UNIVERSIDADE DE LISBOA

FACULDADE DE MEDICINA VETERINÁRIA

STUDIES ON A YET UNASSIGNED FUNCTION OF AN AFRICAN SWINE FEVER VIRUS PROTEIN POTENTIALLY INVOLVED IN THE HOST CELL RESPONSE TO INFECTION

PATRÍCIA FERREIRA GOMES ESTANISLAU VIEIRA

ORIENTADOR(A):
Doutor José Alexandre da Costa Perdigão e Cameira Leitão
COORIENTADOR(A):
Doutora Sílvia Maurício Correia

2021

UNIVERSIDADE DE LISBOA

FACULDADE DE MEDICINA VETERINÁRIA

STUDIES ON A YET UNASSIGNED FUNCTION OF AN AFRICAN SWINE FEVER VIRUS PROTEIN POTENTIALLY INVOLVED IN THE HOST CELL RESPONSE TO INFECTION

PATRÍCIA FERREIRA GOMES ESTANISLAU VIEIRA

DISSERTAÇÃO DE MESTRADO INTEGRADO EM MEDICINA VETERINÁRIA

JÚRI
PRESIDENTE:
Doutor Luís Manuel Morgado Tavares

VOGAIS:
Doutor Fernando Jorge Silvano Boinas
Doutor José Alexandre da Costa Perdigão e Cameira Leitão

ORIENTADOR(A):
Doutor José Alexandre da Costa Perdigão e Cameira Leitão

COORIENTADOR(A):
Doutora Sílvia Maurício Correia

2021

DECLARAÇÃO RELATIVA ÀS CONDIÇÕES DE REPRODUÇÃO DA DISSERTAÇÃO

Nome: Patrícia Ferreira Gomes Estanislau Vieira

Título da Tese ou Dissertação: Studies on a yet unassigned function o fan african swine fever virus protein potentially involved in the host cell response to infection

Ano de conclusão (indicar o da data da realização das provas públicas): 2021

Designação do curso de Mestrado ou de Doutoramento: Mestrado Integrado em Medicina Veterinária

Área científica em que melhor se enquadra (assinale uma):

☐ Clínica ☐ Produção Animal e Segurança Alimentar

☐ Morfologia e Função ☒ Sanidade Animal

Declaro sobre compromisso de honra que a tese ou dissertação agora entregue corresponde à que foi aprovada pelo júri constituído pela Faculdade de Medicina Veterinária da ULISBOA.

Declaro que concedo à Faculdade de Medicina Veterinária e aos seus agentes uma licença não-exclusiva para arquivar e tornar acessível, nomeadamente através do seu repositório institucional, nas condições abaixo indicadas, a minha tese ou dissertação, no todo ou em parte, em suporte digital.

Declaro que autorizo a Faculdade de Medicina Veterinária a arquivar mais de uma cópia da tese ou dissertação e a, sem alterar o seu conteúdo, converter o documento entregue, para qualquer formato de ficheiro, meio ou suporte, para efeitos de preservação e acesso.

Retenho todos os direitos de autor relativos à tese ou dissertação, e o direito de a usar em trabalhos futuros (como artigos ou livros).

Concordo que a minha tese ou dissertação seja colocada no repositório da Faculdade de Medicina Veterinária com o seguinte estatuto (assinale um):

1. ☐ Disponibilização imediata do conjunto do trabalho para acesso mundial;
2. ☒ Disponibilização do conjunto do trabalho para acesso exclusivo na Faculdade de Medicina Veterinária durante o período de ☐ 6 meses, ☒ 12 meses, sendo que após o tempo assinalado autorizo o acesso mundial*;

* Indique o motivo do embargo (OBRIGATÓRIO)

Esta dissertação contém resultados sensíveis, pertencentes a estudos que continuam em curso, e que ainda não foram publicados.

Nos exemplares das dissertações de mestrado ou teses de doutoramento entregues para a prestação de provas na Universidade e dos quais é obrigatoriamente enviado um exemplar para depósito na Biblioteca da Faculdade de Medicina Veterinária da Universidade de Lisboa deve constar uma das seguintes declarações (incluir apenas uma das três):

1. É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE/TRABALHO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Faculdade de Medicina Veterinária da Universidade de Lisboa, 26 de Outubro de 2021

Assinatura: 

Acknowledgements

Desde o dia em que pisei pela primeira vez a Faculdade de Medicina Veterinária até hoje, muitas foram as pessoas que, de uma forma ou de outra, contribuíram para que este dia, o dia em que me torno médica veterinária, fosse possível.

Antes de mais, quero agradecer ao meu orientador, Doutor Alexandre Leitão, pela oportunidade de desenvolver a minha dissertação nesta área. Obrigada pela confiança. Obrigada por todos os ensinamentos, pela paciência e por ter sempre incentivado o meu espírito crítico.

À minha coorientadora, Doutora Sílvia Correia, muito obrigada por todos os materiais que me permitiram desenvolver este projeto. Mas principalmente obrigada pela enorme disponibilidade, pela preocupação e por todos os conselhos.

Tenho também de agradecer a todos os elementos do laboratório de doenças infecciosas. À Andreia, minha companheira de todos os dias no laboratório, por toda a ajuda e pela paciência, e também por todos os momentos de companheirismo e boa disposição. À Claudia, que tanto me ensinou, incentivou e orientou, muito obrigada por todo o apoio científico e técnico, e por toda amizade. Ao Nuno e à Catarina, obrigada pela boa disposição e bom ambiente de trabalho. E ao professor Fernando Boinas, por me ter proporcionado a possibilidade de desenvolver este trabalho nas instalações do LDI.

Obrigada ao Doutor Pedro Bule, não só pelo material disponibilizado, mas também pelos úteis conselhos e discussões. Obrigada também aos elementos do seu laboratório, Catarina e Marlene. Um agradecimento especial à Sara Zuquete, pela inesgotável disponibilidade.

Obrigada aos meus colegas da FMV. Em especial, um obrigada gigante à Bastos, que já me acompanhava quando este curso era apenas um sonho, e que continuou ao meu lado até ao último dia. À Lúcia, uma das primeiras amigas que fiz na faculdade, e que se tornou uma presença tão importante no final. Obrigada pelas palavras sábias. À Margarida, que chegou mais tarde, mas que parece que conheço desde o primeiro dia. E um agradecimento especial à Raquel, porque se não fosse ela esta dissertação não seria possível. Obrigada por todo o apoio, todas as conversas; obrigada pela amizade e por saberes sempre o que dizer, e como dizer. Obrigada amiga!

Muito obrigada à Inês, minha amiga desde sempre, por estares sempre lá para mim.

Não posso deixar de agradecer às minhas amigas da dança, que transformaram os finais de dia em momentos felizes. Um obrigada especial à Diana, a minha maior inspiração.

Muito obrigada ao Miguel. Não há forma de agradecer tudo o que fizeste por mim ao longo destes anos, e particularmente ao longo deste último ano. Obrigada por todas as conversas infindáveis, por todos os conselhos, por todas as palavras de motivação e por me

ajudares sempre a seguir o caminho certo. Obrigada pela inesgotável compreensão. Nada disto seria possível sem ti.

Obrigada à minha família, que sempre esteve presente e sempre me fez sentir tão valorizada.

Por fim, obrigada à minha mãe e ao meu pai. Obrigada por me terem sempre proporcionado a oportunidade de seguir os meus sonhos. Obrigada por me apoiarem incondicionalmente. Obrigada por todas as palavras sábias e pela forma como me educaram. Obrigada por todo o amor. Devo-vos a vocês a pessoa que sou hoje.

ESTUDO DE UMA PROTEÍNA DO VÍRUS DA PESTE SUÍNA AFRICANA POTENCIALMENTE ENVOLVIDA NA RESPOSTA DO HOSPEDEIRO À INFEÇÃO

Resumo

O sistema imunitário inato é a primeira linha de defesa aquando de uma infeção viral e consiste maioritariamente numa resposta de citocinas inflamatórias, em particular de interferões. Este sistema exerce uma pressão seletiva, o que levou à evolução de múltiplos mecanismos de evasão por parte dos agentes patogénicos invasores, conferindo-lhes vantagem. Vários genes virais foram desenvolvidos para manipular o sistema imunitário do hospedeiro, funcionando como "ferramentas prontas a usar" que podem ser exploradas por nós para a modulação do sistema imunitário e para o desenvolvimento de vacinas atenuadas.

A peste suína africana constitui uma ameaça para a suinicultura mundial que não tem, nos dias que correm, vacina nem tratamento disponíveis. O desenvolvimento de uma vacina atenuada através da deleção de genes virais responsáveis pela evasão ao sistema imunitário inato do hospedeiro poderia ser uma solução prática para o controlo desta doença.

Recentemente, um estudo bioinformático do genoma do vírus da peste suína africana (VPSA) identificou uma ORF com uma região de homologia com o domínio SH2. O domínio SH2 é conhecido por desempenhar um papel importante em diversas cascatas celulares de transdução de sinal. Para além disso, o domínio SH2 é uma das estruturas críticas das proteínas SOCS, uma família de proteínas responsável pela supressão de respostas de citocinas, nomeadamente de respostas por interferão.

A via de sinalização do NF-κB é comum a diversas vias de transdução de sinal intracelulares, sendo algumas delas responsáveis pela produção de interferão I. Desta forma, esta dissertação visa avaliar o efeito desta ORF na via de sinalização do NF-κB, como evidência preliminar de um possível efeito da mesma na indução de interferão I.

A hipótese proposta foi confirmada através de ensaios de luciferase, nos quais a presença da referida ORF inibiu a ativação do NFkB e a expressão génica dependente de PRD II. Para além disto, identificou-se ainda o aminoácido Phe92 como um resíduo importante para a função da proteína codificada por esta ORF, visto que a sua mutação pontual reverteu parcialmente o efeito da proteína. Sumarizando, esta dissertação sugere para esta proteína um papel modulador de respostas pró-inflamatórias de citocinas. A sua caracterização aprofundada poderá fornecer uma nova ferramenta para a manipulação de respostas de citocinas, e um candidato para a construção de uma vacina deletada atenuada contra o VPSA.

Palavras-chave: Vírus da peste suína africana, Imunomodulação, Evasão imunitária, Vacina

STUDIES ON A YET UNASSIGNED FUNCTION OF AN AFRICAN SWINE FEVER VIRUS PROTEIN POTENTIALLY INVOLVED IN THE HOST CELL RESPONSE TO INFECTION

Abstract

The innate immune system is the first line of response against viruses and is predominantly a response based on inflammatory cytokines, in particular interferons (IFNs). This host response has provided selective pressure for the evolution of multiple host evasion mechanisms, conferring an advantage to the invading pathogens. Viral genes evolved to manipulate the host immune system provide “readymade tools” that can be used for immune modulation and construction of attenuated virus vaccines.

African swine fever is a worldwide threat for pig production with no vaccine or treatment nowadays. The development of a vaccine attenuated by the deletion of viral host evasion genes responsible for inhibiting innate immunity would be a practical solution for this threatening disease.

A recent bioinformatic analysis of the African swine fever virus (ASFV) genome has identified an ORF that has a homology region to the SH2 domain. The SH2 domain is known for playing a key role in multiple signal transduction pathways. Particularly relevant, the SH2 domain is a critical structure of SOCS proteins. SOCS is a family of proteins responsible for suppressing cytokine responses, namely IFN responses.

The NF-κB pathway is a signalling route common to many intracellular signal transduction routes, some of them responsible for the IFN I induction. Therefore, the main goal of this dissertation is to assess the effect of the referred ORF on the NF-κB signalling pathway, as preliminary evidence of its possible effect on IFN I induction.

To address that question, luciferase gene reporter assays were performed, and the postulated hypothesis was confirmed. It was observed in this study that the presence of this ORF inside the cells inhibits NF-κB activation and thus PRD II dependent gene expression. Additionally, this work has also identified the Phe92 amino acid as a critical residue for the function of the protein encoded by this ORF, since a single-point mutation of that residue partially reversed the protein’s effect. Overall, this dissertation suggests for this recently identified protein a role as a modulator of host pro-inflammatory cytokine responses. Its further characterization must provide a novel tool for the manipulation of cytokine responses, and a candidate for the construction of a gene deleted attenuated ASFV vaccine.

Keywords: African swine fever virus, Immunomodulation, Immune evasion, Vaccine

Table of contents

List of figures

List of abbreviations

ASF – African swine fever

ASFV – African swine fever virus

ATP – Adenosine triphosphate

C-terminal – Carboxil-terminal

CIS or CISH – Cytokine-inducible SH2 domain-containing protein

CLR – C-type lectin receptor

CpG – Cytidine-phosphateguanosine

DC – Dendritic cell

DD – Death domain

dsRNA – Double-stranded RNA

ESS – Extended SH2

FERM – Four-point-one protein, Erzin, Radixin, Moesin)

GAF – Gamma-activated factor

GAS – Gamma-activated site

GFP – Green fluorescence protein

HA – Hemagglutinin

IκB – Inhibitor of NFκB

IFN – Interferon

IFN-α – Interferon-alpha

IFN-β – Interferon-beta

IFN-γ – Interferon-gamma

IFN-λ – Interferon-lambda

IFN I – Type I interferon

IFN II – Type II interferon

IFN III – Type III interferon

IFNAR – Interferon-alpha receptor

IFNBR – Interferon-beta receptor

IFNGR – Interferon-gamma receptor

IGC – Instituto Gulbenkian de Ciência

IKK – IκB kinase

IL – Interleukin

IL-1R1 – IL-1 receptor type I

IL-RAP – IL1R accessory protein

IRAK – IL-1 receptor-associated kinase

IRF – Interferon regulatory factor

ISFG3 – ISG factor 3

ISG – Interferon-stimulated gene

ISRE – Interferon-stimulated response element

JAK – Janus activated kinase

JH – Janus homology

KIR – Kinase inhibitory region

LPS – Lipopolysaccharide

MAPK – Mitogen-activated protein kinases

mDC – Myeloid dendritic cell

MGF – Multigene family

MyD88 – Myeloid differentiation primary response protein

N-terminal – Amino-terminal

NEMO – NFκB essential modifier

NF-κB – Nuclear factor κB

NIK – NFκB-inducing kinase

NK – Natural Killer

NLR – NOD-like receptor

NLS – Nuclear-localization signal

NOD – Nucleotide-binding oligomerization domain

ORF – Open reading frame

PAMP – Pathogen-associated molecular pattern

pDC – Plasmacytoid dendritic cell

pDP146L – DP146L protein

PEST – Proline-, glutamicacid-, serine- and threonine-rich

PRD – Positive regulatory domain

PRR – Pattern recognition receptor

PVDF – Polyvinylidene difluoride

Rbx2 – RING-box-2

RIG – Retinoic acid-inducible genes

RIP – Receptor-interacting protein

RLR – RIG-I-like receptor

rpm – Rotations per minute

RT – Room temperature

ssRNA – Single-stranded RNA

SDS-PAGE – Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

SH2 – Src homology 2

SOCS – Suppressor of cytokine signalling

SODD – Silencer of death domains

STAT – Signal transducers and activators of transcription

TAB – TAK1-binding proteins

TAK – Transforming growth factor beta-activated kinase

TIR – Toll-IL-1 receptor

TIRAP – TIR domain containing adaptor protein

TLR – Toll-like receptor

TNF – Tumor necrosis factor

TNFR – TNF receptor

TRADD – TNFR1-associated death domain

TRAM – TRIF-related adaptor protein molecule

TRAF – TNF receptor-associated factor

TRIF – TIR domain-containing adaptor inducing IFN-β

TYK2 – Tyrosine kinase 2

1. INTRODUCTION

1.1. The immune system

To survive in a competitive environment, all multicellular organisms have developed a set of organs, cells, molecules, and processes that operate to protect them from invading pathogens. It is called the immune system, and to be effective, it must be able to appropriately detect and respond to foreign antigens, such as microorganisms, cancer cells, and toxins (Marshall et al. 2018). In fact, it is the wide variety of pathogens and the requirement for self-defense that has been responsible for the evolution of such a complex system. In simple multicellular life forms, it consists of very ancient defense strategies. But in mammals, it is a much more tangled system, divided into two main mechanisms: the innate immune system and the adaptive immune system (Hirano et al. 2011). Although being often described as two separate arms of the host defense, they are interconnected and cooperate to protect the host against a wide variety of threats, with the innate response representing the first line of defense and the adaptive response acting later (Chaplin 2010). The adaptive immune response will be briefly presented in this dissertation, but the main focus will lie on the innate immune system.

1.1.1. Adaptive Immune System

The adaptive immune system is antigen-dependent, antigen-specific, and has a delay between exposure to the antigen and maximal response. It is triggered by the innate immune system itself. One of the primary features of the adaptive response is the large variety of antigen-specific receptors expressed on the surface of its cells, enabling the organism to deal in a targeted way against a wide spectrum of invading organisms (Murphy and Weaver 2017; Marshall et al. 2018).

The central cells of this system are two types of lymphocytes named B cells and T cells. B cells mediate the humoral immune response, which is an antibody-mediated immunity. Antibodies play an essential role in containing virus proliferation during the acute phase of viral invasion, even if generally not being enough for virus elimination once an infection has been established. In this case, cell-mediated immune mechanisms are critical for host defense in a cellular immune response mediated by T cells, which works mainly by activating antigen-specific cytotoxic T lymphocytes that induce apoptosis of cells exhibiting foreign antigens (Bonilla and Oettgen 2010; Murphy and Weaver 2017; Marshall et al. 2018).

Another capability of the adaptive immune system is the capacity to develop immunologic memory, which enables the host to grow a faster and more efficient immune response upon subsequent exposure to the same antigen (Bonilla and Oettgen 2010; Murphy and Weaver 2017; Marshall et al. 2018). Such ability of the adaptive response is also the cornerstone of vaccination.

1.1.2. Innate Immune System

The innate immune system is the first line of defense of an organism. It is an antigen-independent defense mechanism that acts up to hours after perception of an antigen. It includes phagocytes (macrophages and neutrophils), dendritic cells, mast cells, basophils, eosinophils, natural killer (NK) cells, and innate lymphoid cells (Turvey and Broide 2010; Murphy and Weaver 2017; Marshall et al. 2018).

In the mid-1990s, the established idea that innate immunity only recognizes pathogens in a nonspecific manner fell due to the new concept of innate specific recognition. Innate immunity to pathogens relies on pattern recognition receptors (PRRs), a family of receptors first discovered by Charles Janeway (Janeway 1989). PRRs are proteins that have evolved to recognize common molecules shared by a wide range of foreign pathogens, known as pathogen-associated molecular patterns (PAMPs) (Marshall et al. 2018). This complex system is divided into families according to their tissue-specific expression and localization. It includes transmembrane proteins such as the Toll-like receptors (TLRs) and C-type lectin receptors (CLRs), but also cytoplasmic proteins such as Retinoic acid-inducible genes (RIG)-I-like receptors (RLRs) and nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs), and several cytosolic DNA receptors (Kawai and Akira 2010; Murphy and Weaver 2017). Different PRRs detect different PAMPs, and trigger a specific signalling pathway. Activation of PRRs on sensor cells, like macrophages and neutrophils, amplifies the immune response by producing inflammatory mediators, such as cytokines.

Cytokines consist of a broad category of small proteins. They are signalling molecules secreted and released by immune cells to trigger specific effects on nearby cells. The term “cytokines” includes interferons, interleukins, lymphokines, and tumor necrosis factors, and they are characterized as either pro-inflammatory, anti-inflammatory, or chemokines (Duncan et al. 2017; Schwartz et al. 2017). The interferons (IFNs) are cytokines with strong antiviral properties representing one of the first lines of host defenses. They are the most crucial antiviral cytokines of the vertebrates’ innate immune response (Correia et al. 2013).

1.1.2.1. Interferons

Interferons were first described in 1957, by Isaacs and Lindenmann (Isaacs and Lindenmann 1957), as molecules that could “interfere” with live virus replication. They are a family of cytokines produced by cells in response to pathogens and tumor cells. Their most outstanding attribute is their ability to interfere with viral replication in a nonspecific way, inhibiting intracellular propagation and intercellular transmission of the virus, establishing an antiviral state. IFNs also have a significant role in the activation of the adaptive antiviral immune response. Mammalian IFNs are divided into three types according to their structure, biological activities, and receptors used for signalling (Fensterl and Sen 2009; Correia et al. 2013).

Type I interferon (IFN I) family includes interferon-alpha (IFN-α), interferon-beta (IFN-β), and six less studied IFN I classes (IFN-ω, IFN-τ, IFN-κ, IFN-ε, IFN-δ, and IFN-ζ). IFN-α/β are produced in most cell types in response to viral molecular pattern recognition. They are induced directly by viral infection and are essential in coordinating the innate and adaptive immune responses. Once secreted, they signal through the ubiquitously expressed cell IFN-α/β receptor (IFNAR/IFNBR), amplifying the original IFN I signal (Akira and Takeda 2004; Honda et al. 2006; Correia et al. 2013; McNab et al. 2015).

