UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Novel Fluorescent Cell-Based Sensors for Detection of Viral Pathogens

Daniela Filipa da Cruz Freitas

Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil em Biofísica Médica e Fisiologia de Sistemas

Dissertação orientada por:
Doutor Nuno Miguel Matela
Doutora Ana Sofia de Sousa Valente Coroadinha

2017

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Novel Fluorescent Cell-Based Sensors for Detection of Viral Pathogens

Daniela Filipa da Cruz Freitas

Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil em Biofísica Médica e Fisiologia de Sistemas

Dissertação orientada por:
Doutor Nuno Miguel Matela
Doutora Ana Sofia de Sousa Valente Coroadinha

ACKNOWLEDGEMENTS

First, I would like to express my gratitude to all people directly or indirectly involved in this thesis.

To Dr. Paula Alves, for giving me the opportunity to do my master thesis at Animal Cell Technology Unit at IBET/ITQB NOVA, for the good working conditions offered and for being a strong example of leadership.

To Dr. Ana Sofia Coroadinha, my external advisor, for giving me the opportunity to join the Cell Line Development and Molecular Biotechnology Laboratory, the supervision and availability.

To Dr. Nuno Matela, my internal advisor, for the guidance and help always demonstrated during my course.

A special acknowledgment to Miguel Guerreiro, for his guidance, constant support and encouragement, patience throughout the time and help. It was such a great pleasure to share this last year with such a hard-working scientist. I learnt so much with you, I don't have words to thank you for that.

To all my Animal Cell Technology Unit colleagues, for good work environment and help. A special thanks to all my colleagues of Cell Line Development and Molecular Biotechnology Lab for all the encouragement in the good and bad moments, fruitful discussions and help. To Daniel Mestre and Tiago Vaz for all the support, laughs and always having the right words to say.

Um grande e especial agradecimento a toda à minha família, por todo o apoio e paciência, por acreditarem sempre em mim e, acima de tudo, por não permitirem que eu deixasse de acreditar em mim. Ao Pedro, por todo o carinho e apoio incondicional, por estares sempre ao meu lado nos bons e maus momentos. Um simples obrigado não é o suficiente para agradecer tudo o que fizeram por mim todos estes anos.

Aos meus amigos, Inês Palma, João Mendes, Joana Dias e Raquel Maio, pelo companheirismo e amizade, por todo o apoio e ajuda durante estes anos. You're the best!

Às minhas amigas de sempre, que a distância nunca irá afastar. Tenho-vos a agradecer todo apoio e encorajamento, compreensão, amizade e por todos os momentos engraçados. Sempre juntas!

A todos, um grande obrigada!

ABSTRACT

Viruses are major pathogenic agents that can cause a variety of serious diseases. Indeed, the establishment of cell culture techniques and recombinant virus manipulation contributed for the development of viral-based biotherapies, like gene therapy or vaccines, which require accurate and fast quantification of virus. Despite the numerous titration methods existing nowadays, the majority of them are not able to provide a robust and fast quantification, essential for their clinical application. Moreover, most of them provide indirect measurements of infectious particles, over-estimating virus infectivity, and some rely on virus modification, e.g. by making use of reporter genes (labelled-viruses), which are not allowed when using those virus for clinical applications. As so, the development of a new system capable to cope these drawbacks is of paramount importance for research, diagnostics and industry.

In this work, genetically encoded switch-on fluorescent mammalian cell-based assays for detection and quantification of label-free Adenoviruses, using the adenoviral protease (Adenain) as a trigger of the sensor, were developed. Three different main strategies were designed based on structural distortion of a fluorescent protein (GFP – Green Fluorescent Protein), preventing fluorescence emission: GFP VISENSOR (cGFP), Embedded Split-GFP VISENSOR (eS11) and Circular Split-GFP VISENSOR (cS11). Upon Adenain proteolytic processing, structural distortion is relieved and fluorescence emission is reconstituted.

VISENSORS performance was assessed by optimizing the best combination of backbone structure and cleavage site, initially by a transient screening and later confirmed on a more biological context, where cells stably expressing the sensor were infected by human Adenovirus serotype 5. Despite eS11 and cGFP displaying similar signal to noise ratio (SNR) performances, cS11 strategy seems the most promising, reaching a signal to noise ratio of 3.12 at 72 hours post-infection. Virus detection was accomplished as soon as 24 hours post-infection in all strategies. Moreover, this work validated the use of VISENSORS as an Adenain dependent sensor and specific for Adenovirus. An attempt to reach maximum distortion and improving SNR performances, a parallel strategy was implemented by structurally distorting both split-GFP fragments. However, the results were not promising. A detailed characterization of the best strategy will be performed as future work, using cell clones stably expressing the sensor to assess VISENSORS' applicability to Adenovirus quantification.

VISENSORS show great potential to deliver a fast, reliable and accurate method for virus and viral vector detection and quantification, much needed not only in the development of viral based-biotherapies, but also for diagnostic and clinical applications.

Keywords: Virus detection and quantification, Adenovirus, Label-free virus, Cell based-sensors, Biosensors

RESUMO

Os vírus constituem um dos principais agentes patogénicos responsáveis por uma grande variedade de doenças graves. O estabelecimento de técnicas de cultura de células e a manipulação génica de vírus contribuíram decisivamente para o desenvolvimento de bioterapias baseadas em partículas virais, como a terapia génica ou a vacinação, que requerem uma quantificação precisa e rápida da carga viral. Apesar dos inúmeros métodos de titulação existentes hoje em dia, a sua maioria não é capaz de fornecer uma quantificação robusta e rápida, essencial para sua aplicação clínica. Além disso apresentam uma série de desvantagens, tais como o facto de na sua maioria fornecerem medidas indiretas do número de partículas infeciosas, sobrestimando a infecciosidade do vírus; são morosos e com reduzido potencial de processamento rápido; outros dependem de modificação génica, envolvendo por exemplo o uso de genes repórter (vírus com marcação), o que não é permitido em aplicações clínicas. Como tal, os biossensores virais representam uma excelente alternativa aos métodos de titulação tradicionais, sendo amplamente utilizados em biomedicina no diagnóstico de doenças infeciosas e desenvolvimento de medicamentos. De todos os biossensores existentes, os sensores baseados em células constituem como uma das estratégias mais promissoras, devido à capacidade das células em responder às mudanças ambientais externas de forma rápida e precisa. Tomando partido das células como elementos de deteção, torna-se possível o desenvolvimento de sensores de elevada especificidade e sensibilidade a um grande número de agentes externos, como os vírus, fornecendo assim uma medida direta da sua infecciosidade. Na construção de um biossensor, para além do detetor é também necessário um transdutor do sinal. Para tal, a fluorescência é amplamente usada como transdutor devido à sua alta sensibilidade e seletividade, suficiente resolução espaço-tempo e baixo custo. Assim, e tendo em conta a necessidade de desenvolver um novo sistema capaz de lidar com as desvantagens apresentadas pelos métodos de titulação atuais, uma estratégia que tire partido de proteínas fluorescentes associadas a células para deteção e quantificação de vírus sem marcação revela-se extremamente promissora.

Este trabalho teve como objetivo o desenvolvimento de um biossensor de fluorescência para deteção e quantificação de vírus sem marcação (de acrónimo *VISENSORS*), tomando partido da protease viral, componente responsável pela maturação e processamento das proteínas virais, como ativador do sensor. Como modelo de estudo, e para prova de conceito, foram usados Adenovírus devido à sua importância no desenvolvimento de vacinas e uso como vetor viral em aplicações de terapia génica. Três estratégias de biossensores foram implementadas tendo por base a distorção estrutural de uma proteína fluorescente, a *GFP*, incapacitando-a de emitir fluorescência até que seja aliviada a distorção pela protease adenoviral. A primeira estratégia (denominada *Circular VISENSOR*, cGFP) consiste na circularização da *superfolder-GFP*, tomando partido da singular característica das inteínas (porções de proteína) de se libertarem, fundindo os segmentos a qual se encontravam ligadas. Na tentativa de alcançar máxima distorção para um melhor desempenho do sensor, duas outras estratégias foram desenvolvidas tendo por base a transcomplementação da *split-GFP*, criando distorção estrutural apenas ao nível do fragmento S11 da GFP. Desta forma, a segunda estratégia (denominada *Embedded split-GFP*, eS11) consistiu na inserção do fragmento S11 no *loop* de uma pequena proteína. Por sua vez, a terceira estratégia (denominada *Circular split-GFP*, cS11) consistiu na circularização do fragmento S11, mediada por inteínas, de forma semelhante à cGFP. A fluorescência é reconstituída quando a protease reconhece uma sequência especifica de corte e, por consequência, alivia a distorção estrutural permitindo a transcomplementação dos fragmentos de GFP e a emissão de fluorescência.

A aplicação de diferentes distorções estruturais no biossensor pode alterar o seu desempenho quer ao nível da emissão de fluorescência quando não ativado ou após a sua ativação. Assim, para avaliar o desempenho do sensor começou por se otimizar a melhor combinação de estrutura de sensor e local de clivagem, primeiramente de forma transiente e posteriormente confirmando num contexto mais biológico, infetando com adenovírus células que expressam de forma estável os sensores. Os resultados obtidos confirmaram esta hipótese. Através da adição de glicinas (aminoácido de pequenas dimensões e elevada flexibilidade) verificou-se um aumento da emissão de fluorescência em todas as estratégias, possivelmente provocado por relaxamento da distorção estrutural. Por outro lado, a utilização de sequências de clivagem constituídas na sua maioria por aminoácidos mais hidrofóbicos demonstraram uma diminuição de fluorescência, contrariamente às sequências que possuem na sua constituição aminoácidos maioritariamente hidrofílicos. Sugere-se assim que as diferentes hidrofilicidades das sequências de corte podem ter um impacto ao nível da eficiência de clivagem por parte da protease adenoviral e/ou da correta maturação do cromóforo da proteína fluorescente. Surpreendentemente, em todas as estratégias, foi observado uma diminuição da emissão de fluorescência por parte das células não infetadas ao longo do tempo, o que pode dever-se ao fato das células não infetadas atingirem a fase estacionária (ao contrário de células infetadas cujo ciclo celular é interrompido) em que a expressão de proteína (tal como o sensor) diminui e, consequentemente, a emissão de fluorescência também diminui. Assim, para a performance Sinal/Ruído por parte de cada sensor é de se considerar não só uma contribuição do aumento de emissão de fluorescência devido à ativação do sensor pela protease adenoviral (Sinal), mas também uma contribuição da diminuição da fluorescência por parte das células não infetadas que são usadas como controlo negativo (Ruído). Comparando as três estratégias desenvolvidas, eS11 e cGFP demonstraram possuir desempenhos semelhantes ao nível da razão Sinal/Ruído, enquanto que a estratégia cS11 parece ser a mais promissora, atingindo uma razão Sinal/Ruído de 3.12 às 72 horas após a infeção. Adicionalmente, demonstrou-se que o sensor é especifico, sendo apenas ativado na presença da protease de Adenovírus e não de outras proteases virais.

Para além das três estratégias principais, foi desenvolvida em paralelo uma nova abordagem baseada na distorção estrutural de ambos os fragmentos da *split-GFP*. Esta foi alcançada através da circularização do fragmento S10 da *GFP* (cGFPS10) em combinação com as distorções realizadas nas estratégias eS11 e cS11. Os resultados obtidos, no entanto, não se mostraram promissores em parte devido à incompatibilidade das estratégias usadas. Por exemplo, a circularização dos fragmentos S11 e S10 da *GFP* (cS11 e cGFPS10) apresentou uma razão Sinal/Ruído menor do que a distorção apenas do fragmento S11 (cS11). Este fenómeno poderá ser explicado pelo uso em ambos os processos de circularização do mesmo tipo de inteínas, não evitando assim a trans-circularização de S11 e S10 e a formação de uma *GFP* funcional, sem ser necessária a ativação por parte da protease viral.

Este trabalho foi pioneiro na implementação em células de mamíferos de estratégias baseadas em distorção estrutural de proteínas fluorescentes como biossensores para a deteção e quantificação de Adenovírus sem marcação. Durante esta tese de mestrado foram implementadas, com sucesso, três diferentes estratégias para deteção de Adenovírus. A melhor das estratégias, a cS11, será agora alvo de uma caracterização detalhada, usando clones celulares a expressarem de forma estável o sensor por forma a validar as otimizações realizadas e a avaliar a sua aplicabilidade como método de quantificação de Adenovírus.

Os VISENSORS demonstram assim grande potencial tendo em vista o desenvolvimento de um método rápido, fiável e específico para a deteção e quantificação de vírus e vetores virais sem marcação, bem como uma enorme aplicabilidade não só no estudo e desenvolvimento de bioterapias baseadas em vírus, mas também com aplicações clínicas e em diagnóstico.

Palavras-chave: Deteção e quantificação de vírus, Adenovírus, Vírus sem marcação, Biossensores

PREFACE

This master thesis is within the scope of the project PTDC/EBB-BIO/118615/2010, PTDC/EBB-BIO/118621/2010 and SFRH/BD/90685/2012 (PhD grant); entitled “Development of fluorescent-cell based systems for the detection and quantification of label-free virus for research and diagnostics” funded by the Portuguese Fundação para a Ciência e Tecnologia (FCT).

LIST OF CONTENTS

FIGURE INDEX

TABLE INDEX

LIST OF ABBREVIATIONS

Abs	Absorbance
AdV	Adenovirus
AdV5	Human Adenovirus serotype 5
ANOVA	Analysis of variance
BSA	Bovine serum albumin
cGFP	Circularized superfolder-GFP strategy
CMV	Cytomegalovirus
CPE	Cytopathic effects
cS11	Circularized GFPS11 fragment strategy
DMEM	Dulbecco's modified eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
eGFP	Enhanced GFP
ELISA	Enzyme-linked immunosorbent assay
EMCV.IRES	Encephalomyocaerditis virus-internal ribosome entry site
eS11	Embedded GFPS11 fragment strategy
FBS	Fetal bovine serum
FDA	Food and Drug Administration
gag	Group-specific antigen gene
GaLV	Gibbon ape leukemia virus
GFP	Green fluorescent protein
GFPS10	S1-S10 fragment of Split-GFP
GFPS11	S11 fragment of Split-GFP
HCV	Hepatitis C virus
HEK	Human embryonic kidney
HIV-1	Human immunodeficiency virus type 1
h.p.i.	Hours post-infection
HSD	Honest significant difference
IF	Immunofluorescence
IFA	Immunofluorescence Foci Assay
I.P.	Infectious particles
IPS-1	Mitochondrially-tethered IFN-β promoter stimulator protein 1
ITR	Inverted terminal repeat
LB	Luria broth
LV	Lentivirus
MLV	Murine leukemia virus
MOI	Multiplicity of infection
M-PER	Mammalian protein extraction reagent
NAAT	Nucleic acid amplification tests
***Npu* DnaE**	*Nostoc punctiforme* DnaE
NTA	Nanoparticle tracking analysis
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PEI	Polyethylenimine

PFU	Plaque-forming units
PI	Propidium iodide
pol	Polymerase gene
pro	Protease gene
pVIc	Cleaved 11-residue peptide from C terminus of precursor protein VI
qPCR	Quantitative polymerase chain reaction
Rev	Regulator of expression of viral proteins
RNA	Ribonucleic acid
RRE	Rev responsive element
RSV	Rous sarcoma virus
RT-PCR	Reverse transcription-PCR
SEAP	Secreted alkaline phosphatase
sfGFP	Superfolder-GFP
SIN	Self-inactivated
SNR	Signal to noise ratio
SPR	Surface plasmon resonance
TAE	Tris-Acetate-EDTA
TBST	Tris buffered saline with Tween 20
TCID50	Tissue Culture Infectious Dose – 50
VBBs	Viral-based biopharmaceuticals
VLPs	Virus-like particles
VSV-G	Vesicular stomatitis virus G glycoprotein
WHO	World Health Organization

1. INTRODUCTION

1.1. AN HISTORICAL PERSPECTIVE

According to World Health Organization (WHO), one of the leading causes of death worldwide, especially in low income countries, is infectious diseases. These are caused by pathogenic microorganisms and can be spread directly or indirectly from one person to another. In 2010 over 15 million people died worldwide due to infectious diseases and it is expected that in 2050 the number of deaths continues to be a challenge with 13 million deaths predicted [1]. The prevalence of infectious diseases led to an increase in vaccine market since vaccination has been a key element in reducing mortality attributed to infectious pathogens. In 2014 the global vaccine market revenue hit the 32.2 billion USD mark and a growth of 27 billion USD is estimated until 2020 [2].

The establishment of cell culture techniques and recombinant virus manipulation contributed decisively to the development of viral-based biopharmaceuticals (VBBs). VBBs comprise virus-derived components or virus-based particles able to be used for therapeutic purposes, which includes viral vectors for gene therapy, oncolytic virotherapy and viral vaccines [3]. A vaccine consists of a pathogen-based preparation able to provide acquired immune protection against specific diseases and can be categorized into: live attenuated, inactived and subunit. Live attenuated vaccines make use of live mutated viruses, inept to cause any disease but still able to infect target cells. Inactivated vaccines use killed viruses treated by chemical (formalin or formaldehyde) or physical means (heat or radiation); therefore, the virus is no longer infectious but its immunogenicity persists. Subunit vaccines consist of parts of the virus comprising specific antigens known to induce high immune responses [4]. Recently, a novel type of subunit vaccines has been approved and commercialized, named virus-like particles (VLPs). VLPs combine virus-derived structural antigens with the lack of genetic material, which makes them non-infectious particles [4 5].

VBBs had also a huge impact on the development of gene therapy field. As stated by Jayanant Iemsam-Arng et al.[6], gene therapy can be described as a procedure aimed at replacing, manipulating or supplementing non-functional genes with healthy ones. For gene therapy success, an effective delivery system should be developed without causing major side effects. One extremely effective method for gene delivery takes advantage of viral vectors, since viruses have the ability to express the genes of interest in target cells, using cell's transcription equipment. Through genetic manipulation, viral vectors comprise only the transgene and the minimal number of genes responsible for viral production and replication, reducing virus infectivity [7 8]. The first steps in gene therapy started in the 70s and since then gene therapy has faced an exponential progress. Twenty years later, in the 90s, the first human gene therapy clinical trial for therapeutic purposes using viral vectors was approved by Food and Drug Administration (FDA) [9 10]. According to The Journal of Gene Medicine, in 2016 around 2400 approved clinical trials using viral vectors have been performed worldwide. Adenovirus (AdV), adeno-associated virus, lentivirus (LV) and retrovirus are the most used viral vectors in gene therapy clinical trials (**Figure 1.1**). Specifically, AdV has been the most studied viral vector in gene therapy thanks to their efficient transduction and transgene expression [11 12]. However, these viruses can induce high immune responses in humans, which might have contributed in the last years to an increased interest in LV vectors (**Figure 1.1**). In 2004 the State and Food Drug Administration of China approved the commercialization of the Gendicine®, a gene therapy product based on AdV

serotype 5 vector with the therapeutic gene *p53*, for the treatment of head and neck squamous cell carcinoma [13] [14].

