UNIVERSIDADE DO ALGARVE

RECOMBINANT PROTEINS IN DIFFERENTIATION OF STEM CELLS

Ana Rita Fragoso Nascimento

Dissertação de Mestrado
Mestrado em Ciências Biomédicas

Trabalho efetuado sob a orientação de
Professor Doutor José Bragança
e co-orientação de
Professor Doutor Guilherme Ferreira

2015

UNIVERSIDADE DO ALGARVE

RECOMBINANT PROTEINS IN DIFFERENTIATION OF STEM CELLS

Ana Rita Fragoso Nascimento

Dissertação de Mestrado
Mestrado em Ciências Biomédicas

Trabalho efetuado sob a orientação de
Professor Doutor José Bragança
e co-orientação de
Professor Doutor Guilherme Ferreira

2015

Recombinant Proteins in Differentiation of Stem Cells

Declaração de autoria de trabalho:

Declaro ser a autora deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

_____

Ana Rita Fragoso Nascimento

ACKNOWLEDGMENTS

I am using this opportunity to express my gratitude to everyone who supported me throughout the course of this Master Thesis.

I would like to express my special thanks to my supervisor Professor José Bragança, who has given me the opportunity to work in his laboratory. I would like to thank you for encouraging my research and for allowing me to grow as a research scientist. Your advice and support have been priceless.

To Dr. Guilherme Ferreira, I appreciate the support given during the orientation of this dissertation. The opportunity and knowledge acquired express my gratitude.

I would also like to thank to Professor Rui Martinho, Professor Eduardo Melo and Carlos Lopes, who were essential at some developmental stages of this work.

To my laboratory colleagues and friends from 1.12 and from 3.30, Gisela Oliveira, Ana Jesus, João Charneca, Vanessa Afonso, Eduarda Guerreiro and Carina Silva, for all the support given as well as the good and fun moments through this year. I would like to specially thank to João Santos for all his help and patience, his invaluably constructive criticism and friendly advice during the project work. I'm also very grateful to Catarina Águas, for all her help and incentive in the development of this thesis.

To all my friends for being present and giving me support through all this year, including the entertaining moments which allowed me to put aside the laboratory concerns and overcome them, but also for the patience in certain times along this dissertation year.

Last but not the least, I would like to express my gratitude to my family for believing in me and giving me the opportunity to get my MsC degree. A special word of thanks for the love, encouragement, and tolerance of Rúben, who has made all the difference in my life. Without his patience and sacrifice, I could not have completed this thesis.

ABSTRACT

Cardiomyocytes derived from embryonic stem cells (ESCs) offer a great alternative to generate a large number of cells with potential application in biomedical research and drug or toxicology screening. Although the differentiation efficiency can be improved by the genetic manipulation of ESCs to over-express cardiac-specific transcription factors, current protocols for differentiation are time consuming, have low yield, and lack of reproducibility. Protein transduction has been demonstrated as an alternative approach for increasing the efficiency of ESCs differentiation toward cardiomyocytes.

The objective of this work was to produce a recombinant chimeric protein with the intrinsic ability to transduce into cells. In this work, we successfully produced and purified a novel recombinant protein from bacteria, which is capable of overcome the cellular membrane barrier. This protein is composed by two domains: the human protein CITED2 fused at its N-terminal domain to a Protein Transduction Domain (PTD), rich in arginines, which confers biologically active proteins the capability to translocate across the membrane and deliver them inside the cell.

We showed that CITED2 recombinant protein, added to the culture medium of the cells, was successfully internalized, localized in the nucleus and functional. In mouse embryonic stem cells (mESC), the knockout of *Cited2* gene impairs these cells to differentiate into cardiomyocytes. We showed that the supplementation, on the second day of differentiation, of CITED2 recombinant protein to the culture medium of *Cited2* depleted cells, rescued the cardiogenic defects of these cells. On the other hand, the supplementation at the onset of differentiation, while the cells are still in a pluripotent state, suggested a delay in the overall differentiation process.

Although further studies are still required, the direct application of the developed recombinant protein suggests a capability to replace the endogenous protein effects, further suggesting that CITED2 protein has a relevant role in cardiogenesis and pluripotency.

Keywords: Cardiac differentiation, Cited2, Mouse Embryonic Stem Cells, Protein Transduction Domain, Recombinant protein.

RESUMO

A utilização de cardiomiócitos derivados de células estaminais embrionárias (ESC) permite-nos acesso a um grande número de células para investigação biomédica e uso de fármacos ou ensaios toxicológicos. Apesar de ser possível aumentar a eficiência de diferenciação através de manipulação genética sobre expressando fatores cardíacos específicos, os protocolos atuais são bastante morosos, têm baixo rendimento e são dificilmente reprodutíveis. A transdução proteica é um método alternativo para aumentar a eficácia de diferenciação de células estaminais embrionárias em cardiomiócitos.

O objetivo consistiu na construção de uma proteína quimérica com a capacidade intrínseca de transdução celular. Produzimos e purificámos esta proteína recombinante, com a capacidade de atravessar a membrana celular, sendo composta por dois domínios: a proteína humana CITED2, ligada através do domínio terminal-N a um Domínio de Transdução Proteica (PTD), rico em argininas, que confere a capacidade de transduzir proteínas biologicamente ativas e entregá-las no interior da célula.

Com este trabalho, conseguimos demonstrar que esta proteína recombinante CITED2, quando adicionada ao meio de cultura celular, atravessa a membrana celular, alcança o núcleo, e é funcional.

Em células embrionárias de ratinho (mESC), o *knockout* do gene *Cited2* leva à diferenciação destas células em cardiomiócitos. A suplementação com a proteína recombinante CITED2, no segundo dia de diferenciação, no meio de cultura em que as células tinham sofrido a depleção de *Cited2*, tem a capacidade de compensar os defeitos na diferenciação cardíaca. Por outro lado, se o meio de cultura for suplementado na fase inicial de diferenciação (dia 0), enquanto as células ainda se encontram num estado de pluripotência, sugere um atraso no processo de diferenciação geral.

Apesar de serem necessários estudos mais aprofundados, a aplicação direta da proteína desenvolvida parece ter a capacidade de substituição dos efeitos causados pela proteína endógena, sugerindo assim que a proteína CITED2 desempenha um papel importante tanto em cardiogénese como em pluripotência.

Palavras-chave: Diferenciação cardíaca, Cited2, Células estaminais embrionárias de ratinho, Domínio de Transdução Proteica, Proteína recombinante.

RESUMO ALARGADO

As doenças cardiovasculares são a principal causa de morte nos países desenvolvidos, constituindo um problema de saúde pública que urge minorar. No entanto, o coração tem uma capacidade limitada de regeneração e quando uma lesão ocorre, este não consegue substituir o tecido cardíaco danificado por novo músculo cardíaco, pois os cardiomiócitos (células altamente diferenciadas do coração com propriedades contráteis) morrem quer por apoptose ou necrose do tecido. Assim, essas células são substituídas por fibroblastos que contribuem para a formação de tecido cicatricial, tornando o coração mais frágil, com as suas propriedades contráteis comprometidas.

Um dos principais objetivos da investigação biomédica tem sido o desenvolvimento de métodos para a geração de cardiomiócitos *de novo*, tanto *in vitro* como *in vivo*, tal como novas ferramentas terapêuticas e métodos para regenerar o músculo cardíaco. Tendo em conta a informação anterior, um modo de tentar investigar os mecanismos moleculares que contribuem para o desenvolvimento de doenças do miocárdio, é essencialmente gerar cardiomiócitos funcionais, uma vez que a sua utilização na descoberta de medicamentos e estudos de toxicidade seria altamente benéfico, pois iria permitir testar novas moléculas farmacológicas para o tratamento deste tipo de doença. A solução ideal seria obter novas células, com propriedades contráteis de modo a reparar o tecido danificado.

As células estaminais são consideradas uma das mais promissoras fontes de células para regeneração, sendo que vários tecidos humanos já foram propostos como uma fonte de células estaminais com potencial cardiogénico (e portanto, capaz de gerar novos cardiomiócitos). O transplante autólogo de células estaminais dos próprios pacientes constitui uma das melhores hipóteses propostas para a reparação eficiente de corações fragilizados por doenças cardíacas, principalmente devido ao potencial de populações de células estaminais (como as células progenitoras cardíacas (CPCs)) que podem contribuir para promover a regeneração cardíaca.

Relativamente à diferenciação celular em cardiomiócitos, para gerar células seguras para aplicações terapêuticas é essencial evitar integração viral, e portanto é fundamental reduzir e evitar a manipulação genética, tornando a produção e aplicação de proteínas recombinantes uma alternativa mais viável para a expressão do gene em estudo.

Cited2, um gene que codifica uma proteína nuclear, liga-se diretamente e com elevada afinidade, à primeira região enriquecida com cisteína e histidinas (CH1) do p300 e CBP, estando presente em todos vertebrados. Vários estudos demonstraram que as variações/mutações no gene *Cited2* está relacionada com defeitos cardíacos congénitos.

Foi observado que *Cited2* é essencial na manutenção da pluripotência uma vez que regula diretamente a expressão de fatores de transcrição tal como *Oct4*, *Sox2*, *Nanog*, *Klf4* e *Tbx3*, que fazem parte da rede nuclear dos genes responsáveis pelo estado de pluripotência nas ESC. Uma variação dos níveis de expressão destes reguladores de pluripotência dos níveis de expressão ótimos, levam à diferenciação das ESC. Vários estudos suportam esta hipótese, uma vez que a sobrexpressão de CITED2 em mESC suporta a autorrenovação destas células, mesmo na ausência de *leukemia inhibitory factor* (LIF). Pelo contrário, a ausência de CITED2 durante a diferenciação de mESC leva a um atraso em genes de autorrenovação, ativando vias de diferenciação hematopoiéticas, neuronais e cardíacas.

A entrega intracelular de macromoléculas biológicas (como por exemplo fármacos) constituem um campo importante da medicina, uma vez que muitas macromoléculas biológicas, como as proteínas, péptidos, e ácidos nucleicos têm demonstrado ser úteis para o tratamento de vários problemas de saúde. No entanto, a entrega de macromoléculas nos seus locais-alvo não é fácil, principalmente devido à estrutura da membrana celular eucariótica. Um dos sistemas que podem ser utilizados para superar este problema, é a aplicação de domínios de transdução proteicos (PTDs), que são capazes de transportar a moléculas associadas através da membrana e desta forma permitir a entrada das proteínas biologicamente ativas no interior da célula.

Assim sendo, o objetivo desta tese foi a produção de uma proteína recombinante com a capacidade de transdução celular. Neste trabalho, clonámos o gene *Cited2* num plasmídeo, de modo a produzir uma nova proteína recombinante. Esta proteína consiste na fusão de CITED2 com um péptido rico em argininas, denominado PTD, através do domínio terminal-N. Este PTD tem a capacidade de atravessar a membrana celular e desta forma, seria capaz de transportar a proteína CITED2 para o interior da célula.

Assim, este projeto de investigação pode ser dividido em três etapas gerais:

- Clonagem molecular dos fragmentos de DNA no plasmídeo pGEX-6P-1;
- Produção e purificação das proteínas recombinantes;

- Realização de ensaios celulares para análise da entrada e efeitos na pluripotência e diferenciação cardíaca em células estaminais embrionárias de ratinho.

Tendo em conta o objetivo proposto, começámos por clonar os fragmentos de DNA desejados num plasmídeo que permitiria a posterior produção destas proteínas. Após a produção das proteínas recombinantes TAT-CITED2 e 8R-CITED2, estas foram purificadas através de FPLC: TAT-CITED2 e 8R-CITED2. No entanto, a validação funcional a nível celular foi testada apenas com a proteína 8R-CITED2 por falta de tempo.

Quando suplementada no meio de cultura, a proteína recombinante 8R-CITED2 demonstrou ter sido internalizada nas células. A atividade da mesma foi comprovada através de um ensaio de BiFC, onde se demonstrou a competitividade contra a proteína CITED2 endógena pelo domínio p300. Por último, em ensaios de diferenciação celular, a proteína recombinante apresentou um padrão semelhante à proteína endógena quando esta última é suprimida. Ao estudar o papel desta proteína recombinante, podemos concluir que esta tem a capacidade de restauração dos defeitos causado pelo *knockout* de *Cited2* endógeno em mESC quando aplicada ao meio de cultura no segundo dia de diferenciação, e parece ter alguma influência no atraso da diferenciação cardíaca quando o meio é suplementado no dia 0.

Usando esta abordagem, esta proteína pode ser utilizada para expandir células estaminais embrionárias de ratinho num ambiente controlado e auxiliar no controlo da diferenciação destas células. Este método oferece melhores alternativas do que os métodos utilizados anteriormente, visto que envolvem transfeções que geralmente resultam em sistemas artificiais e níveis subótimos de CITED2. No entanto, pesquisas e ensaios adicionais são necessários para aumentar a eficiência e segurança desta proteína recombinante.

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

α-MHC	α-Myosin Heavy Chain
Amp	Ampicillin
16R	Arginine hexadecamer
4HT	4-hydroxytamoxifen
4R	Arginine tetramer
8R	Arginine octamer
AR-CPPs	Arginine-rich Cell Penetrating Peptides
BIFC	Bifluorescence Complementation Method
bp	Base pairs
BSA	Bovine Serum Albumine
CBP	cAMP-responsive element-biding protein
CH1	Cysteine-histidine rich
CIP	Calf Intestine Phosphatase
Cited2	CBP/p300-Interating transactivators with ED rich tail 2
CPCs	Cardiac progenitor cells
CPPs	Cell-penetrating peptides
CSC	Cancer Stem Cells
cTnT	Cardiac isoform of Troponin-T
D	Aspartic acid
DAPI	4’,6-diamidino-2-phenylindole
DNA	Desoxirribonucleic acid
DTT	Dithiothreitol
E	Glutamic acid
E. coli	*Escherichia coli*
EB	Embryoid Body
ECM	Extracellular matrix
EDTA	Ethylenediamine tetraacetic acid
ESC	Embryonic stem cells
FBS	Fetal Bovine Serum
FHF	First Heart Field
FT	Flow-through

GMEM	Glasgow Minimum Essential Medium
GSH	Reduced Glutathione
GST	Glutathione S-Transferase
hESCs	Human Embryonic Stem Cells
HIF-1	Hypoxia-inducible factor
HRP	Horseradish Peroxidase
InsF	Insoluble Fraction
iPSCs	Induced Pluripotent Stem Cells
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Isl1	Islet1
KO	Knockout
LB	Lysogeny broth
LIF	Leukaemia Inhibitory Factor
MEFs	Mouse Embryonic Fibroblasts
mESCs	Mouse Embryonic Stem Cells
NSCLC	Non-Small Cell Lung Cancer
O.D.	Optical density
O.N.	Overnight
PBS	Phosphate Buffered Saline
pDNA	Plasmid DNA
PMSF	Phenylmethylsulfonyl fluoride
PTDs	Protein Transduction Domains
qRT-PCR	quantitative Reverse Transcriptase-Polymerase Chain Reaction
R	Arginine
RF	Reprogramming factors
RNA	Ribonucleic acid
RT	Room temperature
SDS-PAGE	Sodium Sodecyl Sulphate-Polyacrylamide Gel Electrophoresis
SF	Fraction/injected Sample
SHF	Second Heart Field
TAE	Tris-acetate
TAT-RFs	Reprogramming factors fused with TAT
YFP	Yellow Fluorescent Protein

I. INTRODUCTION

I.1. CARDIOVASCULAR DISEASE

Heart disease is the leading cause of mortality worldwide, resulting in 17.3 million deaths (31.5%) in 2013 up from 12.3 million (25.8%) in 1990, with insufficient therapeutic options and poor prognosis.[1,2] Cardiovascular disease affects older adults, amongst which coronary artery disease and stroke account for 80% of deaths in males and 75% of deaths in females with cardiovascular disease.[3]

Despite significant advances in therapeutic modalities and prevention strategies, the high mortality rates associated with heart disease is fearsome. This has driven research into new therapeutic strategies including cardiac regenerative therapy as a new approach for severe cardiac diseases resistant to the conventional treatments.[1]

To investigate the molecular mechanisms leading to myocardial diseases in humans, it is essential to generate functional cardiomyocytes. Their use in drug discovery and toxicology studies would be highly beneficial, allowing that new pharmacological molecules for the treatment of cardiac disorders to be validated pre-clinically on cells of human origin.[4]

Recent studies from several laboratories have demonstrated that cardiomyocyte turnover occurs throughout life in mammals. However, most of the reports find a remarkably low annual post-natal cardiomyocyte renewal rate of approximately 1%, which increases modestly after injury but declines with age.[2] The regenerative potential of the adult mammalian heart is very limited and when a lesion occurs it cannot replace the damaged cardiac tissue with new function muscle, since the cardiomyocytes die either by apoptosis or necrosis of the tissue. Then these cells are replaced by fibroblasts for scar tissue formation, turning the heart fragile, without good contractile properties.[2,5,6]

An optimal solution is to obtain new cells, with contractile properties to replace the injured tissue. However the access to the human heart tissue is very limited, and although human cardiomyocytes could be isolated from heart through biopsies, the procedure is complicated, invasive, unpractical to obtain viable cell preparations in large quantities (since they lack of proliferation capacity), and the majority of the cells acquired do not beat spontaneously.[6,7]

To overcome this problem, biomedical research fields have been developing methods for the generation of *de novo* cardiomyocytes, both *in vitro* and *in vivo*, together with novel therapeutic tools and approaches to regenerate cardiac muscle for diseased hearts.[2,5,6]

The development of new approaches has resulted in the isolation of cardiomyocytes from various new-born animals or the production of genetically engineered cell lines. Nevertheless, these models have significant limitations since they present basic physiological differences in comparison to human cardiomyocytes as well as high costs and ethical questions.[8,9]

Regarding cardiac cell-based treatments, stem cells are a promising cell source, which are being prioritized by scientists for basic research and clinical trials.[10,11] Several different human tissues have already been proposed as a source of stem cells with cardiogenic potential (therefore capable to generate new cardiomyocytes) (e.g., fetal cardiomyocytes, adult cardiac progenitor cells, skeletal myoblasts, bone marrow-derived stem cells, adipose-derived stem cells, umbilical cord-derived stem cells, and pluripotent stem cells), and some methods to isolate and expand these cells have been developed aiming cardiac regenerative therapy.[7,10–12] In this field, stem cell-derived cardiomyocytes would facilitate the discovery of small molecules promoting cardiomyocyte differentiation that could be used for the activation of endogenous cardiac stem cells in clinical settings.[12]

Stem cell autologous transplantation in patients is one of the candidates approach for efficient repair of impaired heart, due to the potential of stem cell populations (like cardiac progenitor cells (CPCs)) to promote cardiac regeneration and repair in experimental models and in patients with heart disease.[2] Even so, there are complex and challenging pathobiological apprehensions. To make this possible, cell enhancement strategies need to be improved.[13] For instance, patient derived cells undergoing a pre-treatment with small molecules or genetic modification which may contribute to an augmented recruitment of important factors which could enhance differentiation or other beneficial functions.[11]

I.2. STEM CELLS

Stem cells are often described by their capabilities of self-renewal, to be progenitors of tissue or organ specific cells and their ability to functionally reconstitute a given tissue *in vivo*.[14]

These cells are present throughout the different developmental stages of an organism into and during the adult life and can be divided into two groups: embryonic and adult stem cells. Embryonic stem cells (ESC) are found only in the embryo and are classified as pluripotent, as they can give rise to any cell type of the three germ layers. On the other hand, adult stem cells are found in adult organisms, within fully differentiated tissues, and are classified as multipotent because they can only self-renew and differentiate into the different cells of the specific tissue where they reside or from which they were isolated.

Pluripotent ESC are derived from the inner cell mass of the blastocyst. Unlike adult stem cells, ESC have a very high degree of self-renewal capability, at least in part due to the fact that they express high levels of telomerase.[14] This property makes them hold great potential as cell-based therapies to promote vascularization and tissue regeneration.[15]

At day 4 and day 7 to 10 of mouse and human development, respectively, cells of the blastocyst become lineage restricted, which means that they become committed to a germline and subsequently to a specialization within certain tissues. So, when working *in vitro* with ESC cells, we must to have in mind the need to maintain their pluripotency (preventing differentiation) and their self-renewal capacity using specific culture conditions.[14,16]

In the field of cardiac regenerative therapy, it was shown that skeletal myoblasts extracted from patients themselves, have the potential to improve the function of a failing heart when they are transplanted into the heart as myoblast sheets[17]. However, these cells are not able to differentiate into cardiomyocytes, and therefore cannot replace the defective cardiomyocytes in a failing heart. A more likely efficient cell source in cardiac regenerative therapy could be if the transplantable cardiomyocytes could be prepared from ESCs.[17,18]

I.2.1. DIFFERENTIATION

Several strategies have been established to regulate the differentiation of ESC into all the three germ layers: the mesoderm, endoderm, and ectoderm.

There are several methods that can be used to differentiate ESC towards a specific fate (**Figure I.1**). One of these methods is genetic manipulation, which force the expression of some transcription factors and can lead to direct differentiation of ES cells toward specific lineages.

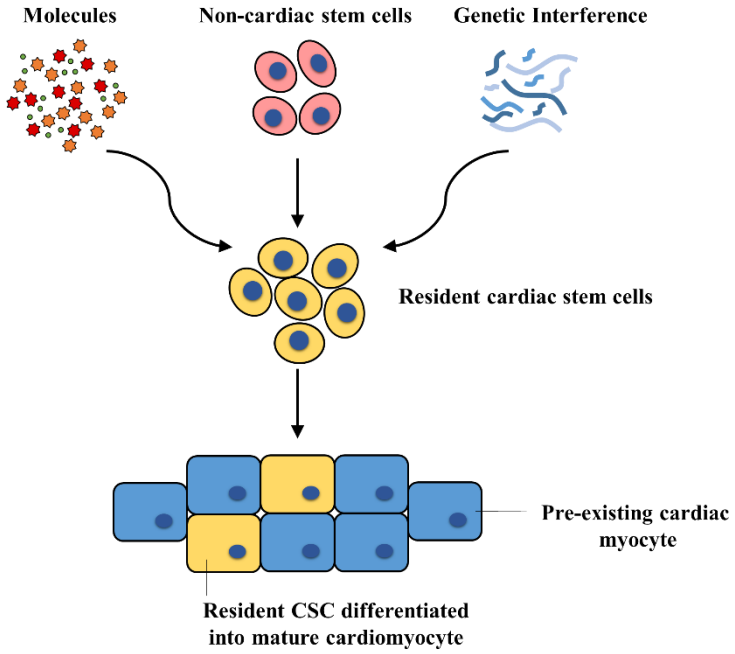


Figure I.1 – Schematic overview of activation of endogenous multipotent cardiac stem cells by various means for myocardial repair. (A) Molecules (e.g. growth factors, cytokines), (B) noncardiac stem cells (e.g. bone marrow), (C) or gene therapy (e.g. micro-RNAs, gene transfer). Upon activation, resident endogenous cardiac stem cells can proliferate and mature into newly formed cardiac myocytes (yellow cardiac myocytes).

