

**EXPRESSION, PURIFICATION AND BIOPHYSICAL
CHARACTERIZATION OF ATP INDEPENDENT LUCIFERASE FROM
*GAUSSIA PRINCEPS***

Dissertation

Zur Erlangung des akademischen Grades Doktor der
Naturwissenschaften (Dr. rer. nat.) der Mathematisch-Naturwissenschaftlichen
Fakultät der Christian- Albrechts- Universität zu Kiel

vorgelegt von

Leslye del Carmen Roca Burgos

Kiel 2018

Referent: Prof. Dr. Axel Scheidig
Koreferent: Prof. Dr. Joachim Grötzinger

Tag der Verteidigung: 19.07.2018

Gez. Prof. Dr. Frank Kempken, Decan

A mis padres Irene y Héctor

Erklärung

Hiermit erkläre ich, dass ich bis zum heutigen Tage weder an der Christian-Albrechts-Universität zu Kiel noch an einer anderen Hochschule ein Promotionsverfahren endgültig nicht bestanden habe oder mich in einem entsprechenden Verfahren befinde. Ich erkläre, dass die Abhandlung nach Inhalt und Form die eigene Arbeit ist und dass ich die Inanspruchnahme fremder Hilfen aufgeführt habe sowie, dass ich die wörtlich oder inhaltlich aus anderen Quellen entnommenen Stellen als solche gekennzeichnet habe. Ich erkläre, dass diese Arbeit weder ganz noch zum Teil schon einer anderen Stelle im Rahmen eines Prüfungsverfahrens vorgelegen hat. Ich erkläre, dass die Arbeit unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden ist.

Kiel, den 31.05.2018

Leslye del Carmen Roca Burgos

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Prof. Axel Scheidig, for his guidance and all the motivation throughout my PhD. I would like to thank him for his continuous support especially in the final stage of my thesis.

My special thanks to Dr. Christoph Plieth, not only for guiding me in many performances of bioluminescence experiments in the laboratory, but also for giving me his comments and valuable suggestions which have improved my thesis.

I would like to thank Frederic Garzoni for his kind support and training with Baculovirus system in insect cells at the EMBL-Laboratory in Grenoble, France.

I am grateful to Dr. Annette Faust for her support, encouragement and friendship. A very special gratitude to my colleague and friend Chrissy, for her friendship and comments to my thesis. I am also grateful to Dr. Marta Kaleta for proofreading my thesis.

I would like to thank to Chrissy, Tina, Brigitte, Sonia, Susi, Sebastian, Christian, Renato, Felix, with whom I spent many hours in the lab having productive discussions and enjoying a friendly environment.

A special thanks to my family for supporting me throughout my PhD. Words cannot express how grateful I am to my parents. My gratitude to Alvaro, my husband, for all his support and encouragement day by day. Finally, I would like to thank my sons, Ignacio and Antonio, who provide me endless inspiration and motivation.

ZUSAMMENFASSUNG

CTZ-abhängige Luciferasen aus verschiedenen Phyla (z. B. von *Aequorea victoria* und *Gaussia princeps*) zeigen keine signifikante strukturelle Homologie. Daher ist die Bestimmung der dreidimensionalen Struktur des GLuc-Proteins erforderlich, um den zugrundeliegenden Reaktionsmechanismus weiter zu erforschen. Kenntnisse über den Aufbau des aktiven Zentrums könnten helfen, Reportervarianten mit höherer Effizienz oder mit veränderten spektralen Eigenschaften zu entwerfen, was die Anwendungsmöglichkeiten in der Zellbiologie erweitern wird.

In dieser Arbeit wurde GLuc kloniert, heterolog in *E. coli*-Zellen exprimiert und ein Protokoll für die Reinigung des Volllängenkonstrukts bis zur Homogenität, gemessen mittels SDS-PAGE, RP-HPLC und DLS, entwickelt. Darüber hinaus wurde die Charakterisierung des löslichen Proteins unter Verwendung der folgenden biophysikalischen Methoden durchgeführt: dynamische Lichtstreuung, Hochleistungsflüssigchromatographie, Magnetische Kernresonanz-Spektroskopie und Biolayer-Interferometrie. Die Expression und Reinigung nach dem optimierten Protokoll erzeugt ein lösliches rekombinantes Luciferase-Enzym mit hoher Biolumineszenz-Aktivität und Stabilität. Zusätzlich erzeugen die initialen Kristallisationsbedingungen kleine Nadelcluster und einige Sphärolithe.

Diese Arbeit bestätigt, dass die *Gaussia*-Luciferase aus zwei sich wiederholenden strukturellen Domänen besteht, die von der cDNA-Sequenz abgeleitet sind. Beide Domänen weisen nach Zugabe von CTZ Biolumineszenz-Aktivität auf. Die katalytischen Eigenschaften beider Domänen wurden näher charakterisiert. Zu diesem Zweck wurde die Volllängen-GLuc mit einer Thrombin-Spaltungsstelle zwischen beiden Domänen sowie die zwei individuellen katalytischen Domänen kloniert und in *E. coli*-Zellen exprimiert. Die Proteine wurden durch Affinitätschromatographie, Ionenaustauschchromatographie und die Ergebnisse mittels SDS-PAGE und HPLC analysiert. Ich stellte fest, dass beide Domänen identische Pufferbedingungen für eine maximale Lichtausbeute benötigen. Die katalytischen Eigenschaften einschließlich optimaler Pufferbedingungen (pH, Temperatur, Salz- und Detergenskonzentration) sind mit der intakten Volllängen-*Gaussia*-Luciferase identisch. Außerdem erzeugte ich neben dem Volllängen-GLuc-Konstrukt zur Expression in *E. coli*-Zellen drei Konstrukte, die in Sf21-Insektenzellen unter Verwendung des Baculovirus-Expressionsvektorsystems (BEVS) zur Proteinkristallisation in lebenden Zellen exprimiert werden können. Hier zeigen lichtmikroskopische Bilder die Bildung von GLuc-Mikrokristallen *in vivo*. Diese Ergebnisse stellen den Ausgangspunkt für eine detaillierte und systematische Untersuchung der zellulären Kristallisationsmechanismen in lebenden Insektenzellen dar. Eine weitere Optimierung der Methoden und eine Verbesserung des Kristallextraktionsverfahrens ist erforderlich, um Röntgen-kristallographische Analysen zu ermöglichen und vollständige strukturelle Informationen über die *Gaussia*-Luciferase und ihre getrennt exprimierten und kristallisierten aktiven Domänen zu erhalten.

ABSTRACT

CTZ-dependent luciferases from different phyla (e.g. from *Aequorea victoria* and *Gaussia princeps*) show no significant structural homology. Therefore, determination of the three dimensional structure of the GLuc-protein is needed to further explore the underlying reaction mechanism. Knowledge about the architecture of the active site may help to design reporter variants with higher efficiency or with modified spectral characteristics and this will complement the cell biology tool box.

In this work, GLuc has been cloned, heterologously expressed in *E. coli* cells, and a protocol was established to purify the full-length construct to homogeneity, as judged by SDS-PAGE. Furthermore, characterization of the soluble protein was performed employing the following biophysical methods: dynamic light scattering, high pressure liquid chromatography, nuclear magnetic resonance and biolayer interferometry. The expression and purification from the optimised protocol produce a soluble recombinant luciferase enzyme with high bioluminescence activity and stability. The tested initial crystallization conditions produced small needle clusters and some spherulites.

This thesis confirms that luciferase consists of two structural domain repeats as deduced from the cDNA sequence. Both domains, when expressed separately, have bioluminescence activity with CTZ. The catalytic properties of both domains were characterized in more detail. For this purpose, the full-length GLuc with a thrombin cleavage site between both domains, and the two individual catalytic domains were cloned and expressed in *E. coli*. The proteins were purified by gel filtration and the results were analysed by SDS-PAGE and HPLC. Both domains were found to require identical buffer conditions for maximal light yield. The catalytic properties including optimal buffer condition (pH, temperature, salt and detergent concentration) are identical to the intact full-length *Gaussia* luciferase. Besides the construction of the individual full-length GLuc construct for expression in *E. coli*, three constructs were generated for expression in *Sf21* insect cells using the baculovirus expression vector system (BEVS) for protein crystallization in living cells. The results show the formation of GLuc microcrystals *in vivo*, which were visualized with a light microscope. These results may represent a starting point for a detailed and systematic investigation of the cellular crystallization mechanisms in living insect cells. Further optimization of the methods and improvements of crystallisation procedures are required to allow for advanced X-ray-based analyses in order to obtain full structural information about the *Gaussia* luciferase and its separately expressed and crystallized active domains.

Abbreviations

AA	Amino acid(s)
AC	Affinity chromatography
ACN	Acetonitrile
ADP	Adenosine diphosphate
Amp	Ampicillin
AMP	Adenosine monophosphate
ATP	Adenosine 5'-triphosphate
BEVS	Baculovirus Expression Vector System
BLI	Bio-Layer Interferometry
BME	Beta-mercaptoethanol
Bp	Base pair
BRET	Bioluminescence resonance energy transfer
BSA	Bovine serum albumin
C-His	Carboxy terminus-histidine
cDNA	Complementary DNA
CTZ	Coelenterazine
CV	Column volume
D1	First domain
D2	Second domain
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
FLuc	Firefly Luciferase
FMNH ₂	Reduced riboflavin phosphate
FPLC	Fast protein liquid chromatography
Fw	Forward
Gen	Gentamicyn
GLuc	<i>Gaussia</i> luciferase
His	Histidine
HPLC	High Performance Liquid Chromatography
HTX	High-Throughput-Crystallisation

IEC	Ion Exchange Chromatography
IPTG	Isopropyl- β -D-thiogalactopyranoside
Kan	Kanamycin
kDa	Kilodalton
LB-Medium	Lysogeny broth-Medium
MPD	2-methyl-2,4 pentanediol
MW	Molecular weight
MWCO	Molecular weight cutoff
N-His	Amino terminus-histidine
N ₂	Liquid nitrogen
NMR	Nuclear magnetic resonance
NTA	Nitrilotriacetic acid
OD	Optical Density (OD ₆₀₀ at the wavelength $\lambda = 600$ nm)
PAA	Polyacrylamide
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pI	Isoelectric point
PMSF	Phenylmethylsulfonyl fluoride
RLU	Relative light units
Rmluc	<i>Renilla muelleri</i> luciferase
RP	Reverse phase
Rv	Reverse
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SFX	Serial femtosecond crystallography
SSX	Serial synchrotron crystallography
TBE	Tris-borate-EDTA
Tet	Tetracycline
TFA	Trifluoroacetic acid
Tris	Tris (hydroxymethyl) aminomethan
UV/Vis	Ultraviolet/Visible

TABLE OF CONTENTS

Erklärung	v
Acknowledgements	vi
Zusammenfassung	vii
Abstract	viii
Abbreviations	ix
Table of Contents	xi
List of Figures	xv
List of Tables	xviii
I. Introduction.....	1
1. Bioluminescence	1
1.1. Marine Bioluminescence	1
1.2. Chemiluminescence, bioluminescence, fluorescence and phosphorescence	1
1.3. Luciferin-luciferase system.....	2
1.4. The five chemical classes of luciferin.....	2
2. <i>Gaussia princeps</i>	3
2.1. <i>Gaussia</i> luciferase.....	4
2.2. Application of <i>Gaussia</i> luciferase.....	7
3. Motivation, Aim and Outline of the work	8
II. Material and methods.....	9
1. Material	9
1.1. Chemicals, consumables and kits	9
1.2. Vectors and expression host strains	9
1.3. Oligonucleotides	10
1.4. Bacterial culture	10
1.5. Machines, instruments, columns and resins.....	11
1.6. Crystallization solutions.....	11
1.7. Computer programs and online databases	11

2. Methods.....	12
2.1. Protein sequence and bioinformatics prediction	12
2.2. Propagation of DNA in <i>E. coli</i>	12
2.2.1. Heat shock transformation and electroporation.....	12
2.2.2. Amplification and plasmid isolation from <i>E. coli</i>	13
2.2.3. DNA quantification	13
2.3. Agarose gel electrophoresis	13
2.4. DNA extraction.....	14
2.5. Restriction digestion and dephosphorylation.....	14
2.6. Ligation reaction	14
2.7. Polymerase chain reaction (PCR)	15
2.7.1. Colony PCR reaction.....	15
2.8. Heterologous protein expression systems	16
2.8.1. Expression in bacteria.....	16
2.8.2. Expression in insect cells.....	17
2.9. Protein purification methods.....	18
2.9.1. Cell disruption	18
2.9.2. Affinity chromatography	19
2.9.3. Affinity chromatography under Strep-tag II matrix	19
2.9.4. Size exclusion chromatography	20
2.9.5. Ion exchange chromatography	20
2.9.6. Purification protocol of full-length GLuc and its catalytic domains	22
2.9.7. Purification protocol of full-length GLuc construct.....	23
2.10. Protein analysis of purity and protein concentration	23
2.10.1. Estimation of protein weight and purity by SDS-PAGE	23
2.10.2. Determination of protein concentration	24
2.11. Bioluminescence activity	25
2.12. Biophysical methods for protein characterization	30
2.12.1. Analytical reversed-phase high-performance liquid chromatography method.....	30

2.12.2. Dynamic light scattering to characterize protein stability	30
2.12.3. Nuclear magnetic resonance	31
2.12.4. Bio-layer interferometry / Octet system	31
2.13. Protein crystallization techniques	32
2.14. Method of hanging drop vapor diffusion for protein crystallization	34
2.15. Method of sitting drop vapor diffusion for protein crystallization	34
2.16. Additive screen	34
2.17. Crystallization <i>in vivo</i>	34
III. Results.....	37
1. Luciferase constructs: cloning, expression, purification and characterization	37
1.1. Construction of the expression vectors for <i>Gaussia</i> luciferase with a thrombin cleavage site and catalytic domains in <i>E. coli</i> cells.....	38
1.1.1. Construction of the expression vector for full-length <i>Gaussia</i> luciferase in <i>E. coli</i> cells.....	41
1.1.2. Generation of recombinant pFastBac1 vector for <i>in vivo</i> crystallization	42
1.2. Cloning results and generation of luciferase expression constructs	47
1.3. Cloning and generation of luciferase constructs for <i>in vivo</i> crystallization.....	50
1.4. Expression and purification of full-length <i>Gaussia</i> luciferase and its catalytic domains	53
1.5. Analytical reversed-phase high-performance liquid chromatography.....	57
1.6. Bioluminescence activity and kinetics.....	60
2. Expression and purification of full-length <i>Gaussia</i> luciferase construct.....	71
3. Biophysical methods for protein characterization	74
3.1. Dynamic Light Scattering to characterize protein stability	75
3.2. Nuclear magnetic resonance spectroscopy	76
3.3. Bio-Layer Interferometry / Octet system.....	79
3.4. Crystallization of <i>Gaussia</i> lucifersase construct.....	81
3.5. <i>In vivo</i> crystallization of <i>Gaussia</i> luciferase.....	84
IV. Discussion	87
V. Conclusions and future perspectives.....	95

VI. Literature..... 99

LIST OF FIGURES

Figure I-1. Five chemical classes of luciferin identified in bioluminescence.....	3
Figure I-2. <i>Gaussia princeps</i>	4
Figure I-3. Amino acid sequence of <i>Gaussia princeps</i> -luciferase.....	5
Figure I-4. Enzymatic reaction catalysed by <i>Gaussia princeps</i>	6
Figure II-1. Phase diagram for protein crystallization.....	33
Figure III-1. Schematic representation of the fusion proteins after expression in <i>E. coli</i> cells	38
Figure III-2. Schematic representation of the cloning strategy of <i>Gaussia</i> luciferase domain constructs for expression in <i>E. coli</i> cells.....	39
Figure III-3. Schematic representation of the cloning strategy of full-length <i>Gaussia</i> luciferase construct for expression in <i>E. coli</i> cells.....	40
Figure III-4. Schematic representation of the protein construct for expression in <i>E. coli</i> cells	41
Figure III-5. Schematic representation of the subcloning strategy of <i>Gaussia</i> luciferase construct for expression in <i>E. coli</i> cells	42
Figure III-6. Schematic representation of the protein constructs of <i>Gaussia</i> luciferase for expression in insect cells	43
Figure III-7. Schematic plasmid maps and amino acid sequence of pFastBac1 expression vector ligated with the insert PCR product to generate constructs for <i>in vivo</i> crystallization	46
Figure III-8. Agarose gel electrophoresis, analysis of PCR amplified DNA products.....	47
Figure III-9. Agarose gel electrophoresis, analysis of digested vectors and PCR- amplified inserts	48
Figure III-10. Agarose gel electrophoresis, analysis of colony PCR from GLuc domain 1.....	48
Figure III-11. Agarose gel electrophoresis, analysis of colony PCR of full-length GLuc thrombin construct for expression in <i>E. coli</i>	49
Figure III-12. Agarose gel electrophoresis, double restriction digestion control test.....	49
Figure III-13. Agarose gel electrophoresis, analysis of colony PCR of GLuc construct for expression in <i>E. coli</i>	50

Figure III–14. Agarose gel electrophoresis, analysis of PCR amplified DNA products for <i>in vivo</i> crystallization constructs	51
Figure III–15. Agarose gel electrophoresis, analysis of colony PCR of GLuc construct for <i>in vivo</i> crystallization.....	52
Figure III–16. Size-exclusion chromatography (SEC) and SDS-PAGE analysis of purified GLucFITrhom.....	54
Figure III–17. Size-exclusion chromatogram phy and SDS-PAGE analysis of purified GLucD1.....	55
Figure III–18. Size exclusion chromatography and SDS-PAGE analysis of purified GLucD2.....	56
Figure III–19. SDS-PAGE analysis of GLucFIThrom after cleavage with thrombin enzyme	57
Figure III–20. Representative analytical reversed-phase chromatography of GLuc construct	58
Figure III–21. Representative analytical reversed phase chromatography for different GLuc constructs	59
Figure III–22. Effect of buffer composition on the stability of <i>Gaussia</i> luciferase as measured by bioluminescence activity.....	61
Figure III–23. Effect of detergents on bioluminescence emission of <i>Gaussia</i> luciferase	62
Figure III–24. <i>Gaussia</i> luciferase luminescence effect of Tween concentration	63
Figure III–25. Bioluminescence-specific activity kinetics of GLuc constructs	64
Figure III–26. Relative luminescence activity of soluble GLuc constructs	64
Figure III–27. Effect of buffer pH on luminescence intensity of <i>Gaussia</i> luciferase and two catalytic domains.....	65
Figure III–28. Effect of salt concentration on bioluminescence emission of <i>Gaussia princeps</i> (GLuc) and <i>Renilla muelleri</i> (RmLuc) luciferase	66
Figure III–29. Effect of salt concentration on full-length GLuc and two catalytic domains	67
Figure III–30. Effect of temperature on <i>Gaussia princeps</i> and <i>Renilla muelleri</i> luciferase activity	68
Figure III–31. Thermostability of <i>Gaussia princeps</i> and <i>Renilla muelleri</i> luciferase.....	69
Figure III–32. SDS-PAGE analysis from a mixture of purified catalytic domains.....	70
Figure III–33. Luminescence emission of a mixture of catalytic domains GLucD1 and GLucD2.....	71

Figure III–34. Schematic plasmid map of recombinant pET28(a)+GLucFl construct for expression in <i>E. coli</i> cells.....	72
Figure III–35. Ion exchange purification chromatogram and SDS-PAGE analysis of purified GLucFl.....	74
Figure III–36. Dynamic light scattering (DLS) experiment of GLucFl	76
Figure III–37. Representative 1D ¹ H-NMR spectra from GLucFl and GLucD1 proteins	78
Figure III–38. Representative 1D ¹ H-NMR spectra of the buffer containing Tris and piperazine	79
Figure III–39. Bio-Layer Interferometry protein interaction experiment of GLuc catalytic domains.....	80
Figure III–40. BLI – Kinetic results of the interaction between both GLuc catalytic domains	81
Figure III–41. Crystallization of recombinant GLucFl construct.....	83
Figure III–42. Light microscopic images of <i>Sf21</i> insect cells with <i>in vivo</i> protein crystal growth	85
Figure IV–1. Amino acids sequence alignment and schematic representation of <i>Gaussia</i> luciferase	89
Figure IV–2. GLuc complementation constructs and ligand-receptor interaction	91

LIST OF TABLES

Table II-1. Kits.....	9
Table II-2. Vectors.....	9
Table II-3. Host strains	9
Table II-4. Primers used in the PCRs to generate expression plasmids for GLuc constructs.....	10
Table II-5. Culture media composition.....	10
Table II-6. Machines and Instruments	11
Table II-7. Columns and resins.....	11
Table II-8. Computer programs and online databases	11
Table II-9. PCR reaction temperature programme	15
Table II-10. Colony PCR reaction temperature programme.....	16
Table II-11. Protocol and buffer composition for 6xHis-tag fusion protein purification	19
Table II-12. Protocol and buffer composition for Strep-tag fusion protein purification	21
Table II-13. Chromatography columns settings	21
Table II-14. Protocol and buffer composition for GLucFl construct protein purification.....	21
Table II-15. SDS-PAGE gel 12%.....	24
Table II-16. Buffers composition for SDS-PAGE.....	24
Table II-17. Buffer assay and reaction mixture	26
Table II-18. Programme assay	27
Table II-19. Buffers used for activity assay.....	28
Table II-20. Detergents used for activity assay	28
Table II-21. Binding kinetic assay protocol.....	32
Table III-1. Constructs designed for recombinant protein expression.....	37
Table III-2. Summary of reversed-phase chromatography parameters	58

INTRODUCTION**1. Bioluminescence****1.1. Marine Bioluminescence**

Bioluminescence is the biological generation of visible light by living organisms (Haddock *et al.*, 2010). Bioluminescence is a widespread phenomenon in nature and has been documented across a wide range of organisms from bacteria and fungi to large squid and fish (Herring 1987, Sverdrup *et al.*, 1942, Harvey 1956, Widder 2001). The majority of bioluminescent organisms are found in the ocean (Harvey, 1956), they usually reside below 800 m depth (Meyer-Rochow, 2001). Marine bioluminescence is mainly blue, and therefore tuned to the wavelengths that travel furthest through the water (Valiadi and Iglesias-Rodriguez, 2013). Bioluminescence is produced by chemical reaction between a substrate (luciferin) and an enzyme (luciferase) in the presence of molecular oxygen (Wilson and Hastings, 1998). Marine organisms utilize bioluminescence for vital functions ranging from defence to reproduction. Therefore, it is considered the most ubiquitous form of inter- communication between the organisms (Herring, 2007).

The function of bioluminescence in the oceans is clearly understood in the context of this dark environment. Below 200 m very little, if any, sunlight filters down (depending on the region) and hence marine organisms mainly communicate by using bioluminescence. For many marine organisms, their primary visual stimulus comes from biologically generated light, rather than direct sunlight (Haddock *et al.*, 2010). Flashes and trails provide information on potential partners, predators, meals and mates (Johnsen *et al.*, 1999; Sverdrup *et al.*, 1942; Harvey, 1956, Lloyd, 1983).

1.2. Chemiluminescence, bioluminescence, fluorescence and phosphorescence

Chemiluminescence is the light emitted by a chemical reaction. Bioluminescence is chemiluminescence catalysed by an enzyme. Consequently, bioluminescence is the emission of light by a living organism (Shimomura, 2006).

Very similar to bioluminescence, fluorescence has also been widely applied in analytical chemistry. The difference between bioluminescence and fluorescence is the need of an excitation source. All bioluminescent light is produced by chemiluminescent reaction (Campbell, 1994), which produces energy allowing electrons to be excited to higher energy levels. This always involves a chemical reaction. In contrast, fluorescence relies on a light source to excite the chromophore (a molecular structure whose presence is responsible for the

colour of the fluorescence light) to an elevated energy level. The chromophore will then relax and return to its ground state, which causes the release of a photon of light (Roda *et al.*, 2008).

One relationship between bioluminescence and fluorescence is that brightly fluorescent proteins, such as green fluorescent protein (GFP), can be co-localized with a luciferase.

1.3. Luciferin-luciferase system

There are three key components of a chemiluminescent reaction. The first component is a luciferin molecule which needs to be oxidised to produce light (substrate). The second component is a luciferase. This is a generic name for protein enzymes that produce light as a by-product when catalysing the oxidation of their specific substrate. The third component is molecular oxygen which is also needed for a chemiluminescent reaction (Haddock *et al.*, 2000). During the bioluminescent reaction, the luciferin is oxidised to oxyluciferin, with the emission of light and carbon dioxide. In addition to the three key components, cofactors can also be involved in the reaction. Typical cofactors include calcium (Ca^{2+}), reduced riboflavin phosphate (FMNH_2^+) and adenosine 5'-triphosphate (ATP). Such cofactors can cause a conformational change in the protein structure (Shimomura and Johnson 1975; Wiles *et al.*, 2005). Aequorin (i.e. the luciferase from jellyfish *Aequorea victoria*) needs Ca^{2+} ions for activity and is therefore widely used as a recombinant cellular [Ca^{2+}]-reporter (Plieth 2006). Firefly luciferase needs ATP as a cofactor, hence it is often used in enzymatic assays for the quantification of [ATP] (Hastings, 1998).

Light emitted from a luciferin-luciferase reaction system can be easily detected, measured and quantified by photomultipliers and photo-counting cameras (Voss *et al.*, 2000; Berthold *et al.*, 2000). Consequently chemiluminescence assays are among the most sensitive quantification methods (Kricka 1991, Lundin 2014).

1.4. The five chemical classes of luciferin

Most bioluminescent reactions need oxygen as an oxidising agent, and an organic substrate. Both are needed to generate an excited state, leading to the emission of light. However, the luminescence reactions in different organisms vary widely, and the enzymes catalysing the reactions (luciferases) and the substrates (luciferins) are also quite distinct.

Five distinct chemical classes of luciferin have been identified: aliphatic aldehydes, benzothiazoles, linear tetrapyrroles, flavins and imidazolopyrazines (Campbell and Herring, 1990) (Fig.I-1).

The most common luciferins in marine organisms are imidazolopyrazines including coelenterazine (CTZ) (Dunstan *et al.*, 2000). CTZ is used by organisms in eight distinct marine phyla: Cnidarians, Ctenophore, Mollusca, Protozoa, Echinodermata, Arthropoda,

Chaetognatha and Chordata (Thomson *et al.*, 1997). Significant levels of CTZ have also been found in some species of non-luminous marine organisms (Shimomura, 2006). Bioluminescent coelenterates, decapods, squid, copepods, radiolarians, some fish and ostracods all use imidazolopyrazine luciferin (Campbell, 2003). The luciferases that use coelenterazine as a substrate, in the presence of molecular oxygen, have a typical reaction that involves the oxidation of the imidazolopyrazine ring with the emission of blue light and carbon dioxide.

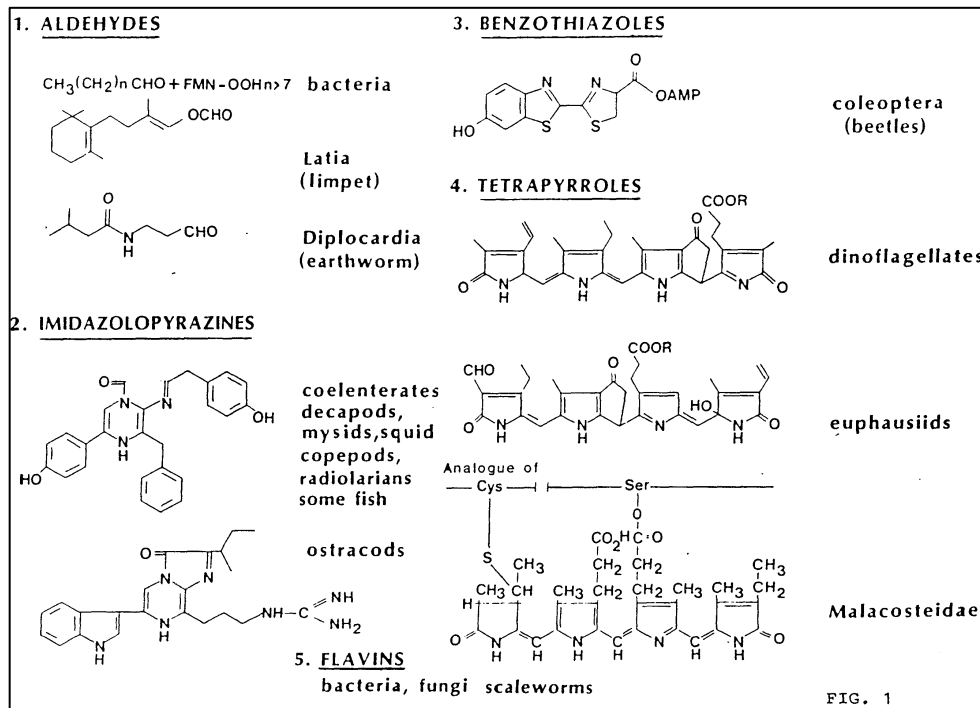


Figure I-1. Five chemical classes of luciferin identified in bioluminescence

Chemical classes of luciferin: aliphatic aldehydes, imidazolopyrazines, benzothiazoles, linear tetrapyrroles and flavins. Source: European patent application. Modified bioluminescent protein and their use.

Source: Campbell A. (2001).

2. *Gaussia princeps*

The marine mesopelagic calanoid copepod *Gaussia princeps* belongs to the Metridinidae family. It is the most widespread of the bioluminescent members of the genus *Gaussia*, first described in 1894 (Scott, 1894).

Gaussia princeps (Fig. I-2) is relatively large and characterized by grey, almost black pigmentation of the cephalothorax mainly in the head. The length of adult females ranges between 10-12 mm and the length of adult males between 9.5-11 mm (Barnes and Case, 1972).

Gaussia princeps is found at tropical temperatures in all oceans of the world. It lives between 200 to 1.000 m in the deep-sea, in the mesopelagic zone (Childress, 1977). During the day it can be found below 400 m and during the night *Gaussia* migrates upwards to an upper limit of 200 m in the sea, indicating a diurnal migratory behaviour of the copepods.

Calanoid copepods luminescence is produced by a secretion of luciferase and luciferin from luminous glands, which are located in the antennae, cephalothorax, thorax and abdomen. The luminous glands respond to different stimulations. Mechanical and electrical stimuli, light and the presence of potential predators can induce a discharge of luminous glands (Clarke *et al.*, 1962; Barnes and Case, 1972; Bannister and Herring, 1989). *Gaussia princeps* produces bright extracellular luminescent blue light, and reacts to the stimuli with a fast backward swimming movement. The cloud of blue light serves mainly as a defence mechanism against predators. It provides a glowing signal and distracts the attention of the aggressor, allowing the copepod to escape (Wiles *et al.*, 2005).



Figure I-2. *Gaussia princeps*

Source: Image from Dr. Steve Haddock, Monterey Bay Aquarium Research Institute (Monterey, CA, USA).

2.1. *Gaussia* luciferase

Gaussia luciferase (GLuc) isolated from marine copepod *Gaussia princeps*, is a monomeric protein composed of 185 amino acids (Fig. I-3) with an N-terminal signal peptide for secretion (the first 17 amino acid residues).

GLuc is the smallest luciferase protein reported so far with a molecular mass of 19.9 kDa. Its theoretical isoelectric point is 6.8 as calculated by ProtParam tools (www.expasy.org/tools).

GLuc has no significant sequence homologies with other coelenterazine-type luciferases (Lorenz *et al.*, 1991). GLuc sequence (GenBank database accession –No. AY015993) was first reported, isolated by cloning and expressed in *E. coli* by Bryan and Gyorgyi (1999). Later Verhaegen and Christopoulos (2002) cloned the luciferase gene from calanoid *Gaussia princeps*. They published the first quantitative analytical study of GLuc and its potential as a new reporter for DNA hybridization.

10	20	30	40	50
MGVKVLFALI	<u>CI</u> AVAEAKPT	ENNEDFNIVA	VASNFAT'TDL	DADRGKLP GK
60	70	80	90	100
KLPLEVLKEM	EANARKAGCT	RGCLICLSHI	KCTPKMKKFI	PGRCHTYEGD
110	120	130	140	150
KESAQGGIGE	AIVDIPEIPG	FKDLEPMEQF	IAQVDLCVDC	TTGCLKGLAN
160	170	180		
VQCSDLLKKW	LPQRCATFAS	KIQGQVDKIK	GAGGD	

Figure I–3. Amino acid sequence of *Gaussia princeps*-luciferase

The protein sequence of *Gaussia* luciferase is composed of 185 amino acids including the signal peptide (underlined letters). The positions of Cys residues are marked in red.

Source: www.ncbi.nih.gov/genbank

GLuc protein is secreted when expressed in mammalian cells (Gaur *et al.*, 2017). *In vitro*, it is possible to quantify this secreted protein from the analysis of cell culture medium alone, thus avoiding the need to lyse the cells. *In vivo*, similar quantification can be undertaken by using fluids such as blood sample (Tannous, 2009).

The secreted luciferase of *Gaussia princeps* catalyses the oxidation of coelenterazine (8 - benzy 1-2-4 hydroxybenzyl) - 6 - (4 - hydroxyphenyl) - imidazo (1, 2-a) - pyrazin -3(7H) - on, C₂₆H₂₁N₃O₃ luciferin; an imidazopiridazinone compound) to produce an excited state of coelenteramide. This emits blue light when relaxing to its ground state (Fig. I-4). Thus, GLuc does not require ATP as a cofactor. The bioluminescence activity of GLuc is highly stable and significantly stronger, between 100 to 1000 times, in comparison to *Firefly* luciferase (FLuc) or *Renilla* luciferase (RmLuc) reported by Remy and Michnick (2006) and Wurdinger *et al.* (2008). GLuc emits light, with a spectral peak of 480 nm, with a broad emission spectrum and detectable signal to 600 nm (Tannous *et al.*, 2005). The light emission spectra of full-length GLuc determined in this dissertation show a peak at 485 nm (Suppl. Fig 4).

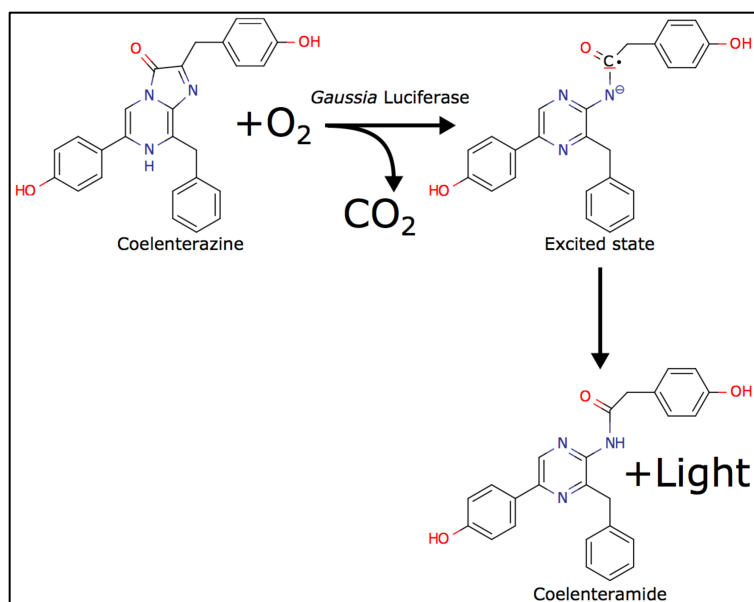


Figure I-4. Enzymatic reaction catalysed by *Gaussia princeps*

Luciferin-luciferase reaction occurs when CTZ is oxidised by O_2 , CO_2 is released and an excited state of coelenteramide is formed. The emission of blue light occurs when the excited coelenteramide relaxes to its ground state.

Source: Nan Wu *et al.* (2015).

GLuc was cloned and expressed in different host organisms such as bacteria (*E. coli*), yeast (*Pichia Pastori*), fungi, baculovirus system/insect cells, mammalian cells and green algae cells (Knappskog *et al.*, 2007, Inouye and Suzuki, 2016).

GLuc is a cysteine-rich protein: the GLuc sequence contains eleven Cys residues (Fig. I-3). Ten of them form disulfide bonds and the eleventh cysteine is located in the pro-sequence (Inouye and Sahara, 2008). Disulfide bonded proteins are difficult to express in *E. coli* (Baneyx and Mujacic, 2004), because the cytoplasm of *E. coli* is constantly maintained as a reduced environment in order to allow thousands of specific reactions. When the cell breaks during protein purification, cysteines are air-oxidized producing non-native disulfide bonds.

Gaussia luciferase was also expressed in cell-free systems (Goerke *et al.*, 2008), where the protein folding environment can be controlled and protein can be recovered from insoluble aggregates with optimization techniques of refolding. However, refolding with five disulfide bonds is a difficult process to recover the functionality of the active protein.

Inouye and Sahara (2008) identified the presence of two domain-like repeats of 71 amino acid residues (27-97, 98-168 aa) in the primary structure of *Gaussia* luciferase. They suggested that the result of this duplication is two catalytic domains working synergistically. In addition, they found that both domains produce bioluminescence when expressed separately. In contrast, Luker and Luker (2014) describe a split *Gaussia* luciferase complementation for imaging of ligand-receptor binding in cells. Here, luciferase complementation requires fusion of inactive amino (N)- terminal and carboxy (C)-terminal enzyme fragments to two different target proteins. Therefore, it is assumed that the

interaction of the target proteins will reconstitute the active enzyme, producing bioluminescence as a quantitative measure.

2.2. Application of *Gaussia* luciferase

A large variety of luciferases have been identified and adopted in cell biology as highly sensitive markers and analytical tools (Badr, 2014). Luciferases are ideal for the development of biological reporter systems because the light they emit can easily be measured and quantified (Maguire *et al.*, 2009). This allows for the tracing and monitoring of the activity of the studied gene, without destroying the cells or tissues. Consequently, luciferases are widely used to study basic cellular processes such as the regulation of transcription and translation. GLuc is also used in biomolecular-imaging applications and for the detection of protein-protein interactions, making the GLuc an invaluable tool for both biomedical and pharmaceutical research (Gaur *et al.*, 2017).

Gaussia luciferase is a very sensitive, non-invasive and non-toxic protein (Welsh *et al.*, 2005). It can be used as a reporter protein in real time, tracking tumour cells and cell metastasis, for monitoring of fungal infections, such as in *Candida albicans* and *Aspergillus fumigatis* (Enjalbert *et al.*, 2009) and the study of viral behaviour within cells (Du Q. Tang *et al.*, 2011). GLuc is ideal for monitoring biological processes directly in culture or *in vivo* in fluids such as blood, serum or urine (Tannous, 2009). While typically *in vivo* imaging experiments are performed using small animal model organisms (mice, rats), secreted luciferases can be applied for imaging larger animal model organisms. GLuc has already been used as a sensitive reporter gene and is useful for monitoring gene transfer in the lungs of mice and sheep (Griesenbach *et al.*, 2011). In addition, GLuc was used to monitor intracellular Ca^{2+} levels (Kim *et al.*, 2009) and for the detection of protein-protein interaction *in vivo* and *in vitro* (Remy *et al.*, 2006).

GLuc expressed in *E. coli* was used as a detection agent in a DNA hybridization assay (Verhaegen and Christopoulos, 2002). Tannous *et al.* (2005) cloned the humanized version of the GLuc enzyme (hGLuc) into a herpes simplex virus-1 (HSV-1) amplicon under a cytomegalovirus (CMV) immediate early (IE) promoter, and evaluated the bioluminescence emitted from DNA-transfected and vector-infected mammalian cells. Wurdinger *et al.*, (2008) evaluated GLuc for imaging after it was cloned into a lentivirus (LV) vector and transduced into human glioma cells that were implanted into mice.

In recent years, bioluminescence resonance energy transfer (BRET) has gained importance as a detection technology for protease activity due to its high sensitivity and low intrinsic background (Branchini, 2016). Li *et al.* (2012) used GLuc as the donor for BRET and enhanced yellow fluorescence protein (EYFP) as the acceptor to detect proteolysis in cells.

GLuc is used in combination with other luciferases (FLuc or RmLuc) to develop double and triple reporters for simultaneous protein-protein interactions within the same cell (Lee, 2008).

3. Motivation, Aim and Outline of the work

The catalytic properties of GLuc make it an extremely sensitive detection reporter. This makes GLuc a potential tool in bio-imaging applications as a reporter protein for *in vivo* tumour growth and signal transduction processes (Maguire *et al.* 2013), and also, when used as a split-reporter protein, for protein-protein interaction studies (Remy and Michnick, 2006).