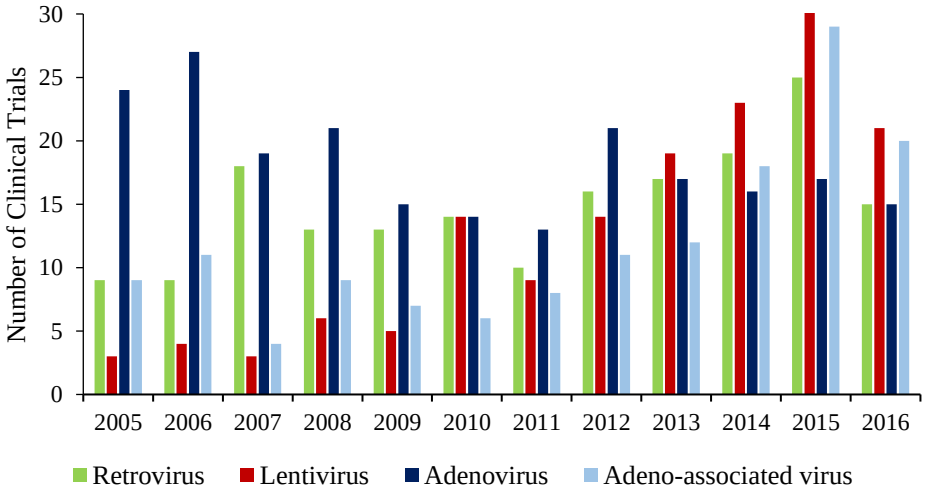

Figure 1.1 - Viral vectors used in gene therapy clinical trials. Data gathered from: The Journal of Gene Medicine, 2016 John Wiley and Sons, Ltd, www.wiley.co.uk/genmed/clinical [Accessed: 08-Aug-2017].

However, and despite the current success of gene therapy, its application must be carefully planned. Indeed, in 1999 gene therapy has experienced a terrible incident with the death of an 18 years old patient. The boy was erroneously administrated with a high titer of AdV vectors, which generated a high immune response leading to his death four days later by multi organ failure [10]. This incident showed that dosage is of paramount importance for a safe and successful application of gene therapy protocols; as such, sensitive and trustworthy titration methods are needed to assure it.

1.2. METHODS FOR VIRUS DETECTION AND QUANTIFICATION

Clinical and research applications based on viruses and viral vectors require reliable, accurate and fast methods for virus detection and quantification. For example, the study of new vaccines designs makes use of virus titration to estimate vaccine efficiency. In gene therapy, it is mandatory having a number of viral genome copies on target cells among a defined and reproducible range. If not, it could lead into an inadequate therapeutic effect or reinforce serious side effects.

Current titration methods for viral vectors can be categorized into: functional and non-functional [15]. Functional methods provide information concerning virus functionality, resulting in a direct estimation of virus infectivity. Virus quantification can be based, for example, on number of viral genome copies on target cells (e.g. Quantitative Polymerase Chain Reaction - qPCR), expression of reporter genes (e.g. Green Fluorescent Protein, GFP) or cytopathic effects (CPE) (e.g. Tissue Culture Infectious Dose 50%, TCID50). Non-functional titration methods provide an indirect estimation of the infectious viruses, which often lead to an over-estimation of virus infectivity. This estimation depends, for example, on the virus component to be measured in the viral preparations. The number of viral genome copies in viral preparations (e.g. qPCR) is an example of such indirect titration methods. Below are described some of the assays widely-used for virus detection and quantification.

1.2.1. DETECTION AND QUANTIFICATION OF INFECTIOUS VIRUS UNITS

Cell culture virus isolation was one of the first methods for virus detection and for many years it was considered an attractive method for the detection of viral human pathogens *in vitro* [16]. Some viruses can cause morphologic changes in host cells, which are named CPE. The virus identification is accomplished by observing differences on host cells during virus culture, such as rounding cells or texture changes [17]. The plaque assay is a widely-used virus titration method providing a direct and quantitative measurement of virus titer based on the number of plaque-forming units (PFU), specific for lytic viruses [18 19]. This assay consists on infecting a monolayer of cells with serial viral dilutions. Infected cells will generate a hole (or plaque) due to the cell lysis promoted by the virus infection. Cells are then fixed and stained and virus titer is determined by counting the number of plaques [17 20]. The main advantages of the use of this method stands on the ease to implement, it's inexpensive and doesn't require advanced materials. Nevertheless, this approach is laborious, time-consuming (could take several days for plaques to be formed) and is operator error-prone. Moreover, only viruses capable of generating plaques on host cells are suitable to be used. TCID50 it's another example of an end point assay performed to determine the dilution able to produce CPE in 50% of the seeded cells. Cells are plated in well-plates and infected with serial dilutions of a viral preparation. Virus titer is calculated by determining the last CPE positive dilution, based on the number of positive wells [17 20 21]. It's also time-consuming (~10 days), and high error due to operator visual inspection. Again, only viruses able to cause CPE can be quantified with this titration method.

Cell culture based techniques can be combined with immunofluorescence (IF) staining for a faster and trustworthy detection. IF consists of a biological assay which makes use of antibodies labelled with a fluorescent dye (also named fluorophore) for detection of specific target antigens under a fluorescent microscope [22 23]. So, IF is a fast (requiring only 20 to 30 minutes), versatile, sensitive and specific method for viral identification. Still, shows some limitations such as variability, because of non-specific binding or cross-reactivity of antibodies, and it is very expensive due to the cost of the antibodies [17]. Another extensively used virus infectivity based method is named immunofluorescence foci assay (IFA). IFA stands as a suitable alternative for the titration of virions in cell lines which do not form plaques or any CPE on host cells, for example the human immunodeficiency virus type 1 (HIV-1) [24]. This technique consists on a modification of the plaque assay, by making use of antibody based staining similar to IF [17 20]. IFA shows improved sensitivity and speed (2 days) regarding other TCID50 and plaque assays; however, introduces the same disadvantages as IF assay.

The use of reporter genes for the detection and quantification of virus pathogens is extensively used and easy to perform. Several studies report the use of reporter genes for detection and quantification of multiple virus, such as influenza virus [25], retrovirus [26] or herpes simplex virus [27]. In research laboratories, such as in the Cell Line Development and Molecular Biotechnology Laboratory (part of ACT Unit IBET/ITQB NOVA), recombinant LV vectors harbouring a fluorescence protein reporter as transgene are regularly titrated by means of flow cytometry. This method allows determining the number of infectious particles depending on the percentage of GFP-positive infected cells. The use of reporter genes provides a fast and sensitive method for detection and quantification of viruses. However, one of the major limitations of this method is the use of reporter genes, which are not tolerated in viral vectors for clinical use, due to safety issues. Additionally, it requires previous virus or viral vector modification, which, besides being time-consuming can alter virus biology and life cycle.

1.2.2. QUANTIFICATION OF VIRAL NUCLEIC ACID

Advances in technology have provided new and better tools for the detection of virus, like the introduction of molecular techniques able to detect viral nucleic acid. Polymerase chain reaction (PCR) consists on an enzyme-mediated *in vitro* assay, where the deoxyribonucleic acid (DNA) polymerase exponentially amplifies a target sequence of DNA for various cycles, using specific primers [28]. PCR led to important improvements on the field of molecular diagnosis, especially on the development of nucleic acid amplification tests (NAAT). NAAT comprises several amplification techniques widely used for virus detection, such as nucleic acid sequence-based amplification, transcription-mediated amplification, qPCR and reverse transcription-PCR (RT-PCR). The latter assays are usually used for assessing viral titers [29]. The major advantage of NAAT is that they require low amounts of viral nucleic acid or even viable viruses to detect viral pathogens. In addition, NAAT have shown higher analytical and clinical sensitivity and specificity than cell culture or IF methods [16]. qPCR method is also usually used to determine the number of integrated proviral DNA copies in target cells, by labelling the amplified DNA portion with fluorescence [15 30]. This technique is widely used for assessing virus load because it's a fast, sensitive, high-throughput and reproducible method. However, imperfect sample purification can lead to PCR inhibition. Besides the quantification of viral DNA, NAAT can also be used for the quantification of viral ribonucleic acid (RNA) by adding a reverse transcriptase enzyme to convert RNA to DNA. The combination of these two approaches, as in RT-PCR, can improve the detection sensitivity [31].

1.2.3. QUANTIFICATION OF VIRUS PROTEINS

The specific relation between antibody and antigen has become a key element to design a diversity of antiviral antibodies detection and quantification assays, for high throughput analysis of several samples, such as immunoblotting and enzyme-linked immunosorbent assay (ELISA).

Immunoblotting techniques (also known as Western-Blotting) make use of electrophoresis for the separation of proteins in polyacrylamide gels. The proteins are then transferred onto chemically resilient membranes and incubated with a primary antibody specific for the protein target. Later, the membrane is incubated with the secondary antibody combined with a marker compound for ease detection [32]. Thus, this method allows the detection of single viral proteins from infected cell lysates. This technique is one of the most used confirmatory test for HIV-1 [33] and hepatitis C virus (HCV) [34] infection diagnosis, for example.

ELISA is a widely-used method for virus detection and quantification, for example for the diagnose of HIV-1 infection [33]. There are two main variations of this technique, regarding antigen immobilization and detection: direct or indirect. The antigen is immobilized either directly or by a specific antibody (capture antibody). If the primary antibody is conjugated with an enzyme, it's called direct ELISA. On the other hand, if the primary antibody needs to be combined with an enzyme-conjugated second antibody, it's named indirect ELISA. The detection of the virus is allowed thanks to the interaction of the enzyme and a substrate, which leads to color formation. The viral quantification is made by the measurement of the optical density, which is proportional to the quantity of antigen [35]. Overall, this assay is simple, versatile, sensitive and specific. However, in some cases is expensive, time-consuming and presents additional variability, due to cross-reactivity and non-specific biding.

1.2.4. QUANTIFICATION OF VIRUS PARTICLES

Relating particle size and Brownian motion, nanoparticle tracking analysis (NTA) method is able to determine particle size distribution on viral preparations by laser light scattering microscopy coupled with a camera [36]. Kramberger et al. [37] showed fast estimation of virus concentration in different samples of AdV and influenza virus, however AdV titer was underestimated. So, this method is fast and allows sample visualization during the quantification process; still, provides an indirect measure of infectious particles because it quantifies the number of total virus particles in the sample.

1.3. BIOSENSORS FOR VIRUS DETECTION AND QUANTIFICATION

Conventional systems like the ones presented above have several disadvantages: the majority of them provide indirect measurements of infectious particles, which results in an over-estimation of the virus infectivity; and may rely on vector modification, by making use of reporter genes (labelled-viruses), which is not allowed when using those vectors for clinical trials. Therefore, there's an evident need for the development of simple, fast and reliable quantification methods for label-free viruses. As so, virus biosensors already represent a promising alternative for traditional titration methods extensively used in biomedicine to diagnose infectious diseases and drug development. A biosensor consists of a device able to convert a biological response into an electrical signal, by a transducer [38]. Usually, it involves three basic components: a biological recognition component (enzyme, antibody, DNA, etc.), a sensor element for signal acquisition (electrical, optical and thermal) and an element for amplification/signal processing [39]. The difference between a biosensor and a regular sensor is the presence of bio-recognition elements (also referred as affinity reagents).

Virus biosensors can be classified according to their affinity reagents and viral targets into: immuno- (antibody-), DNA-, antigen- and cell-based biosensor [40]. Immunosensing assays lean on the specific interactions between antigens and antibodies to get a quantifiable response. Antibodies are produced by the host as an immune response to the presence of viral pathogens, representing one of the most widely used affinity reagent in viral biosensors [41]. DNA-biosensors operation is based on nucleic acid hybridization [42]. Recent studies have demonstrated that the use of peptide nucleic acids exhibit high stability and can accomplished a fast, specific and strong hybridization on viral DNA detection [43] [44]. Antigen-based biosensors rely on the detection of whole virus particles and surface antigens, such as the capsid and envelope [40]. Like other serological methods, their performance depends on the number of antibodies produced during the infection. Of all existing virus biosensors, cell-based sensors are the most promising approach. Living cells are capable to preserve vital functions by responding to external environmental changes quickly and precisely. Taking advantage of living cells as sensing elements offers the chance to develop sensors with high specificity and sensitivity to a vast number of external agents, like viruses, responsible for affecting cells' activity [45] [46]. Additionally, cell-based sensors can provide a direct measure of virus infectivity.

There are several transducers utilized on biosensors, which can be divided into three major groups: electrochemical, piezoelectric and optical. Electrochemical biosensors are often used as virus detection methods because they offer a fast, high sensitive, practical, selective and economic detection [47]. The bio-recognition process of these biosensors can provide different responses and according to these, biosensors based on electrochemical transducers are categorized into: conductometry, potentiometry, voltammetry, and impedance [40]. The latter ones are the most

used for virus detection. Voltammetric biosensors measure the current on varying the potential [40]. Several of voltammetric techniques have been successfully used for virus detection, for example differential pulse voltammetry [48], square wave voltammetry [49] and cyclic voltammetry [50] [51]. Impedance-based biosensors have also been extensively used for viral pathogens detection [52] [53] [54] [55]. They measure the electrical opposition to current flow at the interface [56]. Therefore, impedance biosensors are a very flexible method, thanks to their ability to detect a diversity of targets just by changing the probe used [57]. Additionally, this approach, also named electrochemical impedance spectroscopy, provides sensitive and non-invasive measurements.

Optical biosensors turned out to be a promising approach for a fast, sensitive and real-time monitoring method for virus detection. Surface plasmon resonance (SPR) is one of the most developed label-free and real time detection methods and relies on the variation of the refractive index of the sample [58]. This biosensor owns a series of applications in the detection of several virus, such as hepatitis B virus [59], HIV-1 [60] and dengue virus [61]. Despite the advantages, some limitations should be considered. For example, small molecules generate a weak signal, hampering the differentiation of true signal from noise. In addition, the interpretation of the binding kinetics is not always direct, because in some cases the published biosensor data does not match to the simple bimolecular interaction binding model, which questions the biosensor validity [62]. Fluorescence-based biosensors comprise another optical detection biosensor system. Thanks to its high sensitivity and selectivity, sufficient spatiotemporal resolution and low cost, fluorescence is extensively used in biological imaging [63]. Mainly, this method measure the fluorescent signals from fluorophores conjugated to bio-recognition elements or viral targets [40]. Some of the existing fluorescence biosensors are based on fluorescence resonant energy, which results on great sensitivity for virus detection and quantification [64] [65]. This phenomenon takes place only when two fluorophores are in suitable proximity between them and the excitation spectrum of the acceptor is overlapped by the emission spectrum of the donor, resulting on a non-contact, radiationless and distance-dependent energy transfer method [40] [63].

1.3.1. FLUORESCENT CELL-BASED SENSORS

Recent efforts have been made for the development of fluorescent cell-based sensors, by combining fluorescence elements with molecular biology techniques and protein chemistry. Fluorescent proteins can absorb light and emit it in a longer wavelength. One of the most widely used reporter protein is the GFP from jellyfish *Aequorea victoria*. This well-known fluorescent protein comprises 238 amino acid residues and two peaks of maximum absorption at 395 and 475 nm. On top of that, GFP exhibits the unique property of generating internally the chromophore by an autocatalytic reaction involving three residues (Ser^{65}-Tyr^{66}-Gly^{67}) [66]. This peculiarity allows GFP to be easily genetically modified and cloned into different biological systems, which seems an attractive approach for biosensors [67]. Different versions of fluorescent proteins were developed since GFP was discovered by Osamu Shimomura et al. [68]. The modification of the chromophore structure or the adjacent environment to absorb and emit light at different wavelengths allows the creation of new variants of fluorescent protein colours [67]. Examples of produced colours of fluorescent proteins based on GFP are blue [69] or yellow fluorescence [70]. Moreover, efforts have been made to improve some GFP properties, like folding, solubility or chromophore maturation. Mutations on the chromophore residue Ser^{65} to a threonine and Phe^{64} to a leucine resulted in a GFP variant with improved brightness, the enhanced GFP (EGFP) [71]. Thanks to its enhanced brightness it is widely used as a fluorescent label for quantitative fluorescence microscopy applications. However, studies have shown that it is not suitable for single-molecule tracking

detection for long periods of time [71]. Based on EGFP and folding reporter GFP [72], Pédelacq et al. developed a strongly folded mutant, that folds even when fused to polypeptides, the superfolder-GFP (sfGFP). This GFP variant comprises the cycle-3 mutations from folding reporter GFP, the two mutations of EGFP and six new mutations [73]. sfGFP displays increased resistance to chemical denaturation, faster refolding kinetics and high tolerance to circular permutation. Recently, Kamiyama et al. developed a system for protein tagging named split-GFP [74]. This system relies on GFP cleavage into two fragments, GFP S1-10 (GFPS10) and GFPS11. The fragments alone are non-fluorescent since the conserved residue E222 present in GFPS11 is required for chromophore maturation. However, by transcomplementation GFP fluorescence is reconstituted.

Additionally, the use of fluorescence proteins has the advantage of switching from off-to-on mode when structural distortions are performed [46 67]. Therefore, a fluorescence cell-based switch off-to-on sensor activated by viral enzymes, like the viral proteases, represents an exciting and promising label-free system for virus titration, without requiring vector modification. There are many viruses in which polyprotein processing and maturation is regulated by viral proteases, which provides several advantages for virus lifecycle. This method of genome organization allows genome condensation by erasing genetic features (e.g. promoters) necessary for individual protein expression as well as regulation of protein activity depending on polyprotein cleavage site [75]. The main function of proteases is hydrolyzing the peptide bond between amino acid residues in a polypeptide chain, which makes them suitable as sensor activator. According to their catalyzation mechanism, proteases can be categorized into: aspartic, metalloproteases, cysteine, serine and threonine. One of the main differences between these proteases is the nucleophile of the active site. In the latter three the nucleophile is an amino acid residue (Cys, Ser or Thr, respectively) from which derives the name, while the others use an activated water molecule [76 77]. A few switch off-to-on fluorescent biosensors activated by viral enzymes were already developed and are briefly presented.

In 2009 Schekhawat et al. [78] developed a method based on coiled-coil auto inhibition embedded in split proteins of firefly luciferase. The biosensor consists of a split reporter protein, an antiparallel heterodimeric coiled-coil and a protease cleavable sequence linker. When the linker is cleaved by the protease, the coiled-coil portions can interact, allowing the transcomplementation of firefly luciferase and, consequently, emission of light (**Figure 1.2A**).

In the same year, Iro et al. [79] developed a cell-based biosensor for HCV, consisting on a fused EGFP and secreted alkaline phosphatase (SEAP) with a cleavable linker site for HCV protease. After HCV infection, the linker is cleaved and SEAP is released from the fusion protein and secreted into the extracellular supernatant where it can be detected (**Figure 1.2B**). This system is fast and sensitive however it only allows detection and not virus quantification.

Also for HCV detection, Jones et al. [80] developed a sensitive method able to distinguish infected cells from non-infected cells, in live or fixed samples, based on relocalization of a genetically encoded fluorescence protein. For this purpose, the system makes use EGFP or red fluorescent protein fused with the carboxy-terminal region of mitochondrially-tethered IFN-β promoter stimulator protein 1 (IPS-1) (a cellular substrate of HCV), comprising the HCV protease cleavage site and a mitochondrial targeting sequence. Upon HCV infection, the reporter protein is cleaved and relocalized from the mitochondria membrane (punctuated pattern) to the cytoplasm (diffuse fluorescence), as depicted in **Figure 1.2C**. One of the major limitations of this system is not allowing high throughput, because it's based on single-cell analysis by fluorescence microscopy.