For example, the overexpression of HOXB4, a gene involved in hematopoietic lineage, significantly enhances the hematopoietic potential of mouse ESC (mESC) differentiation *in vitro*. Expression of GATA-6 and GATA-4 genes, involved in myocardial differentiation and function, induces mESC differentiation into the extra-embryonic endoderm, while GATA-4 overexpression alone enhances cardiogenesis and markedly increases the number of terminally differentiated beating cardiomyocytes.[19]

Another way to genetically control the differentiation of these cells is the use of micro-RNAs, since a single miRNA can target multiple pathways simultaneously, being the effects on gene expression very powerful.[20–22] Recently, a combination of microRNAs (miRNAs: 1, 133, 208 and 499) has been identified as capable to reprogram murine fibroblasts to cardiomyocyte-like cells even though the efficiency rate being very low (about 1.5-7.7%).[22] The most recent study combined a subset of four transcription factors (Gata4, Hand1, Tbx5 and myocardin) with two miRNAs (1, 133) to successfully reprogram human fibroblasts into cardiomyocyte-like cells. [23]

The control of cell proliferation by growth factors, being often added to a medium to promote differentiation, is another approach that could affect survival of specific cell types. For instance, this direct differentiation can be implemented to overcome the heterogeneous cell mixture derived from differentiation cultures, and could be achieved by controlling the nutrient intake, the extracellular matrix (ECM), co-culturing with inductive cell types or adding a signaling molecule that have an impact on gene expression and cell proliferation.[23]

However, the formation of ESC aggregates is the most widely used, consisting in the formation of suspended spherical aggregates called Embryoid Bodies (EBs).[18] This structure facilitates multicellular interactions, in which cell-cell contact exists and gap junctions may be established. EB formation is stimulated in the absence of leukemia induced factor (LIF) from the culture media or mouse embryonic fibroblasts (MEFs) feeder layer, changing from two-dimensional monolayer cell cultures to three-dimensional cell based structures in suspension.[19,24] In the case of mESC, spherical ESC aggregates with morula-like structures formed in 2–6 days in suspension culture, and the removal of molecules which promote self-renewal from the culture medium, combined with cellular aggregation is an efficient way to differentiate them.[19,25]

I.2.2. CARDIAC DIFFERENTIATION

The ability of ESC to differentiate into spontaneously contracting cardiomyocyte-like cells has attracted substantial interest from the scientific community. Initially it was a process considered difficult to control, but now cardiomyogenesis *in vitro* is a process which, to a certain extent, could be effectively manipulated and directed in the future.[12]

Heart development involves a series of highly complex morphogenetic processes that are chronologically regulated by multiple phase-specific signals. The heart is the first organ to form in the embryo where its early function is essential to the circulation of nutrients and removal of waste. The construction of the heart is a complex process and involves the integration of different cell populations at distinct site as development proceeds.

Using chicken and mice embryos as models it has been demonstrated that the heart tissue is composed of three major mesoderm-derived cell lineages: the cardiac myocyte, the vascular smooth muscle, and the endothelial cell lineages. Few days after fertilization, the

three embryonic layers form, the endoderm, the ectoderm, and the mesoderm. The primitive streak is formed from primitive endoderm and is the origin of many tissues, being cardiac progenitors formed in the posterior primitive streak.

The cardiac mesoderm gives rise to the endocardium, the first heart field (FHF, which forms the atria, left ventricle, and the nodal conduction system), the secondary heart field (SHF, which forms the right ventricle, outflow tract, and part of the atria), and the proepicardial mesenchyme.[26]

The generation of cardiomyocytes from pluripotent stem cells can be divided into four phases: formation of mesoderm, driving of mesoderm toward anterior mesoderm or cardiogenic mesoderm, formation of cardiac mesoderm, and finally the maturation of early cardiomyocytes. All of these steps are conducted by the expression of transcription factors, primarly T/Brachyury is expressed for primitive streak mesoderm, followed by Mesp-1 for cardiogenic mesoderm, and Nkx2.5, Tbx5/20, Gata-4, Mef2c, and Hand1/2 for cardiac mesoderm.[12,27]

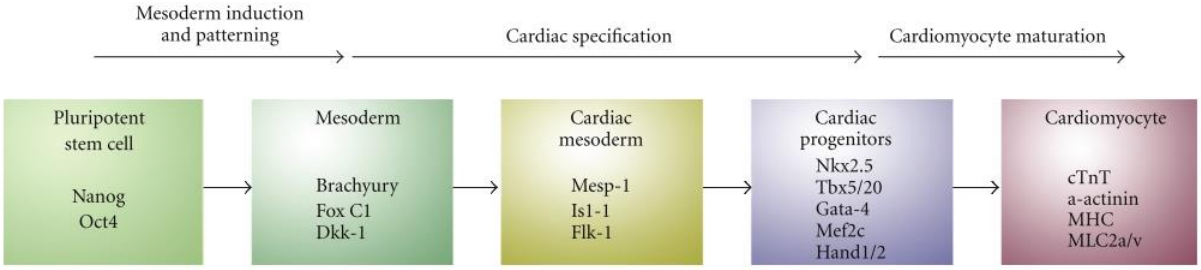


Figure I.2 – Diagram illustrating sequential steps in differentiation of pluripotent stem cells to cardiomyocytes. Pluripotent stem cells differentiate into early mesoderm cells, which further differentiate to cardiac mesoderm. Then, cells become committed to cardiac progenitors and differentiate to functional beating cardiomyocytes. The typical markers for each step are indicated.[12]

Further identification of maturing cardiomyocytes could be done by the expression of cardiac structural proteins such as α-actinin, α-myosin heavy chain (α-MHC), and/or the cardiac isoform of Troponin-T (cTnT). **Figure I.2** shows the sequential steps in the differentiation of pluripotent stem cells to cardiomyocytes.[9,12]

There are several possible approaches for generating and expanding cardiomyocytes from major sources of starting cells like: induced pluripotent stem cells (iPSCs), adult heart-derived cardiac progenitor cells (CPCs), and reprogrammed fibroblasts (**Figure I.3**).[5,12,28]

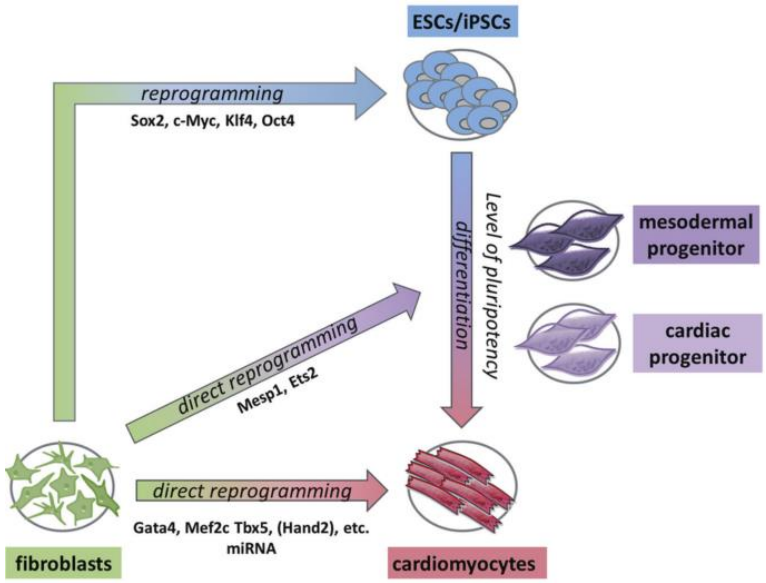


Figure I.3 – Schematic representation of different stem cell sources for cardiac regeneration. ESCs/iPSCs can be generated by reprogramming fibroblasts and then differentiated into cardiac lineages. Another of the approaches rely on the direct conversion of mature somatic cells (like fibroblasts) into other mature cell types like cardiomyocytes or to an intermediate stage between full pluripotency and total maturity: a cardiac progenitor stage.[21]

Besides their potential in clinical applications, these cells can offer more physiologically and clinically relevant reproducible human cell models than the ones presently available. Since they can offer the cardiac phenotype and the functional proper

ties of the pluripotent stem cell-derived cardiomyocytes, they can be used as models to study early events of human cardiogenesis and have the potential to be used in pharmaceutical drug discovery and safety toxicology.[12]

A genetic manipulation approach, like reprogramming somatic cells into cardiac lineage cells, bypassing the pluripotent state, is another possibility. Human embryonic stem cells (hESCs), and the recent discovery of iPSCs generation, have attracted the curiosity of many investigators regarding the potential of these cells, trying to develop strategies to efficiently and reliably direct stem cell differentiation into the cardiovascular lineage. [5]

Differentiation of cardiomyocytes from hESCs has progressed rapidly through a growth factor-mediated approach of Srivastava group, who showed successful direct conversion of fibroblasts into cardiomyocyte-like cells *in vitro* and *in vivo* by a specific combination of cardiac transcriptional factors (Gata4, Mef2c, and Tbx5). [2,25]

Also, Itskovitz-Eldor *et al*. (2000)[29] have demonstrated that contracting cardiomyocytes can be generated from of hPSCs opening the possibility of producing an unlimited number of human cardiomyocytes to rebuild the heart. These advances in embryology and hPSC differentiation have offered significant insights into the mechanisms of cardiopoiesis, building a promising future regarding the repair of injured hearts through clinical applications of these cells. [5,29]

Although the efficiency of differentiation protocols has increased over time, there are some disadvantages using this type of method. One of them relies on the optimization of the methods used, to obtain a better efficiency. Another concern is the fact that a possible risk of viral transduction-mediated tumorigenesis could occur, making this method not the safest one.[2,25]

Another alternative source of *de novo* cardiomyocytes is the direct reprogramming of cardiac fibroblasts and other adult cell types into cardiomyocytes using cardiac-specific transcription factors: Gata4, Mef2c, and Tbx5.[30] Using this approach, reprogramming cardiac fibroblasts *in vivo* for heart regeneration becomes a possibility, with direct application of these transcription factors.[5]

Ultimately, the derivation of cardiomyocytes without viral integration is essential for the generation of safe cells for therapeutic applications, and to avoid genetic manipulation. Therefore the production and application of recombinant proteins constitute an alternative for the over-expression of a desired gene.

I.3. CITED2

CITED2 is a CREB-binding protein (CBP)/p300-interacting transactivator with glutamic acid (E) and aspartic acid (D) enriched tail, previously named melanocyte-specific gene-related gene (MRG)1/p35srj. This gene locates within the 6q23 region in humans and it encodes a nuclear protein which is ubiquitously expressed, and binds directly, and with high affinity, to the first cysteine–histidine-rich (CH1) region of p300 and CBP, being present in all vertebrates.[31–33]

As a CBP/p300-dependent transcription factor, CITED2 regulates gene transcription by interacting with other transcription factors acting as a co-activator or a co-repressor.

Several studies have shown that variations/mutations or abnormal methylation of the CITED2 gene are related with congenital heart defects.[34,35]

In mouse, CITED2 is essential for normal development, since disruption of the gene is embryonic lethal. When CITED2 is not expressed in mice, the embryos die in utero with cardiac and aortic arch malformations, adrenal gland agenesis, small cranial and dorsal root ganglia, exencephaly, neural crest and left-right patterning defects.[32,36] The cardiac malformations are various, such as atrial and ventricular septal defects, double outlet right ventricle, common arterial trunk, transposition of the great arteries, interrupted and aberrant aortic arches and others. [35,37–39]

CITED2 is essential at two different times of mESC fate, first maintaining ESC in pluripotency state via direct regulation of a number of activated signaling pathways that control the expression of the transcription factors *Oct4, Sox2* and *Nanog*, and then leading to differentiation towards cardiomyocytes. [35,40]

Interestingly, a shift of the expression levels of these core pluripotency master regulators from their optimal expression prompt ESC to differentiate. Therefore, the maintenance of pluripotency depends on the stringent control of their expression by forward and feedback regulation loops.[35] Studies supported this hypothesis, since the overexpression of CITED2 in mESC sustains self-renewal of these cells, even in the absence of LIF.[36,41] One of the possible explanations is the fact that CITED2 is a direct target of FoxP1 which mediates pluripotency in mESC, and could be involved in the maintenance of fetal and adult hematopoietic stem cells. In contrast, the absence of CITED2 during mESC differentiation leads to a delay in self-renewal genes activating hematopoietic, neuronal and cardiac differentiation pathways. [40,42]

Recent work developed in our group showed that CITED2-depletion at the onset of differentiation significantly impairs the generation of cardiomyocytes and decreases the expression of early mesoderm markers (*Brachyury, Mesp1*), pro-cardiogenic transcription factors (*Isl1, Gata4, Tbx5*) and secreted molecules (*Wnt5a* and *Fgf10*). In opposition, CITED2 overexpression is sufficient to stimulate the expression of these genes in undifferentiated ESC and to promote cardiac differentiation. Finally it was shown that CITED2 expression is highly associated with Cardiac Progenitor Cells (CPC) populations, particularly cardiac progenitors of the SHF marked by Isl1 expression (Manuscript in preparation: Cited2 synergizes with Isl1

and promotes cardiac differentiation of mouse embryonic stem cells; Ivette Pacheco-Leyva, A. Matias, D. Oliveira, Rita Nascimento, V. Afonso, JMA Santos, E. Guerreiro, Anna C. Michell, Annebel M. van De Vrugt, G. Oliveira, K. Kranc, G. Ferreira, I. and J. Bragança).

I.4. RECOMBINANT PROTEINS

Intracellular delivery of biological macromolecules is a major topic in the field of drug delivery, since many biological macromolecules, like proteins, peptides, and nucleic acids, have proven to be useful for the treatment of various health problems.[43]

However, the delivery of macromolecules to their target sites is not easy, mostly due to the structure of the eukaryotic cell membrane structure. The plasma membrane consists of a lipid bilayer in which proteins and glycoproteins are inserted, and their hydrophobic nature makes the delivery of proteins less efficient, since it prevents the hydrophilic compounds to cross the membrane. [25,44]

One of the systems that can be used to overcome the problem of the entry into the cells, is the application of protein transduction domains (PTDs), also known as cell-penetrating peptides (CPPs). These kind of peptides are capable of transporting cargo across the membrane and delivering biologically active proteins inside the cell (**Figure I.4**). Besides, they can be linked to another cargos such as peptides, proteins, oligonucleotides, pDNA, or liposomes.[25,43,45]

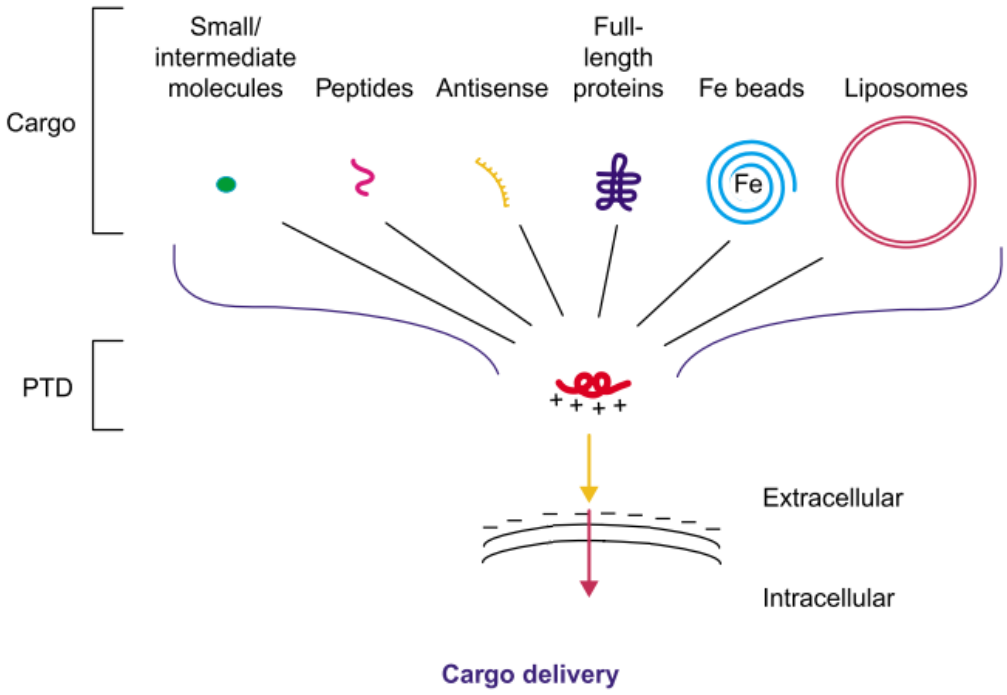


Figure I.4 – Schematic representation of cargo variety covalently linked to arginine-dependent protein transduction domains (PTDs). Adapted from [46].

The initial discovery of PTDs was originated from the independent observation by Green and Frankel in 1988, when a special aminoacid sequence derived from HIV-1 TAT protein, the TAT PTD, showed that could penetrate cells in a receptor-independent, concentration-dependent and activate HIV-1-specific target genes. That peptide could translocate across the plasma membrane by its 11 basic aminoacids (residues 47–57) and to deliver heterogeneous proteins into cells. [25,47–50] This PTD has a higher efficiency for protein delivery into the cells when compared to other PTD signals.[25]

Once added to the culture media, TAT-mediated transduction occurs through a rapid, temperature and energy-independent process, suggesting direct penetration across the lipid bilayer because of the strong binding of the PTD to the cell surface. [50]

One of the hypothesis for the capacity of translocation of these kind of molecules is that their positive charges allow the protein to interact with lipid rafts in a membrane which is negatively charged. That way it can overcome the cell membrane barrier by different mechanisms, like macropinocytosis.[25,47,49] Endocytosis was later suggested as an alternative internalization pathway. It is now thought that endocytosis and direct translocation are two coexisting pathways.[51] In all cases, after endocytic uptake, the internalized CPPs (either alone

or linked to cargos) should escape from the endocytic vesicles to the cytosol to avoid degradation.[43] Within cells, the TAT-fusion proteins are either degraded or refolded by the cellular machinery into functional proteins.[49]

One of the characteristics of the TAT PTD sequence which seems to be responsible for the transduction of these proteins is the enrichment by arginine (R) residues and that the sequences with 6-12 consecutive R residues are functional PTDs.[47] These kind of arginine-rich cell-penetrating peptides (AR-CPPs) are the most widely studied.[43]

These oligoarginine peptides present differential manners of internalization and cellular localization depending on the number of arginine residues in the molecules. For example, a tetramer of arginines (4R) did not show significant internalization, while an octamer (8R) showed efficient internalization and nuclear localization, very similar to that of TAT. Conversely, a larger number of residues of arginine did not result in an increased internalization capability, as an hexadecamer of arginines (16R) showed less efficient internalization compared to TAT, without showing significant nuclear localization.[52]

The transduction of these PTD-based proteins showed to be very efficient working for virtually all types of cells tested. The localization of transduced proteins within the cells depends on the nature of imported proteins, the cell type used and delivery approach.[47]

In 2009, Zhou *et al.* reported the successful reprogramming of mouse somatic cells to pluripotency using recombinant proteins of 11R-Sox2, 11R-Oct4, 11R-Klf4 and 11R-c-Myc. Kim *et al.*(2009) also reported the successful induction of pluripotency in human fibroblast cells using 293T cell extracts that contained the same four reprogramming factors (RF) with the application of a polyarginine peptide C-terminal 9R PTD. [47,53–55]

Following the same point of view, Zhang *et al.* (2012) used reprogramming factors fused with TAT (TAT-RFs) or 11R to induce human foreskin fibroblast and reprogrammed these cells to iPSCs with success. The RFs used were the same as the mentioned by previous groups and one more, Nanog, and all of them were added to the culture medium. Comparing the efficiency of transduction between TAT and 11R, it was found that they were almost the same. However, regarding the reprogramming efficiency, TAT-RFs presented better results than 11R-RFs.[47]

Concerning the use of recombinant proteins in cardiac differentiation, Fonoudi and collaborators (2013) developed a recombinant protein which was added to hESCs cultures, and efficiently penetrated into the cells, enhancing the differentiation into cardiac cells. To do this, they developed a transduction system based on the fusion of TAT with Islet1 (Isl1). Isl1 is a marker of myocardial lineage during mammalian cardiogenesis and marks a common population of progenitors in the heart that can differentiate into cardiomyocytes, smooth muscle and endothelial cells. The intent was to improve the cardiomyocyte differentiation rate under a suspension culture condition and they have successfully demonstrated that the application of TAT-ISL1 increased the differentiation of cardiomyocytes (2–3 folds) without genetic modification.[25]

Since this method lacks of genetic manipulation, these kind of molecules can be easily applied in drug discovery or cell therapy since the risk for the application is diminished.

II. AIMS OF THE PROJECT

In this thesis, it was proposed to combine TAT and an arginine-rich peptide (composed by 8 arginines) with a human protein – CITED2 – which is involved in cardiac differentiation, to evaluate the transduction into mESC.

The specific aims of this study were the following:

- Cloning of two different PTDs into a pGEX vector and sequential cloning of the CITED2 gene
- Production and purification of GST-3C Protease cleavage protein
- Overexpression of recombinant proteins in *E.Coli*
- Purification and cleavage of the recombinant proteins
- Assessment of proteins functionality through their capacity to rescue Cited2 knockout mESCs cardiogenic defects

III. MATERIALS AND METHODS

This practical work aimed the synthesis and application of recombinant proteins to mESC. The process is not simple and involved many phases, since the DNA cloning to the direct application of the protein to the stem cells.

Molecular Cloning	• Restriction digestion of pGEX vector • Ligation of inserts to pGEX DNA • Transformation of DH5α cells • Isolation of pGEX DNA (Mini/MidiPrep) • Sequencing
Recombinant Proteins Production and Purification	• Transformation of BL21 (DE3) with plasmids • Monitor the expression of fusion protein • Harvest cells and lysis • Purification of Gluthation S-Transferase (GST) fusion protein • Cleavage and elution of fusion protein • Charaterization of recombinant protein
Cell assays	• Entry analysis of recombinant proteins into the cells • Evaluation of protein effects of the protein in cardiomyocyte differentiation

Figure III.1 – Schematic representation of the work developed in course of this thesis.