Type II interferon (IFN II) family comprises only one member, the interferon-gamma (IFN-γ), also known as “immune IFN”. Whereas IFN I is ubiquitously produced by infected cells, secretion of IFN-γ is restricted to the immune system cells, such as T cells, NK cells, dendritic cells (DCs), and macrophages. Once secreted, it signals via the ubiquitously expressed IFN-γ receptor (IFNGR) and stimulates critical cell-mediated immune responses against pathogenic intracellular microorganisms. It also has a central role in the development of antitumor immune responses. Furthermore, although the antiviral activity is not its primary biological function, IFN-γ can also amplify the effect of IFN-α/β, working together to eliminate viral infections (Young and Bream 2007; Correia et al. 2013; McNab et al. 2015).

The most recently identified IFN family is type III interferon (IFN III, or IFN-lambda (IFN-λ)), composed by IFN-λ1, IFN-λ2, IFN-λ3 (also known as interleukin (IL)-29, IL-28A, IL-28B, respectively), and IFN-λ4. They share some function similarities with the IFN I family, being also ubiquitously induced by viral infections. However, unlike IFN I that exert antiviral activity on all cell types, IFN III’s action seems to be limited since the expression of their receptor is restricted mainly to epithelial surfaces. Thus, this IFN family plays a prime role in innate antiviral defense at epithelial surfaces, which is a primary portal of entry for many infections (Fensterl and Sen 2009; Correia et al. 2013; McNab et al. 2015).

The IFN system assembles two major activities: the induction of the IFN response by contact with the pathogen, which leads to IFN production; and the impact of secreted IFN on neighboring cells, culminating in the establishment of the anti-viral state.

1.1.2.2. Interferon Induction

There are several ways for a cell to recognize a viral infection and trigger the production of interferons. Upon viral infection, cells recognize PAMPs through PRRs, and several pathways are initiated in order to induce type I IFNs, which finally leads to the transcription of IFN-stimulated genes (ISGs), thus establishing an antiviral state. Depending on its nature, each PAMP can be detected in different cellular compartments: they can be detected in the cytoplasm, by cytosolic receptors; or in the cell membrane and endosomes, by TLR receptors (Figure 1) (Randall and Goodbourn 2008; Fensterl and Sen 2009).

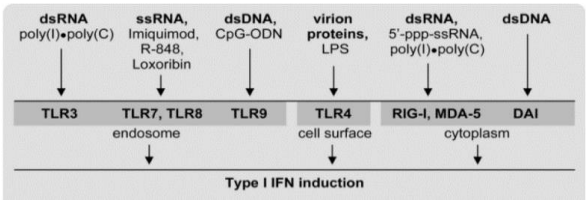

Figure 1. IFN I-inducing PRRs, their ligands, and their cellular localizations (Fensterl and Sen 2009)

Type II IFN production does not occur in direct response to viral PAMPs. It takes place later in the infection process as a second wave of cytokine expression. IFN II secretion is not within the scope of this dissertation, and therefore, it will not be further developed.

Despite different signalling pathways can occur upon viral detection, there are two transcription factors that can be activated and are common to most IFN I induction signal cascades: IRF3 and NF-κB. IRF3 is cytoplasmic before induction. After the appropriate signal, it undergoes a conformational change with dimerization and exposition of a nuclear-localization signal (NLS). It is translocated into the nucleus and activates the IFNβ promoter, among others (Dragan et al. 2007; Randall and Goodbourn 2008). NF-κB is a family of transcription factors playing critical roles in inflammation, immunity, and cell proliferation, differentiation and survival. It consists of a group of dimmers, being the p50/p65 heterodimer the most representative one (Oeckinghaus and Ghosh 2009). In unstimulated cells, NF-κB dimers are retained as an inactive form in the cytosol by association with the inhibitor of NF-κB (IκB). The appropriate signal causes phosphorylation of the IκB by the IκB kinase (IKK) complex, and its subsequent ubiquitination and degradation by proteasome, releasing the NF-κB. The NLS of p65 subunit activates and NF-κB is translocated into the nucleus, inducing transcription of target genes, including IFNβ. Additionally to these two transcription factors, ideal induction of the IFNβ gene requires binding of the c-Jun/ATF-2 heterodimer to the promoter. (Wullaert et al. 2006; Randall and Goodbourn 2008).

The promoter region of IFN-β contains four regulatory elements: the positive regulatory domains (PRDs) I, II, III, and IV. The PRD II domain is the binding site for NF-κB; IRF3 binds to PRD I/III; and c-Jun/ATF-2 binds to PRD IV. All transcription factors linked to the PRDs form the enhanceosome, enabling the promoter to be activated and IFN-β to be secreted. However, promotermapping studies have shown that the individual binding sites can respond independently to stimuli and confer some degree of response. Currently, the consensus view is that IRF3 (or IRF7) is indispensable for induction, but NF-κB and c-Jun/ATF-2 may not be (Randall and Goodbourn 2008).

1.1.2.2.1. Cytosolic Pathways

The cytosolic pathways are TLR-independent pathways owned by cells to detect viral nucleic acids and induce type I IFN production. Such pathways are in most cases activated by viral genomes, both RNA and DNA molecules, present in the cytoplasm at some time point of the infection time course. Viruses become thus susceptible to detection by different types of receptors, depending on the viral nucleic acid type, leading to different signalling pathways (Randall and Goodbourn 2008; Fensterl and Sen 2009).

1.1.2.2.1.1. Intracellular viral RNA

The dsRNA from viral RNA genomes or viral replication intermediates can be detected in the cytoplasm by RIG-I-like receptors (RLRs), a family of receptors consisting of three members: retinoic acid inducible gene I (RIG-I), melanoma differentiation associated gene 5 (MDA5), and laboratory of genetics and physiology 2 (LGP2) (Randall and Goodbourn 2008; Wilkins and Gale Jr 2010; Christensen and Paludan 2017).

After ligand binding, both RIG-I and MDA-5 receptors use the mitochondrial adaptor protein MAVS (mitochondrial antiviral signalling protein) (also known as CARD adaptor inducing IFN-β (Cardif)/virus-induced signalling adaptor (VISA)/IFN-β promoter stimulator protein 1 (IPS-1)) as a platform for the initiation of the signalling transduction (Figure 2) (Randall and Goodbourn 2008; Fensterl and Sen 2009). RIG-I, but not MDA-5, requires ubiquitination by TRIM25 to interact with MAVS (Gack et al. 2007).

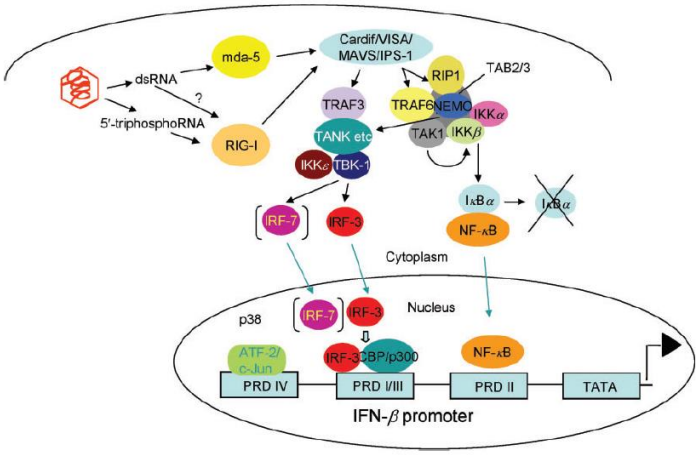

Figure 2. MDA-5 and RIG-I dependent signalling pathways (Randall and Goodbourn 2008)

The recruitment of MAVS triggers recruitment of both TRAF6 and TRAF3, leading signalling through both NF-κB and IRF-3 pathways, respectively. The signalling branch via the TRAF6 complex activates the kinases IKK, p38, and JNK, which results in phosphorylation and activation of the transcription factor NF-κB, and also ARF-2 and C-JUN. In the other branch,

via the TRAF3 complex, the kinases TBK1 and IKKε became activated. They then phosphorylate and activate both the IRF3 and IRF7 transcription factors. Activation of all these transcription factors, that then translocate into the nucleus, culminates in the expression of IFN I, as well as inflammatory cytokines (Randall and Goodbourn 2008; Fensterl and Sen 2009; Wilkins and Gale Jr 2010; Christensen and Paludan 2017).

1.1.2.2.1.2. Cytoplasmic DNA

Other cytosolic PRRs recognize the presence of DNA inside the cells and also trigger for IFN induction. Several cytoplasmic DNA receptors have been proposed, including absent in melanoma (AIM2), gamma-interferon-inducible protein (IFI16), and cyclic GMP-AMP synthase (cGAS), and each receptor triggers a particular pathway (Figure 3) (Christensen and Paludan 2017). Such receptors are mainly present in macrophages and DCs, which are the most common mammalian cells sensing cytosolic DNA (Randall and Goodbourn 2008).

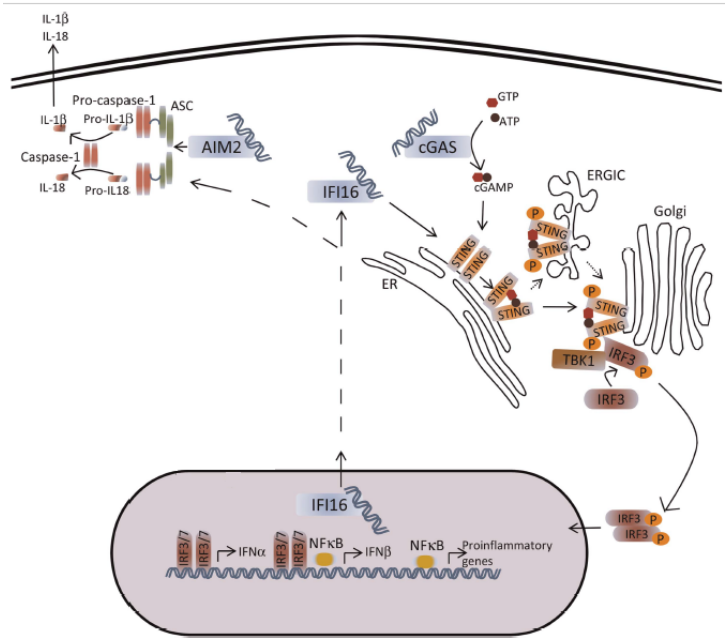

Figure 3. AIM2, IFI16, and cGAS dependent signalling pathways (adapted from Christensen and Paludan 2017)

Detection of viral DNA by cGAS receptors induces its conformational change and activation, triggering the synthesis of the second messenger cyclic GMP-AMP (2'3'-cGAMP). cGAMP binds to STING, which results in the translocation of STING from the ER to perinuclear autophagy-like vesicles. TBK1 is then autophosphorylated and recruited to STING, inducing the phosphorylation of the latter. Such allows the transcription factor IRF3 to dock to the STING's phosphorylated residue, resulting in IRF3 phosphorylation and dimerization. IRF3

thus translocates into the nucleus and induces the transcription of IFN I genes (Christensen and Paludan 2017; García-Sastre 2017).

The IFI16 receptor is known to sense longer and naked DNA, both in the cytoplasm and nuclear compartments. Its signalling pathway is not fully uncovered, but it is currently thought to act in the same route as cGAS, through STING, leading also to the production of IFN I (Christensen and Paludan 2017).

Finally, DNA binding to the AIM2 receptor initiates the assembly and activation of the inflammasome complex. Such leads to the maturation of the proforms of the cytokines IL-1β and IL-18, which stimulates bioactive cytokines production (Christensen and Paludan 2017).

1.1.2.2.2. TLR Signalling Pathway

TLRs are transmembrane proteins that recognize different types of PAMPs, derived from a wide range of microorganisms. Twelve TLRs are known so far, each with distinct functions in terms of PAMP recognition and immune response (Kawai and Akira 2010).

TLRs can be divided in two subgroups. TLR1, TLR2, TLR4, TLR5, TLR6 and TLR11 are expressed on the cell surface and recognize mainly microbial membrane components. TLR3, TLR7, TLR8 and TLR9 are exclusively expressed in intracellular compartments, such as endosomes, lysosomes, and endolysosomes. They detect nucleic acids taken into cells by phagocytosis, receptor-mediated endocytosis or micropinocytosis. (Kawai and Akira 2010; Murphy and Weaver 2017). TLR3 can also be expressed on the plasma membrane in some cells, like fibroblasts. TLR3, TLR7, and TLR9 are the main TLR responsible for the detection of viral PAMPs upon viral infections (Kawai and Akira 2010).

TLR3 recognizes double-stranded RNA (dsRNA), which is a replicative intermediate of many types of viruses. It shows a relatively wide tissue distribution (Randall and Goodbourn 2008), which enable the detection of both extracellular and intracellular viral nucleic acids. TLR3 triggers the production of type I interferon and inflammatory cytokines, revealing the importance of this activation route in antiviral defense (Tabeta et al. 2004; Randall and Goodbourn 2008; Kawai and Akira 2010; Murphy and Weaver 2017).

TLR7 and TLR9 are all exclusive endosomal receptors, highly expressed by plasmacytoid DCs (pDCs). TLR7 recognizes endosomal single-stranded RNA (ssRNA) derived from RNA viruses, and produces large amounts of type I IFN (Diebold et al. 2004; Kawai and Akira 2006; Randall and Goodbourn 2008). TLR9 serves as a sensor of DNA virus infection, recognizing endosomal “foreign” unmethylated 22-deoxyribo(cytidine-phosphateguanosine) (CpG) DNA, frequently present in bacteria and viruses (Tabeta et al. 2004; Randall and Goodbourn 2008; Kawai and Akira 2010).

Additionally, TLR4 and TLR2 induce IFN-α and IFN-β in response to extracellular viral particles detection. Nonetheless, and unlike the most usual pathway, this particular pathway of IFN I production is not triggered by nucleic acids (Murphy and Weaver 2017).

Upon PAMP recognition, two cytoplasmic adaptor molecules interact with TLR's intracellular domain in order to trigger subsequent pathways: MyD88 (Myeloid differentiation primary response protein) and TRIF (TIR domain-containing adaptor inducing IFN-β) (Figure 4). Moreover, TRAM (TRIF-related adaptor protein molecule) and TIRAP (TIR domain containing adaptor protein) function as sorting adaptors that are needed to recruit TRIF and MyD88, respectively, for some TLRs. MyD88 is used by all TLRs except TLR3, and only requires the adaptor TIRAP to TLR4 and TLR2. TRIF is used by TLR3 and TLR4, requiring the adaptor TRAM only for the TLR4 route. Thus, TLR signalling pathways can be largely classified as either MyD88-dependent pathways or TRIF-dependent pathways (Kawai and Akira 2010).

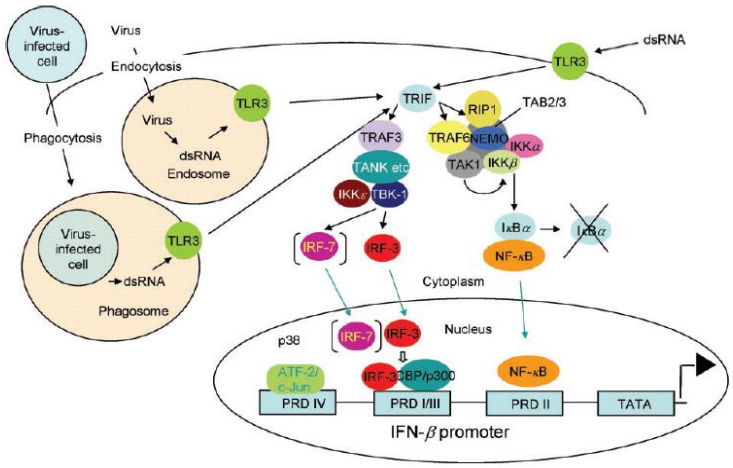

Figure 4. TLR3 route as an example of the TLR signalling pathways (Randall and Goodbourn 2008)

Independently of the adaptor molecules recruited, TLR viral detection triggers the same two signalling branches previously exposed for the intracellular viral RNA receptors pathway (subsection 1.1.2.2.1.1). The first branch is via the TRAF6 complex, and leads to the activation of NF-κB transcription factor, being an NF-κB-dependent pathway. The other branch signals via the TRAF3 complex, leading to activation of IRF-3 and IRF-7 transcription factors. Nuclear translocation of IRF-3 and IRF-7 induces IFN I secretion in an NF-κB-independent manner (Randall and Goodbourn 2008; Kawai and Akira 2010; Murphy and Weaver 2017).

1.1.2.3. Interferon Impact

After being secreted, IFNs move along to neighbor cells in order to amplify the antiviral signal by direct interaction with specific cell-surface receptors. Once ligated to the receptors, IFNs activate the JAK-STAT pathway, a common signalling pathway used by many cytokines.

The JAK-STAT pathway is an intracellular signalling pathway that involves Janus activated kinases (JAK)s and signal transducers and activators of transcription (STAT)s proteins. It plays an essential role in driving biological responses to cytokines (Imada and Leonard 2000).

1.1.2.3.1. Janus Family Tyrosine Kinases

The JAK proteins are the first members of the JAK-STAT Pathway. Each JAK is constituted by seven domains, from Janus homology (JH) 1 to JH7. JH1 is a kinase domain, the functional catalytic one; JH2 is a pseudokinase domain, whose function is still unclear; JH3 is an Src homology 2 (SH2)-like domain; and JH4-7 are the FERM (Four-point-one protein, Erzin, Radixin, Moesin) domain (Figure 5) (Imada and Leonard 2000; Lee and Rhee 2017; Schwartz et al. 2017). The mammalian JAK family has four members: JAK1, JAK2, JAK3, and tyrosine kinase 2 (TYK2); different JAKs associate with different cytokine receptors, initiating a slightly different pathway (Shuai and Liu 2003; Lee and Rhee 2017).

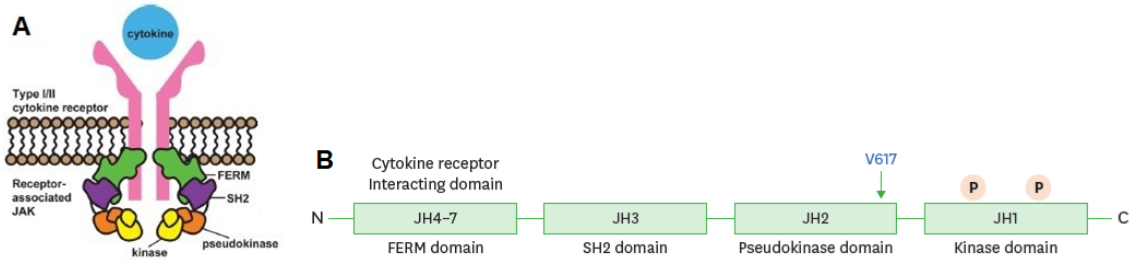

Figure 5. Schematic representation of JAK protein's structure. A. (Schwartz et al. 2017) **B.** (Lee and Rhee 2017)

JAKs are near the cell membrane, associated with the cytokine receptors. Some of them are constitutively associated with the receptors, and in others the association is enhanced by cytokine stimulation (Imada and Leonard 2000; Schwartz et al. 2017). JH6 and 7 are responsible for binding JAKs to the cell receptors (Imada and Leonard 2000; Lee and Rhee 2017; Schwartz et al. 2017). As tyrosine kinases, JAKs' function is to transfer a phosphate from ATP molecules to tyrosine residues on other proteins, like cytokine receptors, JAKs themselves, and downstream signalling molecules. The JH1 domain contains two tyrosines that can be phosphorylated after ligand stimulation (Schwartz et al. 2017).

1.1.2.3.2. Signal Transducers and Activators of Transcription

STAT proteins are the following crucial molecules in the JAK-STAT pathway. When phosphorylated, they translocate to the nucleus, bind DNA and drive gene transcription, finishing the signal transduction (Schwartz et al. 2017). There are seven mammalian STATs:

STAT1, 2, 3, 4, 5A, 5B, and 6, which are highly homologous in several regions, like the SH2 domain, the DNA-binding domain, and the transactivation domain (Shuai and Liu 2003).

Each STAT has a conserved and rapidly phosphorylated tyrosine residue, approximately 700 residues from the amino (N)-terminal. Once phosphorylated by activated JAKs, this phosphotyrosine binds to the SH2 domain of another STAT, forming a dimer. Thus, this SH2 domain is involved in the activation and dimerization of STATs. Moreover, it is also responsible for recognizing the specific phosphotyrosine sequence on the activated cytokine receptor, playing an essential role in the association between STATs and receptors (Shuai et al. 1994; Mikita et al. 1998). The STATs' highly conserved N-terminal domain regulates STATs activity and STAT dimers oligomerization (Xu et al. 1996). The carboxil (C)-terminal region is the transcriptional activation domain (Figure 6) (Wen et al. 1995).

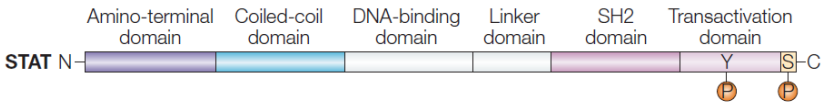

Figure 6. Schematic representation of STAT proteins structure (adapted from Shuai and Liu 2003)

1.1.2.3.3. The JAK-STAT pathway

IFN receptors typically consist of two sub-units and are transmembrane proteins. They have an extracellular cytokine binding domain and an intracytoplasmic domain that binds a JAK protein. In the absence of cytokine, the two receptor chains are not associated. Type I and type II IFN signal through distinct receptors. Depending on the type of receptor activated, slightly different pathways are triggered (Figure 7), with downstream components that can either be unique or common to both signaling pathways (Imada and Leonard 2000).