Callahan et al. [81] implemented a method for the detection of HIV-1, based on transcomplementation of GFP. The GFPS11 was embedded as a surface loop of a small protein, Eglin c, together with a cleavable sequence linker, which leads to its structural distortion and

prevents transcomplementation. The fluorescence is reconstituted once the structural distortion is relieved by protease cleavage and transcomplementation of GFPS11 and GFPS10 is allowed.

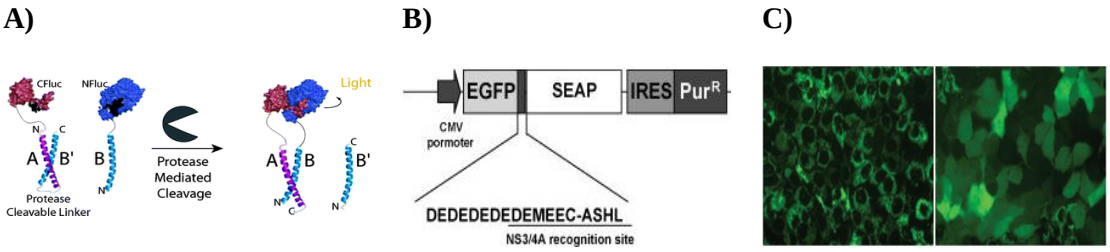

Figure 1.2 - Fluorescence cell-based systems already developed for virus detection. A) "Auto-inhibited Coiled-Coil Design Strategy for Split-Protein Protease Sensor". The C-terminal fragment of firefly luciferase (CFluc) is attached to one of the coiled-coil portions, A, associated with B', an amino acid sequence similar to B, and the protease cleavable sequence linker. The B portion of the coiled-coil is coupled with the N-terminal firefly luciferase fragment (NFluc). After cleavage of the linker by the protease, firefly luciferase is able to transcomplement and emit fluorescence. Adapted from S. S. Shekhawat et al. [78]. **B) Cell-based SEAP reporter assay.** EGFP is linked to SEAP by a protease HCV cleavage site and a spacer (DEDEDEDE). Adapted from Iro et al. [79]. **C) HCV-dependent fluorescence relocalization**. Upon the cleavage of HCV protease, the reporter protein is relocalized from the mitochondria membrane (punctuated pattern) to the cytoplasm (diffuse fluorescence). Adapted from Jones et al. [80].

The strategies presented above [78] [79] [81] [80], as well as other strategies already developed present a series of disadvantages that should be considered, such as low sensitivity, low signal to noise ratio (SNR) performance and most are biochemical assays. Additionally, most of them do not allow virus quantification, only detection. Therefore, a robust strategy for detection and quantification of label-free virus and viral vectors is still a current need.

1.4. FIRST STEPS TOWARDS LABEL-FREE VIRUS DETECTION

Recently, our group at the Cell Line Development and Molecular Biotechnology Laboratory started to tackle these limitations by developing a new promising titration method for retroviral vectors, named Single Step Cloning Screening Method [26]. This method is based on GFP transcomplementation, and relies on co-culture of cells stably expressing the GFPS10 fragment with cells producing retrovirus coding the GFPS11 fragment in the viral transgene. When GFPS11-coding viruses infect target cells, GFPS10 and GFPS11 transcomplement and fluorescence emission is reconstituted (**Figure 1.3**). This system is fast, provides high throughput and a direct measure of infectious particles. Despite all the advantages, it's still not a label-free sensor, since the virus genome encodes the GFPS11 fragment. Therefore, a step further was needed to develop a new method for detection and quantification of label-free viruses, by taking advantage of viral enzymes.

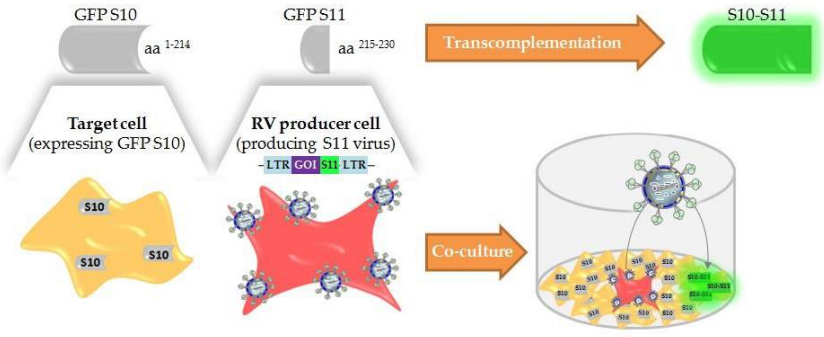

Figure 1.3 - Schematic representation of the Single Step Cloning Screening Method. Cells stably expressing GFPS10 fragment (amino acids 1-214) are infected with retrovirus (RV) encoding GFPS11 fragment (amino acids 215-230) fused to a gene of interest. Fluorescence emission is recovered by transcomplementation. GOI, gene of interest; LTR, long terminal repeat. Adapted from A. F. Rodrigues et al. [26].

As previously referred, AdV is still the "gold standard" vector in gene therapy (**Figure 1.1**). So, it would be of great interest to develop a fluorescent cell-based sensor for detection and quantification of label-free AdV, by taking advantage of the adenoviral protease as a sensor activator.

1.5. BIOLOGY OF ADENOVIRUSES

Adenoviridae family is composed by more than 55 human serotypes and are known for causing respiratory, ocular and intestinal infections worldwide. These viruses comprise non-enveloped virions with an icosahedral protein coat surrounding their DNA-containing core (**Figure 1.4**). The capsid comprises 12 identical fibers each coming from penton base vertices. The genome of the AdV is a linear, double-stranded DNA, coding for two major transcription regions: early and late region. Early region genes encode for non-structural viral proteins responsible for several regulation functions, which will allow late genes' expression. On the other hand, the late region genes encode structural viral proteins essential to virus particles formation. The linear DNA has a length ranging from 26 to 40 kb and consists of compact nucleosome-like structure which at the extremities contains inverted terminal repeat (ITR) sequences comprising the origin of viral DNA replication [11 82 83] .

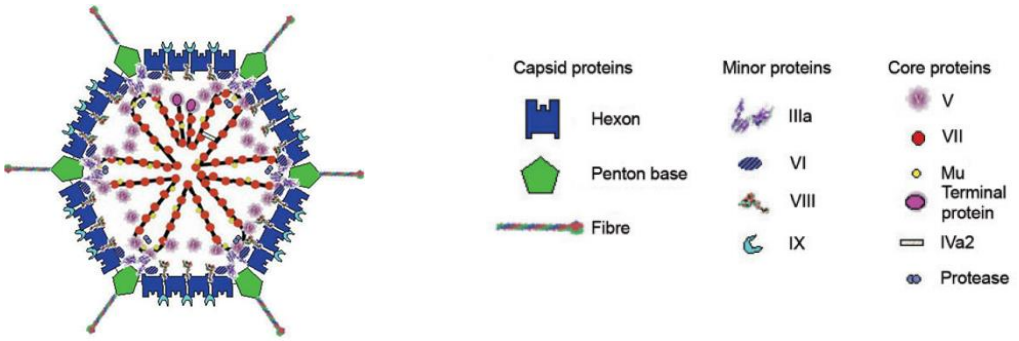

Figure 1.4 - Schematic representation of AdV structure. Adapted from Russel, W. C (2009) [83].

The AdV replication cycle comprises two phases, the early and late phase. Once the virus interacts with the host cell the early phase begins. The virus enters the cell and carries the viral genome to the nucleus, where the transcription and translation of the early genes occurs. The late phase relies on the expression of the late genes, which allows the assembly in the nucleus of the structural proteins and maturation of infectious virus. AdV are classified as lytic viruses since they induce host cells lysis, starting usually 20 to 24 hours post-infection (h.p.i.), allowing virions to be release [84].

The adenoviral protease – commonly known as Adenain – is essential for virus infectivity and maturation, since it catalysis the processing of six structural viral proteins (three capsid proteins and three DNA-associated core proteins) [85]. The processing of these proteins is performed in consensus cleavage sequences: (M/I/L)XGX↓G (type 1) and (M/I/L)XGG↓X (Type 2) [86]. Additionally, this protease also cleaves the cytokeratin 18 late in adenoviral infection (LAGM↓G) [87], contributing to the cell cytoskeleton dissociation and, consequently, cell lysis. The cysteine protease Adenain has about 200 amino acid residues and comprises two domains with the active site at the domain interface. The nucleophile of the active site is localized in the Cys122 residue [85]. Interestingly, the adenoviral protease is expressed in a nearly inactive form, with its full activation only occurring in the presence of the two co-factors. The cleaved 11-residue peptide from the C-terminus of precursor protein VI (pVIc) results in a 120-fold increase in protease activity. Adenain interaction with viral DNA rises its activity 3-10 fold [88] [89]. The activation of the peptide co-factor is performed by Adenain itself, by cleaving pVIc from the precursor protein in a specific cleavage site (IVGL↓GVQS) [85]. Of notice, the cytoplasmatic protein actin contains a C-terminal portion homologous to pVIc [90], that when cleaved can also work as a co-factor.

As of today, there are no approved anti-adenoviral therapeutics available, despite a growing unmet medical need for specific anti-adenoviral drugs, in particular for immunosuppressed patients. Due to its importance in viral maturation and infectivity, Adenain is a viable target for the development of antiviral agents [91].

2. AIM AND STRATEGY

The aim of this project was to develop a mammalian cell-based fluorescent biosensor for detection and quantification of label-free viruses, the VISENSORS, taking advantage of the viral protease activity as a trigger of the biosensor. Since several viruses make use of proteases for virus maturation and life cycle, this enzyme seems a suitable sensor activator.

As proof-of-concept, human AdV serotype 5 (AdV5) were used as study model, due to their extensive value in development of vectored vaccines and gene therapy (**Figure 1.1**).

VISENSORS are mammalian cell-based genetically encoded biological sensors that rely on structural distortion of fluorescent proteins, like GFP, to limit its fluorescence emission. Upon recognition of a specific cleavage sequence by the adenoviral protease, when cells are infected with label-free AdV5, fluorescence emission is recovered, thus a switch-on system. Three different strategies were addressed as shown in **Figure 2.1**.

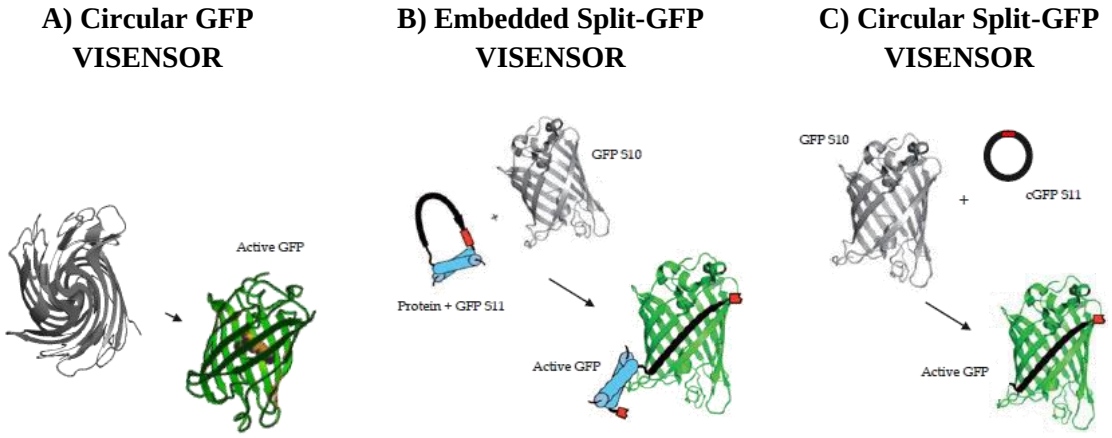

Figure 2.1 - Different strategies considered for VISENSORS development. (A) Circular GFP VISENSOR. Following Zhang et. al [92] work, this strategy consists on a genetically encoded switch-on fluorescent biosensor whose fluorescence is limited by protein circularization promoted by *Npu* DnaE intein. Fluorescence is reconstituted once proteolytic cleavage takes place and circular GFP (cGFP) is converted to a linear form, able to complete its final maturation folding. **(B) Embedded Split-GFP VISENSOR**. This strategy consists on a mammalian optimized version of Callahan et al. [81], to embed GFPS11 fragment (eS11) as a surface loop of the small Eglin c protein (in blue). This implements a structural distortion on GFPS11 until AdV protease cleaves the cleavable linker (in red). **(C) Circular Split GFP VISENSOR**. This approach consists on the cyclization of GFPS11 fragment (cS11) in a similar approach followed by Sakamoto et al.[93]. GFPS11 cyclization is mediated by split-intein protein (*Npu* DnaE intein) splicing reaction and once the viral protease recognizes a specific cleavable linker (in red), structural distortion is relieved and GFPS11 can transcomplement with GFPS10. Schematic figures adapted from B. P. Callahan et al. [81].

3. MATERIALS AND METHODS

3.1. PLASMIDS

All the plasmids used in this work were derived from a lentiviral transgene plasmid – *pRRLSin*, a self-inactivating (SIN) third-generation lentiviral plasmid kindly provided by Miguel Guerreiro (ACT Unit IBET/ITQB NOVA, Oeiras, Portugal). *pRRLSin.CMV.GFPS10.IRES.Zeo.WPRE* and *pRRLSin.CMV.GFPS10.IRES.Puro.WPRE* take advantage of a cytomegalovirus (CMV) promoter to drive expression of the gene of interest (GFPS10, for example) and of encephalomyocaerditis virus-internal ribosome entry site (EMCV.IRES) to drive expression of the *zeocin* or *puromycin* resistance gene. The original *pRRLSin.hPGK.eGFP.WPRE* lentiviral transgene was kindly provided by Dr. Didier Trono through Addgene plasmid repository (Addgene, Cambridge, Massachusetts, USA). These lentiviral transgene plasmids were used not only for VISENSORS' characterization by transient transfection but also for lentiviral production for stable cell lines' development. Information concerning the PCR primers and plasmids constructed during this work is listed in **Tables S1** and **S2**. The remaining plasmids used during this work were kindly provided by Miguel Guerreiro (ACT Unit IBET/ITQB NOVA, Oeiras, Portugal). Bellow, a brief description of the all the plasmids used is presented.

3.1.1. CIRCULAR GFP VISENSORS

For the construction of cGFP plasmids *pRRLSin.CMV.GFPS10.IRES.Zeo.WPRE* was used as backbone plasmid, digesting with NheI and BamHI restriction enzymes to replace GFPS10 by the genes coding the different cGFP VISENSORS. As so: *pRRLSin_cGFP-LRGAG en*codes for the circular version of sfGFP with a LRGA↓G cleavage site; *pRRLSin_cGFP-LRGAG_w/_Met* and *pRRLSin_cGFP-G/LRGAG/G* are similar to *pRRLSin_cGFP-LRGAG* but with the addition of a methionine residue at the C-terminal fragment of sfGFP or the addition of one glycine residue at each side the cleavage sequence, respectively; *pRRLSin_cGFP-IVGLG*, *pRRLSin_cGFP-MGGRG*, *pRRLSin_cGFP-IRGRG*, *pRRLSin_cGFP-NTGWG* and *pRRLSin_cGFP-EEGEG* differ from *pRRLSin_cGFP-LRGAG* only in the Adenain cleavage site (IVGL↓G, MGGR↓G, IRGR↓G, NTGW↓G and EEGE↓G, respectively).

3.1.2. EMBEDDED SPLIT-GFP VISENSORS

For the construction of eS11 plasmids *pRRLSin.CMV.GFPS10.IRES.Puro.WPRE* was also used as backbone plasmid, replacing GFPS10 by the genes coding the different eS11 VISENSORS. As so: *pRRLSin_eS11-G/LRGAG/G* encodes the GFPS11 fragment embedded in Eglin C's loop with a LRGA↓G cleavage site surrounded on each side with one glycine residue; *pRRLSin_eS11-LRGAG* was constructed by removing the glycine spacers from *pRRLSin_eS11-G/LRGAG/G*; *pRRLSin_eS11-IVGLG* and *pRRLSin_eS11-EEGEG* are based on the construction of *pRRLSin_eS11-LRGAG* with different cleavable sites (IVGL↓G and EEGE↓G, respectively).

3.1.3. CIRCULAR SPLIT-GFP VISENSORS

For the construction of cS11 plasmids, *pRRLSin.CMV.GFPS10.IRES.Puro.WPRE* was used as backbone plasmid, replacing GFPS10 by the genes coding the different cS11 VISENSORS. As so:

*pRRLSin_cS11-LRGAG en*codes for the circular version of GFPS11 fragment of Split-GFP with a LRGA↓G cleavage site; *pRRLSin_cS11-G/LRGAG/G* and *pRRLSin_cS11-GG/LRGAG/GG* were derived from *pRRLSin_cS11-LRGAG* by adding one or two glycine spacers at each side the cleavage sequence, respectively; *pRRLSin_cS11-G/IVGLG/G* and *pRRLSin_cS11-G/EEGEG/G* are a modified version of *pRRLSin_cS11-G/LRGAG/G*, harbouring different cleavage sites (IVGL↓G and EEGE↓G, respectively). All the information concerning the plasmids constructed in this work for this strategy is listed in **Table S1**.

For the construction of cGFPS10 plasmid, *pRRLSin.CMV.GFPS10.IRES.Zeo.WPRE* was used as backbone plasmid. *pRRLSIN_cGFPS10-LRGAG* codes for the circular version of GFPS10 fragment of Split-GFP with a LRGA↓G cleavage site, is an analogous construction to cGFPS11-*LRGAG* by replacing the GFPS11 with GFPS10 fragment.

3.1.4. VIRAL PROTEASES

pRRLSin.CMV.GFPS10.IRES.Zeo.WPRE was used as backbone plasmid. For the construction of *pRRLSin_Adenain-MVGLG*-VIc, encoding the wild-type adenovirus protease fused with a MVGL↓G cleavage site and pVIc (a modified version of Balakirev et al. [94] work), GFPS10 was removed and replaced with *Adenain-MVGLG*-VIc coding sequence (often refered as Adenain for simplicity). For the construction of *pRRLSin_Adenain-MVGLG-VIc_mCherry*, the *zeocin* resistance gene from *pRRLSin_Adenain-MVGLG-VIc* was replaced by the *mCherry* gene isolated from *pPuro_mCherry* [95], kindly provided by Ana Formas-Oliveira (ACT Unit IBET/ITQB NOVA, Oeiras, Portugal). *pRRLSin_Adenain-C104G/C122G_mCherry* encodes a non-functional version of the Adenain, due to C122G (active site nucleophile) [94] and C104G (pVIc binding site) [96] mutations, and lacks pVIc fusion. All the information concerning the plasmids constructed is listed in **Table S2**.