III.1. MOLECULAR CLONING

Gene cloning is a complex process that requires several DNA techniques, in which a desired fragment of DNA is inserted in a cloning vector, such as a plasmid, in order to increase the number of copies of the fragment of interest. This is usually achieved by inserting the vector in a host, being the most commonly used bacteria.

This can be done with enzymes which are used to excise the fragments of interest and the same enzymes to open up the plasmid. After obtaining the DNA from the vector and the desired fragment, another enzyme will be needed to integrate the fragment in the host genome (ligation).

The aim for our final cloning was to obtain CITED2 linked to TAT and 8R in a pGEX vector which will permit the further purification of the proteins, so we needed to do four subclonings into the pGEX-6P-1 vector (Amersham Biosciences):

- Insertion of TAT
- Insertion of 8R
- Insertion of CITED2 into each one of the created vectors

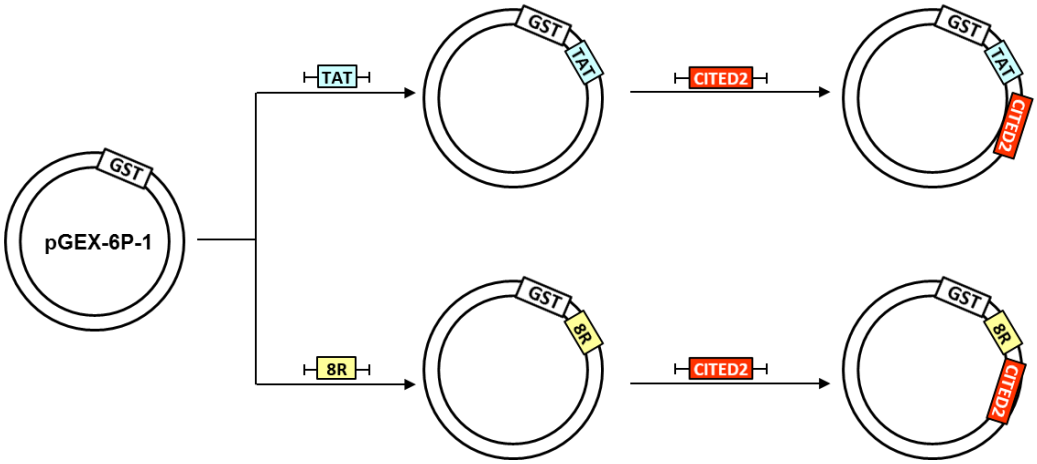


Figure III. 1 - Schematic representation of sequential fragment cloning into pGEX-6P-1 vector.

The pGEX-6P-1 vector (**Appendix 1**) was the expression vector chosen and is designed to express a Glutathione S-Transferase (GST) tag fused to the N-terminal of the protein of interest through a short peptide chain containing a specific endoprotease cleavage site. GST tag will allow the binding of the fusion protein to glutathione affinity columns for subsequent purification. This vector has a T7 promoter system, which is similar to the lac operon in *E. coli* and, being Isopropyl β-D-1-thiogalactopyranoside (IPTG) sensitive, so when IPTG is added to the culture medium, the transcription of protein linked to GST is induced. For bacterial selection, it has a Ampicillin (Amp) resistance gene.

III.1.1. RESTRICTION ENZYME DIGEST

A restriction enzyme is a naturally occurring bacterial endonuclease that allows to cut a DNA sequence, recognizing a specific nucleotide sequence also known as restriction site. This method have turn out to be the most widely used method to selectively transfer a specific

gene or a DNA sequence from one plasmid to another when the size of the plasmid insert is known. Another application is when the size of the plasmid insert and vector backbone are known, this technique can be used to verify and confirm the construct (diagnostic digest).

To diagnostic digests, the samples were prepared accordingly with **Table 1**, at 37ºC for 1h. To isolate inserts, samples were also prepared in the same way but the digestion was performed O.N. (overnight) at 37ºC.

Table 1 – Reaction mix used for the usual restriction digestion of the pDNA.

Reagent	*Volume*
DNA	*x* µL (1 µg)
Enzyme 1	1 µL
Enzyme 2	1 µL
Buffer 10X	5 µL
MilliQ H_2O	Up to 50 µL
Total	50 µL

To perform the restriction enzyme digest the double stranded DNA has to be incubated with the appropriate restriction enzymes in a suitable buffer recommended by the supplier.

For this work, we performed several digestions to accomplish different goals:

- Double digestion of pGEX-6P-1 with BamHI and EcoRV to check the quality and purity of the plasmid
- Digestion of pGEX-6P-1 with BamHI to insert TAT and 8R
- Double digestion of pGEX-6P-1+TAT (pGEXT) with SacII and PstI to check if cloning of TAT was successful
- Double Digestion of pGEX-6P-1+8R (pGEXR) with BamHI and PstI to check if cloning of 8R was successful
- Double Digestion of pGEXT with BamHI and XhoI to insert Cited2
- Double Digestion of pGEXR with BamHI and XhoI to insert Cited2
- Double Digestion of pGEXTC2 and pGEXRC2 with BamHI and XhoI to check if cloning of Cited2 was successful.

The enzymes used in this work and the appropriate buffers are listed in **Table 2**.

Table 2 – List of restriction enzymes used for digestions, with their recognition site and optimal conditions.

Enzyme	*Recognition site*	*Temperature*	*Concentration*	*Producer*
BamHI	5'…G▼GATCC…3'	37ºC	10 units/µL	Fermentas
EcoRV	5'…GAT▼ATC…3'	37ºC	10 units/µL	Fermentas
SacII	5'…CCGC▼GG…3'	37ºC	10 units/µL	Takara
PstI	5'…CTGCA▼G…3'	37ºC	10 units/µL	Promega
XhoI	5'…C▼TCGAG…3'	37ºC	10 units/µL	Fermentas

Buffers were selected considering the enzymes used. In the case of double digestion, the buffer was selected considering the best activity for both enzymes chosen.

III.1.2. AGAROSE GEL ELECTROPHORESIS

Gel electrophoresis is a standard method in molecular biology used to separate molecules accordingly with their size (in this case by length in base pairs) for visualization and purification. Applying an electric charge, negatively charged DNA move through an agarose gel matrix towards a positive electrode. This way, smaller molecules move faster than larger molecules, because they migrate more easily through the matrix.

This process was made to check the DNA from the vector, to purify the desired fragments or to know if the clonings were successful.

For each run was prepared a fresh 1% TAE-gel by dissolving the agarose powder into TAE buffer in a microwave for about 1 to 3 min (see composition at **Table 3**). After it cooled down for 5 min, GreenSafe (NYZTech, Portugal) was added, and the molten agarose was poured into the flat mold for hardening. A comb was placed to form wells for running the samples and let the agarose gel solidify for at least 30 min, removing carefully the comb before the run.

Table 3 – Composition of 1% TAE agarose gel.

Agarose Gel	
Agarose	0.5 g
TAE 1X	50 mL
GreenSafe	2 µL

The gel was then transferred into an electrophoresis unit filled with TAE-buffer. Samples were previously prepared with loading buffer (Orange Loading dye 6X) accordingly with **Table 4**, and carefully loaded into the wells along with a ladder mix (GeneRuler DNA Ladder Mix, Thermo Scientific), under a constant electric potential of 120 V for 40 min.

The posterior visualization of DNA bands was performed in ChemiDoc XRS Molecular Imager (Bio-Rad).

Table 4 – Sample preparation prior agarose gel electrophoresis.

DNA Sample	
DNA	20 µL
Orange Loading Dye 6X	3 µL

III.1.3. GEL PURIFICATION

Gel purification is the method that allows to isolate and purify DNA fragments after a standard agarose gel electrophoresis, cutting out the DNA bands which had the expected size of the agarose gel and executing several steps to purify the DNA samples.

In this work, we performed this method for several occasions: to isolate the plasmid vector pGEX-6P-1 DNA digested with BamHI, the pGEX-6P1+TAT and the pGEX-6P1+8R digested with BamHI and XhoI, and also CITED2 DNA insert from a preexisting plasmid (pSB54), digested with BamHI and XhoI. All the purifications were performed using GeneJET™ Gel Extraction Kit from Fermentas.

Purification was carried out according to the manufacturer's instructions. In summary, the excised fragments were placed in centrifuge tubes and weighted. For each 100 mg of

agarose gel, 100 µL of binding buffer was added, then the mix was incubated at 50-60ºC for 10min, inverting the tube a few times. The binding of DNA was performed by transferring the solution to a column, which was placed in a collection tube (provided by the kit). The sample was centrifuged for 1 min at 12000g and the flow-through discarded. Then, 700 µL of Wash Buffer were added to the column and centrifuged in the same conditions, being the flow-through also discarded. The centrifugation was repeated but with the empty column. Afterwards, the column was placed in a fresh 1.5 mL centrifuge tube and 50 µL of Elution Buffer was added to the column, which allowed the sample to be obtained after a centrifugation for 1 min at 12000g.

III.1.4. DEPHOSPHORILATION

One of the steps prior to the ligation of fragments to the vector, is dephosphorilation. This process will avoid self-closure after the restriction digestion since the pDNA become linearized.

To achieve that, 1 µL of Calf Intestine Phosphatase (CIP) (Finnzymes, Thermo Scientific) was added to the DNA obtained, mixed and incubated for 30 min at 37 ºC in a dry bath (CH-100, Heating/Cooling Dry Block, Biosan). After this time, the tubes needed to be submitted to heat inactivation of the enzyme, leaving it at 65ºC for 30min, to ensure that the enzyme would not interfere in the ligation reaction.

III.1.5. LIGATION REACTION

After restriction of the inserts and the vector, and subsequently dephosphorilation of the last, the ligation reactions were performed in order to the inserts to be integrated into the host vector (**Table 5**). For the ligation to happen, both DNAs needed to be incubated with an enzyme – T4 DNA ligase (Fermentas) – and its buffer (10X T4 Ligase Buffer, Fermentas), for at least 2 hours (sometimes it was necessary to extend the time to 4 hours) at Room Temperature (RT).

TAT and 8R DNA fragments were already prepared by professor J.Bragança (primers used are listed in **Table 6**). Both set of oligonucleotides encoding TAT or 8R were chemically

synthetized to harbor, after annealing, a BglII hemi-site at their 5' end and a BamHI site at their 3' end flanking the TAT or 8R sequence. The BglII hemi-site can hybridize a BamHI hemi-site, however the resulting BglII/BamHI composite site can no longer be digested by neither of these enzymes. Since these oligonucleotides were synthetic, their 5' extremities were phosphorylated with a T4 polynucleotide kinase in the presence of ATP before the annealing.

Table 5 – Ligation mix components.

Reagent	*Volume*
10X T4 Ligase Buffer	1 µL
pGEX-6P-1	3 µL
Insert	5 µL
T4 DNA ligase	1 µL
MilliQ H_2O	Up to 10 µL
Total	10 µL

Table 6 – Gene primers used in to prepare TAT and 8R inserts.

Gene	***Forward 5'-3'***	***Reverse 5'-3'***
TAT	GATCTGGCTACGGCCGCAAGAAACGCC GCCAGCGCCGCCGCGGTG	GATCCACCGCGGCGGCGCTGGCGGCGTT TCTTGCGGCCGTAGCCA
8R	GATCTGGCCGCCGCCGCCGCCGCCGCCG CCGCCGCCGCCGCGGTG	GATCCACCGCGGCGGCGGCGGCGGCGG CGGCGGCGGCGGCGGCCA

After ligation procedure, 5 µL of cloned DNA was used to transform chemically competent *E. coli* cells.

III.2. PREPARATION OF CHEMICALLY COMPETENT CELLS

Competent cells are bacterial cells that possess more easily altered cell walls by which foreign DNA can be passed through easily. The majority of cells cannot take up DNA efficiently unless they have been exposed to special chemical or electrical treatments to make them competent.

In this work *E. coli* cells were subjected to chemical treatment, where they were treated with high cation concentration (in this case $MgCl_2$) and then exposed to a heat shock which makes the cell membrane to become permeable for plasmid DNA.

Cultures were grown in 25 mL of Lysogeny Broth (LB; 10g/L Tryptone, 5g/L Yeast Extract, 10g/L NaCl, pH 7.0) medium at 37ºC in an orbital shaker (Agitorb 200 IC, Aralab®) until they reach an $O.D._{600nm}$ of 0.47. Cells were harvested by centrifugation at 1000 g for 10 min (Centrifuge 5810R, Eppendorf®), discarding supernatant. Accordingly with culture volume, cells were resuspended in 1/10 (v/v) of ice cold TSS buffer, previously prepared (10% Polyethylene glycol 8000, 5% Dimethyl Sulfoxide, 20 mM $MgCl_2$ and 85% autoclaved LB, pH 7.0), and incubated on ice for 10 min. Cells were either used immediately or frozen in 100 µL aliquots.

III.3. TRANSFORMATION OF COMPETENT CELLS

Transformation is the process where foreign DNA (like plasmids) is introduced into a bacterial cell. This process is very common and useful since it allows the replication of a specific plasmid DNA. The majority of the plasmids carry a gene for antibiotic resistance which can be used as a selection tool.

However, not all bacteria strains can be transformed, so they need to enroll a process to become competent, which makes their cell walls more likely to uptake DNA (see section III.2).

The bacterial cells used in this study were from the strain *E.coli* DH5α (for cloning) and BL21(DE3) (for protein expression), previously made competent. All the transformations were performed using the heat-shock method.

In cloning, bacteria was transformed by adding 5 µL of the ligation mix to 50 µL of DH5α and incubated on ice for 30 min. Further the sample was incubated at 42ºC for 1 min and placed on ice for 2 min. To make the cells recover from the heat shock, 250 µL of Super Optimal broth with Catabolite repression (SOC) were added to the cell suspension and incubated at 180 rpm, at 37ºC for 1 hour in an orbital shaker (Agitorb 200 IC, Aralab®). Bacteria were plated in LB agar plates with Amp (100µg/mL) and placed inverted on the

incubator at 37ºC O.N. For the usual transformation with purified and amplified DNA (miniprep, midiprep or protein expression), only 1 µL was needed to make the transformation.

When the colonies were the result from a ligation reaction transformation, we checked if the plasmid was well constructed, by purifying pDNA, double digesting it and analyzing the result by electrophoresis.

III.4. EXTRACTION OF DNA

III.4.1. MINIPREP

In order to extract the pDNA from the recombinant bacteria, to check by electrophoresis if the inserts were cloned in the right way to the cloning vector, the GeneJET™ Plasmid Miniprep (Thermo Scientific) was used.

Several colonies were isolated and transferred from the plate into a 2 mL LB medium, supplemented with Amp (100 µg/mL) and incubated O.N. at 37ºC, with agitation (200 rpm).

To perform the extraction, cultures were placed in a clean and sterile microcentrifuge tube, and spun down at 6800 g for 2 min being the supernatant discarded. To the pelleted cells 250 µL of Resuspension Solution (which contained RNase A) was added and a micropipette was used to resuspend the cells. Then 250 µL of Lysis Solution was used to disrupt the bacterial walls, being the solution mixed by inversion, making the solution clear and viscous. A volume of 350 µL of Neutralization solution was added to the mix, and the tube was inverted 4 to 6 times. The mixture was centrifuged at 14000 g for 10 min (Centrifuge 5810R, Eppendorf®).

The supernatant was transferred carefully to a fresh tube (avoiding the white precipitate) using a micropipette and applied to the Thermo Scientific GeneJET Spin Column. The columns were then centrifuged for 1 min and the flow-through was discarded, since the plasmid DNA binds to the column in this step. Then 500 µL of Wash Solution were added to the column to wash it and the column was centrifuged again for 1 min, being the flow-through discarded. This step was repeated one more time. The empty column was centrifuged for an additional 1 min to remove residual wash buffer.

Finally, the column was placed in a clean 1.5 mL microcentrifuge tube and the DNA was eluted by incubation in 50 µl Elution Buffer for 2 min and centrifugation for 2 min.

Once resuspended, the samples were quantified by spectrophotometry with NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific Inc, USA) and stored at -20ºC until further use.

Clonings which seemed to be successful were sent to StabVida for sequencing (plasmid maps and sequencing results at **Appendix 2** and **3**). Once the identity of the cloned fragments was confirmed, a larger scale preparation of plasmid DNA was required. To achieve this goal, midiprep was performed.

III.4.2. MIDIPREP

In the various stages of this experimental work several DNA constructs needed to be purified such as pGEX-6P-1 vector for the preparation of the ligation reactions, and/or extraction of various products of ligation reactions for the sequencing.

To obtain the plasmid DNA, the QIAGEN® Plasmid Midi Kit was used according to the manufacturer's instructions.

Previously, *E. coli* was transformed with the desired DNA and the day after, a colony was picked and inoculated in 100 mL of LB medium supplemented with Amp (100 µg/mL), allowing growth to occur O.N. at 37ºC in an incubator with agitation (200 rpm).

Cell suspension was then transferred to 50 mL conical centrifugal tubes and centrifuged for 15 min at 6000 g, at 4ºC. Supernatant was discarded and the pellet resuspended with 4 mL of Buffer P1 (which contained RNase) until no cell clumps were visible. After, 4 mL of Buffer P2 was added and all components were mixed thoroughly by inversion and incubated for 5 min at RT. Buffer P3 was prechilled and 4 mL were added to the suspension, mixed by inversion and incubated on ice for about 15 min. The sample was centrifuged for 30 min at 4ºC at 18000 g (Centrifuge 5810R, Eppendorf®).

Meanwhile, the resin column (Qiagen-tip 100) was assembled and needed to be equilibrated in a suitable buffer which was provided by the kit. A volume of 4 ml of Buffer QBT was added to the column and let through the column, being the flow through discarded.

Supernatant was transferred to the column and the solution was allowed to pass through the resin by gravity flow, discarding the flow-through.

For the washing step, 10 mL of Buffer QC was added to the column and once again passed within the column by gravity flow, being the flow through discarded. This step was repeated one more time.

The columns were transferred to fresh 15 ml conical centrifuge tubes and 5 ml of Buffer QF were added to release pDNA from the resin. Columns were discarded and 3.5 mL of isopropanol were added to the flow-through, being the solution mixed and centrifuged at 15000 g for 30 min at 4ºC. Supernatant was discarded, and 1 mL of 70% ethanol was used to wash the pellet, and the solution was transferred to a 1.5 mL centrifuge tube. The following step was a centrifugation at 15000 g, at 4ºC, for 10 min. Supernatant was discarded once more and pellet was allowed to dry at RT until no ethanol was visible.

Plasmid DNA was resuspended in 250 µL of TE buffer, and the sample was submitted to NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific Inc, USA) measurement in order to check the DNA concentration and stored at -20ºC until further use.

III.5. RECOMBINANT PROTEIN PRODUCTION

Plasmids were obtained ant they were used to transform *E. coli* BL21 (DE3) for expression of the proteins.

BL21(DE3) strain was chosen for the fact that is the most widely used host for target gene expression, because it has good characteristics such as protease deficiency, solubility enhancement and others. This strain with DE3, a λ prophage, carries the T7 RNA polymerase gene and the $lacI^q$. Expression of T7 promoter containing plasmids is repressed until induction of IPTG or Lactose.

III.5.1. PRODUCTION OF GST-8R-C2 AND GST-TAT-C2 PROTEINS

The day after transformation of the cells, a colony was picked from the plate and inoculated in 50 mL of LB medium (with Amp) O.N., 37ºC at 180 rpm.

Next day, $O.D._{600nm}$ was measured in a spectrophotometer (Pharmaspec UV-1700, Shimadzu) and the volume necessary to reach an $O.D._{600nm}$ of 0.1 was diluted into 500 mL of LB medium (with Amp). Culture was incubated with shaking (180 rpm) until $O.D._{600nm}$ reached 0.6-0.8, and at that point 0.2 mM of IPTG was added to the culture. The incubation continued for 4 hours in the orbital shaker but at a lower temperature (20ºC) at 180 rpm.

Samples of 1 mL were collected before and after induction by IPTG to be analyzed by Sodium Sodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-Page) and confirm the expression of the protein of interest.

Culture was transferred for containers which allowed to centrifuge that kind of volume in a swing-bucket rotor at 4000 rpm for 20 min at 4ºC. Supernatant was discarded and pellet was stored at -80ºC until purification procedure.

III.5.2. PRODUCTION OF GST-3C PROTEASE

GST-3C is a fusion protein of GST and human rhinovirus 3C protease.

The protease specifically recognizes the amino acid sequence Leu-Glu-Val-Leu-Phe-Gln▼Gly-Pro, cleaving between the Gln and Gly residues. Since the protease is fused to GST, it is easily removed from cleavage reactions using Glutathione Sepharose. This protease is maximally active at 4°C, and when cleavage is performed, it should be done at low temperatures, improving the stability of the target protein.

Plasmid DNA for the production of GST-3C was a gift from Dr Kamel El Omari, University of Oxford. For this protein, the production procedure was similar to the one of GST-8R-C2 and GST-TAT-C2 described before (see section III.4.1), and the differences rely on the concentration of IPTG applied to the medium (0.5 mM in this case), and the incubation time after induction (O.N. instead of 4 hours).

III.6. SDS-PAGE

SDS-PAGE (Sodium Sodecyl Sulphate-Polyacrylamide Gel Electrophoresis) is commonly used method to separate protein-mixtures according to their molecular weight,

based on their differential rates of migration through a matrix (a gel) under the influence of an applied electrical field. SDS, an anionic detergent, is added to the protein solution to linearize the protein, to cover the intrinsic charge of the different proteins and to apply a constant negative loading per unit mass. Additionally all samples are heated to 95 °C for 5 min for better linearization and denaturing of the proteins.

The velocity of the proteins depends on ionic strength, viscosity of the gel and temperature. A molecular weight marker (ladder) has to be added to determine the molecular mass of the proteins.

This technique was used to monitor two situations: if the induction of the protein production with IPTG in BL21 (DE3) was done successfully, and to evaluate the fractionated samples after the purification process.

Accordingly with the nature of samples, they were prepared to run on the gel with different methods. The collected samples (1 mL) from the bacterial growth, before and after induction with IPTG, were centrifuged at 8000 rpm for 10 min at 4ºC (supernatant was discarded). Pellet was then resuspended in SDS loading buffer 1X accordingly with samples absorbance at the time they were collected, and this way they could be compared in the gel. The samples from the purification process, were prepared in different ways (see **Table 7**), since the insoluble fraction (InsF), soluble fraction/injected sample (SF), flow-through (FT) and Wash were more concentrated comparatively with the digested and the eluted samples from the column.