The IFN I (α/β) receptor is a heterodimer with two subunits: IFNα receptor (IFNAR) 1 and IFNAR2. Their cytoplasmic domains are associated with inactive JAKs, IFNAR1 with TYK2, and IFNAR2 with JAK1. IFNAR2 is also ligated to a STAT2 protein, with a STAT1 protein weakly bound. When IFN I binds the receptor, its dimerization and conformational change induces phosphorylation of IFNAR1 by TYK2, originating a docking site for STAT2. Then, TYK2 phosphorylates STAT2, and JAK1 phosphorylates STAT1. Phosphorylated STATs form a stable and active heterodimer. At this point, the STAT1-STAT2 dimer associates with IRF-9, creating a complex known as the IFN-stimulated gene (ISG) factor 3 (ISFG3). This complex then translocates into the nucleus, binds to ISRE (IFN-stimulated response element) (present in the promoter region of most ISGs), and induces the transcription of several hundreds of ISGs (Randall and Goodbourn 2008; Correia et al. 2013; Murray 2014; Dixon et al. 2019).

The IFN II (γ) receptor is also a heterodimer glycoprotein with two subunits: IFNGR1, and IFNGR2, weakly pre-associated in unstimulated cells. Their cytoplasmic domains are

connected with inactive JAKs, IFNGR1 with JAK1 and IFNGR2 with JAK2. When IFN II binds to the receptor, the signalling pathway is initiated by triggering receptor dimerization. This brings JAK1 and JAK2 into close proximity, resulting in the activation of JAK2, which in turn trans-phosphorylates and activates JAK1. Then, activated JAKs phosphorylate IFNGR1, creating docking sites for STAT1. Two STAT1 molecules interact via SH2 domains with IFNGR1 and are phosphorylated, resulting in their activation and dissociation from the receptor. Activated STAT1 molecules dimerize, forming a STAT1-STAT1 homodimer called GAF (gamma-activated factor). Such translocates into the nucleus (without the need for IRF9) and binds to GAS (gamma-activated site) motifs present in ISGs on target genes, inducing their transcription (Randall and Goodbourn 2008; Correia et al. 2013; Murray 2014).

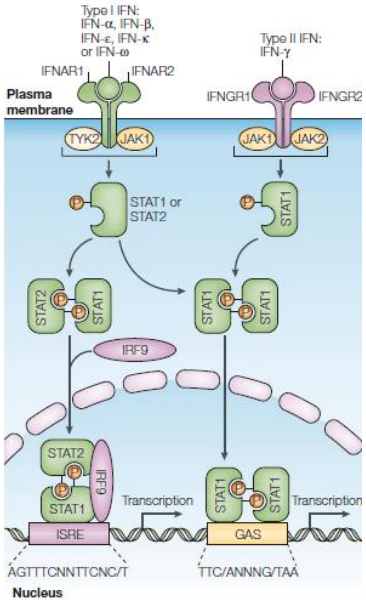

Figure 7. The JAK-STAT Pathway (Platanias 2005)

At the end of the signal transduction, both IFNs lead to the transcription of over 300 ISGs which mediate various biological responses. Their effect is to inhibit the viral cycle, thus stopping virus replication and transmission to other cells and establishing an antiviral state (Randall and Goodbourn 2008).

1.1.3. Immune regulation

As described above, the primary effect of interferons is antiviral protection through inducing ISGs and enhancing immune responses. They are capable of setting up a large-scale inflammatory response. However, various cytokines and stimuli are permanently present in the microenvironment of immune cells. Incorrect activation of immune responses and excessive production of cytokines and other inflammatory mediators can lead to dysregulation of host

cytokine signalling and ultimately cause tissue or organ dysfunction. Depending on the pathogen, the host, and the context, IFNs can either have protective effects in viral infection or contribute to immunosuppression or immunopathology (McNab et al. 2015; Huang et al. 2020). Therefore, in order to avoid damage to the host, the response to IFN must be tightly regulated and terminated once the threat is over. In the last years, many negative-feedback systems were discovered, and one of the most significant is the suppressor of cytokine signalling (SOCS) family of proteins (Yoshimura et al. 2007).

1.2. SOCS Proteins

SOCS are a family of intracellular proteins, discovered in 1997, and is composed by eight members: seven SOCS proteins (named SOCS 1-7) and a cytokine-inducible SH2 domain-containing protein (CIS or CISH) (Linossi et al. 2013; Huang et al. 2020).

SOCS proteins are currently known as key physiological regulators of both innate and adaptive immunity. They play an essential role in regulating signalling pathways at the intracellular level, acting in a classical negative-feedback loop to inhibit cytokine signal transduction and thereby avoiding body damage caused by its excessive secretion (Yoshimura et al. 2007; Huang et al. 2020). Early studies showed that transcription of the CIS, SOCS 2, and 3 genes were induced by many different cytokines and that expression of SOCS 1 and 3 could inhibit signaling from multiple receptor complexes (Linossi et al. 2013). Nevertheless, their degradation and regulation mechanisms are still not fully known (Yoshimura et al. 2007).

In the past two decades, the understanding of the structure and function of SOCS proteins has grown. The following subsections will review the present knowledge on the topic.

1.2.1. General Structure and Function

Each SOCS protein contains an N-terminal domain of variable length and composition, a central SH2 domain, and a conserved C-terminal 40 amino acids module, also called SOCS box (Figure 8). SOCS proteins can be subdivided into two groups based on their N-terminal region length: CIS and SOCS 1-3 have a short N-terminal region, while SOCS 4-7 hold a long N-terminal region with 270-385 residues (Linossi et al. 2013).

The SOCS-SH2 domain contains a short conserved α-helical extension, known as the N-terminal extended SH2 domain (ESS). The ESS contributes to the binding and stabilizes the modular SH2-subdomain (Babon et al. 2006; Linossi et al. 2013). The N-terminal region of SOCS3 is also involved in its protein stability (Sasaki et al. 2003; Yoshimura et al. 2007).

A proline-, glutamicacid-, serine- and threonine-rich (PEST) motif is present within the SH2 domain of SOCS 1, 3, 5, 7, and CIS (Huang et al. 2020). Such is not required for protein function but increases SOCS turnover and affects its protein stability (Babon et al. 2006).

Another particular structure, only present on SOCS 1 and 3, is the kinase inhibitory region (KIR) motif, located exactly N-terminal to the ESS-SH2 (Sasaki et al. 1999; Linossi et al. 2013).

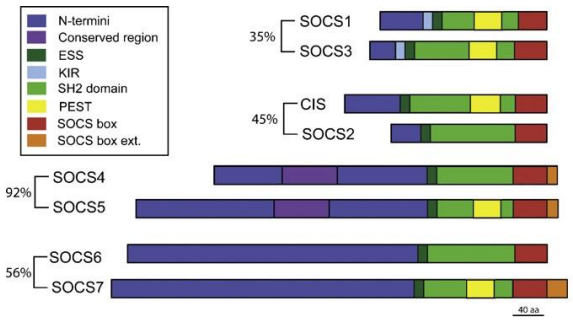

Figure 8. SOCS proteins' structure (adapted from Linossi et al. 2013)

SOCS proteins act as cytokine signalling regulators in various ways, as each protein domain has its role to achieve the end goal effect. The SH2 domain and the SOCS box work together for a process called ubiquitination. The SH2 domain has as its primary function determining the target of each SOCS protein. It interacts with the target by binding specific phosphorylated tyrosine residues on its preferred substrate. The presence of the ESS structure enhances this substrate interaction (Liu et al. 2006; Huang et al. 2020). The SOCS box has an Elongin B and C binding site that interacts with the Cullin 5 scaffold, which in turn recruits a RING-box-2 (Rbx2) protein, creating the Elongin B and C/Cullin 5/Rbx2 complex. Once SH2 binds to the chosen substrate, the SOCS box interacts with that complex, which recruits an E2 ubiquitin-conjugating enzyme. In this way, the SH2-bound substrate is brought into close proximity with the ubiquitinating machinery, and thus SOCS proteins assist the ubiquitination of target proteins, marking them for degradation via proteasome (Figure 9). So, SOCS proteins, as well as other SOCS-box-containing molecules, seem to act as E3 ubiquitin ligases. However, the specific role and function of the SOCS box are in need of further exploration (Akhtar and Benveniste 2011; Linossi et al. 2013). Lastly, a particularity of SOCS 1 and 3 is given by their KIR motif, a pseudosubstrate that can inhibit tyrosine kinase activity, being a supplementary attribute for suppressing cytokine signals (Yoshimura et al. 2007).

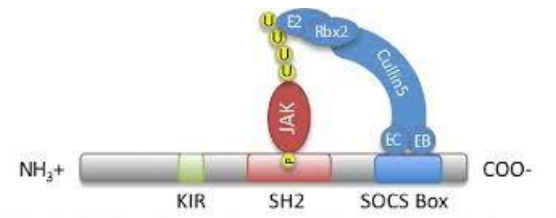

Figure 9. The SH2 domain and SOCS box ubiquitination process (Akhtar and Benveniste 2011)

Collectively, this complex and multifaceted structure provides for an extensive range of beneficial biological effects. Most research done to date focus on CIS and SOCS 1-3. Accordingly, SOCS 4-7 are less known, and their precise biological roles are yet to be fully uncovered (Linossi et al. 2013).

1.2.2. Mechanism of Action

SOCS proteins have been demonstrated to negatively regulate cytokines' response. They can inhibit both IFN induction and impact pathways, as critical steps to prevent excessive inflammatory response and consequent damage to the host. SOCS proteins act through several mechanisms, all of which rely in part on their SH2 domains (Linossi et al. 2013).

As explained earlier, the SOCS box has a ubiquitination role, acting as an E3 ubiquitin ligase. All SOCS proteins are able to ubiquitinate bound proteins, such as JAKs and cytokine receptors, leading them to proteasomal degradation (Linossi et al. 2013). It is important to remember that the SH2 domain also has a vital role in this system since it is responsible for signalling the target substrates for ubiquitination.

Besides that, SOCS proteins can also bind to JAK-STAT-involved proteins and block their activity. The SH2 domain of SOCS 1, together with the KIR structure, attaches directly to the JAK's activation loop, thus preventing JAK phosphorylation (Kubo et al. 2003; Yoshimura et al. 2007). However, the SH2 domain of SOCS 2, 3, and CIS does not have this high affinity to the activation loop of JAKs. It can only bind to phosphorylated tyrosine residues on the cytokine receptors and block STAT recruitment. Nevertheless, although the SH2 domain of SOCS 3 does not have a high affinity for JAK compared to SOCS 1, its KIR structure has a higher affinity for the kinase domain of JAK2. It can cover the substrate-binding groove of JAK2 and block its association with the substrate, thus preventing the perpetuation of the signalling pathway (Sasaki et al. 1999; Yoshimura et al. 2007; Huang et al. 2020).

CIS and SOCS2 can block the binding site of STATs on the receptor. Their SH2 domain bind to JAK-phosphorylated cytokine receptors and block STATs recruitment, thus competing with them and inhibiting STATs activation (Yoshimura et al. 1995; Huang et al. 2020)

SOCS1 can directly bind to IFN I receptors, ensuring that SOCS1 has a very efficient suppressive effect on IFN I signalling even at low expression levels. It appears to interact with the TYK2/IFNAR1 complex directly (Piganis et al. 2011; Linossi et al. 2013). Besides, SOCS1 regulates IFNAR1-specific but not IFNAR2-specific signals, thereby eliminating tyrosine phosphorylation of STAT1 and minimizing the antiviral gene expression (Mansell et al. 2006; Yoshimura et al. 2007). However, despite having a beneficial inhibitory effect on the IFNγ pathway, recent studies have shown that SOCS1 can not directly bind to IFNγ receptors. Instead, complete suppression of IFNγ signalling by SOCS1 requires phosphorylation of the IFNGR1 subunit by its SH2 domain (Qing et al. 2005; Yoshimura et al. 2007).

Concerning IFN I induction pathways, a recent study by Zhou et al. (2015) showed that both SOCS1 and SOCS3 affect TLR pathway, directly interacting with TRAF6 and degrading it through polyubiquitination (Zhou et al. 2015). Moreover, a study by Yu et al. (2018) demonstrated that SH2 domains of SOCS1 and SOCS3 promote the proteasome-mediated degradation of IRF7 through polyubiquitination (Yoshimura et al. 2007; Linossi et al. 2013; Yu et al. 2018; Huang et al. 2020). SOCS3 also inhibits the activation of TRAF6 and TAK1, which are key factors to induce and signal TLR responses (Frobøse et al. 2006).

1.2.2.1. Regulation of the NF-κB-dependent Pathway

Besides the mechanisms of action already described, SOCS proteins have also an inhibitory role in the NF-κB-dependent signalling route, one that is particularly relevant for the work developed in this dissertation.

Many mechanisms have already been proposed for the effect of SOCS1 in the NF-κB pathway. Firstly, SOCS1 can directly interfere in the signal transduction of the NF-κB pathway by binding to the p65 subunit of NF-κB, thus promoting its degradation by ubiquitination (Ryo et al. 2003; Yoshimura et al. 2007; Huang et al. 2020). Besides that, SOCS1 also targets tyrosine-phosphorylated TIRAP for programmed degradation, thereby suppressing the TIRAP-mediated phosphorylation of the p65 subunit of NF-κB, and thus NF-κB transactivation (Mansell et al. 2006; Yoshimura et al. 2007; Huang et al. 2020).

Finally, SOCS3 can also negatively regulate the NF-κB-dependent proinflammatory cytokine response. Phosphorylated SOCS3 binds directly to the NF-κB inhibitor IκB, masking its phosphorylation site, and thus delaying its phosphorylation and subsequent degradation. Therefore, SOCS3 prevents the activation and nuclear translocation of the NF-κB transcription factor, impairing the extension of the transduction signalling pathway (Nair et al. 2011; Huang et al. 2020).

1.3. Viral Strategies for Immune Evasion

During the course of evolution, the continuous interaction between viruses and their hosts has shaped and determined the survival strategies used by both. Vertebrate hosts have evolved an elaborate system of innate and adaptive immune responses to recognize and destroy infected cells, interfering with viral replication. In turn, the selective pressure imposed by host defenses led viruses to develop mechanisms to evade host immune machinery. To replicate and spread in a host, a virus depends on highly specific interactions between viral host evasion proteins and infected cells in a well-balanced operation, since a complete inhibition of host defenses would lead to host death and viral elimination (Engel and Angulo 2012; Correia et al. 2013).

Viruses use two main strategies to evade the host defense mechanisms: to manipulate the components of the innate immune system (such as cytokines, chemokines, apoptosis, and others) or to avoid recognition by the adaptive immune system. The list of viral mechanisms to manipulate the host immune response is extensive. However, downregulation of the IFN system seems to be a priority for most viruses. They can interfere with the induction and impact of IFN, or they can block molecules that mediate the "antiviral state" (Correia et al. 2013). Given that the induction of IFN proceeds in a cascade-like manner, viruses have evolved molecular mechanisms that act at different steps in the transduction pathways. Almost 50% of the studied viruses interfere at multiple steps in the IFN response (Versteeg and García-Sastre 2010). Modulation of the STAT proteins' activity and inhibition of the JAK kinases, thus inhibiting signalling through it, are very common viral strategies as well (Weber and Haller 2007; Correia et al. 2013).

Due to the immunomodulatory effect of SOCS proteins and its regulatory role in cytokines (especially IFNs), it is predictable that viruses could also manipulate the host's SOCS proteins to escape innate immunity. In fact, a growing number of studies has shown that pathogens can induce SOCS1 and SOCS3 to evade deleterious host immune responses (Duncan et al. 2017). For example, Porcine Reproductive and Respiratory Syndrome virus hijack the host's SOCS1 protein to inhibit the expression of IFN-β and promote viral replication (Luo et al. 2020). Furthermore, when Hepatitis C Virus and Herpes Simplex Virus 1 infect human hepatoma cells and amniotic cells, respectively, they cancel IFNα/β signalling by enhancing SOCS3 protein expression (Bode et al. 2003; Yokota et al. 2004). Lastly, Influenza Virus infection mostly inhibits the JAK-STAT pathway by up-regulating the expression of SOCS1 and SOCS3, disrupting host antiviral defense mechanisms (Jia et al. 2010).

The knowledge of these evasion strategies provides "ready-made tools" for gene manipulation, which can result in a deeper understanding of immune and inflammatory responses. Besides that, the study of these mechanisms can lead to new approaches to manage viral infections, namely the design of novel and effective antiviral treatments and attenuated virus vaccines.

1.4. African Swine Fever Virus

African swine fever (ASF) is a highly contagious viral hemorrhagic disease affecting the family *Suidae*. The African swine fever virus (ASFV) is transmitted by direct contact between infected animals or by ingestion of infected pork or other contaminated materials (Guinat et al. 2014). Naturally, the close proximity between wild boars and domestic pigs in extensive production systems may favor this contact. There is also an indirect way of transmission by fomites (such as clothes, transport trucks, and feed supplies), in which human practices have significant responsibility (Dixon et al. 2020; Salguero 2020). Soft tick vectors of the genus

Ornithodoros can also play a role in ASFV transmission, including *O. moubata* in Africa and *O. erraticus* in the Iberian Peninsula (Basto et al. 2006; Jori and Bastos 2009; Jori et al. 2013).

ASF was first described in East Africa (Montgomery 1921) in domestic pigs. Now we know that the source of the infection was an ancient sylvatic cycle between African wild suids (mostly Warthogs (*Phacochoerus Africanus*)) and soft ticks. Since then, the disease was described in most sub-Saharan African countries. In 1957/1960, it was introduced in Portugal and then spread to other European countries, the Caribbean and Brazil. The disease was eradicated from all these countries outside Africa, except the Italian island of Sardinia, where it remains endemic. It reappeared in 2007 in Georgia (Caucasus), and a new transmission era began. The virus rapidly spread across the Caucasian region and neighboring countries and is now present across Central and Eastern Europe, including some Union European countries. It also expanded to the east, having reached Russia and, more recently, to China in 2018, where it is currently widespread. However, European and Asian situations are distinct: the former mostly involves outbreaks in wild boars, while for the latter domestic swine are the most affected. Nowadays, ASF is present worldwide, except for the Americas, Australia, and New Zealand (Acharya and Wilson 2020; Dixon et al. 2020; Kedkovid et al. 2020).

ASFV is spreading very rapidly, and without vaccine or treatment, the only available methods for disease control are the quarantine of affected areas and the slaughter of infected and susceptible animals. Rapid detection is urgent when ASFV enters new territory to minimize the consequences of these drastic measures (Dixon et al. 2020). It is a disease with high economic impact due to business decline across the whole pig production chain, high costs for disease control, and consequent lack of commerce. ASF currently poses a serious threat to European pig production, and thus the development of a vaccine is of extreme urgency. Yet, more than immunity, there is also a present-day need to create a vaccine that includes markers for differentiation between infected and vaccinated animals, which naturally constitutes an added challenge. Lastly, a vaccine for wild boars, given as bait, should also be helpful in managing the spread of ASF (Arias et al. 2017; Galindo and Alonso 2017).

1.4.1. Viral structure and genome organization

The African swine fever virus is a large, enveloped, genetically complex virus with a double-stranded DNA genome that varies between 170 and 193kpb. It belongs to the family *Asfarviridae*, and some aspects of its genome structure and replication strategies are shared with other large DNA virus families, most notably the Poxviruses. It is the only known DNA arbovirus (Dixon et al. 2019; Rock 2021). For many years, ASFV was the only member of *Asfarviridae*. However, in 2009, a virus infecting a marine dinoflagellate was detected, and its partial sequencing revealed it likely belongs to *Asfarviridae*. New amoeba viruses have been

identified since then. Today we can say that despite not being the only member of its family, ASFV is the only one affecting vertebrate hosts (Karki et al. 2021).

The ASFV particle is composed of multiple concentric layers with an icosahedral morphology (Figure 10). It has a genome-containing nucleoid (first layer), a thick protein layer named core shell (second layer), an inner lipid envelope (third layer), and an icosahedral protein capsid containing over 50 proteins (fourth layer) (Alejo et al. 2018; Salas and Andrés 2013). Extracellular particles also have an external envelope that they acquire after budding through the plasma membrane (Andrés et al. 2001). The extracellular ASFV particle is 260 to 300nm in diameter (Wang et al. 2019).