CTZ-dependent luciferase from *Gaussia princeps* shows no significant sequence homologies with other luciferases even when the enzymes come from the same organism. Therefore, the determination of the three-dimensional structure of the GLuc protein is needed to further explore the underlying reaction mechanism. Knowledge about the architecture of the active site may help to design reporter variants with higher efficiency or with modified spectral characteristics. This would complement the cell biology tool box. In order to elucidate the GLuc structure by X-ray structure analyses, pure and functionally active protein needs to be produced for crystallization and biochemical characterization. The scope of this thesis was to design a full-length GLuc construct, and to establish an expression and purification protocol that produces large amounts of pure and active luciferase protein. Moreover, a full-length GLuc with a thrombin cleavage site between the two catalytic domains and the characterization of the two separately expressed domains was planned. This would allow for investigation of the luminescence properties in more detail and to establish the optimum buffer conditions (e.g. pH, salt, temperature, detergent and reducing conditions) for bioluminescence activity.

Apart from the generation of the two full-length GLuc constructs for expression in *E.coli*, three constructs were designed for expression in Sf21 insect cells using the baculovirus expression vector system (BEVS) to obtain protein crystals in living cells.

II

MATERIAL AND METHODS

1. Material

1.1. Chemicals, consumables and kits

Coelenterazine (CTZ) was obtained from NanoLight™ Technology (Prolume Ltd., Arizona, USA). A 5 mM coelenterazine stock solution was prepared by dissolving it in methanol. Aliquots of 250 µl were stored at -20 °C in the dark until needed. For longer storage, aliquots of 1.250 nMol were dried down using a SpeedVac and re-dissolved in methanol immediately before usage.

Table II–1. Kits

Description	Source
NucleoSpin Gel and PCR Clean-up Kit	Macherey-Nagel, Düren
QIAprep Spin Miniprep Kit	Qiagen, Hilden
QIAGEN Plasmid Midiprep Kit	Qiagen, Hilden

1.2. Vectors and expression host strains

Table II–2. Vectors

Vector	Promoter	Selection	Tag	<i>E. coli</i> strains	Source
pQE30	T5/lac	Amp	N-His	M15[pREP4]; SG13009 [pREP4]	Qiagen
pQE31	T5/lac	Amp	N-His	M15[pREP4]; SG13009 [pREP4]	Qiagen
pET28a+	T7/lac	Kan	N-His and C-His	BL21(DE3)	Novagen
pFastBac1	polyhedrin	Amp, Gen	None	DH10Bac	Life Technologies

Table II–3. Host strains

Strain	Description	Use	Antibiotic	Supplier
XL1-Blue	----	non-expression host; general purpose cloning; plasmid propagation	None	Stratagene
DH5a	----	non-expression host; general purpose cloning; plasmid propagation	None	Invitrogen
BL21 (DE3)	deficient in <i>lon</i> and <i>ompT</i> proteases	general purpose expression host	None	Stratagene
SG 13009 and M15 [pREP4]	contains pREP4	expression host used in combination with most pQE-vectors (Qiagen)	Kan	Qiagen
DH10Bac	contains parent bacmid bMON14272 and helper plasmid pMON7124	host for pFastBac vectors; generation of recombinant bacmid DNA for use in Baculovirus system	Kan + Tet	Life Technologies

1.3. Oligonucleotides

Oligonucleotides used as primers for PCR were designed using the Serial Cloner software (2.61) and synthesized by the Eurofins MWG GmbH (Ebersberg, Germany). The salt-free lyophilized powder was diluted according to the manufacturer's instructions in double distilled autoclaved water to a final concentration of 10 pmol/ μ l stock solutions and frozen at -20 °C.

Table II–4. Primers used in the PCRs to generate expression plasmids for GLuc constructs

Description	Sequence
GLucFIThrom_forward	5'-GAATTCGAGCTCACATATGAAACCAACTGAAAACAATGAAGAT-3'
GLucFIThrom_reverse	5'-CCTAGGAAGCTTGGATCCCTACTTCTCGAACTGCGGGTGGCTCCAATCAC-3'
GLucD1_forward	5'-GAATTCGAGCTCACATATGAAACCAACTGAAAACAATGAAGAT-3'
GLucD1_reverse	5'-GTCTCAGGTACCGCTGCCGCGCGGCACCAGGCCGCTACCGCCGTCACAATAGCCTCTCC-3'
GLucD2_forward	5'-GTCTCAGGTACCGGTGGCACCGGTATTCTGAAATTCCTGGG-3'
GLucD2_reverse	5'-CCTAGGAAGCTTGGATCCCTACTTCTCGAACTGCGGGTGGCTCCAATCAC-3'
GLucFI-C-His_forward	5'-GTCTCATCATGATTCATCACCACCATCATCACAACCAACTGAAAACAATGAAGATTCAAC-3'
GLucFI-C-His_reverse	5'-CAGAGTCTCGAGCTACTTCTCGAACTGCGGGTGG-3'
GLucFI_forward (Insect cells)	5'-AAACCAACTGAAAACAATGAAGATTCAAC-3'
GLucFI_reverse (Insect cells)	5'-ATCACCACCGGCACCCTTTATTTGTCCAC-3'

1.4. Bacterial culture

Lysogeny broth (LB), (Carl Roth, Karlsruhe; Germany), is the rich medium most commonly used to grow *E. coli* cell cultures. The medium components were mixed in distilled water and autoclaved prior to use for 20 min at 121 °C and 2 bar.

Table II–5. Culture media composition

LB medium (Luria-Miller)		LB Agar (Luria-Miller)	
Composition	1/L	Composition	1/L
Tryptone	10g	Tryptone	10g
Yeast extract	5g	Yeast extract	5g
Sodium chloride	10g	Sodium chloride	5g
		Agar - Agar	15g

Components purchased from Carl R oth (Karlsruhe, Germany).

For long-term storage of bacterial strains, bacterial suspension was thoroughly mixed with equal volumes of glycerol and LB-medium. This mixture was aliquotted in 2 ml screw-cap vials, snap frozen in liquid nitrogen and stored at -80 °C.

1.5. Machines, instruments, columns and resins

Table II–6. Machines and Instruments

Machine	Manufacturer
FPLC- Äkta Design and Purifier	GE Healthcare Europe, Freiburg, Germany
DLS- Zetasizer Nano Series comfort	Malvern, Worcestershire, UK
PCR iCycler	Biorad, München, Germany
PCR T100 Thermal Cycler	Biorad, München, Germany
INFORS HAT Multitron II	Infors, Einsbach, Germany
Avanti J-26 XP	Beckman Coulter, Krefeld, Germany
Tecan Infinite M200 Pro	Tecan Trading AG, Switzerland
HPLC System	Agilent Technologies, California, USA
Gen Pulser II	Biorad, München, Germany
Sonoplus ultrasonic homogenizers	Bandelin, Berlin, Germany

Table II–7. Columns and resins

Description	Manufacturer
Ni- NTA Agarose	Qiagen, Hilden, Germany
Strep-Tactin II	IBA, Göttingen, Germany
Pierce Centrifuge Columns, 0.8 ml	Thermo Scientific, Waltham, MA, USA
Superdex 200 10/300 GL	GE Healthcare Europe, Freiburg, Germany
Hitrap QHP Anion Ion exchange	GE Healthcare Europe, Freiburg, Germany
HiTrap Desalting	GE Healthcare Europe, Freiburg, Germany

1.6. Crystallization solutions

Crystallization solutions for the initial screening were used from prefabricated kits (Hampton Research, Aliso Viejo, USA). Other crystallization solutions were prepared in-house. Each stock solution of the individual components was prepared at a concentration of 0.5 M or 1 M and sterile filtered (0.2 μm diameter). 2-Methyl-2,4-pentanediol (MPD) (Qiagen, Hilden, Germany) stock solution was prepared with 60% (v/v) using deionized/MilliQ water as a solvent. This stock solution was not filtered. The crystallization batches were pipetted together immediately before the crystallization screen.

1.7. Computer programs and online databases

Table II–8. Computer programs and online databases

Description	Source
NCBI Genbank	www.ncbi.nih.gov/genbank
SerialCloner V2.6.1	www.serial-cloner.com-about.com
RCSB-PDB	www.pdb.org
FolIndex©	www.biportal.weizmann.ac.il/fldbin/findex.html
ProtParam	www.expasy.org/tools
PSIPRED	www.bioinfo.cs.ucl.ac.uk/psipred
UniProt	www.uniprot.org
ExpASy	www.expasy.ch

2. Methods

2.1. Protein sequence and bioinformatics prediction

Gaussia luciferase (GLuc) is a monomeric enzyme composed of 185 amino acids, including a signal peptide sequence for secretion, with a coding sequence of 555 bp (NCBI GenBank accession number AAG54095 [AY015993], Suppl. Fig. 1).

A secondary structure prediction of GLuc was obtained using the Protein Structure Prediction Server (PSIPRED), a bioinformatic tool provided by the University College London (Suppl. Fig. 2).

The fold prediction analysis, to predict whether a given protein sequence is intrinsically unfolded, was carried out with FolIndex© – a bioinformatic tool provided by Weizmann Institute of Science (Suppl. Fig. 3).

In order to obtain information about the important parameters for expression and purification (MW and pI value) the ExPASy tools (available at <https://www.expasy.org/>) were used.

2.2. Propagation of DNA in *E. coli*

2.2.1. Heat shock transformation and electroporation

E. coli cells were transformed with plasmid DNA using heat shock and electroporation methods. For both methods, suspensions of competent *E. coli* cells were thawed on ice. 10 ng plasmid DNA was mixed with 50 µl of suspension.

For the heat shock transformation method, the mixture of DNA and chemically-competent cells was incubated for 30 min on ice followed by a heat shock for 90 sec at 42 °C. Immediately thereafter, the cell suspension was incubated for 2 min on ice. Subsequently, the cells were re-suspended in 930 µl of fresh LB medium (without antibiotics) and the resulting mixture was incubated for 1 hour at 37 °C, with gentle shaking at 650 rpm followed by centrifugation at 1500g for 3 min. Cells were re-suspended after centrifugation and streaked on agar plates supplemented with the appropriate antibiotic. Agar plates were incubated overnight at 37 °C.

For the electroporation method, 50 µl aliquots of the mixture of DNA and electro-competent cells were placed quickly in a pre-cooled electro-competent cuvette. The settings for electroporation (Gene Pulser 25UF) were 1800 V, 200 Ω and capacity 25 µF to reach a time constant of $RC = 5$ ms. After the electroporation, the cells were re-suspended in 300 µl of fresh LB medium. The resulting mixture was incubated for 1 hour at 37 °C. After incubation, the cells were streaked out on agar plates supplemented with the appropriate antibiotic. The incubation of the plates was carried out overnight at 37 °C.

2.2.2. Amplification and plasmid isolation from *E. coli*

Vectors harbouring the DNA coding for GLuc were amplified in XL1-Blue or DH5 alpha competent *E. coli* cells. Transformation was performed by heat shock transformation or electroporation (section 2.2.1). A single colony from an overnight cultured LB agar plate was used to inoculate 5 ml autoclaved LB medium. The bacterial cells were cultured and incubated overnight at 37 °C, with constant shaking at 225 rpm. Plasmid DNA was isolated from the bacterial culture in accordance with the manufacturer's kit instructions (QIAprep Spin miniprep).

2.2.3. DNA quantification

The concentration of the DNA solution with distilled and autoclaved water was calculated through measuring the absorbance at 260 nm with a spectrophotometer. Distilled water was used as the blank value. DNA concentration was calculated by the rule of proportion according to the $A_{260} = 1 \leftrightarrow [\text{DNA}] = 50 \mu\text{g/ml}$.

2.3. Agarose gel electrophoresis

The separation of DNA fragments was carried out using agarose gel electrophoresis. The agarose concentration of the gels was adjusted according to the molecular size to be analyzed. Generally, 1% agarose gels were used. For smaller fragments (<1000 bp) 1.5% gels were used. The appropriate amount of agarose was boiled in 0.5x Tris-borate-EDTA (TBE) buffer. After the boiled agarose solution was cooled down to about 50 °C, ethidium bromide was added to the agarose solution to a final concentration of 0.003%. A gel casting device with a pocket comb was filled with the boiled agarose solution and allowed to cool down and set. The set agarose gel was placed in a gel chamber filled with 0.5x TBE buffer. The 0.5x TBE buffer composition was: 44.5 mM Tris (pH 7.6), 44.5 mM boric acid and 2 mM EDTA.

The DNA samples to be analyzed were mixed with 6x loading buffer and filled into the gel pockets with a DNA size standard marker (GeneRuler™ 1 kb DNA ladder, Thermo Scientific). The electrophoretic separation of the molecules was carried out at a constant voltage of 120 V for 30 min. Ethidium-bromide was used for the visualization of the DNA bands under a UV/VIS light absorption gel documentation unit (Gel Doc 2000, Gel Documentation System, Bio-Rad Laboratories UK Ltd).

The 6x loading buffer contained 10 mM Tris-HCl (pH 7.6), 60 % (v/v) Glycerol, 60 mM EDTA, 0.3 % (w/v) Bromophenol blue, 0.3 % (w/v) Xylencyanol.

2.4. DNA extraction

DNA fragments were excised from the agarose gel with a clean and sharp scalpel. The gel slice was weighed in a tube and was purified using the NucleoSpin Gel Extraction Kit from Macherey-Nagel (Düren, Germany) following the manufacturer's instructions.

2.5. Restriction digestion and dephosphorylation

The digestion of DNA with restriction endonucleases allow the DNA to be cut at specific sequences, thereby enabling cloning and determination of fragment sizes as well as their location in the vector. The digestion of DNA fragments was performed using Type II restriction endonucleases purchased from New England BioLabs (NEB, Frankfurt am Main). These enzymes recognize defined palindromic DNA sequences of 4 to 6 bp length and cut them at a specific position within their recognition sequence. This produces either sticky-ends or blunt-ends. In this work, both DNA ends were used. The restriction digestions were performed as double digests or two subsequent single digests. The selection of the buffer system was carried out according to the manufacturer's recommendations, so that maximum activity could be ensured for both enzymes.

The linearized DNA vector was dephosphorylated with Shrimp alkaline phosphate (SAP) to prevent re-ligation. To the reaction mixture containing 1 pmol of 5' ends of DNA, 1 U of SAP was added. Further, the reaction mixture was incubated at 37 °C for 30 min. The enzyme was inactivated at 65 °C for 15 min. The dephosphorylated vector DNA was isolated from the remaining reaction by preparative agarose gel electrophoresis. The corresponding band was cut from the gel and purified using gel extraction.

For a digestion reaction, 0.5-1 µg DNA was mixed with 1 U of restriction enzymes and incubated for 1 h at 37 °C in the recommended buffer. The incubation time of the digestion was adapted to the amount of DNA used. After the digestion reaction, the mixture was incubated for 20 min at 80 °C to stop the enzymatic reaction.

The purification of digested DNA was performed using commercially available kit NucleoSpin Gel extraction Kit from Macherey-Nagel (Düren, Germany). Digested vector DNA was purified from a 1% agarose gel by gel extraction (section 2.4).

2.6. Ligation reaction

The ligation reaction was performed in a final volume of 20 µl. Between 50-80 ng of linearized vector DNA were mixed with the DNA to be inserted in a 1:4 ratio. To calculate the required amount of insert DNA, the following formula was used:

$$\text{ng insert} = \frac{(\text{bp insert}) \cdot (\text{ng vector})}{\text{bp vector}} \cdot 4$$

Vector and insert were mixed with ligase buffer and the reaction mixture was complemented by adding 1 U of T4 DNA ligase (Thermo Scientific, Waltham, MA, USA). The reaction was carried out overnight at 16 °C. For heat shock transformation into DH5a or XL1-Blue competent *E. coli*, between 1-5 µl of ligation mixture were used per 100 µl competent cells.

2.7. Polymerase chain reaction (PCR)

Polymerase chain reaction is a technique widely used to amplify small amounts of DNA. GLuc DNA was amplified by using oligonucleotide primers described in Table II-4. PCR was carried out in a thermocycler (iCycler, Bio-Rad, Munich). The GLuc sequences used in the course of the cloning experiments were synthesized using the Phusion High-Fidelity DNA polymerase (Thermo Scientific, Waltham, MA, USA). The parameters of the PCR were adjusted to the primers' melting temperatures, the construct length and the specific properties of the enzyme used according to the manufacturer's recommendation. To determine the optimal annealing temperature, a gradient PCR programme was carried out. The PCR programme used for standard PCR-reaction is shown in Table II-9. Three different annealing temperatures (between 50-60 °C) were tested for each fragment. The volume of an individual PCR reaction mixture was 25 µl.

Table II-9. PCR reaction temperature programme

Step	Time	Temperature	Cycles
Initial denaturation	5 min	94°C	
Denaturation	30 sec	94°C	} x30
Annealing	30 sec	50-60°C*	
Extension	1 min	72°C	
Final extension	10 min	72°C	
Cold phase	∞	4°C	

* The annealing temperature depends on the sequence of the primers.

2.7.1. Colony PCR reaction

To identify positive clones after the transformation of ligation approaches, a colony PCR was performed. Cells of each colony selected from an LB agar plate were re-suspended directly in 10 µl PCR reaction mixture. In the course of the initial denaturation, these cells were lysed. The released plasmid DNA subsequently functioned as a template for the amplification of the target sequence. The primers were chosen such that they would complement the DNA sequence of the cloned insert. The PCR program used is shown in Table II-10.

Table II–10. Colony PCR reaction temperature programme

Step	Time	Temperature	Cycles
Initial denaturation	5 min	95°C	} x30
Denaturation	30 sec	95°C	
Annealing	30 sec	55°C	
Extension	45 sec	72°C	
Final extension	10 min	72°C	
Cold phase	∞	4°C	

A detailed description of the cloning strategy of individual constructs from *Gaussia* luciferase protein is given in the results section (Chapter 3).

2.8. Heterologous protein expression systems

2.8.1. Expression in bacteria

The full-length construct with thrombin cleavage site, and the constructs with separated catalytic domains from *Gaussia* luciferase, were sub-cloned into the vector pQE30 and pQE31 and expressed in *Escherichia coli* strain SG13009(pREP4) (Qiagen, Hilden, Germany). The final plasmid constructs were introduced into competent cells using electroporation (section 2.2.1). For protein expression, a single bacterial colony from an LB plate was cultured in 50 ml LB medium containing ampicillin (100 µg/ml), and incubated overnight at 37 °C with continuous shaking at 160 rpm. The following day, 20 ml of this pre-culture was transferred into 1 L of LB medium to yield an OD₆₀₀ of about 0.1. The expression culture was cultivated for 2 h in a baffled 5 L Erlenmeyer-flask at 37 °C and with shaking at 90 rpm (Infors HT Multitron II, Einsbach, Germany). When the OD₆₀₀ was in a range between 0.5-0.7, the expression was induced with 1 mM IPTG. After 1 h of further shaking at 37 °C, the temperature was reduced to 18 °C for shaking overnight at 100 rpm. The cells were harvested by centrifugation for 10 min at 10,000g and 4 °C. The pellet was frozen and stored at -80 °C until protein extraction and purification.

The full-length GLuc construct sub-cloned into the vector pET28a+ was expressed in *E. coli* strain BL21 (DE3) (Stratagene, San Diego, CA, USA). The construct was introduced into the cells using heat shock transformation (section 2.2.1). The bacterial colony from an LB agar plate was cultured in 50 ml LB medium containing kanamycin (50 µg/ml) at 37 °C with continuous shaking at 160 rpm overnight. This seed culture was transferred into a 1 L flask with LB medium, incubated and induced as described above. One hour after IPTG induction at 37 °C, the incubation temperature was lowered to 20 °C with continued shaking at 100 rpm overnight. The culture was harvested, frozen and stored as described above.

2.8.2. Expression in insect cells

The most commonly used cell lines for the baculovirus expression system are derived from *Spodoptera frugiperda*, so-called *Sf9* and *Sf21*. In this work the *Sf21* insect cell line was used. In general, this cell lines have a generation time of between 24 and 48 h when incubated at 27 °C in culture media. Insect cells were grown in suspension culture to a desired cell density of 1.5×10^6 cells/ml, before they were transfected with a recombinant baculovirus genome containing the gene of interest, called a bacmid.

i) Transposition

To generate recombinant bacmid DNA, the three recombinant plasmids of pFastBac1GLucFl were transformed into DH10Bac competent *E. coli* cells (Invitrogen, Carlsbad, CA, USA). The gene of interest was transposed into the bacmid inducing lacZ gene disruption. Selections of positive colonies were performed via blue-white screening. White colonies containing the recombinant bacmid were selected on LB agar plates supplemented with kanamycin (50 µg/ml), gentamycin (7 µg/ml), tetracycline (10 µg/ml), X-gal (20 µg/ml) and IPTG (40 µg/ml). An overnight culture was prepared from a single white bacterial colony transferred into 10 ml of fresh autoclaved LB medium supplemented with antibiotics (kanamycin, tetracycline and gentamycin at the same concentration as above), to generate low copy number plasmid of high molecular weight DNA. The culture was continuously shaken at 225 rpm at 37 °C for a minimum of 16 h. The DNA was isolated using a Qiagen Plasmid miniprep kit (Qiagen, Hilden). Bacmid DNA was then used for transfection of *Sf21* insect cells.

ii) DNA transfection

Once the recombinant bacmid contained the gene of GLuc, and the correctness of the construct sequence was confirmed by sequencing, the next step was the transfection to produce recombinant baculovirus. The bacmid DNA was transfected into the insect cells using the transfection reagent X-Treme Gene HP (Roche, Mannheim, Germany). Healthy cells with a viability greater than 95% (analysis performed with the trypan blue viability test, Thermo Fischer Scientific, Waltham, USA), and growing in a logarithmic phase with a density of 1.5×10^6 cells/ml were prepared in suspension culture before the transfection.

Under a sterile hood, a 6-well tissue culture plate (Greiner Bio-One, Frickenhausen, Germany) was prepared with about 1×10^6 cells/ml *Sf21* ovary cells per well from the Fall Army worm *Spodoptera frugiperda* (Invitrogen, USA). The wells were filled with fresh SFM4Insect media with L-glutamine, poloxamer 188 and sodium bicarbonate (Thermo Fischer scientific HyClone # SH 30913.02) to a total volume of 3 ml. Two wells were seeded for each DNA plasmid and one well contained only medium to be used as a cell control.

The transfection reagent X-Treme Gene HP (10 μ l) and the bacmid DNA (1 μ g/ μ l per transfection) were diluted in 200 μ l SFM4Insect medium. The resulting mixture of DNA and the transfection reagent was then dropped carefully onto the cells on the prepared 6-well tissue culture plate. The plate was incubated 60 min under the sterile hood at room temperature. After an incubation period of 1 h, the cells were attached to the bottom of the plate. The supernatant was replaced with fresh medium. Cell culture plates were then incubated for another 48-60 hours at 27 °C until signs of viral infection appeared.

iii) Isolation of virus stocks

Transfected cells usually show signs of infection three days post-infection when the cells start to appear larger when observed under the inverted microscope.

At this point, the supernatant from each well (3 ml) containing the P1 virus stock was collected into a sterile 15 ml falcon tube, wrapped in aluminum foil and stored at 4 °C. To generate a high titer (virus stock P2), the initial virus stock P1 was used to infect freshly prepared insect cells. For this purpose 3 ml of cell culture with a density of 0.7 to 1 x 10⁶ cells/ml (viability of 80-90%) as determined using a Neubauer haemocytometer counter chamber, were infected with 200 μ l of virus stock P1 in each well from the 6-well tissue culture plate (Greiner, Sigma Aldrich). The plate was incubated at 27 °C for 48-60 hours. The supernatant was harvested three days post infection and then was transferred into a sterile 15 ml tube. This was the P2 virus stock, stored in the dark at 4 °C. The former procedure was repeated to generate the virus stock P3.

2.9. Protein purification methods

2.9.1. Cell disruption

Frozen bacterial pellets were re-suspended in 35 ml of the pre-cooled lysis buffer (the composition is given in Table II-11) and disrupted by sonication pulses of 0.6 s with a Bandelin sonoplus sonicator UW2200 (Berlin, Germany). Pulsed sonication was performed 6 times for each 30 sec with 30 % power and with 1 min pauses in an ice water bath. The lysates were clarified by centrifugation at 50,000g (Centrifuge AvantiTM J25 with JA-25.50 rotor, Beckman Coulter, Krefeld, Germany) for 1 h at 4 °C. The supernatant containing the soluble overexpressed protein was filtered using a membrane filter of 0.22 μ m pore size before protein purification.

Three different methods of chromatography were used to purify the target proteins, namely affinity chromatography (AC), size exclusion chromatography (SEC) and ion exchange chromatography (IEC).

2.9.2. Affinity chromatography

Affinity chromatography (AC) utilizes the propensity of deprotonated histidine side chains to complex divalent metal cations. Recombinant proteins tagged with a ‘polyhistidine tag’ can be separated from the majority of non-tagged proteins by adsorbance on a solid phase exposing divalent transition metal cations (*e.g.* Ni²⁺ immobilized with nitrilotriacetic acid (NTA) on agarose beads). The bound protein can be eluted through competitive displacement by application of high concentrations of imidazole.

The designed constructs (Table III-1) contain a sequence coding for 6xHis-tag, which was utilized separate the target protein from other proteins. This purification step was carried out in an Econo-Pac gravity flow cartridge (Biorad, München, Germany) filled with Ni²⁺/NTA agarose (Qiagen, Hilden, Germany) pre-equilibrated with wash buffer. Proteins were bound to the pre-equilibrated matrix, followed by washing and elution steps. Buffer compositions and volumes are shown in Table II-11. The protein was eluted with a concentration step of imidazole elution buffer to a final concentration of 500 mM imidazole in standard buffer. One aliquot was saved during each step for quality control analysis (SDS-PAGE gel).

Table II–11. Protocol and buffer composition for 6xHis-tag fusion protein purification

Recommended Buffer	Concentration of ingredients	Buffer column volume
Lysis buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl 50 mM PMSF, 1 mg/ml Lysozyme 20 mM Imidazole	---
Equilibration buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl 20 mM Imidazole	10 CV
Wash buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl 20 mM Imidazole	10 CV
Elution buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl 50 mM, 75 mM, 500 mM Imidazole	8 x 0.5 CV

2.9.3. Affinity chromatography under Strep-tag II matrix

Apart from the hexahistidine tag (6xHis-tag), there are other affinity sequence tags like the octapeptide called Strep-tag (WSHPQFEK) that binds to streptavidin. Recombinant proteins could be purified with high affinity over a Strep-tag II matrix (IBA Göttingen). This matrix is a synthetic streptavidin variant most widely used for affinity chromatography.

The GLuc protein construct with a Strep-tag II affinity sequence tag was purified through a 1 ml Strep-tag II matrix pre-equilibrated with Strep-tag II washing buffer. Further purification steps are described in Table II-12.

2.9.4. Size exclusion chromatography

Size exclusion chromatography (SEC) is a method that separates molecules according to their hydrodynamic volume (their size). This method is based on the differing probability of molecules traversing the pores of the matrix; large molecules do not enter the pores and are eluted in the column's void volume.

Purification of the target proteins was performed with the Äkta FPLC Purifier system by size exclusion chromatography with a Superdex 200 10/30 HR column (GE Healthcare, Little Chalfont, UK) pre-equilibrated in SEC buffer (Table II-15). The proteins were detected by monitoring their UV absorbance at 280 nm. The sample protein to be purified was applied via a sample loop (1 ml). The higher the molecular weight of the protein, the shorter the retention time on the column. The collected fractions were analyzed by SDS-PAGE. 50 µl aliquots of purified protein were flash-frozen in liquid nitrogen and stored at -80 °C.

2.9.5. Ion exchange chromatography

Ion exchange chromatography (IEC) involves two principal steps. First the binding of the protein to a charged resin, and second the elution of the protein from the charged resin using a concentration gradient of 0 to 1 M NaCl. The separation of the protein mixtures depends on the different surface charges. The charge of a protein at a given pH depends on its isoelectric point (pI). For example, in cation ion exchange chromatography, a sample protein with a pI 8.2 must run in a mobile phase buffer of pH 6.0. In anion ion exchange chromatography a molecule with a pI 6.8 must run in mobile phase buffer at pH 8.0.

The sample protein can be eluted from the IEC column with a gradient of salt buffer (*e.g.* NaCl). The higher the charge of the protein, the higher the ionic strength of the buffer is needed for elution. The anion IEC column was connected to the Äkta FPLC Purifier system and the proteins were detected by monitoring their UV absorption at 280 nm. Equilibration of the column was performed with a buffer without salt, the sample was injected and eluted using a NaCl elution gradient from 0 to 1 M over 40-60 CV.

Table II–12. Protocol and buffer composition for Strep-tag fusion protein purification

Recommended Buffer	Concentration of ingredients	Buffer column volume
Lysis buffer	100 mM Tris HCl pH 8.0, 150 mM NaCl 1 mM PMSF, 1 mg/ml Lysozyme	----
Equilibration buffer	100 mM Tris HCl pH 8.0, 150 mM NaCl 1 mM EDTA	2 CV
Wash buffer	100 mM Tris HCl pH 8.0, 150 mM NaCl 1 mM EDTA	10 CV
Elution buffer	100 mM Tris HCl pH 8.0, 150 mM NaCl 1 mM EDTA, 2.5 mM desthiobiotin	6 x 0.5 CV
Regeneration buffer	100 mM Tris HCl, 150 mM NaCl 1 mM EDTA, 1 mM HABA (hydroxy-azophenyl-benzoic acid), pH 8.0	3 x 5 CV

Table II–13. Chromatography columns settings

Column	Column volume	Flow rate	Sample volume	Buffer composition
Superdex 200 10/300 GL	23.6 ml	0.5 ml/min	25 - 500 µl	50 mM Tris-HCl pH 8.0 and 300 mM NaCl
HiTrap QHP Anion Ion exchange	1 ml	1 ml/min	8 - 10 ml	50 mM Tris-HCl pH 8.0 and 50 mM NaCl
HiTrap Desalting	5 ml (void volume 1.5 ml)	1 - 10 ml/min	0.1 - 1.5 ml	50 mM Tris-HCl pH 8.0

Table II–14. Protocol and buffer composition for GLucFI construct protein purification

Recommended Buffer	Concentration of ingredients	Buffer column volume
Lysis buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl 50 mM PMSF, 1 mg/ml Lysozyme 3 U/ml DNaseI, 5 mM MgCl ₂ 20 mM Imidazole	----
Equilibration buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl 20 mM Imidazole 20 mM MgCl ₂ 50 mM KCl	15 CV
Wash buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl 20 mM Imidazole 5 mM ATP	12 CV
Elution buffer	50 mM Tris-HCl, pH 8.0, 300 mM NaCl 50 mM, 75 mM, 500 mM Imidazole	10 x 0.5 CV

2.9.6. Purification protocol of full-length GLuc and its catalytic domains

The expression test results revealed that 37 °C and 1 mM IPTG were the optimal conditions for the expression of all three constructs. A large-scale protein production was performed as described in section 2.8.1. Immediately after cultivation and harvesting, the *E. coli* cells were pelleted and flash frozen at -80 °C. The cell pellets were subsequently resuspended in lysis buffer and disrupted by sonication as described in the protocol in section 2.9.1.

GLuc constructs were exclusively purified from the supernatant (soluble fraction) according to the following procedure. The lysate was cleared of cell debris by centrifugation (1 h, 50,000 g at 4 °C). After the centrifugation the cleared raw extract from GLucD1 was filtered through a 0.22 µm pore size filter to prevent clogging in the subsequent purification steps. The filtered supernatant was loaded onto the Econo-Pac cartridge with 2 ml of Ni²⁺/NTA superflow pre-equilibrated with washing buffer. After the washing steps (see protocol in section 2.9.2, Chapter 2) the bound GLuc was eluted stepwise with elution buffer 50 mM Tris-HCl (pH 8.0), 300 mM NaCl and imidazole from 50 mM to 500 mM. The resulting eluted fractions were analysed using 12% SDS-PAGE under reducing conditions. The gel was stained with Coomassie brilliant blue R-250 dye and the intensity of the bands was proportional to the amount of protein. The purity of the protein and the expected protein size were estimated with the 12% SDS-PAGE analysis.

GLucFITHrom and GLucD2 were purified with a Strep-Tactin II matrix. Purification protocols are in section 2.9.3 and further purification steps are shown in Table II-12. Elution fractions from the proteins purified by affinity chromatography were analysed by 12 % SDS-PAGE.

In order to improve the purity of the protein and remove imidazole or desthiobiotin after affinity chromatography (AC) using Ni²⁺/NTA or Strep-Tactin matrix, a further purification step was assessed based on analytical size exclusion chromatography (SEC) used as the final purification step.

The SEC was performed for all three expression constructs with elution samples from the affinity chromatography purification, either with Ni²⁺/NTA or with the Strep-Tactin matrix. The eluted fraction from each protein was loaded onto a SEC column Superdex 200 10/300 GL (GE Healthcare Europe Freiburg, Germany) using an FPLC system (Äkta Purifier, GE Healthcare, Little Chalfont, United Kingdom). Fractions containing protein (as determined by OD₂₈₀) were analyzed by SDS-PAGE and the fractions were pooled and concentrated using centrifuge-driven filter devices with adequate molecular weight size cut-offs (MWCO 3000 Da). Aliquots of 50 µl were flash-frozen in liquid nitrogen and stored at -80 °C.

2.9.7. Purification protocol of full-length GLuc construct

For purification of GLucFl protein construct, the cell pellets were resuspended in lysis buffer and disrupted on ice by sonication using a Bandelin sonicator (protocol in section 2.5.1). The lysate was further clarified by centrifugation (1 h, 50,000 g, at 4 °C) and the raw extract was passed through a 45 µm filter to prevent clogging in the subsequent purification steps. The filtered soluble supernatant was then loaded onto a pre-equilibrated Ni²⁺-NTA agarose matrix (Qiagen, Hilden) in an Econo-Pac gravity flow cartridge. After several washing steps with the wash buffer (protocol section 2.9.2, Table II-14) GLuc protein was eluted stepwise with the elution buffer containing imidazole concentrations from 20 mM to 500 mM (adjusted with the elution buffer). Eluted fractions were collected and analysed by 12% SDS-PAGE. The fractions with the highest luciferase activity and purity, as indicated by the SDS-PAGE analysis, were further processed. These fractions were desalted on a 5 ml HiTrap-column (settings, Table II-13) prepacked with Sephadex G-25 (GE Healthcare Europe, Freiburg, Germany) into a storage buffer (50 mM Tris-HCl pH 8.0), with the purpose of removing salt and imidazole from the sample, using an FPLC system (Äkta purifier, GE Healthcare Europe, Freiburg, Germany).

For further analytical separations, ion-exchange linear gradient elution was performed (protocol section 2.9.5). The desalted sample protein was directly injected into a pre-equilibrated anion ion-exchange column (HiTrap QHP, GE Healthcare, Germany). The column was equilibrated with low ionic strength buffer (mobile phase) 50 mM Tris-HCl. Subsequently, the gradient elution was started with increasing salt concentration from 0 to 1 M NaCl over 60 column volumes (CV). The flow rate recommended for the ion-exchange stationary phase was 1 ml/min. After 50 CV from the linear gradient elution, 10 CV of isocratic elution were used to elute all the remaining sample protein attached to the matrix. Fractions containing protein (as determined by OD₂₈₀) were analyzed by SDS-PAGE and the purest fractions were pooled and concentrated using centrifuge-driven filter devices with adequate molecular weight size cut-offs. To increase protein concentration, a centrifugal concentrator with MWCO 3000 Da was used for GLuc. Aliquots of 50 µl were flash-frozen in liquid nitrogen and stored at -80 °C.

2.10. Protein analysis of purity and protein concentration

2.10.1. Estimation of protein weight and purity by SDS-PAGE

A common method for separating proteins is the SDS-polyacrylamide gel electrophoresis. The SDS-PAGE, when run under fully denaturing conditions, allows for the estimation of a protein's molecular weight and purity independently of protein folding and amino acid composition.

The protein is denatured by heat in the presence of reducing agents and sodium dodecyl sulfate (SDS) prior to electrophoresis. SDS is a detergent with a strongly anionic head group

which binds the main chain of protein. After protein binding to SDS, the denatured protein's negative charge is directly proportional to its molecular weight. The denatured proteins are pulled by the electric field through an alkaline polyacrylamide gel. The electrophoretic mobility of proteins varies according to their size and charge. The molecular weight is estimated by comparison to molecular weight standards running on the same gel.

The SDS-PAGE analysis was carried out under reducing conditions using 12% acrylamide gels (Rotiphorese Gel 30 (37.5:1); Carl Roth, Karlsruhe, Germany). The SDS-PAGE protocol is given in Table II-15. Samples were mixed at a 4:1 ratio with 4x SDS-PAGE reducing buffer as shown in Table II-16, followed by heating at 95 °C for 5 min and subsequent cooling on ice prior to loading into the gels. The electrophoresis was performed at constant 250 V for 35 min. The gels were then stained with Coomassie Brilliant Blue R-250 (Carl Roth, Karlsruhe) according to standard protocols. Composition of the staining buffer is given in Table II-16.

Table II–15. SDS-PAGE gel 12%

Description	Separating Gel	Stacking Gel
Distilled water	3.3 ml	2.75 ml
1.5 M Tris pH 8.8	2.5 ml	----
1 M Tris pH 6.8	----	0.5 ml
Rotiphorese Gel 30	4 ml	0.65 ml
SDS 10 % (w/v)	100 µl	40 µl
APS 10 % (w/v)	100 µl	40 µl
TEMED	15 µl	5 µl

Table II–16. Buffers composition for SDS-PAGE

Buffer	Concentration of ingredients
4x Loading Buffer	240 mM Tris-HCl pH 6.8, 40 % (v/v) Glycerin 7.5 % (v/v) 2- Betamercaptoethanol, 6.7 % (w/v) SDS, 16 % Bromophenol Blue
10 x Running Buffer	25 mM Tris , 192 mM Glycin, 1 % SDS
Staining buffer	0.01 % (w/v) Coomassie Brilliant Blue R-250 10% (v/v) Acetic acid, 30% (v/v) Ethanol
Destaining buffer	10% (v/v) Acetic acid, 30% (v/v) Ethanol

2.10.2. Determination of protein concentration

i) Bradford method

The Bradford assay is used to determine the protein concentration of a solution. The method is based on the formation of a complex between the dye, the Coomassie Brilliant Blue G-250 (Carl Roth, Karlsruhe), and the protein in the sample solution. The protein-dye complex

produce a shift in the absorption of the dye from 465 to 595 nm and the absorbance at 595 nm is proportional to the concentration of the protein. The protein assay was carried out in 96-well flat-bottom microplate (Greiner Bio-One, Frickenhausen, Germany) using the plate reader Tecan Infinite M200 Pro (Männedorf, Switzerland) following the established protocol by Ernst and Zor (2010). The protein concentration was calculated using a standard curve created using BSA (bovine serum albumin, Pierce Rockford, IL) as a reference protein (Suppl. Fig 5).

ii) Spectrophotometer method

An alternative method of calculating the protein concentration is by measuring the absorbance at 280 nm with a spectrophotometer. The aromatic amino acids absorb light in the ultraviolet range of 280 nm (tryptophan and tyrosine) and 260 nm (phenylalanine). The molar extinction coefficient of a protein at 280 nm is dependent on its aromatic amino acids. It can be determined experimentally or theoretically calculated by the bioinformatics tool ProtParam/ExPASy. Absorbance measurement of the sample protein (3 μ l) was carried out at 280 nm in a spectrophotometer (NanoDrop UV/Vis). The protein concentration was calculated following the Lambert-Beer Law using the molar extinction coefficient of the respective protein. The spectrophotometer was previously calibrated with water or the corresponding buffer as a blank.

2.11. Bioluminescence activity

The bioluminescence activity assay of luciferase protein was performed in the microplate reader Tecan infinite M200 Pro (Männedorf, Switzerland), supported by the i-control software. The luminometer temperature was held constant at 28 °C for all the reactions. The amount of luciferase protein was adjusted to ensure that the light emission did not saturate the detector. All kinetic measurements were carried out using 96-well flat-bottom transparent microplates (Greiner Bio-One, Frickenhausen, Germany).