3.2. CLONING PROCEDURES

All PCR reactions were performed in a *Biometria® T3 Personal Thermocycler* (Biometria, Göttingen, Germany) under suitable conditions for amplification, using a proof-reading *Phusion® High Fidelity* DNA Polymerase (Finnzymes Oy, Vantaa, Finland). The custom-made oligonucleotides (**Table S1** and **S2**) were synthesised by Sigma-Aldrich (St. Louis, Missouri, USA). All restriction reactions were incubated at least 1 hour at 37º C using the indicated endonuclease restriction enzymes (New England Biolabs, Ipswich, Massachusetts, USA). DNA fragments were isolated by agarose gel (NZYTech, Lisboa, Portugal) prepared in 1x Tris-Acetate-EDTA (TAE) buffer (Qiagen, Hilden, Germany) and *RedSafe™ Nucleic Acid Staining Solution* (iNTrON Biotechnology, South Korea). Agarose gel concentration was dependent on the size of the DNA fragments to isolate. Loading buffer (*Gel Loading Dye Purple*) (New England Biolabs) and a standard *NZYDNA Ladder III* (NZYTech) were used during gel electrophoresis. Agarose gels were analysed using *GelDoc™ XR+* system (Bio-Rad, Hercules, California, USA) and purified with *Gel Band Purification Kit* (GE Healthcare, Little Chalfont, UK). DNA quantification was performed using *Nanodrop ND-2000c* (Thermo Scientific™, Waltham, Massachusetts, USA).

Ligation reactions (**Table S1 and S2**) were performed using *In-Fusion® HD Cloning Kit* (Clontech Laboratories Inc., Mountain View, California, USA) following manufacturer's instructions, with a molar ratio of insert to vector of 4:1.

3.3. BACTERIAL STRAINS AND CULTURE MEDIA

Bacteria transformation was performed using *Escherichia coli (E. coli)* Stellar™ (Clontech Laboratories Inc.) and One Shot® Stbl3™ (Life Technologies, Carlsbad, California, USA) competent cells. Agar and liquid cultures were carried in *Luria Broth* (LB) media (Fast-Media® LB) (Invivogen, San Diego, California, USA) supplemented with the proper antibiotic and following manufacturer's instructions.

3.4. PLASMID PURIFICATION AND QUALITY CONTROL

The constructed plasmids were purified at small-scale (yields up to 20 µg of DNA) using *GeneJET Plasmid Miniprep Kit* (Thermo Scientific™) and large-scale (yields up to 500 µg of DNA) with *Genopure Plasmid Maxi Kit* (Roche Applied Science, Penzberg, Germany), according to manufacturer's instructions, and stored at -20º C. Plasmid quantification and purity (evaluated by Abs_{260nm}/Abs_{280nm} and Abs_{260nm}/Abs_{230nm} ratios) was performed using *Nanodrop ND-2000c* (Thermo Scientific). Plasmids restriction pattern was evaluated by enzymatic restriction and agarose gel electrophoresis. Purified plasmids were sequenced using GATC Biotech services (Constance, Germany) for nucleotide sequence validation. Working bacteria banks were established and stored at -80º C in 15% (v/v) glycerol (Sigma-Aldrich).

3.5. CELL LINES AND CULTURE CONDITIONS

HEK 293 derived cell line (ATCC CRL-1573) stably expressing the E1 gene was used for AdV5 production and for the establishment of cell lines stably expressing the VISENSORS. HEK 293T (ATCC CRL-11268) is a HEK 293 derived cell line expressing SV40 large T antigen and was used for transient lentiviral vector production and for transient transfection screenings for VISENSORS' optimization of backbones and cleavable sequences. HEK 293 FLEX GFPS11 [26] is a HEK 293 derived cell line producing murine leukemia virus (MLV) based vector coding for a GFPS11 gene, pseudotyped with gibbon ape leukemia virus (GaLV) and were used to titrate LV vectors coding for GFPS10. All cell cultures were maintained in Dulbecco's Modified Eagle's Medium (DMEM) (Gibco™, Carlsbad, UK) supplemented with 10% (v/v) of Fetal Bovine Serum (FBS) (Gibco™) at 37 ºC inside an incubator with a humidified atmosphere containing 5% CO_2.

3.6. CELL CONCENTRATION AND VIABILITY

Trypan blue exclusion assay was carried to determine cell concentration and viability, using 0.1% (v/v) of Trypan Blue (Sigma-Aldrich) solution prepared in Phosphate Buffer Saline (PBS) (Gibco™). Cell counting was performed in a Fuchs-Rosenthal haemocytometer (Marienfeld-Superior, Lauda-Königshofen, Germany) using an inverted microscope (Olympus, Tokyo, Japan).

3.7. TRANSIENT TRANSFECTIONS FOR VISENSORS' CHARACTERIZATION

Initial characterization of VISENSORS was performed by transient transfection. Briefly, HEK 293T cells were seeded at $8x10^4$ cell/cm^2 in 24 well-plates (Nunc, Waltham, Massachusetts, USA). 24 hours later, cells were transfected with a total of 5 µg of total DNA per million cells, using

polyethylenimine (PEI, Linear 25 kDa) (Polysciences, Inc., Warrington, Pennsylvania, USA) at 1:1.5 (w/w) DNA:PEI ratio. The following conditions were used: for cGFP strategy, cGFP sensor plus Adenain (positive condition) or mCherry (coded by pPuro_mCherry plasmid) (mock, negative control) were transfected; for split-GFP systems, GFPS11 (eS11 or cS11) plus GFPS10 (wild-type or cGFPS10) plus Adenain or mCherry were transfected. For cGFP specificity analysis, additionally to the functional Adenain, a non-functional Adenain C104G/C122G or *pMDLg/pRRE* plasmid (the latter coding for HIV-1 viral protease) were used. 48 hours after transfection cells were observed by fluorescence microscopy in a Leica DMI6000 inverted microscope (Leica, Wetzlar, Germany), harvested and analysed for GFP fluorescence intensity by flow cytometry (CyFlow® space) (Partec GmbH, Görlitz, Germany).

3.8. PROTEIN EXTRACTION, QUANTIFICATION AND WESTERN BLOT ANALYSIS

For cGFP strategy validation, transient transfections were performed as previously described, but using a cell inoculum of $1x10^5$ cell/cm^2 in 75 cm^2 tissue flasks. 48 hours post-transfection, cells were harvested, pelleted at 300g for 10 minutes at 4º C and then washed in PBS. After a new centrifugation, cells were lysed in Mammalian Protein Extraction Reagent (M-PER) (Thermo Scientific™) supplemented with cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche Applied Science) at a concentration of $3x10^6$ cells per 100 µL of lysis solution, following manufacturer's instructions. Aliquots were frozen at -20º C. Total protein quantification was performed using the Pierce™ BCA Protein Assay Kit (Thermo Scientific™), following manufacturer's instructions.

Protein electrophoresis separation was performed in NuPAGE® electrophoresis system (Life Technologies), under denaturing conditions and loading the same amount of protein in all wells. Samples were resolved on a NuPAGE® 12% (w/v) Bis-Tris gel with NuPAGE® MES (for GFP detection only) or MOPS SDS Running Buffer at 180 V. Protein transfer was conducted into a nitrocellulose membrane in a Trans-Bot® Turbo™ Transfer System (Bio-Rad). All procedures followed manufacturer's instructions. For immunoblotting, membranes were incubated over-night at room temperature with mild agitation with the following primary antibodies: mouse anti-GFP (#2955) (Cell Signalling Technology, Danvers, Massachusetts, USA), diluted 1:1000 in 5% (w/v) Bovine Serum Albumin (BSA) (Sigma-Aldrich) in Tris Buffered Saline (Sigma-Aldrich) with 0.1% (w/v) Tween 20 (Merck) (TBST); rabbit anti-Myc tag (#2278) (Cell Signaling Technology), diluted 1:1000 in 5% (w/v) BSA in TBST; mouse β-actin (#8226) (Abcam, Cambridge, UK) and mouse α-tubulin (#T6199) (Sigma-Aldrich), diluted 1: 2000 and 1:5000, respectively, in 5% (w/v) BSA in TBST. Secondary antibody incubation was performed with sheep anti-mouse ECL (NA931V) or donkey anti-rabbit ECL (NA93AV) (GE Healthcare) diluted 1:5000 in blocking solution (5% (w/v) skim milk powder (Merck, Darmstadt, German) in TBST) for 2 hours at room temperature with mild agitation. Finally, membranes were incubated with Amersham™ ECL™ prime western blotting detection reagent (GE Healthcare), following manufacturer's instructions and the emitted chemiluminescence was detected in ChemiDoc XRS System (Bio-Rad).

3.9. TRANSIENT LENTIVIRAL VECTOR PRODUCTION

For the transient production of LV vectors the third generation lentiviral packaging system was used [97]. HEK 293T cells were seeded at $8x10^4$ cells/cm^2 in tissue-flasks and transiently transfected with lentiviral expression cassettes kindly provided by Dr. Didier Trono (Addgene, Cambridge, Massachusetts, USA): *pMDLg/pRRE* (encoding HIV-1 Gag-Pro-Pol polyprotein and RRE), *pMD$_2$G* (encoding the envelope G glycoprotein of the vesicular stomatitis – VSV-G) and *pRSV-REV* (encoding HIV-1 Rev protein). For the transgene cassettes, we used the *pRRLSin*-derived plasmids (coding for the

VISENSORS) mentioned above. Respectively, a ratio of 1.0:0.25:0.9:2.5 was used. 24 hours post-transfection culture medium was replaced and reduced to half of the original volume. LV vectors were harvested 48 hours after transfection, filtered through a 0.45 µm-pore-size cellulose acetate filter (Sarstedt) for clarification and aliquots were stored at -80º C.

3.10. LENTIVIRAL VECTOR TITRATION

For titration of LV vectors coding for GFPS10, HEK 293 FLEX GFPS11 cells were seeded at $5x10^4$ cells/cm^2 in 24 well-plates. After 24 hours, cells were counted, culture medium was removed and 0.2 mL of volume of LV suspension prepared at several dilutions in fresh DMEM supplemented with 10% (v/v) FBS and 8 µg/mL of polybrene (Sigma-Aldrich) was added in duplicates. Cells were incubated at 37º C for 4 hours after which 0.5 mL of fresh supplemented DMEM was added. 48 hours after infection, cells were harvested and analyzed for GFP fluorescence by flow cytometry (Partec GmbH). Viral titers, defined as the number of infectious particles (I.P.) per mL, were determined considering the percentage of GFP-positive cells using the equation below. Only dilutions rendering 2-20% of GFP-positive infected cells were considered.

$$Titer\ (I.P./mL) = \frac{\%\ GFP\ positive\ cells \times number\ of\ cells\ at\ infection}{100} \times \frac{viral\ dilution\ factor}{Volume\ of\ infection} \qquad \textbf{(Eq. 3.1)}$$

3.11. ESTABLISHMENT OF STABLE CELL LINES EXPRESSING THE VISENSORS

For stable production of the biosensor in mammalian cells, VISENSORS genes were delivered by lentiviral transduction. HEK 293 cells were seeded at $8x10^4$ cells/cm^2 in 6 well-plates. After 24 hours, lentiviral transduction was performed by removing the medium and infecting cells with 0.6 mL of viral suspension harboring fresh DMEM supplemented with 10% (v/v) FBS and 8 µg/mL of polybrene with a multiplicity of infection (MOI) of 5. Cells were incubated at 37º C and 1.4 mL of fresh supplemented DMEM was added four hours later. 48 hours post-infection, cells were amplified to 75 cm^2 tissue-flasks in fresh supplemented DMEM with the respective selection antibiotic. Culture medium was exchanged every 3-4 days until selection was complete. Cell banks were made in a freezing solution containing fresh FBS supplemented with 10% (v/v) dimethyl sulfoxide (DMSO) (Sigma-Aldrich) and stored at -80º C.

3.12. ADENOVIRUS STOCK PREPARATION AND TITRATION

A recombinant E1-deleted strain of AdV5 expressing a gene of interest for vaccination purposes (provided by Dr. Geneviève Libeau, CIRAD - UMR Contrôle des Maladies, Montpellier, France) without a fluorescent label (label-free) was used in this work. The AdV5 viral stock was prepared by infecting monolayers of HEK 293 cells (80% confluency) using a MOI of 5. 42 hours after infection, cells were harvested with a cell scrapper and the cell pellet was lysed with 0.1% (v/v) Triton X-100 (Sigma-Aldrich) in 10 mM Tris-HCl Buffer at pH 8 (Calbiochem, Billerica, Massachusetts, USA). Intracellular viral particles were collected after centrifuging ressuspended cells in lysis buffer for 10 min at 3000 g, 4º C, purified by $CsCl_2$ gradients according to standard procedures, followed by desalting using the AKTA system with a Hiprep 26/10 desalting column equilibrated with sterile 10 mM Tris-HCl Buffer at pH 8, 2 mM $MgCl_2$. Aliquots were stored at -80º C in a cryopreservation solution containing 10 mM Tris-HCl Buffer at pH 8, 2 mM $MgCl_2$ and 0.5 M Trehalose [98].

AdV5 titration was performed by the end-point dilution method (TCID50) using 96-well plates and HEK 293 cells. Viral titer was calculated according to the method of Spearman and Kraber [99] - 1.2 x 10^{10} TCID50/mL.

3.13. ADENOVIRUS INFECTION FOR VISENSORS' CHARACTERIZATION

To evaluate VISENSORS response to AdV5 infection, HEK 293 cells stably expressing the VISENSORS were seeded at $1x10^5$ cells/cm^2 in 24 well-plates. After 24 hours, culture medium was removed and cells were infected with an MOI of 5 in 0.2 mL of fresh non-supplemented DMEM with mild agitation. After 1 hour, 0.3 mL of supplemented DMEM was added. At the given time points after infection, cells were observed by fluorescence microscopy in a Leica DMI6000 inverted microscope (Leica) and harvested and analyzed for GFP fluorescence intensity by flow cytometry (CyFlow® space, Partec).

3.14. STATISTICAL ANALYSIS

In each experiment, each condition was conducted at least in triplicates and SNR performance was evaluated by flow cytometry. Signal stands for the GFP geometric mean fluorescence intensity when the sensor was activated, either by the transiently transfection of the Adenain or by AdV infection. In the other hand, Noise (also referred as background fluorescence) stands for the GFP geometric mean fluorescence intensity in an off-state, which corresponds in the transient screening to the mock condition (containing mCherry fluorescent reporter) and to non-infected cells in stable cells. The use of SNR gives an estimation of how much the sensor fluorescence has increased when activated.

Statistical analyses were performed using R Statistical Software. ANOVA F-test (normal distribution assumption) was conducted to evaluate the mean value of the different backbones in cGFP and cS11 and cleavage sites of all strategies. This test was applied for equal mean values and in case of significant differences the HSD of Tukey test was performed. Unpaired two-tailed t-test (normal distribution assumption) was also conducted for equal mean values of different backbones in eS11 and combined distortion of GFPS10 and GFPS11 fragments strategy.

4. RESULTS

4.1. CIRCULAR GFP VISENSOR

In order to achieve the goal of developing a fluorescent cell-based biosensor for detection of label-free AdV, our first approach relied on structural distortion of a GFP by means of protein circularization, wherein referred as cGFP. This strategy, which was based on Zhang et al. [92] work for detection of caspase-3-like activity, consists on a switch-on fluorescent sensor based on circularized sfGFP generated by means of *Nostoc punctiforme* (*Npu* DnaE) intein splicing, as depicted in **Figure 4.1**. The truncated N- and C-termini of a sfGFP were linked with a cleavable linker recognized by the AdV protease. Additionally, new N- and C-terminal ends at sfGFP residues A154 and D155 were generated and joined with the Dc- and Dn-fragments, respectively, of the *Npu* DnaE split-intein. Intein-mediated *cis*-ligation at its natural exteins (CFN and AEY) leads to circularization of the sfGFP and fluorescence emission inhibition. Fluorescence is reconstituted once proteolytic cleavage takes place at the cleavable linker and cGFP is converted to a linear form, able to complete its final maturation folding and emit fluorescence.

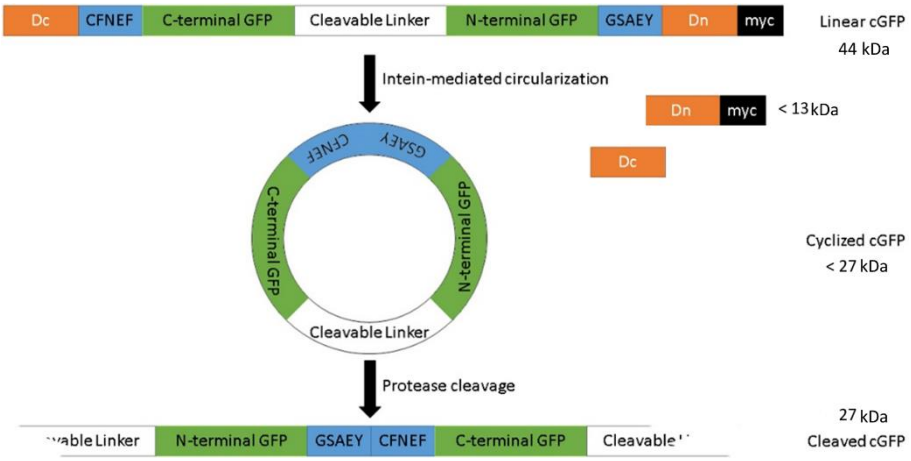

Figure 4.1 - Circularization process of cGFP biosensor. sfGFP was fragmented and its order inverted. The original N- and C-termini were linked with a cleavable linker specific for the AdV viral protease. Circularization is a *Npu* DnaE split-intein-mediated process by cis-ligation of its exteins (CFN and AEY, respectively). After protease cleavage, cGFP is converted to a linear protein able to produce fluorescence. Dc, C-fragment of *Npu* DnaE intein; Dn-myc, N-fragment of *Npu* DnaE intein fused with a myc tag; CFN and AEY are the conserved extein residues for the splicing by *Npu* DnaE intein; EF and GS are residues from EcoRI and BamHI endonuclease restriction sites.

4.1.1. OPTIMIZATION OF BACKBONE AND CLEAVABLE SEQUENCE

Different degrees of structural distortion performed in the fluorescent biosensor affect the fluorescence background level and cleavage efficiency. Therefore, it is essential to understand the effect of different backbones – the amino acid residues supporting the cleavable linker - as well as of different cleavable sequences on these parameters.

For an initial characterization of the cGFP system, transient co-transfections of AdV protease and cGFP sensor were made. First, a direct adaptation of Zhang et al. [92] system was tested, by replacing the cleavable linker with the synthetic sequence LRGA↓G [86] [100] (arrow denotes cleavage site by Adenain) - cLRGAG w/ Met. Additional backbones were assessed by removing a methionine residue located at the C-terminal end of sfGFP, for increased structural distortion (denoted as cLRGAG), and

adding a glycine spacer on each side of the cleavage site, for increased exposure of the cleavable sequence to Adenain (denoted as cG/LRGAG/G).

As seen by fluorescence microscopy, **Figure 4.2A**, co-transfection of cGFP sensor with Adenain led to an increase in GFP fluorescence emission when compared to co-transfection with a mock plasmid, confirming that our switch-on strategy is functional. Additionally, we could observe that addition of flexible glycine spacers in cG/LRGAG/G led to a higher fluorescence emission, both with or without addition of Adenain. Flow cytometry analysis showed that cLRGAG backbone had the highest SNR performance (1.85±0.09), very similar to cG/LRGAG/G (1.82±0.05). cLRGAG w/ Met showed the lowest SNR performance (1.41±0.06) (**Figure 4.2B**).