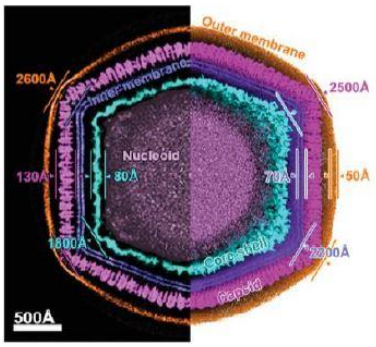

Figure 10. ASFV viral particle structure (Wang et al. 2019)

Nearly half of the ASFV genome has not been experimentally checked at the protein level, and most of it has not been functionally characterized well enough (Zhu et al. 2020), so there is still a considerable amount of work to be done in this field. The number of open reading frames (ORFs) encoded by the ASFV genome varies between 151 and 167, depending on the viral isolate (Galindo and Alonso 2017). The differences in genome length and gene number are mainly due to the gain or loss of ORFs from the multigene families (MGFs) encoded by the virus (Dixon et al. 2012). A striking feature of the ASFV genome is its many MGFs, comprising approximately 30% of it. The multigene family proteins are a large group of proteins with undefined functions in ASFV and are the most abundant genes in its genome. There are 5 ASFV MGFs: MGF-100, MGF-110, MGF-300, MGF-360, and MGF-505. They are mainly encoded at either 3' or 5' ends of the genome, and most of them have copies at both ends. MGF proteins within the same MGF family tend to have similar structures, locations, and functions. Reportedly, they play important roles in diverse stages of viral infection associated with virulence, antigenicity, and immune escape. The large investment in coding for multiple copies of several MGFs implies that they probably confer a selective advantage to the virus and may represent mechanisms of immune evasion (Dixon et al. 2012; Zhu et al. 2020).

1.4.2. Pathogenesis and host immune response

ASF's pathogenesis is very complex. ASFV primary target cells are monocytes and macrophages, although replication in dendritic cells has also been reported (Salguero et al. 2002; Dixon et al. 2020). In Africa's original wild suid hosts, the infection with ASFV results in mild or no clinical signs with long-term persistent infections. They are mostly asymptomatic and have low viremia titers. The virus' replication is efficiently contained, resulting in low levels of cell and tissue damage (Jori and Bastos 2009; Jori et al. 2013). However, in wild boars and domestic pigs, clinical courses go from peracute to chronic forms. Peracute clinical form is considered to be a hemorrhagic fever, characterized by massive apoptosis of lymphocytes and hemorrhagic pathology with extensive vascular damage. This form of disease has a fatality rate approaching 100% (Salguero 2020). Chronic forms of disease caused by low virulent isolates are characterized by lower fatality rates and by the absence of vascular lesions (Sánchez-Vizcaíno et al. 2015; Dixon et al. 2020).

Immunity to ASFV is highly elaborate and remains ill defined. It is mediated by multiple mechanisms of the innate and adaptive immune systems, both humoral and cellular, and by the recognition of multiple antigenic targets (Correia et al. 2013). Virus replication in monocytes and macrophages induces activation of this cell population, increasing the secretion of pro-inflammatory cytokines at the early stages of infection, leading to the "cytokine storm" responsible for a massive induction of lymphocytes apoptosis (Salguero et al. 2005).

1.4.3. ASFV strategies for immune evasion

The replication of ASFV in macrophages, a fundamental cell in both innate and adaptive immune systems, constitutes a serious threat to a virus. Given that the IFN system is a key player against viruses, in particular acute and persistent ones, ASFV must have evolved multiple genes for manipulating this immune response. Furthermore, ASFV is adapted to infect both vertebrate and invertebrate (soft ticks) hosts. Only the innate response is common to these two hosts, particularly cytokine responses, so one expects the virus to have developed a strategy to evade this system, allowing it to succeed in both hosts (Correia et al. 2013). Indeed, in macrophages infected with virulent ASFV, the interferon response is suppressed; and many studies to date have identified several genes that can interfere with IFNs induction and IFNs signalling pathway (Correia et al. 2013; Dixon et al. 2019).

The multigene families MGF-360 and MGF-505 include multiple genes capable of this function, such as A238L, A276R, A528R, I329L, MGF-505-7R, and MGF-360-12L. A238L was one of the first to be described and has a domain with homology to IκB. It interacts with the p65 subunit of the NF-κB transcription factor, replacing IκB after its degradation, and thereby inhibiting NF-κB activation and translocation to the nucleus (Powell et al. 1996). A276R impairs the induction of IFN-β through targeting IRF3, but not IRF7, in an NF-κB independent manner

(Correia et al. 2013). A528R has evolved to inhibit IFN induction by inhibiting the NF-κB transcription factor and also to reduce the impact of both types I and II IFNs (Correia et al. 2013). I329L is a viral TLR3 antagonist, so it inhibits TLR3-mediated activation of NF-κB and thus reduces the effectiveness of the host type I IFN antivirus response. This inhibition is at the level of TRIF before the bifurcation into the TRAF3/TRAF6 pathway (de Oliveira et al. 2011; Correia et al. 2013). MGF505-7R efficiently inhibits IFN-β promoter activation via the cGAS-STING pathway (a signalling pathway that follows PAMP recognition by a cytosolic DNA receptor). Thus, it blocks dsDNA-triggered induction of downstream antiviral genes (Li et al. 2021). Lastly, MGF-360-12L binds to nuclear transport proteins, blocking NF-κB nuclear translocation in an IκB independent way (Zhuo et al. 2020).

Deleting these proteins from the wild type virus may strengthen the host IFN response and constitute a good starting point to the development of an attenuated live virus vaccine. It could restrict virus spread after the initial infection, thereby allowing for the proper development of a protective adaptive immune response.

1.5. ASFV DP146L

The search for ASFV proteins capable of evading the host's immune response is still ongoing. Using bioinformatic screening of the ASFV genome, the ORF DP146L was recently identified by Instituto Gulbenkian de Ciência (IGC) Infection and Immunity Group, led by Michael Parkhouse, as a new potential immune evasion gene (unpublished work). DP146L, also known as MGF-100-3L, belongs to the ASFV MGF-100. It encodes for a protein with 146 amino acids, 438 nucleotides, with a molecular weight estimated at 17kDa. Bioinformatic analysis of DP146L shows strong evidence of structural homology to an SH2 domain (Figure 11).

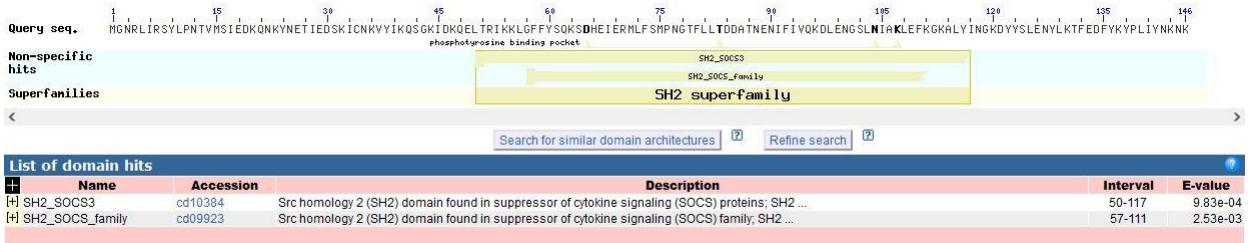

Figure 11. DP146L homology with SH2 superfamily. Results taken from the "Conserved Domains" program, at NCBI, showing a putative SH2 domain in the amino acid sequence of DP146L (National Center for Biotechnology Information (NCBI) 2021)

The SH2 domain is a short (approximately 100 amino acid) conserved structure present in many intracellular signalling proteins. It is known to recognize phosphorylated tyrosine residues on target proteins. Phosphorylation of protein tyrosine residues is an important

mechanism for the formation of signalling complexes. Thus, the SH2 domain plays a crucial role in several cellular mechanisms, including signal transduction pathways stimulated by cytokines (Russell et al. 1992; Murphy and Weaver 2017). Notably, the SH2 domain is present in the above described SOCS proteins.

For that reason, DP146L is a putative suppressor of cytokine responses (SOCS). It can possibly have a role as a modulator of host pro-inflammatory responses, inhibiting the production of pro-inflammatory cytokines by degrading components of both the induction and impact IFN pathways. If this hypothesis is confirmed, DP146L may provide a novel tool for (1) The construction of a DP146L deletion mutant vaccine for ASV and also for (2) The downregulation of the inflammatory response in general.

2. OBJECTIVES

The aim of this work is the preliminary study of the yet unstudied ASFV DP146L. The primary objective is the functional analysis of DP146L through luciferase reporter assays to assess its impact in the NF-κB signalling pathway. A parallel goal is the single-point mutation of the Phe92 DP146L protein (pDP146L)'s amino acid, by site-directed mutagenesis, in order to extend protein's functional analysis through luciferase assays.

A secondary objective of this dissertation is the analysis of the cellular location of the transfected DP146L through immunofluorescence assay, and the evaluation of its cellular expression by western blot.

Finally, this dissertation also aims the pDP146L production and purification, as a resource for future studies on this protein.

3. MATERIALS AND METHODS

3.1. Cell cultures

Human embryonic kidney 293T (HEK 293T) cells were maintained in Dulbecco's modified Eagle's medium (DMEM) (BioWest) supplemented with 10% (v/v) fetal bovine serum (FBS) and 1% (v/v) penicillin-streptomycin (Pen/Strep). African green monkey Vero cells were maintained in Dulbecco's modified Eagle's medium (DMEM) (Gibco) supplemented with 10% FBS, 1% Pen/Strep, and 1% non-essential amino acids (NEAA). All cell lines were maintained in a controlled environment, with 5% CO_2, at 37ºC.

3.2. Plasmids

The pcDNA3 empty expression vector, in frame with an haemagglutinin (HA) tag, was used as a negative control in immunofluorescence, western blot and luciferase reporter assays. The ASFV ORF K205R cloned into the pcDNA3-HA tag vector was used as a positive

control in immunofluorescence, and western blot assays. The ASFV ORF DP146L cloned into the pcDNA3-HA tag vector was the plasmid vector of the gene of study used in all assays (schematic representation in Annex 2). All these three plasmids were previously constructed by Dr. Sílvia Correia, the co-promoter of this dissertation, at the IGC Infection and Immunity group, led by Michael Parkhouse, since the work here described was done in collaboration between the two groups.

The ASFV ORF P1192R cloned into the pIC113 plasmid vector, in frame with a green fluorescence protein (GFP) ORF (pIC113-p1192R-GFP), was used as a positive control in the immunofluorescence assay and was a gift from Dr. João Coelho.

The pCMVβ plasmid contains a β-galactosidase gene under the control of human cytomegalovirus immediate early promoter, and was used as an internal transfection control in luciferase reporter assays. The reporter plasmid for the PRD II promoter [p(PRD2)5tkΔ(-39)lucter] (pPRD II) was used to study the effect of the DP146L gene on PRD II dependent expression in luciferase assays. Both plasmids were also previously existing constructs from Dr. Sílvia Correia and were used in the frame of the collaboration mentioned above.

The pET-28a cloning plasmid vector was used for cloning DP146L for protein production and purification, and it was kindly provided by Dr. Pedro Bule. The pJET1.2/blunt cloning vector used as an intermediate in the cloning procedure is from Thermo Scientific.

The pcDNA3-HA-DP146L$_{mut}$ single point mutated plasmid, used for luciferase assay, was produced using Q5 Site-Directed Mutagenesis kit (New England Biolabs (NEB)).

3.3. Competent cells preparation

To prepare competent DH5α *E. coli* cells, a glycerol stock was inoculated in Lysogeny Broth (LB) liquid medium and incubated overnight in an orbital shaker at 37ºC, with constant shaking (180 rotations per minute (rpm)). The next day, 3mL of culture were inoculated in 300mL LB medium pre-heated to 37ºC and incubated until it reached an optical density (OD) of 600. Then, the culture was centrifuged at 7500g, for 2 minutes, at 4ºC. The supernatant was removed, and the pellet was resuspended in 150mL of cold $MgCl_2$ 0,1M. Centrifugation was repeated, and then the pellet was resuspended in 150mL of cold $CaCl_2$ 0,1M. After 20 minutes of incubation on ice, the same centrifugation was repeated. The supernatant was discarded, and the pellet was finally resuspended in 6mL of $CaCl_2$ 0.1M/Glycerol 15% mix. Competent cells were aliquoted and stored at -80ºC.

3.4. DNA plasmids amplification

All plasmids used in this dissertation were produced by transformation in DH5α *E. coli* competent cells from 3.3. subchapter. In a transformation tube (1.5mL), 50µL of DH5α cells

and 1µL of the desired plasmid DNA were mixed. The mixture was incubated on ice for 30 minutes, followed by a 42ºC heat shock for precisely 90 seconds, and a new incubation of 2 minutes on ice. Then, 950µL of SOC medium was added, and the mixture was incubated for 1 hour at 37ºC, 220rpm, for the bacteria to recover and express the antibiotic resistance marker encoded by the plasmid. After incubation, centrifugation at 4700g for 2 minutes was performed, and the pellet was resuspended in 150µL of SOC medium. The volume was plated onto LB agar plates with the appropriate antibiotic (100µg/mL), and plates were stored at 37ºC overnight (12-16 hours). Single colonies were cultured on liquid LB medium, with the same antibiotic, overnight at 37ºC, at 220rpm. The resulting plasmid DNA was extracted from cells using a Miniprep or Midiprep kit (NZYTech), following the manufacturer's instructions.

3.5. Analysis of DP146L expression

3.5.1. Transfection of mammalian cells

Two transfection reagents were used as an attempt to reach an ideal transfection efficiency: TransIT®-LT1 Transfection Reagent (Mirus) and Lipofectamine® LTX & PLUS Transfection Reagent (Thermo Scientific).

Cells were seeded on a 6-well plate ($2.5x10^5$cells/well for western blot and $2x10^5$cells/well for immunofluorescence) in 2mL DMEM medium and were incubated overnight. For immunofluorescence, three sterile coverslips were placed in each well. The next day, cells were transfected with 2.5µg of plasmid DNA. pcDNA3-HA-K205R and pIC113-p1192R-GFP were used for positive control; empty pcDNA3 for negative control; and pcDNA3-HA-DP146L was the plasmid of study. A plate well of untransfected cells was always left as an additional negative control.

TransIT®-LT1: a mix was made with 7.5µL of TransIT reagent and 250µL of Opti-MEM (serum-free medium) (Gibco) and was incubated at room temperature (RT) for 5 minutes. The DNA plasmid was added, and the final mixture was incubated at RT for 30 minutes. After incubation, the mix was added dropwise to the cells, and cells were incubated for 48h. This protocol is for a 1:3 DNA:transfection reagent ratio. We also tested 1:4 and 1:5 ratios, increasing the transfection reagent volume to 10µL and 12.5µL, respectively.

Lipofectamine® LTX & PLUS: a mix was made with 50µL of Opti-MEM and 12.5µL of Lipofectamine reagent; and a DNA mix was made with 50µL of Opti-MEM, 2.5µL of PLUS reagent, and the DNA plasmid. Diluted DNA was added to diluted Lipofectamine, and the final mixture was incubated for 5 minutes at RT. Then the DNA-lipid complex was added to cells and was incubated for 48h. In order to optimize transfection efficiency, we also increased the transfection reagent volume to 15.5µL.

3.5.2. Immunofluorescence assay

For immunofluorescence assays, both HEK293T and Vero cell lines were transfected with TransIT®-LT1, and HEK293T was also transfected with Lipofectamine® LTX & PLUS. Different DNA:transfection reagent ratios were tested in order to optimize protocols.

Forty-eight hours after transfection, the cells were washed with PBS and fixed with 4% paraformaldehyde for 30 minutes. After washing with PBS, the cells were permeabilized with PBS + 0.1% Triton X-100 for 15 minutes. The cells were washed with PBS + 0.05% Tween20 and then blocked with PBS + 0.05% Tween20 + 5% normal goat serum (blocking buffer) for 30 minutes. After washing with PBS + 0.05% Tween20 again, two staining protocols were tested: (1) coverslips were incubated for 1 hour with anti-HA high-affinity antibody (Roche) diluted 1:100 in blocking buffer, then washed two times with PBS + 0.05% Tween20, and incubated again, for 1 hour with Alexa Fluor 594 goat anti-mouse secondary antibody (Life Technologies) diluted 1:1000 in blocking buffer. (2) In parallel, another coverslip was incubated for 1h30 with conjugated antibody anti-HA FITC diluted 1:100 in blocking buffer. After rewashing two times and incubating with DAPI (200ng/µL) for 5 minutes, coverslips were mounted on Slow Fade mounting medium (Life Technologies) and examined under a Leica DMR fluorescence microscope.

For cells transfected with the pIC113-p1192R-GFP plasmid, the procedure is slightly different. After fixation and washing with PBS, the cells were permeabilized for only 5 minutes, and coverslips followed directly for incubation with DAPI and montage.

3.5.3. Western blot

For western blot assays, only HEK293T cells were transfected with the TransIT®-LT1 transfection reagent, with the 1:3 ratio.

Forty-eight hours after transfection, the cells were placed on ice and lysed. They were scraped from wells with 1mL of PSB, transferred to pre-cooled microcentrifuge tubes, and centrifugated at 300g, for 5 minutes, at 4ºC. The supernatant was discarded, and the pellet was resuspended in 30µL Lysis Buffer (Cell Signalling Lysis Buffer, from Sigma) supplemented with PMSF protease inhibitor. The cell lysate was clarified by centrifugation, for 20 minutes, at 16000g, 4ºC. The supernatant was transferred to a clean tube, and 6x SDS sample buffer (0.35 M Tris-HCl pH 6.8, 10.28% (w/v) SDS, 36% (v/v) glycerol, 0.6 M DTT, 0.012% (w/v) bromophenol blue) was added. Samples were boiled at 95ºC for 5 minutes and run in a 14% sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Separated proteins were transferred to a polyvinylidene difluoride (PVDF) membrane (Bio-rad) and blocked with 5% (w/v) non-fat milk powder for 1 hour at RT. The membrane was probed with an anti-HA-HRP conjugated antibody (Thermo Scientific) at 4ºC overnight. According to the manufacturer's instructions, the membrane was developed by chemiluminescence detection

with Immobilon® Forte Western HRP substrate (Millipore), and exposed at ChemiDoc XRS+ (Bio-rad) imaging system to visualize bands.

3.6. DP146L functional analysis

3.6.1. Luciferase reporter gene assay

HEK-293T cells ($6x10^4$cells/well, in a 24 well plate) were co-transfected with 100ng of pPRD II luciferase reporter plasmid, 25ng of pCMVβ internal control plasmid, and 300ng of either pcDNA3-HA-DP146L or the empty pcDNA3-HA. Transfections were performed according to the TransIT-LT1 and Lipofectamine LTX & PLUS protocols exposed on 3.5.1. subchapter. Forty-eight hours post-transfection, the cells were either stimulated with 100ng/mL IL-1 (PeproTech) or 100ng/mL TNFα (PeproTech) for five hours, or left untreated. Each transfection was made in triplicate. After stimulation, the cells were lysed with 100µL luciferase lysis buffer (1/10 potassium phosphate buffer, pH 7.8 (18.16mL K_2HPO_4 1M, 1.84 KH_2PO_4 1M, water up to 200mL), 1% Triton X-100, 1mM DTT, 2mM EDTA).

Luciferase activity was measured using the luciferase assay system kit (Promega) according to the manufacturer's protocol. The β-galactosidase activity was measured using the Galacton-Plus kit (Thermo Scientific). Variations in transfection efficiency between cells were normalized by dividing luciferase values by β-galactosidase values.

3.6.2. Site-directed mutagenesis (SDM)

The 92nd amino acid of DP146L was mutated from a Phenylalanine (Phe) to an Aspartate (Asp), by substituting the nucleotide TTT into GAT. This single point mutated plasmid was produced using the Q5 Site-Directed Mutagenesis kit (NEB). The particularity of this kit is the type of primers it requires. They are non-overlapping primers, with the desired nucleotide change in the center of the forward primer, including at least ten complementary nucleotides on the 3' side of the mutation; the reverse primer has to be designed so that the 5' ends of the two primers anneal back-to-back. The primers used for this SDM were designed using the NEBaseChanger™ online design software, and were: forward – TGAAAATATA**GAT**ATTGTTCAAAAAGATCTGGAAAACG (mutated nucleotides highlighted in bold); and reverse – TTGGTTGCATCGTCCGTA.

3.6.2.1. Exponential amplification

The Q5 Site-Directed Mutagenesis kit protocol is divided into three steps. The first one is the exponential amplification by polymerase chain reaction (PCR). A PCR mixture was assembled with 12.5µL Q5 Hot Start High-Fidelity 2X Master Mix, 1.25µL 10µM forward primer, 10µM reverse primer, 1µL 10ng/µL of template DNA (pcDNA3-HA-DP146L), and 9µL of

nuclease-free water. The tubes were placed on a thermocycler and subjected to the following conditions: initial denaturation at 98ºC for 30 seconds, followed by 25 cycles of denaturation at 98ºC for 10 seconds, annealing at 61ºC for 30 seconds, and extension at 72ºC for 150 seconds. Finally, a final extension of 2 minutes at 72ºC.

3.6.2.2. Kinase, Ligase, DpnI (KLD) treatment

The second step is treatment with a mixture of Kinase, Ligase, and DpnI (KLD) enzymes. This step allows phosphorylation and ligation of the PCR product at once and degradation of template DNA by the DpnI restriction enzyme. This step consists of the assembly of 1µL PCR product, 5µL 2X KLD reaction buffer, 1µL 10X KLD enzyme mix, and 3µL of nuclease-free water, and incubation at RT for 5 minutes.