Bioluminescence activity and kinetic assays for each GLuc construct were performed using native coelenterazine (CTZ, NanoLight Technology) as a substrate. The coelenterazine solution for each experiment was prepared freshly and covered with aluminium foil to protect it from light, since CTZ is very sensitive to light and oxidizing environments (Saleh and Plieth, 2010). Therefore, CTZ solution was also deoxygenated with N₂ gas. To prepare the CTZ for experimental use, a 5 mM CTZ aliquot from the freezer was re-dissolved in 50 μ l ice-cold (-20 °C) methanol. One aliquot of 50 μ l (stock solution) was diluted in 5 ml assay buffer (50 mM Tris-HCl pH 8.2 and 200 mM NaCl) already oxygen-free (saturating the buffer with liquid N₂) and protected from light with aluminium foil. The final concentration of CTZ in the assay buffer was 50 μ M. Dry CTZ is very stable and storage under oxygen-free

conditions is not necessary. Nevertheless, it must be stored ice-cold ($-20\text{ }^{\circ}\text{C}$), preferably in the dark (Plieth, 2006).

The basic procedure for the luminescence reaction was as follow: 3 μl of purified protein solution from each construct concentrated in range between of 20 to 70 $\mu\text{g/ml}$ was mixed with 150 μl assay buffer (Table II-17). The resulting mixture was transferred into a 96-well microplate and then placed in the luminometer plate reader. The luciferase reaction was started by injecting 50 μl of 50 μM coelenterazine (50 mM Tris-HCl, pH 8.2; 200 mM NaCl) into the microplate well. Luminescence was recorded for ca. 1 min using an integration time of 1 s and displayed as counts per second.

In general, the experiments were performed at $28\text{ }^{\circ}\text{C}$, with a kinetic cycle of 60 sec (the programme settings are shown in Table II-18). The luminescence intensity was measured as luminescence counts per seconds. A negative control with assay buffer was used in order to calculate differences in light output between buffer controls and protein samples.

Bioluminescence raw data were extracted from the results in the Tecan infinite M200 Pro. A spreadsheet in Excel 2011 (Microsoft Corporation) was used to prepare the graphs of each individual experiment. Results were analysed with the average background luminescence subtracted from each protein sample. Averages and standard deviations were calculated over three measurements.

To determine the protein concentration in the purified protein samples, a Bradford Coomassie Brilliant Blue assay was carried out using bovine serum albumin (BSA) as a standard (Bradford, 1976). The standard curve of the Bradford protein assay was performed using a 96-well flat-bottom microplate (Greiner Bio-One, Frickenhausen, Germany) according to a protocol from Ernst and Zor (2010).

Table II–17. Buffer assay and reaction mixture

Buffer condition for activity assay
50 mM Tris-HCl pH 8.2, 200 mM NaCl and 0.1% Tween (deoxygenated with N ₂)
Reaction mixture
50 μl Coelenterazine 50 μM (in 50 mM Tris-HCl pH 8.2 and 200 mM NaCl)
150 μl luciferase protein (20 to 70 $\mu\text{g/ml}$) in the buffer for activity assay

Table II–18. Programme assay

Application: Tecan i-control	
Device: infinite 200Pro	
Wait (Plate)	On Target Temperature: 28 °C
Wait (Plate Temperature)	Valid Range: 27 - 30 °C
List of actions in this measurement script:	
Measurement well-wise	
Well Kinetic	
Luminescence	Label1
Condition: Execute at Cycle:	12
Channel A	Inject (Standard)
Volume	50 µl
Speed	200 µl/sec
Refill speed	100 µl/sec
Injection refill volume	600 µl
Label: Label1	
Mode	Luminescence
Well Kinetic Measurement	
Kinetic Cycles	60
Interval Time	Minimal
Attenuation	AUTOMATIC
Color for OD2 Attenuation	
Integration Time	1000 ms
Settle Time	0 ms
Part of Plate	A1-A12
Start Time:	

The following experiments were carried out to determine the bioluminescence activity of *Gaussia* luciferase and its two catalytic domains. In addition, the enzymatic properties of GLuc were compared in some experiments with a purified luciferase from *Renilla muelleri* (RmLuc) provided by Dr. Christoph Plieth.

i) Effect of buffers and detergents on bioluminescence activity of GLuc

To determine the best buffer condition for the activity assay, various buffers and detergents were used as shown in Table II-19 and II-20. Protein targets were purified and concentrated in µg/ml. Aliquots of purified protein were mixed in 100 µl of various buffers. The mixture was placed into a 96-well microplate for further luminescence activity measurements. The enzymatic activity started when the 50 µl CTZ was injected into the samples.

Table II–19. Buffers used for activity assay

Buffer condition for activity assay
100 mM Tris-HCl pH 8.0, 200 mM NaCl
100 mM Tris-HCl pH 8.0, 200 mM NaCl and 0.1% Tween
100 mM Tris-HCl pH 8.0, 200 mM NaCl and 50 mM DDT
100 mM NaPP pH 8.0
100 mM NaPP and 0.1% Tween
100 mM NaPP and 50 mM DTT
100 mM PBS pH 7.4
100 mM PBS pH 7.4 and 0.1% Tween
100 mM PBS pH 7.4 and 50 mM DTT

Table II–20. Detergents used for activity assay

Detergent in the buffer condition for activity assay
100 mM Tris-HCl pH 8.0, 200 mM NaCl and 0.1% Tween
100 mM Tris-HCl pH 8.0, 200 mM NaCl and 0.1% SDS
100 mM Tris-HCl pH 8.0, 200 mM NaCl and 0.1% Triton X
100 mM Tris-HCl pH 8.0, 200 mM NaCl and 0.1% Silvet L 77
100 mM Tris-HCl pH 8.0, 200 mM NaCl and 0.1% Brij 35
100 mM Tris-HCl pH 8.0, 200 mM NaCl and 0.1% Brij 56
100 mM Tris-HCl pH 8.0, 200 mM NaCl and 0.1% CTAB

ii) Determination of the bioluminescence activity of GLuc

To compare the light signals emitted by each sample protein, the bioluminescence of each construct was measured. The relative activity of luciferase was normalized to protein concentration $RLU_N = RLU/\mu\text{g}$. Protein concentration (in $\mu\text{g}/\text{ml}$) was estimated by the Bradford assay to achieve the specific activities of recombinant proteins purified from soluble fractions.

For optimal assay response and to estimate the ideal dilution, luciferase was prepared in 12 serial dilutions with assay buffer in a 96-well microplate. The light reaction was started by the injection of 50 μl coelenterazine (50 μM) into each well. The final reaction volume in each well was 200 μl . Luminescence was then read using the microplate reader Tecan infinite M200 Pro (Männedorf, Switzerland).

Assay buffer without the purified protein served as a negative control. To determine the standard deviation, each individual protein sample was measured in triplicate.

iii) Determination of the effect of buffer pH

To identify the optimum pH for the bioluminescence kinetics, a final buffer volume of 350 ml was prepared containing 50 mM Tris-HCl, 200 mM NaCl and 0.1% Tween. 40 ml aliquots of this buffer were titrated to pH values in the range of 10.5 - 7.0 with intervals of pH 0.5. The titration started at basic pH 10.5. To obtain lower pH values drops of HCl (37% v/v) were

dispensed into the buffer. Luciferase activity was measured by mixing 1 μ l of sample protein diluted in 150 μ l of buffer at various pH values and 50 μ l of CTZ solution.

iv) Effect of NaCl concentration

The effect of NaCl on the luminescence of GLuc was carried out by preparing serial dilutions of the salt with concentrations in the range of 0 mM to 4.096 mM NaCl. The serial dilutions were prepared with the assay buffer mixed with the respective samples of luciferase. The bioluminescence reaction was started with the injection of 50 μ l of CTZ into the 96-well plate in the plate reader. The kinetic reaction was then measured.

v) Effect of different temperature on bioluminescence activity

To determine the effect of different temperatures on bioluminescence activity, the luciferase assay was conducted at different temperatures from 15 °C to 45 °C (at 5 °C intervals) pre-setup in the microplate reader. Measurements of bioluminescence were carried out using the same basic procedure as described in Table II-17.

vi) Effect of thermostability on bioluminescence activity

Thermostability of all samples of the purified proteins was examined. Luciferase activity was measured in triplicate. Before the measurement of bioluminescence, samples were mixed in assay buffer (50 mM Tris-HCl pH 8.3; 200 mM NaCl and 0.1% Tween). The resulting mixtures were incubated at different temperatures in the range from 36 °C to 58 °C, then 60 °C to 82 °C and finally temperature was increased from 78 °C to 100 °C in a thermocycler gradient-PCR machine (Bio Rad DNA Engine Tetrad 2, München, Germany) for 1 h. A 50 μ l aliquot from the samples was kept as a control at room temperature. After heating by PCR the samples were loaded into the 96-well microplate, followed by luminescence recording.

vii) Substrate effect between the two catalytic domains

To determine if a cooperative substrate effect exists between the two catalytic domains, a kinetic experiment was set up as follows: a mixture of the two catalytic domains (1:1) were placed in the microplate reader and the luminescence reaction was started after the injection of 50 μ l CTZ solution. The photon emission was then recorded. The kinetics were measured within 8 hours of experimental time at constant temperature of 28 °C. At the same time, luminescence from a full-length GLuc sample was measured and used as a control.

2.12. Biophysical methods for protein characterization

2.12.1. Analytical reversed-phase high-performance liquid chromatography method

Reversed-phase high-performance liquid chromatography (RP-HPLC) was carried out using an Agilent 1100 system with UV detector (Agilent Technologies, Santa Clara, California USA) controlled by ChemStation software version B.01.03 (Agilent Technologies, Santa Clara, California USA). Detection of the samples in RP-HPLC was carried out at 214.8 nm, which is specific for the detection of peptide bonds. Detection of peptides or proteins at 280 nm corresponds to the absorption of the aromatic amino acids tryptophan and tyrosine. A reverse-phase Jupiter (C5 Phenomenex) column was used. The C5 column characteristics are: 5 µm particle size, 300 Å pore size, 4.6 mm id. (internal diameter) and 250 mm length. RP-HPLC was performed with a mobile phase using a gradient programme: sterile filtered water with 0.1% trifluoroacetic acid (TFA) for eluent A and acetonitrile (CH₃CN) with 0.1% TFA in water (95:5) for eluent B. Both eluents were filtered (0.22 µm filter) and degassed prior to use in order to remove particulates and prevent the formation of air bubbles that could block solvent lines or the column. RP-HPLC was performed with an acidic mobile phase with TFA. TFA is the most commonly used additive because its volatility can help to provide pH control. Acetonitrile is a commonly used organic solvent, which exhibits high optical transparency in the detection wavelengths used for protein analysis.

The flow rate to equilibrate the column was 1 ml/min and the experiment was set up at room temperature. The column was equilibrated to the initial aqueous solvent composition over 20 min. Once a stable baseline in the HPLC was obtained, 20 µl double deionized water was injected in the column as a “blank” run, prior to loading the sample protein. The samples of the SEC-purified proteins were filtered before the injection. Subsequently, every single purified sample in a volume of 20 µl was injected in the HPLC loop in gradient mode of increasing organic solvent concentration from 20% B to 80% B within 30 min with 1 ml/min flow rate at room temperature. Retention times were compared between all the samples of GLuc.

2.12.2. Dynamic light scattering to characterize protein stability

Dynamic light scattering (DLS) was measured using a Zetasizer Nano Series Nano-S (Malvern, Worcestershire, UK). Prior to measurement, the sample was centrifuged for 10 min at 10,000g in order to remove possible dust or other particles that may interfere with the measurement. Once the protein sample was ready, it was carefully loaded into a polystyrene micro cuvette to avoid bubbles on the walls of the cuvette. The cuvette was closed with a plastic lid and was placed into a sample holder on the optics block. The volume of the sample protein to be analysed was minimum 70 µl with a starting concentration of 1 mg/ml. At least

50 cycles were averaged with an acquisition time of 10 seconds. After the measurements, the protein samples were recovered for use in crystallization screens and other experiments.

2.12.3. Nuclear magnetic resonance

Nuclear magnetic resonance (NMR) spectroscopy is a method that relies on the intrinsic spin properties of the nuclei. Spins are exposed to constant magnetic fields and then they are excited by a radiofrequency pulse sequence. The one-dimensional ^1H -NMR spectrum signals allow for a chemical shift analysis of the peptide backbone.

1D proton experiment provides a chemical shift of the hydrogen atoms in a biomolecule. The different ^1H nuclei in a protein resonate at different frequencies and a plot of intensity against resonance frequency is shown in a 1D ^1H -NMR NMR spectrum. The spectrum contains one peak for each backbone amide proton.

To characterize the stability of *Gaussia* luciferase samples a one-dimensional (1D) proton experiment using spectroscopic NMR measurements was performed.

Hydrogen-deuterium exchange experiments were carried out by mixing of purified protein samples (450-500 μl) loaded in 5 mm NMR tubes with 10% of D_2O , and the solvent (H_2O) immediately before data collection. Different concentrations of sample proteins (1.5 mg/ml, 2.5 mg/ml and 3 mg/ml) and two different buffers containing 10 mM and 20 mM Tris-HCl (pH 7.8 and pH 6.4), with 50 mM NaCl were tested. The samples were concentrated using centrifugal concentrator devices (MWCO 3000 Da) at 4 $^\circ\text{C}$ before the experiments to reach a concentration of 1.5 mg/ml to 3 mg/ml. Then, the 1D ^1H -NMR experiment was performed and the spectra were recorded. After the measurements, the samples were transferred back into Eppendorf tubes and stored at -20 $^\circ\text{C}$ for further analysis.

The experiment for ^1H spectra was carried out on a Fourier transform-NMR (FT-NMR) spectrometer AV III 600 (Bruker) at a frequency of 600.1 MHz. The spectrometer was equipped with a triple resonance cryogenic probe, using a temperature of 280 K and trimethylsilyl propionic acid (TSP) as an internal standard substance. Measurements were exclusively carried out by the research group of Prof. Sönnichsen at the Institute of Organic Chemistry, Kiel University.

2.12.4. Bio-layer interferometry / Octet system

Bio-layer interferometry (BLI) is a label-free technology to provide real time information on protein interactions. An Octet Red 96 instrument (fortéBIO, London UK) was used for the experiments. BLI is an optical analytical technique that analyses the interference pattern of white light reflected at two surfaces: a layer of immobilized protein on the biosensor tip and an internal reference layer. The binding between a ligand immobilized on the biosensor tip surface and an analyte in solution produce an increase in optical thickness at the biosensor

tip. The increment in refractive index causes a wavelength shift $\Delta\lambda$. This spectral shift is monitored by the detector and reported as a sensorgram (nm shift). Interactions are monitored in real time (rates of association, dissociation or concentration) with precision and accuracy. Kinetic data were obtained using the Octet data acquisition software and a microplate format.

To determine the molecular interaction of the *Gaussia* luciferase split between GLucD1 and GLucD2, a kinetic assay was performed, assay buffer (50 mM Tris-HCl, pH 8.0; 200 mM NaCl), regeneration solution, the ligand (GLucD2SASStrep-tag II) and the analyte (GLucD1) were prepared in advance. The biosensor was hydrated for at least 10 min and the buffer was loaded into the assay plate followed by the ligand, again the buffer assay and finally the analysed protein sample. The plate was placed in the Octet system RED96 from Pall *fortéBio* and the plate layout was defined. Protein samples were recovered on ice after the kinetics assay for further experiments.

Before starting the kinetic reaction in the BLI, the high-affinity monoclonal antibody purified mouse anti-Strep-tag II Mab (StrepMAB-Immo, IBA, Germany) was preloaded onto the anti-mouse IgG Fc Capture Biosensor tip (AMC, *fortéBio*) in the Octet system. The kinetic measurement was performed at room temperature. Protocol composition is shown in Table II-21.

Table II–21. Binding kinetic assay protocol

Description	Composition
Sensor	AMC soaked in PBST + 0.1 % BSA
Preload	StrepMAB at 10 µg/ml in PBST + 0.1 % BSA + 0.05 % P20
Ligand	GLucD2SASStrep-tag II at 10 µg/ml in Tris-HCl (pH 8.0) and 200 mM NaCl
Analyte	GLucD1 at 10 µM in Tris-HCl (pH 8.0) and 200 mM NaCl

2.13. Protein crystallization techniques

To determine the three-dimensional structure of a protein by X-ray crystallography, single protein crystals (>10-50 µm) are required. Pure, soluble and homogenous protein can be crystallized in a single crystal. Different methods to grow crystals can be used. The goal of each method is to bring the protein to a supersaturated state by the addition of a controlled range of precipitating chemicals and/or by changing protein concentration.

A detailed description of the theory and practice behind X-ray crystallography is beyond the scope of this thesis. Therefore, only a basic description of crystallization theory and methods is presented here, following the work of Rupp (2007) and Giacovazzo *et al.* (2011).

Crystallization is the process of arranging atoms and molecules that are in a fluid or solution state into an ordered solid state. This process occurs in two steps, nucleation and growth.

Nucleation occurs at the seed of the crystal, and crystal growth is a dynamic process occurring in equilibrium where dissolved molecules precipitate out of the solution.

There are many different crystallization techniques used for protein crystallization, e.g. vapour diffusion, batch crystallization under oil, dialysis and free-interface diffusion (Fig. II-1). In this dissertation, a vapour diffusion method (hanging drop and sitting drop) was used exclusively for crystallization screening. In this method, the protein solution was mixed with the precipitant solution on a cover slip and was stored in an airtight container either hanging on the container lid (“hanging drop vapor diffusion”) or sitting on an elevated pedestal (“sitting drop vapor diffusion”). A drop composed of a mixture of protein solution and precipitant solution is placed in vapour equilibration over a large volume of the reservoir solution (“mother liquor”). The concentration of the precipitant is higher than that in the drop. Due to the concentration gradient, the water from the drop evaporates and the protein concentration increases, until an equilibrium of vapor pressure is reached. The resulting increase of protein and precipitant concentration in the drop leads eventually to the formation of protein crystals. However, amorphous precipitates can also arise with no crystallization at all.

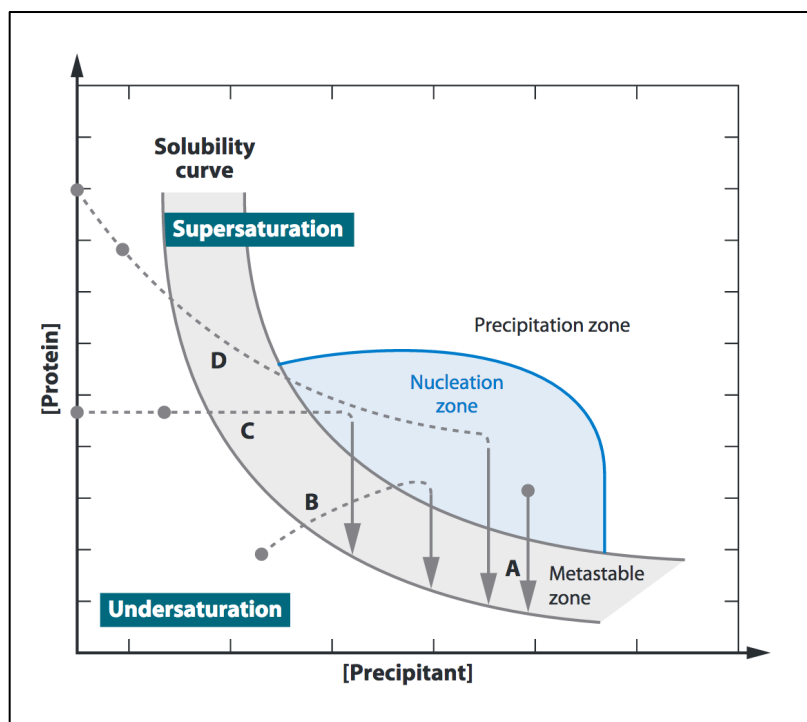


Figure II-1. Phase diagram for protein crystallization

Protein crystallisation starts with nuclei formation in a phase of supersaturation which can be reached by different methods like Microbatch (A), Vapor diffusion (B), Microdialysis (C) and free- interface diffusion (D). Crystals only form in the zone of supersaturation (on the right of the solubility curve), where the concentration of the protein in the precipitant solution is above its solubility. In the zone of undersaturation (on the left of the solubility curve), crystals dissolve and cannot form.

Source: Li and Ismagilov, 2010.

2.14. Method of hanging drop vapor diffusion for protein crystallization

Initial screening for crystallization was performed in 24-well plates (Limbro). To obtain a hermetically sealed system, the edges from each well were previously greased with silicone paste. 500 μ l or 1 ml reservoir solution was deposited in each well. 1 μ l protein and 1 μ l reservoir solution were mixed on a circular cover slip. Each well was closed with the cover slip (22 mm diameter) and the labelled plates were store at 18 °C. The drops were inspected regularly under the microscope.

2.15. Method of sitting drop vapor diffusion for protein crystallization

Initial screening for optimization of crystallization conditions were prepared with a liquid handling robot (Mosquito, TTP Labtech, Melbourn, UK). Plates with 96 wells (Swissci plates 3-drop UVP sitting drop plates, TTP Labtech, Melbourn, UK) were used with a reservoir volume of 100–200 μ l. Protein samples were mixed 1:1 (50 nl protein plus 50 nl reservoir solution) and placed in each well for crystallisation. Labelled plates were stored at 18 °C.

2.16. Additive screen

A screening of additives was performed either with a commercial kit or with buffers prepared in-house in order to optimise an already known reproducible crystallization condition. Additives can be a variety of substances, *e.g.* salts, detergents, amino acids, co-factors, reducing agents, alcohols, which help in optimizing the crystallization process or increasing the packing symmetry of crystals. 450 μ l of reservoir and 50 μ l of additive (Additive screen, Hampton Research) were mixed for the crystallization screen. The screening was performed at the High-throughput Crystallization Laboratory at the EMBL-Grenoble.

2.17. Crystallization *in vivo*

Protein crystallization may occur as a native process and crystals may appear in living cells (Koopmann *et al.*, 2012 and Redecke *et al.*, 2013). For *in vivo* crystallization, a recombinant baculovirus expression system is required (chapter 2.8.2). After the construction of the recombinant baculovirus containing the gene of GLuc (validated by PCR and sequencing), the next step was the transfection. This is where the baculovirus takes control of the insect cells by inhibiting the transcription of the cellular genes until the infected cells die. In this process P1 virus stock is obtained (detailed description is in section 2.8.2). For a more efficient *in vivo* crystallization from recombinant protein, Duszenko *et al.* (2015) suggest to use a high-titer virus stock. Therefore, P3 virus stock was used to infect 70% confluent monolayer cell culture of Sf21 insect cells (cell density 1×10^6 cells/ml with 80-90% viability). The cell culture plate was incubated at 27 °C. Inspections were performed after

three or five days under light microscopy using a 200-fold to 400-fold final magnification. Pictures were recorded with a digital camera (Pentax MX1).

III

RESULTS

1. Luciferase constructs: cloning, expression, purification and characterization

The recombinant *Gaussia* luciferase enzyme was successfully expressed in *Escherichia coli* cells, purified and characterized. In the course of this work, seven constructs were produced for the heterologous expression of *Gaussia* luciferase (GLuc) (Table III-1).

Two construct variants of the entire region were built:

The first construct of GLuc with a thrombin cleavage site between amino acids 97/98 (pQE31GLucFIThrom) was generated to study the two catalytic domains, and to investigate the luminescence properties including the optimum conditions (pH, salt and temperature) for bioluminescence activity.

The second construct (pET28a(+))GLucFI) of full-length GLuc was produced to obtain soluble and highly pure protein for crystallization, in order to determine the three-dimensional structure using X-ray from a single crystal diffraction. In addition, this construct was used to functionally characterize the enzymatic activity with various biochemical and biophysical methods.

The two catalytic domains (pQE31GLucD1Throm and pQE30GLucD2) were expressed separately to investigate their catalytic function and to elucidate whether both domains have luminescence activity. In addition, kinetic experiments were performed to investigate whether the mixture of both catalytic domains acts synergistically in terms of light yield.

Table III–1. Constructs designed for recombinant protein expression

Construct	Expression product
pQE31GLucFIThrom	GLuc fused with N-terminal 6xHis-tag, thrombin site between 97/98 aa C- terminal Strep-tag II
pQE31GLucD1Throm	GLucDomain1 fused with N-terminal 6xHis-tag and C-terminal thrombin site
pQE30GLucD2	GLucDomain2 fused with N-terminal 6xHis-tag and C-terminal Strep-tag II
pET28a(+))GLucFI	GLuc fused with C- terminal 6xHis-tag
pFastBac1_ERGLucFI	GLuc with C- terminal ER localization signal
pFastBac1_ER-KDELGLucFI	GLuc with C- terminal ER localization and retention signal
pFastBac1_SKLGLucFI	GLuc with C- terminal Peroxisomal localization signal

* Note. ER (endoplasmic reticulum).

In addition to the individual constructs of full-length GLuc for expression in *E.coli*, three constructs were generated for expression in *Sf21* insect cells using the baculovirus expression vector (pFastBac1) system (BEVS) for protein crystallization in living cells (Table III-1).

1.1. Construction of the expression vectors for *Gaussia* luciferase with a thrombin cleavage site and catalytic domains in *E. coli* cells

The pQE expression vector was used to express *Gaussia* luciferase (GLucFITHrom) and two separate catalytic domains (GLucD1 and GLucD2). A pET19b vector with the full-length gene construct was used as a template for PCR.

An internal thrombin cleavage site with the amino acid sequence Leu-Val-Pro-Arg-/Gly-Ser (LVPRGS) was placed into a region of GLucFI between the amino acids (aa) 97/98 (numbering based on the amino acid sequence without a signal peptide and affinity tags). Additional base triplets coding for flexible linker (GGSG and GGTG) were introduced via primers between the luciferase domains and the thrombin site. Fig. III-1 shows the schematic representation of the expression constructs of GLuc.

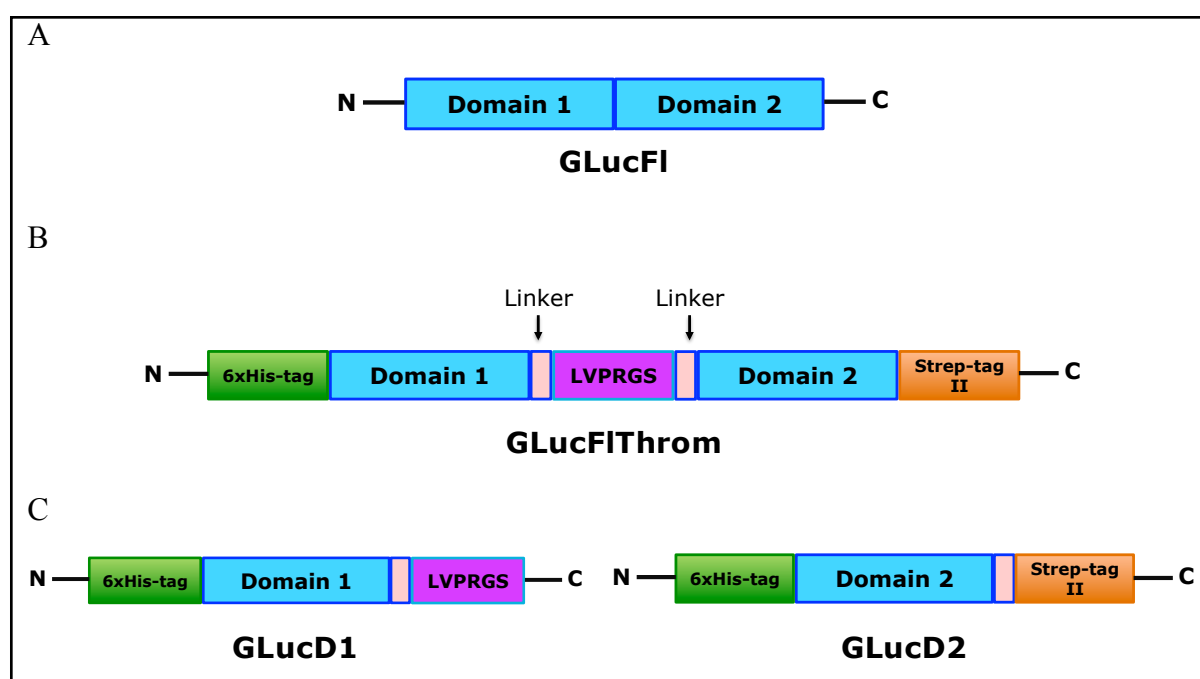


Figure III-1. Schematic representation of the fusion proteins after expression in *E. coli* cells

A) The full-length GLuc protein (GLucFI). B) Fusion protein, which is expressed from the construct pQE31GLuc full-length containing a central thrombin cleavage site (GLucFITHrom) and additional flexible linkers between amino acids (aa) 97/98 fused at its N-terminus with 6xHis-tag and at its C-terminus with Strep-tag II. C) Full-length GLuc split in two individual catalytic domains: GLucD1 fused with 6xHis-tag in the N-terminus and the thrombin site in the C-terminus; and GLucD2 fused with 6xHis-tag in the N-terminus and the Strep-tag II in the C-terminus.

The DNA coding for the two catalytic domains were amplified by PCR and the resulting fragments were cloned separately as shown in Fig. III-2. The first domain (GLucD1) was extended at the 3'-end by a sequence coding for the thrombin site (LVPRGS).

The resulting construct harbouring this PCR-fragment was named pQE31GLucD1Throm (aa 1-97). The second domain (GLucD2) was extended at the 3'-end by a sequence coding for a Strep-tag II (WSHPQFEK). The resulting construct was named pQE30GLucD2 (aa 98-168) (Fig. III-2).

The full-length GLuc thrombin construct was produced by subcloning the GLucD1 fragment into the vector harbouring the GLucD2 fragment and/or the subcloning of GLucD2 into the vector harbouring GLucD1 (Fig. III-3).

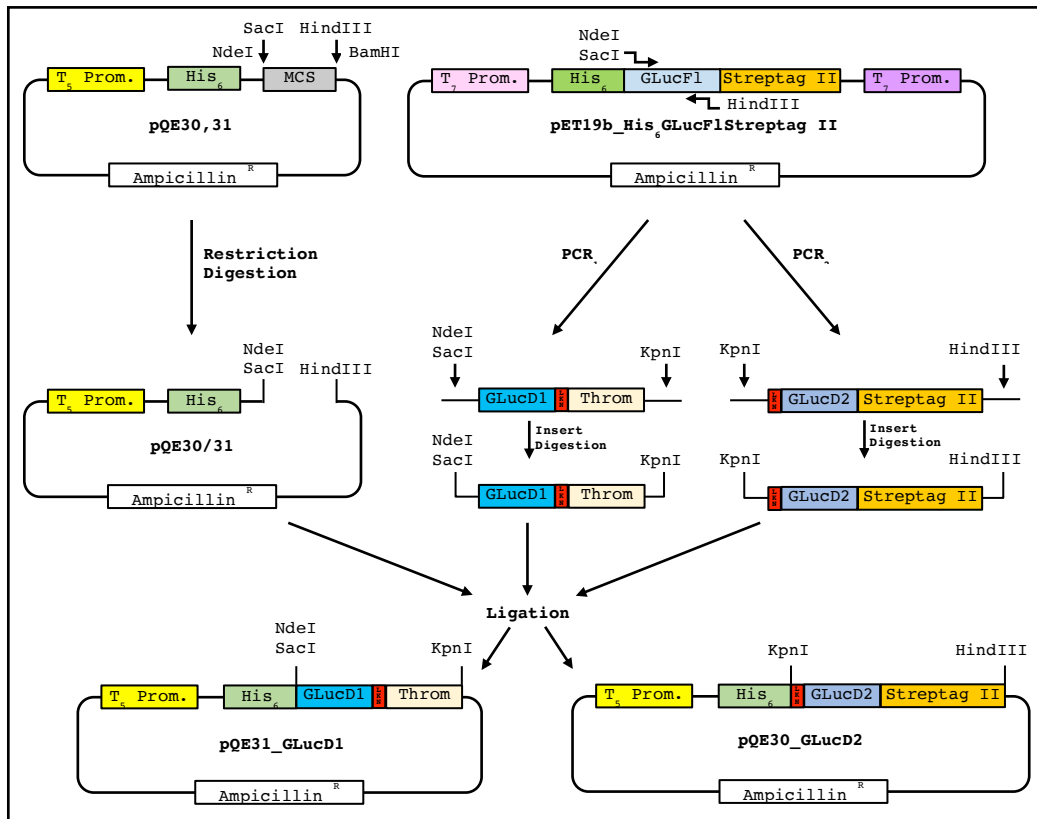


Figure III-2. Schematic representation of the cloning strategy of *Gaussia* luciferase domain constructs for expression in *E. coli* cells

Cloning strategy of the individual domains of GLuc with the internal thrombin cleavage site designed for expression in *E. coli* cells.

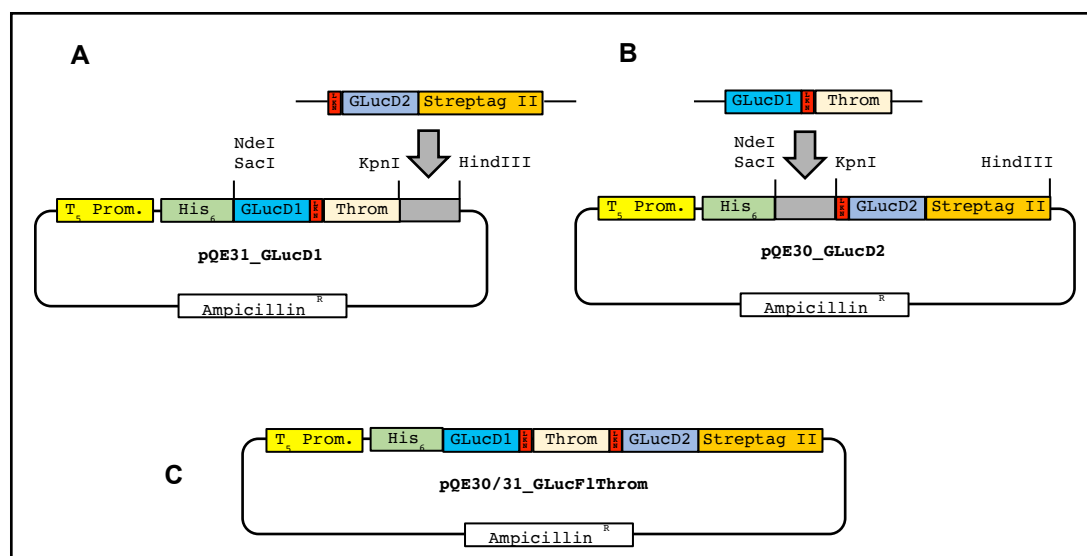


Figure III–3. Schematic representation of the cloning strategy of full-length *Gaussia* luciferase construct for expression in *E. coli* cells

Cloning strategy of the full-length GLuc protein construct designed for expression in *E. coli* with the internal thrombin cleavage site and the subcloning steps to introduce the catalytic domains. A) The subcloning of GLucD2 into the vector harbouring GLucD1 fragment. B) The subcloning of the GLucD1 fragment into the vector already harbouring GLucD2. C) The full-length GLuc construct with the internal thrombin cleavage site.

DNA amplification of *Gaussia* luciferase was performed by PCR reaction in a thermocycler using the primers listed in Table II-4, Chapter 2. After the PCR reaction (programme settings are shown in Table II-9, Chapter 2), the separation of different sizes of DNA fragments was achieved by agarose gel electrophoresis. DNA Gel imaging documentation system (Gel Doc 2000 documentation system, Bio-Rad Laboratories, UK) was used for visualisation of the of the nucleic acid fragments. In order to isolate the DNA from the agarose gel, the corresponding visible bands were excised from the gel for further extraction and purification. The amplified PCR products were purified from the gel using the NucleoSpin Gel and PCR Clean-up Kit (Machery-Nagel, Düren, Germany). The PCR fragments digested with the respective restriction enzymes (*SacI*, *KpnI* and *HindIII*, Fermentas, now part of Thermo Fischer Scientific, Waltham, USA), were inserted into the pQE vector (Qiagen, Hilden, Germany), (see Suppl. Fig. 9). After appropriate single or double digestion the insert fragments were ligated into the digested pQE vector using a T4-DNA ligase (Thermo Fischer Scientific, Waltham, USA) overnight at 18 °C. The host *E. coli* strain XL1-Blue was used for DNA amplification. The correctness of the expression cassette was verified by sequencing analysis at Eurofins Genomics (MWG, Ebersberg, Germany). The resulting protein constructs were named as follows: pQE31GLucF1Throm fused with 6xHis-tag sequence at its N-terminus and Strep-tag II sequence at its C-terminus, pQE31GLucD1 fused with 6xHis-tag at its N-terminus and the thrombin cleavage site at its C-terminus, and pQE30GLucD2 fused with 6xHis-tag at its N-terminus and Strep-tag II at its C-terminus (Circular plasmid map of ligation PCR inserts into vector pQE31 vector is shown in Suppl. Fig. 10).

1.1.1. Construction of the expression vector for full-length *Gaussia* luciferase in *E. coli* cells

The Serial Cloner 2.61 software was used to develop a cloning strategy for a full-length GLuc construct in pET28a(+) vector with a His-tag at the C-terminus of the protein (Suppl. Fig. 6). The gene coding for *Gaussia* luciferase was amplified by PCR (the primers are shown in Table II-9) using the pQE30GLucFl construct (created by Dr. Christoph Plieth at the Centre for Biochemistry and Molecular Biology, Kiel University) as a template. The cDNA of GLuc was subcloned into *NcoI/BspHI* and *XhoI* (Fermentas, Thermo Fischer Scientific, Waltham, USA) restriction sites of pET28a(+) vector (Merck, Darmstadt, Germany). Digested insert and the linearized expression vector were ligated overnight at 18 °C with T4-DNA ligase (Thermo Scientific, Waltham USA). Fig. III-5 shows the schematic representation of the cloning strategy for the GLucFl construct for expression in *E. coli*. The host *E. coli* strain XL1-Blue was used for DNA amplification. The correctness of the GLuc sequence insertion into the vector was verified by DNA sequencing. The resulted construct was named pET28a(+)GLucFl with a 6xHis-tag at its C-terminus (Fig. III-4).

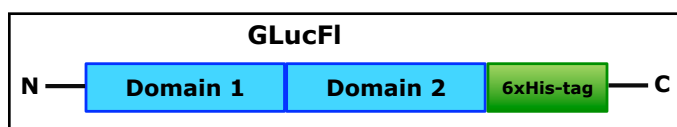


Figure III-4. Schematic representation of the protein construct for expression in *E. coli* cells
Full-length GLuc construct was tagged with 6xHis-tag at its C-terminus.

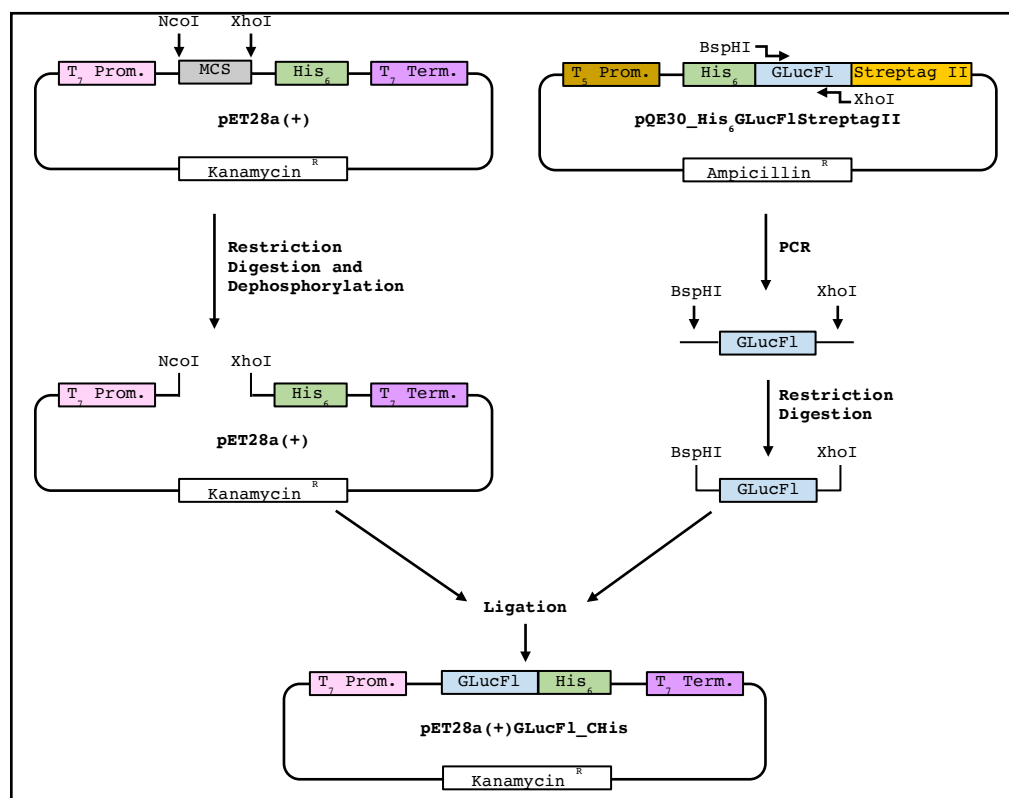


Figure III-5. Schematic representation of the subcloning strategy of *Gaussia* luciferase construct for expression in *E. coli* cells

Subcloning strategy designed of full-length GLuc construct for expression in *E. coli* cells using pET28a(+) vector digested with *Nco*I and *Xho*I restriction enzymes. *Bsp*HI used as *Nco*I cleavage within GLuc insert. This enzyme produces compatible ends to *Nco*I cleavage fragments.