Different cleavable sequences may induce different structural distortions levels as well as different recognition efficiencies by the Adenain protease. Therefore, alternative cleavable sites were tested (**Table 4.1)** in the best backbone (cLRGAG, as previously shown).

Table 4.1 Cleavable sequences recognized by Adenoviral protease. The arrow (↓) represents proteolytic cleavage site.

Cleavable Sequence	Origin/Occurrence
IVGL↓G	Natural (pVI, site 239) [101 102]
LRGA↓G	Synthetic [86 100]
MGGR↓G	Natural (pTP, 183) [101 102]
IRGR↓G	Natural (pVIII, site 131) [101 102]
NTGW↓G	Natural (pVII, site 13) [101 103]
EEGE↓G	Natural (L1 52/55K, site 66) [101 104]

As seen by fluorescence microscopy and flow cytometry, **Figure 4.2**, indeed different cleavable sequences can have an important impact in structural distortion and/or in sensor cleavage by Adenain. cIVGLG cleavable sequence presented the lowest fluorescence emission and SNR performance (1.20±0.01). On the opposite spectrum, cEEGEG presented the highest fluorescence emission, but with a SNR performance on par with other cleavable sequences. cMGGRG, cIRGRG and cNTGWG sensors showed similar GFP fluorescence emission to cLRGAG (data not shown). Again, cLRGAG exhibited the highest SNR (1.85±0.09), therefore, it was chosen to proceed in the cGFP strategy characterization.

4.1.2. VALIDATION OF cGFP SYSTEM

To validate cGFP system and assess that the increase on GFP fluorescence emission was due to sensor activation mediated by cleavage of the cleavable linker by Adenain, protein extracts from cells transiently co-transfected with cLRGAG sensor and active protease (Adenain) or a mutated non-active version (Adenain C104G/C122G) were analysed by Western Blot. Sensor specificity was also evaluated by transient co-transfection with HIV-1 viral protease.

Western blotting analysis (**Figure 4.3**) revealed full circularization of the sensor by means of *Npu* DnaE inteins, suggested by the release of the myc-tagged Dc-fragment (13 kDa band) and the assembly of the circularized cGFP (<27 kDa). Only cells co-transfected with the sensor and active Adenain protease (and not non-active Adenain or HIV-1 protease) were able to produce the cleaved cGFP (linear form with 27 kDa, indicated by an arrow), which migrates slower in the gel than the circularized counterpart. Additionally, we also observed a band at <54 kDa, only detected by anti-GFP antibody, which may consist of a circularized bimolecular cGFP able to function as sensor similarly to single molecule cGFP.

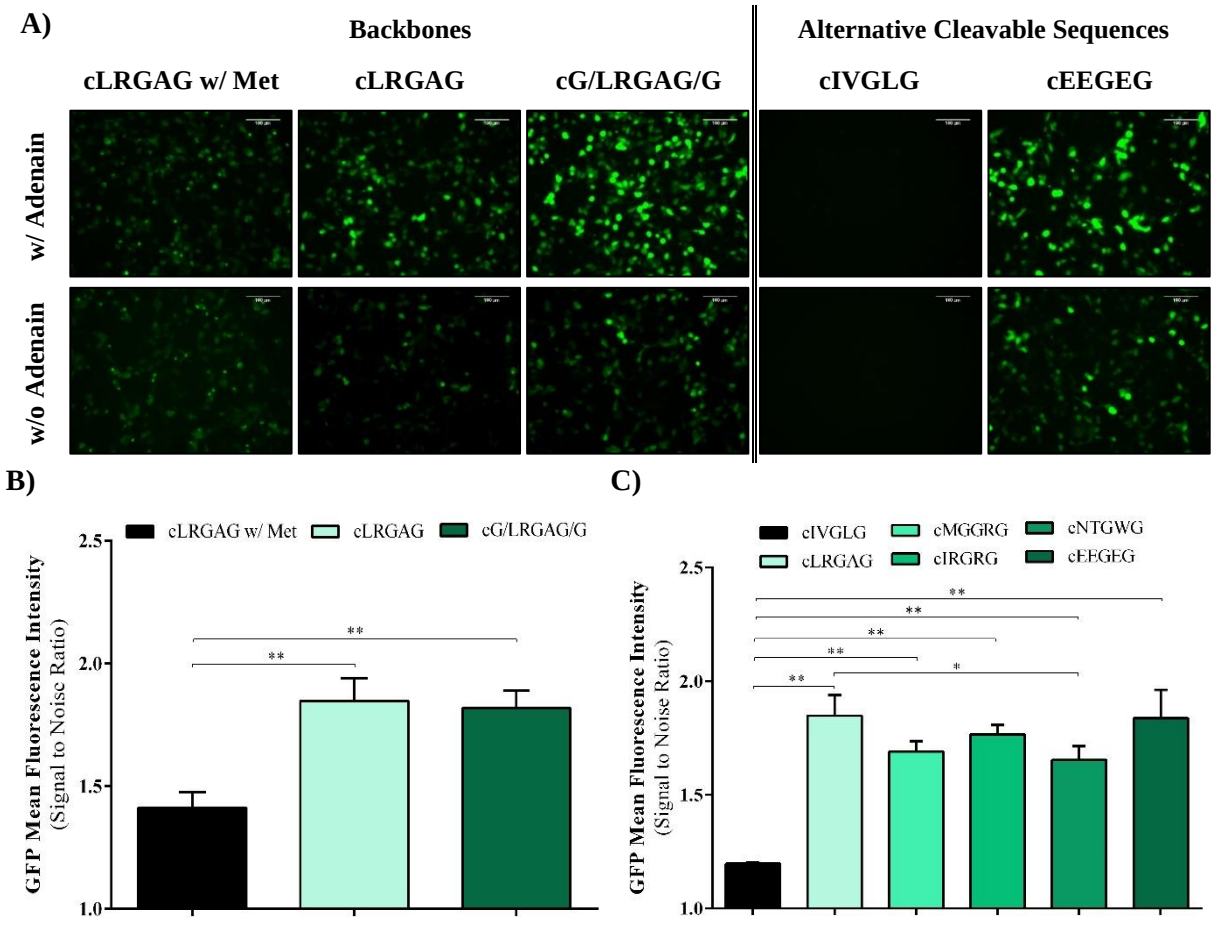

Figure 4.2 – Backbones and alternative cleavable sequences optimization for cGFP system. HEK 293T cells were transiently co-transfected with different cGFP backbones - cLRGAG w/ Met, direct adaptation of Zhang et al. [92] system by replacing DEVD↓G cleavable sequence by a specific AdV protease cleavable sequence, LRGA↓G; cLRGAG, deletion of a methionine residue; cG/LRGAG/G, cLRGAG with the addition of a glycine spacer by each side of the cleavage site - and cleavable sequences with AdV protease or a mock plasmid (containing mCherry fluorescent reporter). **(A)** Fluorescence microscopy images were acquired 48 hours post-transfection. Scale Bar: 100 µm. **(B)** SNR performance was evaluated by flow cytometry 48 hours post-transfection. SNR shown as average + standard deviation of at least triplicates (n=3). Statistical analyses using ANOVA F-test (HSD of Tukey), * $p < 0.05$ and ** $p < 0.005$.

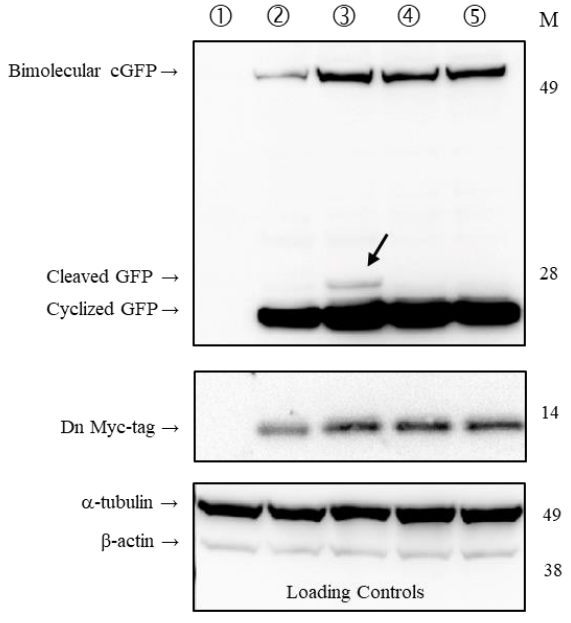

Figure 4.3 - Proteolytic processing analysis of the cGFP sensor. Western blotting analysis of the proteolytic processing of the cLRGAG sensor by ③ Adenain, ④ mutated non-active Adenain (C104G/C122G) and ⑤ HIV-1 protease. Extracts of HEK 293T cells transiently co-transfected with different proteases and cLRGAG sensor were examined by immunoblotting with anti-GFP and anti-myc tag antibodies. ① Negative control, not HEK 293 cells transfected. ② Co-transfection of a mock plasmid (containing mCherry fluorescent reporter). Circularized bimolecular of cGFP (<54 kDa); Cleaved cGFP (27 kDa); Cyclized cGFP (< 27 kDa); Dn Myc-tag, myc-tagged Dc intein fragment (13 kDa); M, SeeBlue® Plus2 Pre-Stained Protein Standard (Invitrogen, Carslbad, California, USA). α-tubulin (50 kDa) and β-actin (40 kDa). Schematic representation of the resulting proteins after protease cleavage of cGFP sensor is depicted in **Figure 4.1**

4.1.3. HEK 293 CELL LINE STABLY EXPRESSING cGFP SYSTEM

After initial characterization by transient co-transfection, a more comprehensive characterization was necessary to assess the sensor response to AdV infection. Therefore, HEK 293 cells stably expressing different backbones and alternative sequences were established by lentiviral transduction. Regarding backbone optimization, similar results to the transient screenings in terms of GFP fluorescence intensity were obtained, with cG/LRGAG/G generating higher fluorescence emission than cLRGAG w/ Met and cLRGAG (**Figure 4.4A**). Contrary to transient screening (**Figure 4.2B**) similar SNR performances were observed (**Figure 4.4B**) between all backbones. As for the alternative cleavable sequences, a similar behaviour was observed in stable cells (**Figure 4.4C/D**) when comparing with transient screening (**Figure 4.2C**). Again, cEEGEG and cIVGLG sensors presented the highest and lowest GFP fluorescence emission intensities, respectively (**Figure 4.4C**). Surprisingly, for all cGFP sensors the fluorescence of non-infected cells (background fluorescence) decreased over time (dashed lines, **Figure 4.4A/C**). Once again cLRGAG showed higher performance over time, reaching a SNR of 2.14±0.07 at 48 h.p.i. and 2.74±0.39 at 72 h.p.i. (**Figure 4.4D**).

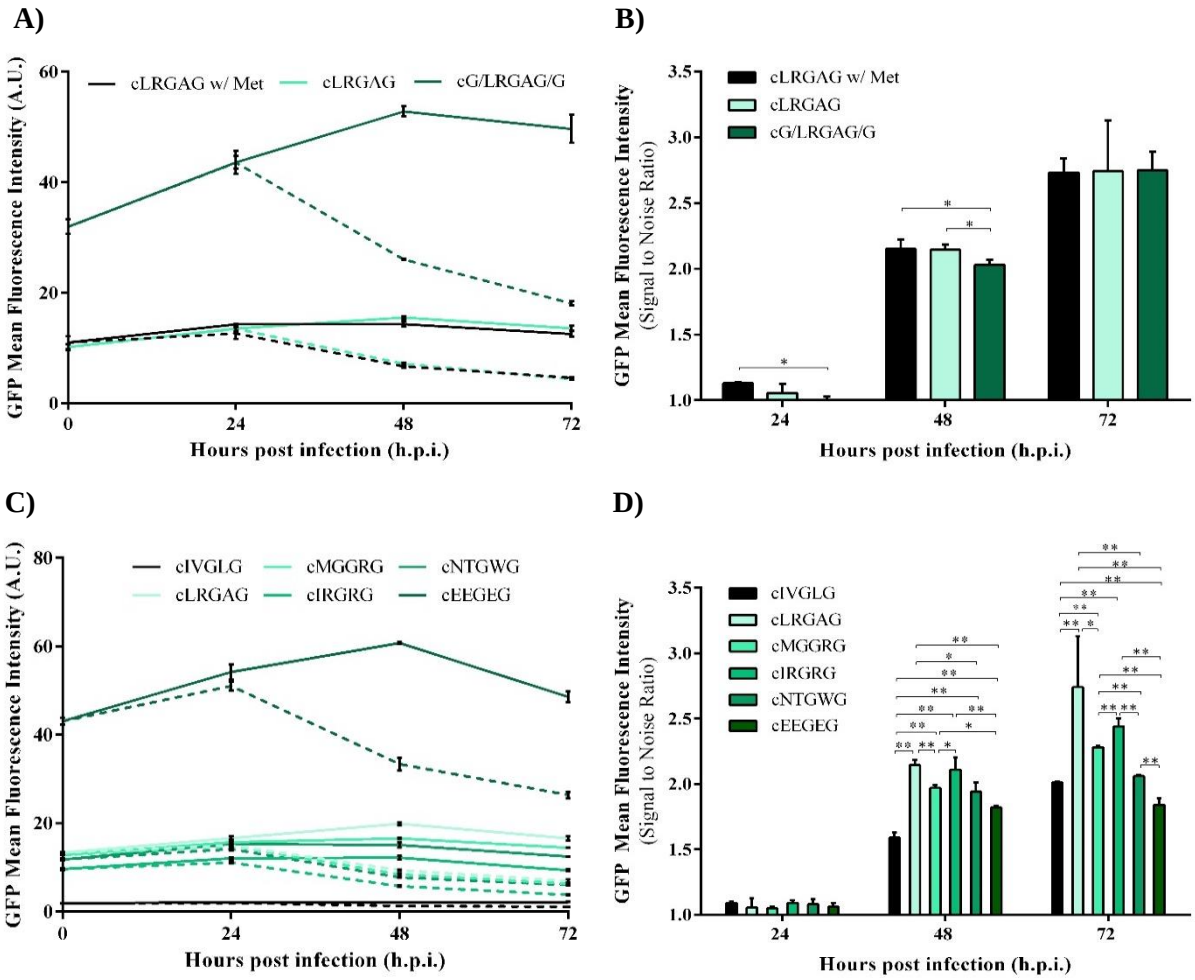

Figure 4.4 - Characterization of HEK 293 cells stably expressing cGFP sensor. HEK 293 cells stably expressing different (**A**) backbones or (**C**) cleavable sequences (in cLRGAG backbone) were infected with AdV at a MOI of 5. GFP fluorescence intensity of live cells (non-labelled with Propidium Iodide – PI) was evaluated along time - 0, 24, 48 and 72 h.p.i. - by means of flow cytometry. Solid lines represent GFP mean fluorescence intensity of infected cells; dashed lines represent GFP mean fluorescence intensity of non-infected cells. SNR performance evaluation by flow cytometry of (**B**) the backbones and (**D**) alternative cleavable sequences along time after AdV infection. Statistical analyses of triplicates (n=3) using ANOVA F-test (HSD of Tukey, * $p < 0.05$ and ** $p < 0.005$.

In conclusion, cGFP system activated upon infection as monitored by the increase in fluorescence and was validated for protease cleavage by western blotting analysis. Nonetheless, it is observed GFP fluorescence noise emission from non-infected cells (background fluorescence), suggesting that the performed structural distortions are not be enough to completely abolish the fluorescence emission. Therefore, in order to tackle background fluorescence emission, maximizing SNR performances, two other systems were developed based on deformation of Split-GFP, eS11 and cS11.

4.2. EMBEDDED SPLIT-GFP VISENSOR

A new strategy was followed by replacing the natural loop of a small protein, Eglin c, by the GFPS11 fragment to produce structural distortion. This structural distortion hampers the transcomplementation of split-GFP fragments, reducing fluorescence emission. Upon proteolytic cleavage, the structural distortion is relieved and, consequently, fluorescence emission is reconstituted by transcomplementation of GFPS10 and GFPS11 fragments. Our eS11 system consists on a mammalian optimized version of Callahan et al. [81], harbouring Eglin c backbone protein and an AdV specific cleavable sequence, as depicted in **Figure 4.5**.

4.2.1. OPTIMIZATION OF BACKBONE AND CLEAVABLE SEQUENCE

An initial characterization of the biosensor was performed by means of transient co-transfection (eS11 sensor, GFPS10 and AdV protease or a mock plasmid) in order to optimize SNR performance. First, we started by testing different backbones, with or without the addition of glycine residues spacers surrounding the cleavage site (denoted as eG/LRGAG/G-S11 and eLRGAG-S11, respectively), as in **Figure 4.5**.

Again and as seen by fluorescence microscopy, upon co-transfection of sensor with Adenain there is an increase in fluorescence emission, compared to the negative control (co-transfection with a mock plasmid), suggesting that the system is indeed functional (**Figure 4.6A**). Flow cytometry analysis showed that eLRGAG-S11 backbone had a marginally higher SNR performance (2.25±0.11) than eG/LRGAG/G-S11 backbone (2.07±0.02) (**Figure 4.6B**). Therefore, we proceeded with the screening of alternative cleavable sequences with eLRGAG-S11 backbone (without glycine spacers surrounding the cleavable sequence).

Taking in consideration the knowledge gathered from cGFP strategy, to compare alternative cleavable sequences to the synthetic LRGA↓G, we screened only IVGL↓G and EEGE↓G sequences since both led to dramatic differences in fluorescence emission and SNR performance (**Figure 4.2**).

Similarly to cGFP strategy, the usage of different cleavable sequences induced dramatic alterations in noise biosensor fluorescence emission, with IVGL↓G sequence decreasing it and EEGE↓G increasing it (**Figure 4.6A**). Despite the differences in fluorescence emission, surprisingly, SNR performance was found to be similar between all the tested sequences (IVGL↓G 2.22±0.04, LRGA↓G 2.25±0.11 and EEGE↓G 2.28±0.03) (**Figure 4.6C**). Therefore, we chose to proceed with eLRGAG-S11 as the best backbone and cleavable sequence combination.

Figure 4.5 – Schematic representation of eS11 system. The GFPS11 (denoted in green) together with the cleavable sequence (denoted in red) were inserted as a surface loop on the small Eglin c protein (denoted in blue), in order to create a structural distortion on GFPS11 fragment, preventing transcomplementation with GFPS10 fragment until viral protease cleavage. To increase the flexibility of the system, glycine residues were added (spacers denoted in black).