3.6.2.3. Plasmid amplification and DNA sequencing

The last step of SDM is the amplification of the final plasmid. It was transformed into DH5α *E. coli* competent cells 5µL of the KLD mix (as explained in 3.4. subchapter), and plasmid DNA was extracted using the NZYMiniprep kit. Sanger sequencing of DNA plasmid for confirmation of the single-point mutated DP146L was performed by Eurofins Genomics, using a T7 forward primer (TAATACGACTCACTATAGGG) and a pcDNA3_rev primer (GGCAACTAGAAGGCACAGTC).

After pcDNA3-HA-DP146L$_{mut}$ confirmation, luciferase reporter assay was repeated using this new plasmid, the pcDNA3-HA-DP146L for effect comparison, and the empty pcDNA3 as a negative control. The protocol was the same as described in 3.6.1. subchapter.

3.7. Cloning DP146L into pET28a plasmid vector

The DP146L ORF was cloned into the pET28a cloning vector in frame with an N-terminal His-tag originally present in the plasmid. After multiple unsuccessful attempts to directly clone DP146L into pET28a, DP146L was firstly cloned into pJET1.2/blunt cloning vector as an intermediate step for final cloning into pET28a.

3.7.1. PCR amplification

Primers for PCR amplification of the DP146L gene from the pcDNA3-HA-DP146L plasmid vector were designed containing NdeI and XhoI specific restriction sites. The primers used were: forward – GGAATTC**CATATG**GGGAACCGTCTTATCAG; and reverse – GCCG**CTCGAG**TTACTTATTTTTATTGTAAATTAGTGGGTATTTG (restriction sites highlighted in bold). A PCR mixture was set containing 20µL 5X Phusion HF buffer, 2mM dNTPs, 5µM forward primer, 5µM reverse primer, 1µL Phusion DNA polymerase, 20ng of

template DNA, and nuclease-free water up to 100µL. The tubes were placed on a thermocycler and subjected to the following conditions: initial denaturation at 98ºC for 2 minutes, followed by 30 cycles of denaturation at 98ºC for 30 seconds, annealing at 62ºC for 30 seconds (the ten first cycles), and 68ºC for 30 seconds (the following 20 cycles), and extension at 72ºC for 60 seconds. Finally, an extension of 5 minutes at 72ºC. The PCR product was migrated in a 1% agarose gel electrophoresis in TAE buffer, with GreenSafe (4µL/100mL) (NZYTech), to verify the correct amplification of the fragment.

The amplification was followed by DNA precipitation of the PCR product. 1:10 volume of 3M sodium acetate (pH 5.2), and 2.5 volumes of ice-cold 100% ethanol were added to the PCR product. The mixture was incubated overnight at -20ºC to promote DNA precipitation. The next day, it was centrifuged at 15000g, for 15 minutes, at 4ºC. The supernatant was removed, and the pellet was air-dried for about 10 minutes to remove any residual fluid. The pellet was resuspended in 50µL of Milli-Q water. The final sample was analyzed on NanoDrop™ One (Thermo Scientific), for concentration measurement, and A260/230 and A260/280 ratios verification. Such rations are DNA quality evaluation parameters: A260/230 is used for the detection of an issue in the DNA extraction procedure, usually due to the presence of organic compounds such as phenol, guanidine, or glycogen in the DNA sample; A260/280 indicates the purity of the extracted DNA, which can be impaired by the presence of both RNA or proteins (Koetsier and Cantor 2019).

3.7.2. Cloning into pJET plasmid vector

After fragment amplification and DNA purification, DP146L was cloned into a pJET1.2/blunt cloning vector (Thermo Scientific). According to the manufacturer's instructions, 25ng of DNA are necessary for each 500bp fragment. Since the DP146L fragment is 461bp long (including restriction sites), 23.05ng are needed. However, a 5µL reaction was performed instead of the 20µL suggested in the protocol. This way, 6ng of DNA fragment was used in the ligation reaction. A mixture was set with 2.5µL 2X reaction buffer, 1µL of fragment DNA (6ng/µL), 0.25µL pJET cloning vector, 0.25µL T4 DNA ligase, and 1µL of milli-Q water. It was incubated for 5 minutes at RT and was then transformed into DH5α competent cells (as described in 3.4. subchapter).

After transformation, a colony PCR was performed to confirm pJET cloning. A mixture was assembled with 1µL 10X Taq buffer, 0.2µL dNTPs mix, 0.5µL pJET forward sequencing primer, 0.5µL pJET reverse sequencing primer, 0.05µL DreamTaq DNA polymerase, and 6.75µL of milli-Q water. A colony was picked and inoculated in the PCR mix. A PCR reaction was initiated with an initial denaturation at 95ºC for 3 minutes, followed by 25 cycles of denaturation at 95ºC for 30 seconds, annealing at 55ºC for 30 seconds, and extension at 72ºC

for 1 minute. A final extension was performed for 5 minutes at 72ºC. The PCR products were analyzed in a 1% agarose gel electrophoresis as described in the previous subchapter.

After confirmation of proper cloning, the plasmid DNA from a positive colony was extracted using an NZYMiniprep kit.

3.7.3. Restriction reaction

For final cloning into pET28a cloning vector, DP146L gene was removed from pJET plasmid by restriction with the enzymes NdeI and XhoI. However, since pJET plasmid has an XhoI restriction site near the NdeI restriction site of the fragment, these restrictions were performed sequentially in two steps: first, the restriction with NdeI and electrophoresis for linearization confirmation; and next, the digestion with XhoI. Restriction reactions included 5µL 10X CutSmart buffer, 1µL restriction enzyme, 5µg pJET plasmid, and nuclease-free water up to 50µL. Incubation time was 2 hours for NdeI and overnight for XhoI, both at 37ºC.

The plasmid pET28a was also linearized by restriction with NdeI and XhoI simultaneously, for 2 hours, at 37ºC.

3.7.4. DNA purification

In order to purify both the pET28a plasmid vector and the DP146L insert, after restriction the DNAs were run with 6X purple gel loading dye (NEB) on a 1% agarose gel with GreenSafe. Appropriate bands were excised under UV light illumination in a transilluminator using a clean scalpel. Both insert and vector DNAs were purified using the NZYGelpure kit (NZYTech), following the manufacturer's instructions, and DNA concentrations were measured on Nanodrop.

3.7.5. Ligation

After DNA purification, the DP146L fragment was ligated with the pET28a vector in an exact 1:3 vector:insert ratio, calculated using the online ligation calculator NEBioCalculator. The ligation mixture included 2µL 10X T4 DNA ligase buffer, 1µL T4 DNA ligase, 50ng of linearized pET28a, 12.59ng of DP146L fragment, and nuclease-free water up to 20µL. It was incubated for 5 hours at room temperature. Only 10µL of the ligation product was transformed in DH5α as described in 3.4. subchapter.

3.7.6. Confirmation of pET28a-DP146L clones

For cloning tracking, a colony PCR was performed as described in 3.7.2. subchapter, using the T7 forward and T7 reverse standard primers, and PCR products were run in an

agarose gel electrophoresis, allowing identification of possibly positive clones. Plasmid DNA extracts from these clones were then obtained using the NZYMiniprep kit.

Final confirmation of proper cloning and correct DNA sequence was done through Sanger sequencing by Eurofins Genomics, using T7 forward and T7 reverse primers.

3.8. Small-scale protein production and purification

3.8.1. Bacterial transformation and induction of expression

The pET28a-DP146L-His-tag was first transformed in BL21 *E. coli* competent cells kindly provided by Dr. Pedro Bule. A mixture of 100µL of BL21 cells and 2µL of plasmid DNA, in a transformation tube (1.5mL), was incubated on ice for 30 minutes, followed by a 42ºC heat shock for precisely 45 seconds, and a new incubation of 2 minutes on ice. It was added 500µL of SOC medium (prewarmed to 37ºC), and the mixture was incubated for 1 hour at 37ºC, 200rpm. After incubation, centrifugation at 1000g for 1 minute at 20ºC was performed, and the pellet was resuspended in 100µL of SOC medium. The volume was spread on solid LB agar plates supplemented with kanamycin (100µg/mL), and plates were stored at 37ºC overnight (12-16 hours).

The next day, an auto-induction medium was prepared by adding 50g/L auto-induction LB broth (NZYTech), 10g/L glycerol 100%, Milli-Q water up to 1L, and kanamycin (1:1000 dilution). This auto-induction medium is used for low production plasmids and does not require the addition of IPTG nor consequently to monitor cell growth. Four lines of colonies were scraped from the plate and inoculated into the medium. The liquid culture was incubated for 3 hours, at 37ºC, 220rpm, and then overnight, at 19ºC, 120rpm, for gene expression and thus protein production.

3.8.2. Protein purification

The day after induction, cells were harvested by centrifugation at 5000g, for 15 minutes, at 4ºC. The pellet was resuspended in 10mL of 10mM imidazole and sonicated for 5 minutes for bacterial lysis. The lysate was centrifuged for 30 minutes at 17000g, 4ºC. Some of the pellet (to which 100µL of 100mM imidazole has been added) and 100µL of the supernatant were collected to be analyzed later. The rest of the pellet was discarded. The rest of the supernatant was filtered with a 0.45µm filter.

Protein purification was done by affinity chromatography with nickel ions (Ni^{2+}), since His-tag binds to nickel, and thus DP146L protein becomes retained in the column. A His GraviTrap Column (GE Healthcare, Life Sciences) was prepared by washing with 10mL of 10mM imidazole solution. The filtered supernatant was applied to the column, followed by three different concentration imidazole solutions: 10mM, 35mM, and 60mM. For further analysis,

100µL of each flow-through were collected and stored. The imidazole concentration gradient allows the column to release nonspecifically bound proteins before the final delivery of our protein of interest. Finally, 6mL of 300mM imidazole solution were applied to the column, and 1mL of the eluate was collected to a different collection tube. Imidazole competes with the His-tag and clings to nickel, causing the His-tag protein to be eluted.

3.8.3. Protein analysis in polyacrylamide gel

The protein produced was analyzed by electrophoresis in a 14% polyacrylamide gel (SDS-PAGE). The screening samples were: the initial supernatant, the diluted pellet, the three flow-throughs, and the two more concentrated final eluates. The SDS-PAGE protocol was the same as described at the beginning of 3.5.3. subchapter.

For nonspecific detection of proteins in the SDS-PAGE gel, Silver Blue staining was performed (Silver Blue is an adaption of the Coomassie Blue staining). For that purpose, the gel was put in Elix water and boiled in the microwave for 2 minutes. Then, the water was removed, and Silver Blue staining solution was added. The gel was boiled in the staining solution for 1-2 minutes and then incubated with agitation for 10 minutes at RT. Finally, the staining solution was removed, and the gel was incubated overnight in Elix water at RT. On the following day, the gel was exposed at ChemiDoc XRS+ to visualize bands.

4. RESULTS

4.1. Analysis of cellular expression and localization of transfected DP146L

The first task of this dissertation consisted on the transfection of DP146L into mammalian cells to investigate its cellular localization. The pcDNA3-HA-DP146L plasmid was transfected into HEK293T cells using the TransIT®-LT1 transfection reagent. The pcDNA3-HA empty vector was transfected in parallel as a negative control. The cellular distribution of the fusion HA-DP146L was assessed by immunofluorescence assay, using anti-HA antibodies for staining, both in a direct (FITC conjugated primary antibody) and indirect (primary and secondary antibodies) techniques. As expected, no fluorescence signal was detected in the pcDNA3-HA transfected cells. However, no fluorescence significant differences were observed between DP146L and the negative control (Figure 12). Indeed, some DP146L transfected cells became fluorescent, but they were very rare (less than one transfected cell per microscope field), which does not allow for any conclusions to de drawn.

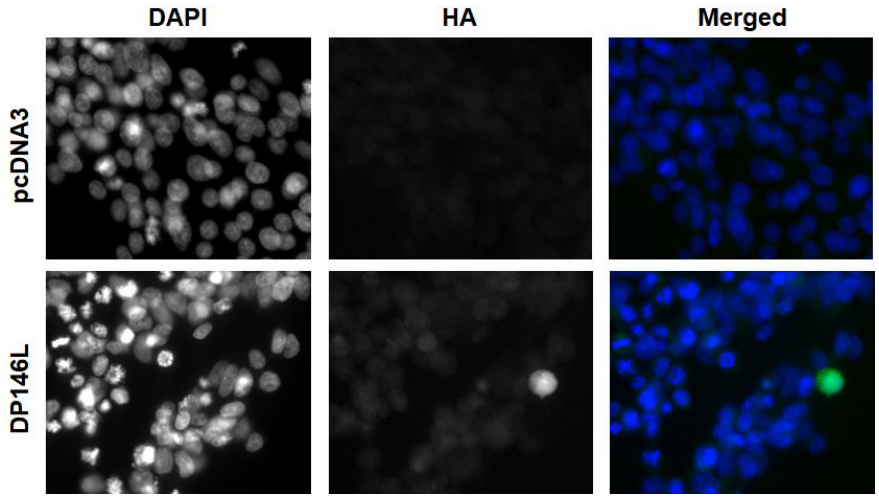

Figure 12. Immunofluorescence assay results of the transfected pcDNA3-HA-DP146L and its negative control. HEK293T cells were transfected with 2.5µg of empty pcDNA3-HA (upper panels) or 2.5µg of pcDNA3-HA-DP146L (lower panels), with the TransIT®-LT1 transfection reagent (Mirus). The DP146L protein was stained using an anti-HA FITC conjugated antibody diluted 1:100 in blocking buffer. DAPI solution was used for nuclei staining. Image taken at 200x magnification

To address this issue, one must understand whether the problem is specific to pcDNA3-HA-DP146L or to the transfection technique itself. For this purpose, two well-known plasmids were selected as positive controls to be compared with pcDNA3-HA-DP146L, namely pcDNA3-HA-K205R and pIC113-P1192R-GFP. Both positive control plasmids were successfully transfected and, as expected, both localized in the cytoplasm (Figure 13A). However, their transfection efficiency was not as good as previously obtained in the lab, and pcDNA3-HA-DP146L fluorescence cells were still very rarely observed. We may thus assume that the transfection process is defective.

A transfection optimization process was then initiated by testing different transfection conditions. Two cell lines were tested: HEK293T and Vero cells. The plasmid used to test transfection efficiency was pIC113-P1192R-GFP, since it does not require the use of antibodies, thereby reducing the number of variables in study. The transfection efficiency (the number of transfected cells relative to the total number of cells) was actually different between cell lines, with 50% of transfected cells in HEK293T and 20% in Vero cells (Figure 13B). HEK293T cells were therefore chosen for subsequent studies.

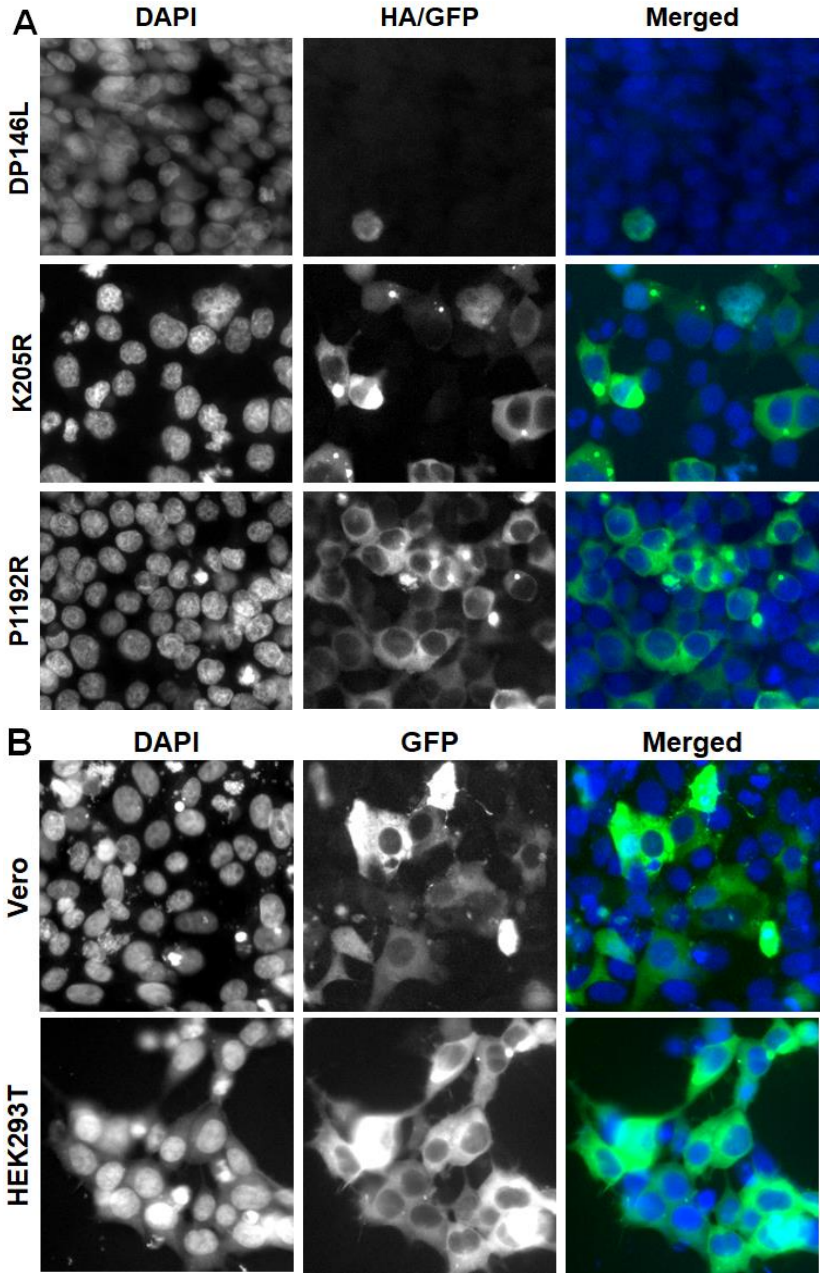

Figure 13. A. Immunofluorescence assay results comparing transfection efficiency between different plasmids. HEK293T cells were transfected with 2.5µg of pcDNA3-HA-DP146L, pcDNA3-HA-K205R, or pIC-P1192R-GFP plasmids, using the TransIT®-LT1 transfection reagent (Mirus) 1:3 DNA:reagent ratio. The first two proteins were stained using an anti-HA FITC conjugated antibody diluted 1:100 in blocking buffer. DAPI solution was used for nuclei staining. Image taken at 200x magnification. **B. Fluorescence microscopy results comparing transfection efficiency between HEK293T and Vero cell lines.** Cells were transfected with 2.5µg of pIC-P1192R-GFP plasmid using the TransIT®-LT1 transfection reagent (Mirus) 1:3 DNA:reagent ratio. DAPI solution was used for nuclei staining. Image taken at 200x magnification

Proceeding with the optimization of the transfection process, a second transfection reagent was tested: Lipofectamine® LTX & PLUS. Different DNA:reagent ratios were also tested for both transfection reagents, as explained in section 3.5.1. Results for control plasmids

suggested slightly higher transfection efficiencies for Lipofectamine® LTX & PLUS than for TransIT®-LT1, although the number of assay repetitions was not enough for an exact quantification. However, the pcDNA3-HA-DP146L transfected with this reagent was still very poorly seen by immunofluorescence (results not shown). Therefore, no transfection conditions were identified that resulted in a significant and conclusive visualization of this protein by immunofluorescence.

Nevertheless, the problem could rather be with the immunofluorescence detection instead of with the transfection process itself. To deeper investigate whether pcDNA3-HA-DP146L was being transfected, a western blot assay was performed. The pcDNA3-HA empty vector was used as a negative control, and pcDNA3-HA-K205R as positive control. For detection, an HRP conjugated anti-HA tag antibody was used. As expected, a 25kDa band corresponding to the K205R protein was detected, and no band could be seen in the negative control lane (Figure 14). However, in agreement with the immunofluorescence results, no band could be seen for DP146L either.

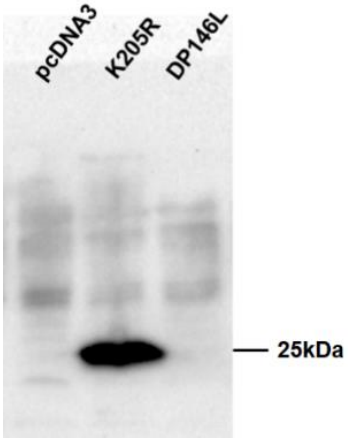

Figure 14. Western blot assay results of the transfected pcDNA3-HA-DP146L and respective controls. HEK293T cells were transfected with 2.5µg of empty pcDNA3-HA (negative control), 2.5µg of pcDNA3-HA-K205R (positive control), or 2.5µg of pcDNA3-HA-DP146L, using the TransIT®-LT1 transfection reagent (Mirus). DP146L and K205R proteins were detected using an anti-HA-HRP conjugated antibody

4.2. The ASFV gene DP146L inhibits the NF-κB signalling pathway

Because of its structural homology with an SH2 domain, it was hypothesized that the pDP146L could interfere with the IFN induction and impact pathways. As a preliminary study for this protein's characterization, this dissertation investigated the pDP146L interference in the NF-κB pathway, since such is present in the IFN induction pathway. To address that question, luciferase gene reporter assays were performed in order to determine if the NF-κB activation in DP146L transfected cells was modified. HEK293T cells were co-transfected either with the pcDNA3-HA-DP146L plasmid or the pcDNA3-HA empty vector (used as a control); a

luciferase reporter plasmid containing the NF-κB binding site of the IFNβ promoter, the PRD II; and the β-Galactosidase internal control plasmid. The cells were then stimulated with TNFα or IL-1 ligands or left untreated (medium). Both Lipofectamine® LTX & PLUS and TransIT®-LT1 were tested in these assays.