1.1.2. Generation of recombinant pFastBac1 vector for *in vivo* crystallization

An important component of the recombinant Baculovirus expression system (BEVS) is the pFastBac1 vector, where the genes of GLuc were subcloned into, following the Bac-to-Bac manual from Invitrogen, USA. Here, this system is used to express the target protein in insect cells for *in vivo* crystallography screenings.

The oligonucleotide primers were designed with non-stop codon using the Serial Cloner software 2.61 to introduce the GLuc sequence (Suppl. Fig. 18) into the pFastBac1 vector (Invitrogen, USA), which is a donor plasmid used in the baculovirus expression system (Fig. III-7). The pFastBac1 vector is a non-fusion vector (*i.e.* non-fusion tags are present in the vector; Suppl. Fig. 17). Prof. Dr. Lars Redecke, Institute of Biochemistry, University of Lübeck, provided the pFastBac1 vectors to our group. In this experiment, three pFastBac1 constructs containing different localization signals were used. These include pFastBac1-ER with the endoplasmic reticulum signal localized at the N-terminus, pFastBac1-ER-KDEL (ER-sequence at the N-terminus and the ER-retention sequence KDEL at the C-terminus) and pFastBac1-SKL (peroxisomes signal at the C-terminus) (Fig. III-6). In many resident ER proteins a signal KDEL (*i.e.* the tetrapeptide lysine-aspartate-glutamate-leucine) motif is

located at the C-terminus. The peroxisomal targeting signal is the tripeptide sequence SKL (serine-lysine-leucine) attached at the C-terminus (Tang and Teng, 2005).

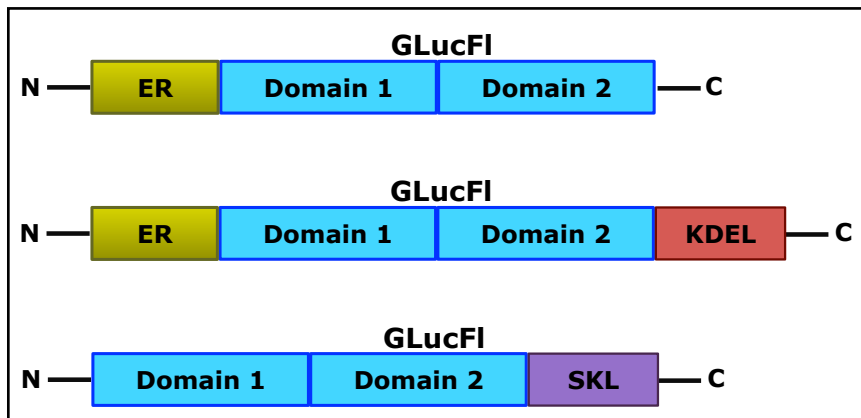
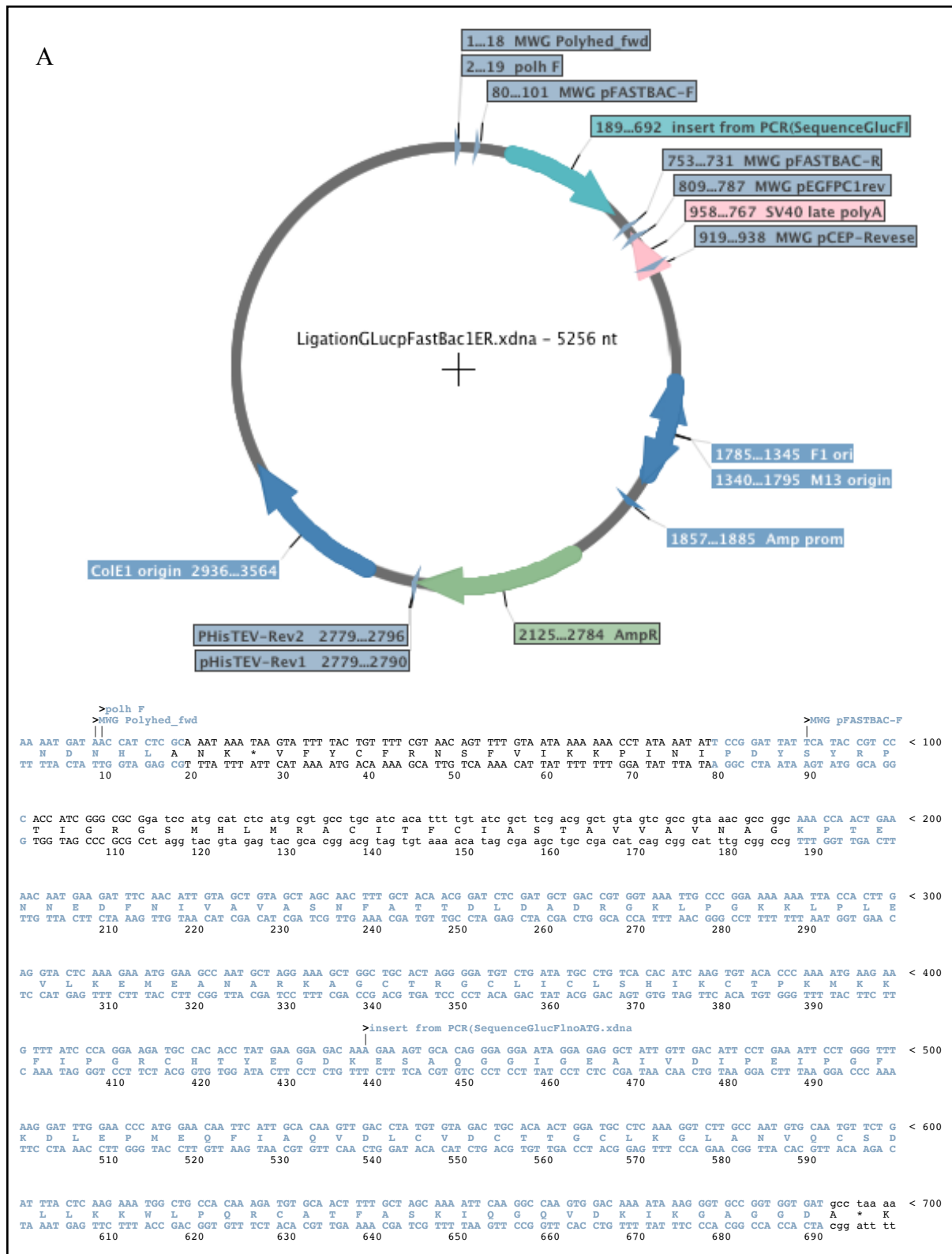
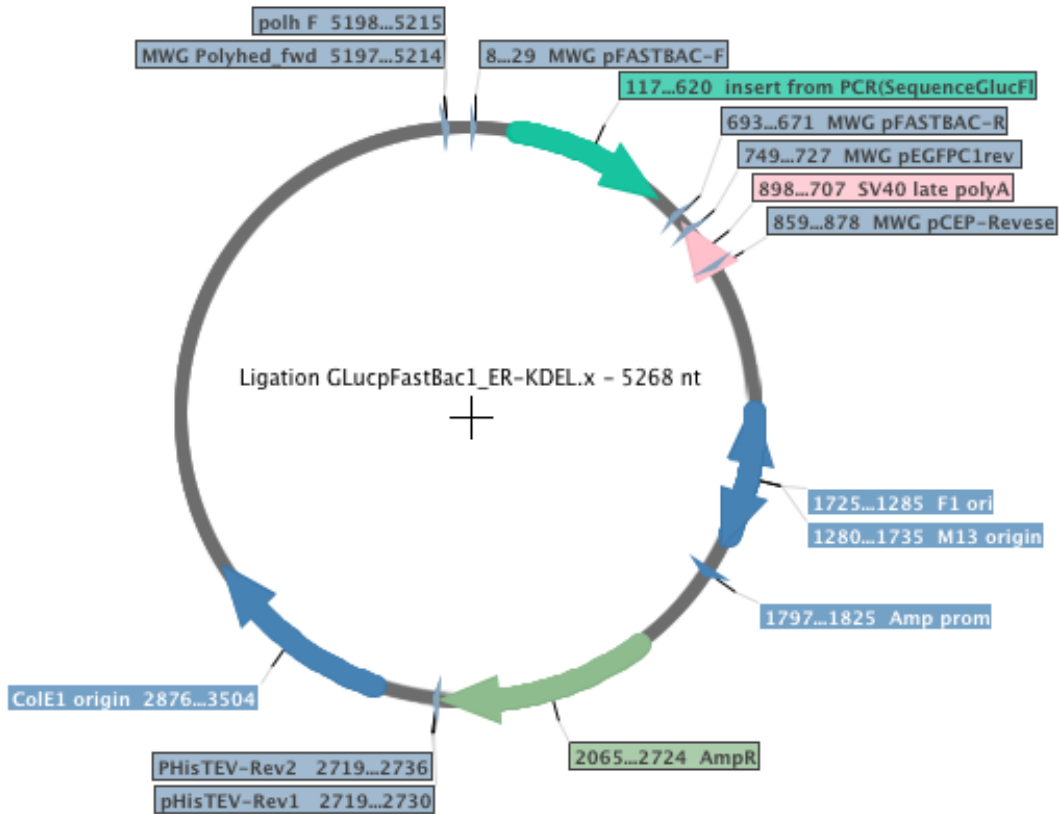


Figure III-6. Schematic representation of the protein constructs of *Gaussia luciferase* for expression in insect cells

Representation of the three fusion protein constructs of full-length GLuc expressed in *Sf21* cells using the pFastBac1 vector that contains different localization signals. The localization signals are: ER-endoplasmic reticulum signal localized at the N-terminus, KDEL (Lys-Asp-Glu-Leu) and SKL (Ser-Lys-Leu) signal localized at the C-terminus. These constructs were used for *in vivo* crystallization.



B



```

>MWG pFASTBAC-F
TA AAT ATT CCG GAT TAT TCA TAC COT CCC ACC ATC GGG CGC Gga tcc atg eat ctc atg egt gcc tgc atc aca ttt tgt atc gct tgc acg get gta gt < 100
K N I P D Y S Y R P P T I G R G S M H L M R K A C I T F C I A S T A V V
AT TTA TAA GGC CTA ATA AGT ATG GCA GGG TGG TAG CCC GCG Cct agg tac gta gag tac gca cgg acg tag tgt aaa aca tag cga agc tgc cga cat ca
    10      20      30      40      50      60      70      80      90

c gcc gta aac gcc ggc AAA CCA ACT GAA AAC AAT GAA GAT TTC AAC ATT GTA GCT GTA GCT AGC AAC TTT GCT ACA ACG GAT CTC GAT GCT GAC CGT GGT < 200
A V N A G K P T E N N E D F N I V A V A S N F A T T D L D A D R G
g cgg cat ttg cgg ccg TTT GGT TGA CTT TTG TTA CTT CTA AAG TTG TAA CAT CGA CAT CGA TCG TTG AAA CGA TGT TGC CTA GAG CTA CGA CTG GCA CCA
    110     120     130     140     150     160     170     180     190

AAA TTG CCC GGA AAA AAA TTA CCA CTT GAG GTA CTC AAA GAA ATG GAA GCC AAT GCT AGG AAA GCT GGC TGC ACT AGG GGA TGT CTG ATA TGC CTG TCA C < 300
K L P G K K L P L E V L K E H E A N A R K A G C T R G C L I C L S H
TTT AAC GGG CCT TTT TTT AAT GGT GAA CTC CAT GAG TTT CTT TAC CTT CGG TTA CGA TCC TTT CGA CCG ACG TGA TCC CCT ACA GAC TAT ACG GAC AGT G
    210     220     230     240     250     260     270     280     290

AC ATC AAG TGT ACA CCC AAA ATG AAG AAG TTT ATC CCA GGA AGA TGC CAC ACC TAT GAA GGA GAC AAA >insert from PCR(SequenceGlucFlnoATG.xdna < 400
F I C T P K M K F I P G R C H T Y E G D K E S A Q G G I G E A I
TG TAG TTC ACA TGT GGG TTT TAC TTC AAA TAG GGT CCT TCT ACG GTG TGG ATA CTT CCT CTG TTT CTT TCA CGT GTC CCT CCT TAT CCT CTC CGA TA
    310     320     330     340     350     360     370     380     390

T GTT GAC ATT CCT GAA ATT CCT GGG TTT AAG GAT TTG GAA CCC ATG GAA CAA TTC ATT GCA CAA GTT GAC CTA TGT GTA GAC TGC ACA ACT GGA TGC CTC < 500
V D I F E I P G F K D L E F M E Q F I A Q V D L C V D L C T T G C L
A CAA CTG TAA GGA CTT TAA GGA CCC AAA TTC CTA AAC CTT GGG TAC CTT GGT AAG TAA CGT GTT CAA CTG GAT ACA CAT CTG ACG TGT TGA CCT ACG GAG
    410     420     430     440     450     460     470     480     490

AAA GGT CTT GCC AAT GTG CAA TGT TCT GAT TTA CTC AAG AAA TGG CTG CCA CAA AGA TGT GCA ACT TTT GCT AGC AAA ATT CAA GGC CAA GTG GAC AAA A < 600
K G L A N V Q C S D L L K K W L P Q R C A T F A S K I Q G Q V D K I
TTT CCA GAA CGG TTA CAC GTF ACA AGA CTA AAT GAG TTC TTT ACC GAC GGT GTT TCT ACA CGT TGA AAA CGA TCG TTT TAA GTT CCG GGT CAC CTG TTT T
    510     520     530     540     550     560     570     580     590

TA AAG GGT GCC GGT GGT GAT gcc aaa gat gaa ctg taa aag ctT GTC GAG AAG TAC TAG AGG ATC ATA ATC AGC CAT ACC ACA TTT GTA GAG GTT TTA CT < 700
K G A G G D A K D E L * K L V E K Y * R I I I S H T T F V E V L L
AT TTC CCA CGG CCA CTA cgg ttt cta ctt gac att ttc gaA CAG CTC TTC ATG ATC TCC TAG TAT TAG TCG GTA TGG TGT AAA CAT CTC CAA AAT GA
    610     620     630     640     650     660     670     680     690
    
```

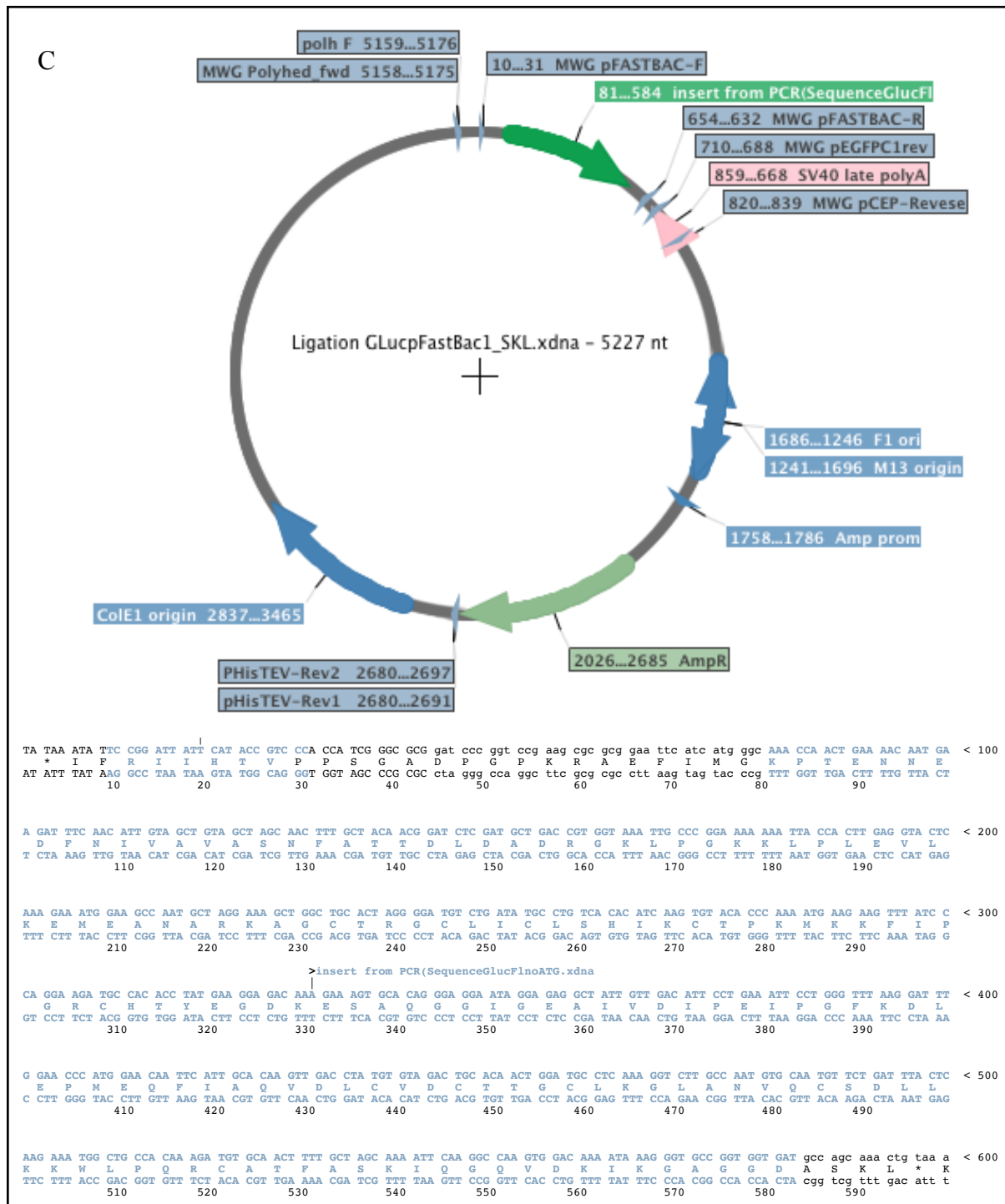


Figure III-7. Schematic plasmid maps and amino acid sequence of pFastBac1 expression vector ligated with the insert PCR product to generate constructs for *in vivo* crystallization

A) Plasmid map of ligation product: pFastBac1-ER with the insert from a PCR GLucFI. B) Plasmid map of ligation product: pFastBac1-ER-KDEL with the insert from a PCR GLucFI. C) Plasmid map of ligation product: pFastBac1-SKL with the insert from a PCR GLucFI. The plasmid maps were constructed using the Serial Cloner 2.61 software.

1.2. Cloning results and generation of luciferase expression constructs

The constructs for heterologous expression of GLuc in *E. coli* cells are based on the commercially available vectors pQE (Qiagen, Hilden, Germany), pET (Merck, Darmstadt, Germany) and pFastBac1 vector (Invitrogen, USA).

The DNA vector used for cloning was digested with specific restriction enzymes (NEB, Frankfurt am Main, Germany) via single or double digestion for each designed construct (Fig. III-9). In a parallel reaction, the dephosphorylation of the linearized plasmid DNA was performed using alkaline phosphatase to minimize re-circularization of the vector and maximize the ligation between the insert and the vector. The dephosphorylated vector DNA was isolated from the remaining reaction mix by preparative agarose gel electrophoresis. The corresponding band was cut from the gel and purified via gel extraction.

The full-length GLuc and the individual domain inserts were amplified by PCR (Fig. III-8A and 8B) using the created constructs pET19bGLucFl and pQE30GLucFl as a DNA template for the PCR reaction. The primers used for the corresponding PCR products are listed in section 1.3, Chapter 2.

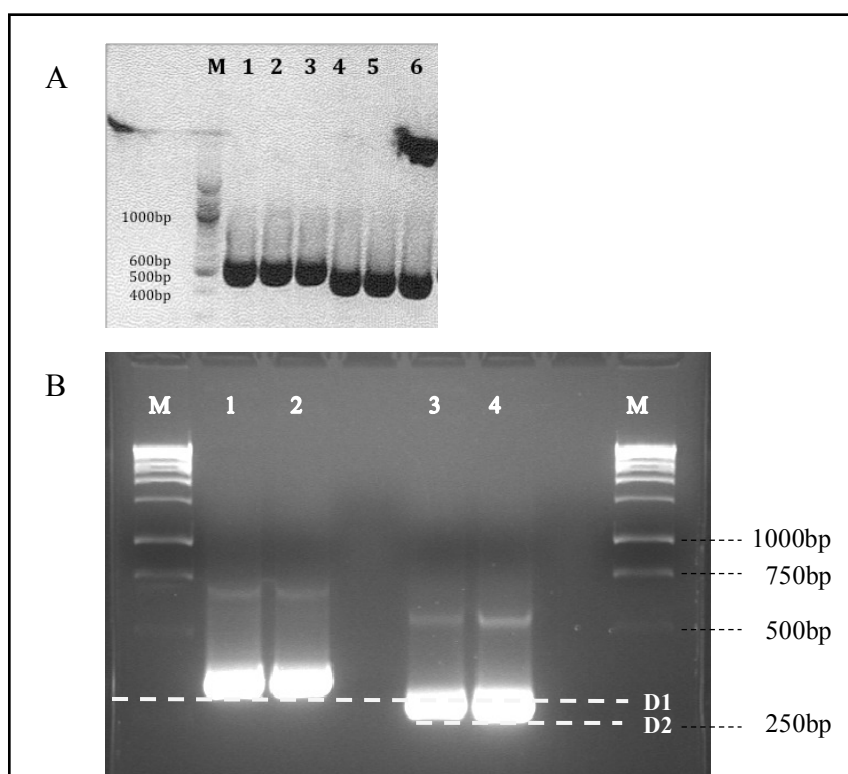


Figure III-8. Agarose gel electrophoresis, analysis of PCR amplified DNA products

Lanes M contain the molecular size (bp) marker. A) Lane 4, 5 and 6 correspond to the PCR products of full-length GLuc with the correspondent primers sequences (Table II-4, Chapter 2). Lanes 1 to 3 do not belong to the experiment (different constructs). B) Lanes 1 and 2 correspond to the PCR products of the DNA fragment GLucD1 using primers as as shown in Table II-4. B) Lanes 3 and 4 correspond to the PCR products of the DNA fragment GLucD2 (primers sequences are shown in Table II-4 section 1.3 Chapter 2).

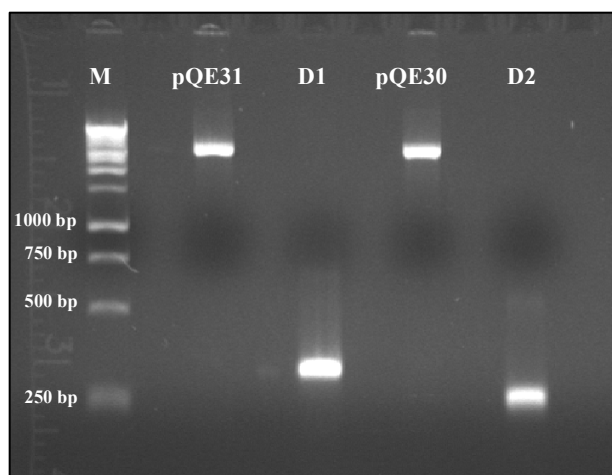


Figure III-9. Agarose gel electrophoresis, analysis of digested vectors and PCR-amplified inserts

Lane M contain the molecular size (bp) marker (GeneRuler 1kb DNA marker; Fermentas). Lanes 1 and 3 correspond to pQE31 and pQE30 vector, respectively, digested with *SacI/KpnI* and *KpnI/HindIII* restriction enzymes. Lane 2 shows GLucD1 digested with *SacI/KpnI* restriction enzymes and lane 4 GLucD2 digested with *KpnI/HindIII* restriction enzymes.

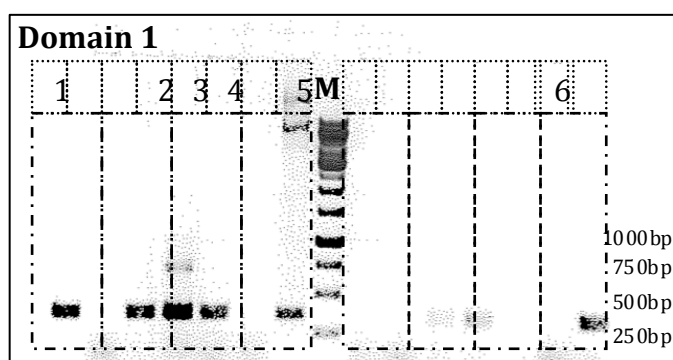


Figure III-10. Agarose gel electrophoresis, analysis of colony PCR from GLuc domain 1

Lane M is the molecular size (bp) marker (GeneRuler 1kb DNA marker Fermentas). Lanes 1, 2, 3, 4, 5 and 6 are positive colonies from pQE31GLucD1 construct (primers sequences are shown in Table II-4 section 1.3 Chapter 2).

PCR products were purified and digested via single or double digest (Fig. III-9). For the ligation reaction a 1:4 ratio of vector to insert was used followed by the incubation at 18 °C overnight. The resulting mixture was introduced via heat shock transformation into the competent XL1-Blue *E. coli* cells. Following the transformation, the resulting bacterial colonies were screened by colony PCR to identify positive clones (Fig. III-10 and 11). Positive clones were amplified and the plasmid DNA was isolated to verify the sequence correctness via sequencing. A restriction digestion control test was performed for the construct pQE31GLucFIThrom digested with *SacI* and *HindIII* (Fig. III-12). (DNA chromatograms and sequencing results are shown in Suppl. Fig. 11-16).

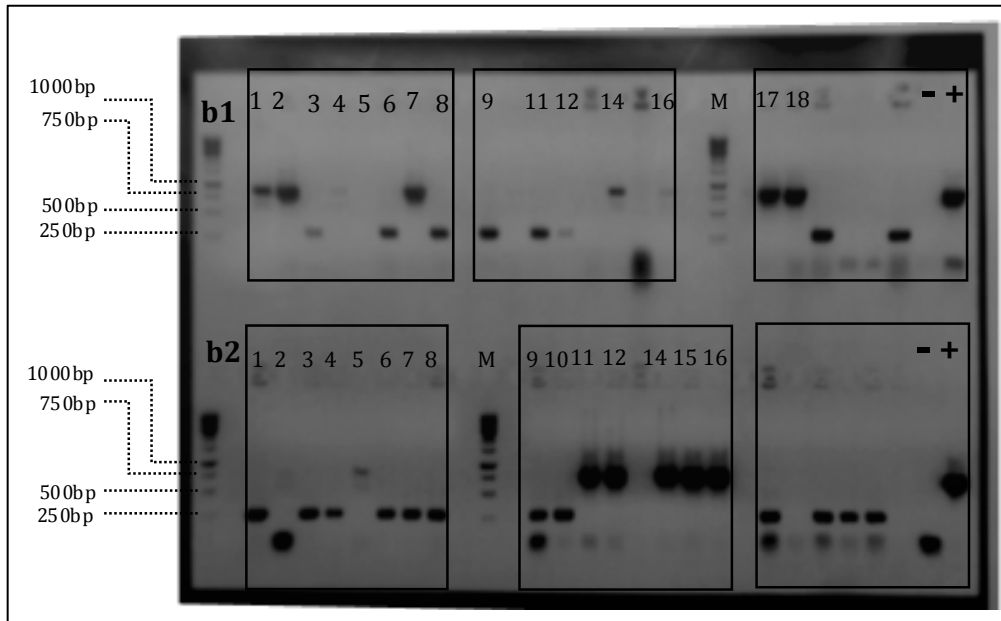


Figure III-11. Agarose gel electrophoresis, analysis of colony PCR of full-length GLuc thrombin construct for expression in *E. coli*

Lanes M contain the molecular size (bp) marker (GeneRuler 1kb DNA marker, Fermentas). Lanes 1,2,7,11,12,14,15,16,17 and 18 are positive colonies from pQE31GLucFIThrom construct using primers sequences as shown in Table II-4, Chapter 2.

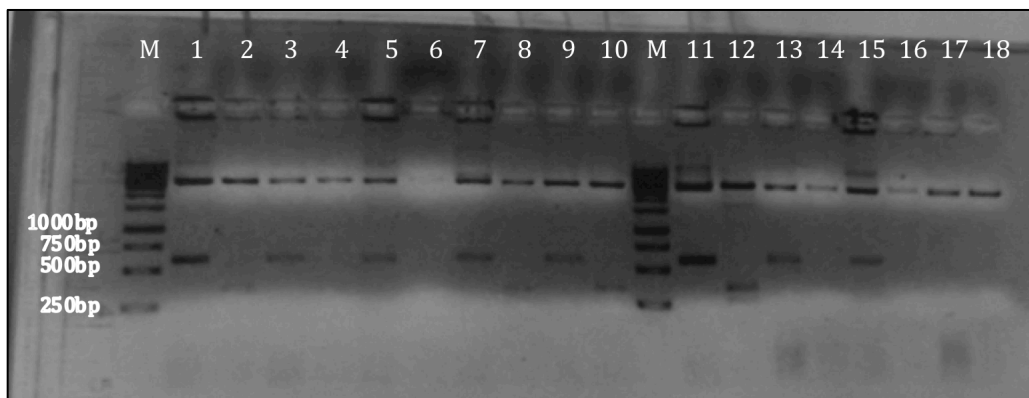


Figure III-12. Agarose gel electrophoresis, double restriction digestion control test

Lanes M contain the molecular size (bp) marker (GeneRuler 1kb DNA marker Fermentas). Lanes 1, 3, 5, 7, 9, 11, 13 and 15 show construct pQE31GLucFIThrom from positive clones digested with *Sacl*/*Hind*III restriction enzymes.

Similar cloning experiments were performed for the pET28a(+)*GLuc*Fl construct. The full-length *GLuc* DNA insert was amplified by PCR (Fig. III-7A). After the PCR reaction, the insert product was purified and digested via a single or double reaction with specific restriction enzymes. The DNA vector pET28a(+) used for cloning was digested with specific restriction enzymes (NEB, Frankfurt am Main, Germany) via single or double digestion for each designed construct (Fig. III-5). The resulting mixture was introduced via heat shock transformation into the competent XL1-Blue *E. coli* cells. Following the transformation, the

resulting bacterial colonies were screened by colony PCR to identify positive clones (Fig. III-13). Positive clones were amplified and the plasmid DNA isolated to validate the sequence via sequencing (chromatogram and sequencing results see Suppl. Fig 7 and 8).

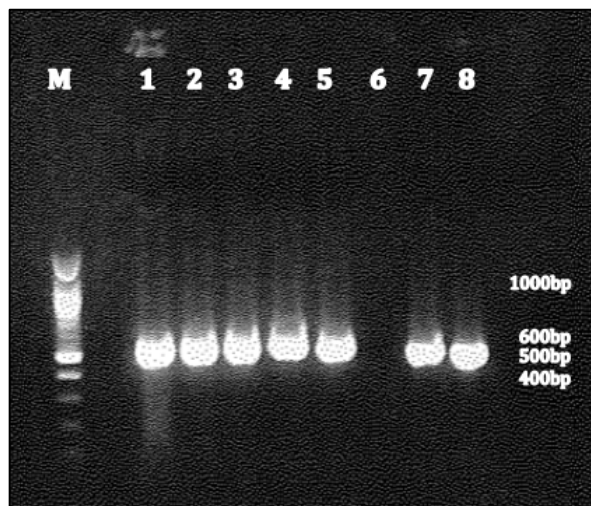


Figure III-13. Agarose gel electrophoresis, analysis of colony PCR of GLuc construct for expression in *E. coli*

Lane M contains the molecular size (bp) marker (New England BioLabs Quick-Load 100bp DNA mark). The bands of the standard reveal the corresponding size estimation of the colony PCR products with their base pair of the standard. Lane 7 and 8 correspond to positive colonies from pET28a(+)GLucFI fused with C-terminal 6xHis-tag construct with the correspondent primers sequences see Table II-4, Chapter 2. Lanes from 1 to 6 and lane 9 correspond to another experiment.

1.3. Cloning and generation of luciferase constructs for *in vivo* crystallization

The gene coding for the full-length GLuc without a signal peptide was amplified by PCR using primers for blunt end cloning, shown in Table II-4, Chapter 2. After the PCR reaction (the programme is shown in Table II-9, Chapter 2) the separation of different sizes of DNA fragments was analyzed by agarose gel electrophoresis. The detection of DNA was performed using a UV/VIS gel documentation system (Gel Doc 2000 documentation system, Bio Rad Laboratories, UK). The amplified and purified PCR products (as shown in Fig. III-14) were digested with *SfoI* enzyme (NEB, Frankfurt am Main, Germany), which allowed for the production of the DNA insert fragments.

The multiple cloning site of the pFastBac1 vector contains a central *SfoI* restriction site. The three pFastBac1 vectors were linearized with the restriction enzyme, *SfoI*, leaving blunt-ends. Then the plasmid was subjected to agarose gel electrophoresis for analysis. The gel slice containing the linearized plasmid was excised and purified with a commercial PCR Clean-up Kit (Machery-Nagel, Düren) following the manufacturer's instructions.

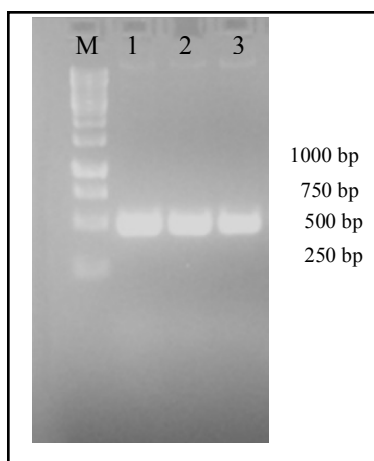


Figure III-14. Agarose gel electrophoresis, analysis of PCR amplified DNA products for *in vivo* crystallization constructs

Lanes M contain the molecular size (bp) marker (GeneRuler 1kb DNA marker, Fermentas). A) Lanes 1, 2 and 3 correspond to the PCR products of full-length GLuc with the appropriate primers sequences (Table II-4, Chapter 2).

The digested fragments were ligated into the *SfoI* site of the linearized pFastBac1 vectors using the T4-DNA Ligase enzyme (Life Technologies, Waltham USA), overnight at room temperature. The ligation solution was transformed into competent *E. coli* cells XL1-Blue (Invitrogen, Carlsbad, CA, USA) by heat shock transformation (section 2.2, Chapter 2). After transformation, the mixture was incubated overnight at 37 °C. Cells were streaked onto fresh LB agar plates containing ampicillin (100 µg/ml) and tetracycline (12 µg/ml).

Colony PCR was performed to identify positive colonies (protocol is given in Table II-10). Results from colony PCR are shown in Fig. III-15.

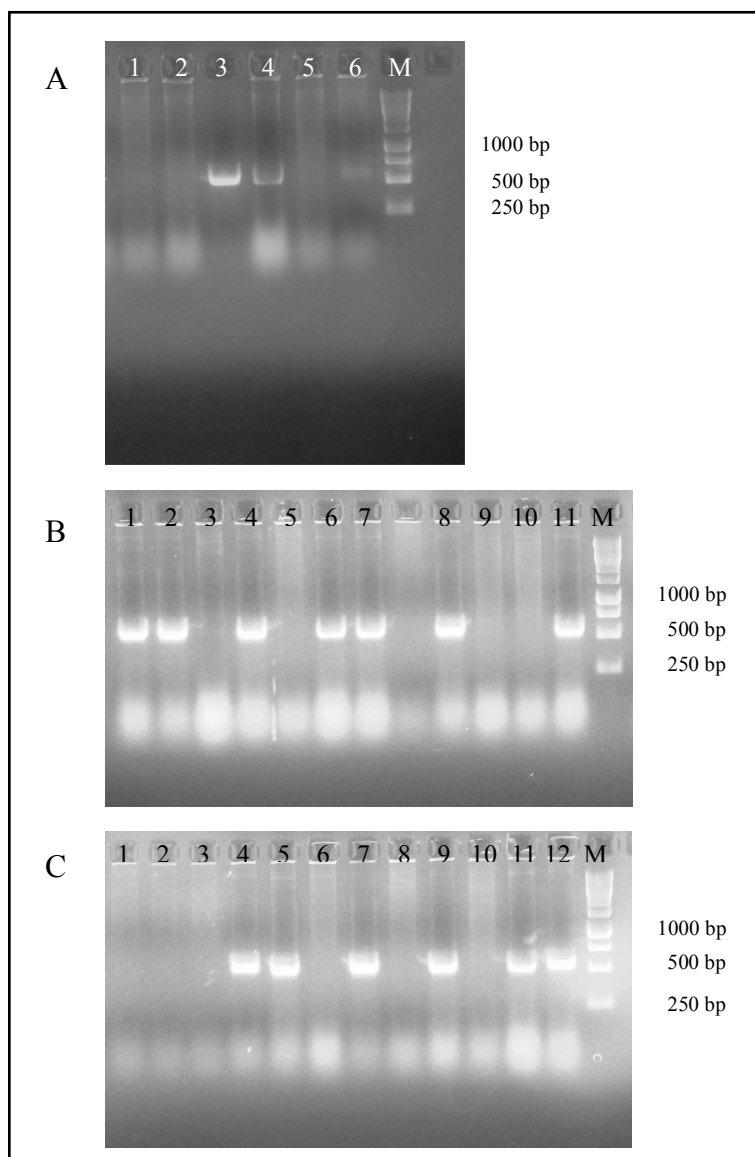


Figure III-15. Agarose gel electrophoresis, analysis of colony PCR of GLuc construct for *in vivo* crystallization

Lane M contain the molecular size (bp) marker (GeneRuler 1kb DNA marker Fermentas). A) Lanes 3 and 4 correspond to positive colonies from pFastBac1-ERGLucFI construct with the correspondent primers sequences (Table II-4, Chapter 2). B) Lanes 1, 2, 4, 6, 7, 8 and 11 correspond to positive colonies from pFastBac1-ER-KDELGLucFI construct with the correspondent primers sequences see Table II-4, Chapter 2. C) Lanes 4, 5, 7, 9, 11 and 12 correspond to positive colonies from pFastBac1-SKLGLucFI construct with the correspondent primers sequences see Table II-4, Chapter 2.

Cells from one positive colony identified by PCR were inoculated into 10 ml LB media with ampicillin (100 $\mu\text{g/ml}$) and tetracycline (12 $\mu\text{g/ml}$). The correctness of the sequence orientation of pFastBac1GLucFL constructs was verified by sequencing (Suppl. Fig. 19 -24).

1.4. Expression and purification of full-length *Gaussia* luciferase and its catalytic domains

Gaussia luciferase was expressed in *E. coli* without its signal peptide using a pQE expression vector system. A construct of full-length GLuc was generated with the thrombin cleavage site between amino acids (97/98) and named pQE31GLucFLThrom (Fig. III-3).

To investigate the catalytic function of the repeat regions corresponding to the two catalytic domains, two constructs (aa 1-97 and aa 98-168) were produced and named pQE31GLucD1 and pQE30GLucD2, respectively. Briefly, the constructs were transformed in a host *E. coli* strain SG10039 (Qiagen, Hilden, Germany). Further steps of protein production are previously described in section 2.8, Chapter 2. Fig. III-16A, III-17A and III-18A show the chromatogram for the GLucFLThrom and the two catalytic domains, respectively. However, the chromatograms of the purification from all three protein constructs reveal irregularities in the peak symmetry when the sample was eluted from the SEC column. This might be because the proteins are polydisperse, meaning that proteins are not in a single state all the time such as monomers or dimers. Thus, if monomers, dimers or higher oligomers of a protein exist side-by-side, the protein will not elute in a single peak from the SEC column due to the different retention times of monomers, dimers and oligomers. In all three cases, the protein eluted as two partially separated peaks detected by absorption spectroscopy at 280 nm. One main larger peak (indicated by a grey bar in Fig. III-16, 17 and 18 respectively) was detected together with a potential shoulder in the chromatograms. According to the retention time, the main peak of all the SEC results corresponds approximately to the molecular weight from full-length GLuc and the two individual domains (based on the reference chromatograms provided by GE Healthcare, Little Chalfont, United Kingdom). Samples were collected from the main peaks and from the shoulders from each construct to be analysed by SDS-PAGE and to confirm the purity of individual fractions (In Fig. III-16B, III-17B and III-18B the peaks are highlighted in grey). The clear bands presented in the SDS-PAGE (Fig. III-16B) correspond to approximately 23 kDa for GLucFLThrom and 15 kDa and 11 kDa theoretical molecular weight (indicated by a red arrow) from the GLucD1 and GLucD2, as shown in Fig. III-17B and III-18B, respectively. Samples collected from the shoulders were discarded because, according to the SDS-PAGE results, impurities and a relative low amount of protein were present compared to the samples from the main peak. In addition, the bioluminescence activity of the fractions from the shoulders is noticeably lower than that of the main peak fractions (Fig. III-15). Based on the results from SDS-PAGE and activity assay, the elution fractions from GLucFLThrom (B₁ to C₃) were collected. Likewise, fractions from B₁ to C₃ of the purified GLucD1 and fractions from B₂ to C₄ of the purified GLucD2 were collected. A sample from each pooled fraction (indicated by the grey bar) obtained from analytical SEC was used for subsequent reverse-phase HPLC analysis and functional characterization of the enzymatic activity.