Table 7 – Sample preparation prior SDS-Page electrophoresis.

	InsF (µL)	*SF (µL)*	*FT (µL)*	*Wash (µL)*	*Digested/Eluted (µL)*
Sample	5	5	5	5	20
MilliQ H_2O	15	15	15	15	-
4X Loading Buffer	5	5	5	5	5

Since proteins are mostly colorless, to monitor the expression of the protein we used a staining method which consists in the immersion of the gel in Coomassie Brilliant Blue

solution (10% acetic acid, 25% methanol, 0.2% Coomassie Brilliant blue G250) and subsequent destaining.

In the case of the results of purification process we used the Coomassie Brilliant Blue solution and Western Blot to confirm the identity of the protein.

Therefore, for each run an acrylamide gel need to be prepared, and it was composed by two parts: the stacking gel and the resolving gel (**Table 8**). Stacking gel is buffered by Tris-HCl at pH 6.8, while the resolving gel is also buffered by Tris-HCl but at pH of 8.8. The stacking gel has a low concentration of acrylamide and the resolving gel a higher concentration capable of retarding the movement of the proteins.

Table 8 – Composition of each layer of a 12% acrylamide SDS-Page gel.

	Resolving gel (12% acrylamide)	*Stacking gel (5% acrylamide)*
MilliQ H_2O	2.15 mL	1.45 mL
1.5 M Tris (pH 8.8)	1.25 mL	-
1 M Tris (pH 6.8)	-	0.25 mL
Acrylamide 40%	1.5 mL	0.25 mL
10% SDS	50 µL	20 µL
10% APS	50 µL	20 µL
TEMED	5 µL	2 µL

After the gel being polymerized and placed in a electrophoresis unit Mini Protean 3 (Bio Rad), a running buffer solution (25 mM Tris, 250 mM glycine, 0.1% (m/v) SDS, pH 8.3) was poured into the unit, and the samples were loaded into the wells as long as 2.5 µL of a proper molecular weight marker of a ladder (PageRuler^{TM} Prestained Protein Ladder, Fermentas Life Sciences).

The runs were performed at 140 V during approximately 2 hours. After the run, the gel was stained O.N. with Coomassie Brilliant Blue solution at RT and then in destaining solution (40% methanol, 10% acetic acid and 50% of MQ H_2O) until the gel is colorless except bluish protein bands, and kept in MilliQ H_2O.

III.7. PROTEIN PURIFICATION

III.7.1. GST-3C PROTEASE

Frozen pellet was thawed on ice and resuspended in 5 mL of ice cold Phosphate Buffered Saline (PBS) 1X, with 1mM of Dithiothreitol (DTT) and 1 mM of phenylmethylsulfonyl fluoride (PMSF). Sample was then sonicated on ice for a total time of 1 min (On 7'', Off 10'' cycle, amplitude of 60%) in a ultrasonic processor (VCX 130, VibraCell). Centrifugation was performed at 18000 g for 40 min at 4ºC to pellet debris. The pellet was discarded and the supernatant was filtrated through 0.45 µm sterile filter (Sarstedt) and then applied to a XK16 column packed with 6 mL Glutathione Sepharose Fast Flow (GE Healthcare) resin, using affinity chromatography in a fast protein liquid chromatography (FPLC) ÄKTA™ (Amersham Biosciences). All the process was monitored by Unicorn software.

The column was pre-equilibrated with approximately 5 column volumes of binding buffer (1X PBS, pH 7.3) and the sample was injected into the system. The flow rate of injection was 0,1-0,2 mL/min. After all the sample being loaded into the column, 5–10 column volumes of binding buffer was passed through the system (or until no material appears in the flow-through). Then, the fusion-protein-bound Glutathione Sepharose medium was eluted with Elution buffer (50 mM Tris-HCl, 10 mM Reduced Glutathione (GSH), pH 8.0).

Since the protein of interest was eluted with GSH, it was necessary to remove it from the samples. Dialysis is the method most used to remove this component.

A dialysis membrane (Regenerated Cellulose Membrane, Biotech) was cut and hydrated in distillated water, and sealed in both ends with clamps as the sample was applied. The membrane was immersed in 4000 mL of dyalisis solution (50 mM Tris HCl, 150 mM NaCl, 1mM, pH 8.0) O.N. at 4ºC with agitation.

The fractions were analyzed by SDS-Page and stored at -80ºC.

Quantification of the protein was done by measuring them with Bradford assay (see section III.8), using the dialysis solution as blank.

III.7.2. GST-8R-C2 AND GST-TAT-C2

The purification process was adapted from Frangioni and Neel (1993).

Frozen pellet was thawed on ice and resuspended in 10 mL of ice cold STE Buffer (10 mM Tris-HCl, pH 8.0;1 mM Ethylenediamine tetraacetic acid (EDTA);150 mM NaCl). Then, 5 mM of DTT, 1 mM of PMSF and 1.5 mL of 10% Sarkosyl (final concentration 1,5%) were added. All the components were mixed thoroughly by inversion and sonicated for a total time of 1 min (On 7'', Off 10'' cycle; model 75186, VibraCell). Then 8 mL of 10% Triton X-100 was added, and the volume was adjusted to a total volume of 20 mL with STE Buffer. At this point, the final concentration of Sarkosyl and Triton X-100 was 0.7% and 4%. The solution was incubated at RT for 1 h, and then centrifuged at 12000 rpm for 30 min to pellet debris. The pellet was discarded and the supernatant was applied to a XK16 column packed with 6 mL Glutathione Sepharose Fast Flow (GE Healthcare) resin using affinity chromatography in a fast protein liquid chromatography (FPLC) ÄKTA^{TM} (Amersham Biosciences), and all the process was monitored by Unicorn software.

The column was pre-equilibrated with approximately 5 column volumes of binding buffer (1X PBS, pH 7.3) and the sample was injected into the system. The flow rate of injection was 0,1-0,2 mL/min. After all the sample being loaded into the column, 5–10 column volumes of binding buffer was passed through the system (or until no material appears in the flow-through). Then, the fusion-protein-bound Glutathione Sepharose medium was washed with Cleavage buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.5) until the values of absorbance and conductivity were stable.

For each mL of Glutathione Sepharose volume, a mixture of 80 µl of GST-3C Protease and 920 µl of Cleavage buffer was prepared and loaded into the column. That process allowed to incubate this protease with the recombinant protein at 4ºC O.N., which permitted the cleavage of the sample. That way, the eluate will contain the protein of interest, while the GST portion of the fusion protein and the GST-3C Protease will remain bound to the Glutathione Sepharose medium.

Following incubation, cleavage buffer was applied to the column to elute all the cleaved protein (shown by the absorbance peak at 280 nm), being the eluate collected in fractions of 2 mL.

After the stabilization of absorbance, the GST-tag and GST-3C bound protein were eluted with Elution buffer (50 mM Tris-HCl, 10 mM reduced glutathione, pH 8.0).

The eluted fractions were analyzed by SDS-Page and/or Western Blot, and stored at -80ºC.

III.8. PROTEIN QUANTIFICATION: BRADFORD ASSAY

In order to quantify the purified protein, Bradford method was used, which is based on the application of a dye, Coomassie Brilliant Blue G250 (Bio-Rad), at concentration of known protein solutions (BSA), allowing the quantification of the unknown samples.

To prepare protein samples were added to 798 µL MQ H_2O, 2 µL of sample and 200 µL dye. The determination of the samples concentration was performed by reading the Abs_{595nm} in a spectrophotometer (Pharmaspec UV-1700, Shimadzu), as well as a straight calibration was made in triplicate from standard solutions of known concentrations of BSA (0-1500 mg/mL), prepared under the same conditions as the samples.

III.9. WESTERN BLOT

Blotting is a method to transfer proteins or nucleotides (DNA or RNA) onto a membrane carrier like nitrocellulose or nylon. Next the proteins or nucleotides are visualized by colorant staining (e.g. silver staining of proteins, labelling with chromophoric antibodies or radioactive molecules).

In this thesis, Western Blotting method was used to characterize the product of protein purification procedure and to confirm its entrance into the cells after application into them. With this method, proteins were separated by SDS-Page on a 12% acrylamide/bis-Acrylamide (29:1; NYZTech) gel electrophoresis and then transferred to a nitrocellulose membrane (Hybond-C Extra nitrocellulose, Amersham Biosciences) by using an electric current to pull the proteins from the SDS-PAGE gel into the membrane. For this to happen, the gel and the membrane were previously placed in transfer buffer (192 mM glycine, 20% methanol, 25 mM Tris) and subsequently between two pieces of Whatman paper, and two sponges. The assembly was inserted into a transfer cassette and the system placed in a container with

transfer solution and a cold plate. The electrodes were connected to a power source, and the transfer occurred within 60 min, at a fixed amperage of 250 mA and to a maximum of 100 V.

Afterwards, the nitrocellulose membrane had to be blocked with blocking solution (6% dry milk in PBS, 0.2% Tween 20) O.N. at 4ºC to reduce the unspecific reactions in the final product of the Western Blot. Then, the membrane was washed with washing solution (PBS, 0.2% Tween 20) 3 times for 10 min at RT and under agitation. To stain the transferred proteins, the membrane was incubated with a primary antibody - mouse monoclonal JA22 against CITED2 (ab5155, Abcam) at 1:2000 dilution - specific to the amino acid sequence of interest, for 1 hour at RT with agitation.

After, 3 washes were performed for 10 min with washing solution. Finally, an anti-mouse secondary antibody coupled to a chromophoric protein (HorseRadish Peroxidase-HRP) that binds to the primary antibody for 1 hour at RT with agitation. The washes were repeated after incubation. Revelation of the bands was made using SuperSignal® West Pico kit (Pierce), accordingly with manufacture's protocol, and the reaction was detected by image acquisition on ChemiDocTM XRS+ Imaging System (BioRad).

III.10. CELL ASSAYS

III.10.1. EMBRIONIC STEM CELL LINES

To test the functionality of the recombinant protein, we performed several experiments with three different cell lines.

The $C2^{\Delta/\Delta}$[MG5] *Cited2*-null ESC cells, described elsewhere by Kranc *et al*.(2015), were used to study the cellular entrance of the recombinant 8R-CITED2 protein using methods like Western Blot and Immunofluorescence. These cells have a deletion of both alleles of Cited2 gene and therefore do not express this protein[16], making them ideal to detect intracellular recombinant protein. These cells are hereafter referred as MG5 cells.

The $C2^{fl/fl}$[Cre] cells, described elsewhere by Kranc *et al*.(2015), were used to study the influence of the recombinant protein in the differentiation of ESC cells into cardiomyocytes. This mESC line harbour the CITED2 alleles flanked by LoxP sites, and constitutively express a tamoxifen-inducible Cre recombinase.

E14/T ESC were used for Bifluorescence complementation assays (BiFC). This mESC cell line was derived from E14.Tg2a, which was derived by Hooper *et al*. from 129/OLA in 1987[56] and originally developed as a mouse model of Lesch-Nyhan disorder with a deficiency of hypoxanthine phosphoribosyltransferase (HPRT). This ESC line grows fast and steadily in both feeder and feeder-free systems and is an ideal system for genetic engineering.

mESC lines were cultured on 0.1% (SIGMA) gelatin-coated plates in GMEM (Glasgow MEM, Pan Biotech) medium supplemented with 10% Fetal Bovine Serum (FBS) (Sigma), 4 mM L-Glutamin (200 mM GIBCO®, Invitrogen™), Penicillin/Streptavidine mix (GIBCO®, Invitrogen™), 1X Non-essential aminoacids (GIBCO®, Invitrogen™), 0.05 mM β-mercaptoethanol (GIBCO®, Invitrogen™) and 1 mM Sodium Pyruvate (GIBCO®, Invitrogen™). Additionally 1000 U/mL of Leukemia Inhibitory Factor (LIF) (ESGRO mLIF Meddium Supplement, Millipore) was added to the medium, since it is an interleukin 6 class cytokine that affects cell growth by inhibiting differentiation.

Cultures were maintained in culture at 37ºC with 5% CO_2.The cultures were monitored under inverted microscope.

III.10.2. CELLULAR ENTRANCE OF RECOMBINANT PROTEIN

$5x10^4$ MG5 cells were seeded in a 24 well plate, previously prepared with 0.1% gelatin and incubated at 37ºC, 5% CO_2, O.N.

Protein solutions were thawed and diluted with the culture medium and stored at 4ºC prior to supplementation of cell cultures.

Next day, the medium in contact with the cells was replaced by the medium where the protein had been dissolved at the desired concentration (20 µg/mL).

At different time points (4 h, 8 h, 12 h and 24 h), cells were recovered, by washing them with PBS 1X, adding 80 µL of trypsin (a proteolytic enzyme which breaks down proteins) to dissociate adherent cells from the plate. Since long exposure to trypsin can affect the cells, when the cells appeared to be detached, 160 µL of medium with serum was added to deactivate the enzyme. The mixture was then transferred to a 1.5 mL centrifuge tube and

centrifuged at 15000 rpm for 1 min, being the supernatant discarded and the pellet stored at -80ºC.

Protein lysates were prepared in cold lysis buffer (Tris 50 mM (pH7.5), 180 mM NaCl, 0.5% NP-40) supplemented with complete protease inhibitors 25X (Roche), 1 mM DTT, 0.2 µM Sodium Ortho-Vanadate (Na_3VO_4, Sigma) and 100 µM PMSF (PhenylMethylSulfonyl Fluoride, Sigma). Final protein concentrations were measured using Bradford protein assay as previously described.

Approximately 30 µg of protein lysates were applied to a SDS-Page gel and the process of Western Blotting was done as previously described (see section III.7).

III.10.3. FLUORESCENCE IMMUNOCYTOCHEMISTRY

In order to detect if 8R-CITED2 would enter the cells and the nuclei, a fluorescence immunocytochemistry assay was performed in complementation with a Western Blot.

This technique is composed by 6 steps: coverslip preparation, fixation of cells, permeabilization of cells, blocking, antibody incubation and DAPI staining.

First, is necessary to prepare the cells to the visualization in slides, so the cells needed to be placed into coverslips. These coverslips required to be prepared in order to the cells adhere (the cells don't adhere to glass), so they were washed with 70% ethanol in a 50 mL conical tube, followed by a wash with PBS 1X. Using clean forceps, coverslips were hold vertically to dry and placed inside a well in a 6-well plate. To each of them, 200 µL of 0.1% gelatin covered the surface of the coverslip for 10 min and then removed. MG5 cells were plated into the coverslips ($5x10^4$ cells/coverslip) in a total volume of 300 µL of medium per coverslip. 8R-CITED2 was added to medium with the cells at 20 µg/mL concentration, and let incubate at 37ºC O.N.

The day after, we proceeded to fixation of cells. For that, the media from the coverslips was removed and the cells were washed with 1 mL PBS. PBS was removed and 1 mL of 2% paraformaldehyde (Sigma) diluted in PBS was added and incubated for 20 min at RT.

To allow the antibodies to enter the cells, it was necessary to make them permeable. For that, the fixation solution was removed, and 1 mL of Permeabilization solution (0.1% (v/v) Triton X-100 (Sigma)) was added and incubated for 20min at RT. Then, permeabilization solution was removed and 1 mL of Washing solution (.,02% Tween in PBS) was added for 5 min to wash the cells. The washing solution was then removed and the blocking step was performed.

To each coverslip, 1 mL of Blocking solution, containing 100 mM glycine (Merck), 2 mg/mL Bovine Serum Albumin (BSA; NZYTech) and 0.02% sodium azide (Sigma) in PBS, was added and incubated at least for 30 min and keep at 4ºC until use.

The following step was antibody incubation. First, 150 µL of the primary antibody - rabbit polyclonal anti-CITED2 (H-220, Santa Cruz Biotechnology) - was added at 1:100 dilution in blocking solution, for 2 h at RT. Next, coverslips were washed three times in 0.02% Tween (Sigma) diluted in PBS for 5 min and incubated with 150 µL of secondary antibody - AlexaFluor-488 donkey anti-rabbit antibody (Invitrogen) - at 1:2000 dilution in blocking solution for 1 h at RT. The coverslips were then washed three times with washing solution for 5 min.

To make the nuclei visible, it was needed to stain it with DAPI. A volume of 500 µL of DAPI at 1 µg/mL diluted in PBS was used for coverslip and incubated for 5 min. Then, the coverslips were washed with washing solution for 5 min, and placed onto slides using Fluoromount (Sigma).

Fluorescence microscopy was performed using an Axioimager Z2 microscope (Carl Zeiss) at 100X magnification. Negative controls were used to set up exposure conditions for detection of a specific signal.

Treatment of resulting images was done using software ImageJ, respecting the same parameters in control cells and recombinant protein treated cells.

III.10.4. BIMOLECULAR FLUORESCENCE COMPLEMENTATION (BIFC)

Protein interactions are a fundamental mechanism and can be studied by different methods which make the protein-protein interactions visible, not only revealing the interactions, but also show where those interactions take place. This is of great importance to study the interactions among transcription factors.[57]

BIFC analysis is one of the methods which enables direct visualization of protein interactions in living cells. The principle is based on the formation of a fluorescent complex, when two non-fluorescent fragments of a fluorescent protein (like YFP), interact to each other. The interaction between the fusion proteins facilitates the association between the fragments of the fluorescent protein making the fluorescence appear. [23]

The vectors used in this assay were described previously by the group[58], using pPyCAGIP vector modified and obtaining the following vectors:

- VEC-CITED2, which is a vector expressing the C-terminal fragment (residues 156-239) of Venus-YFP fused in frame to a flag-tagged CITED2 at its N-terminal region.
- VEN-CH1, which is a vector expressing the N-terminal fragment (residues 1-155) of Venus-YFP fused in frame to the CH1 domain of p300 (high affinity domain of interaction with CITED2).

Fluorescence detection of transfected cells will be originated by the close proximity of the residues 1-155 and residues 156-239 of Venus-YFP.

First, cells had to be prepared the day before transduction, so we plated $5x10^4$ of E14/T ESC for well in a 24-well plate, and let the cells incubate at 37ºC, 5% CO_2 for 24 h. At the same time, protein solutions were thawed and diluted with the culture medium 24 h and stored at 4ºC prior to supplementation of cell cultures.

Next day, the transfection solution was prepared to each well. 25 ng of DNA of each of the vectors was diluted in medium and the solution of Lipofectamin 2000 (Invitrogen) was carefully added to the DNA mix, drop by drop. Before application of the transfection solution, the cells were washed with medium and finally the solution was applied to the wells. After 4 h, 2 mL of fresh complete medium were added to each well.

After 24 h, the medium was changed and at that time we added 8R-CITED2 at 20 µg/mL and 40 µg/mL concentration.

After days, pictures were taken at 10X in a confocal microscope, and cells were counted using ImageJ software.

III.10.5. CARDIAC DIFFERENTIATION - HANGING DROP METHOD

To induce ESC differentiation, hanging drop method was the chosen one. To this experiment, GMEM was prepared with supplements as referred before, but without the addition of LIF (since this factor maintains pluripotency of the cells and we wanted them to differentiate), and raising the concentration of FBS to 20%.

$C2^{fl/fl}$[Cre] cells were prepared in a cell suspension containing a defined cell number (5×10^4 cells/plate) and placed in a 100 mm bacteriological petri dish in 20 µL volume drops (Day 0). Then plates were inverted (upside down), and 1 mL of PBS was added to the bottom part of the plate to maintain the humidity inside them. Plates were cultivated in hanging drops at 37ºC with 5% CO_2 for 2 days. The cells aggregated and form one EB per drop.

At day 2, EBs were formed, and PBS was removed. Plates were then inverted and 4 mL of GMEM medium was added. After 3 days, EBs were transferred onto 60 mm tissue culture dishes, previously coated with gelatin.

The first beating cardiomyocytes appear in two to three days after plating of 5 days old EBs.

To study the influence in cardiomyocyte differentiation by the produced recombinant proteins, we had to compare them with different conditions to the controls.

At day 2 of differentiation, the plates were inverted and 8R-Cited2 was added to the medium in concentrations of 5-10 µg/mL where $C2^{fl/fl}$[Cre] had suffered a knock-down by the addition of 1 µM 4-hydroxytamoxifen (4HT). As controls we used $C2^{fl/fl}$[Cre] with ethanol (4HT is diluted in ethanol), and $C2^{fl/fl}$[Cre] with 4HT.

We observed the differentiation of cells in culture by light microscopy, since it is a simple and efficient way to assess the morphological changes of the cells and to observe the

emergence of beating foci. Percentage of colonies with beating foci and average number of beating foci per colonies with beating foci were counted 8 to 10 days after the initiation of differentiation.

III.11. STATISTICAL ANALYSIS

The Two-way analysis of variance (ANOVA) with the multiple comparison test (Tukey test) were performed to compare groups. All analyses were run using the GraphPad Prism® (version 6.1) software.

IV. RESULTS AND DISCUSSION

IV.1. CONSTRUCTION OF EXPRESSION VECTORS

The empty pGEX-6P-1 vector was digested by the endonuclease BamHI to linearize the vector. After dephosphorilation with CIP to avoid self-closure of the vector during the ligation process, a TAE agarose gel was performed for further purification of the band.

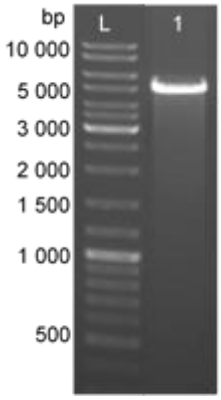


Figure IV.1 - Visualization of pGEX-6P-1 digestion product by BamHI in a 1% agarose gel. Gel was stained with GreenSafe. Lane L corresponds to the ladder GeneRuler DNA Ladder Mix (Thermo Scientific). Lane 1 shows the pGEX-6P-1 digestion product.

The linearized vector has an expected length of 4984 bp, therefore the band visualized on agarose gel electrophoresis which migrates close to 5000 bp indicates a correct digestion of the plasmid (**Figure IV.1**).

This band was then cut off the gel, purified with GeneJET^{TM} Gel Extraction Kit and the DNA resulting from the extraction was quantified using a Nanodrop 2000.

Next, we inserted at the BamHI site of the pGEX-6P-1 vector, double stranded oligonucleotides encoding either the TAT or 8R peptides. As mentioned in methods section, these oligonucleotides were designed to place TAT or 8R in frame with the GST protein expressed by pGEX-6P-1.