As expected, the empty vector control showed significantly increased luciferase activity in stimulated cells when compared to unstimulated ones. Stimulation with TNFα results in greater luciferase activity than with IL-1. We might also observe that, in comparison with the control vector, the pDP146L has a visible impact on the expression dependent of PRD II. There is a clear inhibitory effect on the luciferase activity after TNFα and IL-1 stimulation, with values significantly lower than those of the control plasmid (Figure 15).

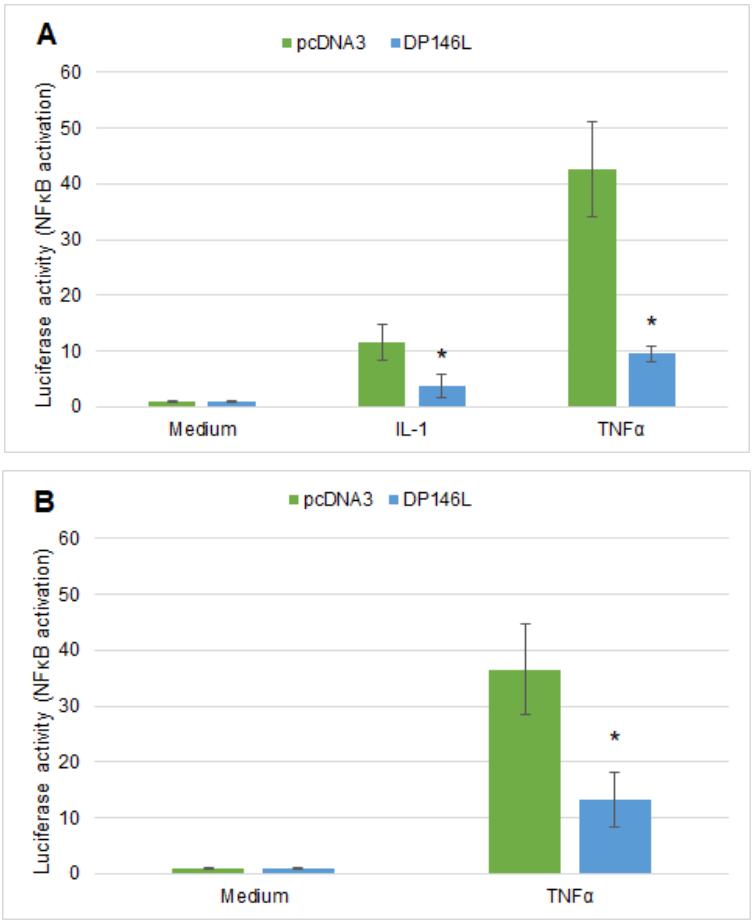

Figure 15. Luciferase reporter gene assay results of pDP146L. HEK293T cells were co-transfected with 100ng of pPRD II luciferase reporter, 25ng of internal control pCMVβ, and 300ng of pcDNA3-HA empty vector or pcDNA3-HA-DP146L. Transfection was performed using **(A)** Lipofectamine® LTX & PLUS (Thermofisher Scientific) or **(B)** TransIT®-LT1 (Mirus). Forty-eight hours post-transfection, the cells were either stimulated with 100ng/mL of IL-1 or TNFα for five hours, or left untreated (medium). Luciferase assay was normalized to β-galactosidase activity as a control for transfection efficiency. Data is expressed as the average ± standard deviation of luciferase activity from triplicate samples. Statistically significant difference was assessed by Student's t-test and is represented as p≤0,05 (*)

4.3. Phe92Asp amino acid single point mutation of DP146L

Several pDP146L amino acids were identified as potentially essential residues for pDP146L function. The phenylalanine at position 92 was the selected amino acid for testing in this preliminary study because of its location in the SH2 homology region (as can be seen in Figure 16). In order to assess its importance, the original phenylalanine amino acid was substituted by a completely different one, codifying for an aspartic acid. The amino acid substitution was performed by site-directed mutagenesis, constructing a mutated pcDNA3-HA-DP146L$_{Phe92Asp}$ plasmid, changing the nucleotides 274-276 from TTT to GAT, which encodes for a single point mutated pDP146L^{Phe92Asp}. The successfully mutated plasmid was confirmed by Sanger sequencing, as can be seen in figure 17.

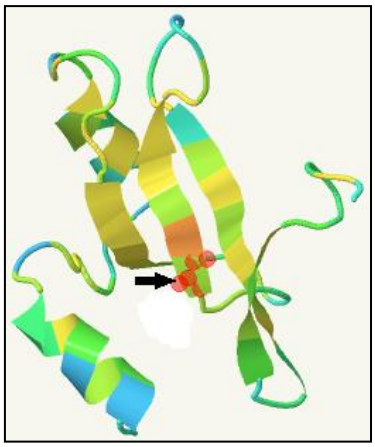

Figure 16. Phe92 amino acid location in the SH2 domain homology region of the pDP146L. Results taken from the "Phyre2" web portal for protein modeling, prediction and analysis, by Structural Bioinformatics Group (Imperial College, London) (Kelley et al. 2015). The figure shows the three-dimensional structure prediction of the pDP146L region of homology with the SH2 domain. The Phe92 amino acid location inside the SH2 homology region is identified in the figure by the black arrow

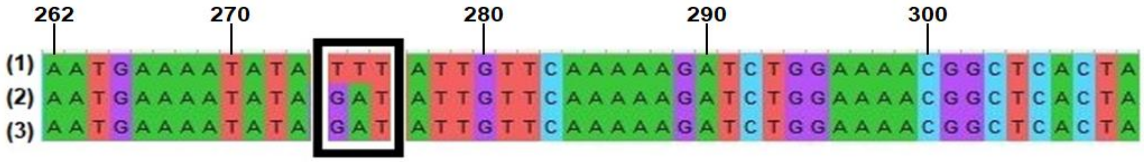

Figure 17. Sanger sequencing results of the F92 amino acid single point mutated pDP146L^{Phe92Asp}. 92nd pDP146L amino acid was mutated by site-directed mutagenesis by replacing a phenylalanine (TTT) with an aspartate (GAT) (nucleotides positions 274-276). The figure exposes the Sanger sequencing results of the region surrounding the amino acid mutation. (1) pcDNA3-HA-DP146L original plasmid sequenced with T7 primer. (2) pcDNA3-HA-DP146L$_{Phe92Asp}$ mutated plasmid sequenced with T7 primer. (3) pcDNA3-HA-DP146L$_{Phe92Asp}$ mutated plasmid sequenced with pcDNA3_rev primer (resulting sequence reversed). The black box highlights the mutated amino acid in each sequence

4.4. The Phe92Asp mutation of pDP146L partially reverses its function

In order to evaluate the impact of the pDP146L^{Phe92Asp} in the NFkB signalling pathway, the luciferase reporter assay (see sub-section 4.2) was repeated with the empty vector control (pcDNA3-HA) and the expression plasmids for pDP146L and mutated pDP146L^{Phe92Asp}. Only the TNFα stimulus was used since it has provided higher luciferase activity values in the 4.2 sub-section, and is thus likely to allow for more straightforward interpretation of the results. HEK293T cells were transfected with Lipofectamine® LTX & PLUS to maximize transfection efficiency.

We can observe that the value of the luciferase activity corresponding to the mutated pDP146L^{Phe92Asp} construct is significantly lower than the empty vector control value (Figure 18). Yet, its value is slightly higher than the one observed for the wild-type pDP146L, being the difference between both statistically significant.

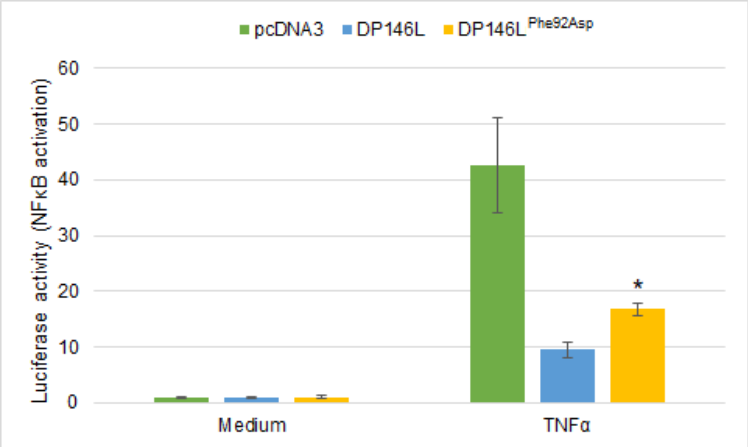

Figure 18. Luciferase reporter gene assay results of the mutated pDP146L^{Phe92Asp}. HEK293T cells were co-transfected with 100ng of pPRD II luciferase reporter, 25ng of internal control pCMVβ, and 300ng of pcDNA3-HA empty vector, pcDNA3-HA-DP146L original plasmid, or pcDNA3-HA-DP146L$_{Phe92Asp}$ mutated plasmid. Transfection was performed using Lipofectamine® LTX & PLUS (Thermofisher Scientific). Forty-eight hours post-transfection, the cells were either stimulated with 100ng/mL of TNFα for five hours, or left untreated (medium). Luciferase assay was normalized to β-galactosidase activity as a control for transfection efficiency. Data is expressed as average ± standard deviation of luciferase activity from triplicate samples. Statistically significant difference was assessed by Student's t-test and is represented as p≤0,05 (*)

4.5. Cloning and pDP146L protein production

The final task of this dissertation was the small-scale protein production and purification of the pDP146L. For this purpose, the DP146L ORF was first cloned into the pET28a cloning vector.

After several unsuccessful attempts of direct cloning into the pET28a plasmid, the CloneJET system was chosen as a new strategy. It is an advanced cloning system for the highly efficient cloning of PCR products. It has the advantage of containing a lethal gene that

is disrupted by the ligation of a DNA insert into the cloning site. This way, only cells with recombinant plasmids are able to propagate, which theoretically means that, after *E. coli* transformation, all colonies are positive clones. Therefore, after DP146L amplification from the pcDNA3-HA-DP146L plasmid by PCR reaction, the fragment was cloned into the pJET1.2/blunt cloning vector as an intermediate step for the final cloning.

After pJET cloning, the DP146L ORF was extracted by restriction and inserted into pET28a on NdeI and XhoI cloning sites. It was followed by an E. coli transformation and subsequent colony PCR of several resulting colonies. Electrophoretic agarose gel run revealed only one band with the expected molecular size of 687bp (Figure 19), corresponding to the only possible positive clone. Plasmid DNA was then extracted and purified, with the fidelity of the resulting clone assessed by Sanger sequencing, which confirmed the proper cloning of DP146L into the pET28a plasmid vector.

Figure 19. Electrophoresis agarose gel run of the colony PCR results. After ligation of DP146L fragment into the pET28a vector on NdeI and XhoI cloning sites, the resulting plasmid DNA was transformed into DH5α E. coli competent cells. A PCR reaction was performed with the several resulted colonies, and PCR products were run in an electrophoresis agarose gel. Of the five resulted bands, only one has the expected molecular size of 687bp (identified with *), the only possibly positive clone. It was then Sanger sequenced, which allowed confirmation of the successful pET28a-DP146L final clone

The DP146L ORF was cloned so that it becomes in frame with the N terminal His-tag already carried by the plasmid vector, which can be used for affinity purification of the cloned protein. pET28a plasmid vector map and schematic representation of the resulting pET28a-DP146L plasmid are available in chapter 8 (Annex 3 and 4, respectively).

For the small-scale pDP146L protein production, a first attempt was done using a common LB medium and inducing protein production with IPTG. However, the cell pellet obtained after the first centrifugation of bacterial culture was very undersized, which allowed the conclusion that an auto-induction LB medium would be beneficial. A second attempt was then performed using the NZYTech auto-induction LB medium, resulting in a considerable-sized pellet. Cells were then lysed, and the pDP146L was purified through nickel affinity

chromatography columns. Final protein analysis was done by electrophoretic SDS-PAGE gel run (Figure 20). The cell debris pellet that resulted from the last centrifugation was also run. Two well-known proteins, produced in similar conditions to pDP146L, were run alongside in the gel as a control.

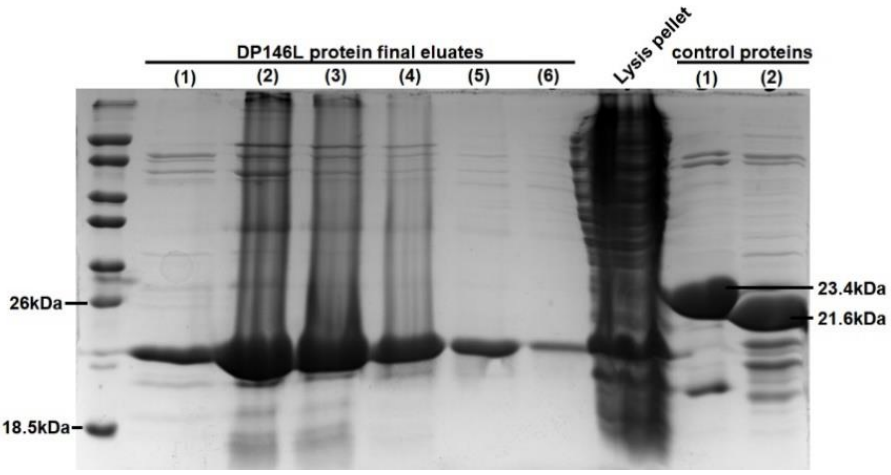

Figure 20. Analysis of the purified pDP146L through electrophoretic SDS-PAGE gel run. The produced DP146L protein was purified using nickel affinity chromatography columns, and the six final eluates ((1) to (6)) were analyzed in a 14% SDS-PAGE gel electrophoretic run followed by silver blue staining. The cell debris pellet resulting from the cell lysis centrifugation was also analyzed (diluted in 100µL of 100mM imidazole). Two well-known proteins were also run (control proteins (1) and (2)) alongside in the gel as a control (23.4kDa and 21.6kDa, respectively). The pDP146L can be observed in all final eluates' lanes in a region compatible with the 19kDa molecular weight predicted by the ProtParam online tool, at ExPASy (Duvaud et al. 2021). The lysis pellet lane also contains a band in the same region, suggesting that some pDP146L remained stuck in the cell debris

As shown in figure 20, the pDP146L is present in all six final eluates, with a higher concentration in eluates 2, 3, and 4. In the pellet lane, we can also see a band with the same molecular weight as pDP146L, suggesting that some of the protein remained stuck in the cell lysis pellet. Comparing the molecular marker bands and the two proteins used as a control, the pDP146L bands seem to have the expected size of 19kDa (pDP146L plus His tag size, according to the ProtParam online tool, at ExPASy online site (Duvaud et al. 2021)). All different concentration imidazole solution flow-throughs stored during the purification process were also run in an SDS-PAGE gel. Neither of them presented a band compatible with the predicted pDP146L molecular weight (result not shown).

5. DISCUSSION

5.1. An unsuccessful attempt to identify the cellular location of pDP146L

This dissertation consists on a preliminary study of the yet uncharacterized ASFV pDP146L. As a first step of the protein's study, a DP146L expressing vector was transfected

into mammalian cells in order to identify its intracellular location. Since the DP146L ORF is cloned in the pcDNA3 plasmid in frame with a hemagglutinin (HA) tag, protein detection was performed using anti-HA tag antibodies.

It was not possible to detect pDP146L through immunofluorescence assay, even after several attempts of protocol optimization (which included the testing of two different cell lines (HEK293T and Vero), two different transfection reagents (TransIT®-LT1 Transfection Reagent (Mirus) and Lipofectamine® LTX & PLUS Transfection Reagent (Thermo Scientific)), and different DNA:transfection reagent ratios). In order to exclude the hypothesis of a technical problem during the transfection process, two other known plasmids were transfected in parallel and both were detected by immunofluorescence, unlike the pDP146L. HEK293T cells are known to be easy to transfect, with a theoretical transfection efficiency of around 80% (Gavrilescu and van Etten 2007). The highest efficiency achieved in this dissertation was not only considerably lower than the reported value, but also lower than the percentages already obtained for these two well-known plasmids in the lab. We thus understand that transfections carried out in the lab in that period of time were defective. Some factors that could affect the transfection quality are: the welfare of transfected cells, the quality of the culture medium, DNA purity and concentration, possible culture contaminations (such as fungi or mycoplasma), incubator temperature, and/or CO_2 concentration (Mirus Bio 2021). Some of these conditions are very difficult to control, even with optimization strategies.

Several explanations for the non-detection of pDP146L in these assays can be postulated. Since the pDP146L was not detected both in immunofluorescence and western blot assays, a possible explanation was that the pcDNA3-HA-DP146L plasmid might not be being transfected. However, this possibility was rejected when luciferase assays were performed. Such assays evidenced the expression of the pDP146L, since a clear inhibition of the luciferase values in DP146L transfected cells activated by the ligand was observed, proving a successfully transfection of the plasmid. Another possible explanation for the non-detection of DP146L is the HA tag nucleotide sequence being damaged. In order to assess this possibility, the plasmid was Sanger sequenced and the accuracy of the HA tag sequence was confirmed (result not shown), so this hypothesis was also discarded. An alternative explanation is that the protein may be being degraded by the cell after correct plasmid transfection and protein synthesis, which could explain the lack of protein detection. Indeed, considering the postulated function of the pDP146L as a putative suppressor of cytokine responses (SOCS), and since SOCS proteins are known as targeting molecules for degradation, it seems very logical to predict that the pDP146L could proceed together with its target proteins and end up being degraded along with them. However, this hypothesis still needs further exploration. Finally, there is also the possibility that the protein is being correctly synthesized inside the cell albeit at very low concentrations, making it difficult to be detected.

In conclusion, it was not possible to detect intracellular pDP146L after several immunofluorescence and western blot assays, making it impossible to pinpoint its cellular location. A new strategy that should increase detection efficiency consists in the use of an anti-pDP146L antibody. This way, the pDP146L could be directly detected, instead of relying on the associated HA tag. Better detection efficiency, conjugated with an improvement in transfection conditions, could allow for intracellular localization of the pDP146L.

As an initial task to address that issue, the pDP146L was successfully produced in an *E. coli*-based protein expression system, and purified through nickel affinity chromatography columns. In the final SDS-PAGE protein analysis we observed that part of the protein remains trapped in the cell lysis pellet, suggesting some insolubility. Despite this, a substantial amount of protein is present in final eluates, indicating that the ASFV pDP146L is at least partially soluble. Protein solubility is an important thermodynamic parameter for protein studies.

With the pDP146L produced in this dissertation, we are now one step ahead to immunize laboratory mice and obtain specific anti-pDP146L antibodies, which should improve the detection of pDP146L expression, which was a problematic task in this dissertation.

5.2. The yet unstudied ASFV pDP146L is potentially involved in the host cell response to infection

Macrophages are central cells for the trigger of immune response and are also the target cells of the ASFV. Once infected, these cells rapidly develop a pro-inflammatory response, mainly due to cytokine induction, establishing an antiviral state. It would be difficult, or even impossible, for a virus to persist in this cell environment and establish an infection without a strategy to interfere with the innate immune system. ASFV would greatly benefit from having a cytokine response evasion system since it is one of the first responses to invading organisms (Correia et al. 2013; Dixon et al. 2019).

Some ASFV genes interfering with both IFN induction and impact pathways have already been discovered (Correia et al. 2013; Zhuo et al. 2020; Li et al. 2021). However, the search for more host immune evasion genes is still ongoing. Recently, some bioinformatic studies developed at the IGC Infection and Immunity Group, led by Michael Parkhouse, identified a new, unstudied gene that could belong to this group: the DP146L ORF. It has strong homology to the SH2 domain, which is present in many intracellular signalling proteins, namely in the SOCS proteins. Thus, DP146L possibly encodes for a protein with similar functions to the SOCS family of proteins, and so a putative suppressor of cytokine signalling, namely the IFN I induction and impact. This discovery launched a new line of research about ASFV's immune evasion proteins.

This dissertation is a preliminary study of the ASFV pDP146L's function. Its main goal is to assess DP146L's effect on the NF-κB signalling pathway, since it is one of the target

pathways of the SOCS proteins. Furthermore, such is also one of the possible routes to IFN I induction, namely through the TLR signalling pathway and the RIG-I and MDA5 cytosolic pathways, and so it is also a probable target of pDP146L. In the NF-κB signalling pathway, the NF-κB transcription factor is activated and then translocates into the nucleus and binds to the PDR II promoter binding site. Accordingly, in this dissertation, the pDP146L effect on the NF-κB pathway was assessed through luciferase reporter gene assays, with the luciferase gene under the control of PRD II. The NF-κB signalling pathway is not exclusive of the IFN induction pathway, being common to many other signal transductions. The tumor necrosis factor (TNF)α and the IL-1 were the chosen stimuli used in this dissertation's luciferase assays, since both of them also induce NF-κB activation and its binding to the PRD II region.