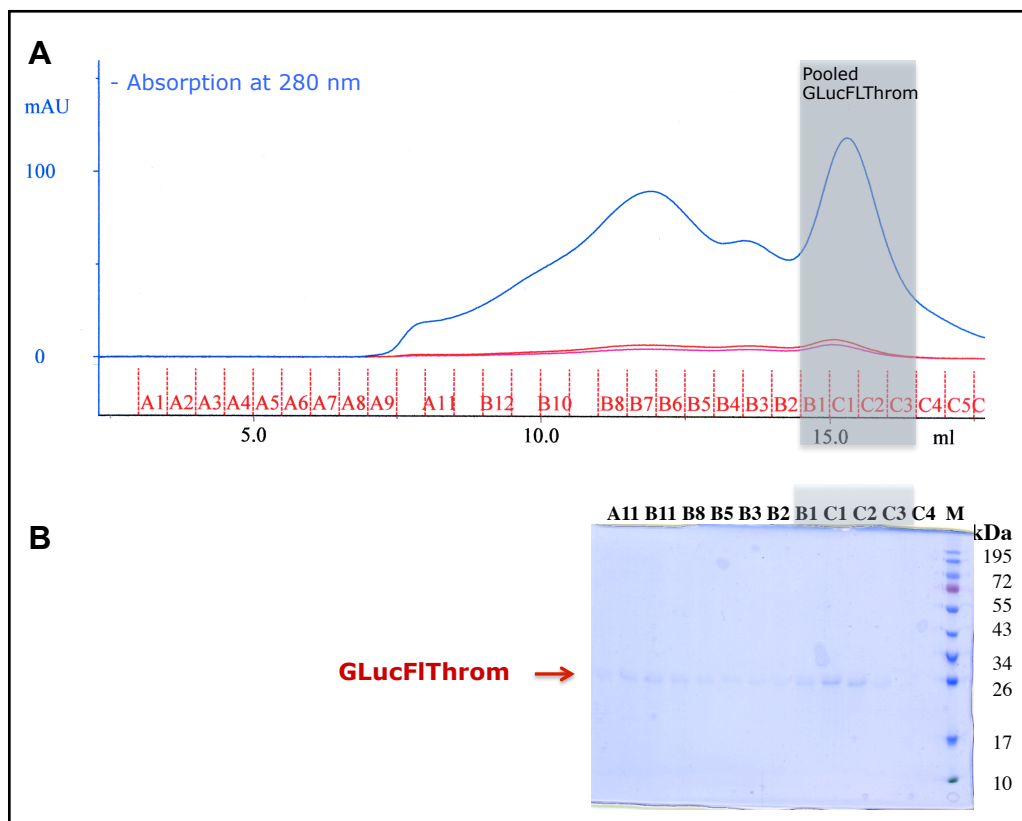


Figure III-16. Size-exclusion chromatography (SEC) and SDS-PAGE analysis of purified GLucFIThrom

A) Chromatogram of the purification of the GLucFIThrom by SEC using an analytical Superdex 200 10/300 column. Denoted in blue are the mili absorbance units (mAU) of UV at $\lambda = 280$ nm. The red horizontal lines (above the x-axis) denote the collection fractions from GLucFIThrom. The protein elutes in two peaks from the column. The corresponding pooled elution fractions (B1 to C3) from the main peak are shown in grey. Fractions were subsequently analysed by SDS-PAGE, HPLC experiments and luminescence activity assays. B) Analysis of the fractions from the SEC purification by 12% SDS-PAGE. The 12% SDS-PAGE was prepared under reducing conditions and the expected molecular weight of GLucFIThrom is indicated by the red arrow. Collected elution fractions were loaded into the gel and the molecular mass numbers on the right correspond to the molecular weight standard in kDa (LMW marker, Fermentas Thermo Fischer Scientific, Waltham, USA).

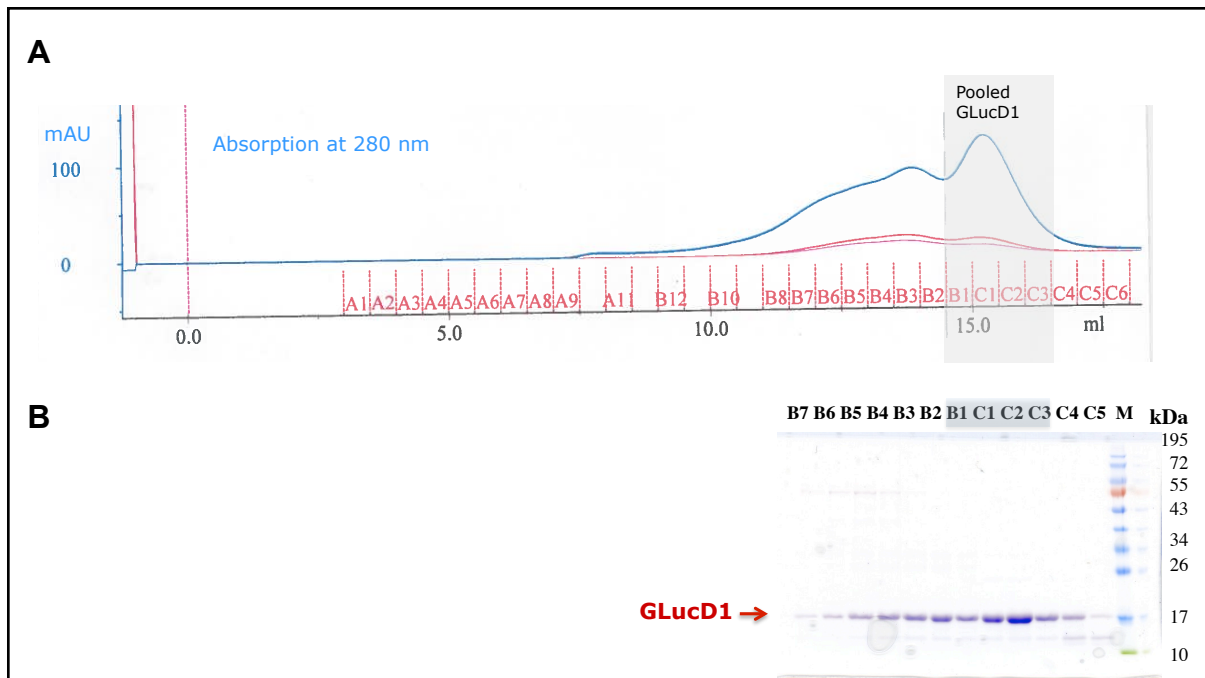


Figure III-17. Size-exclusion chromatogram and SDS-PAGE analysis of purified GLucD1

A) Chromatogram of the purification of GLucD1 by SEC using an analytical Superdex 300 10/300 column. Denoted in blue are the milli absorbance units (mAU) of UV at $\lambda = 280$ nm. The red horizontal lines (above the x-axis) denote the collection fractions from GLucD1. The protein elutes in two peaks from the column. The corresponding pooled elution fractions (B1 to C3) from the main peak are shown in grey. Fractions were subsequently analysed by SDS-PAGE, HPLC experiments and luminescence activity assays. B) Analysis of the fractions from the SEC purification by 12% SDS-PAGE. The 12% SDS-PAGE was prepared under reducing conditions and the expected molecular weight of GLucD1 is indicated by the red arrow. Collected elution fractions were loaded into the gel and the molecular mass numbers on the right correspond to the molecular weight standard in kDa (LMW marker, Fermentas, Thermo Fischer Scientific, Waltham, USA).

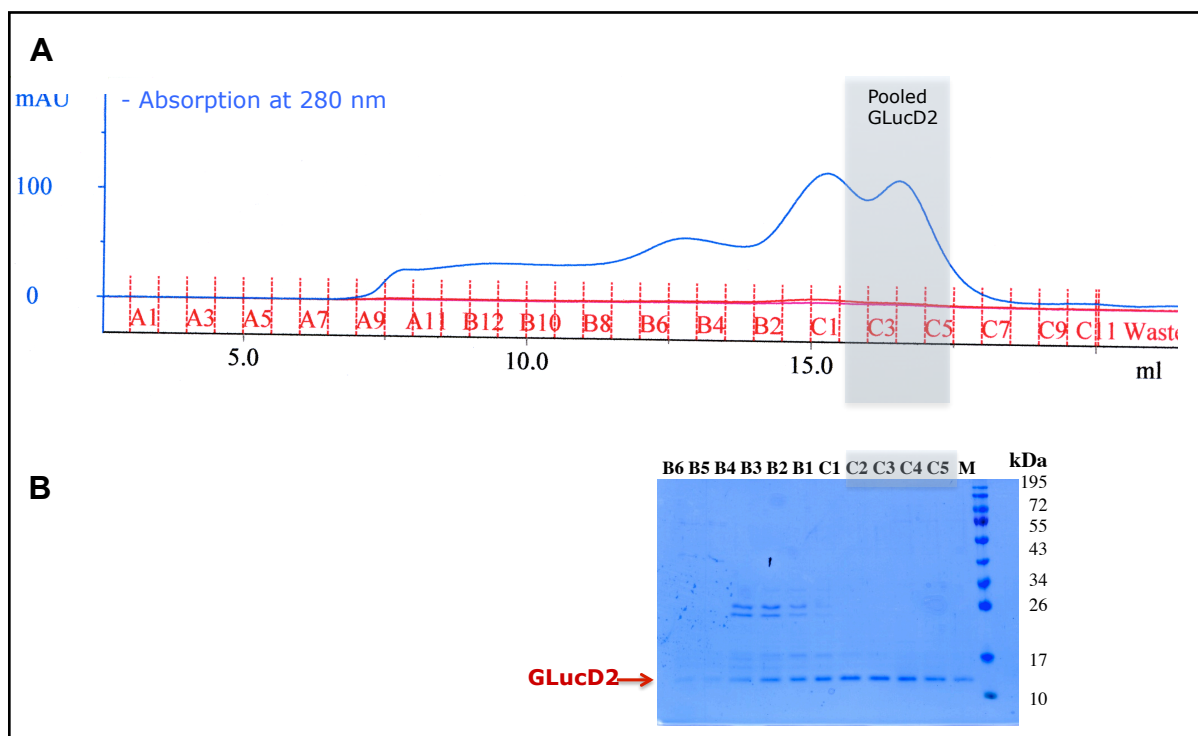


Figure III-18. Size exclusion chromatography and SDS-PAGE analysis of purified GLucD2

A) Chromatogram of the purification of the GLucD2 by SEC using an analytical Superdex 300 10/300 column. Denoted in blue are the milli absorbance units (mAU) of UV at $\lambda = 280$ nm. The red horizontal lines (above the x-axis) denote the collection fractions from GLucD2. The protein elutes in two peaks from the column. The corresponding pooled elution fractions (C2 to C5) are shown in grey. Fractions were subsequently analysed by SDS-PAGE, HPLC experiments and luminescence activity assays. B) Analysis of the fraction from the SEC purification by 12% SDS-PAGE. The 12% SDS-PAGE was prepared under reducing conditions and the expected molecular weight of GLucD2 is indicated by the red arrow. Collected elution fractions were loaded into the gel and the molecular mass numbers on the right correspond to the molecular weight standard in kDa (LMW marker, Fermentas, Thermo Fischer Scientific, Waltham, USA).

Active collected pooled proteins were transferred to fresh buffer 50 mM Tris-HCl (pH8.0) and 300 mM NaCl, then concentrated using centrifugal filter concentrators (MWCO 3000 Da, Merck, Darmstadt, Germany). Protein concentration in samples was measured using Bradford assay as described in section 2.10.3, Chapter 2. Subsequently, 50 μ l aliquots of each construct protein were flash-frozen in liquid nitrogen and stored at -80 °C.

The GLucFIThrom was cleaved into two domains with thrombin protease. The conditions for cleavage were optimized. The protocol optimization was performed by testing different temperatures, times of incubation and concentrations of protease. The most efficient parameters for the procedure were: the digestion with protease concentration of 15 NIH units and overnight incubation at 10 °C. After incubation, the reaction mixture was loaded onto Strep-tag II matrix pre-equilibrated with the protein buffer (for protocol see Table II-12). Subsequently, the individual fractions from the flow-through were analysed by 12% SDS-PAGE. The flow-through fractions after cleavage showed bands in the SDS-PAGE corresponding to GLucD1 and GLucD2 as given in Fig. III-19. The cleavage reaction of the

entire region of GLuc was successful, but it was not possible to separate the individual domains from each other.

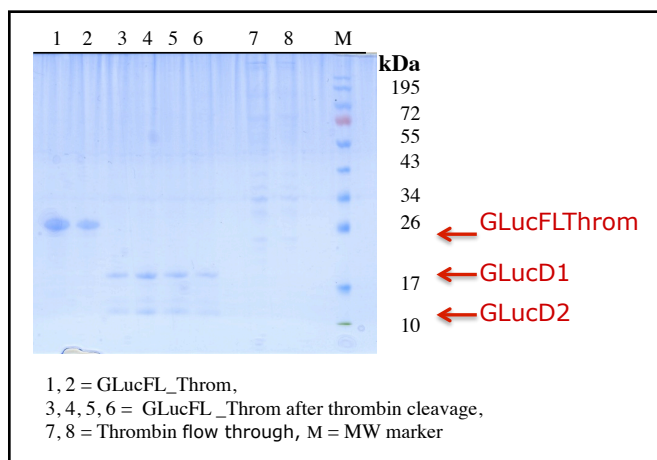


Figure III–19. SDS-PAGE analysis of GLucFITHrom after cleavage with thrombin enzyme

The 12% SDS-PAGE gel under reducing conditions shows the GLucFITHrom before cleavage (lanes 1 and 2), the GLucFITHrom after cleavage with the thrombin enzyme (lanes 3 to 6) and a sample of thrombin enzyme as a control (lanes 7 and 8). The numbers on the right correspond to the standard molecular protein weight in kDa. GLucFITHrom, GLuc1 and GLuc2 fractions indicated by red arrows. (LMW marker, Fermentas Thermo Fischer Scientific, Waltham, USA).

1.5. Analytical reversed-phase high-performance liquid chromatography

Reversed-phase high-performance liquid chromatography (RP-HPLC) involves the separation of molecules on the basis of hydrophobicity. The RP-HPLC for protein analysis consists of an n-alkylsilica-based sorbent from which solutes are eluted with gradients of increasing concentrations of organic solvent such as acetonitrile (CH₃CN), containing an ionic modifier such as trifluoroacetic acid (TFA) (Aguilar, 2004).

The aim of this part of the project was to analyse the purity and homogeneity of the SEC-purified pooled protein using RP-HPLC. It required finding the initial separation conditions and subsequent optimization of the appropriate experimental parameters: the selection of the column, the ionic additive, the organic solvent, the mobile phase flow rate, the gradient time and the operating temperature. This process helped to achieve a high resolution of the sample protein. A rigorous protocol was established to ensure that each sample was processed in the same way. The experimental procedure for the analysis of RP-HPLC is described in detail in section 2.12.1, Chapter 2. All the RP-HPLC experiments were carried out in the same running conditions and on the same day. This technique has the capability of generate a single homogeneous peak related to protein disulphide bonds.

The retention time in reverse-phase chromatography is highly sensitive to protein structure and has been used to distinguish between different disulphide bond patterns in proteins containing multiple cysteines (Weissman and Kim, 1991).

In order to compare a homogeneous sample from GLuc with a non-homogeneous or proteoform sample (Smith *et al.*, 2013) a RP-HPLC experiment was performed. A full-length GLuc construct cloned into pQE31 vector with 6xHis-tag at its N-terminal, Strep-tag at its C-terminal and with HRV3C cleavage site was used as a proteoform sample. The proteoform sample was only used for this experiment, therefore, there is no further description of it. Fig. III-20 shows the analysis of recombinant protein by RP-HPLC of pooled fractions of a proteoform pQE31GLucFIHRV3C sample after nickel affinity chromatography purification, without optimization of the protocol. The denoted blue line indicates the UV-Vis absorption at 214 nm. The graphic represent the elution profile with several overlapping peaks, this suggests a proteoform sample considered that all peaks represent a mixture of protein species. In contrast, Fig. III-21 shows typical RP-HPLC chromatograms of homogenous protein corresponding to the purified samples (GLucFIThrom before and after cleavage and the two catalytic domains respectively) where the large majority of components are elutes in a large single peak within the gradient time. The analytical profiles show the retention times around 11 min for all the samples. These results strongly suggest that all fractions pooled from the purification by analytical SEC are homogeneous and probably feature the correct disulphide bonds. These results are in line with some other RP-HPLC studies of *Gaussia* luciferase (Rathnayaka *et al.*, 2011).

Table III–2. Summary of reversed-phase chromatography parameters

Time	%B	%C	%D	Flow	Max Press
1	20	0.0	0.0	1.000	200
40	80	0.0	0.0	1.000	200

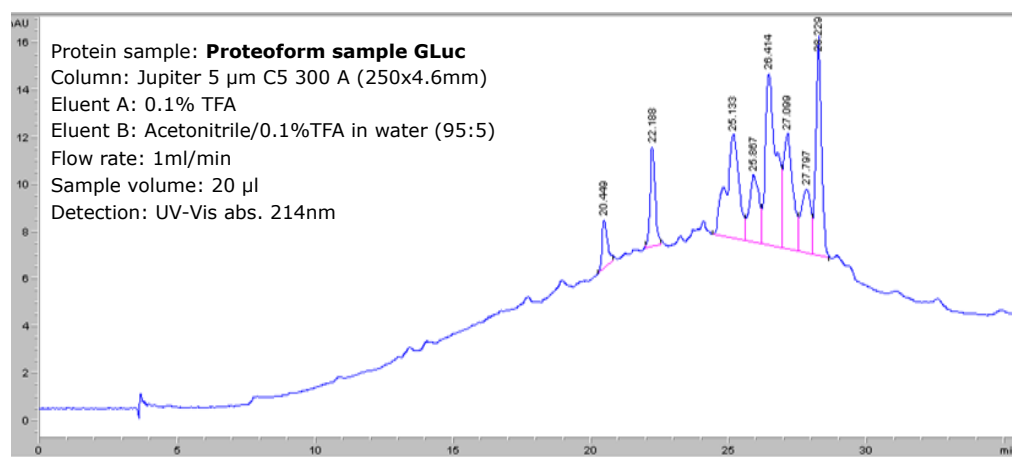


Figure III–20. Representative analytical reversed-phase chromatography of GLuc construct

Analysis of recombinant protein by RP-HPLC of pooled fractions of a proteoform GLuc sample after nickel affinity chromatography purification, without optimization of the protocol. The denoted blue line indicates the UV-Vis absorption at 214 nm. Graphic represent the elution profile of several overlapping peaks of the sample of GLuc using a Jupiter (C5 Phenomenex) 250 mm x 4.6 mm id, 5 µm particle size, 300 A pore size, column. The percentage of the solvent B was increased linearly from 20% to 80%. Conditions: AB linear gradient, where Eluent A is 0,1% TFA and Eluent B is acetonitrile/0.1% TFA in water; flow-rate, 1ml/min over 40 min at room temperature.

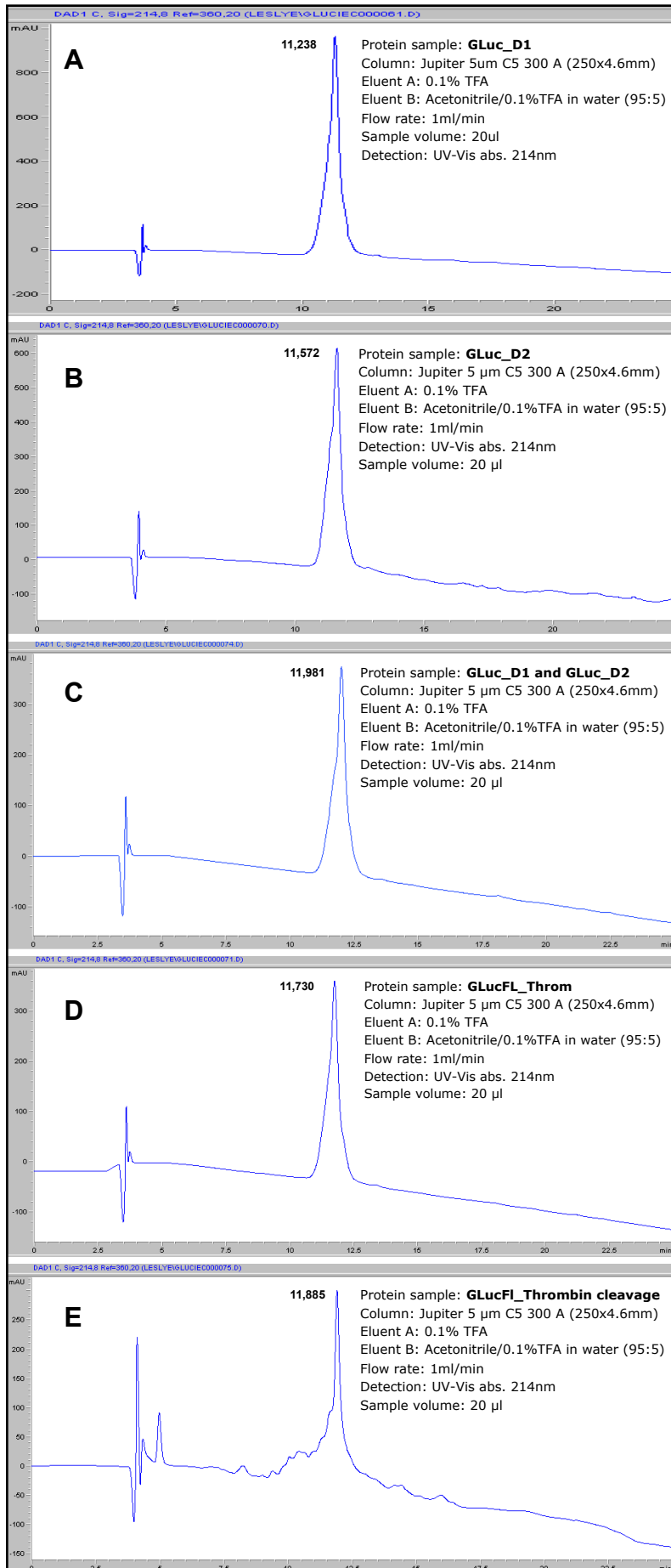


Figure III-21. Representative analytical reversed phase chromatography for different GLuc constructs

RP-HPLC elution profile illustrating the analysis of recombinant GLuc constructs of pooled fractions after size exclusion chromatography purification. The denoted blue line indicates the UV-Vis absorption at 214 nm. There is a large single peak in all constructs from GLuc. Graphics represent the elution profile from GLucFIThrom (A), GLucD1 (B), GLucD2 (C), GLucFIThrom after thrombin cleavage (D), and mixed GLucD1-GLucD2 (E) using a Jupiter (C5 Phenomenex) 250 mm x 4.6 mm id, 5 μm particle size, 300 A pore size, column. The percentage of the solvent B was increased linearly from 20% to 80%. Conditions: AB linear gradient, where Eluent A is 0,1% TFA and Eluent B is acetonitrile/0.1% TFA in water; flow-rate, 1ml/min over 40 min (A) and 30 min (B to E) at room temperature.

1.6. Bioluminescence activity and kinetics

To study the functional basics and properties of the enzymatic activity from *Gaussia princeps*, several experiments were performed using soluble GLuc protein, which forms the correct disulphide bonds as demonstrated by the RP-HPLC results.

The bioluminescence activity was determined by kinetic experiments using GLuc protein purified by SEC. The enzymatic reactions were performed to characterize the bioluminescence properties of GLucF1Throm and confirm the luminescence activity of the two catalytic domains with the CTZ substrate. Also, the enzymatic reactions were used to compare the luminescence intensity between GLucF1Throm and its domains. In addition, light emission kinetics were compared between the GLucF1Throm with other constructs from GLucF1 and *Renilla muelleri* (RmLuc). RmLuc is an extensively investigated luciferase, which also needs CTZ as a substrate to produce light. Purified samples of RmLuc were provided by Dr. Christoph Plieth.

Luminescence activity was measured in the microplate reader Tecan Infinite M200 Pro (Männedorf, Switzerland). The detailed protocol is described in section 2.11.

Various parameters were tested to optimize the light emission kinetics of GLuc constructs. Different experiments were performed and the results are explained below.

1.4.1 Effect of buffers and detergents on bioluminescence activity of GLuc

To determine the effect of different buffers and detergents on light emissions kinetics and to establish the best condition for the luciferase activity assays, each luciferase protein was diluted in three buffers: (1) 100 mM Tris-HCl (pH 8.0), 300 mM NaCl; (2) PBS (pH 7.4) and (3) 100 mM NaPP (pH 8.0) - all individually mixed with 0.1 % Tween or 50 mM Dithiothreitol (DTT) and no detergents, respectively. Diluted GLuc samples in the buffers were mixed with CTZ and kinetic assays were performed. The enzymatic reaction was monitored over the time of 0, 30, 60, 180, 360, 510 min and overnight, as described in section 2.11, Chapter 2.

The resulting light emissions and decays from the samples mixed with different buffers are shown in Fig. III-22. For full-length GLuc, the intensity of the emitted light decreases much more rapidly when the enzyme is diluted in PBS with 0.1% Tween instead of when it is diluted in 100 mM Tris-HCl (pH 8.0) with 0.1 % Tween, thus, the GLuc sample diluted in 100 mM Tris-HCl (pH 8.0) supplemented with 0.1% Tween shows a stable light emission for a longer period of time. Bioluminescence signal was inhibited when the GLuc samples were dissolved in NaPP buffer.

GLuc which was dissolved in buffers Tris and PBS supplied with 0.1% Tween shows high and stable luminescence activity in comparison to the luminescence activity of the GLuc dissolved in buffers Tris and PBS supplemented with 50 mM DTT and dissolved in buffer

without detergent. Hence, DTT reduces the remaining free dissolved oxygen in the buffers see Fig. III-22. More likely is that DTT reduces the luciferase and can inactivate it. The importance of this process (i.e. DTT consumes O_2) is likely, since the data obtained under DTT conditions is more scattered than the data from DTT-free conditions, which points to a fluctuation in dissolved O_2 -concentration. The kinetic reaction with a buffer containing DTT shows a strong decay of bioluminescence activity. This result is similar to the results from Inouye and Sahara (2008). They found that the treatment of GLuc samples with 1 mM DTT and 0.01% BME the luminescence activity of *Gaussia* luciferase was completely lost.

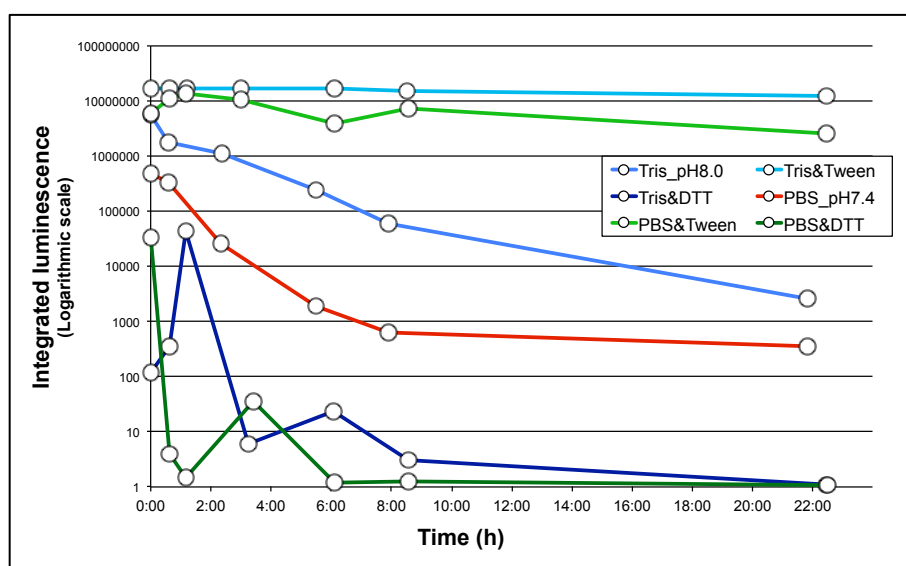


Figure III-22. Effect of buffer composition on the stability of *Gaussia* luciferase as measured by bioluminescence activity

Time course (0, 30, 60, 180, 360, 510 min and O/N) of the bioluminescence intensities by full-length GLuc with different buffers and detergents. The graphic shows the decay of light emission intensity as integrated luminescence of full-length GLuc with 50 mM (Tris-HCl, PBS) pH 8.0 and PBS pH 7,4 buffers supplied with detergents (0.1% Tween and 50 mM DTT, respectively). Sample protein was mixed in Tris-HCl and PBS buffer without detergent as a control. Bioluminescence was initiated by rapid injection of CTZ (50 μ l, 50 μ M in assay buffer) and measured immediately at room temperature. The points on the plots are averages of three independent measurements (n=3).

This result suggests that the Tris-HCl buffer is the best option, which leaves the question of whether Tween is the best detergent for the kinetics. To answer this question, another experiment was set up using several detergents at a concentration of 0.1%, diluted in buffer containing 100 mM Tris-HCl (pH 8.0). The experiment was set up in different time-points, the reaction was measured after 2 h, 6 h, 9 h and three samples O/N. The list of detergents used is presented in the Table II-20. As shown in Fig. III-23, Tween, Brij 35 and Brij 56 produced the best signal to background ratio. Tween causes the lowest kinetic decay and produces a stable luminescence emission over a long period of time, stabilizing the activity overnight.

Tris-HCl buffer with Triton gives low luminescence signal and buffer with SDS causes a massive kinetic decay of the luminescence (falling by a factor of 3 to 6 within the first 2 minutes). Similar results were observed when the buffer was supplied with Silwet, the activity decayed by a factor of 8 within the first 3 h. The buffer supplied with CTAB shows bioluminescence inhibition and higher signal decay. Detergents can perhaps interfere in the activity by denaturing the protein or affecting the interaction between the protein and the substrate. In summary, the results show that the presence of the detergents Tween, Brij 35 and Brij 56 in the buffer assay is a good option to stabilize the light emission in the kinetics experiments. Furthermore, this result is important for the application of the GLuc in bio-imaging.

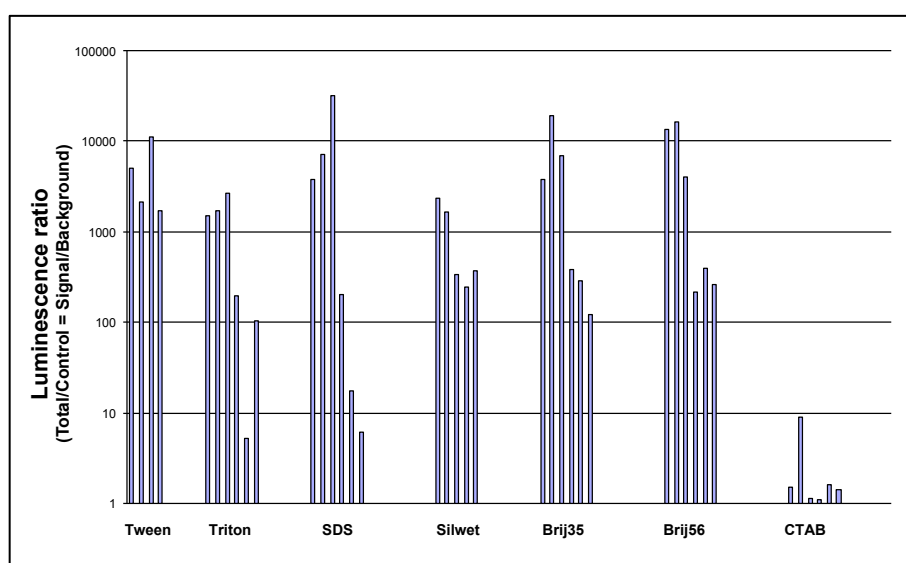


Figure III-23. Effect of detergents on bioluminescence emission of *Gaussia luciferase*

Luminescence ratio signal/background of GLuc with different detergents. Sample protein was mixed in 100 Tris-HCl (pH 8.0) and 0.1% of different detergents measured at different time-points (2 h, 6 h, 9 h, O/N₁, O/N₂ and O/N₃). As shown in the graphic Tween, Brij 35 and Brij 56 produce the best signal background. In contrast, Triton and Silwet makes a luminescence decay. SDS makes a massive decay of luminescence and with CTAB the bioluminescence activity is lost.

In order to elucidate which concentration of Tween gives the highest and most stable luminescence emission, an experiment was set up using the buffer Tris-HCl (pH 8.0) supplemented with different concentrations of Tween (Tween 4%, 2%, 1%, 0,5%, 0,250%, 0,125%, 0,0625% and 0,0312%). The bioluminescence intensity was measured at different time points (0, 1h, 2h, 6h and O/N). The results of this experiment show that the luminescence intensity is Tween-dependent and the maximal luminescence activity and least decay could be achieved with the assay buffer supplemented with around 0.0312% and 0.01255% of Tween (Fig. III-24). These results suggest that the addition of non-ionic detergent (Tween) plays a key role in stabilizing light emission of the GLuc samples. Thus, all subsequent experiments (unless stated otherwise) were carried out in the presence of

0.1 % Tween. Welsh *et al.* (2009) reported that the presence of Tween in the assay buffer extends the light emission of GLuc protein and might prevent its aggregation.

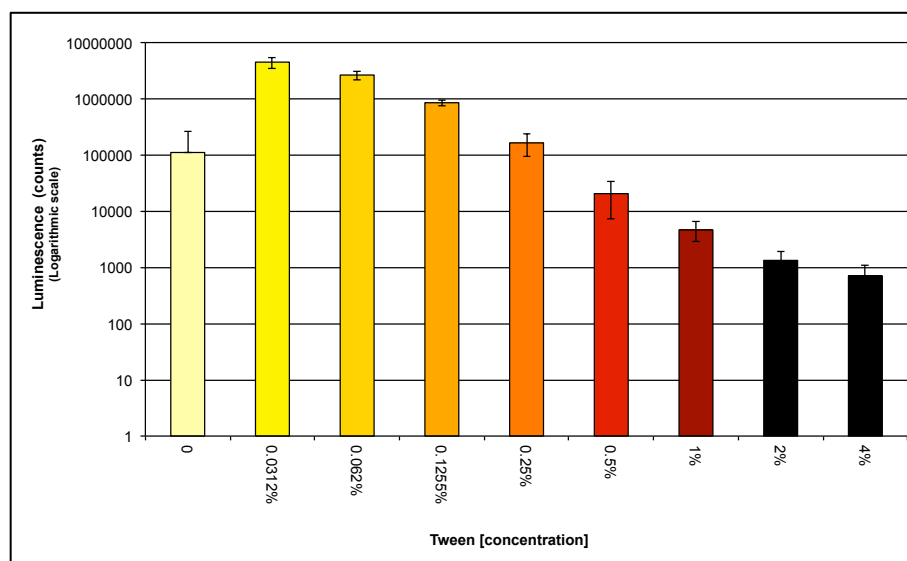


Figure III–24. *Gaussia* luciferase luminescence effect of Tween concentration

The assay buffer was supplemented with different concentrations of Tween, between 0 and 4%. The maximal GLuc luminescence (counts) and least decay of the luminescence intensity is observed at around 0.031% and 0.125%. Averages and standard deviations are calculated over three measurements.

1.4.2 Determination of the bioluminescence activity of GLuc constructs

The next experiment was performed to compare bioluminescence from all purified GLuc constructs. The protocol is described in the section 2.11, Chapter 2. Bioluminescence emission intensity was expressed in luminescence counts. Significant difference was found between the specific activity of full-length GLuc (highly active) and the activity of full-length GLuc with the thrombin. However, there is not too much difference between the specific activity from the construct with the thrombin, before and after cleavage with the thrombin protease. The digested Thrombin-construct still emits ten times more light than the mixture of GLucD1 and GLucD2, and emits just a little less light than the undigested GLucFLThrom construct. Thus, the digestion with Thrombin enzyme seemingly fails to fully separate the two domains (Fig. III-25). Further, Fig. III-25 shows the specific activity of two individual domains (GLucD1 and GLucD2). Thus, both domains have the catalytic function to emit light independently and they have similar light emission kinetics, but less luminescence counts in comparison with the full-length GLuc.

On the other hand, the mixture of the two domains GLucD1 and GLucD2 shows luminescence activity similar to the single domain and the kinetic activity is not increased as expected. The two domains do not work synergistically in producing bioluminescence when they are mixed and thus do not display to the full activity of the full-length construct.

Fig. III-26 shows the relative luminescence for all constructs normalized to the corrected logarithmic luminescence of the full-length GLuc. The luminescence of GLucFITHrom exhibits 53% of the maximum activity; the sample of GLucFITHrom after cleavage shows around 35% of activity and both catalytic domains around 20% and their mixture show 13% of activity. This proves that introduction of the thrombin cleavage site into the full-length GLuc protein significantly diminishes bioluminescence activity.

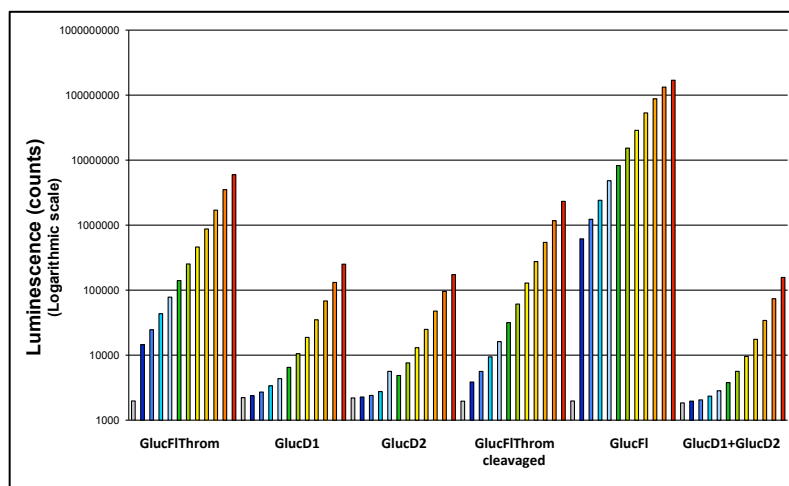


Figure III-25. Bioluminescence-specific activity kinetics of GLuc constructs

Analysis of the luminescence specific activity from GLuc constructs in counts. Protein samples were assayed using a microplate reader right after automatic injection of CTZ into a series of twelve 1:2 dilutions of the protein samples mixed with 50 μ l of 50 μ M CTZ. The different colours of the bars represent the different dilutions of the samples (grey bars indicate buffer control, red and purple bars the lowest and highest dilution respectively). The maximum bioluminescence intensity is observed in samples indicated with red bars.

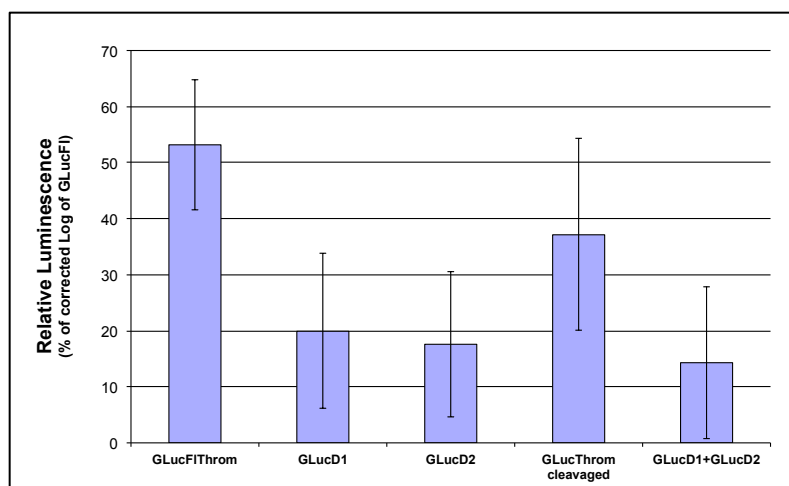


Figure III-26. Relative luminescence activity of soluble GLuc constructs

The diagram shows the relative luminescence intensity units (RLU) of the different GLuc constructs normalised to the average of the corrected logarithmic luminescence of the full-length GLuc. The intensities were determined in triplicate (n=3). Error bars represent the standard deviation.