The ligation between the inserts and the linearized vector was performed and the products of the ligation were transformed into DH5α cells, which were plated and incubated overnight at 37°C. A negative control (a digested and dephosphorylated plasmid without the insert) was performed at the same time to estimate the background of the transformation due

to an auto-ligation of the empty vector. No colonies were observed in the control plate (the linearization and dephosphorilation steps were successful), while four single colonies were present in pGEX+TAT transformed cells plate. The positive colonies were amplified in liquid medium and the plasmid DNA of these colonies was purified and subjected to a diagnostic digestion with SacII and PstI.

Restriction enzymes for the diagnostic digest were chosen based in their possible restriction sites in the vector or insert. PstI only has a unique restriction site in the host vector (position 1937bp), which would linearize the vector. Meanwhile, the SacII enzyme only has a unique restriction site localized in the insert. Thus, a double digestion of the expected construct, in other words, with the TAT inserted in the correct orientation, would generate two DNA fragments of 996 bp and 4033 bp. The insertion of the TAT expressing oligonucleotides in the wrong orientation would give rise to two fragments of 1027 bp and 4002 bp, while the absence of insert would give rise to a unique fragment with a size of 4984 bp.

After digestion by PstI and SacII of the four clones, the resulting fragments were separated and visualized in an agarose gel (**Figure IV.2**).

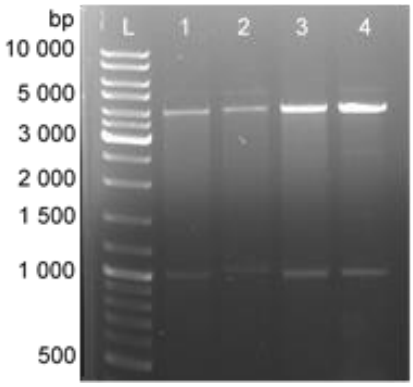


Figure IV.2 - Visualization of the plasmid DNA isolated from positive colonies to pGEX-6P-1+TAT, digested with SacII and PstI in a 1% agarose gel. Gel was stained with GreenSafe. Lane L corresponds to the ladder GeneRuler DNA Ladder Mix (Thermo Scientific). Lane 1 to 4 shows double digestion of 4 possible successful clones of pGEX-6P-1+TAT.

As we can observe three of the four clones, 1, 3 and 4, present two bands at approximately 1000 bp and 4000 bp (theoretically corresponding to bands of 996 bp and 4033 bp) suggesting they underwent an appropriate ligation. Clone number 2 showed two bands

(**Figure IV.2**), approximately at the same sizes, but the band at 1000 bp is slightly above the others, which suggested that the insert is inverted (theoretically corresponding to bands of 1027 bp and 4002 bp).

For further experiments, we selected the clone number 4 because the concentration of plasmid DNA was higher than the others (42 ng/µl).

For the pGEX+8R construction (**Figure IV.3**), no colonies were observed in the control plate and two single colonies were present in the plate of cells transformed with the ligation products resulting from the ligation of the plasmid and the insert. The plasmid DNA of the 2 positive clones was amplified, purified and subjected to a diagnostic digestion with BamHI and PstI. A correct insertion of the oligonucleotides encoding the 8R in the vector should give rise to two bands of 996 bp and 4033 bp after a double digestion by BamHI and PstI.

As we can observe only one (clone 1) of the two colonies had an appropriate plasmid DNA digestion profile on an agarose gel (**Figure IV.3**). Clone number 1 was selected and clone number 2 was discarded, since it showed only one band, which means that pGEX-6P-1 entered the cells allowing them to grow as a colony (due to a working Amp resistance gene) but didn't incorporate the insert.

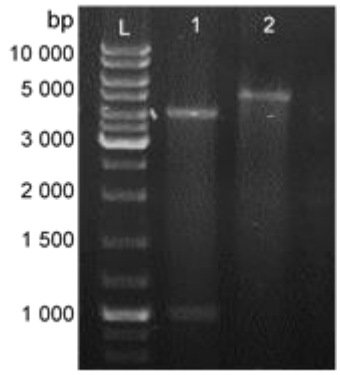


Figure IV.3 – Visualization of the plasmid DNA isolated from positive colonies to pGEX-6P-1+8R, digested with SacII and PstI in a 1% agarose gel. Gel was stained with GreenSafe. Lane L corresponds to the ladder GeneRuler DNA Ladder Mix (Thermo Scientific). Lane 1 and 2 show double digestion of the possible successful clones of pGEX-6P-1+8R.

Next, we needed to isolate the CITED2 DNA sequence from an existing expression vector (pSB54[31]) and insert it into the two previous prepared vectors: pGEX-6P-1+TAT (referred hereafter as pGEXT) and pGEX-6P-1+8R (referred hereafter as pGEXR).

The CITED2 insert was excised from pSB54, using the BamHI and XhoI restriction enzymes. The same enzymes were used to digest the pGEXT and pGEXR vectors in order to perform the cloning. Importantly, the insertion of the CITED2 fragment excised by BamHI and XhoI in these vectors will locate the CITED2 cDNA in frame with the GST and the TAT or 8R peptide sequences.

To control the restricted endonuclease reaction and to select the cut out polynucleotides a TAE agarose gel was performed.

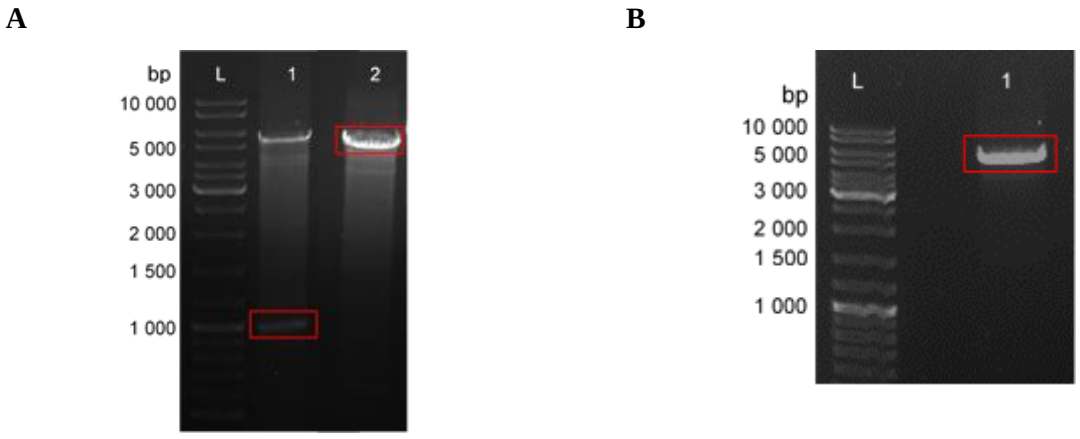


Figure IV.4 – Visualization of pSB54, pGEXT and pGEXR double digested with BamHI and XhoI, in a 1% agarose gel in order to isolate and purify inserts and vectors. Gels were stained with GreenSafe. Gel A corresponds to the double digestion of pSB54 (lane 1) and pGEXT (lane 2). Gel B corresponds to the double digested pGEXR (lane 1). Lane L corresponds to the ladder GeneRuler DNA Ladder Mix (Thermo Scientific) and red rectangles correspond to the fragments excised to further purification of the DNA.

The length of the vector pSB54 is 6391 bp and the double digestion by BamHI and XhoI is expected to originate two fragments: one of the fragments with a length of 923 bp corresponds to the CITED2 fragment and the other fragment with a length of 5468 bp results from the remaining plasmid DNA.

The digestion of pGEXT and pGEXR vectors by BamHI and XhoI should originate a 5005 bp fragment corresponding to the linearized vectors and a fragment of 24 bp long, which will not be detected on a conventional agarose gel.

On an agarose gel, we separated the fragments resulting from the digestion of pSB54 vector, and determined that a fragment corresponding to the vector itself (above 5000 bp) was present with the fragment corresponding to CITED2 DNA, which had about 1000 bp, according to the ladder on the gel (lane 1 from **Figure IV.4.**A). In lane 2 (pGEXT), the expected fragment had the right length (around 5000 bp) while the cut out DNA fragment of 24 bp is not visible. The same happened with in pGEXR (lane 1 from **Figure IV.4**.B).

The relevant DNA fragments were excised from the gel and purified as referred in section III.1.3.

After that, we needed to insert the CITED2 DNA fragment into pGEXT and pGEXR vectors by ligation (for 2 h and 4 h). The ligation products were transformed into DH5α cells, which were plated and incubated O.N. at 37°C. The negative control (plasmid without the insert) transformed cells was also obtained at the same time to verify if the vector would self-close.

A **B**

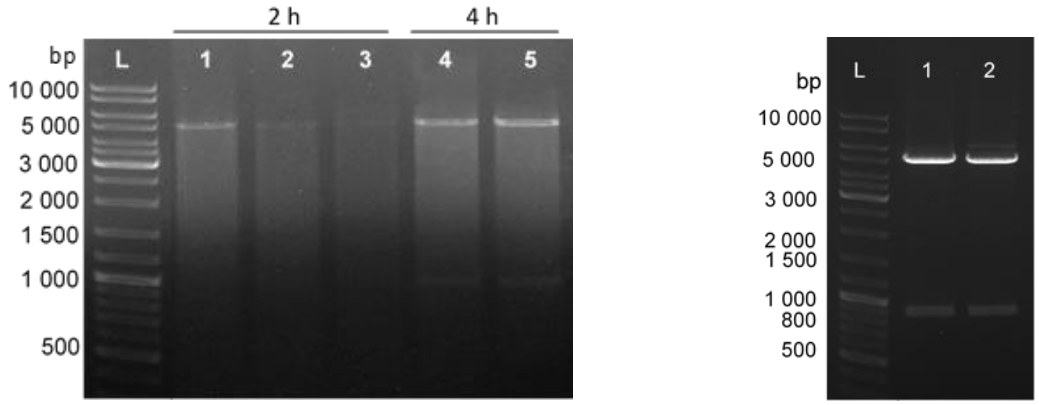


Figure IV.5 - Double digestion with BamHI and XhoI of pGEXTC2 (A) and pGEXRC2 (B) in a 1% agarose gel in 1X TAE, stained with GreenSafe, to confirm the insertion of CITED2 gene. (A) Lanes 1 to 3 correspond to DNA isolated from colonies transformed with plasmid DNA ligated for 2 h. Lanes 4 and 5 correspond to DNA isolated from colonies transformed with plasmid DNA ligated for 4 h. (B) Lanes 1 and 2 correspond to DNA isolated from colonies transformed with plasmid DNA ligated for 4 h. Lane L corresponds to the ladder GeneRuler DNA Ladder Mix (Thermo Scientific).

Three single colonies arose in pGEXT-CITED2, which suffered ligation reaction for 2 hours, and two colonies from DH5α transformed with DNA ligated for 4 hours. All colonies were picked from transformed cells plate, amplified and the plasmid DNA was purified and used for a diagnostic test digestion with BamHI and XhoI.

In the case of pGEXR-CITED2, the ligation reaction was done for 4 h since it seemed to increase the efficiency of the reaction. After transformation of DH5α, plasmid DNA from two colonies was extracted to perform the same test diagnostic digestion. The expected bands were 1000 bp corresponding to *Cited2*, and 5000 bp corresponding to the vector itself (pGEXT or pGEXR).

Observing the results (**Figure IV.5**), we can conclude that neither of the 3 clones ligated for 2 h (lane 1 to 3) was a successful one, since none of them present the expected bands, showing only one band at 5000 bp, and therefore not having the insert.

Opposing to these results, the DNA fragments (pGEXT and CITED2) which suffered the ligation reaction for a longer time (4 h), presented a band at 1000 bp in both of the picked colonies. This band corresponds to the insert CITED2 which was successfully cloned into pGEXT.

In the gel corresponding to the diagnostic digest from pGEXR-CITED2 colonies, we can distinguish two good defined bands at 1000 bp and 5000 bp, which confirmed the presence of the insert. After this selection, cloned plasmid DNA was sent to sequencing.

IV.2. RECOMBINANT PROTEIN PRODUCTION

For expression of recombinant proteins, a suitable production system had to be established. Initially we faced some problems, such as the poor solubilization of the bacterially expressed protein, forcing its accumulation in inclusion bodies, which leads to a low yield in the production of the protein.

Many expression screenings were performed to overcome the problems and optimize the expression, like modifying the starting point of the induction, varying IPTG concentrations, duration of induction and temperature of expression.

IV.2.1. PRODUCTION AND PURIFICATION OF 3C PROTEASE

The 3C protease protein was produced as described in the methodology section, using BL21 (DE3) cells. Samples were collected before and after induction by IPTG to be analyzed

by SDS-Page to confirm the expression of the protein of interest after total protein stain (**Figure IV.6**).

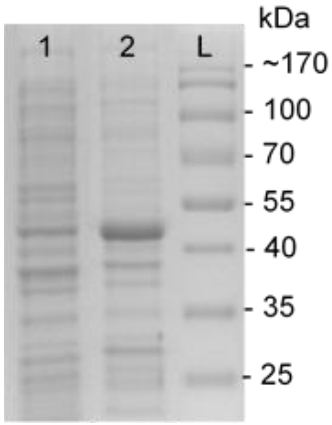


Figure IV.6 - SDS gel stained with Coomassie Blue to verify the expression of the recombinant protein: GST-3C. 1 represents the protein extract from cell lysate before IPTG induction and 2 represents the protein extract from cell lysate after IPTG induction. L corresponds to the ladder used: PageRuler^{TM} Prestained Protein Ladder (Fermentas Life Sciences).

The expression of a protein of 46 kDa was stimulated in the bacteria transformed with the GST-3C protease expression vector and treated with IPTG (**Figure IV.6**, lane 2), suggesting that GST-3C protease is expressed in our extracts.

We further proceeded to the purification of the protein of interest (GST-3C) from the other proteins produced by BL21 (DE3). To separate these components, we resorted to chromatographic affinity methods, which are based on the specific interaction of the sample with a substrate (GST) in stationary phase (chromatographic column).

The frozen cells transformed with the GST-3C expression vector and treated with IPTG were resuspended in PBS and disrupted by sonication. After the separation of the supernatant and the cell lysates, the supernatant was applied to a XK16 column packed with Glutathione Sepharose Fast Flow resin to perform an affinity chromatography in a FPLC ÄKTA^{TM}. The basis of this method is the interaction of GST fusion protein with the resin of the GSTrap column.

All the process was monitored by Unicorn software. This software allowed to register the alterations in several conditions like Abs 280nm and % of Elution Buffer injected into the column, and this way we could obtain a chromatogram of the process.

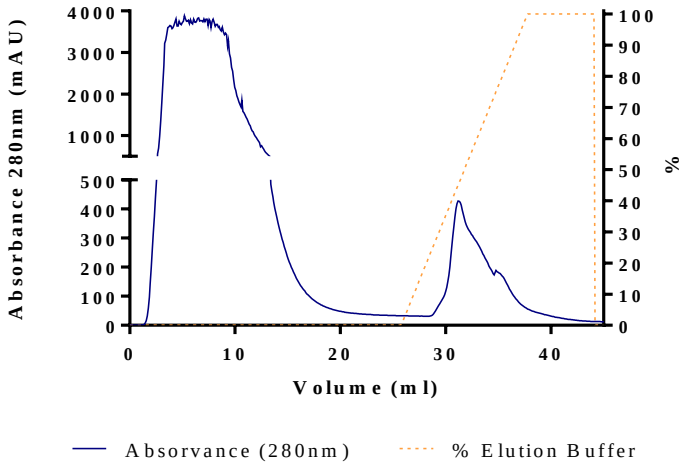


Figure IV.7 – Purification chromatogram of the recombinant protein GST-3C. 10 mL of soluble extraction was injected into a XK 16/20 column packed with Glutathione Sepharose 4 Fast Flow, and isolated by affinity chromatography to GST-GSH. Elution was performed with 10 mM of GSH.

Based on **Figure IV.7**, it is possible to evaluate the stages of the chromatographic purification process of GST-3C protein. The Abs_{280nm} is an indicator of the presence of protein in the collected solution.

At the start of the chromatographic process, where the sample is injected into the column, we can observe a remarkable increase in absorbance corresponding to the proteinaceous material that was not adsorbed.

At this step, GST fused proteins were retained, due to their affinity to the Glutathione (GSH) coupled to the resin, while the remaining was entrained in the mobile phase.

After selective immobilization and the collection of flow-through, the absorbance stabilized. Then, the protein of interest could be eluted from the column when in the presence of a solution of reduced GSH, a competitive molecule. The addition of this Elution Buffer to the column can be observed by the dotted yellow line in the chromatogram, and should be gradual in order to try to obtain the protein with higher grade of purity. The protein sample output increases the absorbance and at this point, the purified protein was collected in 7 fractions of 2 mL.

Samples from the injected solution, flow-through and the fractions of interest were analyzed by SDS-Page.

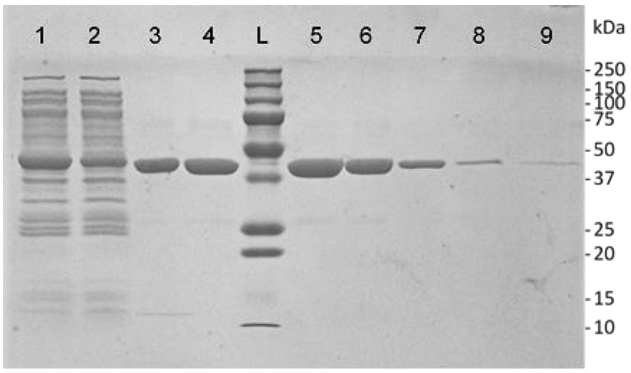


Figure IV.8 – SDS-Page analysis of sample purification steps stained with Coomassie Blue. Lane 1 corresponds to the soluble fraction/injected sample, lane 2 shows the flow-through and lanes 3 to 9 shows the eluted fractions of GST-3C. L corresponds to the ladder used: Precision Plus Protein™ Unstained Standard (Bio Rad).

The analysis of the SDS-Page gel (**Figure IV.8**), indicates that the injected sample, corresponding to the mix of all the soluble proteins produced by the transformed BL21 (DE3), has a ticker band which corresponds to the protein of interest. In flow-through fraction, it is possible to see that our protein is still present.

The eluted fractions had sharp and well defined bands migrating between the 50 kDa and 37 kDa bands of molecular weight markers, corresponding to our protein of interest which has 46 kDa.

No other bands were visible in the eluted fractions, meaning that the purification process was efficient, and allowed the preparation of GST-3C with a high grade of purity.

IV.2.2. PRODUCTION AND PURIFICATION OF GST-TAT-CITED2 AND GST-8R-CITED2 PROTEIN

These recombinant proteins were produced as described in the methodology section, using BL21 (DE3) cells.

Samples were collected before and after induction by IPTG to be analyzed by SDS-Page to confirm the expression of the protein of interest (**Figure IV.9**).

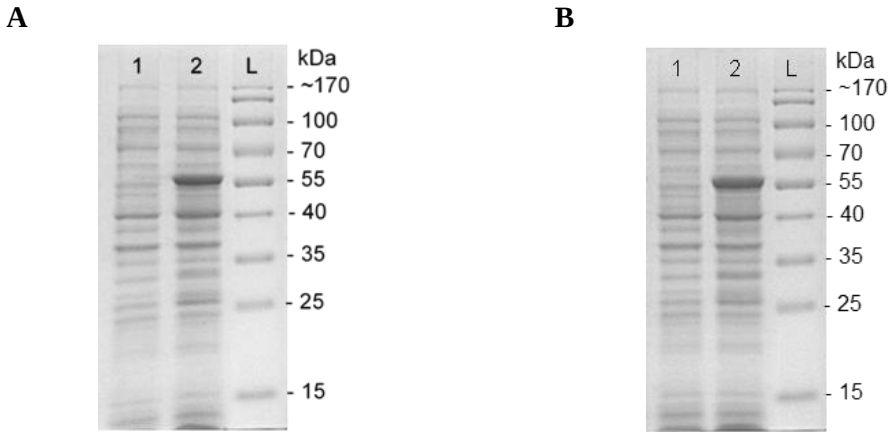


Figure IV.9 – SDS-Page gels stained with Coomassie Blue to verify the expression of the recombinant proteins (A) GST-TAT-CITED2 and (B) GST-8R-CITED2. 1 represents the protein extract from cell lysate before IPTG induction and 2 represents the protein extract from cell lysate after IPTG induction. L corresponds to the ladder used: PageRuler^{TM} Prestained Protein Ladder (Fermentas Life Sciences).

The analysis of the SDS-PAGE results indicates an increase in the production of GST-TAT-CITED2 and GST-8R-CITED2 after the induction by IPTG (**Figure IV.9**). GST-TAT-CITED2 has a predicted molecular mass of 57 kDa and examining the presented gel, we can confirm that the band induced by IPTG has the expected molecular weight. Analyzing the gel representing the production of GST-8R-CITED2, we can also observe a band with the expected molecular weight which is approximately 56 kDa.

Next, we proceeded to the purification of these proteins of interest using an approach similar to the one used for the purification of GST-3C, employing chromatographic methods.

In both cases, frozen bacterial pellets were thawed on ice and resuspended in ice cold STE Buffer with DTT, PMSF and Sarkosyl, and disrupted by sonication on ice. Triton X-100 was added to lysate, and the mixture was left to incubate, before performing a centrifugation in order to separate the supernatant and the cell debris. The soluble fractions were applied to the XK16 column packed with Glutathione Sepharose Fast Flow resin using affinity chromatography in a FPLC ÄKTA^{TM}. The basis of this method was described before for GST-3C protease purification.

All the process was monitored by Unicorn software to obtain a chromatogram of the process. Unfortunately, the chromatogram of the purification of TAT-CITED2 could not be obtained since the software failed to save the data, however that chromatogram was

essentially similar to the chromatogram obtained during the purification of 8R-CITED2 (**Figure IV.10**).