TNFα is a potent inflammatory cytokine that signals through two distinct receptors, TNF receptor 1 (TNFR1) and TNFR2. Signalling through TNFR1 leads to the activation of NF-κB. Upon TNF binding to TNFR1, it undergoes a conformational change that leads to dissociation of the inhibitory protein SODD (silencer of death domains) from the intracellular TNFR1 death domain (DD). It promotes the adaptor protein TRADD (TNFR1-associated death domain) to bind to the TNFR1 DD and initiate signalling. Then TRADD recruits TRAF2 and the serine-threonine RIP kinase. Therefore, IKK is recruited by TRAF2 and activated by RIP and phosphorylates IκB. IκB is degraded and releases the transcription factor NF-κB, which translocates to the nucleus to promote transcription of proteins involved in crucial cell activities (Silke 2011; Urschel and Cicha 2015). IL-1 is also a cytokine with a critical role in the immune and inflammatory responses. Its receptor's cytoplasmic tail shares highly significant homology with the cytoplasmic region of TLRs. Both of them contain TIR domains. Response to IL-1 starts when it binds to the extracellular domain of the IL-1 receptor type I (IL-1R1). It allows the recruitment of a second receptor subunit, the IL-1R accessory protein (IL-1RAP). Formation of the receptor heterodimer induces signalling through the recruitment of MyD88, IRAKs, and TRAF6. This complex next engages with another multiprotein complex involving NIK (NF-κB-inducing kinase) and IKKs. IKK is responsible for IκB phosphorylation and consequent degradation, leading to the release of NF-κB, which translocates to the nucleus and binds to the PRD II region of target genes (O'Neill and Greene 1998; Sims and Smith 2010).

Luciferase assays results demonstrate that the ASFV pDP146L strongly inhibits PRD II expression. With the presented results, we can conclude that the pDP146L has an inhibitory effect in the NF-κB signalling pathway. The NF-κB signalling pathway regulates many important cellular mechanisms, in particular in inflammatory responses, and is one possible route to the induction of IFN I. Therefore, these preliminary results show a possible function for pDP146L as an inhibitor of the IFN I induction. However, these results do not determine which specific pathway step this protein interferes with, nor its specific mechanism of action.

This is a preliminary result and will require additional functional assays to deepen the knowledge about this protein.

It is also important to note that this dissertation only demonstrate an effect of pDP146L on TNFα and IL-1 signalling pathways, not ensuring an effect in the IFN I induction pathways (even sharing the same NF-κB pathway). As described by Randall and Goodbourn (2008), individual binding sites can respond independently to stimuli and confer some degree of response. IRF3 (or IRF7) are the only known indispensable transcription factors for IFN-β induction. Activation of NF-κB may not be essential (Randall and Goodbourn 2008). Therefore, pDP146L impact on PDR II expression may not be directly proportional to its impact in the final IFN-β secretion. It is thus necessary to repeat the same luciferase assay with the IFN-β gene promoter reporter, in order to assess the concrete impact of the ASFV pDP146L on IFN-β induction.

To further extend the study of the ASFV pDP146L, several amino acids were identified by bioinformatic analysis as potentially essential residues for protein's function, and the amino acid at position 92 was chosen as the target of this preliminary study. It is a phenylalanine and localizes in the region showing high homology to the SH2 protein domain, and was thus hypothesized as a relevant residue for the interaction with target proteins, being important for protein's function as a suppressor of cytokine responses. The modification of a critical amino acid would lead to a variance in the protein's effect.

To address this question, the 92nd pDP146L amino acid was changed and a mutated pDP146L^{Phe92Asp} was engineered. The single-point mutation was successfully performed by site-directed mutagenesis. The 92nd amino acid was originally a phenylalanine encoded by the TTT nucleotide triplet. Phenylalanine is a non-polar, hydrophobic, non-charged amino acid, which belongs to the aromatic R group family. It was substituted by an aspartate encoded by the GAT nucleotide triplet. Aspartate, unlike the previous one, is a polar, hydrophilic, acidic, negatively charged amino acid (Nelson and Cox 2008). The swap of amino acids with totally opposite characteristics is very likely to affect that residue's function.

The pDP146L^{Phe92Asp} was then submitted to the function analysis previously performed for the wild-type pDP146L. Luciferase assays results demonstrate that the mutated protein is still capable of inhibiting PDR II expression. However, its inhibitory effect appears to be less efficient than the one demonstrated by the wild-type pDP146L, suggesting a partial reversal of the protein's effect. Such result suggests that Phe92 is an important residue for pDP146L's function. However, the fact that the Phe92Asp single-point mutation did not completely reverse the wild-type protein's effect indicates that more pDP146L residues contribute to its action. It is likely that a multiple-point mutation would be required for total reversion of protein's effect, modifying all or a considerable amount of its critical amino acids.

Overall, these findings support the hypothesis proposed above that the Phe92 amino acid is one of the critical residues for pDP146L function. Nonetheless, further experiments need to be done in order to identify other critical residues for pDP146L function. Moreover, since the mutation of the protein led to an increasing of the PRD II expression, these results also support the previous conclusion that the ASFV pDP146L inhibits the PRD II expression, interfering with the NF-κB pathway, since its modification led to a partial functional reversal.

This dissertation presents a preliminary result accordant to the postulated hypothesis, suggesting for pDP146L a role as a modulator of host pro-inflammatory cytokine responses. More investigation needs to be done in order to fully understand this protein and its roles during ASFV infection. Even so, pDP146L could be capable of diminishing the host's immune response, thus allowing more time for viral replication and establishment of infection.

6. CONCLUSION AND FUTURE PERSPECTIVES

ASF is a worldwide highly contagious viral hemorrhagic disease that is rapidly spreading throughout the European Union countries, presenting a severe threat from a health and economic perspective. The development of a live virus vaccine attenuated by the deletion of non-essential viral genes that confer an advantage to the virus over host defenses seems to be a practical solution for ASFV. For that purpose, a new line of investigation has emerged in order to identify genes that have evolved for host immune system evasion.

A recent bioinformatic analysis of the ASFV complete genome identified the DP146L ORF as a possible member of this group of host immune evasion genes. The pDP146L has a structural homology to SOCS proteins, since they both possess an SH2 domain, and thus it is now thought to be able to suppress cytokine responses, particularly IFN responses.

This dissertation is a preliminary study of the ASFV pDP146L, focusing primarily on its functional characterization. The results supported the initially proposed hypothesis. It was found that pDP146L interferes with the NF-κB signalling pathway, since it was demonstrated to inhibit PRD II expression, the promoter binding site of NF-κB. This study thus suggests for pDP146L a role as a modulator of host pro-inflammatory cytokine responses. However, this dissertation's results only demonstrate an inhibitory effect of the pDP146L in TNFα and IL-1 pathways, not assuring the same impact on the TLR signalling pathway that results in IFN I induction, even sharing the same NF-κB route. However, and since NF-κB activation and PRD II expression are crucial factors for maximal levels of IFN I expression, it is very likely to predict that pDP146L could also have an inhibitory effect in IFN I induction. With this being true, the pDP146L could reduce IFN I production, repressing innate immune host response and giving extra time for the virus to replicate and establish an infection. More investigation needs to be done in order to confirm this hypothesis.

Moreover, this dissertation also identified a critical residue for the ASFV pDP146L protein's function. It is the 92nd amino acid of pDP146L, a phenylalanine, and is located in the homology SH2 region. The single-point mutation of Phe92Asp partially reverses the pDP146L effect, proving this amino acid as a relevant one for this protein's function, and affirming the plausibility of the postulated pDP146L function.

However, this dissertation is just an initial study on the ASFV pDP146L, and there is still a lot of work to be done in order to know this protein fully. The following logical step is to use the pDP146L purified in this dissertation to produce specific anti-pDP146L antibodies in order to optimize protein detection. This would allow determining protein's intracellular location, a task that was not accomplished in this study. Specific anti-pDP146L antibodies could also be used in the future to stain pDP146L at different time points during infection, in order to investigate its temporal expression and intracellular location during virus infection. At the same time, due to the possibility that the reason for not detecting pDP146L is due to proteasome degradation, it would be necessary to repeat both the immunofluorescence and western blot assays in the presence of MG132, an inhibitor of the proteasome, to evaluate this hypothesis.

In order to deepen the functional understanding of the pDP146L, more luciferase gene assays still need to be performed with additional luciferase gene reporters. The IFN-β reporter should be tested to determine the real impact of pDP146L in the IFN-β-induction pathway. An ELISA assay could also be performed to quantify the impact in the IFN-β production. Repeating the same luciferase assays with gene reporters from the JAK-STAT pathway could also be interesting, in order to assess if pDP146L also interferes with IFN impact, other than IFN induction. Following that, searching for pDP146L intracellular interacting partners by immunoprecipitation, for example, would allow determining in which specific steps pDP146L interacts.

Moreover, as mentioned in the 5.2. subsection, more residues need to be tested, similar to what was done with Phe92Asp in this dissertation, in order to identify additional critical residues of pDP146L. It will allow the multiple-point mutation of pDP146L, which will more likely make it totally lose its effect.

Finally, after the complete protein study, the construction of a mutant virus deleted in the pDP146L will be required to study pDP146L's actual impact on ASFV infection. Such will allow us to investigate any difference in the course of virus infection and viral replication for an ASFV lacking the pDP146L. This step might also be the ultimate trial for assessing pDP146L as a candidate for constructing a deleted vaccine.

The main focus of the pDP146L study is its potential value as a candidate for the construction of an ASFV attenuated vaccine. Yet, genes that can evade the host immune system are "readymade tools" to manipulate cytokine responses. Thus DP146L, as a gene capable of controlling the induction and impact of IFN, would be very useful for controlling

diseases characterized by their excessive inflammatory responses. Accordingly, the understanding of the ASFV DP146L ORF would provide, apart from a candidate target for vaccine development, a novel tool for the manipulation of cytokine responses in health and disease.

7. BIBLIOGRAPHY

Acharya KP, Wilson RT. 2020. Pig production is at risk from African Swine Fever (ASF) in Nepal. Transbound Emerg Dis. 2020(May). doi:10.1111/TBED.13720.

Akhtar LN, Benveniste EN. 2011. Viral Exploitation of Host SOCS Protein Functions. J Virol. 85(5):1912–1921. doi:10.1128/jvi.01857-10.

Akira S, Takeda K. 2004. Toll-like receptor signalling. Nat Rev Immunol. 4(7):499–511. doi:10.1038/nri1391.

Alejo A, Matamoros T, Guerra M, Andrés G. 2018. A proteomic atlas of the African swine fever virus particle. J Virol. 92(23):1-18. doi: 10.1128/jvi.01293-18.

Andrés G, García-Escudero R, Viñuela E, Salas ML, Rodríguez JM. 2001. African swine fever virus structural protein pE120R is essential for virus transport from assembly sites to plasma membrane but not for infectivity. J Virol. 75(15):6758-6768. doi:10.1128/jvi.75.15.6758-6768.2001.

Arias M, de la Torre A, Dixon L, Gallardo C, Jori F, Laddomada A, Martins C, Parkhouse RM, Revilla Y, Rodriguez F, et al. 2017. Approaches and perspectives for development of African swine fever virus vaccines. Vaccines. 5:35. doi:10.3390/vaccines5040035.

Babon JJ, McManus EJ, Yao S, DeSouza DP, Mielke LA, Sprigg NS, Willson TA, Hilton DJ, Nicola NA, Baca M, et al. 2006. The Structure of SOCS3 Reveals the Basis of the Extended SH2 Domain Function and Identifies an Unstructured Insertion That Regulates Stability. Mol Cell. 22:205–216. doi:10.1016/j.molcel.2006.03.024.

Basto AP, Nix RJ, Boinas F, Mendes S, Silva MJ, Cartaxeiro C, Portugal RS, Leitão A, Dixon LK, Martins C. 2006. Kinetics of African swine fever virus infection in Ornithodoros erraticus ticks. J Gen Virol. 87:1863–1871. doi:10.1099/vir.0.81765-0.

Bode JG, Ludwig S, Ehrhardt C, Erhardt A, Albrecht U, Schaper F, Heinrich PC, Häussinger D. 2003. IFN-α antagonistic activity of HCV core protein involves induction of suppressor of cytokine signaling-3. FASEB J. 17:488–490. doi:10.1096/fj.02-0664fje.

Bonilla FA, Oettgen HC. 2010. Adaptive immunity. J Allergy Clin Immunol. 125(Suppl 2):S33-40. doi:10.1016/j.jaci.2009.09.017.

Chaplin DD. 2010. Overview of the immune response. J Allergy Clin Immunol. 125(2 Suppl 2):S3-23. doi:10.1016/j.jaci.2009.12.980.

Christensen MH, Paludan SR. 2017. Viral evasion of DNA-stimulated innate immune responses. Cell Mol Immunol. 14(1):4–13. doi:10.1038/cmi.2016.06.

Correia S, Ventura S, Parkhouse RM. 2013. Identification and utility of innate immune system evasion mechanisms of ASFV. Virus Res. 173(1):87–100. doi:10.1016/j.virusres.2012.10.013.

Diebold SS, Kaisho T, Hemmi H, Akira S, Reis E Sousa C. 2004. Innate Antiviral Responses by Means of TLR7-Mediated Recognition of Single-Stranded RNA. Science.

303(5663):1529–1531. doi:10.1126/science.1093616.

Dixon LK, Chapman DAG, Netherton CL, Upton C. 2012. African swine fever virus replication and genomics. Virus Res. 173(2013):3–14. doi:10.1016/j.virusres.2012.10.020.

Dixon LK, Islam M, Nash R, Reis AL. 2019. African swine fever virus evasion of host defences. Virus Res. 266(2019):25–33. doi:10.1016/j.virusres.2019.04.002.

Dixon LK, Stahl K, Jori F, Vial L, Pfeiffer DiU. 2020. African Swine Fever Epidemiology and Control. Annu Rev Anim Biosci. 8:221–246. doi:10.1146/annurev-animal-021419-083741.

Dragan AI, Hargreaves V V., Makeyeva EN, Privalov PL. 2007. Mechanisms of activation of interferon regulator factor 3: the role of C-terminal domain phosphorylation in IRF-3 dimerization and DNA binding. Nucleic Acids Res. 35(11):3525–3534. doi:10.1093/nar/gkm142.

Duncan SA, Baganizi DR, Sahu R, Singh SR, Dennis VA. 2017. SOCS proteins as regulators of inflammatory responses induced by bacterial infections: A review. Front Microbiol. 8(DEC):1–15. doi:10.3389/fmicb.2017.02431.

Duvaud S, Gabella C, Lisacek F, Stockinger H, Ioannidis V, Durinx C. 2021. Expasy, the Swiss Bioinformatics Research Portal, as designed by its users. Nucleic Acids Res. 49:W216-W227. doi: 10.1093/nar/gkab225

Engel P, Angulo A. 2012. Viral immunomodulatory proteins: Usurping host genes as a survival strategy. Adv Exp Med Biol. 738:256–276. doi:10.1007/978-1-4614-1680-7_15.

Fensterl V, Sen GC. 2009. Interferons and viral infections. BioFactors. 35(1):14–20. doi:10.1002/biof.6.

Frobøse H, Rønn SG, Heding PE, Mendoza H, Cohen P, Mandrup-Poulsen T, Billestrup N. 2006. Suppressor of cytokine signaling-3 inhibits interleukin-1 signaling by targeting the TRAF-6/TAK1 complex. Mol Endocrinol. 20(7):1587–1596. doi:10.1210/me.2005-0301.

Gack MU, Shin YC, Joo C, Urano T, Liang C, Sun L, Takeuchi O, Akira S, Chen Z, Inoue S, et al. 2007. TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. Nature. 446:916–920. doi:10.1038/nature05732.

Galindo I, Alonso C. 2017. African swine fever virus: A review. Viruses. 9(5):103. doi:10.3390/v9050103.

García-Sastre A. 2017. Ten Strategies of Interferon Evasion by Viruses. Cell Press. 22:176-184. doi:10.1016/j.chom.2017.07.012.

Gavrilescu LC, van Etten RA. 2007. Production of replication-defective retrovirus by transient transfection of 293T cells. J Vis Exp. 10:550. doi:10.3791/550.

Guinat C, Reis AL, Netherton CL, Goatley L, Pfeiffer DU, Dixon L. 2014. Dynamics of African swine fever virus shedding and excretion in domestic pigs infected by intramuscular

inoculation and contact transmission. Vet Res. 45:93. doi:10.1186/s13567-014-0093-8.

Hirano M, Das S, Guo P, Cooper MD. 2011. The Evolution of Adaptive Immunity in Vertebrates. 1st ed. Elsevier inc.

Honda K, Takaoka A, Taniguchi T. 2006. Type I Inteferon Gene Induction by the Interferon Regulatory Factor Family of Transcription Factors. Immunity. 25(3):349–360. doi:10.1016/j.immuni.2006.08.009.

Huang S, Liu K, Cheng A, Wang M, Cui M, Huang J, Zhu D, Chen S, Liu M, Zhao X, et al. 2020. SOCS Proteins Participate in the Regulation of Innate Immune Response Caused by Viruses. Front Immunol. 11. doi:10.3389/fimmu.2020.558341.

Imada K, Leonard WJ. 2000. The Jak-STAT pathway. Mol Immunol. 37(2000):1–11. doi:10.1016/s0161-5890(00)00018-3.

Isaacs A, Lindenmann J. 1957. Virus interference. I. The interferon. Proc R Soc London Ser B - Biol Sci. 147(927):258–267. doi:10.1098/rspb.1957.0048.

Janeway CA. 1989. Approaching the Asymptote? Evolution and Revolution in Immunology. Cold Spring Harb Symp Quant Biol. 54:1–13. doi:10.1101/sqb.1989.054.01.003.

Jia D, Rahbar R, Chan RWY, Lee SMY, Chan MCW, Wang BX, Baker DP, Sun B, Malik Peiris JS, Nicholls JM, et al. 2010. Influenza virus non-structural protein 1 (NS1) disrupts interferon signaling. PLoS One. 5(11):e13927. doi:10.1371/journal.pone.0013927.

Jori F, Bastos ADS. 2009. Role of Wild Suids in the Epidemiology of African Swine Fever. Eco Heal. 6(2):296–310. doi:10.1007/s10393-009-0248-7.

Jori F, Vial L, Penrith ML, Pérez-sánchez R, Etter E, Albina E, Michaud V, Roger F. 2013. Review of the sylvatic cycle of African swine fever in sub-Saharan Africa and the Indian ocean. Virus Res. 173(2013):212–227. doi:10.1016/j.virusres.2012.10.005.

Karki S, Moniruzzaman M, Aylward FO. 2021. Comparative Genomics and Environmental Distribution of Large dsDNA Viruses in the Family Asfarviridae. Front Microbiol. 12(March):1–13. doi:10.3389/fmicb.2021.657471.

Kawai T, Akira S. 2006. Innate immune recognition of viral infection. Nat Immunol. 7(2):131–137. doi:10.1038/ni1303.

Kawai T, Akira S. 2010. The role of pattern-recognition receptors in innate immunity: Update on toll-like receptors. Nat Immunol. 11(5):373–384. doi:10.1038/ni.1863.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 10(6):845-858. doi: 10.1038/nprot.2015-053.

Kedkovid R, Sirisereewan C, Thanawongnuwech R. 2020. Major swine viral diseases: an Asian perspective after the African swine fever introduction. Porc Heal Manag. 6:20. doi:10.1186/s40813-020-00159-x.

Koetsier G, Cantor E. 2019. A Practical Guide to Analyzing Nucleic Acid Concentration and Purity with Microvolume Spectrophotometers [Internet]. Ipswich: New England Biolabs; [accessed 2021 Aug 7]. https://www.neb.com/-/media/nebus/files/application-notes/technote_mvs_analysis_of_nucleic_acid_concentration_and_purity.pdf?rev=c24cea043416420d84fb6bf7b554dbbb.

Kubo M, Hanada T, Yoshimura A. 2003. Suppressors of cytokine signaling and immunity. Nat Immunol. 4(12):1169–1176. doi:10.1038/ni1012.

Lee M, Rhee I. 2017. Cytokine signaling in tumor progression. Immune Netw. 17(4):214–227. doi:10.4110/in.2017.17.4.214.

Li D, Yang W, Li L, Li P, Ma Z, Zhang J, Qi X, Ren J, Ru Y, Niu Q, et al. 2021. African Swine Fever Virus MGF-505-7R Negatively Regulates cGAS–STING-Mediated Signaling Pathway. J Immunol. 206(8):1844–1857. doi:10.4049/jimmunol.2001110.

Linossi EM, Babon JJ, Hilton DJ, Nicholson SE. 2013. Suppression of cytokine signaling: The SOCS perspective. Cytokine Growth Factor Rev. 24(3):241–248. doi:10.1016/j.cytogfr.2013.03.005.

Liu BA, Jablonowski K, Raina M, Arcé M, Pawson T, Nash PD. 2006. The Human and Mouse Complement of SH2 Domain Proteins - Establishing the Boundaries of Phosphotyrosine Signaling. Mol Cell. 22:851–868. doi:10.1016/j.molcel.2006.06.001.