1.4.3 Determination of the effect of buffer pH

To determine the optimal pH for light emission kinetics, GLuc samples were mixed with buffer 50 mM Tris-HCl, 100 mM NaCl and 0.1 % Tween with various pH values in the range of pH 6.5 to pH 10.5 and the luminescence assay was performed as described previously. The intensity of light emission from luciferase samples appears to be highly pH dependent. Fig. III-27 shows the optimal pH for all constructs, including the single domains, is at around 8.2 ± 0.1 . The bioluminescent activity of samples significantly decreased in basic buffers at $\text{pH} > 8.5$.

The results suggest that the catalytic residues in the oxidation reaction might be similar in the two catalytic domains of *Gaussia* luciferase.

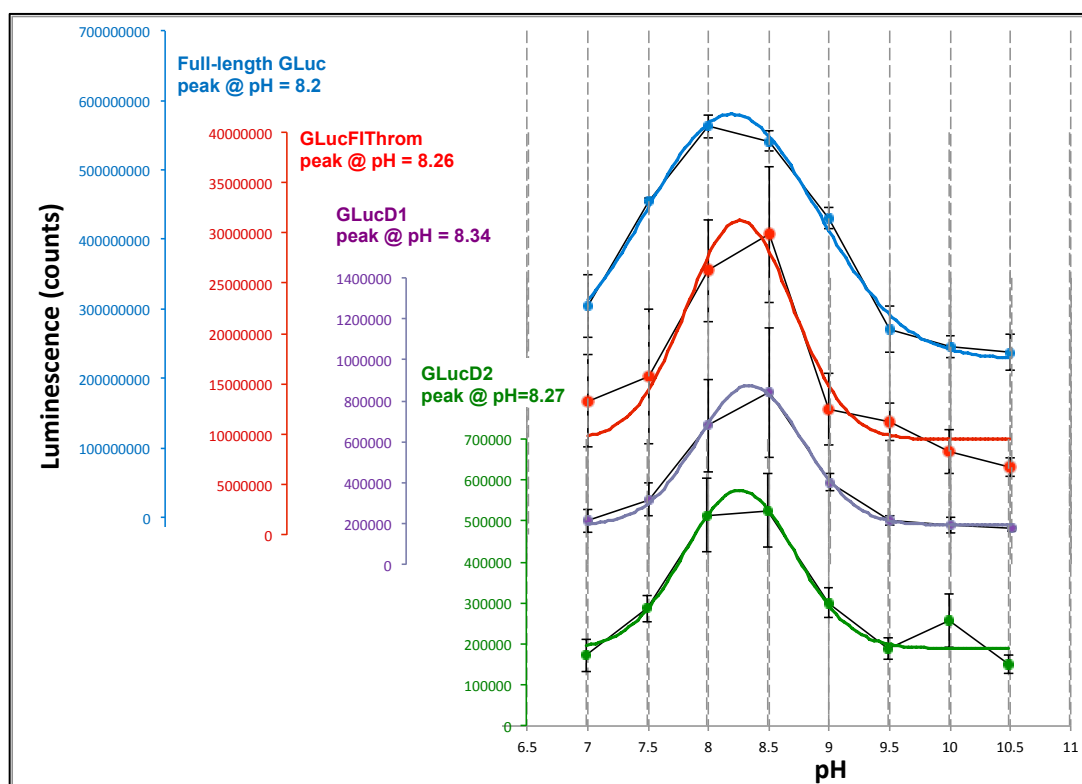


Figure III-27. Effect of buffer pH on luminescence intensity of *Gaussia* luciferase and two catalytic domains

The effect of pH on luminescence intensity was measured in the assay buffer containing 50 mM Tris-HCl, (with different pH ranging from 6.5 to 10.5) supplemented with 100 mM NaCl and 0.1% Tween. The maximum bioluminescence activity was observed at pH around 8.2 ± 0.1 for all the samples: full-length GLuc (blue), GLucFIThrom (red), GLucD1 (purple) and GLucD2 (green). The data represent the means and standard deviations of three repeats.

1.4.4 Effect of NaCl concentration

To elucidate the effect of salt concentration on full-length GLuc, on its individual catalytic domains and on RmLuc bioluminescence activity, samples of each of protein were diluted in

buffer (100 mM Tris-HCl (pH 8.2) and 0.01% Tween) containing various concentrations of NaCl as described in section 2.11, Chapter 2, and then measured for their enzymatic activity.

The highest signal intensity was observed in the presence of 100 mM NaCl for GLuc and at twice higher concentration of 200 mM NaCl for RmLuc (Fig. III-28).

Since the optimum NaCl concentration for maximal bioluminescence stimulation was 100 mM, this concentration was used in buffer composition for further GLuc purification.

As shown in Fig. III-29, bioluminescent emissions of GLucD1 and GLucD2 respond to the NaCl concentration in a similar manner. Although the mechanism of the increased bioluminescence with NaCl is unknown, these results are consistent with the notion that the GLuc is secreted into and functions in a marine environment that contains about 470 mM Na^+ .

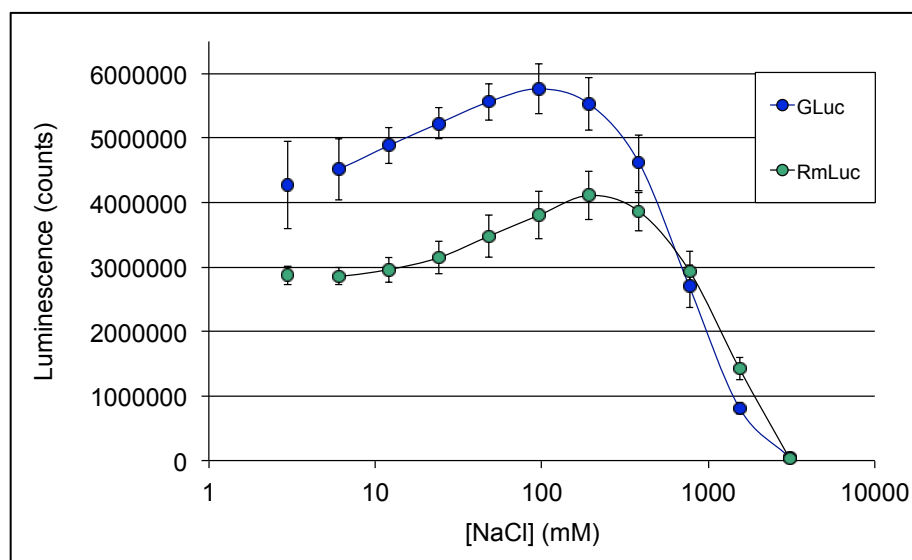


Figure III-28. Effect of salt concentration on bioluminescence emission of *Gaussia princeps* (GLuc) and *Renilla muelleri* (RmLuc) luciferase

Integrated luminescence effect of NaCl concentration on full-length GLuc (blue) and RmLuc (green). Results represent maximum luminescence (counts) versus NaCl concentration. The highest luminescence was observed at 100 mM NaCl for GLuc and 200 mM NaCl for RmLuc. The experiment was performed in buffer 50 mM Tris-HCL (pH 8.2). The points are averages of three measurements (n=3).

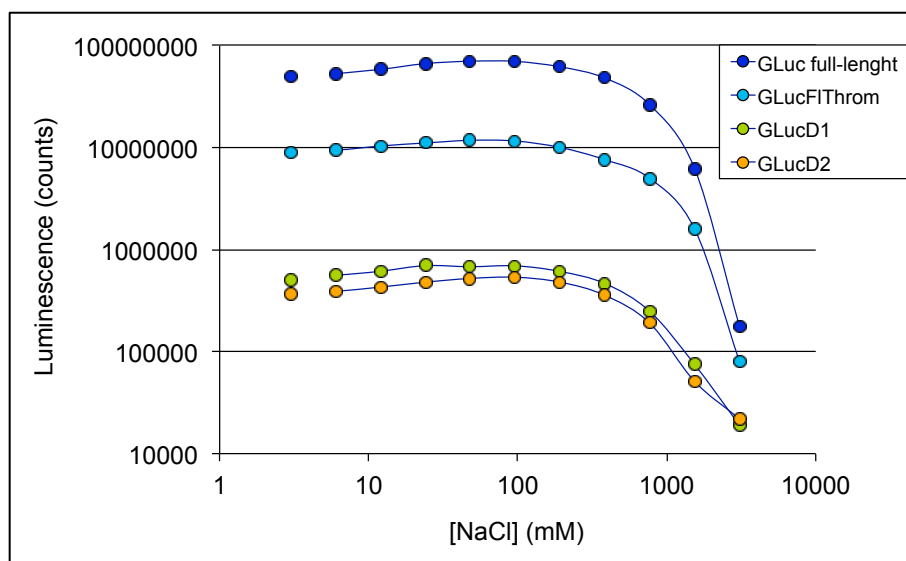


Figure III-29. Effect of salt concentration on full-length GLuc and two catalytic domains

Integrated luminescence effect of NaCl concentration on full-length GLuc (blue), GLucFIThrom (turquoise), GLucD1 (green) and GLucD2 (orange). Results represent maximum luminescence counts versus NaCl concentration. The highest luminescence was observed at 100 mM for all the samples. The salt effect was measured in buffer 50 mM Tris-HCl (pH 8.2).

1.4.5 Effect of different temperatures on bioluminescence activity

The luminescence yield of luciferases increases with the temperature. In contrast, under cold conditions the luminescence yield decreases (Plieth, 2006). To determine the bioluminescence activity at different temperatures, the next experiment was performed using an assay buffer (50 mM Tris-HCl pH 8.2, 100 mM NaCl and 0.1% Tween) containing the GLuc constructs protein samples and an injection of 50 μ l of CTZ. The temperature of the Tecan plate reader was set up at different values (from 15 $^{\circ}$ C to 45 $^{\circ}$ C, with 5 $^{\circ}$ C intervals) as described in section 2.11, Chapter 2. The results presented in Fig. III-30 show the highest signal intensity around 27 $^{\circ}$ C to 28 $^{\circ}$ C for the GLucFl and around 26 $^{\circ}$ C for the individual domains (GLucD1 and GLucD2). From these results, it can be concluded that GLuc luminescence is temperature-dependent and the catalytic properties of the domains GLucD1 and GLucD2 are almost identical to the GLucFl. In contrast, results from RmLuc produced different signals and the kinetics show that the optimum temperature is above 40 $^{\circ}$ C.

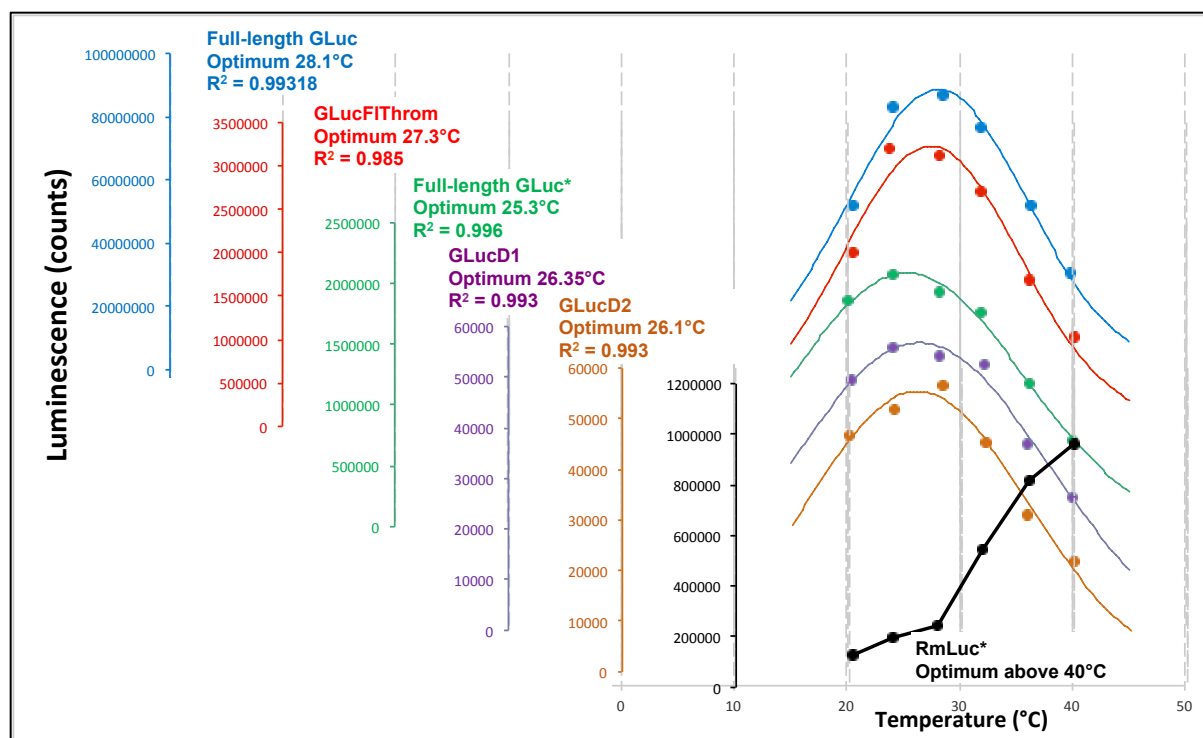


Figure III-30. Effect of temperature on *Gaussia princeps* and *Renilla muelleri* luciferase activity Integrated luminescence effect of different temperatures in the enzymatic reaction on full-length GLuc (blue), GLucFITHrom (red), full-length GLuc* sample purified by Dr. Plieth (green), GLucD1 (purple), GLucD2 (orange) and RmLuc (black). The enzyme activity was measured at different temperatures in the range of 15 °C to 45 °C. Maximum bioluminescence was observed at around 27 °C ± 1 °C for all the GLuc samples. In contrast, the optimum temperature for RmLuc was above 40 °C. The fit Gaussian Fit Model functions are shown by the coloured lines.

1.4.6 Effect of thermostability on bioluminescence activity

The bioluminescence activity of luciferases is sensitive to changes in temperature. The effect of temperature on luciferase activity was compared between *Gaussia* and *Renilla* luciferases. An assessment of the thermal stability of the bioluminescence activity was performed with purified GLuc and RmLuc proteins. Before measuring the bioluminescence, the samples were pre-heated for 2 h in a thermo-cycler (PCR machine) to different temperatures, first from 36 °C to 58 °C, then from 60 °C to 82 °C and finally from 78 °C to 100 °C. The luminescence enzymatic reaction was carried out afterwards at 28 °C using the plate reader. The assay was performed twice with three independent replicates to minimize the deviation error.

In the course of the experiment, the bioluminescence activity of the GLuc protein samples remained detectable and stable even after incubation at 70 °C and only pretreatment with temperatures above 80 °C caused inactivation of GLuc (Fig. III-31). GLuc retained its maximal luminescence activity after 2 h of heat-treatment with temperatures up to 70 °C. Heat half-inhibition of GLuc is at around 88.5 °C (Fig. III-31A). Also, the bioluminescence activity of GLuc is only gradually decreasing with increasing pre-treatment temperatures above 70 °C. In contrast, the bioluminescence of RmLuc decreases more rapidly, suggesting

a half-denatured state already at around 34.4 °C (Fig. III-31B). The results show, that GLucFl has a robust thermostability in comparison to RmLuc samples. This represents an advantage, because GLuc might be suited for novel applications under harsh conditions.

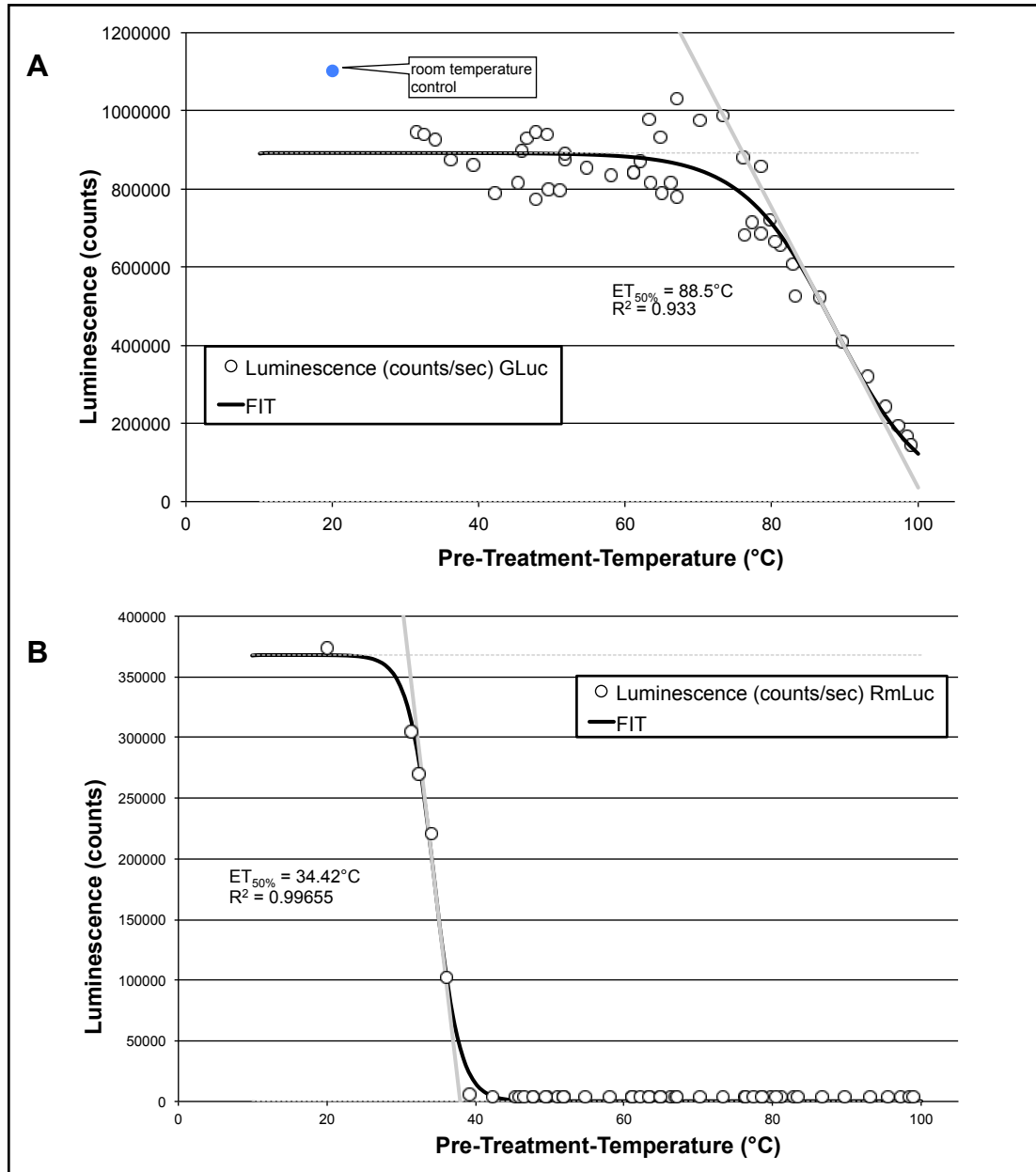


Figure III-31. Thermostability of *Gaussia princeps* and *Renilla muelleri* luciferase
Plot of the luciferase activity in counts versus the pre-treatment-temperature (°C) of GLuc (figure A) and RmLuc (figure B). The curved line indicates a Boltzmann fit. The points on the plots are averages of three measurements (n=3).

1.4.7 Substrate effect between the two catalytic domains

Tzertzinis *et al.* (2012), suggested the existence of a positive cooperative substrate effect between the two catalytic domains in *Gaussia* luciferase. Therefore, to test the possibility of a substrate effect, and perhaps a complex formation between the two domains, a sample of the SEC-purified GLucD1 was mixed with GLucD2 (1:1) and incubated for 1 h 30 min at 4 °C. After incubation, the mixture was loaded onto a pre-equilibrated Strep-Tactin super-flow, to extract the complex formed from the individual domains. Since GLucD1 construct is not StrepII-tagged, this protein will not be bound to the Strep-Tactin super-flow and will elute into flow through. Recovered elution fractions were directly loaded into a 12% SDS-PAGE gel under reducing conditions. Fig. III-32 shows the bands from the SDS-PAGE analysis corresponding to the mixture of the domains. To investigate the hypothesis of a substrate effect between the two domains of GLuc, a bioluminescence assay was performed with the mixture of both domains (1:1). The enzymatic reaction was monitored for 8 h at constant temperature of 28 °C. At the same time, luminescence from a GLucFl sample was measured and used as a control. The result of the bioluminescence measurements reveals no increment in the activity from the mixture of the two domains, as shown in Fig.III-33. This suggests that there is no synergetic effect between the two catalytic domains and there is no a substrate cooperative effect in the bioluminescence with the constructs from GLucD1 and GLucD2 domains presented in this work.

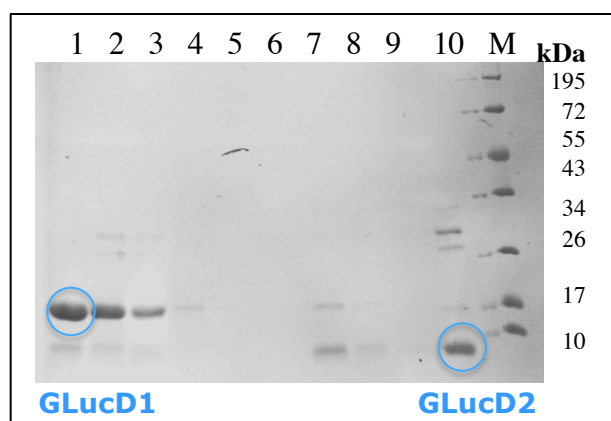


Figure III–32. SDS-PAGE analysis from a mixture of purified catalytic domains

The 12% SDS-PAGE analysis under reducing conditions shows a double band in lane 7 from the mixture of the two catalytic domains after 1 h 30 min incubation at 4 °C, after passing through equilibrated Strep-Tactin super-flow. Lane 1 shows the GLucD1 purified by SEC, lane 2 the flow-through from Strep-Tactin super-flow from the mixture (GLucD1 and GLucD2). Lanes 3 to 6: wash steps flow-through, lanes 7 to 9: elution from the column, and lane 10 GLucD2 purified by SEC. M: molecular weight marker in kDa (Fermentas, Thermo Fischer Scientific, Waltham, USA).

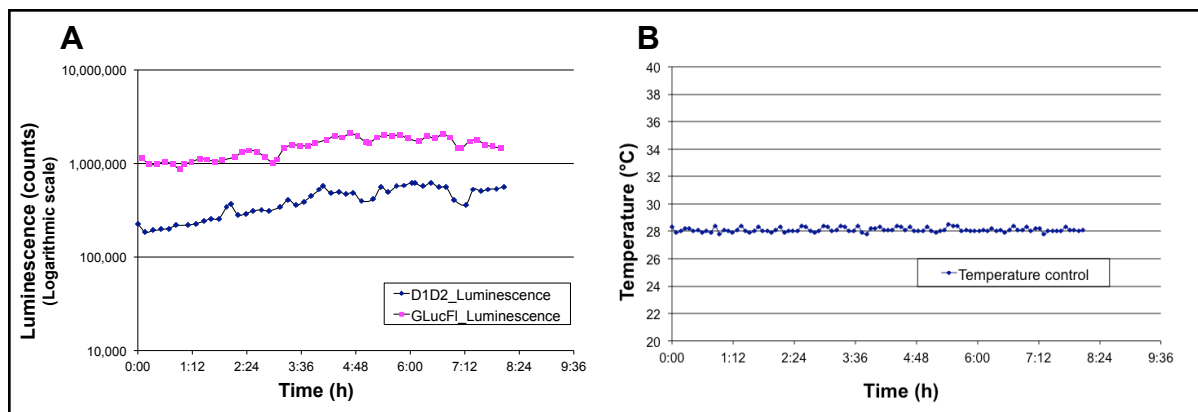


Figure III-33. Luminescence emission of a mixture of catalytic domains GLucD1 and GLucD2
A) Bioluminescence intensity responses of a mixture of GLucD1 and GLucD2 (1:1) (blue) and GLucFl (pink) as a control. There is no increase of the bioluminescence activity in the mixture of the two catalytic domains. B) Temperature recording for 8 h of the experiment (minimal fluctuations around 28 °C).

2. Expression and purification of full-length *Gaussia* luciferase construct

Previous studies have shown that the expression of GLuc in *E. coli* directs the freshly produced protein to inclusion bodies, and therefore a refolding step is required to recover active protein (Inouye and Sahara, 2008). Further, it is crucial to determine the optimal fusion tag, expression condition, suitable *E. coli* strain, and best purification protocols. This is to improve the quantity, solubility and stability of the luciferase protein containing ten cysteine residues, because the enzymatic activity strongly depends on a correct formation of intramolecular disulphide bonds. However, several constructs for GLuc expression were designed and tested in order to obtain a highly active, stable and pure protein, appropriate for crystallization and structural analysis. As a result, the most soluble protein construct was the recombinant pET28(a)+GLucFl construct expressed by *E. coli* strain BL21 (DE3) (Stratagene, San Diego, CA USA) (Fig. III-34).

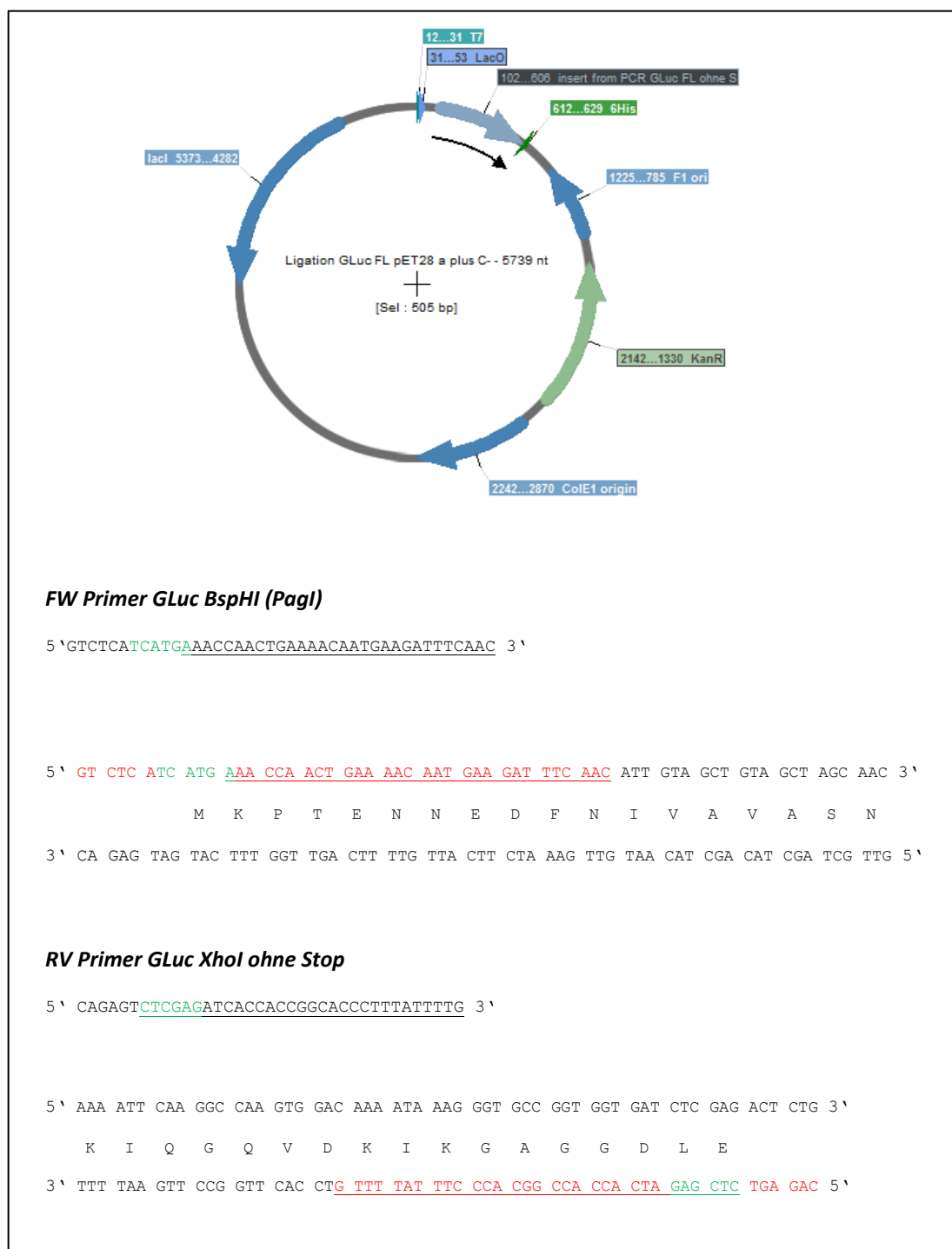


Figure III–34. Schematic plasmid map of recombinant pET28(a)+GLucFI construct for expression in *E. coli* cells

Plasmid map of recombinant pET28(a)+GLucFI construct with 6x-His-tag at the C-terminus for expression in *E. coli*. Black arrow shows the insert PCR from the full-length GLuc. The plasmid map was constructed using the Serial Cloner software 2.61.

Full-length GLuc fused with a C-terminal 6xHis-tag affinity sequence was successfully expressed *E. coli* cells. The protein production was performed according to the protocol described in section 2.8.1, Chapter 2.

Fig. III-35A shows the chromatogram of the anion ion exchange chromatography of GLucFl protein and the gradient elution conditions purification. The sample was applied to the column via injection, the injection mark denoted at 0 CV (pink vertical line in the chromatogram). After injection, the elution gradient started after 20 CV with 0 M salt buffer and then salt concentration was gradually increased (green line). A sharp peak started to elute at approximately 50 mM salt concentration denoted in the elution fractions from the chromatogram with grey mark (A₂, A₃, A₄). A small shoulder appears at 300 mM salt concentration denoted in the elution fractions from the chromatogram from B₁₁ to B₅. This shoulder corresponds to some impurities which are also observed in the SDS-PAGE gel (Fig. III-35B). A second peak appears at 1 M salt concentration (this peak corresponds to some impurities, remains of very tightly bound molecules attached to the column). The chromatography run was carried out at a flow rate of 1 ml/min and nine fractions were collected within 60 min. Protein elution was monitored during all the process (60 min) and the relative absorption value was recorded at 280 nm (blue line in the chromatogram). The corresponding anion ion-exchange experiment shows that the majority of the protein can be detected in the elution from the column (Fig. III-35A). Analysis of the collected fractions after purification through anion ion-exchange was performed using 12% SDS-PAGE (Fig. III-35B). The elute from the main peak (fractions highlighted in grey) shows a strong single band at approximately 20 kDa in the SDS-PAGE, which corresponds to the theoretical molecular weight of 19.36 kDa. There are no other bands visible above or under the strong band, indicating high purity of the sample. In summary, these results show that the overall purification strategy was successful, yielding highly pure protein. Fractions highlighted in grey (A₂, A₃, A₄) were pooled together and concentrated in a centrifugal concentrator (MWCO 3000 Da) up to 25 mg/ml in 50 mM Tris-HCl (pH 8.0) and 50 mM NaCl buffer. Aliquots of 50 µl were frozen in liquid nitrogen and stored at -80 °C. Later, these aliquots were used for further characterization and for crystallization trials.

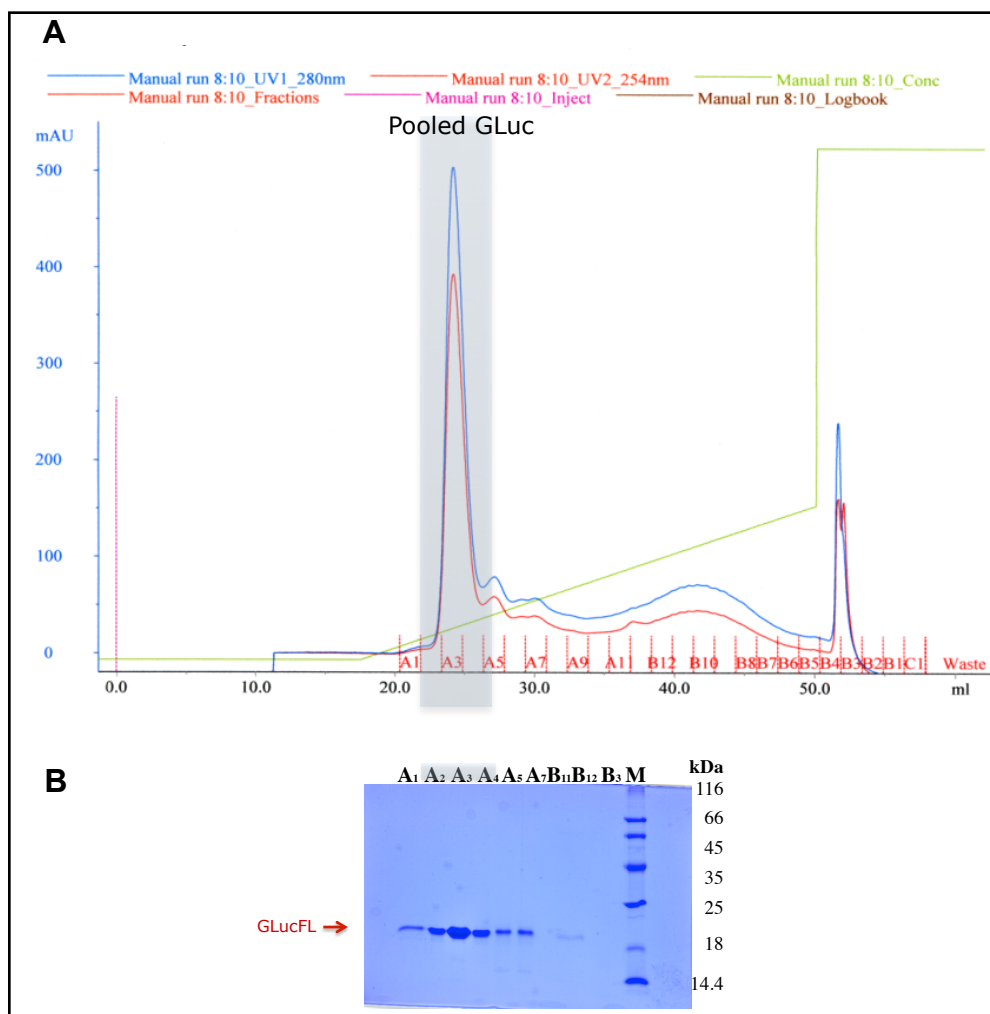


Figure III-35. Ion exchange purification chromatogram and SDS-PAGE analysis of purified GLucFI

A) Chromatogram of the purification of GLucFI construct by anion ion exchange chromatography column. Denoted in pink the injection of the sample, in blue the milli absorbance units (mAU) of UV at $\lambda = 280$ nm, the red line corresponds to the mAU of UV at $\lambda = 254$ nm. GLucFI elutes in a large single peak from the column. The green line shows the salt concentration from the gradient elution, a small shoulder in the chromatogram elutes later and a second peak elutes with a high concentration of salt. The grey highlight indicates the pooled fractions. B) The pooled fractions (A₂, A₃, A₄) were used for further assays and used for analysis of the molecular mass by 12% SDS-PAGE gel. M indicates the molecular weight marker in kDa.

3. Biophysical methods for protein characterization

High purity and homogeneity of the sample are crucial for crystallization to be successful and to achieve a proper characterization of GLuc. Biophysical methods such as dynamic light scattering (DLS) and nuclear magnetic resonance (NMR) were used to characterize the presence of aggregates or the polydispersity state, stability and whether or not the protein was correctly folded.

3.1. Dynamic Light Scattering to characterize protein stability

Dynamic light scattering (DLS) analysis is a technique to evaluate the hydrodynamic diameter via the Stokes-Einstein equation, which defines a relationship between the diameter of a particle and the diffusion coefficient due to Brownian motion in solution (Shiba *et al.*, 2010).

Small particles diffuse faster than large ones. This effect is exploited with DLS. DLS experiments give information about size distribution of light scattering particles. The position of the distribution peak gives information about the mean particle size and the width of the peak correlates with the degree of polydispersity in the sample.

Consequently, a distinction of different particle species can only be obtained with low polydispersity and when the hydrodynamic radius of the particles in solution differ by more than 50%.

To determine the hydrodynamic radius of the protein molecules in solution, as well as the homogeneity or monodispersity of the purified protein, a DLS experiment was performed following the protocol described in section 2.8.2, Chapter 2.

Elution fractions of GLucFl construct from the anion ion-exchange purification were pooled and concentrated from 1 mg/ml to achieve 11 mg/ml. Size distributions were measured with a Zetasizer NanoZS (Malvern Instruments, UK) based on DLS. After the measurement, the GLuc sample was recovered for crystallization screening and other biophysical experiments.

The results from the DLS measurements of the GLucFl protein sample are shown in Fig. III-36. A peak with a mode 5.6 nm for the diameter is observed, accounting for 91.7% of the total measured signal intensity (Fig. III-36A). The same peak with a mode 4.8 nm for the diameter is observed, accounting for 100% of the total measured mass distribution by volume (Fig. III-36B). Based on the intensity and mass distribution data the sample was classified as polydisperse by the Malvern DLS analysis software with a relative polydispersity of 35.7% by intensity and 28.4% by volume. The polydispersity correlated with the average width of the peak provides information about the molecular weight. The software estimated the molecular weight of the sample at 47.2 kDa by intensity distribution and at 28 kDa by mass distribution. The molecular weight values are only an approximation. The purity of the sample was proved by the DLS results because no other species were detected.

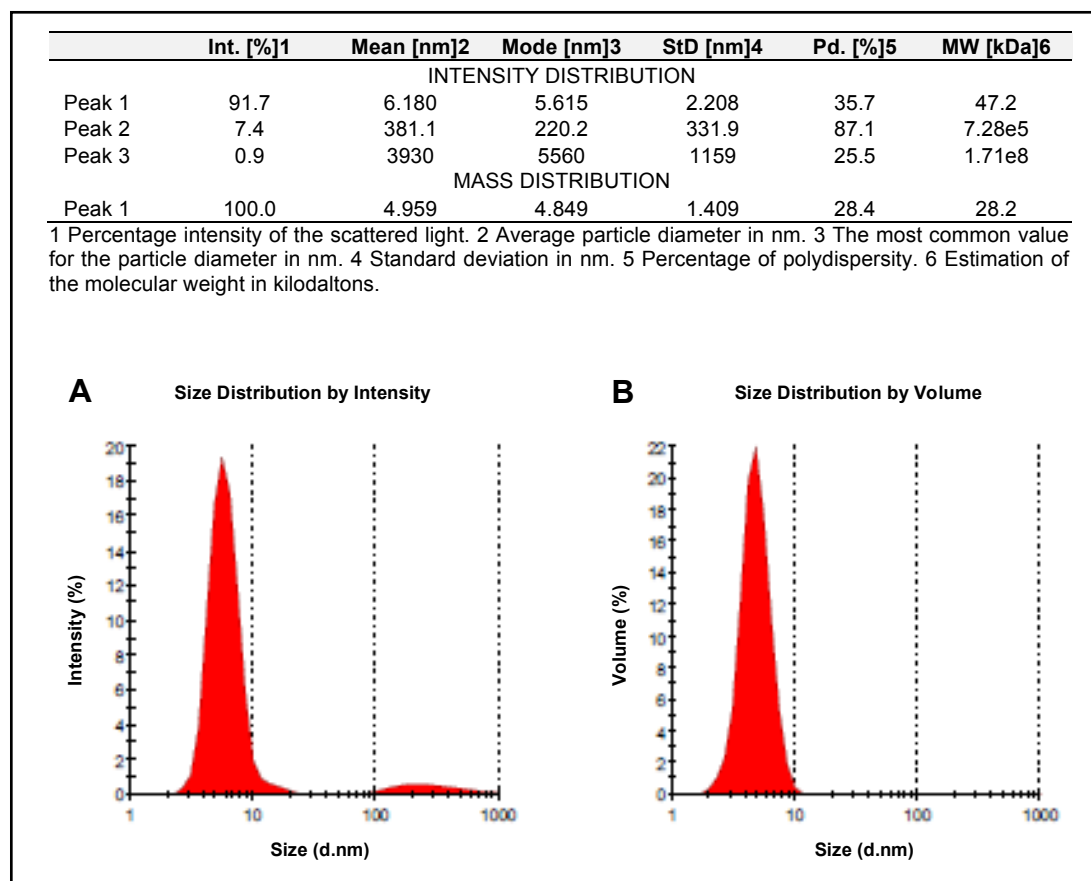


Figure III–36. Dynamic light scattering (DLS) experiment of GLucFI

Representation of the size distribution by intensity and volume. A) Size distribution corresponding to the intensity of the scattered light. The diameter of the molecules in nm was plotted against the intensity of scattered light as a percentage. B) Size distribution of the molecules in accordance with the volume. The particle diameter in nm was plotted versus the volumetric proportion of the species in percentage.