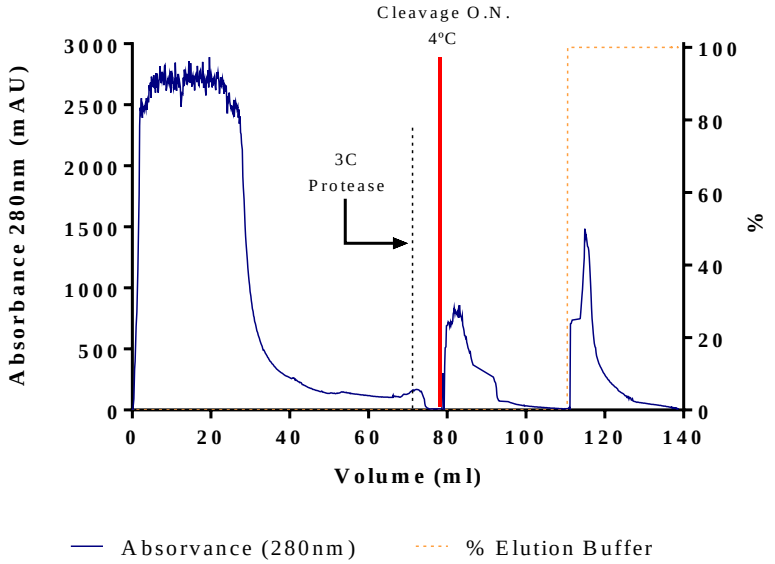


Figure IV.10 - Purification chromatogram of the recombinant protein 8R-CITED2 with on-column cleavage. 20 mL of soluble extraction was injected into a XK 16/20 column packed with Glutathione Sepharose 4 Fast Flow, isolated by affinity chromatography to GST-GSH, and cleaved O.N. at 4ºC with GST-3C protease. 8R-CITED2 (first peak after cleavage) was then released from the column and recovered in fractions. Further elution of GST-tag and GST-3C was performed with the application of 10 mM of reduced GSH.

The samples were injected into the column (20 mL), and the GST fused proteins were retained, due to their affinity for the resin, while the remaining proteins exited the system (flow-through sample was collected). After this selective immobilization the absorbance stabilized.

At this point, we injected the GST-3C protease to the column to cleave the GST tag from the protein of interest TAT-CITED2 or 8R-CITED2. We performed this on-column. On-column cleavage (see **Figure IV.11**) is generally recommended as the method of choice since many potential contaminants can be washed out and the target protein eluted with a higher level of purity.

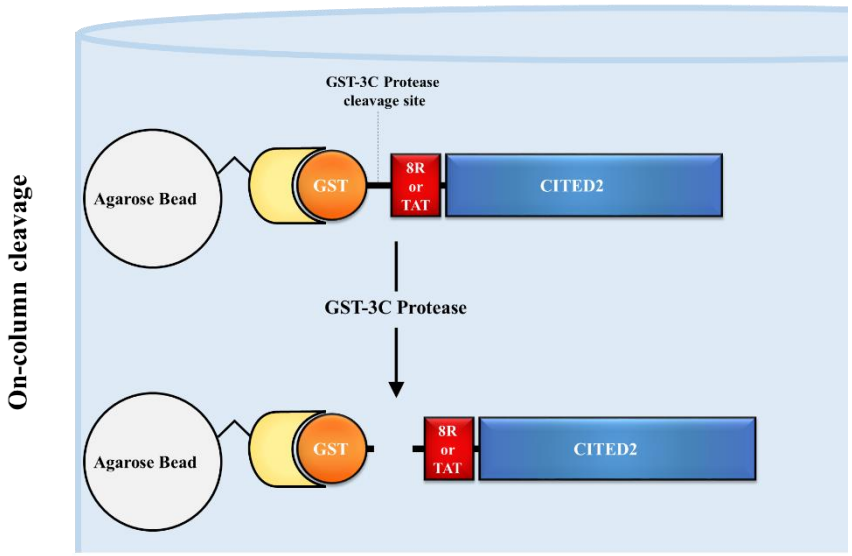


Figure IV.11 - Flow chart of the affinity purification procedure and GST-3C Protease cleavage of GST fusion proteins.

Before addition of GST-3C protease, the GST-fusion protein of interest was equilibrated in the protease cleavage buffer. For this purpose, we washed the column with 10 column volumes of GST-3C cleavage buffer before loading the GST-3C protease mix onto the column and sealing both ends of the column. We let the protein of interest in contact with the protease O.N. at 4º C.

Then, the protein of interest (TAT-CITED2 or 8R-CITED2) was eluted from the column when the flow was restored and the eluates were collected in fractions of 2 mL. The elution was monitored in the chromatogram of the absorbance levels. To remove the GST moiety of the fusion protein and the GST-3C protease, we applied a competitive solution containing reduced GSH molecules. The addition of this Elution Buffer to the column can be observed by the dotted yellow line in the chromatogram (**Figure IV.10**). The output of these molecules can be observed again in the chromatogram by the increase in the absorbance.

Samples from the injected solution, flow-through and the fractions of interest were analyzed by SDS-PAGE (**Figure IV.12** and **IV.13**) and Western Blotting (**Figure IV.14**).

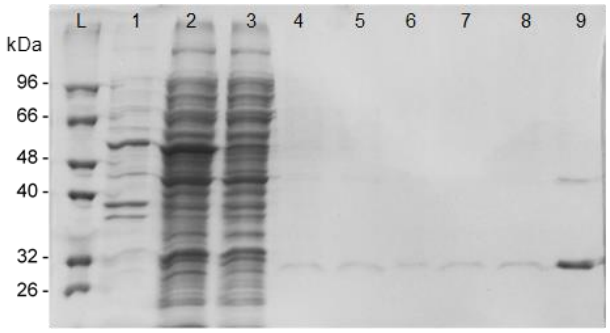


Figure IV.12 - SDS-PAGE gel stained with Coomassie Blue to monitor the purification process of TAT-CITED2. Lane 1 represents the insoluble fraction of protein extract, lane 2 represents the soluble fraction which was injected into the column and lane 3 corresponds to flow-through from the column. Purified protein can be observed in lane 4 to 8, while the GST tag and GST-3C protease are represented in lane 9. L corresponds to the ladder used: Precision Plus Protein™ Unstained Standard (Bio Rad).

The SDS-PAGE gel from the purification of TAT-CITED2 revealed that this protein was partially lost in the insoluble fraction since the band is also strong in that fraction (**Figure IV.12**). Following the purification process, we can verify that the protein was successfully bond to the GST beads, since in the lane 3 the protein is significantly less detected. After digestion with GST-3C protease, protein was collected in 2 mL fractions (lane 4-8) and the band observed is around 32 kDa, which is what we expected (57 kDa recombinant protein – 25 kDa GST tag = 32 kDa). In the last lane we can observe the eluted GST-tag and the GST-3C protease (46 kDa).

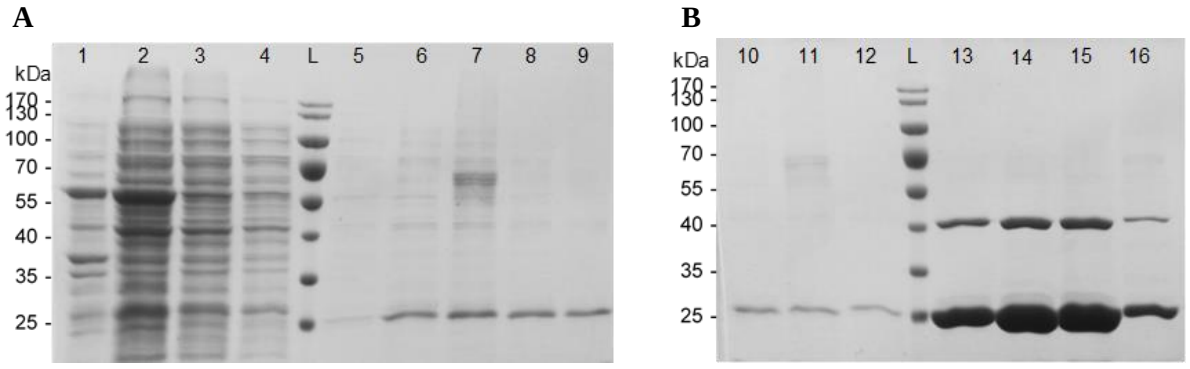


Figure IV.13 – SDS-Page gel stained with Coomassie Blue to monitor the purification process of 8R-CITED2. Lane 1 represents the insoluble fraction of protein extract, lane 2 represents the soluble fraction which was injected into the column and lane 3 corresponds to flow-through from the column. Lanes 4 and 5 represent the washings steps performed before the application of GST-3C protease. Purified protein can be observed in lane 6 to 12, while the GST tag and GST-3C protease are represented in lanes 13 to 16. L corresponds to the ladder used: PageRuler™ Prestained Protein Ladder (Fermentas Life Sciences).

The 8R-CITED2 purification was also relatively successful. Indeed, in the SDS-PAGE (**Figure IV.13**), we distinguish a stronger band in lanes 1 and 2, slightly above the 55 kDa, which corresponds to the molecular weight of the protein of interest. However, like TAT-CITED2, a fraction of the 8R-CITED2 protein was lost in the process of separation of soluble and insoluble fractions in protein extraction.

The recombinant protein was successfully isolated from the remaining proteins as we can see in the lanes 3 to 5, since there was a substantial decrease in the intensity of the interest band. After the digestion step, 2 mL fractions were collected (lane 6-12), and the band obtained was located between 25 and 35 kDa, corresponding the expected molecular weight (56 kDa recombinant protein – 25 kDa GST tag = 31 kDa). The protein was not completely purified as indicated by the presence of other bands in lanes 6 and 7. These fractions were not used in further studies.

In the last lanes (13-16) we observed the eluted GST-tag and the GST-3C protease (46 kDa).

After the isolation and purification of the recombinant proteins TAT-CITED2 and 8R-CITED2, we confirmed their identity by Western Blotting using a specific antibody against CITED2.

Purified proteins were quantified using the Bradford method and 3 µg of protein extracts were loaded and separated on SDS-PAGE gel. After the transfer of the proteins to a nitrocellulose membrane and blockage of this membrane with milk proteins, the mouse monoclonal JA22 antibody against CITED2 (ab5155, Abcam) was used as primary antibody at 1:2000 dilution, followed by the incubation with anti-mouse HRP.

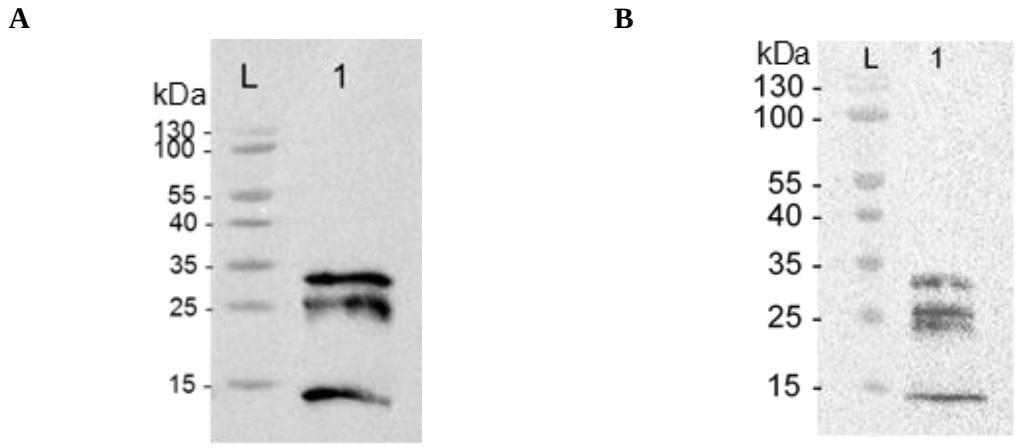


Figure IV.14 – Western Blots to confirm the identity of the purified proteins: (A) TAT-CITED2 and (B) 8R-CITED2. 1 represents the analysed protein and L corresponds to the ladder used: PageRuler^{TM} Prestained Protein Ladder (Fermentas Life Sciences).

Analyzing the Western Blots results in **Figure IV.14**, we observed three or more bands. A band with a lowest migrating ability located between 25 and 35 kDa corresponds to the expected size of the 8R-CITED2 and TAT-CITED2 protein. The lower molecular weight bands, in both cases, suggest that the purified protein suffered a degradation of some sort, which could be caused from poor handling of the samples prior to analysis, like shifts in the temperature. When the purification procedure was done, a sample of the fractions was kept at -20ºC while the fractions themselves were stored at -80ºC to avoid degradation. To run the gel we used the samples kept at -20ºC, which could explain the degradation presented in the Western Blots.

IV.3. CELLULAR ENTRANCE OF RECOMBINANT PROTEIN

To analyze the stability and test the ability of 8R-CITED2 protein to enter cells in culture, we supplemented MG5 mESC in culture with 20 µg/mL 8R-CITED2 or the equivalent volume of elution buffer (used as control). MG5 were used because they lack of the expression of endogeneous Cited2.[16] Therefore, if any signal was detected inside the cells by anti-CITED2 after 8R-CITED2 supplementation, it would indicate that the recombinant protein penetrated the cell membranes to enter the cells. MG5 cells were harvested after 4, 8, 12 and 24 h after supplementation of recombinant purified 8R-CITED2.

The presence of 8R-CITED2 inside MG5 mESC was analyzed in whole cell extracts by Western Blotting as described above (**Figure IV.15**) and further confirmed by immunostaining analysis (**Figure IV.16**).

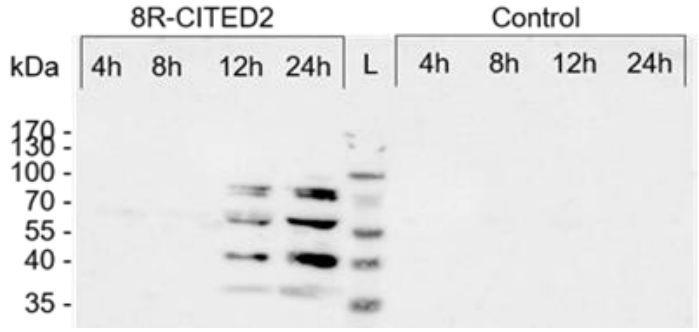


Figure IV.15 – Western Blot to evaluate the penetration of the recombinant protein 8R-CITED2. Cellular uptake of the recombinant protein was detectable in cell lysates of 12 h and 24 h. L corresponds to the ladder used: PageRuler™ Prestained Protein Ladder (Fermentas Life Sciences).

Temporal analysis showed that the 8R-CITED2 protein was detectable in cells only after 12 h of being added to the culture media. The detection was stronger in the extracts harvested at 24 h.

However, extra several bands with higher molecular sizes were detected, suggesting that the recombinant 8R-CITED2 protein has the tendency to form aggregates in the cell, like we observed *in vitro* during the purification, or the protein undergoes some modifications, such as phosphorylation or ubiquitination, as it has been reported.[58,59] These aggregates formation may impair the optimal capacity of the recombinant 8R-CITED2 to translocate into the cells.

IV.4. FLUORESCENCE IMMUNOCYTOCHEMISTRY

Cellular entrance of recombinant 8R-CITED2 was further confirmed by immunostaining analysis of MG5 cells in culture treated with 20 µg/mL 8R-CITED2 protein or elution buffer as control O.N. at 37ºC, since the protein would need at least 12 h to enter the cells (as indicated by the results obtained in the Western Blot).

Before fixation, cells were carefully washed to ensure the removal of all 8R-CITED2 proteins that were loosely bound to the cell surfaces. Then, cells were permeabilized, blocked

with BSA and finally incubated with anti-CITED2 antibody as a primary, and AlexaFluor-488 antibody as secondary antibody. The nuclei were stained with DAPI.

As expected no or a residual fluorescence was detected in the control, since MG5 cells lack Cited2 expression, while a fluorescent signal was detected in the cells transduced with the recombinant protein (**Figure IV.16**) .

Overall, these experiments indicated that 8R-CITED2 enters mESC and translocate into the nucleus. However, the conditions need to be optimized to avoid the formation of 8R-CITED2 aggregates which might decrease the active fraction of 8R-proteins in the cells.

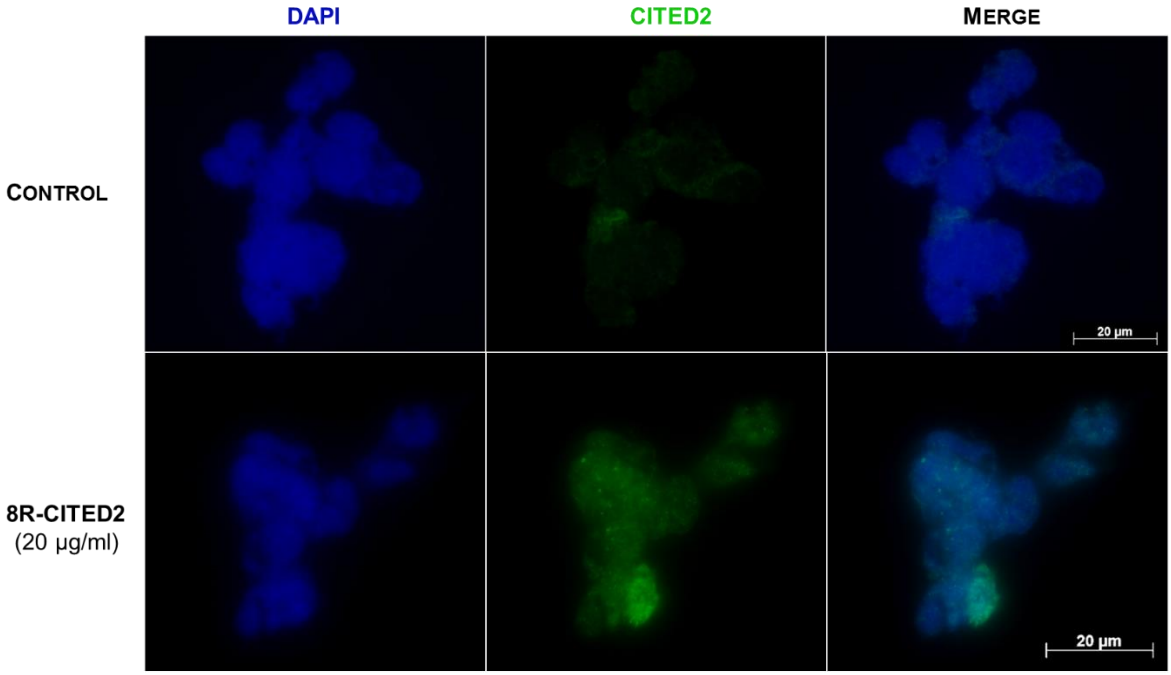


Figure IV.16 – Fluorescence microscopy of MG5 cells treated with 20 µg/mL of 8R-CITED2 detected by immunocytochemical reaction against CITED2, 24 h after supplementation, at 100X magnification. Negative controls were used to set up exposure conditions for detection of a specific signal.

IV.5. COMPETITION ESSAY BETWEEN 8R-CITED2 AND ENDOGENOUS CITED2 FOR P300 INTERACTION

CITED2 has a very high affinity for the CH1 domain of p300.[31,33] We have previously set an *ex vivo* assay to monitor CITED2-CH1 interaction based on BIFC. For this assay, cells

are transfected with vectors expressing either the CH1 domain of p300 fused to the N-terminal domain of the VENUS fluorescent protein (VEN-CH1) or CITED2 fused to the C-terminal domain of VENUS (VEC-CITED2), as described previously.[58] The interaction between CITED2 and the CH1 domain will bring the VENUS domains in close proximity, which will allow the emission of fluorescence. To test the functionality of 8R-CITED2 inside the cells, we evaluated the capacity of 8R-CITED2 protein to interfere/disrupt the endogenous interaction between VEC-CITED2 and the VEN-CH1 domain.

The 8R-CITED2 was added to the culture media of E14/T ESC, at two different concentrations (20 and 40 µg/mL), expressing VEC-CITED2 for VEN-CH1 and the detection of fluorescence in these conditions was compared to the control cells (E14/T with no recombinant protein) (**Figure IV.17**). The supplementation of 8R-CITED2 recombinant protein led to a decrease of the fluorescence detected in a dose-dependent manner, indicating that 8R-CITED2 competes with VEC-CITED2 for the interaction with VEN-CH1.

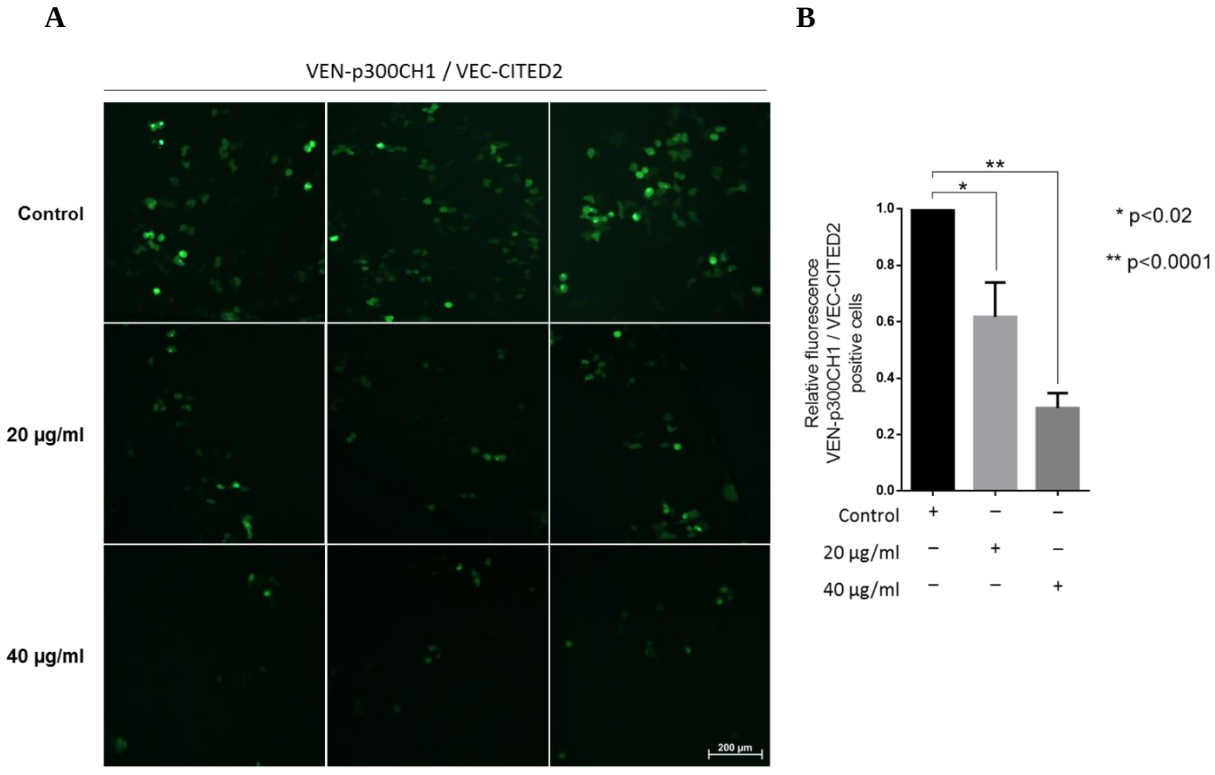


Figure IV.17 – Representative fields of E14/T ESC co-transfected cells with plasmids expressing VEN-CH1 and VEC-CITED2 and quantification of the fluorescence detected with different concentrations of 8R-CITED2. The detection of fluorescence was visualized 48 h after transfection (A). The number of fluorescent cells is presented relative to the fluorescence detected in the control condition (VEN-CH1/VEC-CITED2) set to 1. All values are represented as Mean±SEM (n=5) followed by a Two-way ANOVA treatment (B).