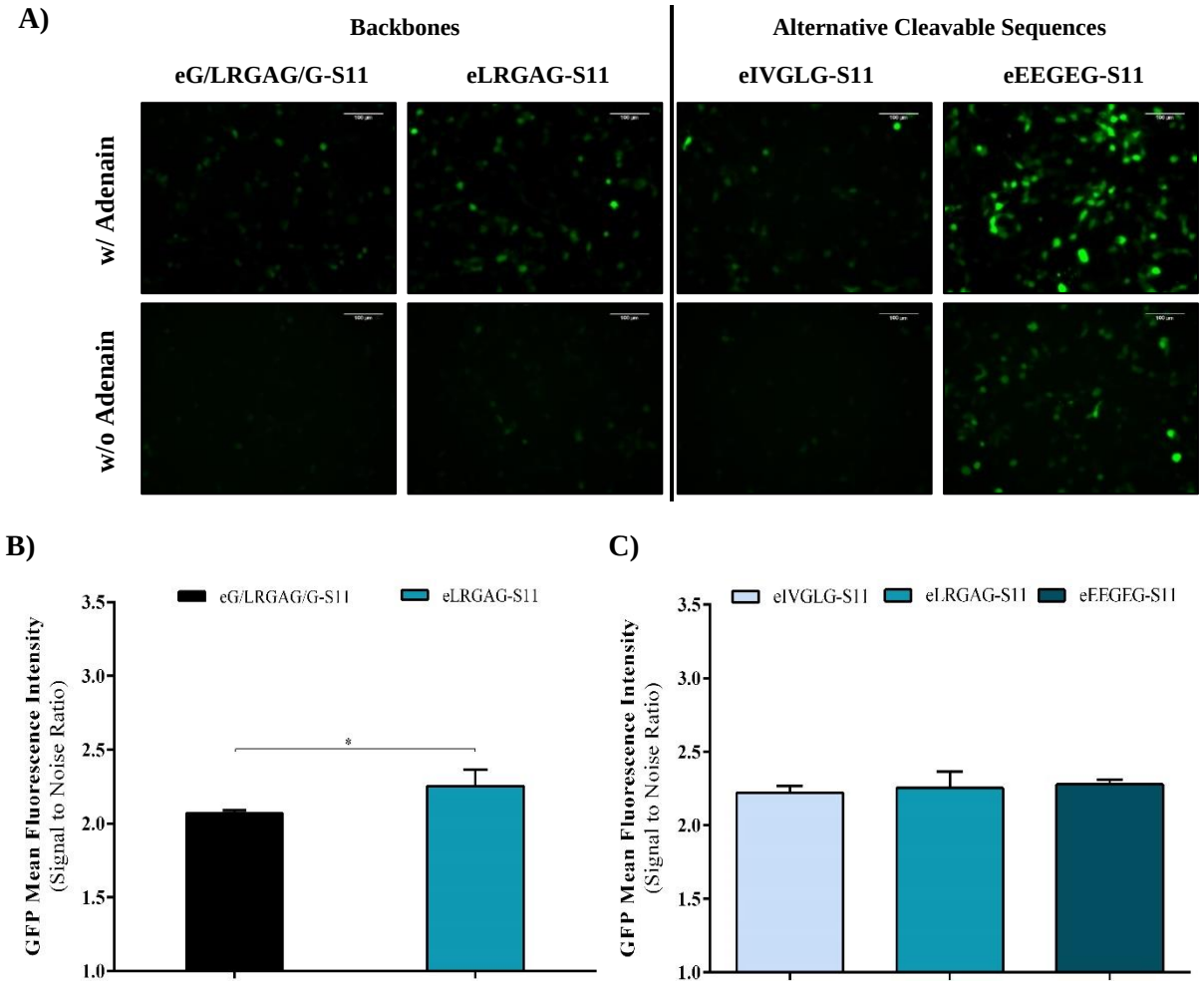

Figure 4.6 – Backbones and alternative sequences optimization of eS11 system. HEK 293T cells were transiently co-transfected with GFPS10, different eS11 backbones - eG/LRGAG/G-S11 with glycine spacer surrounding the cleavage site; eLRGAG-S11 without glycine spacers surrounding the cleavage site - and alternative sequences - IVGL↓G and EEGE↓G – with AdV protease or a mock plasmid (containing mCherry fluorescent reporter). **(A)** Fluorescence microscopy images were acquired 48 hours post-transfection. Scale Bar: 100 µm. **(B)** SNR performance was evaluated by flow cytometry 48 hours post-transfection. SNR shown as average ± standard deviation of at least triplicates (n=3). Statistical analyses using ANOVA F-test (HSD of Tukey).

4.2.2. HEK 293 CELL LINE STABLY EXPRESSING eS11 SYSTEM

After initial optimization by transient co-transfection, a comprehensive characterization was done to assess eS11 biosensor response to AdV infection. Since alternative cleavable sequences did not improve SNR performance, we established by lentiviral transduction HEK 293 cells stably expressing GFPS10 and each one of the different eS11 backbones with LRGA↓G cleavable sequence (**Figure 4.7**).

Similarly to what was observed by transient screening (**Figure 4.6**), eS11 sensors stably expressed in HEK 293 cells showed a similar pattern of fluorescence intensity over time (**Figure 4.7A**). Also, for both sensors the fluorescence emission of non-infected cells (background fluorescence) decreased over time (dashed lines). Confirming the results obtained by transient co-transfection (**Figure 4.6B**), eLRGAG-S11 sensor had the best SNR performance within this eS11 strategy, reaching 2.17±0.04 at 72 h.p.i (**Figure 4.7B**).

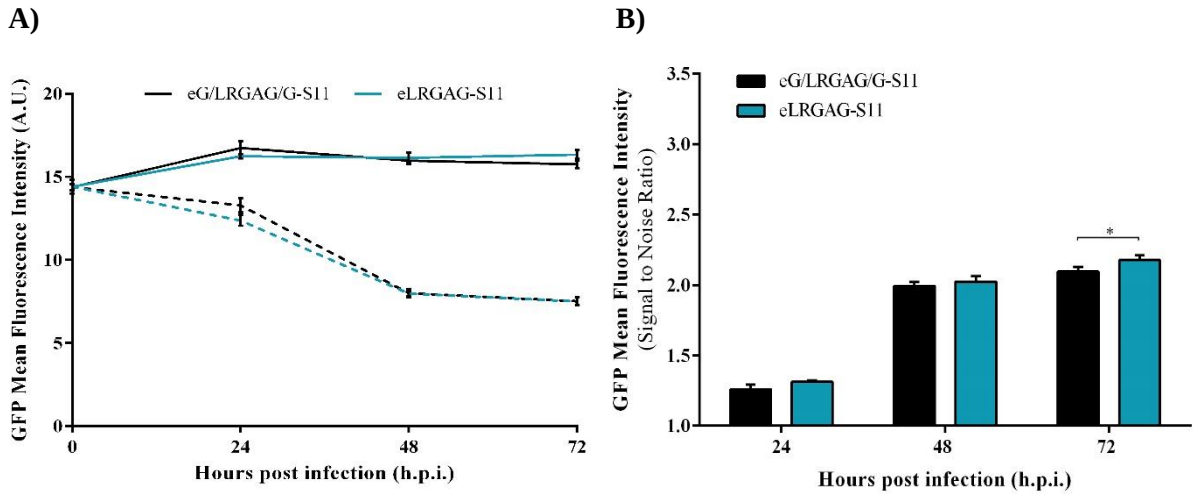

Figure 4.7 - Characterization of HEK 293 cells stably expressing eS11 sensor. HEK 293 cells stably expressing two different eS11 backbones, both with LRGA↓G cleavable sequence, were infected with AdV5 at a MOI of 5. **(A)** GFP fluorescence intensity of live cells (non-labelled with PI) was evaluated along time - 0, 24, 48 and 72 h.p.i. - by means of flow cytometry. Solid lines represent GFP mean fluorescence intensity of infected cells; dashed lines represent GFP mean fluorescence intensity of non-infected cells. **(B)** SNR performance evaluation by flow cytometry along time after AdV infection. Statistical analyses of triplicates (n=3) using non-paired t-test, * $p < 0.05$

4.3. CIRCULAR SPLIT-GFP VISENSOR

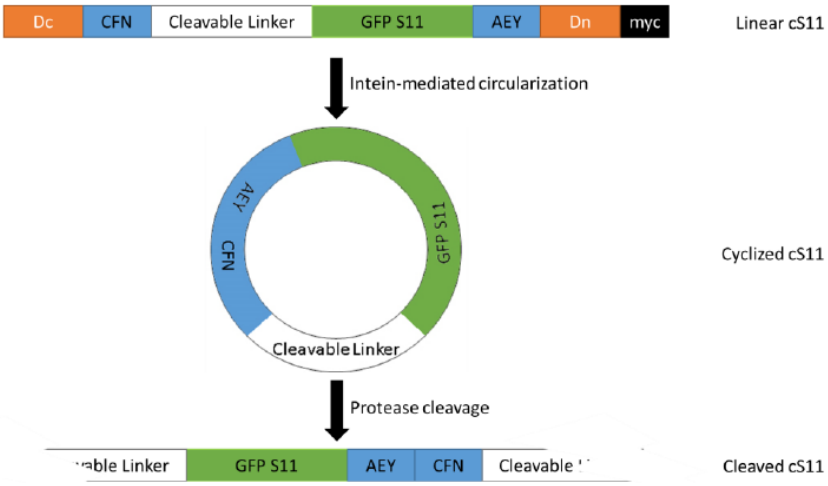

Figure 4.8 - Circularization process of cS11 system. Similarly to the system developed by Sakamoto et al. [93], GFPS11 fragment was inserted between the *Npu* DnaE split-intein (Linear cS11). A cis-splicing reaction at the level of the exteins leads to the excision of the inteins, circularizing the polypeptide (cyclized cS11). Once the distortion is removed by the cleavage of the cleavable linker by the viral protease, GFPS11 is able to transcomplement with GFPS10 fragment, recovering fluorescence emission. Dc, C-fragment of *Npu* DnaE intein; Dn-myc, N-fragment of *Npu* DnaE intein fused with a myc tag; CFN and AEY are the conserved extein residues for the splicing by *Npu* DnaE intein.

Our second strategy for structural distortion of fluorescent split-proteins relied on circularization of GFPS11 fragment mediated by *Npu* DnaE split-intein splicing reaction. This structural distortion diminishes self-association with GFPS10 fragment, preventing fluorescence emission. Once proteolytic cleavage takes place, the two fragments from split-GFP can transcomplement and reconstitute fluorescence emission. Originally, Sakamoto et al. [93] approach was developed for the detection of caspase-3 like protease. We performed a direct adaptation by replacing the DEVD↓G cleavable linker with the synthetic LRGA↓G recognized by the AdV protease, as depicted in **Figure 4.8**.

4.3.1. OPTIMIZATION OF BACKBONE AND CLEAVABLE SEQUENCE

Similarly to the optimizations performed for the former strategies, the first step was to assess the best backbone structure surrounding the cleavage linker by transient co-transfection (cS11 sensor, GFPS10 and AdV protease or a mock plasmid). Three different backbones were considered: cLRGAG-S11, cG/LRGAG/G-S11 and cGG/LRGAG/GG-S11. Again, addition of glycine spacers was evaluated for impact on sensor fluorescence emission and/or cleavage efficiency (as seen by effect on SNR performance).

As seen by fluorescence microscopy (**Figure 4.9A)** progressively higher levels of noise fluorescence intensity were observed when adding one or two glycine residues in each side of the cleavable linker. Nevertheless, all cS11 backbones showed similar SNR performance (**Figure 4.9B**). To further characterize this strategy, we chose to proceed with the backbone cG/LRGAG/G-S11 (SNR 2.73±0.24). To study the impact of different cleavable sequences in sensor performance, a similar approach to eS11 sensor strategy was followed by comparing only IVGL↓G, LRGA↓G and EEGE↓G cleavable sequences. As seen by fluorescence microscopy (**Figure 4.9A)**, EEGE↓G cleavable sequence led to an increase in noise fluorescence emission. Additionally, it had a negative impact on SNR performance (2.40±0.10) when compared with LRGA↓G sequence (2.73±0.24), as depicted in **Figure 4.9C**. No significant differences were observed in SNR performance between IVGL↓G and LRGA↓G sequences (**Figure 4.9C**).

4.3.2. HEK 293 CELL LINE STABLY EXPRESSING cS11 SYSTEM

At this step, HEK 293 cell lines stably expressing GFPS10 and each one of the three cS11 backbones with LRGA↓G sequence (since alternative cleavable sequences did not further improved SNR performance) were established by lentiviral transduction and further characterized for their response to AdV infection. Confirming the results obtained in transient co-transfection (**Figure 4.9A**) addition of glycine spacers lead to an increase in sensor fluorescence emission (**Figure 4.9D**). Again, and for all backbones, GFP fluorescence of non-infected cells (background fluorescence) decreased over time (dashed lines).

Comparing the performance of the cS11 backbones over time, a similar SNR performance was observed until 48 h.p.i. (**Figure 4.9E**). However, at 72 h.p.i. cLRGAG showed the lowest SNR, with no significant differences between cG/LRGAG/G (3.12 ± 0.18) and cGG/LRGAG/GG (3.25 ± 0.06).

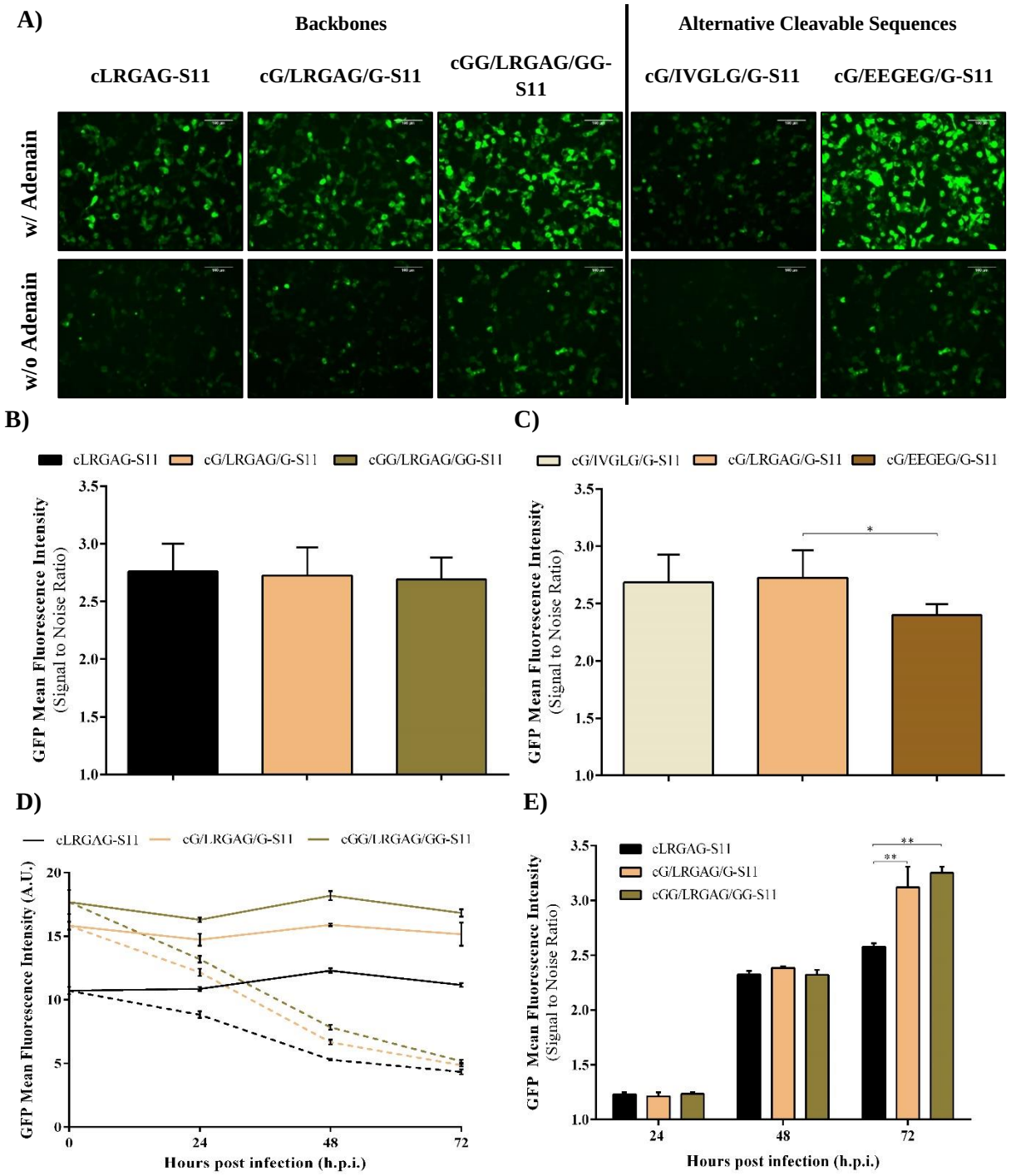

Figure 4.9 – Backbones and alternative sequences optimization and characterization of cS11 sensor. HEK 293T cells were transiently co-transfected with GFPS10, different cS11 backbones - cLRGAG-S11, cG/LRGAG/G-S11 with addition of a glycine spacer by each side of the cleavage site, and cGG/LRGAG/GG-S11 with addition of two glycine spacers - or alternative sequences - IVGL↓G and EEGE↓G – with Adenain or a mock plasmid (containing mCherry fluorescent reporter). **(A)** Fluorescence microscopy images were acquired 48 hours post-transfection. Scale Bar: 100 µm. **(B, C)** SNR performance was evaluated by flow cytometry 48 hours post-transfection. SNR shown as average ± standard deviation of at least triplicates (n=3). Statistical analyses using ANOVA F-test (HSD of Tukey), * $p < 0.05$. HEK 293 cells stably expressing three different cS11 backbones, all with LRGA↓G cleavable sequence, were infected with AdV at a MOI of 5. **(D)** GFP fluorescence intensity of live cells (non-labelled with PI) was evaluated along time - 0, 24, 48 and 72 h.p.i. - by means of flow cytometry. Solid lines represent GFP mean fluorescence intensity of infected cells; dashed lines represent GFP mean fluorescence intensity of non-infected cells. **(E)** SNR performance evaluation by flow cytometry along time after AdV infection. Statistical analyses of triplicates (n=3) using ANOVA F-test (HSD of Tukey), ** $p < 0.005$.

4.4. COMBINED DISTORTION OF GFPS10 AND GFPS11 FRAGMENTS

On a parallel strategy, and to further reduce background fluorescence levels (due to insufficient structural distortion and/or unspecific transcomplementation of GFPS11 and GFPS10 fragments), structural distortion on both GFP fragments was evaluated. A preliminary transient co-transfection screening was performed with cG/LRGAG/G-S11 (best cS11 sensor) and cGFPS10 (**Figure 4.10**). cGFPS10 consists on the circularization of GFPS10 fragment by means of *Npu* DnaE split-inteins, similarly to the cGFP strategy. Additionally, eLRGAG-S11 (best eS11 sensor) in combination with cGFPS10 was also tested. Both strategies were evaluated against wild-type GFPS10 fragment (as used previously).

Surprisingly, cG/LRGAG/G-S11 sensor combined with cGFPS10 resulted in a drastic reduction in SNR performance (1.85±0.06) when compared with the previous strategy using wild-type GFPS10 (SNR 2.73±0.24) (**Figures 4.9B** and **4.10B**). This reduction on SNR might be related to the increase of GFP fluorescence emission of the non-activated sensor (i.e., in absence of Adenain), as seen by fluorescence microscopy (**Figure 4.10A**). On the other hand, no fluorescence emission was detected on eLRGAG-S11 combined with cGFPS10 (SNR 1.01±0.02) (**Figure 4.10**).