3.2. Nuclear magnetic resonance spectroscopy

Nuclear magnetic resonance (NMR) is a tool for studying the structure, function and dynamics of biological macromolecules. With NMR spectroscopy experiments, using hydrogen-deuterium exchange methodologies, it is possible to analyse the structure of intermediates of protein folding and solubility. The one-dimensional (1D) ^1H -NMR spectrum is the most common experiment used in NMR, which shows signals for each of the hydrogen atoms in a biomolecule.

The purified proteins (GLuc1D and GLucFI construct), concentrated to 1.5 mg/ml and 3 mg/ml, respectively, were analysed by 1D ^1H -NMR spectroscopy to verify the folded state and solubility of the sample.

The 1D ^1H -NMR spectra results allowed the almost complete assignment of backbone amide resonances for both proteins (GLucD1 and GLucFI). The dispersion of the 1D ^1H NMR

signals in the regions of the methyl protons (-0.5 to 1.5 ppm), alpha-protons (3.5 to 6 ppm) and amide protons (6 to 10 ppm) provides the main indicators of folded protein.

NMR spectroscopy is a useful tool of structural genomics, because of its unique ability to provide rapid information about the folded state of the protein structure before the crystallization screening, particularly in situations where proteins are multimeric or form nonspecific aggregations (Page *et al.*, 2004). 1D ^1H -NMR spectrum provides a “fingerprint” of the protein, which informs on the suitability of a sample for structural studies and also the folding state of the protein sample.

The result of the ^1H -NMR spectrum for GLucD1 shows clear evidence of chemical shift dispersion especially in the region from 0 to 0.5 ppm (Fig. III-37A). The displayed line broadening over the spectra is very similar to the random coil values, which is caused by aggregation or oligomerization state. This is a clear indication of the presence of unfolded flexible residues. It is probably because it is a short peptide domain and perhaps the region where it was cut to make this construct was not well defined. Another reason could be that the sequence from GLucFl protein contains 10 cysteine residues that need to form native disulfide bonds.

In contrast, the spectrum of full-length GLuc show a dispersion of resonance lines in the spectral regions (Fig. III-37B). These resonances are well separated, sharp lines, and the spectrum shows resonance also in the regions 0 to 2.0 ppm and 6.5 to 9.5 ppm. The sharp intense peaks near 3.5 ppm come from 10 or 20 mM Tris buffer. With these results, it is possible to conclude that different concentrations of Tris buffer have no critical influence on the spectrum. This result is in accordance with Baryshnikova *et al.* (2008), who shows a non-invasive technique for determining pH in biomolecular NMR samples using buffer components like Tris, whose chemical shifts are sensitive to pH in a range from 2.5 to 9.8. The 1D ^1H -NMR chemical shift for Tris pH 6.7 was 3.725 and pH 9.8 was 3.516 ppm (Fig. III-38).

GLuc sample tested with a buffer with lower pH (6.4) suggest that the protein is very stable and structured. These results clearly identify that the spectrum from GLucFl sample protein is likely folded and will require further analysis and additional optimization.

NMR techniques can give information about specific interactions between residues or partial structures during protein folding or unfolding that is not possible to obtain using other methods. The information obtained from the use of pulsed hydrogen-deuterium exchange to examine the formation of hydrogen bonds during folding may be crucial to solve the folding problem.

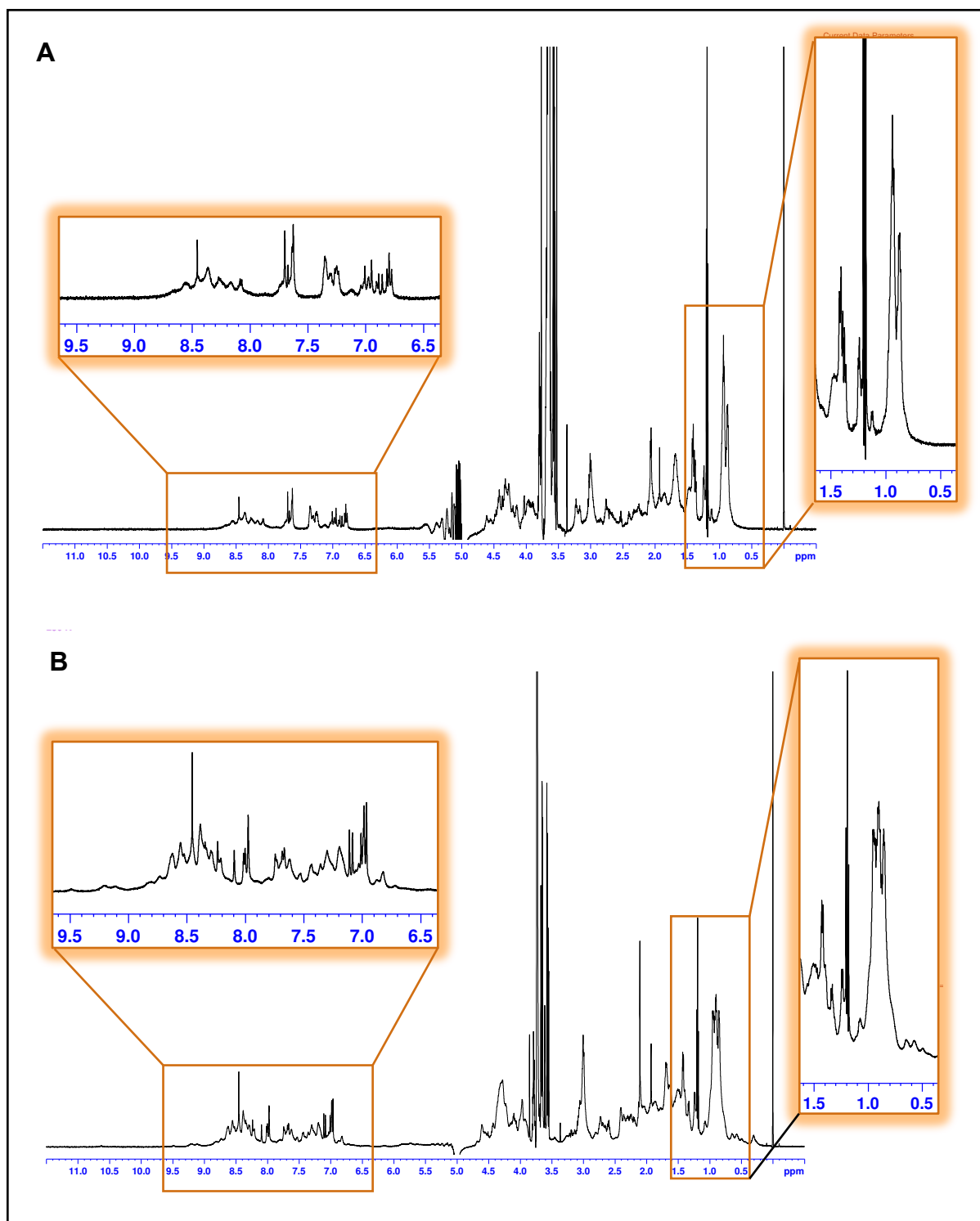


Figure III-37. Representative 1D ^1H -NMR spectra from GLucFI and GLucD1 proteins

1D ^1H -NMR spectrum of the recombinant GLuc constructs. A) 1D ^1H -NMR spectrum of the GLucD1 sample. The spectrum was recorded in 10 % D_2O at 280 K in buffer 20 mM Tris-HCl (pH 7.8). B) 1D ^1H -NMR spectrum from the full-length GLuc construct sample. The spectrum was recorded in 10 % D_2O at 280 K in buffer 10 mM Tris-HCl pH 6.4. The sharp intense peaks near 3.5 ppm come from 10 and 20 mM Tris-HCl buffer.

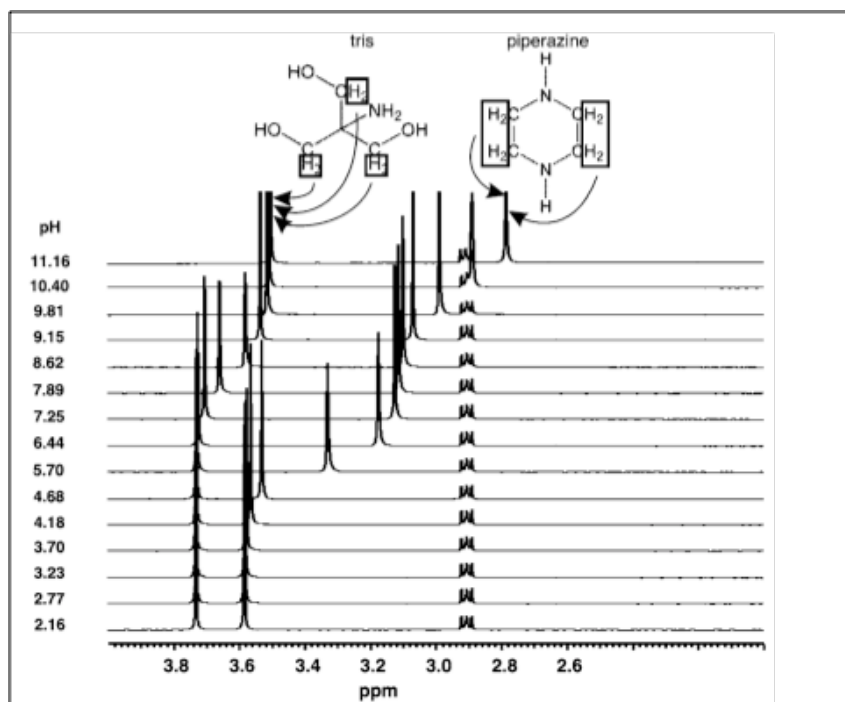


Figure III-38. Representative 1D ^1H -NMR spectra of the buffer containing Tris and piperazine
 1D ^1H -NMR spectra of the Tris and piperazine buffer and internal pH covered from 2.5 to 9.8. Protons that give rise in 1D ^1H -NMR spectrum are shown in boxes and connected to the corresponding resonances with arrows. The triplet corresponds to the resonance from DSS served as a reference compound.

Source: Baryshnikova et al. (2008).

3.3. Bio-Layer Interferometry / Octet system

Bio-Layer Interferometry (BLI) is a label-free technology for measuring biomolecular protein interactions. BLI is an optical analytical technique that analyses the interference pattern of white light reflected on two surfaces: a layer of immobilized protein on the biosensor tip and an internal reference layer. The binding between a ligand immobilized on the biosensor tip surface and an analyte in solution produce an increase in optical thickness at the biosensor tip. The increment in optical thickness result in a wavelength shift $\Delta\lambda$. This spectral shift is monitored at the detector and reported on a sensorgram as a change in wavelength (nm shift). Rates of association, dissociation or concentration of interactions are monitored in real time. Kinetic data were obtained during a pilot experiment using the Octet data acquisition software and a microplate format in the Octet system RED96 – Fig. III-39 (thanks to Pall fortéBio).

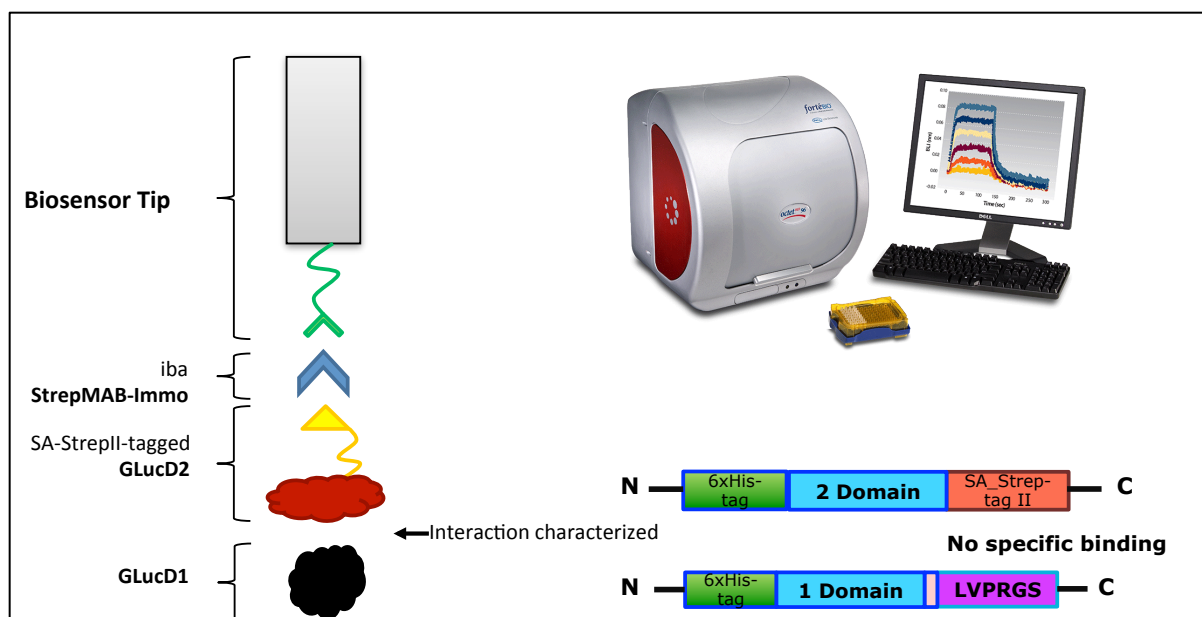


Figure III-39. Bio-Layer Interferometry protein interaction experiment of GLuc catalytic domains

Illustration of the biosensor tip, the fiber and the surface matrix on which a StrepMAB-Immobilized is immobilized. The capturing molecules (GLucD2) – the ligands, are then attached to the tip. The tip is dipped into the sample containing a target molecule (GLucD1) the analyte. The molecular interaction of the two catalytic domains is monitored with the Octet system RED96 from Pall *forteBio* when the wavelength shift changes over time.

The assay steps were programmed using the Octet data acquisition software before running the assay. The kinetic measurement was performed at room temperature. A detailed protocol is shown in Table II-21, Chapter 2.

Fig. III-40A shows kinetic assay data from the experiment using an optimal range of analyte concentration (GLucD1 0, 50, 100, 200 and 400 μM). The sensorgram shows that the StrepMab antibody was loaded onto the AMC sensors (the left side of the sensorgram). After a baseline step, the middle part shows a good binding of the ligand GLucD2 at a concentration of 10 $\mu\text{g}/\text{ml}$. The right part shows non-specific binding of GLucD1 even when the buffer containing the analyte was supplemented with 0.1% BSA and 0.05% P20 to reduce the non-specific binding.

Fig. III-40B shows the sensorgram of the reference subtraction from association and dissociation steps of the kinetic assay and shows very little binding between GLucD1 and GLucD2 with an affinity of 220 μM . However, it needs to be noted that due to this very low affinity and high concentration of GLucD1 used this could be due to various non-specific interactions. This non-specific binding became less pronounced when the detergent 0.05 % P20 was added to the buffer assay and even less pronounced with the addition of 0.1 % BSA.

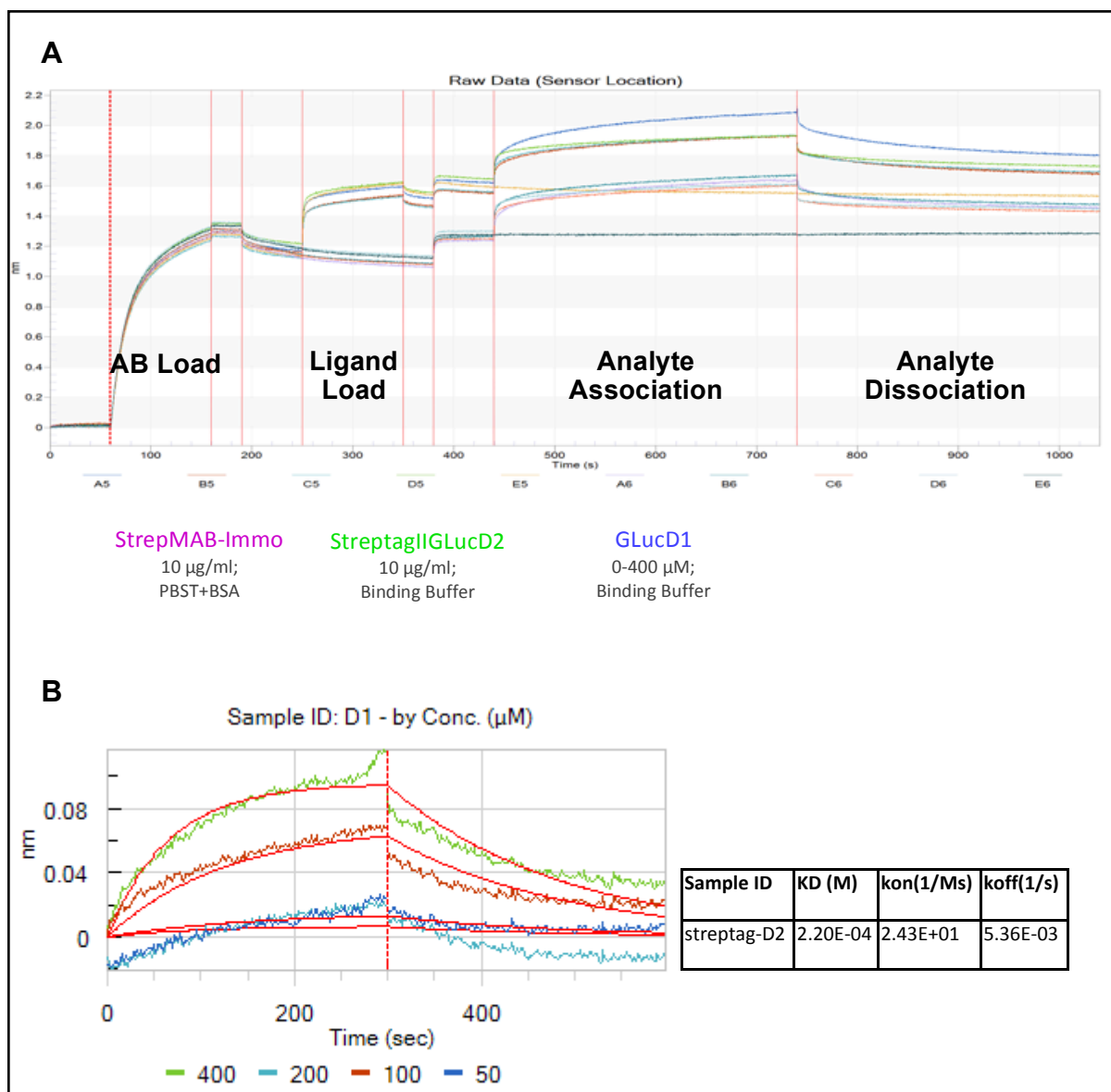


Figure III–40. BLI – Kinetic results of the interaction between both GLuc catalytic domains

A) Raw data sensorgram for ligand and analyte binding showing real-time data acquisition for each step of the kinetic assay. The sensorgram shows on the left part that the StrepMab antibody was loaded onto the AMC sensors. After a baseline step, the middle part shows a good binding of the ligand GLucD2 at a concentration of 10 $\mu\text{g}/\text{ml}$. The right side shows non-specific binding of GLucD1.

B) The sensorgram of the reference subtraction from association and dissociation steps of the kinetic assay. It indicates very little binding between GLucD1 and GLucD2 with an affinity of 220 μM of the analyte sample in solution.

3.4. Crystallization of *Gaussia lucifersase* construct

Full-length GLuc construct was successfully cloned in the pET28a+C-6xHis-tag vector, expressed in *E. coli* and purified by anion ion-exchange. The results of GLuc purification and kinetics experiments show that a highly active enzyme of high purity was obtained. The enzyme was in a homogeneous, low polydispersity state, making the protein suitable for crystallization. Even if a high polydispersity from GLucF1 (Pd. 35%) was observed in the

DLS experiment, it is very likely to crystallize during the screening, as suggested by Borgstahl (2007).

The first crystallization experiments were performed with the aid of high-throughput crystallization (HTX) facility at EMBL-Grenoble, France.

The GLuc protein concentration sent to HTX was 11 mg/ml. The results of the initial crystallization screening in Grenoble showed spherulite-like forms, observed in a condition containing: Methyl-pentanediol (MPD) 60% and 100 mM Sodium Citrate (pH 4.0). Spherulites were observed after four days from the initial screen (Fig. III-41A). After the initial screening, optimization of growth conditions was necessary to obtain crystals suitable for diffraction experiments. Once the spherulites were identified, manual trials, reproduction and optimization of the conditions were carried out by variation of different parameters. These included pH, ionic strength, temperature, protein concentration, presence of various salt, ligands or additives, the type of precipitants and the crystallization method used (hanging/sitting drop vapour diffusion). After several attempts to improve the conditions, an additional screen (Opti-salt, Qiagen) was set up in Grenoble, the sample concentration was 25 mg/ml in the reservoir containing 0.1 M Sodium citrate pH 4.0 and 60% MPD. After two days microcrystals and small clusters were observed in the crystallization drops using visible (Vis) light imaging. The reservoir was composed of 0.1 M calcium chloride and 0.01 M Tris-HCl (pH 8.5). After 8 days, the ultraviolet (UV) light imaging of the crystallization drop obtained in HXT-Grenoble showed white and bright small clusters (Fig. III-41B). UV light imaging can differentiate between protein crystals and salt crystals, as the visualization of protein crystals is based on the presence of intrinsic tryptophan, thus protein crystals are determined by localised white and bright tryptophane fluorescence.

The conditions identified at the HTX-Grenoble, were reproducible and optimized in-house. In addition, a protein sample was co-crystallized with 1 mg/ml CTZ (the substrate was previously dissolved in 10 % v/v Methanol); however since the CTZ is quickly oxidized under normal environmental conditions, no influence on the protein crystallization was detectable (Fig. III-41C). Furthermore, powder diffraction tests have been performed at the EMBL-operated beamline P13 of the PETRA III synchrotron source at the German Electron Synchrotron (DESY) in Hamburg, Germany. Some microcrystals were tested on beamline P13, which is one of the most brilliant synchrotron source available to date. Since clusters were very small, very few spots could be missed with a small rotation of 1 degree. For this reason the microcrystals were tested with a rotation range of 5 degrees. The result showed a diffraction pattern of several rings. To improve the diffraction quality and obtain single spots, further crystallisation optimization should be performed.

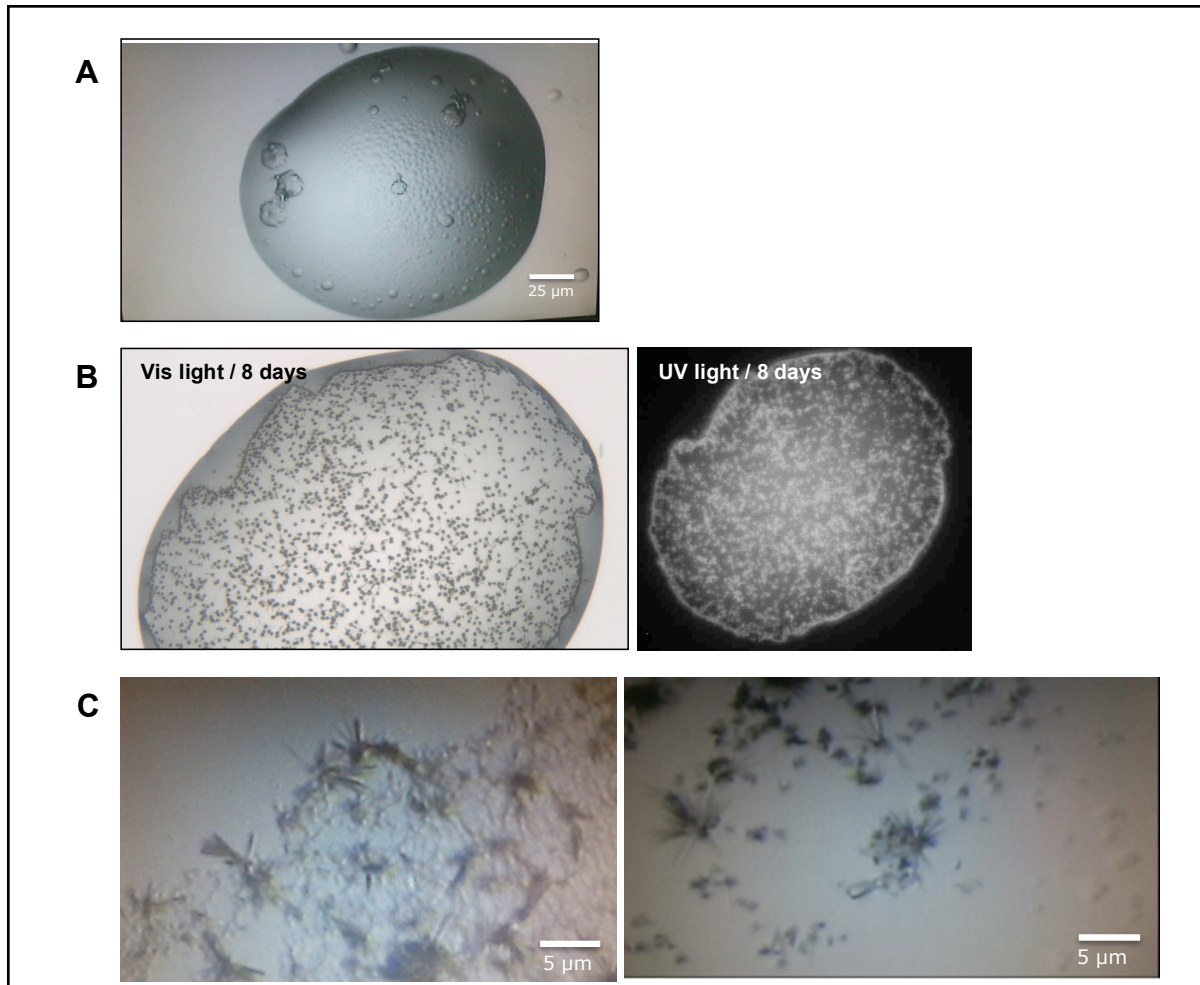


Figure III-41. Crystallization of recombinant GLucFl construct

Representative images of GLucFl microcrystals. A) Initial crystallization conditions: protein concentration of 11 mg/ml. Images of the drop in visible light show spherulite growth after 1 week. B) Small crystal cluster growth after 8 days under optimized conditions (0.1 M Calcium Chloride, 0.01 M Tris-HCl (pH 8.5), protein concentration 25 mg/ml). Vis and UV imaging was performed at the HTX-EMBL Grenoble. C) Visual inspection of the crystallization plates performed manually in house.

The crystallization conditions in which microcrystals have been obtained had to be optimized to obtain a single crystal to determine the three dimensional structure using X-ray diffraction experiments at a synchrotron beamline. Besides, the microcrystals could be tested by serial synchrotron crystallography (SSX) techniques. Moreover, the crystal structure from GLuc is a necessary prerequisite to further explore the underlying reaction mechanism. Knowledge about the architecture of the active site can help design variants as markers or receptors with higher efficiency or modified spectral characteristics, which will be useful tools in cell biology.

3.5. *In vivo* crystallization of *Gaussia* luciferase

Doye and Poon (2006), provided evidence that certain proteins can crystallize in living cells. Examples include crystallisation of peroxidase in peroxisomes, insulin in pancreatic cells and storage proteins in plant seeds. In the last years, more examples of *in vivo* crystal formation from heterologous proteins are related to baculovirus–infected insect cells (Schönherr *et al.*, 2015). The crystallization process is dynamic and occurs in different cellular locations when the baculovirus expression vector system (BEVS) is used (Duszenko *et al.*, 2015). In this project a protocol for the *in vivo* crystallization of GLuc in *Sf21* insect cells using the baculovirus system was established.

Gaussia luciferase was successfully subcloned into the pFastBac1 expression vector plasmid (Invitrogen) and the vector was transformed into DH10Bac competent *E.coli* cells (Invitrogen), following the manufacturer's instructions. The construction of recombinant baculovirus, and subsequent generation of high-titer virus stocks (P3) were performed according to protocol described in section 2.9.4, Chapter 2. After transfection, the baculovirus takes control of the insect cell by inhibiting transcription of cellular genes until the infected cell dies. The initial virus stock P1 was used to generate higher-titer virus stock P3. Recombinant high-titer virus stock (P3) was used to infect 70%-confluent monolayer cultures of *Sf21* insect cells. 72 to 96 h after the infection with P3 virus stock, needle-shaped protein microcrystals appeared in the living *Sf21* cells. Microcrystal formation in these living cells was visualized with a light microscope (Fig. III-42B, C and D). Microcrystal growth continued up to day 5, regularly exceeding the normal dimension of the *Sf21* cells which are 10 μm in length. The microcrystal size and morphology of Firefly luciferase grown and crystallized *in vivo* in *Sf9* insect cells (Schönherr *et al.*, 2015) were similar to GLuc microcrystals observed in this project.

Recombinant high-titer virus stock P3 from the cathepsin B (TbCatB) protein target from *Trypanosoma brucei* cysteine protease was used to infect 70% confluent monolayer culture of *Sf21* insect cells, grown in serum-free medium at 27 °C. This served as a positive control for crystallization *in vivo*. Approximately 72 to 96 h after the infection, the formation of needle-shaped microcrystals started to be visible within cells under the light microscopy (Fig. III-42A). Several microcrystals had grown to size of 15 to 20 μm in length. Additionally, crystals were detected floating in the solution. A similar phenomenon was also observed for heterologous expression of Firefly luciferase, in the same expression system (BEVS) (Koopmann *et al.*, 2012).

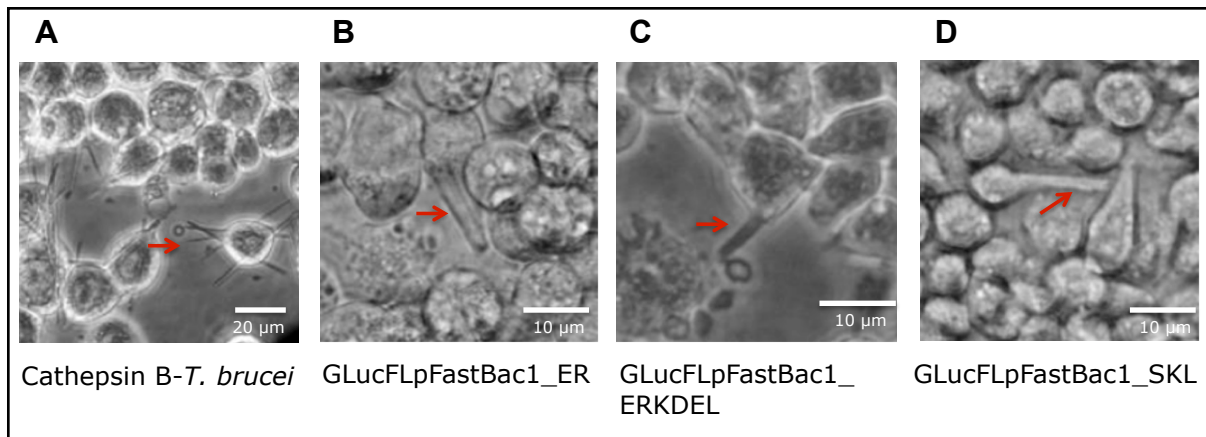


Figure III-42. Light microscopic images of *Sf21* insect cells with *in vivo* protein crystal growth
Light microscopic images showing *Sf21* insect cells containing crystals (indicated by red arrows). A) *In vivo* crystals of Cathepsin B-*T. brucei* after 96 h of *Sf21* cells infected with high virus titer used as a positive control in the laboratory. B) C) and D) Corresponding GLuc crystals formation in *Sf21* living cells infected with P3 virus stocks.

IV

DISCUSSION

Luciferases are ideal for the development of biological reporter systems because the light they emit can easily be measured and quantified. Consequently, luciferases are widely used to study basic cellular processes such as tumour growth, interactions of proteins and gene regulation, transcription and translation.

The catalytic properties of GLuc make it an extremely sensitive detection reporter and makes GLuc a potential tool in bio-imaging applications as reporter protein for *in vivo* tumour growth and signal transduction processes (Maguire *et al.* 2013) or, when used as a split-reporter protein, for protein-protein interaction studies (Remy and Michnick, 2006).

GLuc has no significant sequence homologies with other luciferases including coelenterazine-type luciferases from *Renilla reniformis* (Lorenz *et al.*, 1991). So far, the options of redesigning GLuc variants are limited due to the lack of structural data. Therefore, the main work carried out in this PhD project was to produce a construct of the full-length GLuc to obtain soluble and pure protein for crystallization in order to elucidate the GLuc 3D-structure by X-ray structure analysis and NMR spectroscopy.

Moreover, a full-length GLuc with an internal thrombin cleavage site between the two catalytic domains (GLucF1Throm) and the characterization of the two separately expressed domains (GLucD1 and GlucD2) were generated to investigate the luminescence properties and to establish the optimum buffer conditions (e.g. pH, salt, temperature) for stimulation of the bioluminescence activity.

Three constructs of full-length GLuc in *Sf21* insect cells using the baculovirus expression vector system (BEVS) for crystallization *in vivo* were also generated and expressed.

Plasmids construction of *Gaussia* luciferase with a thrombin cleavage site and catalytic domains for the expression in *E. coli* cells

In this PhD project, one construct of the full-length GLuc with thrombin cleavage site between amino acids residues 97/98 without the signal peptide, and two constructs of the individual catalytic domains (aa 1-97 and aa 98-168, respectively) were produced. The position for the insertion of the thrombin cleavage site was determined based on research by Inouye *et al.* (2008). They found two short individual sequences of 71 amino acid residues containing 4 conserved cysteine residues in each individual sequence, suggesting that this sequence corresponds to two structural domains. Consequently, they split the full-length GLuc sequence between amino acid residues 97/98 without the signal peptide (Fig. IV-1). In

addition, they found that both domains produce bioluminescence when expressed separately. However, Inouye *et al.* (2008) purified GLuc mainly from inclusion bodies.

In contrast, Luker and Luker (2014) defined a split *Gaussia* luciferase complementation for imaging ligand-receptor binding in HEK cells and used the GLuc enzyme fragments divided between amino acid residues 93/94 (omitting the 16 amino acid leader sequence) to form NGLuc and CGLuc with amino acids 1-93 and 94-169, respectively (Fig. IV-2). In this case, luciferase complementation requires fusion of inactive amino (N)-terminal and carboxy (C)-terminal enzyme fragments to two different target proteins. It is assumed that interaction of the target proteins will reconstitute the active enzyme, producing bioluminescence as a quantitative measure.

Hashimoto *et al.* (2011) and Kim *et al.* (2009) identified other potential sites for splitting the GLuc (amino acid residues 92/93 or amino acid residues 105/106, respectively). In both papers, the authors measured the luminescence of each domain separately and the domains did not show any luminescence. Consequently, the light emission resumes only when both fragments aggregate to a full-length GLuc enzyme. However, in these papers the activity of GLuc protein was measured in cell suspension (without cell disruption) instead of supernatant and/or purified protein.

In this PhD project, the three constructs (GLucF1Throm, GLucD1 and GLucD2) were successfully cloned and expressed in *E. coli* SG13009(pREP4) cells. In contrast to the above works, the protein from these constructs was further purified from the soluble fraction (supernatant) using Ni-afinity chromatography and analytical SEC. The concentration of GLuc proteins was measured by Bradford assay before the measurement of luminescence activity.

Because GLuc contains disulphide bonds, all purification procedures were carried out under non-reducing conditions (i.e. without including DTT or b-mercaptoethanol). Most secreted proteins contain cysteine residues and disulfide bonds and GLuc is no exception. The preparation of large quantities of homogeneous recombinant GLuc with a high and stable bioluminescence activity is a difficult process, because of the presence of 10 cysteines that are assumed to be involved in formation of native disulphide bonds (Goerke *et al.*, 2008). Mutational changes of cysteine residues in GLuc sequence lead to significant alteration of bioluminescence response. For instance, Takenaka *et al.* (2012) reported that the substitution of cysteine residues with alanines in domains 1 and 2 of MpLuc1 from *Metridia pacifica* resulted in a complete loss of activity when it was expressed in *E. coli*, suggesting an essential role of the cysteine residues in luciferase activity.

In this PhD project, to confirm a correct formation of intramolecular disulphide bonds and to verify the homogeneity of the purified protein by SEC, fractions from each purified protein construct were analysed by RP-HPLC before running the bioluminescence kinetics experiments. The RP-HPLC pattern of soluble GLuc constructs exhibited a large single peak,

suggesting that these samples are homogeneous. Following the purification and the RP-HPLC analysis, several kinetic experiments were performed in order to establish the optimum buffer conditions (e.g. pH, salt, temperature, detergents and reducing agents) for bioluminescence activity.

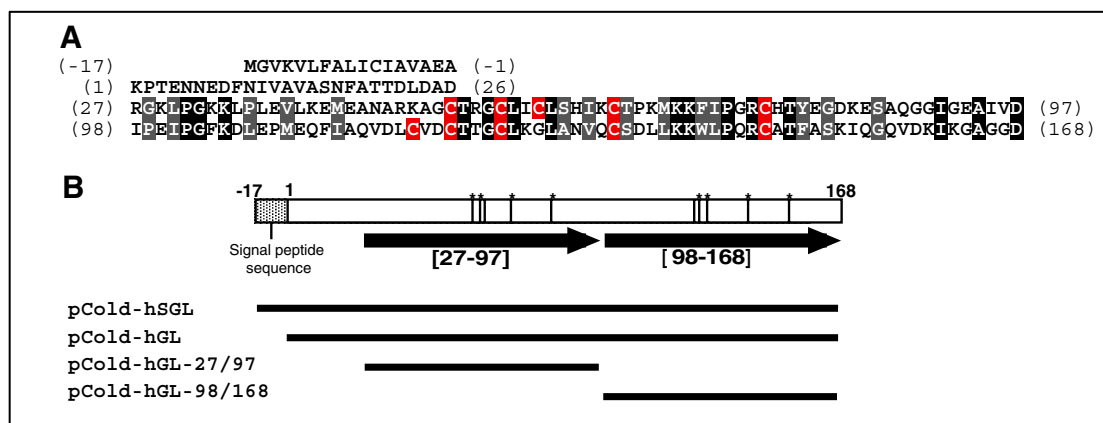


Figure IV-1. Amino acids sequence alignment and schematic representation of *Gaussia luciferase*

A) Alignment of two doimans of *Gaussia luciferase*. Red boxes indicate the conserved cysteine residues. B) Schematic representation of the GLuc gene. The two arrows with positions indicated in square brackets show the individual domains of GLuc.

Source: Inouye *et al.* (2008).

Bioluminescence activity from GLuc is highly stable in the optimal buffer condition described in this dissertation (section 1.4.1, Chapter III). According to the results presented here, the best buffer was 100 mM Tris-HCl (pH 8.2) containing 100 NaCl, the detergent Tween 0.1% and without reducing agents. This buffer plays a key role in stabilizing the GLuc samples light emission. Welsh *et al.* (2009) found that the presence of Tween in the assay buffer prolongs the light emission from GLuc protein.

The determination of the right buffer for the luminescence allowed further experiments to study the functional basis and the enzymatic properties of GLuc and its two catalytic domains. Thus, the specific activity kinetics performed in this PhD project revealed that the full-length GLuc protein was highly active in comparison with the full-length GLuc with internal thrombin cleavage site before cleavage. Furthermore, it was found that both individual domains have the catalytic ability to emit light, but it was of lower luminescence counts in comparison with the luminescence from the full-length GLuc. Moreover, the bioluminescence of the full-length GLuc protein is around 10 times higher than the sum of the individual domains (GLucD1 and GlucD2).

GLucFITHrom was cleaved into two individual domains with thrombin protease. The cleavage reaction of the entire region of GLuc was successful, but it was not possible to separate the individual domains from one another. Perhaps this is because there is a

conformational flexible loop between the two conserved sequence regions of GLuc. The kinetics assay showed that the luminescence activity of GLucF1Throm after cleavage was reduced by half in comparison to the full-length GLuc. This result is in accordance with Maguire *et al.* (2009) who produced different GLuc variant designs with point mutations to prolong the bioluminescence activity. For example, they introduced a point mutation with a single nucleotide change which replaced methionine at position 43 with isoleucine. However, the luminescence activity of this mutation was lower than that the wild type GLuc.

In this PhD project the mixture of the two domains GLucD1 and GLucD2 shows luminescence activity similar to the sum of the single domains. However, the kinetic activity is not increasing as expected, because it seems that there is no luminescence activity complementation between both domains. The activity assay from GLucD1 and GLucD2 does not work synergistically when both domains are mixed and thus do not add up to the full activity of the full-length construct. This may be due to the fact that the fusion tags (6xHistag or Strep tag II) from the constructs interfere with the complementation of both domains, and/or perhaps it is necessary to follow a systematic screening for finding the best pair of luciferase fragments (Luker, G. and Luker, K. 2011).

Options for better redesigning of GLuc mutations are limited due to the lack of structural data. Mutational analysis can serve to reveal the structural and functional characteristics of a protein. Welsh *et al.* (2009) has characterized several mutations of full-length GLuc which show differences in activity and decay of the luminescence signal. They found the strongest effect on decay of the luminescent signal is observed when both M43L and M110L mutations are present simultaneously, as opposed to the individual mutations.

Another luminescence property studied in this PhD project was the effect of buffer pH on GLuc activity. This effect was determined with various pH values and the intensity of light emission from luciferase samples appears to be highly pH dependent. The results suggest that the optimum pH for all sample constructs including the single domains is at 8.2 +/-0.1. These results are supported by other studies (Inouye *et al.*, 2008 and Takenaka *et al.*, 2008), which show that even subtle pH changes affect the enzymatic reaction, influencing the activity or shifting the spectra. Prominent examples of this behaviour are RmLuc (Matthews *et al.*, 1977) and NanoLuc (Hall *et al.*, 2012). In addition, Law *et al.* (2006) reported optimal pH values around 8.0 in five mutant luciferases from *Photinus pyralis*. Moreover, natural habitat of *Gaussia princeps* is deep-sea cold water, where pH values are around 8.01 (Waters and Millero, 2013).

The effect of ions on luminescence intensity of GLuc was investigated using buffers containing different concentrations of salt (NaCl). The results show that GLuc bioluminescence is salt concentration dependent and the highest signal intensity was observed in the presence of 100 mM NaCl for the full-length GLuc and for the two catalytic domains. The highest signal intensity for RmLuc is observed at a higher concentration of 200 mM NaCl. Salt ions might affect the protein conformation, association and/or folding to

modulate their enzymatic activities (Takenaka *et al.*, 2008). Although the mechanism of the increased bioluminescence with NaCl is unknown, the results obtained here are consistent with the notion that GLuc is secreted into and functions in a marine environment that contains about 470 mM Na⁺. Moreover, Inouye and Sahara (2007) reported that it is likely that the bioluminescent intensity from GLuc is salt concentration dependent and the monovalent ion Cl⁻ significantly stimulates the luminescence activity.

The bioluminescence activity reaction of luciferases is sensitive to changes in temperature (Takenaka *et al.*, 2008). The results obtained in this project prove that GLuc luminescence is temperature dependent and the catalytic properties of the domains GLucD1 and GLucD2 are almost identical to the GLucFl at around 28°C. In contrast, results from RmLuc produced different signals and the kinetics experiments show that the optimum luminescence is above 40 °C.

In the same way, the effect of thermal stability of the bioluminescence activity was compared between *Gaussia* and *Renilla* luciferases. The results show that GLucFl has a robust thermostability in comparison to RmLuc samples – heat half-inhibition of GLuc is at 88.5°C and of RmLuc at 34.4°C. This represents an advantage, because GLuc might be suitable for novel applications under harsh conditions.

Furthermore, other reporter proteins (e.g. *Photinus pyralis* luciferase) are of limited use because of their low thermophilic properties (Law *et al.* 2006 and Branchini *et al.*, 2007). Similar results were reported for other luciferases from the marine copepod, where a striking feature of the Mplucs (*Metridia pacifica* luciferases) was its thermostability (Takenaka *et al.*, 2008). In addition, Baggett *et al.* (2004) reported that the use of thermostabilized luciferases might allow the monitoring of micro-metastases and early stages of tumour growth.

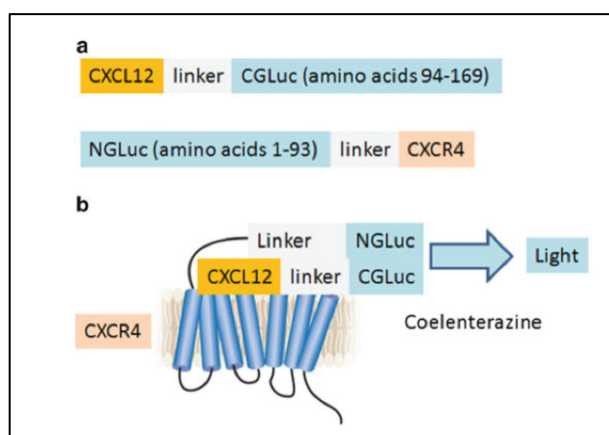


Figure IV-2. GLuc complementation constructs and ligand-receptor interaction

A) Schematic diagrams of GLuc complementation constructs fusion of amino (N)-terminal and carboxy (C)-terminal enzyme fragments. B) Binding of CXCL12-CGLuc to NGLuc-CXCR4 reconstitutes GLuc activity to oxidize the substrate CLZ and the light product.

Source: Luker and Luker, (2014).

Expression and purification of full-length *Gaussia* luciferase construct

To the best of my knowledge, this dissertation is the first to report a protocol to purify GLuc with native disulfide bonds (expressed in *E. coli* cells) from the soluble fraction. Previous research focused on expression in other systems (cell-free, mammalian or insect cells) and purification of GLuc from insoluble fractions (inclusion bodies). The protocol presented here allows to obtain sufficient amounts of highly active, stable and pure protein.

The full-length GLuc construct was successfully cloned into pET28a+C-6xHis-tag vector, expressed in *E. coli* cells and purified by anion ion exchange chromatography. The yield of high purity protein after purification was 3-4 mg obtained from 1 L cell culture. This amount of purified protein was sufficient and suitable for performing biophysical and biochemical studies, as well for crystallization screening experiments.

Thus, after purification, the soluble and high purity protein obtained was used for crystallization screening experiments. The microcrystals obtained after optimization trials (such as cluster formations) should be improved to obtain single spots in the X-ray diffraction pattern or for use in serial femtosecond crystallography (SFX) or serial synchrotron crystallography (SSX) methods. Further optimization of crystallization conditions to obtain single crystals should be performed.

The results obtained during the GLuc purification have been supported by the DLS analysis. The purification of GLuc fused with 6xHis-tag at the C-terminus by IEC together with HPLC measurements gives protein with sufficient purity for crystallization trials. The purity of the sample was reflected by the DLS results because no other species were detected. Based on the intensity and mass distribution data, the sample was classified as polydisperse by the Malvern DLS analysis software, showing a relative polydispersity of 35.7% by intensity and 28.4% by volume. The monodispersity index obtained from the DLS analysis is often used as a parameter for prediction of formation of protein crystals.

Additional biophysical 1D ^1H -NMR spectroscopy of the full-length GLuc showed well-dispersed sharp peaks meaning that the protein is likely to be structured, folded and stable. Nevertheless, further experiments should include the optimization of further conditions that play important role in NMR experiments (e.g. temperature, pH, salt).

In contrast, the spectrum from GLucD1 presents random coil values (a clear indication of the presence of unfolded flexible regions). Probably because it is a short peptide domain and there is a lot of uncertainty about the length of the individual domains (Inouye *et al.*, 2008; Luker and Luker, 2014; Cassonnet *et al.*, 2011; Remy and Michnick, 2006; Kim *et al.*, 2009). Or perhaps the sequence region where it was cut (aa 97/98) contains a flexible linker that impedes the dissection of the entire GLuc region. Another reason of the presence of the unfolded GLucD1 could be that the GLucD1 sequence contains four conserved cysteine residues that need to form disulfide bonds.

This result is in accordance to previous NMR investigations with other GLuc variants identified by Rathnayaka *et al.* (2011). They found that the $^1\text{H-NMR}$ spectrum of recombinant GLuc exhibits well disperse and sharp peaks typical of a natively folded protein of this size.

To study the protein-protein interaction, a pilot experiment in a Bio-Layer Interferometry (BLI) was performed. BLI is a label-free technology for measuring biomolecular protein interactions. The results suggest a low affinity between GLucD2 and GLucD1. However, it cannot be ruled out that the observed binding signal is due to non-specific binding caused by the high analyte concentration. On the other hand, it remains an open question whether there is a binding signal when the experiment is set up in the other way around binding first GLucD1(ligand) and subsequently GLuCD2 (analyte) proteins. According to Hattori and Ozawa (2014), the order of fusion proteins is important for protein complementation assay. Further experiments should test whether the interaction of both domains is dynamic and if there is a synergetic effect, which could be used for complementation assay.

***In vivo* crystallization of *Gaussia* luciferase**

Heterologous expressed proteins are also able to form crystals within cells. In this project three constructs were generated using the baculovirus *Sf21* expression system in insect cells. The *in vivo* crystallization experiments for GLuc were successful. Approximately 72 h after infection, the formation of needle-shaped crystals began to be visible by light microscopy in *Sf21* insect cells infected with recombinant baculovirus that contained the gene encoding the full-length GLuc. Microcrystal growth continued up to day 5, regularly exceeding the normal dimension of the *Sf21* cells which are 10 μm in length. The GLuc microcrystal size and morphology grown *in vivo*, reported in this dissertation were similar to the microcrystals size and morphology from firefly luciferase using *in vivo* crystallization in *Sf9* insect cells (Schönherr *et al.*, 2015).

During the progress of the *Sf21* cell infection, the number of GLuc crystals continuously increased, but no free luciferase crystals floating in the medium were detected under the microscope. In contrast, the *T. Brucei* cathepsin B crystals produced as a control were observed outside and either remained attached to cell remnants or floated freely within the medium (Duszenko *et al.*, 2015).

Successful protein crystallization within living cells, compared to conventional crystallization methods, has the advantage of avoiding time-consuming experiments to optimize the crystallization conditions.

CONCLUSIONS AND FUTURE PERSPECTIVES

Gaussia luciferase is attracting much attention as a potential reporter protein. Studies on GLuc are relatively recent, and the structure and biophysical properties of this protein remain to be fully investigated.

Constructs of *Gaussia* luciferase with a thrombin cleavage site and catalytic domains for the expression in *E. coli* cells

In this PhD project, the full-length GLuc with a thrombin cleavage site, as well as two constructs of the individual catalytic domains, were expressed in *E. coli* strain SG10039. These proteins were successfully purified with a size exclusion chromatography system. The homogeneity of the purified protein was verified by RP-HPLC. Soluble GLuc constructs (GLucFITHrom, GLucD1 and GLucD2) were purified to homogeneity by reversed phase HPLC. The HPLC elution pattern of the samples exhibited a large single peak suggesting that these are homogenous proteins. Therefore, based on these results it was possible to study the functional basis of enzymatic activity from *Gaussia princeps* and its bioluminescence activity.

Further, bioluminescence kinetic experiments were performed and the optimal assay buffer (50 mM Tris-HCl pH 8.2, 100 mM NaCl, 0.1% Tween and without reducing agents) for GLucFITHrom and its catalytic domains were determined. It was therefore demonstrated that the pure full-length GLuc is highly active at optimal conditions.

It was also observed that the full-length GLuc protein is composed of two domains (GLucD1 and GLucD2) and, according to the kinetic experiments, each individual domain exhibits bioluminescent activity with CTZ. Investigation of the protein properties showed that optimal conditions (pH, temperature, salt concentration) for the full-length GLuc and both individual domains are identical. Other important properties of GLucD1 and GLucD2 as individual domains are: (i) smaller size (15 and 11 kDa, respective theoretical molecular weight) than those of MpLuc1 (22.7 kDa) and MpLuc2 (20.3 kDa) luciferase domains, (ii) ability to express active forms in *E. coli* (iii) simple reactions in which luciferin and luciferases are the only two requirements for light emission. These features have not been previously reported in studies of *G. princeps* luciferase expressed in *E. coli* cells. Furthermore, GLuc bioluminescence is salt concentration dependent and the highest signal intensity was reached at 100 mM NaCl.

Other important feature of the full-length GLuc is its thermostability. This PhD project shows that the luciferase enzyme is highly active and stable over a wide range of temperatures.

GLuc activity was detectable and stable up to 70 °C and the pre-treatment with temperatures above 80 °C caused inactivation (heat half-inhibition $ET_{50\%} = 88.5$ °C). In contrast, the thermostability experiment of RmLuc shows that the luminescence is stable up to 30 °C and is half-heat-inhibition at 34.4 °C.

Moreover, the kinetics experiments conducted using the mixture of the individual GLuc domains showed that there is no activity complementation between the GLuc domains 1 and 2 because there is no synergistic effect between them. Therefore, the activity of the full-length GLuc (i.e. reconstitution) is not possible to achieve with the constructs generated here. Further experiments for protein complementation assays should test whether dynamic interactions can be obtained using split GLuc forms. A future task should develop constructs of the split-GLuc (GLucD1 and GLucD2) fused to individual protein partners of interest, to study the protein fragments complementation. It is important to consider the possible advantage that the two catalytic domains are small and likely less apt to interfere with functions of the fusion protein partner.

Construct of full-length *Gaussia* luciferase for the expression in *E. coli* cells

This PhD project has also determined the functional characterization of the enzymatic activity of the full-length GLuc construct expressed in *E. coli* BL21 (DE3) strain using various biochemical and biophysical methods. The established and optimized protocol for purification by IEC allowed GLuc protein in the soluble fraction to be obtained, and its purity was confirmed by 12% SDS-PAGE. The high purity, stability and homogeneity of the GLuc was confirmed by the DLS results and the protein folding state was assessed by 1D $^1\text{H-NMR}$ experiments. The biophysical characterization results of the purified GLuc indicate that the protein is highly stable. Consequently, these results suggest that the purified GLucFl protein sample is fully active and it can be successfully used for crystallization screening experiments and for biophysical and biochemical studies.

In order to elucidate the structural basis for bioluminescence, it is important to determine the structure of the full-length GLuc, as well as the structures of each individual domain. These shorter constructs increase the likelihood of successful crystallization experiments. Based on the individual structures, it would be possible to elucidate the underlying reaction mechanism. Furthermore, it could be possible to initiate a rational protein design approach for modifications of the enzyme towards an even higher quantum yield, broader pH-stability and reduced oxidation sensitivity. Moreover, the elucidation of the 3-D structure of GLuc by NMR experiments should also be considered, because NMR experiments have the advantage of avoiding time-consuming experiments to optimize the crystallization conditions.

The preliminary results from the BLI experiments seem to indicate a low affinity interaction between GLucD2 to GLucD1 protein. However, it cannot be ruled out that the observed

binding signal is present due to non-specific binding caused by the high analyte concentration. Further experiments using BLI should be performed.

Constructs of *Gaussia* luciferase for *in vivo* crystallization expressed *Sf21* insect cells

In this PhD project, three constructs of full-length GLuc were generated for testing an *in vivo* crystallization method. The cloning and expression of the full-length GLuc in pFastBac1 vector using *Sf21* insect cells were successful. The experiments resulted in a successful crystallization *in vivo*. For all of the three constructs, crystals were observed under light microscopy.

Further experiments should focus on the isolation of crystals from infected cells grown in the monolayer culture. Crystals need to be harvested, washed and analysed by a serial femtosecond crystallography (SFX) method using ultra-intense X-ray pulses from free-electron lasers (XFELs). This method might bring new opportunities to enable high-quality data collection from a flowing suspension of crystals of micrometre size. In fact, Redecke *et al.* (2013) demonstrated the applicability of SFX on *in vivo* grown crystals through the study of cathepsin B. Thus, crystals grown within insect cells offer new possibilities for crystallization of proteins that do not form suitable crystals *in vitro*.

VI

LITERATURE

- Badr, CE. (2014). Bioluminescence Imaging: Basics and Practical limitations. *Methods Mol Biol*;1098:1-18. doi: 10.1007/978-1-62703-718-1_1.
- Baggett, B., Gillies, RJ., Momen, S., Morgan, S., Morse, D., Roy, R. and Tisi, L. (2004). Thermostability of firefly luciferases affects efficiency of detection by *in vivo* bioluminescence. *Molecular imaging*, 3 4, 324-32.
- Baneyx, F., Mujacic, M. (2004). Recombinant protein folding and misfolding in *Escherichia coli*, *Nat. Biotechnol.* 22 1399–1408.
- Bannister, NJ, Herring, PJ. (1989). Distribution and structure of luminous cells in four marine copepods. *J. Mar. Biol. Assoc. U.K.* 69: 523-533.
- Barnes, AT, Case, JF. (1972). Bioluminescence in the mesopelagic copepod, *Gaussia princeps* (T. Scott). *Journal of Experimental Marine. Biology and Ecology.* 8, 53–71
- Baryshnikova, O.K., Williams, T.C. and Sykes, B.D. (2008). Internal pH indicators for biomolecular NMR. *Journal of Biomolecular NMR*, Volume 41, Issue 1, pp 5-7. Doi:10.1007/s10858-008-9234-6
- Biggley, WH., Swift, E.,Buchaman R. J., Seliger HH. (1969). Stimulable and spontaneous bioluminescence in the marine dinoflagellate, *Pyrodinium bahamense*, *Gonyaulax polyedra*, and *Pyrocystis lumula*. *Journal of General Physiology* 54:96-122.
- Berthold F; Herick K, Sieve R. (2000). Luminometer design and low light detection. In: Ziegler MM, Baldwin TO, eds. *Methods in Enzymology* San Diego, London, Boston: Academic Press 62-87.
- Borgstahl, G. E. O. (2007). How to use dynamic light scattering to improve the likelihood of growing macromolecular crystals. In *Methods in Molecular Biology* (Vol. 363, pp. 109-129). (Methods in Molecular Biology; Vol. 363). DOI: 10.1385/1-59745-209-2:109
- Bowie, LJ, Irwin, R, Loken, M, De Luca, M, Brand, L. (1973). Excited-state proton transfer and the mechanism of action of firefly luciferase. *Biochemistry*; 12(10): 1852–1857
- Bradford, MM. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* 72:248–254.
- Branchini, B. (2016). Firefly Luciferase-Based Sequential Bioluminescence Resonance Energy Transfer (BRET)-Fluorescence Resonance Energy Transfer (FRET) Protease Assays. In: Kim S. (eds) *Bioluminescence. Methods in Molecular Biology*, vol 1461. Humana Press, New York, NY. DOI 10.1007/978-1-4939-3813-1_1.
- Bryan B. J., Szent-Gyorgyi C. S., (1999). Luciferases, fluorescent proteins, nucleic acids encoding the luciferases and fluorescent proteins and the use thereof in diagnostics, high throughput screening and novelty items, *Int. Patent Appl.* WO 99/49019.

- Campbell, AK (1994). *Rubicon, the fifth dimension of biology*. Duckworth, London.
- Campbell, AK (2003). Save those molecules! Molecular biodiversity and life. *Journal of Applied Ecology*, 40; 193-203.
- Cassonnet P, Rolloy C, Neveu G, Vidalain PO, Chantier T, Pellet J, Jones L, Muller M, Demeret C, Gaud G, Vuillier F, Lotteau V, Tangy F, Favre M, Jacob Y. (2011). Benchmarking a luciferase complementation assay for detecting protein complexes. *Nat Methods*. Nov 29;8(12):990-2. doi: 10.1038/nmeth.1773.
- Childress, J.J. (1977). Effects of pressure, temperature and oxygen on the oxygen-consumption rate of the midwater copepod *Gausia princeps*. *Mar. Biol.*;39:19–24. doi:10.1007/BF00395588
- Clarke, GL., Conover, RJ., David, CN., Nicol, J. AC. (1962). Comparative studies of luminescence in copepods and other pelagic marine animals. *Journal of the Marine Biological Association of the United Kingdom*, 42, 541–564.
- Doye, JPK and Poon, W. (2006). Protein crystallization *in vivo*. *Current opinion in colloid & interface science*, vol 11, no. 1, pp. 40-46. DOI: 10.1016/j.cocis.2005.10.002
- Du ,Q. J., Tang, Y., Huang, W., Rodney, E., Wang, L, Liang, G. (2011). Development of recombinant rabies viruses vectors with *Gaussia luciferase* reporter based on Chinese vaccine strain CTN181, *Virus Res*. 160 82–88.
- Dunstan, SL, Sala-Newby, GB, Fajardo, AB, Taylor, KM, Campbell, AK. (2000). Cloning and expression of the bioluminescent photoprotein pholasin from the bivalve mollusc *Pholas dactylus*. *The journal of biological chemistry*, 275, 9403- 9409.
- Duszenko, M., Redecke, L., Mudogo, C.N., Sommer, B.P., Mogk, S., Oberthuer, D., Betzel, C. (2015). *In vivo* protein crystallization in combination with highly brilliant radiation sources offers novel opportunities for the structural analysis of post-translationally modified eukaryotic proteins. *Acta Crystallogr F Struct Biol Commun*. Aug;71(Pt 8):929-37.
- Enjalbert, B., Rachini, A., Vedyappan, G., Pietrella, D., Spaccapelo, R., Vecchiarelli, A., Brown, A.J., Enfert, C. (2009). A multifunctional, synthetic *Gaussia princeps luciferase* reporter for live imaging of *Candida albicans* infections, *Infect. Immun*. 77 4847–4858.
- Ernst, O., Zor T. (2010). Linearization of the Bradford Protein Assay. *Journal of Visualized Experiments* (38) e1918.
- Gaur, S., Bhargava-Shah, A., Hori, S., Afjei, R., Sekar, TV., Gambhir, SS., Massoud, TF., Paulmurugan, R. (2017). Engineering intracellularly retained *Gaussia Luciferase* reporters for improved biosensing and molecular imaging applications. *ACS Chemical Biology*. Vol 12(9) PMID 28767220 DOI: 10.1021/acscchembio.7b00454
- Giacovazzo, C., Catti, M., Monaco, HL., Artioli, G., Viterbo, D. (2011). *Fundamentals of Crystallography*. 3rd ed. Oxford University Press.
- Goerke, AR., Loening, AM., Gambhir, SS. and Swartz, JR. (2008) Cell-free metabolic engineering promotes high-level production of bioactive *Gaussia princeps luciferase*. *Metab. Eng*. 10, 187-200.

- Griesenbach, U., Vicente, CC., Roberts, MJ., Meng, C., Soussi, S., Xenariou, S., Tennant, P., Baker, A., Baker, E., Gordon, C., Vrettou, C., McCormick, D., Coles, R., Green, AM., Lawton, AE., Sumner- Jones, SG., Cheng, SH., Scheule, RK., Hyde, SC., Gill, DR., Collie, DD., McLachlan, G., Alton, EW. (2011). Secreted *Gaussia* luciferase as a sensitive reporter gene for *in vivo* and *ex vivo* studies of airway gene transfer. *Biomaterials* 32(10): 2614– 2624. doi:10.1016/j.biomaterials.12.001
- Haddock, SHD., McDougall, CM., Case, JF. (2000) The Bioluminescence Web Page. <http://lifesci.ucsb.edu/~biolum/2000>.
- Haddock, SHD., Moline, MA., Case, JF. (2010). Bioluminescence in the sea. *Ann. Rev. Mar. Sci.* 2,443-493.
- Hall, MP., Unch, J., Binkowski, BF., Valley, MP., Butler, BL., Wood, MG., Otto, P., Zimmerman, K., Vidugiris, G., Machleidt, T., Robers, MB., Benink, HA., Eggers, CT., Slater, MR., Meisenheimer, PL., Klaubert, DH., Fan, F., Encell, LP., Wood, KV. (2012). Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem Biol* 7:1848–1857.
- Harvey, EN. (1956). Evolution and bioluminescence. *The Quarterly Review of biology* 31 (4):270-287.
- Hashimoto T., Adams KW., Fan Z., McLean PJ., Hyman BT. (2011) Characterization of oligomer formation of amyloid-beta peptide using a split-luciferase complementation assay. *J Biol Chem.* 2011 Aug 5; 286(31):27081-91. doi: 10.1074/jbc.M111.257378.
- Hastings, JW. (1998). Bioluminescence. In *Cell Physiology*, 2nd edn, Sperelakis N (ed.). Academic Press: New York; 984–1000.
- Hastings, JW. (1983). Biological diversity, chemical mechanisms and the evolutionary origins of bioluminescent systems. *Journal of Molecular Evolution* 19:309-321.
- Hattori M. and Ozawa T. (2014). Split Luciferase Complementation for Analysis of Intracellular Signalling. *Analytical sciences*, Vol.30(5), pages 539-544.
- Herring, PJ., Campbell, AK., Whitfield, M., Maddock, L. (1990). *Light and Life in the Sea*. Cambridge, UK: Cambridge Univ. Press. 365 pp.
- Herring, PJ., Cope, C. (2005). Red bioluminescence in fishes: on the suborbital photophores of *Malacosteus*, *Pachystomias* and *Aristostomias*. *Mar. Biol.* 148:383–94
- Herring, PJ. (2007). Sex with the lights on? A review of bioluminescent sexual dimorphism in the sea. *J. Mar. Biol. Assoc. UK* 87:829–42
- Herring, PJ., Campbell, AK., Whitfield, M., Maddock, L. (1990). *Light and Life in the Sea*. Cambridge, UK: Cambridge Univ. Press. 365 pp.
- Herring, PJ. (1987). Systematic distribution of bioluminescence in living organisms. *Journal of Bioluminescence and Chemiluminescence* 1:147-163
- Inouye, S. and Suzuki, T. (2016) Protein expression of preferred human codon-optimized *Gaussia* luciferase genes with an artificial open-reading frame in mammalian and bacterial cells. *Protein Expr Purif.* Dec;128:93-100. doi: 10.1016/j.pep.2016.08.006.

- Inouye, S., Sahara-Miura, Y., Sato, J., Iimori R., Yoshida, S., Hosoya, T. (2013) Expression, purification and luminescence properties of coelenterazine-utilizing luciferases from *Renilla*, *Oplophorus* and *Gaussia*: comparison of substrate specificity for C2- modified coelenterazines, *Protein Expr. Purif.* 88 150–156.
- Inouye, S. and Sahara, Y. (2008) Identification of two catalytic domains in a luciferase secreted by the copepod *Gaussia princeps*. *Biochem Biophys Res Commun.* 365, 96-101.
- Inouye, S. and Suzuki, T. (2016). Protein expression of preferred human codon-optimized *Gaussia* luciferase genes with an artificial open-reading frame in mammalian and bacterial cells. *Protein Expression and Purification*, 128, pp.93-100.
- Lee, JY., Kim, S., Hwang, W., Jeong, JM., Chung, JK., Lee, MC., Lee DS. (2008). Development of a dual-luciferase reporter system for *in vivo* visualization of MicroRNA biogenesis and posttranscriptional regulation, *J. Nucl. Med.* 49 285–294.
- Lorenz W.W., McCann R.O., Longiaru M., Cormier M.J. (1991) Isolation and expression of a cDNA encoding *Renilla reniformis* luciferase, *Proc. Natl. Acad. Sci. USA* 88 4438–4442.
- Johnsen, S., Balsler, E., Fisher, E.C. and Widder, EA. (1999). Bioluminescence in the deep-sea cirrate octopod *Stauroteuthis syrtensis* Verrill (Mollusca: Cephalopoda). *The Biological Bulletin* 197(1) 26-39.
- Kim SB, Sato M, Tao H. (2009). Split *Gaussia* luciferase-based bioluminescence template for tracing protein dynamics in living cells. *Anal. Chem.* 81:67–74.
- Kim, SB., Suzuki, H., Sato, M. and Tao, H. (2011). Superluminescent variants of marine luciferases for bioassays, *Anal. Chem.* 83, 8732–8740.
- Kim, SB. (2012). Labor-effective manipulation of marine and beetle luciferases for bioassays. *Protein Eng Des Sel* 25:261–269
- Knappskog, S., Ravneberg, H., Gjerdrum, C., Trosse, C., Stern, B. and Pryme, I.F. (2007) The level of synthesis and secretion of *Gaussia princeps* luciferase in transfected CHO cells is heavily dependent on the choice of signal peptide. *J. Biotechnol.* 128, 705-715.
- Koopmann, R., Cupelli, K., Redecke, L., Nass, K., DePonte, D. P., White, T. A., Duszenko, M. (2012). *In vivo* protein crystallization opens new routes in structural biology. *Nature Methods*, 9(3), 259–262. <http://doi.org/10.1038/nmeth.1859>.
- Kricka LJ. (1991). Chemiluminescent techniques, *Clinical Chemistry.* 37(9): 1472-81.
- Law, EGH., Gandelman, OA., Tisi, LC., Loebe, ChR. and Murray, JAH. (2006). Mutagenesis of solvent-exposed amino acids in *Photinus pyralis* luciferase improves thermostability and pH-tolerance. *Biochem. J.* 397, 305–312.
- Li F., Yu J., Zhang Z., Cui Z., Wang D., Wei H., Zhang X.E. (2012). Buffer enhanced bioluminescence resonance energy transfer sensor based on *Gaussia* luciferase for *in vitro* detection of protease, *Anal. Chim. Acta* 724 104–110.
- Li L. and Ismagilov RF. (2010) Protein crystallization using microfluidic technologies based on valves, droplets and Slip Chip. *Annu Rev Biophys* 39: 139-158. Doi:10.1146/annurev.biophys. 050708.133630

- Lloyd, J. E. (1983). Bioluminescence and communication in insects. *Annual Review of Entomology* 28:131-60
- Lorenz, WW., McCann, RO., Longiaru, M., and Cormier, MJ. (1991). Isolation and expression of a cDNA encoding *Renilla reniformis* luciferase. *Proc Natl Acad Sci U S A* Vol.88, No.10: pp.4438-4442.
- Luker, K. E., and Luker, G. D. (2014). Split Gaussia Luciferase for Imaging Ligand–Receptor Binding. *Methods in Molecular Biology* (Clifton, N.J.), 1098, 59–69. http://doi.org/10.1007/978-1-62703-718-1_5.
- Luker, G. and Luker, K. (2011). Luciferase protein complementation assays for bioluminescence imaging of cells and mice. *Methods Mol Biol.*; 680:29–43.
- Lundin A. (2014). Optimization of the Firefly Luciferase Reaction for Analytical Purposes. *Bioluminescence: Fundamentals and Applications in Biotechnology-Volume 2*. Berlin, Heidelberg: Springer Berlin Heidelberg 2014:31-62.
- Maguire, CA., Deliolanis, NC., Pike, L., Niers, JM., Tjon-Kon-Fat, LA, Sena-Esteves, M., Tannous, BA. (2009). Gaussia luciferase variant for high- throughput functional screening application. *Anal Chem* 81:7102–7106.
- Matthews, JC., Hori, K., Cormier, MJ. (1977). Purification and properties of *Renilla reniformis* luciferase. *Biochemistry* 16:85–91.
- Meyer-Rochow, VB. (2001). Light of my life-messages in the dark. *Biologist* 48 (4): 163-167.
- Michellini, E., Cevenini, L., Mezzanotte, L., Ablamsky, D., Southworth, T., Brachini, BR. and Roda, A. (2008). Combining intracellular and secreted bioluminescent reporter proteins for multicolor cell-based assays. *Photochem. Photobiol. Sci.* 7, 212-217.
- Morin, JG. (1983). Coastal bioluminescence: Patterns and functions. *Bull. Mar. Sci.* 33: 787–817.
- Nan, Wu., Rathnayaka, T., Kuroda, Y., (2015). Bacterial expression and re-engineering of Gaussia princeps luciferase and its use as a reporter protein
- Neufeld, HA. and Towner, RD. (1975). A rapid method for determining ATP by the firefly luciferin–luciferase system. *Experientia*; 31(3): 391–392.
- Nishihara, R., Abe, M., Nishiyama, S., Citterio, D., Suzuki, K., Kim, SB. (2017). Luciferase-Specific Coelenterazine Analogues for Optical Contamination-Free Bioassays. *Sci Rep.* Apr 19; 7(1):908. doi: 10.1038/s41598-017-00955-6.
- Page, R., Peti, W., Wilson, I. A., Stevens, R. C., and Wüthrich, K. (2005). NMR screening and crystal quality of bacterially expressed prokaryotic and eukaryotic proteins in a structural genomics pipeline. *Proceedings of the National Academy of Sciences of the United States of America*, 102(6), 1901-1905.
- Plieth, C. (2006). Aequorin as a reporter gene. *Methods Mol. Biol. (Methods in molecular biology* (Clifton, N.J.), 323, 307–327. PMID: 16739587 In: Salinas J; Sanchez- Serrano JJ. (Eds.) “*Arabidopsis Protocols*” 2nd Edition, Humana Press-ISBN 1-58829-395-5.

- Rathnayaka, T., Tawa, M., Nakamura, T., Sohya, S., Kuwajima, K., Yohda, M. and Kuroda, Y. (2011). Solubilization and folding of a fully active recombinant *Gussia luciferase* with native disulfide bonds by using a SEP-Tag. *Biochim. Biophys. Acta* 1814, 1775-1778.
- Remy, I. and Michnick, S.W. (2006). A highly sensitive protein-protein interaction assay based on *Gussia luciferase*. *Nat. Methods*, 3, 977-979.
- Roda, A., Guardigli, M., Michelini, E., & Mirasoli, M. (2009). Bioluminescence in analytical chemistry and *in vivo* imaging. *TrAC Trends in Analytical Chemistry*, 28(3), 307-322. Elsevier Ltd. doi: 10.1016/j.trac.2008.11.015.
- Rupp, B. (2007). *Biomolecular crystallography: Principles, practice, and application to Structural Biology* (1st ed. Taylor & Francis Ltd).
- Saleh, L. and Plieth, C. (2010). A coelenterazine-based luminescence assay to quantify high-molecular-weight superoxide anion scavenger activities. *Nature protocols*, 5, 1635–1641. PMID: 20885375.
- Schönherr, R., Klinge, M., Rudolph, J. M., Fita, K., Rehders, D., Lübber, F., Schneegans, S., Duszenko, M., Betzel, C., Martinez-Costas, J., Duden, R., Redecke, L. (2015). Real-time investigation of dynamic protein crystallization in living cells. *Structural Dynamics*, 2(4), 041712.
- Shiba, K., Niidome, T., Katoh, E., Xiang, H., Han, L. and Mori, T. (2010). Polydispersity as a Parameter for Indicating the Thermal Stability of Proteins by Dynamic Light Scattering. *Anal. Sci.* 26, 659-663.
- Shimomura, O. (2006). *Bioluminescence chemical principles and methods*. World Scientific Publishing.
- Shimomura, O., Johnson, F., and Kohama, Y. (1972). Reactions Involved in Bioluminescence Systems of Limpet (*Latia neritoides*) and Luminous Bacteria. *Proc. Nat. Acad. Sci. USA*. Vol. 69, No. 8, pp. August 2086-2089.
- Shimomura, O., Johnson, F. (1972). Structure of the light-emitting moiety of aequorin. *Biochemistry* 11:1602–1608.
- Shimomura, O., Johnson, F.H. (1975). Chemical nature of bioluminescence systems in coelenterates. *Proc. Natl Acad. Sci. USA*; 72(4): 1546–1549.
- Smith, L.M., Kelleher, N.L. (2013). Proteoform: a single term describing protein complexity. Consortium for P. Top Down. *Nat. Methods* 10, 186–187
- Smith, S.A., de Alwis, A.R., Kose, N., Harris, E., Ibarra, K.D., Kahle, K.M., Crowe, J.E. (2013). The Potent and Broadly Neutralizing Human Dengue Virus-Specific Monoclonal Antibody 1C19 Reveals a Unique Cross-Reactive Epitope on the bc Loop of Domain II of the Envelope Protein. *mBio*, 4(6), e00873–13.
- Sverdrup, H., Johnson, M. and Fleming R. (1942). *The Oceans: Their Physics, Chemistry and General Biology*. Englewood Cliffs: Prentice-Hall Inc.
- Takenaka, Y., Masuda, H., Yamaguchi, A., Nishikawa, S., Shigeri, Y., Yoshida, Y., Mizuno, H. (2008). Two forms of secreted and thermostable luciferases from the marine copepod

- crustacean, *Metridia pacifica*. *Gene*. 2008 Dec 1; 425(1-2): 28–35. doi: 10.1016/j.gene.2008.07.041
- Takenaka, Y., Yamaguchi, A., Tsuruoka, N., Torimura, M., Gojobori, T., Shigeri, Y. (2012). Evolution of bioluminescence in marine planktonic copepods. *Mol Biol Evol* 29:1669–1681
- Tang, BL. and Teng, FYH. (2005). Concepts of protein sorting or targeting signals and membrane topology in undergraduate teaching. *Biochem. Mol. Biol. Educ.*, 33: 188–193. doi:10.1002/bmb.2005.494033032448
- Tannous BA., Kim DE. , Fernandez JL. , Weissleder R. , Breakefield X.O. (2005). Codon-optimized *Gaussia* luciferase cDNA for mammalian gene expression in culture and *in vivo*. *Mol Ther.*;11(3):435–43.
- Tannous, BA. (2009). *Gaussia* luciferase reporter assay for monitoring biological processes in culture and *in vivo*. *Nature protocols*, 4(4), 582-91. doi: 10.1038/nprot.2009.28.
- Thomson, CM., Herring, PJ, and Campbell, AK. (1997). The widespread occurrence and tissue distribution of the imidazolopyrazine luciferins. *Journal of Bioluminescence and Chemiluminescence* 12:87-91.
- Thomson, CM., Herring, PJ., Campbell, AK. (1997). The widespread occurrence and tissue distribution of the imidazolopyrazine luciferins. *J. Biolum. Chemilum.* 12:87–91
- Tinn, O., Oakley, T. (2008). Erratic rates of molecular evolution and incongruence of fossil and molecular divergence time estimates in Ostracoda (Crustacea). *Mol. Phylogenet. Evol.* 48:157–67
- Valiadi, M., Iglesias-Rodriguez, D. (2013). "Understanding Bioluminescence in Dinoflagellates—How Far Have We Come?" *Microorganisms* 1, no. 1: 3-25.
- Verhaegen, M. and Christopoulos, TK. (2002). Recombinant *Gaussia* luciferase. Overexpression, purification, and analytical application of a bioluminescent reporter for DNA hybridization. *Anal. Chem.* 74, 4378-4385.
- Voss KJ. (2000). Physics of low light level detectors. In: Ziegler MM, Baldwin TO, eds. *Methods in Enzymology*. San Diego, London, Boston: Academic Press; 53-61.
- Waters, JF. and Millero, FJ. (2013). The free proton concentration scale for seawater pH. *Marine Chemistry*, 149, 8-22. DOI: 10.1016/j.marchem.2012.11.003.
- Welsh, DK and Kay, SA. (2005). Bioluminescence imaging in living organisms, *Curr. Opin. Biotechnol.* 16,73–78.
- Welsh, JP., Patel, KG., Manthiram, K. and Swartz JR. (2009) Multiply mutated *Gaussia* luciferases provide prolonged and intense bioluminescence, *Biochem. Biophys. Res. Commun.* 389, 563–568.
- Widder, E.A. (2001). Bioluminescence and the pelagic visual environment. *Marine and Freshwater Behaviour and Physiology* 35: 1-26.
- Wiles, S., Ferguson, K., Stefanidou, M., Young, DB., Robertson, BD. (2005). Alternative Luciferase for Monitoring Bacterial Cells under Adverse Conditions. *Applied and Environmental Microbiology*. 71(7):3427-3432. doi:10.1128/AEM.71.7.3427-3432.2005.

- Wilson, T. and J. W. Hastings (1998). Bioluminescence. *Annual Review of Cell and Developmental Biology* 14:197-230.
- Wurdinger, T., Badr, C., Pike, L., de Kleine, R., Weissleder, R., Breakefield, YO, Tannous, BA. (2008). A secreted luciferase for ex vivo monitoring of *in vivo* processes. *Nat Methods* 5(2): 171–173. doi:10.1038/nmeth.1177.