The number of cells emitting fluorescence was quantified using ImageJ software. Since the number of cells plated were the same in each condition, we counted the number of fluorescent cells in the control well. (**Figure IV.17.**B).

Cell exposure to 20 µg/mL of 8R-CITED2 recombinant protein led to a fluorescence drop to approximately 0.6, while exposure to 40 µg/mL recombinant protein caused a greater decrease in fluorescence. That fact can be explained by the 8R-CITED2 entrance in the cells, which reduced the probability of the ligation of VEC-CITED2 and VEN-CH1, causing consequently a decrease in the interaction of both of the residues of Venus-YFP in transfected cells. Taking into account the presented results, 8R-CITED2 seems to be constitutively active, since it competed for p300/CH1 domain.

IV.6. 8R-CITED2 RESCUES CARDIAC DIFFERENTIATION DEFECTS CAUSED BY CITED2-KNOCKOUT

To induce cardiac differentiation and further study the functionality of the recombinant 8R-CITED2 protein, we assessed the capacity of the 8R-CITED2 protein to rescue the cardiac differentiation defects caused by the depletion of Cited2 in mESC. We used the hanging drop method to differentiate mESC $C2^{fl/fl}$[Cre] cells as described elsewhere[16]. At the beginning of differentiation (day 0), mESC are still in a pluripotent state, but during mESC differentiation, spontaneously contracting foci emerge and indicate that the cells have eventually differentiated into cardiomyocytes. These beating foci were counted from 8 to 10 days after the initiation of differentiation (see **Figure IV.18**).

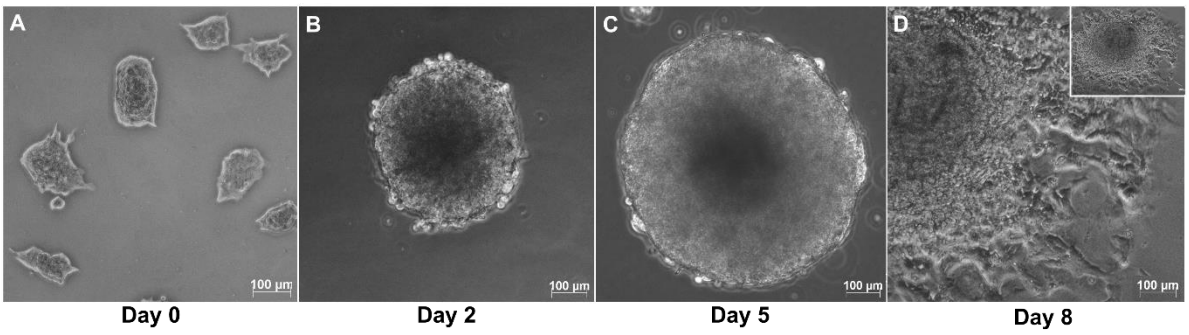


Figure IV.18 - Stages of mESC development. (A) Undifferentiated mESCs cultured in maintenance medium in the presence of LIF represented at Day 0. (B) Hanging drop culture from day 2 of differentiation. (C) Embryoid body at day 5 of differentiation in suspension culture. (D) Embryoid body outgrowth at day 8 of differentiation in gelatine coated plate.

To understand if 8R-CITED2 could compensate the lack of endogenous Cited2, the recombinant protein was added at different differentiation days (day 0 and day 2) and at a concentration between 5 and 10 µg/mL to $C2^{fl/fl}$[Cre], after Cited2 knockout was triggered at day 0 (treatment with 4HT). The process of emerging beating foci was monitored and compared to control cells (treated with ethanol).

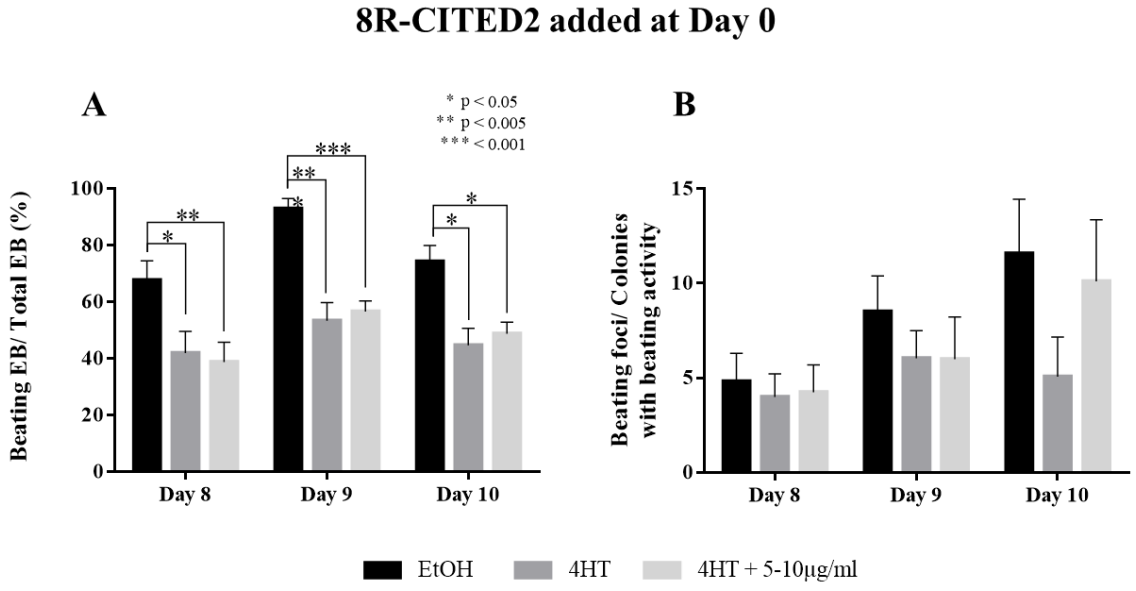


Figure IV.19 – Comparison between $C2^{fl/fl}$[Cre] treated with EtOH, 4HT and 4HT+8R-CITED2. (**A**) Percentage of colonies with beating foci and (**B**) average number of beating foci per beating EB, counted 8 to 10 days after the initiation of differentiation in cell cultures derived from $C2^{fl/fl}$[Cre] mESC treated with 5-10 µg/mL of 8R-CITED2 at Day 0. All values are represented as Mean±SEM (n=3) followed by a Two-way ANOVA treatment.

Analysing the graphs obtained (**Figure IV.19**), it was shown in both that there were no significant differences between the cells treated only with 4HT and 4HT supplemented with 8R-CITED2 at day 0, which means that the recombinant protein (at least in this range of concentrations) was not capable to rescue the ability to differentiate into cardiomyocytes.

These results can be explained by recent data obtained by the group (unpublished data) relative to the *Cited2* expression. Undifferentiated aggregates of mESC were subjected to differentiation in hanging drops and quantitative Reverse Transcriptase-Polymerase Chain Reaction (qRT-PCR) were performed, being the data gently given by João Santos (**Figure IV.20**).

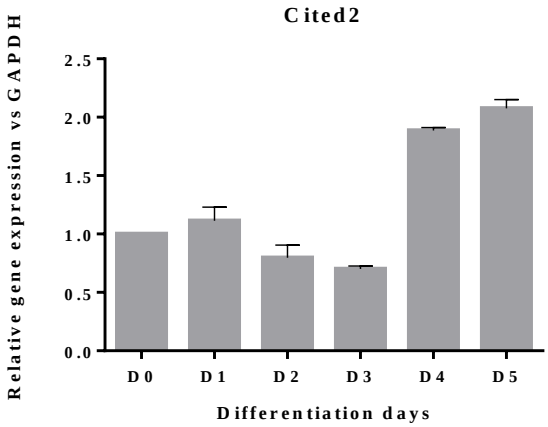


Figure IV.20 – Daily qRT-PCR analysis in aggregate differentiation of mESCs. Undifferentiated aggregates of mESCs were subjected to differentiation in hanging drops. Samples were collected since day 0 to day 5. Data show the maximum expression of Cited2 normalized by the reference gene GAPDH. All values are represented as Mean±SEM (n=3). Results given by João Santos.

Analyzing Cited2 RNA expression during the first few days of differentiation, we observed that the expression increased slightly at day 1, reached the lowest levels between day 2 and 3 of differentiation and increased again from day 3 to 5.

Cited2, previously associated with pluripotency, interacts with *Nanog* and *Oct4* promoting self-renewal and pluripotency maintenance[36,41]. Since the media was supplemented with 8R-CITED2 at day 0 of differentiation, it would not be unexpected that the protein could be promoting pluripotency instead of stimulating cardiogenic pathways. After the peak at day 1, we believe that the protein shifts towards a more cardiac interaction with another genes such as *Isl1*, therefore promoting cardiogenesis.

In fact, preliminary results of mESC supplemented at Day 0 with an augmented concentration (20 µg/mL), resulted in a great reduction in the number of beating EBs and with 40 µg/mL, that number diminished to zero. Those results leads to believe that increasing the amount of CITED2 protein in mESC cells contributes effectively to the maintenance of a pluripotency state as reported previously by Chen and co-workers.[36]

Taking this information into consideration, we believed that administering the protein when CITED2 expression raises again (day 2) could have a more cardiogenic effect instead of day 0.

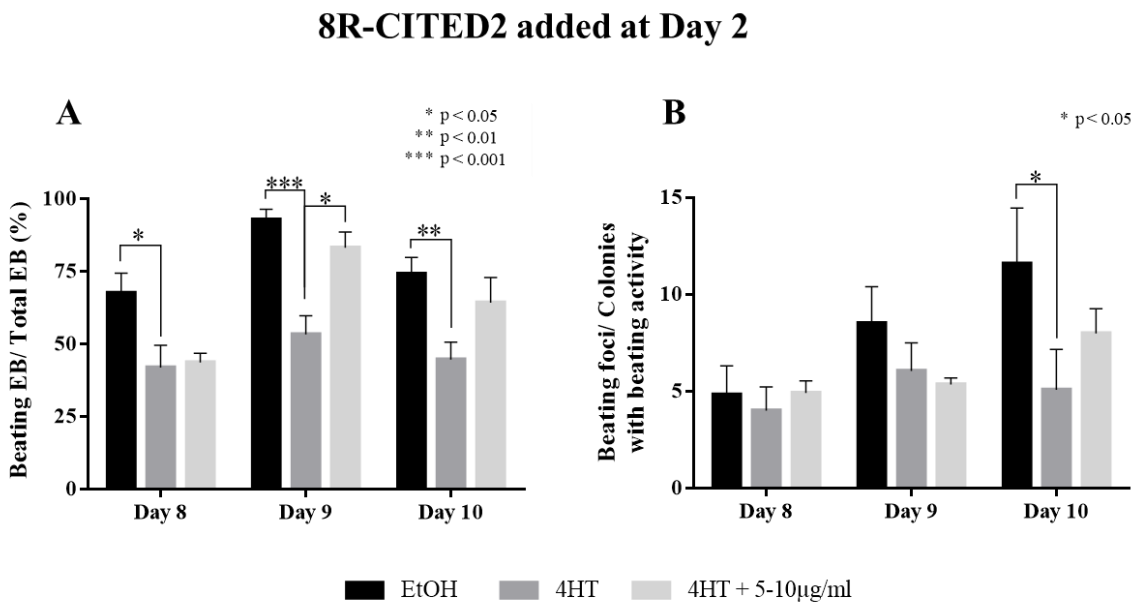


Figure IV.21 – Comparison between C2$^{fl/fl}$[Cre] treated with EtOH, 4HT and 4HT+8R-CITED2. (A) Percentage of colonies with beating foci and (B) average number of beating foci per beating EB, counted 8 to 10 days after the initiation of differentiation in cell cultures derived from C2$^{fl/fl}$[Cre] mESC treated with 5-10 µg/mL of 8R-CITED2 at Day 2. All values are represented as Mean±SEM (n=3) followed by a Two-way ANOVA treatment.

In **Figure IV.21**, it is possible to observe that the supplementation of 8R-CITED2 on Day 2 of differentiation at the same concentration in the culture medium of C2$^{fl/fl}$[Cre] ESC treated with 4HT, resulted in the increase of the number of beating EBs along the differentiation days comparing them to the C2$^{fl/fl}$[Cre] ESC treated with 4HT only. Regarding the average number beating foci in beating colonies, it is also noticeable, at Day 10, a slightly increase when the cells were supplemented with 8R-CITED2.

This means that 8R-CITED2 at 5-10 µg/mL concentration proved to be capable of restore the number of beating EB to levels comparable to those of control cells, as we expected since the supplementation was made at Day 2.

V. CONCLUSION

Cardiovascular disease is one of the main causes of death in the western world and novel therapeutic strategies need to be developed. One decade ago the first clinical trials addressed stem cell based therapies as a potential alternative therapeutic strategy for myocardial regeneration and repair.[21]

The differentiation of pluripotent stem cells toward cardiomyocytes is a dynamic process involving complex signaling networks, and it is still inefficient when compared with other cell fates.[60] Although ESC could be genetically modified to over-express cardiac-specific transcription factors and therefore improving the efficiency of their differentiation, these differentiated cells are not considered safe enough to be applied in cell therapy. For that reason, protein transduction has been demonstrated as an alternative approach towards increasing the differentiation efficiency unto cardiomyocytes.

The use of recombinant proteins associated with transcription factors can be transduced and in turn, stimulate over-expression of the genes that code for that protein. This way, they could trigger the specific pathway that is needed for differentiation toward a particular cell fate [25].

In this work we developed a novel recombinant protein that combined CITED2, a protein involved during two different times of mESC fate (first, it maintains ESC in a pluripotency state and later it is involved in differentiation toward cardiomyocytes), and a PTD (8R), a domain enriched with arginine residues, which is capable of overcoming the cellular membrane barrier.

Through different methods, it was shown that the purified 8R-CITED2, when applied into the studied cells, was successfully internalized, constitutively active and it presented a similar pattern to the endogenous protein when suppressed. While studying the role of this recombinant protein, it was possible to conclude that it can rescue the defects of *Cited2* knockout in mESC differentiation when applied to the media during the second day of differentiation, and it has some influence in the maintenance of pluripotency when media was supplemented at day 0.

Through this approach, the studied protein could be used to expand mESCs in a controlled environment, maintaining its pluripotent state, and to better control their

differentiation in future applications in cell therapy and drug discovery. This method offers better alternatives to previously used methods of transfection that usually resulted in artificial systems and suboptimal levels of CITED2. However, further research is necessary to additionally increase the efficiency and safety of this method.

VI. FUTURE PROSPECTS

The 8R-CITED2 revealed to be a promising molecule that fulfills the endogenous protein functions in mESCs. Due to limitation of time, it was impossible to explore every possibility of this recombinant protein.

In this work, the culture medium was supplemented at day 0 and day 2 of differentiation, but other days could also be considered, like for example the administration at day 1 and day 3, in order to identify the optimal rescue day. Nevertheless the influence at pluripotency and cardiac differentiation, we cannot exclude its importance in differention later stages (day 7 forwards).

As referred before, the overexpression of *CITED2* sustains self-renewal and proliferation in mESC cultured even in the absence of LIF[41], but the expression levels of this protein is hard to determine. Since it is possible to make a knockout of *Cited2* in C2fl/fl[Cre] mESC, another possibility is to investigate the concentration of protein necessary to maintain the cells in a pluripotent state.

The next steps could also be focused in finding possible interactions of 8R-CITED2 with other genes involved either in pluripotency or in cardiac differentiation. For example, using qRT-PCR we could evaluate whereas applying the protein at day 0 would promote pluripotency genes while adding at day 2 would promote an increase in differentiation genes. In the first case, we would investigate the markers usually involved in pluripotency, such as *Nanog*, *Sox2* and *Oct4*, in cells treated with different concentrations of the recombinant protein. This way, we could confirm if the entrance of the protein effectively contributes to the maintenance of pluripotency or even enhances it. On the other hand, our next step would be evaluate the expression of cardiac genes like *Gata4, α-MHC, troponin, Nkx2.5* and *Isl1*, when the cells were subjected to the treatment at day 2 of cardiac differentiation.

Then, the next logical step to the development of this work could be the application of the produced protein to hESC, in order to verify if the results are comparable to the obtained in mESC.

If the results were similar, we could extend our studies to animal models like Cited2$^{-/flox}$; Nkx2.5Cre *Mus musculus*[61] , that have a specific deletion of one of the alleles of *Cited2*. The phenotypical characteristics of this mouse model are: ventricular septal defects, layer of

ventricles and decrease in the number of capillaries to larger vessels. We could use the recombinant protein directly and indirectly to investigate if it could regenerate some of these malfunctions.

One promising approach would be to induce heart injuries in healthy mice and study the role of 8R-CITED2, which ideally could stimulate the cardiomyocyte differentiation in the injured muscle instead of scar tissue formation.

Taken together these approaches could contribute to evaluate the regeneration capacity of the heart cells, induced by the application of 8R-CITED2.

In addition to cardiac defects, *Cited2*-knockout embryos die in utero with, neural tube, liver, lung, eye, skeletal, thymus, gonadal, hematopoietic failures, among others, as well as left-right patterning and neural crest cells migration defects.[35,37–39] It would be interesting to investigate the potential of CITED2 overexpression in the survival and self-renewal of other tissue stem cells, activating or rejuvenating stem cells in damaged tissues or tissues undergoing degenerative processes.

Besides, CITED2 might be also useful against some cancer cells, since it has been reported as a factor with both oncogenic and tumour suppressive properties.[35,62] Since it has been reported that CITED2 in ESC can sustain their self-renewal, an irregular increase of CITED2 expression might contribute to uncontrolled self-renewal and proliferation of stem cells existent in tumour tissues and contribute to originate Cancer Stem Cells (CSC). Otherwise, in Non-Small Cell Lung Cancer (NSCLC), CITED2 was demonstrated to repress the expression of CSC markers in NSCLC-stem cells and enhance their sensitivity to ionizing radiations when in combination with butyrate treatment. So, this recombinant protein could be applied to these tissues to enhance the treatment to this type of cancer.[35,63]

VII. BIBLIOGRAPHY

1. Ford ES, Capewell S. Proportion of the decline in cardiovascular mortality disease due to prevention versus treatment: public health versus clinical care. *Annu. Rev. Public Health* 2011;32:5-22. doi:10.1146/annurev-publhealth-031210-101211.

2. Sahara M, Santoro F, Chien KR. Programming and reprogramming a human heart cell. 2015;34(6):710-738.

3. Mendis S, Puska P, Norrving B. Global atlas on cardiovascular disease prevention and control. *World Heal. Organ.* 2011:2-14.

4. Di Pasquale E, Song B, Condorelli G. Generation of human cardiomyocytes: a differentiation protocol from feeder-free human induced pluripotent stem cells. *J. Vis. Exp.* 2013;(76):1-6. doi:10.3791/50429.

5. Burridge PW, Keller G, Gold JD, Wu JC. Production of de novo cardiomyocytes: human pluripotent stem cell differentiation and direct reprogramming. *Cell Stem Cell* 2012;10(1):16-28. doi:10.1016/j.stem.2011.12.013.

6. Chow MZ, Boheler KR, Li RA. Human pluripotent stem cell-derived cardiomyocytes for heart regeneration , drug discovery and disease modeling : from the genetic , epigenetic , and tissue modeling perspectives. *Stem Cell Res. Ther.* 2013;4(4):1. doi:10.1186/scrt308.

7. Zhang Y, Mignone J, MacLellan WR. Cardiac Regeneration and Stem Cells. *Physiol. Rev.* 2015;95(4):1189-1204. doi:10.1152/physrev.00021.2014.

8. Meyer T, Sartipy P, Blind F, Leisgen C, Guenther E. New cell models and assays in cardiac safety profiling. *Expert Opin. Drug Metab. Toxicol.* 2007;3(4):507-517. doi:10.1517/17425225.3.4.507.

9. Jiang J, Han P, Zhang Q, Zhao J, Ma Y. Cardiac differentiation of human pluripotent stem cells. *J. Cell. Mol. Med.* 2012;16(8):1663-1668. doi:10.1111/j.1582-4934.2012.01528.x.

10. Masumoto H, Yamashita JK. Pluripotent Stem Cells for Cardiac Cell Therapy : The Application of Cell Sheet Technology. 2013.

11. Dimmeler S, Burchfield J, Zeiher AM. Cell-based therapy of myocardial infarction. *Arterioscler. Thromb. Vasc. Biol.* 2008;28(2):208-216. doi:10.1161/ATVBAHA.107.155317.

12. Rajala K, Pekkanen-Mattila M, Aalto-Setälä K. Cardiac differentiation of pluripotent stem cells. *Stem Cells Int.* 2011;2011(1):383709. doi:10.4061/2011/383709.

13. Chavakis E, Koyanagi M, Dimmeler S. Enhancing the outcome of cell therapy for cardiac repair: Progress from bench to bedside and back. *Circulation* 2010;121(2):325-335. doi:10.1161/CIRCULATIONAHA.109.901405.

14. Roobrouck VD, Ulloa-Montoya F, Verfaillie CM. Self-renewal and differentiation capacity of young and aged stem cells. *Exp. Cell Res.* 2008;314(9):1937-1944. doi:10.1016/j.yexcr.2008.03.006.

15. Yang F, Cho S-W, Son SM, et al. Genetic engineering of human stem cells for enhanced angiogenesis using biodegradable polymeric nanoparticles. *Proc. Natl. Acad. Sci. U. S. A.* 2010;107(8):3317-22. doi:10.1073/pnas.0905432106.

16. Kranc KR, Oliveira D V., Armesilla-Diaz A, et al. Acute Loss of Cited2 Impairs Nanog Expression and Decreases Self-Renewal of Mouse Embryonic Stem Cells. *Stem Cells* 2015;33(3):699-712. doi:10.1002/stem.1889.

17. Memon I a., Sawa Y, Fukushima N, et al. Repair of impaired myocardium by means of implantation of engineered autologous myoblast sheets. *J. Thorac. Cardiovasc. Surg.* 2005;130(5):1333-1341. doi:10.1016/j.jtcvs.2005.07.023.