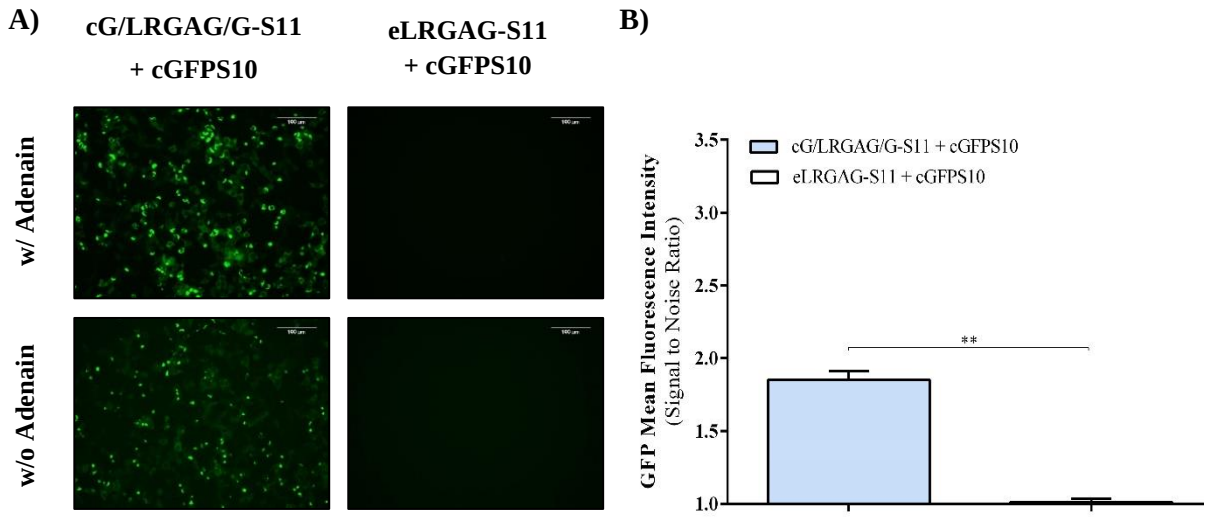

Figure 4.10 - Evaluation of combined distortion of GFPS10 and GFPS11 fragments. HEK 293T cells were transiently co-transfected with cG/LRGAG/G-S11 with cGFPS10 or eLRGAG-S11 with GFPS10/cGFPS10 - and AdV protease or a mock plasmid (containing mCherry fluorescent reporter). cGFPS10, circular distortion of GFPS10 fragment. **(A)** Fluorescence microscopy images were acquired 48 hours post-transfection. Scale Bar: 100 µm. **(B)** SNR performance was evaluated by flow cytometry 48 hours post-transfection. SNR shown as average ± standard deviation of at least triplicates (n=3). Statistical analyses using unpaired two-tailed t-test, ** $p < 0,005$.

5. DISCUSSION AND CONCLUSIONS

Due to the great potential of VBBs, over the last few years there has been an increased interest in the development of new therapies based on viral vectors, like gene therapy or vectored vaccines. Therefore, reliable and fast quantification methods of viral vectors are of outmost importance. However, existing titration methods fail on providing robust and fast quantification of viral vectors since, in general, they overestimate infectious viral titer, show lack of high-throughput potential and are time-consuming. Furthermore, many of them require vector modification such as the use of reporter genes, which may impair virus biology and, most importantly, are not permitted for clinical use due to safety issues. As so, the work developed during this thesis aimed at developing genetically encoded switch-on fluorescent cell-based assays for detection (and quantification) of label-free viruses. As a proof-of-concept, these assays were developed for label-free AdV, the most used viral vector in gene therapy (**Figure 1.1**), and took advantage of the adenoviral protease, Adenain, as a trigger for fluorescence switch-on. Three different main strategies were designed based on structural distortion of a fluorescent protein, hampering fluorescence emission until it is relieved upon proteolytic processing by Adenain (**Figure 2.1**). The first strategy, consists on a circularized GFP mutant, cGFP, generated by means of split-inteins, as shown in **Figure 4.1**. Inteins are protein segments able to excise itself and join the remaining portions (the exteins) with a peptide bond in a process named protein splicing, analogously to intron RNA splicing [105 106]. In order to achieve maximum distortion with concomitant decrease in background fluorescence emission, as well as high cleavage site accessibility towards improved SNR performances, two other strategies were developed taking advantage of the split-GFP system. Embedded GFPS11 (eS11) strategy uses the natural loop of Eglin c protein to create structural distortion on GFPS11 fragment (**Figure 4.5**). On the other hand, circular GFPS11 (cS11) takes advantage of split-inteins splicing reaction to produce a circularized GFPS11 (**Figure 4.8**). Both distortions prevent self-assembly of the GFP fragments, diminishing fluorescence emission.

Different structural distortions on the sensor may affect not only background fluorescence emission but also linker cleavage efficiency. For that reason, an initial characterization by transient co-transfection was performed to study use of different backbone structures aiming to reach maximum distortion and to lower background (non-activated sensor) fluorescence emission, without compromising the GFP chromophore or diminishing cleavage site accessibility. A direct adaptation of Zhang et al. [92] system was initially tested in cGFP strategy by replacing the cleavable linker with the synthetic sequence LRGA↓G [86 100] (cLRGAG w/ Met). Since SNR performance was low (1.41±0.06), and to improve sensor performance, a methionine residue located at the C-terminal end of sfGFP was removed (cLRGAG) and an increase SNR of 30% was observed (1.85±0.09) (**Figure 4.2B**). This suggested that the deletion of the methionine residue influenced sensor distortion, since less amino acids could lead to higher distortion; therefore, background fluorescence is lower resulting on a higher SNR performance. Additional backbones were tested with the methionine deletion and the addition of flexible linkers surrounding the cleavage site. Glycine is a small and flexible amino acid [107 108], hence we hypothesized that the addition of glycine spacers by each side of the cleavage site would increase cleavable site flexibility and exposure to adenoviral protease. Fluorescence microscopy demonstrated that, for all strategies, the addition of glycine residues increased sensor fluorescence emission both in the presence/absence of Adenain, suggesting that the increased sensor flexibility leads to a decreased sensor structural distortion (**Figures 4.2A**, **4.6A** and **4.9A**). This effect resulted in an intensification of background fluorescence emission affecting SNR performance, as seen in eS11 (**Figure 4.6A**) and cS11 (with two glycine spacers, **Figure 4.9A**) strategies. Therefore, in cGFP (**Figure 4.2B**) and eS11 (**Figure 4.6B**) strategies the best SNR performance was achieved when no glycine spacers were added to the backbone structure (cLRGAG and eLRGAG-S11, respectively).

In addition to backbone structure, differences in the cleavable sequence itself may also have an impact in background level. Adenain recognizes different cleavable sequences, which differ by sensitivity and specificity. Usually, Adenain cleaves at consensus cleavage sequences: (M/I/L)XGX↓G (type 1) and (M/I/L)XGG↓X (Type 2), showing preferential cleavage for type 1 sequences [86]. Thus, natural cleavable sequences, present in the AdV viral proteins, as well as a synthetic sequence LRGA↓G [86 100] (**Table 4.1**) were evaluated by transient co-transfection. Indeed, in all strategies, different cleavable sequences induced impressive variations in biosensor fluorescence emission, with sensors harbouring the IVGL↓G cleavable sequence displaying the lowest fluorescence emission and EEGE↓G leading to an increase in fluorescence emission. Consequently, this effect influenced SNR performance in cGFP (**Figure 4.2C**) and cS11 (**Figure 4.9C**) strategies, when compared with LRGA↓G cleavable sequence. A possible explanation for these results lies on the hydrophilicities of the amino acids that compose IVGL↓G and EEGE↓G sequences. Residues like Isoleucine (I), Valine (V) and Leucine (L) are hydrophobic amino acids. Whereas, Arginine (R), Lysine (K) and Glutamic acid (E) are hydrophilic [109]. Therefore, we may assume that the prevalence of hydrophobic amino acids in IVGL↓G cleavable sequence had a negative impact in sensor folding and/or cleavable sequence recognition, generating a bigger structural distortion and low cleavage site exposure. On the other hand, EEGE↓G has in its composition more hydrophilic amino acids, whereby the sensor's folding may be less affected by distortion and/or the cleavable sequence could be more exposed. Still, LRGA↓G sequence has shown to be the best approach for further studies in all strategies.

To validate structurally distorted VISENSORS were being activated by Adenain (relief of structural distortion with concomitant increase in fluorescence emission), and also to confirm their specificity, cell extracts from transient co-transfection were made and analyzed by Western Blot (**Figure 4.3**). This validation was made for cGFP strategy, so we hypothesize it is also occurring for the other strategies, but it should be confirmed. The results demonstrated full circularization of cGFP, with the release of the 13 kDa Dn myc-tagged split-intein and absence of the 44 kDa linear cGFP. Moreover, the appearance of a band corresponding to cleaved cGFP (27 kDa) when using a functional Adenain but not when using a non-functional mutant Adenain or HIV-1- protease validates the use of VISENSORS as a specific Adenain dependent sensor. Surprisingly, in all conditions we observed a band at <54 kDa only detectable with anti-GFP antibody. We hypothesize that is a cyclized bimolecular cGFP that, theoretically, could also work as a biosensor like cGFP. However, its larger size may be leading to a lower structural distortion and thus contributing to a higher background fluorescence level.

The results obtained from the initial transient transfection screening were promising, nonetheless a more comprehensive characterization was needed in a closer biological context with AdV infection. Therefore, HEK 293 cells stably expressing the VISENSORS were stablished by lentiviral transduction. In this context and regarding optimizations of backbones and cleavable sequences, similar results were obtained to transient screenings (**Figures 4.2**, **4.6** and **4.9**). As so, the best SNR performance was obtained with cLRGAG (as backbone and cleavable sequence) reaching a SNR of 2.14 ±0.07 at 48 h.p.i. and 2.74 ± 0.39 at 72 h.p.i. for cGFP strategy (**Figure 4.4**). Characterization in HEK 293 cells stably expressing the eS11 and cS11 strategies combined with IVGL↓G and EEGE↓G sequences was not performed since no further improvement of SNR performance were displayed. Surprisingly, in all strategies, we observed a decrease in the fluorescence emission of non-infected cells (background fluorescence) along time (**Figures 4.4A/C**, **4.7A** and **4.9D**). We can only hypothesize that, in contrast with AdV infected cells whose growth is arrested, non-infected cells continue to grow, reaching a stationary phase. At this phase, protein expression (such as our sensors) is decreased and, consequently, fluorescence emission may decrease. Therefore, for SNR performance not only we have a contribution of an increase in fluorescence emission due to sensor being cleaved by Adenain (Signal) but also a decrease in background fluorescence of control cells (Noise). More importantly, SNR performances were better in stable cells with AdV infection than in transient screenings. In the latter, Adenain might

not be fully activated, despite the fusion with pVIc, since there is no viral DNA to achieve a full activation [88 89]. Indeed, by Western Blotting analysis (**Figure 4.3**) its observable that only a small portion of cyclized GFP is cleaved. In stable cells with AdV infection, we can hypothesize that AdV protease has a higher catalytic activity since it is in presence of both pVIc and viral DNA co-factors, allowing its full activation.

With the aim of achieving maximum distortion and improving SNR performances, the two split-GFP-based strategies were developed (eS11 and cS11). And indeed, despite eS11 and cGFP displaying similar SNR performances, cS11 strategy seems the most promising. Comparing against cLRGAG, eLRGAG-S11 and cG/LRGAG/G-S11 have 22% and 48% higher SNR performance in transient screening (**Figures 4.2, 4.6 and 4.9,** respectively). In stable screening, however, eLRGAG-S11 displays a decrease of 13% of SNR performance (**Figure 4.7B**) and cG/LRGAG/G-S11 shows a modest 14% increased performance (**Figure 4.9E**) against cLRGAG strategy (**Figure 4.4B**).

As a parallel strategy, we performed structural distortion on both split-GFP fragments. To that end, eS11 and cS11 were co-transfected with cyclized GFPS10 instead of wild-type GFP. To our surprise, SNR performance in this double distortion strategy was lower (**Figure 4.10B**). Regarding eS11 with cGFPS10, no fluorescence emission was observed (**Figure 4.10A**). A possible explanation for these results relies on strategy incompatibility or cGFPS10 circularization process impairing the GFP chromophore maturation during transcomplementation. In the other hand, the combination of cS11 and cGFPS10 shows an increase of sensor fluorescence (both in the presence or absence of Adenain), which resulted on a decrease of SNR performance (**Figure 4.9B and 4.10B**). As the cyclization process of both GFP fragments was mediated by *Npu* DnaE split-inteins, we hypothesize the split-inteins, instead of *cis*-splicing of GFPS10 and GFPS11, are trans-splicing both fragments resulting in an active GFP, able to produce fluorescence [110]. In order to avoid intein cross reactivity, inteins who shared limited sequence similarity (for example, orthogonal inteins), such as *Npu* DnaE and *Ssp* DnaB, could be used.

As future work, a detailed characterization of the best strategy must be performed, using HEK 293 cell clones stably expressing cG/LRGAG/G-S11. Since the ratio of GFPS11 and GFPS10 fragments could affect sensor performance, a screening of several cell clones should be conducted with the purpose of finding the cS11 clones displaying higher SNR performance than HEK 293 cS11 cell population, herein used (where SNR performance represents an average of all the different cells). Having picked the best cell clone in terms of SNR performance, the in-deep characterization should consist on: studying clone permissiveness to AdV replication, to confirm if the sensor does not impair AdV life cycle and can be used for live-cell monitoring of AdV infection; sensor activation kinetics, in order to know the optimal time point of analysis after infection and how it responds to different AdV MOIs. Moreover, in order to investigate sensor applicability as a quantification method, the elaboration of a calibration curve using different AdV MOIs should also be performed.

In conclusion, the work conducted during this master thesis delivered three different strategies for the detection of label-free viruses, based on structural distortion of fluorescent proteins and switch-on fluorescence emission upon viral protease processing. To the best of our knowledge, this is the first study developing these structural distortion strategies as a mammalian cell-based system for detection of label-free AdV. Additionally, the approaches herein characterized might be extended for other replicative label-free viruses. Therefore, VISENSORS have the potential to deliver a fast, reliable and accurate method for virus and viral vector detection and quantification, much needed not only in the development of gene therapy and vaccination fields, but also for diagnostic and virus research.

REFERENCES

1. Dye, C. After 2015: infectious diseases in a new era of health and development. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 369, 20130426 (2014).
2. Global Crop Protection Chemicals Market - Industry Analysis, Size, Share, Trends, Segment and Forecast 2014-2020. *Zion Research Analysis 2016* at <http://www.marketresearchstore.com/report/vaccine-market-z53045>
3. Rodrigues, A. F., Carrondo, M. J. T., Alves, P. M. & Coroadinha, A. S. Cellular targets for improved manufacturing of virus-based biopharmaceuticals in animal cells. *Trends Biotechnol.* 32, 602–607 (2014).
4. Rodrigues, A. F., Soares, H. R., Guerreiro, M. R., Alves, P. M. & Coroadinha, A. S. Viral vaccines and their manufacturing cell substrates : New trends and designs in modern vaccinology. 2–17 (2015). doi:10.1002/biot.201500387
5. Chackerian, B. Virus-like particles : flexible platforms for vaccine development. *Expert Rev. Vaccines* 6, 381–390 (2007).
6. Iemsam-Arng, J., Kong, X., Schätzlein, A. G. & Uchegbu, I. F. in *Fundamentals of Pharmaceutical Nanoscience* 493–510 (Springer New York, 2013). doi:10.1007/978-1-4614-9164-4_18
7. Verma, I. M. & Weitzman, M. D. GENE THERAPY: Twenty-First Century Medicine. *Annu. Rev. Biochem.* 74, 711–738 (2005).
8. Ponder, K. P. in *An Introduction to Molecular Medicine and Gene Therapy* 77–112 (John Wiley & Sons, Inc.). doi:10.1002/0471223875.ch4
9. Misra, S. Human gene therapy: a brief overview of the genetic revolution. *J. Assoc. Physicians India* 61, 127–133 (2013).
10. Wirth, T., Parker, N. & Ylä-herttuala, S. History of gene therapy. *Gene* 525, 162–169 (2013).
11. Curiel, D. *Adenoviral Vectors for Gene Therapy*. (Academic Press).
12. Robbins, P. D. & Ghivizzani, S. C. Viral Vectors for Gene Therapy. 80, 35–47 (1998).
13. Pearson, S., Jia, H. & Kandachi, K. China approves first gene therapy. *Nat. Biotechnol.* 22, 3–4 (2004).
14. Peng, Z. Current status of gendicine in China: recombinant human Ad-p53 agent for treatment of cancers. *Hum. Gene Ther.* 16, 1016–27 (2005).
15. Geraerts, M., Willems, S., Baekelandt, V., Debyser, Z. & Gijsbers, R. Comparison of lentiviral vector titration methods. *BMC Biotechnol.* 6, 34 (2006).
16. Leland, D. S. & Ginocchio, C. C. Role of cell culture for virus detection in the age of technology. *Clin. Microbiol. Rev.* 20, 49–78 (2007).
17. Pankaj, K. Methods for Rapid Virus Identification and Quantification. *Labome* 3, 1–7 (2013).
18. Baer, A. & Kehn-Hall, K. Viral Concentration Determination Through Plaque Assays: Using Traditional and Novel Overlay Systems. *J. Vis. Exp.* e52065 (2014). doi:10.3791/52065
19. Lieber, D. & Bailer, S. M. Determination of HSV-1 infectivity by plaque assay and a luciferase reporter cell line. *Methods Mol. Biol.* 1064, 171–181 (2013).
20. Flint, J., Racaniello, V. R., Rall, G. F. & Skalka, A. M. *Principles of Virology, Volume I: Molecular Biology*. *2014* (American Society of Microbiology, 2015). doi:10.1128/9781555818951
21. Smither, S. J. *et al.* Comparison of the plaque assay and 50 % tissue culture infectious dose assay as methods for measuring filovirus infectivity. *J. Virol. Methods* 193, 565–571 (2013).
22. Fritschy, J. & Hartig, W. Immunofluorescence. *Encycl. Life Sci.* 1–7 (1999). doi:10.1038/npg.els.0001174
23. Beutner, E. H. Immunofluorescent Staining: the Fluorescent Antibody Method. *Bacteriol. Rev.* 25, 49–76 (1961).
24. Kvinesdal, B. B., Nielsen, C. M., Poulsen, A. G. & Højlyng, N. Immunofluorescence assay for detection of antibodies to human immunodeficiency virus type 2. *J. Clin. Microbiol.* 27, 2502–4 (1989).
25. Lutz, A., Dyall, J., Olivo, P. D. & Pekosz, A. Virus-inducible reporter genes as a tool for detecting and quantifying influenza A virus replication. *J. Virol. Methods* 126, 13–20 (2005).