Luo X, Chen X, Qiao S, Li R, Xie S, Zhou X, Deng R, Zhou E, Zhang G. 2020. Porcine Reproductive and Respiratory Syndrome Virus Enhances Self-Replication via AP-1–Dependent Induction of SOCS1. J Immunol. 204:394–407. doi:10.4049/jimmunol.1900731.

Mansell A, Smith R, Doyle SL, Gray P, Fenner JE, Crack PJ, Nicholson SE, Hilton DJ, O'Neill LAJ, Hertzog PJ. 2006. Suppressor of cytokine signaling 1 negatively regulates Toll-like receptor signaling by mediating Mal degradation. Nat Immunol. 7(2):148–155. doi:10.1038/ni1299.

Marshall JS, Warrington R, Watson W, Kim HL. 2018. An introduction to immunology and immunopathology. Allergy, Asthma Clin Immunol. 14(Suppl 2):49. doi:10.1186/s13223-018-0278-1.

McNab F, Mayer-Barber K, Sher A, Wack A, O'Garra A. 2015. Type I interferons in infectious disease. Nat Rev Immunol. 15(2):87–103. doi:10.1038/nri3787.

Mikita T, Daniel C, Wu P, Schindler U. 1998. Mutational analysis of the STAT6 SH2 domain. J Biol Chem. 273(28):17634–17642. doi:10.1074/jbc.273.28.17634.

Mirus Bio. 2021. TransIT®-LT1 Full Transfection Protocol [Internet]. Madison (WI): Mirus Bio; [accessed 2021 Jun 21]. https://www.mirusbio.com/assets/protocols/ml001_transit_lt1_transfection_reagent.pdf

Montgomery RE. 1921. On A Form of Swine Fever Occurring in British East Africa (Kenya Colony). J Comp Pathol Ther. 34:159–191. doi:10.1016/s0368-1742(21)80031-4.

Murphy K, Weaver C. 2017. Janeway's Immunobiology. 9th ed. New York: Garland Science.

Murray PJ. 2014. The JAK-STAT Signaling Pathway: Input and Output Integration. J Immunol. 178:2623–2629. doi:10.4049/jimmunol.178.5.2623.

Nair S, Pandey AD, Mukhopadhyay S. 2011. The PPE18 Protein of Mycobacterium tuberculosis Inhibits NF-κB/rel–Mediated Proinflammatory Cytokine Production by Upregulating and Phosphorylating Suppressor of Cytokine Signaling 3 Protein. J Immunol. 186(9):5413–5424. doi:10.4049/jimmunol.1000773.

National Center for Biotechnology Information (NCBI) [Internet]. 2021. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information. [accessed 2021 Jun 15]. https://www.ncbi.nlm.nih.gov/

Nelson DL, Cox MM. 2008. Lehninger Principles of Biochemistry. 5th editio. Ahr K, Rossignol R, editors. New York (NY): W. H. Freeman & Company.

Oeckinghaus A, Ghosh S. 2009. The NF-kB family of transcription factors and its regulation. Cold Spring Harb Perspect Biol. 1(4):1–14. doi:10.1101/cshperspect.a000034.

de Oliveira VL, Almeida SCP, Soares HR, Crespo A, Marshall-Clarke S, Parkhouse RME. 2011. A novel TLR3 inhibitor encoded by African swine fever virus (ASFV). Arch Virol. 156(4):597–609. doi:10.1007/s00705-010-0894-7.

O'Neill LAJ, Greene C. 1998. Signal transduction pathways activated by the IL-1 receptor family: ancient signaling machinery in mammals, insects, and plants. J Leukoc Biol. 63(6):650–657. doi:10.1002/jlb.63.6.650.

Piganis RAR, De Weerd NA, Gould JA, Schindler CW, Mansell A, Nicholson SE, Hertzog PJ. 2011. Suppressor of Cytokine Signaling (SOCS) 1 inhibits type I interferon (IFN) signaling via the interferon α receptor (IFNAR1)-associated tyrosine kinase Tyk2. J Biol Chem. 286:33811–33818. doi:10.1074/jbc.M111.270207.

Platanias LC. 2005. Mechanisms of Type-I- and Type-II-Interferon-Mediated Signalling. 5:375–386. doi:10.1038/nri1604.

Powell PP, Dixon LK, Parkhouse RME. 1996. An IkB homolog encoded by African swine fever virus provides a novel mechanism for downregulation of proinflammatory cytokine responses in host macrophages. J Virol. 70(12):8527–8533. doi:10.1128/jvi.70.12.8527-8533.1996.

Qing Y, Costa-Pereira AP, Watling D, Stark GR. 2005. Role of tyrosine 441 of Interferon-γ receptor subunit 1 in SOCS-1-mediated attenuation of STAT1 activation. J Biol Chem. 280(3):1849–1853. doi:10.1074/jbc.M409863200.

Randall RE, Goodbourn S. 2008. Interferons and viruses: An interplay between induction, signalling, antiviral responses and virus countermeasures. J Gen Virol. 89(1):1–47.

doi:10.1099/vir.0.83391-0.

Rock DL. 2021. Thoughts on African Swine Fever Vaccines. Viruses. 13(5):943. doi:10.3390/v13050943.

Russell RB, Breed J, Barton GJ. 1992. Conservation analysis and structure prediction of the SH2 family of phosphotyrosine binding domains. FEBS Lett. 304(1):15–20. doi:10.1016/0014-5793(92)80579-6.

Ryo A, Suizu F, Yoshida Y, Perrem K, Liou YC, Wulf G, Rottapel R, Yamaoka S, Lu KP. 2003. Regulation of NF-κB Signaling by Pin1-Dependent Prolyl Isomerization and Ubiquitin-Mediated Proteolysis of p65/RelA. Mol Cell. 12(6):1413–1426. doi:10.1016/S1097-2765(03)00490-8.

Salas ML, Andrés G. 2013. African swine fever virus morphogenesis. Virus Res. 173(2013):29-41. doi: 10.1016/j.virusres.2012.09.016.

Salguero FJ. 2020. Comparative Pathology and Pathogenesis of African Swine Fever Infection in Swine. Front Vet Sci. 7(May):12–14. doi:10.3389/fvets.2020.00282.

Salguero FJ, Ruiz-Villamor E, Bautista MJ, Sánchez-Cordón PJ, Carrasco L, Gómez-Villamandos JC. 2002. Changes in macrophages in spleen and lymph nodes during acute African swine fever : expression of cytokines. Vet Immunol Immunopathol. 90(2002):11–22. doi:10.1016/S0165-2427(02)00225-8.

Salguero FJ, Sánchez-Cordón PJ, Núñez A, Fernández de Marco M, Gómez-Villamandos JC. 2005. Proinflammatory cytokines induce lymphocyte apoptosis in acute African swine fever infection. J Comp Pathol. 132(4):289–302. doi:10.1016/j.jcpa.2004.11.004.

Sánchez-Vizcaíno JM, Mur L, Gomez-Villamandos JC, Carrasco L. 2015. An update on the epidemiology and pathology of African swine fever. J Comp Pathol. 152:9-21. doi: 10.1016/j.jcpa.2014.09.003.

Sasaki A, Inagaki-Ohara K, Yoshida T, Yamanaka A, Sasaki M, Yasukawa H, Koromilas AE, Yoshimura A. 2003. The N-terminal truncated isoform of SOCS3 translated from an alternative initiation AUG codon under stress conditions is stable due to the lack of a major ubiquitination site, Lys-6. J Biol Chem. 278(4):2432–2436. doi:10.1074/jbc.C200608200.

Sasaki A, Yasukawa H, Suzuki A, Kamizono S, Syoda T, Kinjyo I, Sasaki M, Johnston JA, Yoshimura A. 1999. Cytokine-inducible SH2 protein-3 (CIS3/SOCS3) inhibits Janus tyrosine kinase by binding through the N-terminal kinase inhibitory region as well as SH2 domain. Genes to Cells. 4:339–351. doi:10.1046/j.1365-2443.1999.00263.x.

Schwartz DM, Kanno Y, Villarino A, Ward M, Gadina M, O'Shea JJ. 2017. JAK inhibition as a therapeutic strategy for immune and inflammatory diseases. Nat Rev Drug Discov. 17(1):78. doi:10.1038/nrd.2017.267.

Shuai K, Horvath CM, Huang LHT, Qureshi SA, Cowburn D, Darnell JE. 1994. Interferon

activation of the transcription factor Stat91 involves dimerization through SH2-phosphotyrosyl peptide interactions. Cell. 76(5):821–828. doi:10.1016/0092-8674(94)90357-3.

Shuai K, Liu B. 2003. Regulation of JAK-STAT signalling in the immune system. Nat Rev Immunol. 3(11):900–911. doi:10.1038/nri1226.

Silke J. 2011. The regulation of TNF signalling: what a tangled web we weave. Curr Opin Immunol. 23:620–626. doi:10.1016/j.coi.2011.08.002.

Sims JE, Smith DE. 2010. The IL-1 family: regulators of immunity. Nat Rev Immunol. 10(2):89–102. doi:10.1038/nri2691.

Tabeta K, Georgel P, Janssen E, Du X, Hoebe K, Crozat K, Mudd S, Shamel L, Sovath S, Goode J, et al. 2004. Toll-like receptors 9 and 3 as essential complonents of innate immune defense against mouse cytomegalovirus infection. Proc Natl Acad Sci U S A. 101(10):3516–3521. doi:10.1073/pnas.0400525101.

Turvey SE, Broide DH. 2010. Innate immunity. J Allergy Clin Immunol. 125(Suppl 2):S24-32. doi:10.1016/j.jaci.2009.07.016.

Urschel K, Cicha I. 2015. TNF-α in the cardiovascular system: from physiology to therapy. Int J Interf Cytokine Mediat Res. 7:9–25. doi:10.2147/IJICMR.S64894.

Versteeg GA, García-Sastre A. 2010. Viral tricks to grid-lock the type I interferon system. Curr Opin Microbiol. 13(4):508–516. doi:10.1016/j.mib.2010.05.009.

Wang N, Zhao D, Wang Jialing, Zhang Y, Wang M, Gao Y, Li F, Wang Jingfei, Bu Z, Rao Z, et al. 2019. Architecture of African swine fever virus and implications for viral assembly. Science. 366(6465):640–644. doi:10.1126/science.aaz1439.

Weber F, Haller O. 2007. Viral suppression of the interferon system. Biochimie. 89(2007):836–842. doi:10.1016/j.biochi.2007.01.005.

Wen Z, Zhong Z, Darnell JE. 1995. Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation. Cell. 82:241–250. doi:10.1016/0092-8674(95)90311-9.

Wilkins C, Gale Jr M. 2010. Recognition of viruses by cytoplasmic sensors. Curr Opin Immunol. 22(1):41–47. doi:10.1016/j.coi.2009.12.003.

Wullaert A, Heyninck K, Janssens S, Beyaert R. 2006. Ubiquitin: tool and target for intracellular NF-κB inhibitors. Trends Immunol. 27(11):533–540. doi:10.1016/j.it.2006.09.003.

Xu X, Sun Y-L, Hoey T. 1996. Cooperative DNA binding and sequence-selective recognition conferred by the STAT amino-terminal domain. Science. 273:794–797. doi:10.1126/science.273.5276.794.

Yokota SI, Yokosawa N, Okabayashi T, Suzutani T, Fujii N. 2004. Induction of suppressor of cytokine signaling-3 by herpes simplex virus type 1 contributes to inhibition of the Interferon signalling pathway. Virology. 78(12):6282–6286.

doi:10.1128/JVI.78.12.6282-6286.2004.

Yoshimura A, Naka T, Kubo M. 2007. SOCS proteins, cytokine signalling and immune regulation. Nat Rev Immunol. 7(6):454–465. doi:10.1038/nri2093.

Yoshimura A, Ohkubo T, Kiguchi T, Jenkins NA, Gilbert DJ, Copeland NG, Hara T, Miyajima A. 1995. A novel cytokine-inducible gene CIS encodes an SH2-containing protein that binds to tyrosinephosphorylated interleukin 3 and erythropoietin receptors. EMBO J. 14(12):2816–2826. doi:10.1002/j.1460-2075.1995.tb07281.x.

Young HA, Bream JH. 2007. IFN-γ: recent advances in understanding regulation of expression, biological functions, and clinical applications. Curr Top Microbiol Immunol. 316:97–117. doi:10.1007/978-3-540-71329-6_6.

Yu C-F, Peng W-M, Schlee M, Barchet W, Eis-Hübinger AM, Kolanus W, Geyer M, Schmitt S, Steinhagen F, Oldenburg J, et al. 2018. SOCS1 and SOCS3 Target IRF7 Degradation To Suppress TLR7-Mediated Type I IFN Production of Human Plasmacytoid Dendritic Cells. J Immunol. 200(12):4024–4035. doi:10.4049/jimmunol.1700510.

Zhou X, Liu Z, Cheng X, Zheng Y, Zeng F, He Y. 2015. Socs1 and Socs3 degrades Traf6 via polyubiquitination in LPS-induced acute necrotizing pancreatitis. Cell Death Dis. 6:1–9. doi:10.1038/cddis.2015.342.

Zhu Z, Chen H, Liu L, Cao Y, Jiang T, Zou Y, Peng Y. 2020. Classification and characterization of multigene family proteins of African swine fever viruses. Brief Bioinform. 22(4):1–11. doi:10.1093/bib/bbaa380.

Zhuo Y, Guo Z, Ba T, Zhang C, He L, Zeng C, Dai H. 2020. African Swine Fever Virus MGF360-12L Inhibits Type I Interferon Production by Blocking the Interaction of Importin α and NF-κB Signaling Pathway. Virol Sin. 36(2):176–186. doi:10.1007/s12250-020-00304-4.

8. ANNEXES

Annex 1 - pcDNA3 vector map

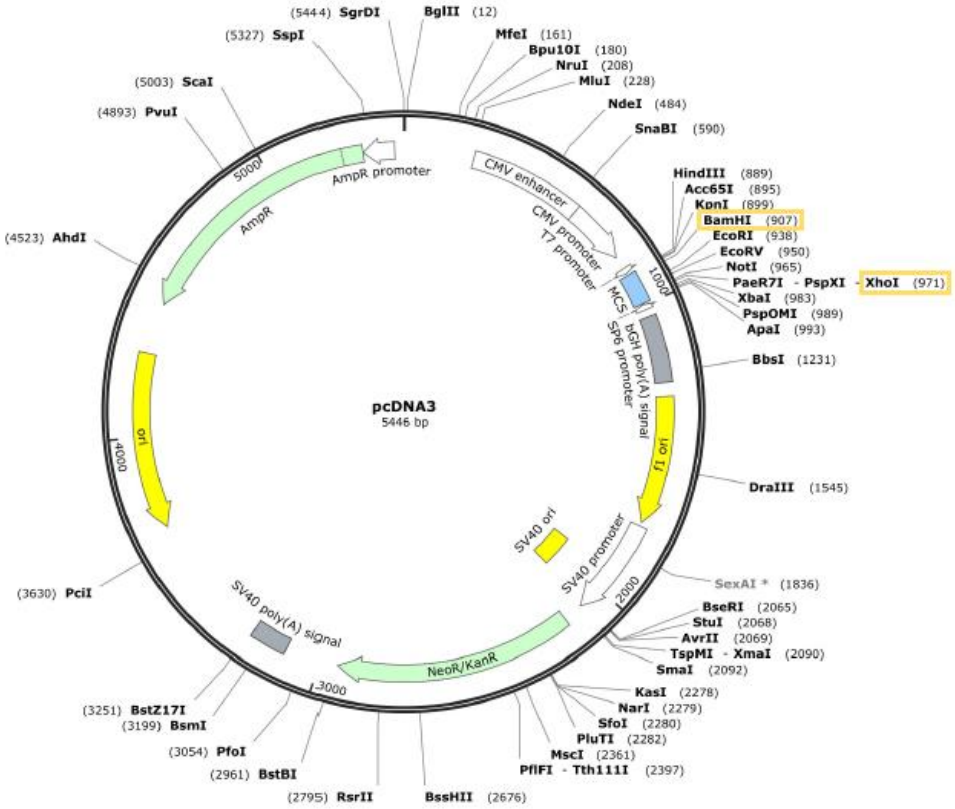

Figure 21. pcDNA3 vector map. Figure extracted from SnapGene.

Annex 2 - Schematic representation of the pcDNA3-HA-DP146L plasmid

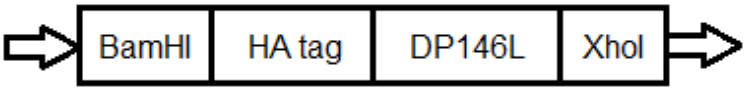

Figure 22. Schematic representation of the pcDNA3-HA-DP146L plasmid. The ASFV DP146L ORF and the HA tag are cloned in the pcDNA3 plasmid in the BamHI and XhoI restriction sites.

Annex 3 - pET28a vector map

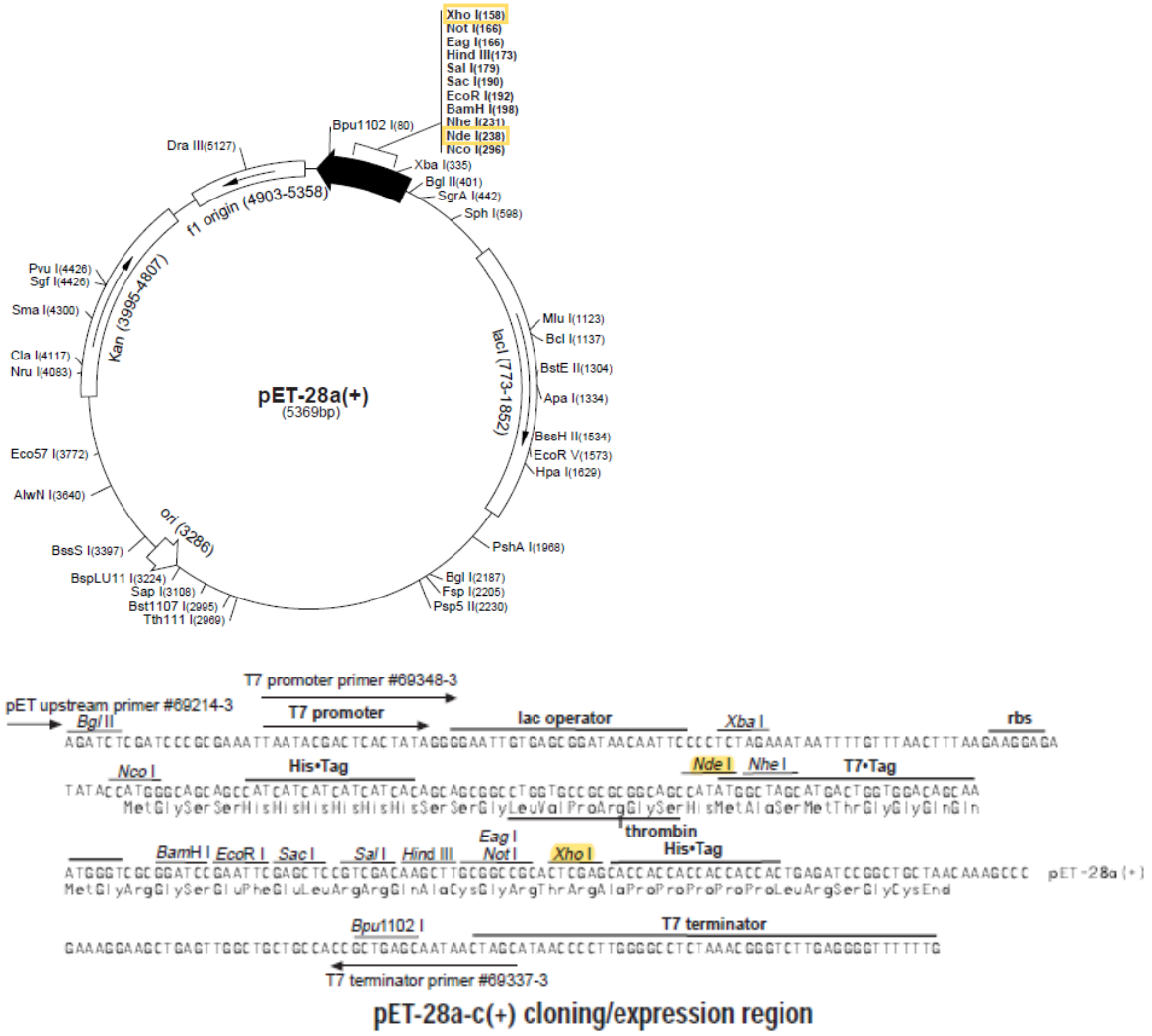

Figure 23. pET28a vector map. The restriction sites used in this dissertation are highlighted in yellow. Figure extracted from Novagen.

Annex 4 - Schematic representation of the pET28a-DP146L plasmid

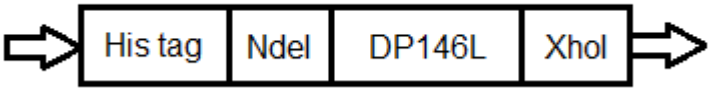

Figure 24. Schematic representation of the pET28a-DP146L plasmid. The ASFV DP146L ORF was cloned in the pET28a plasmid in the NdeI and XhoI restriction sites, in frame with the N-terminal His tag, which is originally present in the plasmid.