18. Sasaki D, Shimizu T, Masuda S, et al. Mass preparation of size-controlled mouse embryonic stem cell aggregates and induction of cardiac differentiation by cell patterning method. *Biomaterials* 2009;30(26):4384-4389. doi:10.1016/j.biomaterials.2009.05.003.

19. Kurosawa H. Methods for inducing embryoid body formation: in vitro differentiation system of embryonic stem cells. *J. Biosci. Bioeng.* 2007;103(5):389-398. doi:10.1263/jbb.103.389.

20. Li Y, Zhang M, Wen H, et al. Comparing the diagnostic values of circulating microRNAs and cardiac troponin T in patients with acute myocardial infarction. *Clinics* 2013;68(1):75-80. doi:10.6061/clinics/2013(01)OA12.

21. Doppler S a., Deutsch MA, Lange R, Krane M. Cardiac regeneration: Current therapies-future concepts. *J. Thorac. Dis.* 2013;5(5):683-697. doi:10.3978/j.issn.2072-1439.2013.08.71.

22. Jayawardena TM, Egemnazarov B, Finch EA, et al. MicroRNA-Mediated In Vitro and

In Vivo Direct Reprogramming of Cardiac Fibroblasts to Cardiomyocytes. *Circ. Res.* 2012;110(11):1465-1473. doi:10.1161/CIRCRESAHA.112.269035.

23. Nam Y, Song K, Luo X, et al. Reprogramming of human fi broblasts toward a cardiac fate. *Proc. Natl. Acad. Sci.* 2013;110(14):5588-5593.

24. Desbaillets I, Ziegler U, Groscurth P, Gassmann M. Embryoid bodies: an in vitro model of mouse embryogenesis. *Exp. Physiol.* 2000;85(6):645-651.

25. Fonoudi H, Yeganeh M, Fattahi F, et al. ISL1 protein transduction promotes cardiomyocyte differentiation from human embryonic stem cells. Milstone DS, ed. *PLoS One* 2013;8(1):e55577. doi:10.1371/journal.pone.0055577.

26. Buckingham M, Meilhac S, Zaffran S. Building the mammalian heart from two sources of myocardial cells. *Nat. Rev. Genet.* 2005;6(11):826-835. doi:10.1038/nrg1710.

27. Noseda M, Peterkin T, Simões FC, Patient R, Schneider MD. Cardiopoietic factors extracellular signals for cardiac lineage commitment. *Circ. Res.* 2011;108(1):129-152. doi:10.1161/CIRCRESAHA.110.223792.

28. Vidarsson H, Hyllner J, Sartipy P. Differentiation of human embryonic stem cells to cardiomyocytes for in vitro and in vivo applications. *Stem Cell Rev. Reports* 2010;6(1):108-120. doi:10.1007/s12015-010-9113-x.

29. Itskovitz-Eldor J, Schuldiner M, Karsenti D, et al. Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. *Mol. Med.* 2000;6(2):88-95. doi:10859025.

30. Ieda M, Fu J-D, Delgado-Olguin P, et al. Direct Reprogramming of Fibroblasts into Functional Cardiomyocytes by Defined Factors. *Cell* 2010;142(3):375-386. doi:10.1016/j.cell.2010.07.002.

31. Bhattacharya S, Michels CL, Leung MK, Arany ZP, Kung AL, Livingston DM. Functional role of p35srj, a novel p300/CBP binding protein, during transactivation by HIF-1. *Genes Dev.* 1999;13(1):64-75. doi:10.1101/gad.13.1.64.

32. Yin Z, Haynie J, Yang X, et al. The essential role of Cited2, a negative regulator for HIF-1alpha, in heart development and neurulation. *Proc. Natl. Acad. Sci. U. S. A.* 2002;99(16):10488-93. doi:10.1073/pnas.162371799.

33. Freedman SJ, Sun Z-YJ, Kung AL, France DS, Wagner G, Eck MJ. Structural basis for negative regulation of hypoxia-inducible factor-1alpha by CITED2. *Nat. Struct. Biol.*

2003;10(7):504-512. doi:10.1038/nsb936.

34. Bamforth SD, Bragança J, Farthing CR, et al. Cited2 controls left-right patterning and heart development through a Nodal-Pitx2c pathway. *Nat. Genet.* 2004;36(11):1189-96. doi:10.1038/ng1446.

35. Bragança J, Santos J. Is CITED2 a general regulator of stem cell functions? *Stem Cell Transl. Investig.* 2015;(APRIL). doi:10.14800/scti.755.

36. Chen CM, Bentham J, Cosgrove C, et al. Functional Significance of SRJ Domain Mutations in CITED2. *PLoS One* 2012;7(10). doi:10.1371/journal.pone.0046256.

37. Combes AN, Spiller CM, Harley VR, et al. Gonadal defects in Cited2-mutant mice indicate a role for SF1 in both testis and ovary differentiation. *Int. J. Dev. Biol.* 2010;54(June 2009):683-689. doi:10.1387/ijdb.092920ac.

38. Chen Y, Haviernik P, Bunting KD, Yang Y. Cited2 is required for normal hematopoiesis in the murine fetal liver Cited2 is required for normal hematopoiesis in the murine fetal liver. 2009;110(8):2889-2898. doi:10.1182/blood-2007-01-066316.

39. Bamforth SD, Bragança J, Eloranta JJ, et al. Cardiac malformations, adrenal agenesis, neural crest defects and exencephaly in mice lacking Cited2, a new Tfap2 co-activator. *Nat. Genet.* 2001;29(4):469-474. doi:10.1038/ng768.

40. Li Q, Ramírez-Bergeron DL, Dunwoodie SL, Yang YC. Cited2 gene controls pluripotency and cardiomyocyte differentiation of murine embryonic stem cells through Oct4 gene. *J. Biol. Chem.* 2012;287(34):29088-29100. doi:10.1074/jbc.M112.378034.

41. Pritsker M, Ford NR, Jenq HT, Lemischka IR. Genomewide gain-of-function genetic screen identifies functionally active genes in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 2006;103(18):6946-6951. doi:10.1073/pnas.0509861103.

42. Bai S, Kerppola TK. Opposing roles of FoxP1 and Nfat3 in transcriptional control of cardiomyocyte hypertrophy. *Mol. Cell. Biol.* 2011;31(14):3068-3080. doi:10.1128/MCB.00925-10.

43. El-Sayed A, Futaki S, Harashima H. Delivery of macromolecules using arginine-rich cell-penetrating peptides: ways to overcome endosomal entrapment. *AAPS J.* 2009;11(1):13-22. doi:10.1208/s12248-008-9071-2.

44. Prochiantz A. Protein transduction, from technology to physiology. *J. Soc. Biol.*

2006;200(3):209-212.

45. Wadia JS, Dowdy SF. Protein transduction technology. *Curr. Opin. Biotechnol.* 2002;13(1):52-6. doi:10.1016/S0958-1669(02)00284-7.

46. Lisziewicz J, Lőrincz O. HIV-specific immunotherapy with DermaVir, the first pDNA/PEIm pathogen-like nanomedicine. *Eur. J. Nanomedicine* 2012;4(2-4):81-87. doi:10.1515/ejnm-2012-0011.

47. Zhang H, Ma Y, Gu J, et al. Reprogramming of somatic cells via TAT-mediated protein transduction of recombinant factors. *Biomaterials* 2012;33(20):5047-55. doi:10.1016/j.biomaterials.2012.03.061.

48. Nagahara H, Vocero-Akbani a M, Snyder EL, et al. Transduction of full-length TAT fusion proteins into mammalian cells: TAT-p27Kip1 induces cell migration. *Nat. Med.* 1998;4(12):1449-1452. doi:10.1038/4042.

49. Schwarze SR, Hruska KA, Dowdy SF. Protein transduction: unrestricted delivery into all cells? *Trends Cell Biol.* 2000;10(7):290-295. doi:10.1016/S0962-8924(00)01771-2.

50. Wadia JS, Stan R V, Dowdy SF. Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat. Med.* 2004;10(3):310-315. doi:10.1038/nm996.

51. Walrant A, Vogel A, Correia I, et al. Membrane interactions of two arginine-rich peptides with different cell internalization capacities. *Biochim. Biophys. Acta - Biomembr.* 2012;1818(7):1755-1763. doi:10.1016/j.bbamem.2012.02.024.

52. Nakase I, Niwa M, Takeuchi T, et al. Cellular uptake of arginine-rich peptides: Roles for macropinocytosis and actin rearrangement. *Mol. Ther.* 2004;10(6):1011-1022. doi:10.1016/j.ymthe.2004.08.010.

53. Zhou H, Wu S, Joo JY, et al. Generation of Induced Pluripotent Stem Cells Using Recombinant Proteins. *Cell Stem Cell* 2009;4(5):381-384. doi:10.1016/j.stem.2009.04.005.

54. Yang Y, Liu B, Dong J, Zhang L, Pang M, Rong L. Proteins reprogramming: present and future. *ScientificWorldJournal.* 2012;2012:453185. doi:10.1100/2012/453185.

55. Kim D, Kim C-H, Moon J-I, et al. Generation of Human Induced Pluripotent Stem Cells by Direct Delivery of Reprogramming Proteins. *Cell Stem Cell* 2009;4(6):472-476. doi:10.1016/j.stem.2009.05.005.

56. Hooper M, Hardy K, Handyside a, Hunter S, Monk M. HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* 1987;326(6110):292-295. doi:10.1038/326292a0.

57. Pastva A, Estell K, Schoeb TR, Atkinson TP, Lisa M. Bimolecular Fluorescence Complementation (BiFC) Analysis as a probe of protein interactions in living cells. 2010;172(7):4520-4526. doi:10.1146/annurev.biophys.37.032807.125842.BIMOLECULAR.

58. Machado-Oliveira G, Guerreiro E, Matias AC, Facucho-Oliveira J, Pacheco-Leyva I, Bragança J. FBXL5 modulates HIF-1α transcriptional activity by degradation of CITED2. *Arch. Biochem. Biophys.* 2015;576:61-72. doi:10.1016/j.abb.2015.04.012.

59. Chen Y, Neve RL, Liu H. Neddylation dysfunction in Alzheimer's disease. *J. Cell. Mol. Med.* 2012;16(11):2583-2591. doi:10.1111/j.1582-4934.2012.01604.x.

60. Willems E, Bushway PJ, Mercola M. Natural and Synthetic Regulators of Embryonic Stem Cell Cardiogenesis. *Pediatr. Cardiol.* 2009;30(5):635-642. doi:10.1007/s00246-009-9409-2.

61. Qu W, Qin S, Ren S, Jiang X, Zhuo R, Zhang X. Peptide-Based Vector of VEGF Plasmid for E ffi cient Gene Delivery in Vitro and Vessel Formation in Vivo. 2013.

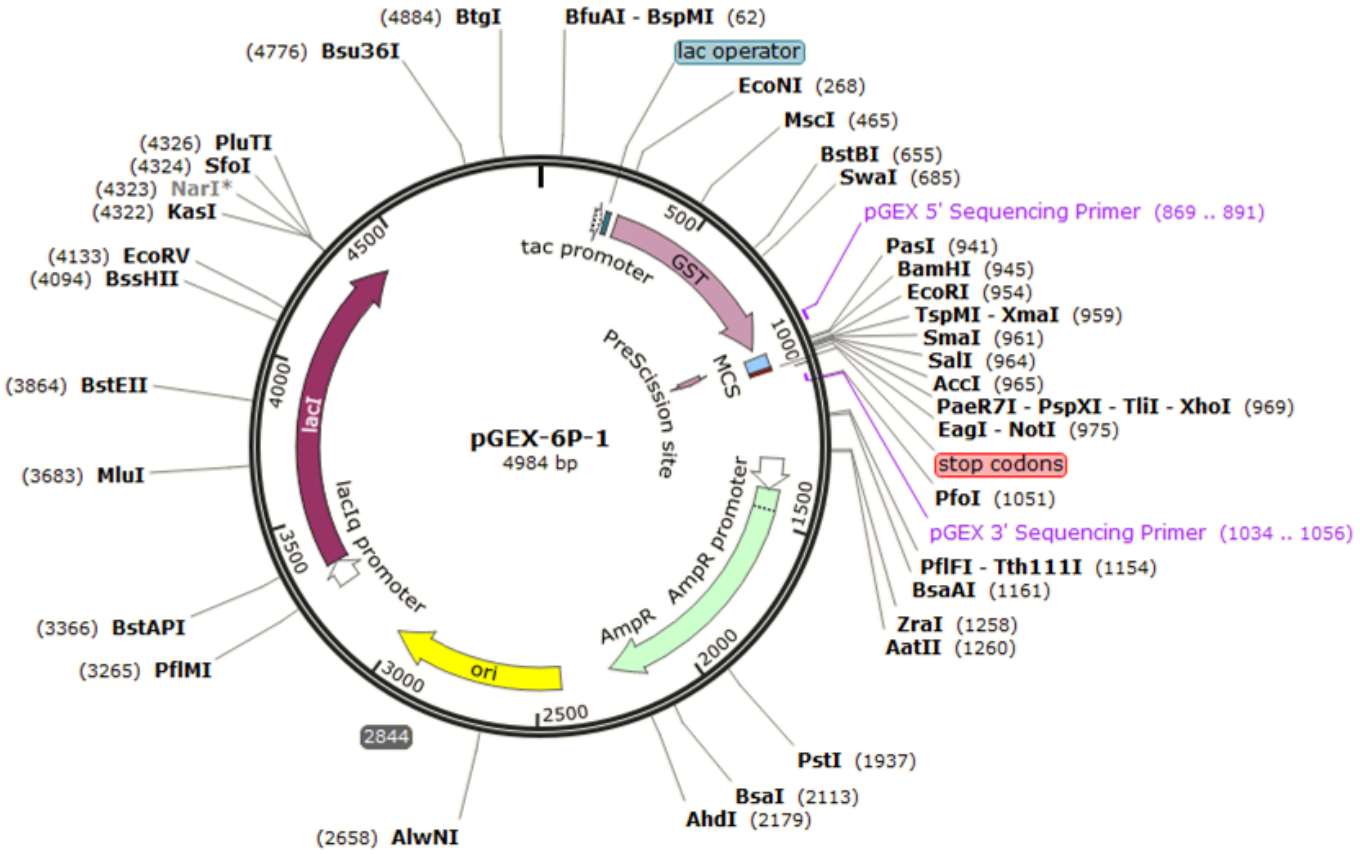
62. Du J, Yang Y-C. Cited2 in hematopoietic stem cell function. *Curr. Opin. Hematol.* 2013;20:301-7. doi:10.1097/MOH.0b013e3283606022.

63. He Z, Leong DJ, Tang J, Majeska RJ, Kalnicki S, Sun HB. A Novel Regulatory Role of CITED2 in Non-Small Cell Lung Cancer (NSCLC) Stem Cells and in Modulating the Effect of HDACi Treatment on NSCLC. *Int. J. Radiat. Oncol.* 2013;87(2):S662. doi:10.1016/j.ijrobp.2013.06.1753.

VIII. APPENDICES

APPENDIX A

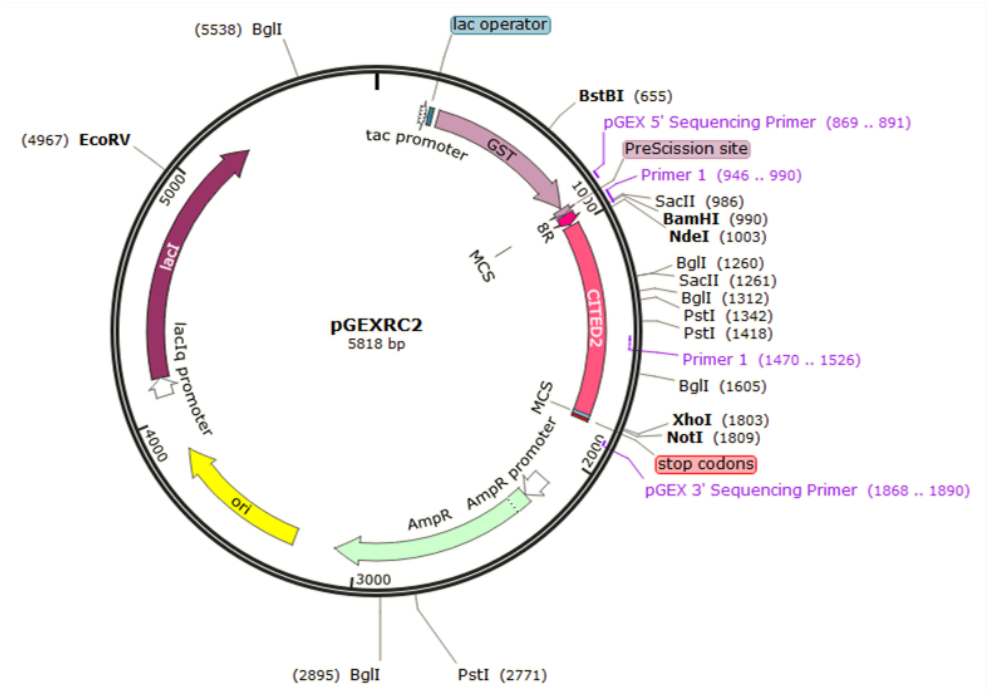
pGEX-6P-1 Plasmid Map



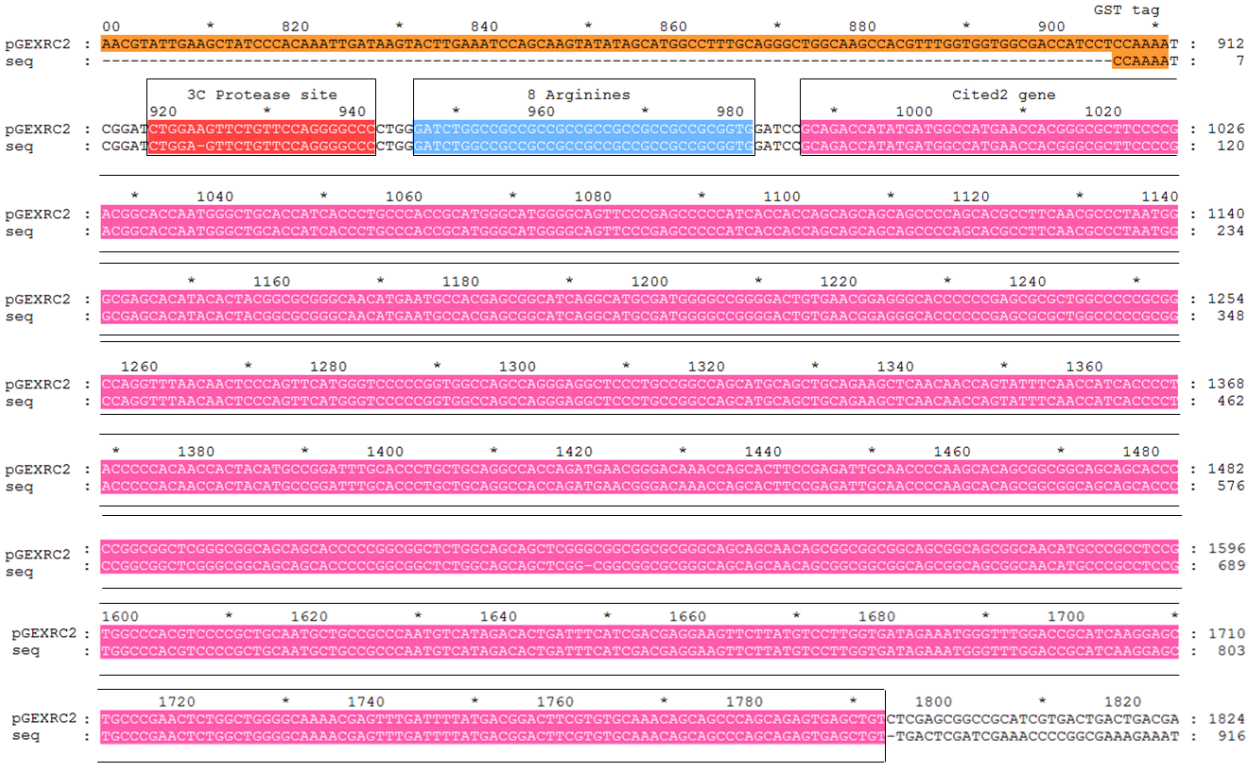
APPENDIX B

pGEXRC2 Plasmid Map

(pGEX-6P-1+8R+Cited2)



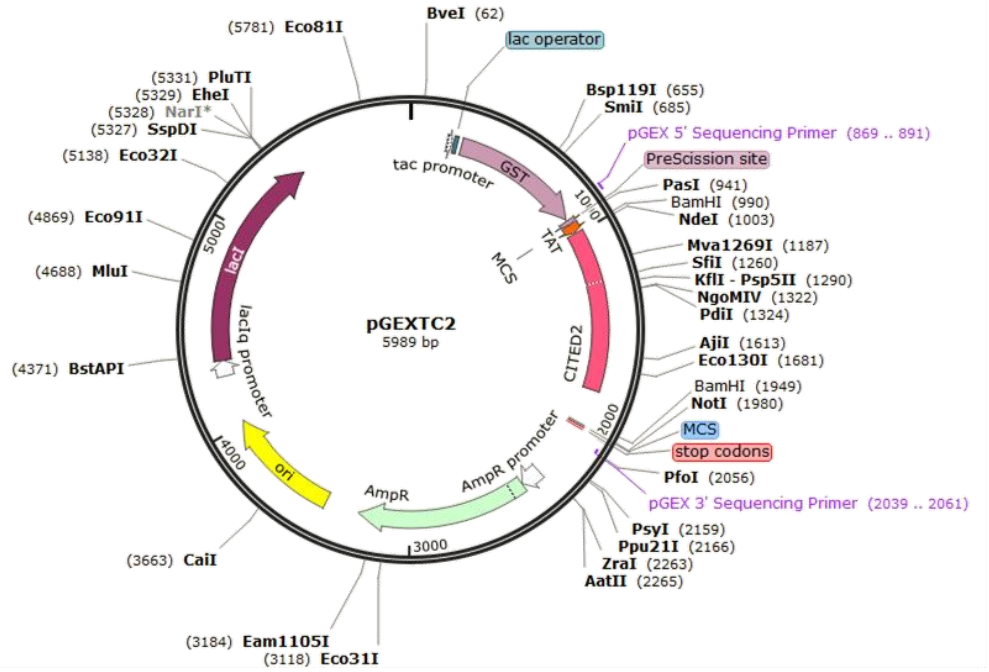
Sequencing Results:



APPENDIX C

pGEXTC2 Plasmid Map

(pGEX-6P-1+TAT+Cited2)



Sequencing Results: