrate injection.

Figure 8. (a) The extended emission kinetics of tVluc with vargulin substrate. (b) The half-life was shown to extend beyond two hours.

Discussion

Within, we have presented a method for the soluble expression and successful purification of a truncated variant of *Vargula* luciferase (tVluc) from a prokaryotic (bacterial) system. The unique composition of the Vluc molecule, specifically the presence of numerous disulfide bonds within its tertiary structure, required that some specific considerations be made for the successful isolation of the tVluc reporter. These considerations included the choice of expression plasmid, expression temperature, fusion partners, etc. For example, the cold shock expression system, pColdI, was chosen for production of tVluc in *E. coli* due to the fact that growth at reduced temperature will slow the rate of protein production, thereby encouraging proper folding and reducing aggregation[22]. By using this expression vector with three different strains of *E. coli*, we examined various routes for maximum efficiency of tVluc production.

From SDS-PAGE analysis of NEB Express system as well as from the lack of the bioluminescent activity we concluded that N-terminal 6xHis tag might have been sequestered internally. On the other hand, the OrigamiTM 2 and SHuffle[®] Express gave promising results based on the resulting high bioluminescence activity.

The tVluc product obtained from Origami™ 2 was dramatically more tolerant of immobilized-metal affinity chromatography (IMAC), and stayed bound to the Ni-NTA agarose until eluted with a high concentration of imidazole giving the overexpression band (Figure 2). This improvement was attributed to the thioredoxin reductase (*trxB*) and glutathione oxidoreductase (*gor*) mutations present in Origami™ 2, which likely contributed to proper disulfide formation for the expression of active tVluc. Although the bioluminescent activity was high, the elution fractions contained numerous contamination bands, demonstrating that background protein expression was still a significant impediment to obtaining usable protein in this *E. coli* strain. Additionally, Origami™ 2 is derived from *E. coli* K12, and it is possible that it exhibits higher protease activity than other strains.

As such, a third strain of *E. coli* was attempted for expression of tVluc. Unlike the previously used cell lines, SHuffle® Express is an *E. coli* B strain which, in addition to being a *trxB*/*gor* mutant, also constitutively expresses a chromosomal copy of the disulfide bond isomerase DsbC. This enzyme is usually found in the periplasm of Gram-negative bacteria and functions to aid in the rearrangement of incorrectly formed disulfides[23]. Expression from this strain produced active tVluc in an exceptionally pure form from the soluble fraction (Figure 3), yielding approximately 1.5 mg of pure protein per 1 L culture.

The similarity in predicted and observed protein structure composition lend validity to the assessment that tVluc is folding in a relatively consistent and

predictable manner within the bacterial cytoplasmic environment. Although the remaining portion of the secondary structure was classified as disordered, this can result from solubility issues in low ionic strength buffer, and the truncation itself may lead to a more fluid tertiary structure in the absence of substrate. Further investigation using synchrotron radiation CD (SRCD) would be warranted only if tertiary structure variability was evident analytically.

In addition to the structural data obtained for this bacterial-expressed tVluc variant, the bioluminescence properties of the protein were also characterized. The bioluminescence activity of the newly isolated and purified tVluc has a 45% loss in activity compared to the full-length Vluc (Figure 5b). This observation is congruent with findings in the literature which reported similar bioluminescence emission differences between the full-length and truncated Vluc variants[10]. The analysis of the bioluminescence emission spectrum revealed a broad peak with an emission maximum occurring at 510 nm (Figure 8a). In comparison to the previously reported value for wild-type Vluc[24], this represents an approximately 50 nm red-shift. As demonstrated in Figure 8b, tVluc exhibited a stable, glow-type emission following the addition of substrate, similar to the kinetics observed in full-length Vluc[11]. From this measurement, it was calculated that the isolated protein possessed an emission half-life of approximately two hours in buffer. Moreover, the ability to express and isolate tVluc in a bacterial system provides many advantageous benefits to its future incorporation into a range of bioluminescence-based reporter systems.

Conclusion

This work represents the first report of a successful expression and purification of a functional *Vargula* luciferase variant in a prokaryotic (bacterial) system. Critically, we have demonstrated that great care must be taken in the design and implementation of tVluc expression to ensure that the cytoplasmic environment and expression conditions are amenable to proper disulfide formation during folding. This bacterially-expressed tVluc was isolated as both soluble and pure, and exhibited a glow-type bioluminescent emission with a half-life of over two hours. This truncated form of the luciferase is slightly smaller than Renilla luciferase (39.5 kDa), providing potential for its use in biosensor-based applications where size is a critical factor[25]. This expression of tVluc in *E. coli* will be a guide for further production of *Vargula* and *Cypridina* luciferase variants, which will have unprecedented value in the design and implementation of multiplexed reporter systems.

Acknowledgements

SKD would like to thank the National Institute of General Medical Science funding (R01GM047915), the State of Florida Department of Health and the National Science Foundation (CHE-1506740). S.D. thanks the Miller School of Medicine of the University of Miami for the Lucille P. Markey Chair in Biochemistry and Molecular Biology.

References

- [1] C. Wu, K. Kawasaki, S. Ohgiya, Y. Ohmiya, Syntheses and evaluation of the bioluminescent activity of (S)-Cypridina luciferin and its analogs. *Tetrahedron Lett* 47 (2006) 753-756.
- [2] E.A. Hunt, S.K. Deo, Bioluminescent stem-loop probes for highly sensitive nucleic acid detection. *Chem Commun* 47 (2011) 9393-9395.
- [3] E.A. Hunt, A. Moutsiopoulou, S. Ioannou, K. Ahern, K. Woodward, E. Dikici, S. Daunert, S.K. Deo, Truncated Variants of Gaussia Luciferase with Tyrosine Linker for Site-Specific Bioconjugate Applications. *Sci Rep* 6 (2016) 26814.
- [4] M. Kumar, D.H. Zhang, D. Broyles, S.K. Deo, A rapid, sensitive, and selective bioluminescence resonance energy transfer (BRET)-based nucleic acid sensing system. *Biosens Bioelectron* 30 (2011) 133-139.
- [5] K.A. Cissell, Y. Rahimi, S. Shrestha, S.K. Deo, Reassembly of a Bioluminescent Protein Renilla Luciferase Directed through DNA Hybridization. *Bioconjugate Chem* 20 (2009) 15-19.
- [6] K.A. Cissell, Y. Rahimi, S. Shrestha, E.A. Hunt, S.K. Deo, Bioluminescence-based detection of MicroRNA, miR21 in breast cancer cells. *Analytical Chemistry* 80 (2008) 2319-2325.
- [7] M. Verhaegen, T.K. Christopoulos, Recombinant Gaussia luciferase. Overexpression, purification, and analytical application of a bioluminescent reporter for DNA hybridization. *Analytical Chemistry* 74 (2002) 4378-4385.
- [8] M. Kumar, L. Kovalski, D. Broyles, E.A. Hunt, P. Daftarian, E. Dikici, S. Daunert, S.K. Deo, Design and development of high bioluminescent resonance energy transfer efficiency hybrid-imaging constructs. *Anal Biochem* 498 (2016) 1-7.
- [9] D. Scott, E. Dikici, M. Ensor, S. Daunert, Bioluminescence and its impact on bioanalysis. *Annu Rev Anal Chem (Palo Alto Calif)* 4 (2011) 297-319.
- [10] Y. Maeda, H. Ueda, J. Kazami, G. Kawano, E. Suzuki, T. Nagamune, Truncation of Vargula luciferase still results in retention of luminescence. *J Biochem* 119 (1996) 601-603.
- [11] E.M. Thompson, S. Nagata, F.I. Tsuji, Cloning and Expression of Cdna for the Luciferase from the Marine Ostracod Vargula-Hilgendorffii. *P Natl Acad Sci USA* 86 (1989) 6567-6571.
- [12] D.H.P.D. Megan Dobbs B.S., Janaki Narahari Ph.D., Jae Choi Ph.D., Georgyi Los Ph.D., Brian Webb Ph.D. (2013), Megan Dobbs B.S., Douglas Hughes Ph.D., Janaki Narahari Ph.D., Jae Choi Ph.D., Georgyi Los Ph.D., Brian Webb Ph.D. (2013) Simultaneous dual-emission detection of luciferase reporter assays, <http://www.bmglabtech.com/en/applications/>: BMG LABTECH, in: A.N. 233 (Ed.).
- [13] L.J. Kricka, Chemiluminescent and Bioluminescent Techniques. *Clinical Chemistry* 37 (1991) 1472-1481.
- [14] A new secreted luciferase optimized to advance cell based reporter technology, <http://www.invivogen.com/review-lucia>.

- [15] E.M. Thompson, S. Nagata, F.I. Tsuji, Cloning and expression of cDNA for the luciferase from the marine ostracod *Vargula hilgendorfii*. *Proc Natl Acad Sci U S A* 86 (1989) 6567-6571.
- [16] B.A. Tannous, D.E. Kim, J.L. Fernandez, R. Weissleder, X.O. Breakefield, Codon-optimized Gaussia luciferase cDNA for mammalian gene expression in culture and in vivo. *Mol Ther* 11 (2005) 435-443.
- [17] S. Inouye, Y. Sahara, Soluble protein expression in *E. coli* cells using IgG-binding domain of protein A as a solubilizing partner in the cold induced system. *Biochem Biophys Res Commun* 376 (2008) 448-453.
- [18] O. Emanuelsson, S. Brunak, G. von Heijne, H. Nielsen, Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2 (2007) 953-971.
- [19] L. Whitmore, B.A. Wallace, Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* 89 (2008) 392-400.
- [20] N. Sreerama, S.Y. Venyaminov, R.W. Woody, Estimation of protein secondary structure from circular dichroism spectra: Inclusion of denatured proteins with native proteins in the analysis. *Analytical Biochemistry* 287 (2000) 243-251.
- [21] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J.E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* 10 (2015) 845-858.
- [22] G.L. Rosano, E.A. Ceccarelli, Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol* 5 (2014) 172.
- [23] P.W. Haebel, D. Goldstone, F. Katzen, J. Beckwith, P. Metcalf, The disulfide bond isomerase DsbC is activated by an immunoglobulin-fold thiol oxidoreductase: crystal structure of the DsbC-DsbD alpha complex. *Embo J* 21 (2002) 4774-4784.
- [24] E.M. Thompson, S. Nagata, F.I. Tsuji, *Vargula-Hilgendorfii* Luciferase - a Secreted Reporter Enzyme for Monitoring Gene-Expression in Mammalian-Cells. *Gene* 96 (1990) 257-262.
- [25] E.A. Hunt, S.K. Deo, Bioluminescent stem-loop probes for highly sensitive nucleic acid detection. *Chem Commun (Camb)* 47 (2011) 9393-9395.

Figure

(a) pCold-I

(b) ptVluc

Figure

Figure

Figure

Figure

a)

b)

Figure

Figure

Figure