26. Rodrigues, A. F. *et al.* Single-step cloning-screening method: A new tool for developing and studying high-titer viral vector producer cells. *Gene Ther.* 22, 685–695 (2015).
27. Kung, S.-H. Reporter Cell Lines for the Detection of Herpes. *DNA Viruses* 292, 73–81 (2005).
28. Mullis, K. B. & Faloona, F. A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* 155, 335–350 (1987).
29. Souf, S. Recent advances in diagnostic testing for viral infections. *Biosci. Horizons* 9, 1–11 (2016).
30. Delenda, C. & Gaillard, C. Real-time quantitative PCR for the design of lentiviral vector analytical assays. *Gene Ther.* 12 Suppl 1, S36-50 (2005).
31. Carmo, M. *et al.* Quantitation of MLV-based retroviral vectors using real-time RT-PCR. *J. Virol. Methods* 119, 115–119 (2004).
32. Magi, B. & Liberatori, S. in *Immunochemical Protocols* 227–253 (Humana Press, 2005). doi:10.1385/1-59259-873-0:227
33. Schochetman, G., Epstein, J. S. & Zuck, T. F. Serodiagnosis of infection with the AIDS virus and other human retroviruses. *Annu Rev Microbiol* 43, 629–659 (1989).
34. Zaaijer, H. L. *et al.* Reliability of the third-generation recombinant immunoblot assay for hepatitis C virus. *Transfusion* 35, 745–749 (1995).
35. Shah, K. & Maghsoudlou, P. Enzyme-linked immunosorbent assay (ELISA): the basics. *Br. J. Hosp. Med.* 77, C98–C101 (2016).
36. Worldwide, M. I. Viruses and Viral Vaccines : Characterization by Nanoparticle Tracking Analysis. 1–12 at <https://cdn.technologynetworks.com/TN/Resources/PDF/WP140912VirusVaccines.pdf>
37. Kramberger, P., Ciringer, M., Štrancar, A. & Peterka, M. Evaluation of nanoparticle tracking analysis for total virus particle determination. *Virol. J.* 9, 265 (2012).
38. Mehrotra, P. Biosensors and their applications - A review. *Journal of Oral Biology and Craniofacial Research* 6, 153–159 (2016).
39. Kizek, R. *et al.* Nanoscale virus biosensors: state of the art. *Nanobiosensors Dis. Diagnosis* 4, 47–66 (2015).
40. Cheng, M. S. & Toh, C.-S. Novel biosensing methodologies for ultrasensitive detection of viruses. *Analyst* 138, 6219–29 (2013).
41. Ueda, H. & Dong, J. From fluorescence polarization to Quenchbody: Recent progress in fluorescent reagentless biosensors based on antibody and other binding proteins. *Biochim. Biophys. Acta - Proteins Proteomics* 1844, 1951–1959 (2014).
42. Zhai, J., Cui, H. & Yang, R. DNA based biosensors. *Biotechnol. Adv.* 15, 43–58 (1997).
43. Nielsen, P., Egholm, M., Berg, R. & Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science (80-.).* 254, 1497–1500 (1991).
44. Abdul Rahman, S. *et al.* Label-Free Dengue Detection Utilizing PNA/DNA Hybridization Based on the Aggregation Process of Unmodified Gold Nanoparticles. *J. Nanomater.* 2014, 1–5 (2014).
45. Taniguchi, A. Live cell-based sensor cells. *Biomaterials* 31, 5911–5915 (2010).
46. Ozkan, C. S. *et al.* in *BioMEMS and Biomedical Nanotechnology* 55–92 (Springer US, 2006). doi:10.1007/978-0-387-25845-4_4
47. Tepeli, Y. & Ülkü, A. Electrochemical biosensors for influenza virus a detection: The potential of adaptation of these devices to POC systems. *Sensors Actuators B Chem.* 254, 377–384 (2018).
48. Zhu, X., Ai, S., Chen, Q., Yin, H. & Xu, J. Label-free electrochemical detection of Avian Influenza Virus genotype utilizing multi-walled carbon nanotubes–cobalt phthalocyanine–PAMAM nanocomposite modified glassy carbon electrode. *Electrochem. commun.* 11, 1543–1546 (2009).
49. Krejcova, L. *et al.* Using of paramagnetic microparticles and quantum dots for isolation and electrochemical detection of influenza viruses' specific nucleic acids. *Int. J. Electrochem. Sci.* 8, 689–702 (2013).
50. Silva, M. M. S. *et al.* A thiophene-modified screen printed electrode for detection of dengue virus NS1 protein. *Talanta* 128, 505–510 (2014).
51. Chung, D.-J., Kim, K.-C. & Choi, S.-H. Electrochemical DNA biosensor based on avidin–biotin conjugation for influenza virus (type A) detection. *Appl. Surf. Sci.* 257, 9390–9396 (2011).

52. Nidzworski, D., Pranszke, P., Grudniewska, M., Król, E. & Gromadzka, B. Universal biosensor for detection of influenza virus. *Biosens. Bioelectron.* 59, 239–242 (2014).
53. Hassen, W. M. *et al.* An impedimetric DNA sensor based on functionalized magnetic nanoparticles for HIV and HBV detection. *Sensors Actuators, B Chem.* 134, 755–760 (2008).
54. Wang, R. *et al.* Evaluation study of a portable impedance biosensor for detection of avian influenza virus. *J. Virol. Methods* 178, 52–58 (2011).
55. Lin, D., Tang, T., Jed Harrison, D., Lee, W. E. & Jemere, A. B. A regenerating ultrasensitive electrochemical impedance immunosensor for the detection of adenovirus. *Biosens. Bioelectron.* 68, 129–134 (2015).
56. Sin, M. L. Y., Mach, K. E., Wong, P. K. & Liao, J. C. Advances and challenges in biosensor-based diagnosis of infectious diseases. *Expert Rev. Mol. Diagn.* 14, 225–44 (2014).
57. Daniels, J. S. & Pourmand, N. Label-Free Impedance Biosensors: Opportunities and Challenges. *Electroanalysis* 19, 1239–1257 (2007).
58. Guo, X. Surface plasmon resonance based biosensor technique: A review. *J. Biophotonics* 5, 483–501 (2012).
59. Zheng, S., Kim, D.-K., Park, T. J., Lee, S. J. & Lee, S. Y. Label-free optical diagnosis of hepatitis B virus with genetically engineered fusion proteins. *Talanta* 82, 803–809 (2010).
60. Lee, J.-H. *et al.* Highly sensitive localized surface plasmon resonance immunosensor for label-free detection of HIV-1. *Nanomedicine Nanotechnology, Biol. Med.* 9, 1018–1026 (2013).
61. Kumbhat, S., Sharma, K., Gehlot, R., Solanki, A. & Joshi, V. Surface plasmon resonance based immunosensor for serological diagnosis of dengue virus infection. *J. Pharm. Biomed. Anal.* 52, 255–259 (2010).
62. Pattnaik, P. Surface Plasmon Resonance: Applications in Understanding Receptor–Ligand Interaction. *Appl. Biochem. Biotechnol.* 126, 079–092 (2005).
63. Tainaka, K. *et al.* Design Strategies of Fluorescent Biosensors Based on Biological Macromolecular Receptors. 1355–1376 (2010). doi:10.3390/s100201355
64. Tian, J., Zhao, H., Liu, M., Chen, Y. & Quan, X. Detection of influenza A virus based on fluorescence resonance energy transfer from quantum dots to carbon nanotubes. *Anal. Chim. Acta* 723, 83–87 (2012).
65. Lu, X. *et al.* A gold nanorods-based fluorescent biosensor for the detection of hepatitis B virus DNA based on fluorescence resonance energy transfer. *Analyst* 138, 642–650 (2013).
66. Orm, M. *et al.* Crystal Structure of the Aequorea victoria Green Fluorescent Protein. *Science (80-.).* 273, 1392–1395 (1996).
67. Crone, D. E. *et al.* GFP-Based Biosensors. *InTech* 3–36 (1996). doi:10.5772/52250
68. Shimomura, O., Johnson, F. H. & Saiga, Y. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea. *J. Cell. Comp. Physiol.* 59, 223–239 (1962).
69. Heim, R., Prasher, D. C. & Tsien, R. Y. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci.* 91, 12501–12504 (1994).
70. Nagai, T. *et al.* A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* 20, 87–90 (2002).
71. Cinelli, R. A. G. *et al.* The Enhanced Green Fluorescent Protein as a Tool for the Analysis of Protein Dynamics and Localization : Local Fluorescence Study at the Single-molecule Level. 71, 771–776 (2000).
72. Waldo, G. S., Standish, B. M., Berendzen, J. & Terwilliger, T. C. Rapid protein-folding assay using green fluorescent protein. *Nat. Biotechnol.* 17, 691–5 (1999).
73. Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* 24, 79–88 (2006).
74. Kamiyama, D. *et al.* Versatile protein tagging in cells with split fluorescent protein. *Nat. Commun.* 7, 1–9 (2016).
75. Yost, S. A. & Marcotrigiano, J. Viral precursor polyproteins: keys of regulation from replication to maturation. *Curr. Opin. Virol.* 3, 137–142 (2013).
76. López-Otín, C. & Bond, J. S. Proteases: Multifunctional enzymes in life and disease. *Journal of Biological Chemistry* 283, 30433–30437 (2008).
77. Clark, D. P., Pazdernik, N. J., Clark, D. P. & Pazdernik, N. J. in *Biotechnology* 295–333 (2016).

doi:10.1016/B978-0-12-385015-7.00009-0
78. Shekhawat, S. S., Porter, J. R., Sriprasad, A. & Ghosh, I. An autoinhibited coiled-coil design strategy for split-protein protease sensors. *J. Am. Chem. Soc.* 131, 15284–15290 (2009).
79. Iro, M. *et al.* A reporter cell line for rapid and sensitive evaluation of hepatitis C virus infectivity and replication. *Antiviral Res.* 83, 148–155 (2009).
80. Jones, C. T. *et al.* Real-time imaging of hepatitis C virus infection using a fluorescent cell-based reporter system. *Nat. Biotechnol.* 28, 167–171 (2010).
81. Callahan, B. P., Stanger, M. J. & Belfort, M. Protease Activation of Split Green Fluorescent Protein. *ChemBioChem* 11, 2259–2263 (2010).
82. Warnock, J. N., Daigre, C. & Al-Rubeai, M. Introduction to viral vectors. *Methods in molecular biology (Clifton, N.J.)* 737, 1–25 (2011).
83. Russell, W. C. Adenoviruses: Update on structure and function. *J. Gen. Virol.* 90, 1–20 (2009).
84. Douglas, J. T. Adenoviral vectors for gene therapy. *Mol. Biotechnol.* 36, 71–80 (2007).
85. Tong, L. Viral proteases. *Chem. Rev.* 102, 4609–4626 (2002).
86. Diouri, M., Keyvani-Amineh, H., Geoghegan, K. F. & Weber, J. M. Cleavage efficiency by adenovirus protease is site-dependent. *J. Biol. Chem.* 271, 32511–32514 (1996).
87. Chen, P. H., Ornelles, D. A. & Shenk, T. The adenovirus L3 23-kilodalton proteinase cleaves the amino-terminal head domain from cytokeratin 18 and disrupts the cytokeratin network of HeLa cells. *J. Virol.* 67, 3507–14 (1993).
88. Mangel, W. F., McGrath, W. J., Toledo, D. L. & Anderson, C. W. Viral DNA and a viral peptide can act as cofactors of adenovirus virion proteinase activity. *Nature* 361, 274–275 (1993).
89. McGrath, J. W. *et al.* Human adenovirus proteinase: DNA binding and stimulation of proteinase activity by DNA. *Biochemistry* 40, 13237–13245 (2001).
90. Brown, M. T. *et al.* Actin can act as a cofactor for a viral proteinase in the cleavage of the cytoskeleton. *J. Biol. Chem.* 277, 46298–46303 (2002).
91. Deu, E., Verdoes, M. & Bogyo, M. New approaches for dissecting protease functions to improve probe development and drug discovery. *Nat. Struct. Mol. Biol.* 19, 9–16 (2012).
92. Zhang, J. *et al.* Visualization of caspase-3-like activity in cells using a genetically encoded fluorescent biosensor activated by protein cleavage. *Nat. Commun.* 4, 1–13 (2013).
93. Sakamoto, S. *et al.* Creation of a caspase-3 sensing system using a combination of split-GFP and split-intein. *Chem. Commun.* 49, 10323 (2013).
94. Balakirev, M. Y., Jaquinod, M., Haas, A. L. & Chroboczek, J. Deubiquitinating function of adenovirus proteinase. *J. Virol.* 76, 6323–6331 (2002).
95. Sofia, A. & Oliveira, F. Application of functional genomics in the production of viral biopharmaceuticals: Vaccines and Vectors for gene therapy. (Faculdade de Ciências, Universidade de Lisboa, 2012).
96. Ruzindana-Umunyana, A., Sircar, S. & Weber, J. M. The effect of mutant peptide cofactors on adenovirus protease activity and virus infection. *Virology* 270, 173–9 (2000).
97. Dull, T. O. M. *et al.* A Third-Generation Lentivirus Vector with a Conditional Packaging System. 72, 8463–8471 (1998).
98. Cruz, P. E. *et al.* Screening of novel excipients for improving the stability of retroviral and adenoviral vectors. *Biotechnol. Prog.* 22, 568–576 (2006).
99. Darling, A. J., Boose, J. A. & Spaltro, J. Virus Assay Methods: Accuracy and Validation. *Biologicals* 26, 105–110 (1998).
100. Ruzindana-umunyana, A., Imbeault, L. & Weber, J. M. Substrate specificity of adenovirus protease. 89, 41–52 (2002).
101. Mangel, W. F. & Martín, C. S. Structure, function and dynamics in adenovirus maturation. *Viruses* 6, 4536–4570 (2014).
102. Doerfler, W. & Böhm, P. *The Molecular Repertoire of Adenoviruses I Virion Structure and Infection*. *Springer* I, (Springer Berlin Heidelberg, 1995).
103. Blanche, F. *et al.* Polypeptide composition of an adenovirus type 5 used in cancer gene therapy. 921, 39–48 (2001).
104. Pérez-berná, A. J. *et al.* Processing of the L1 52 / 55k Protein by the Adenovirus Protease : a New Substrate and New Insights into Virion Maturation. 88, 1513–1524 (2014).
105. Perler, F. B. *et al.* Protein splicing elements : inteins and exteins a definition of terms and

recommended nomenclature. *Nucleic Acids Res.* 22, 1125–1127 (1994).
106. Elleuche, S. & Pöggeler, S. Inteins , valuable genetic elements in molecular biology and biotechnology. *Appl. Microb. Biotechnol.* 479–489 (2010). doi:10.1007/s00253-010-2628-x
107. Chen, X., Zaro, J. L. & Shen, W. C. Fusion protein linkers: Property, design and functionality. *Advanced Drug Delivery Reviews* 65, 1357–1369 (2013).
108. Wehr, M. C. & Rossner, M. J. Split protein biosensor assays in molecular pharmacological studies. *Drug Discovery Today* 21, 415–429 (2016).
109. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157, 105–132 (1982).
110. Demonte, D., Li, N. & Park, S. Postsynthetic Domain Assembly with NpuDnaE and SspDnaB Split Inteins. *Appl. Biochem. Biotechnol.* 177, 1137–1151 (2015).

ANNEXES

Table S1 List of PCR primers, templates and restriction sites of plasmids constructed for cS11 strategy during this work.

Final Plasmid	Insert			Vector	
	Coding for:	Source	Primers		Restriction Site
pRRLSin_cS11-G/IVGLG/G	cG/IVGLG/G-S11	pUC57_cLRGAG-GFPS11	Fw: CGTCAGATCCGCTAGCCGCCACCATGATCAAGATCG (InFs-cVisensor-F) Rv: GCCCAGGCCCACGATGCCGTTGAAGCAGTTGCTGGCG (InFs-cGIVGLG-S11-R) and Fw: ATCGTGGGCCTGGGCGGCAGGGACCACATGGTGCTGC (InFs-cIVGLGG-S11-F) Rv: CGGGCTCGATGGATCTCTGCAGTCACAGGTCCTCTTCGGAGATCA (InFs-cVisensor-R)	pRRLSIN.CMV.GFPS10.IRES.Puro.WPRE	NheI + BamHI
pRRLSin_cS11-G/EEGEG/G	cG/EEGEG/G-S11	pUC57_cLRGAG-GFPS11	Fw: CGTCAGATCCGCTAGCCGCCACCATGATCAAGATCG (InFs-cVisensor-F) Rv: GCCCTCGCCTTCCTCGCCGTTGAAGCAGTTGCTGGCG (InFs-cGEEGEG-S11-R) and Fw: GAGGAAGGCGAGGGCGGCAGGGACCcdcACATGGTGCTGC (InFs-cEEGEGG-S11-F) Rv: CGGGCTCGATGGATCTCTGCAGTCACAGGTCCTCTTCGGAGATCA (InFs-cVisensor-R)	pRRLSIN.CMV.GFPS10.IRES.Puro.WPRE	NheI + BamHI
pRRLSin_cGFPS10-LRGAG	cGFPS10-LRGAG	pUC57_cGFPS10-LRGAG	Fw: CGTCAGATCCGCTAGCCGCCACCATGATCAAGATCG (InFs-cVisensor-F) Rv: CGGGCTCGATGGATCTCTGCAGTCACAGGTCCTCTTCGGAGATCA (InFs-cVisensor-R)	pRRLSIN.CMV.GFPS10.IRES.Zeo.WPRE	NheI + BamHI

Table S2 List of PCR primers, templates and restriction sites of plasmids constructed for viral proteases during this work.

Final Plasmid	Insert			Vector	
	Coding for:	Source	Primers		Restriction Site
pRRLSin_Adenain-MVGLG-VIc	Adenain-MVGLG-VIc	pUC57.Adenain-MVGLG-VIc	Fw: CGTCAGATCCGCTAGCATTCGCCACCATGGGCTC (InFs-Adenain-F) Rv: CGGGCTCGATGGATCCTTAGAAGCATCTCCGC (InFs-Adenain-VIc-R)	pRRLSIN.CMV.GFPS10.IRES.Zeo.WPRE	NheI + BamHI
pRRLSin_Adenain-MVGLG-VIc_mCherry	mCherry	pPuro_mCherry	Fw: ATATCGTCGAACCGGTATGGTGAGCAAGGGCGAG (InFs-mCherry-F) Rv: GAGGTTGATTGTCGACTTACTTGTACAGCTCGTCCATGC (InFs-mCherry-R)	pRRLSIN.CMV. Adenain-MVGLG-VIc.IRES.Zeo.WPRE	AgeI + SalI
pRRLSin_Adenain-C122G-VIc_mCherry	Adenain-C122G-VIc	pRRLSin_Adenain-MVGLG-VIc_mCherry	Fw: CGTCAGATCCGCTAGCATTCGCCACCATGGGCTC (InFs-Adenain-F) Rv: AGAATAGTCCACCGGCGGC (InFs-Adenain-mut-R) and Fw: GCCGCCGGTGGACTATTCT (InFs-Adenain-mut-F) Rv: CGGGCTCGATGGATCCTTAGAAGCATCTCCGC (InFs-Adenain-VIc-R)	pRRLSin_Adenain-MVGLG-VIc_mCherry	NheI + BamHI
pRRLSin_Adenain-C104G/C122G_mCherry	Adenain-C104G/C122G	pRRLSin_Adenain-C122G-VIc_mCherry	Fw: CGTCAGATCCGCTAGCATTCGCCACCATGGGCTC (InFs-Adenain-F) Rv: GCGTTATACCTCGGTCGGG (InFs-Adenain-mut2-R) and Fw: CCCGACCGAGGTATAACGC (InFs-Adenain-mut2-F) Rv: CGGGCTCGATGGATCCTTACATGTTTTTCAAGTGACAAAAAGAAG (InFs-Adenain-R)	pRRLSin_Adenain-MVGLG-VIc_mCherry	NheI + BamHI