UCLouvain

AGRO LOUVAIN

Faculté des bioingénieurs

Effect of an N-terminal tag on the production of SARS-CoV-2 spike protein receptor binding domain in *Nicotiana tabacum* Bright Yellow-2 cells.

Auteur : Joel Skënderaj
Promoteur(s) : Pr. François Chaumont & Marie Peeters
Lecteur(s) : Pr. Henri Batoko & Pr. Patrice Soumillion
Année académique 2023-2024
Mémoire de fin d'études présenté en vue de l'obtention du diplôme de Bioingénieur : chimie et bio-industries

Acknowledgements

I would like to take this opportunity to express my gratitude to everyone who helped me in the realisation of this work.

I am sincerely grateful to Pr. François Chaumont for welcoming me in his team and giving me the opportunity to work on this incredible master thesis. I would like to thank him for his time and his feedback on this work. I would also like to thank Dr. Catherine Navarre for her insightful tips, for her kindness and for her positivity throughout the year. A special thanks to Marie Peeters, my co-supervisor, for her guidance, availability, encouragement and her invaluable help. Thank you for being someone I could always rely on.

I would like to thank the members of the FARM team for their support, practical assistance and for generously lending me resources whenever I was in need. I am thankful for the positive and collaborative environment I found in FYMO group, making work in the lab more pleasurable.

On a personal note, I am endlessly grateful to my parents for their unconditional support throughout my studies. And last but not least, I am thankful to my relatives and friends, for their support, encouragement, interest in my work and their good words.

Finally, I would like to thank everyone who has contributed to this work, directly or indirectly, making this master's thesis a great experience.

Table of Contents

Table of Figures

Table of Tables

List of acronyms

2,4-D: 2,4- dichlorophenoxyacetid acid
2YT RGS: 2x Yeast Extract and Tryptone supplemented with rifampicin, gentamycin and spectinomycin
2YT: 2x Yeast Extract and Tryptone
A. tumefaciens: *Agrobacterium tumefaciens*
APA: anti-pseudomonas aeruginosa
APN: aminopeptidase N
ARDS: acute respiratory distress syndrome
ATP: adenosine triphosphate
BCoV: bat coronaviruses
BHK: Baby Hamster Kidney cells
BSA: bovine serum albumin
BY-2: Bright Yellow-2
CaMV: cauliflower mosaic virus
CEA: carcinoembryonic antigen
CH: central helix
CHO: Chinese Hamster Ovary cells
COVID-19: coronavirus disease 19
CT: cytoplasmic tail
CTD: C-terminal domain
D: aspartic acid
E protein: envelope protein
E: glutamic acid
E. coli: *Escherichia coli*
ELISA: Enzyme-Linked Immunosorbent Assay
EM: extracellular medium
EMA: European Medicines Agency
ER: endoplasmic reticulum
FDA: Food and Drug Administration
FP: fusion peptide
G: glycine
GMP: Good Manufacturing Practices
H: histidine
hACE2: human angiotensin-converting enzyme 2
HEK293: Human Embryonic Kidney 293 cells
hEPO: human erythropoietin
HR: heptad repeat
HRP: horseradish peroxidase
il-10: interleukin-10
IPTG: isopropyl β-D-1-thiogalactopyranoside
K: lysine
L: leucine
LB: left border
LB: Lysogeny Broth
LBA: Lysogeny Broth supplemented with ampicillin
LBS: Lysogeny Broth supplemented with spectinomycin
M protein: membrane protein
M. sativa: *Medicago sativa*
M. truncatula: *Medicago truncatula*
mAbs: monoclonal antibodies
MERS-CoV: middle east respiratory syndrome coronavirus
MHV: murine hepatitis virus
mRNA: messenger RNA
MS: Murashige and Skoog
MSCCK: Murashige and Skoog supplemented with carbenicillin, cefotaxime and kanamycin
MSK: Murashige and Skoog supplemented with kanamycin
N protein: nucleocapsid protein
N: asparagine
N. benthamiana: *Nicotiana benthamiana*
N. tabacum: *Nicotiana tabacum*
nAbs: neutralising antibodies
Nos: nopaline synthase
nptII: neomycin phosphotransferase II
nsps: non-structural proteins
NTA: nitrilotriacetic acid
NTD: N-terminal domain
O/N: overnight
OD_{600}: optical density at 600 nm
PCR: polymerase chain reaction

PDI: protein disulphide isomerase
PEDV: porcine epidemic diarrhea
PFA: *Pyrococcus furiosus* alpha amylase
PMP1: plasma membrane ATPase proteolipid 1
PMSF: phenylmethylsulfonylfluoride
Q: glutamine
R: arginine
RB: right border
RBD: receptor binding domain
RBM: receptor binding motif
RdRp: RNA-dependent RNA polymerase
RT: room temperature
S protein: spike protein
S. cerevisiae: *Saccharomyces cerevisiae*
SAR: Scaffold Attachment Region
SARS-CoV-2: severe acute respiratory syndrome coronavirus 2
SARS-CoV: severe acute respiratory syndrome coronavirus
Sf-21: *Spodoptera furgiperda* cell line 21
Sf-9: *Spodoptera furgiperda* cell line 9
sg: subgenomic
SP: signal peptide
SV5: simian virus 5
T-DNA: transfer DNA
TBS: Tris Buffered Saline
TGEV: transmissible gastroenteritis virus
TM: transmembrane segment
Tn-368: *Trichoplusia ni* cell line 368
tRNA: transfer RNA
TSLP: total soluble leaf proteins
TSP: total soluble proteins
UTR: untranslated
VLP: virus-like particles
VOC: variant of concern
wt: wild type
Y: tyrosine

Chapter I: Introduction

1.1 Coronaviruses and their properties

Coronaviruses (CoVs) infect a large variety of animal species, wild and domestic, including pigs, cattle and poultry, and cause illnesses that range from gastrointestinal tract infections to encephalitis and demyelination, which can often be fatal. CoVs also infect humans (hCoVs) and have been mostly associated with upper respiratory tract infections. They often cause mild symptoms similar to the common cold or pneumonia (Hasöksüz *et al*., 2020; Schoeman *et al*., 2022). However, outbreaks such as the Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) in 2002, the Middle East Respiratory Syndrome Coronavirus (MERS-CoV) in 2012 or the more recent Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in 2020, have shown that hCoVs can also cause more severe illnesses and be lethal (Weiss and Navas-Martin, 2005; Schoeman *et al*., 2022).

It is therefore important to study these viruses, in order to better understand their properties, replication mechanisms, transmissibility and virulence. This could allow to better mitigate a possible (rather probable) future outbreak.

1.1.1 Classification

CoVs belong to the *Coronaviridae* family, in the *Nidovirales* order. This order is composed of three families, the two others being the *Arteriviridae* and *Roniviridae* containing bird and insect pathogens respectively. The *Coronaviridae* family can be further divided into two subfamilies, namely the *Coronavirinae* (coronaviruses) and *Torovirinae* (toroviruses). The latter contains viruses that cause enteric infections to livestock but are not known to infect humans (Weiss and Navas-Martin, 2005; Burrell *et al*., 2017). Finally, the *Coronavirinae* subfamily can be divided into four genera.

The *Alphacoronavirus* genus, includes among others the human coronavirus 229E (hCoV-229E) and the human coronavirus NL63 (hCoV-NL63), which are the most common and are known to cause pneumonia and flu-like symptoms that are usually mild. The genus includes some animal coronaviruses as well.

The *Betacoronavirus* genus, includes some of the most pathogenic hCoVs such as SARS-CoV, MERS-CoV or SARS-CoV-2, but also some other, less severe, human pathogens such as the human coronavirus OC43 (hCoV-OC43) or the human coronavirus HKU1 (hCoV-HKU1). Interestingly, the bat coronaviruses (BCoV), are also found in the *Betacoronavirus* genus. The last two genera are the *Gammacoronaviruses* that mainly infect birds and the *Deltacoronaviruses*, mainly found in pigs (Burrell *et al*., 2017; Schoeman *et al*., 2022).

1.1.2 Structure and genome

CoVs are spherical-shaped viruses of around 100 nm in diameter, enveloped with a lipid membrane derived from the host cell. The proteins found in the membrane, namely the envelope (E) protein, the membrane (M) protein, and especially the spike (S) protein protruding from the membrane give the virus its characteristic halo-like appearance that is visible under electron microscopy, hence its name “corona” which means crown in Latin (Ludwig and Zarbock, 2020; Schoeman *et al*., 2022). A schematic representation of a CoV virion is depicted in Figure 1.

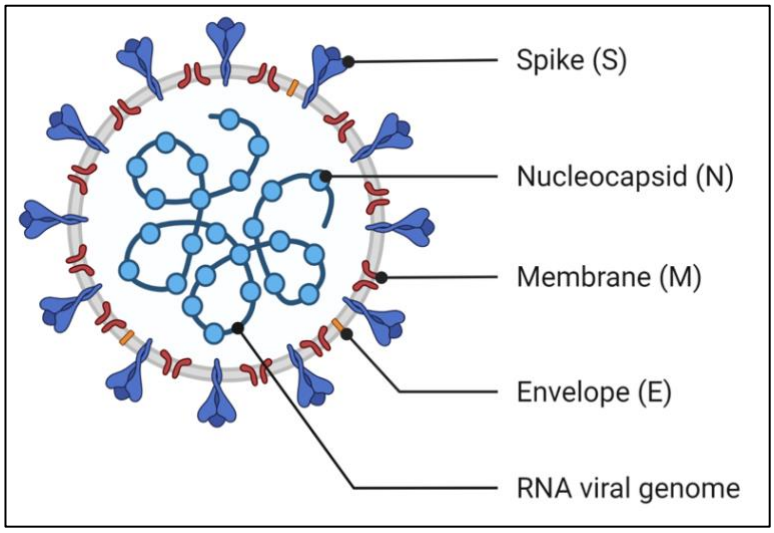

Figure 1: Schematic representation of a CoV virion (Schoeman et al.*, 2022)*

CoVs have a linear, single-stranded, positive sense RNA genome with a size ranging between 26 and 32 kilobases, making it the largest genome of known RNA viruses. The genome is capped at the 5'- end and polyadenylated in the 3'- end. It encodes a number of non-structural proteins (nsps) that are essential for the virus replication and four major structural proteins, namely the envelope (E), membrane (M), spike (S) and the nucleocapsid (N) proteins (Burrell *et al*., 2017; Schoeman *et al*., 2022). The role of each structural protein will be described later.

The genome is organised as follows, 5'- replicase – S – E – M – N - 3', which is a common, conserved order in all CoVs. Two-thirds of the 5'-end of the genome, contains an ORF encoding for a large frameshift polyprotein (ORF1a/ORF1b), implicated in viral replication. The ORFs encoding the structural proteins are separated by ORFs encoding for various nsps, that differ among CoVs in terms of number of genes, location, size and nucleotide sequence (Burrell *et al*., 2017; Schoeman *et al*., 2022).

The expression of individual genes occurs through a complex process, where a set of nested mRNAs that share the same 5'-end are produced. Common to all CoVs, at the 5'-end of the genome, is an untranslated (UTR) sequence of about 65 to 98 nucleotides. Another UTR sequence of about 200 to 500 nucleotides is found at the 3'-end, followed by a poly A tail. Both of the UTR sequences, and the poly A tail are present in the subgenomic (sg) mRNA and are important for the regulation and transcription of the viral RNA (Burrell *et al*., 2017).

1.1.3 Betacoronavirus proteins

The different structural and non-structural proteins are relatively well conserved among CoVs in terms of structure and role. They will be briefly described in this section, whereas the SARS-CoV-2 spike protein, will be described into more details in a dedicated section.

1.1.3.1 Spike protein (S)

S is a trimeric glycoprotein of about 150 kDa found at the surface of the virions, forming peplomers of around 15 to 20 nm. S is a type I fusion protein that mediates the binding and the entry of the virus into the host cell, through its two functional subunits named S1 and S2 found on the N- and C-terminus of the protein, respectively. They are separated by a S1/S2 cleavage

site. The S1-subunit is found on the outside of the folded spike protein and is responsible for the binding of the virus to the host cell, via the Receptor Binding Domain (RBD). The position of the RBD varies depending on the virus. In some CoVs, the RBD is located on the N-terminus of S1, while in SARS-CoV and SARS-CoV-2 the RBD is located on the C-terminus. However, the S1-subunit often contains two domains, each capable of binding to the host cell, the N-terminal domain (NTD) and the C-terminal domain (CTD) (Burrell *et al*., 2017; Hasöksüz *et al*., 2020; Schoeman *et al*., 2022). The S2-subunit is a transmembrane domain mediating the fusion of the viral and host membranes and the internalization of the virus into the host cell.

For some CoVs, the expression of S at the cell membrane stimulates cell-to-cell fusion between infected and neighbouring, uninfected cells. This could be a strategy for the viral spread to uninfected cells. The fusion of infected and uninfected cells could allow the spread of the virus between them without the risk of encountering neutralizing antibodies in the host serum (Schoeman *et al*., 2022). Finally, variations in the S glycoprotein are largely responsible for the host and tissue tropism variability of the CoVs (Burrell *et al*., 2017; Schoeman *et al*., 2022).

1.1.3.2 Envelope protein (E)

E is the smallest, roughly 8-12 kDa, and the least well-known structural protein of CoVs. It is a transmembrane protein, with an N-terminal ectodomain and a C-terminal endodomain. However in some CoVs, the protein can span the membrane twice so that both the N- and C-termini are inside the virion (Weiss and Navas-Martin, 2005; Schoeman *et al*., 2022).

The sequence of E varies among CoVs, but its architecture is usually conserved (Fehr & Perlman, 2015). Even though it is produced in large quantities in the host cell, only a small portion of the protein is present in the viral structure, and most of the protein is localized in the endoplasmic reticulum (ER), Golgi and the ER-Golgi intermediate compartment of the host cell. The main hypothesis for the role of E is that it is implicated in the assembly and budding of the virus, through an interaction between its cytoplasmic tail and the M protein. Other hypotheses include a role in the release of the virion, due to its hydrophobic transmembrane domain, or a role in the virus' pathogenesis (Hasöksüz *et al*., 2020).

1.1.3.3 Membrane protein (M)

M is a relatively small protein, about 25-30 kDa, that defines the shape of the viral particles and is the most abundant structural protein in CoVs (Neuman *et al*., 2011). It has three to four transmembrane domains, with usually an N-terminal glycosylated ectodomain and a C-terminal endodomain. The M proteins of a number of CoVs, including SARS-CoV and SARS-CoV-2, are targeted to the vicinity of the Golgi apparatus and are crucial for the formation of the viral envelope, using the host cellular membranes to form new viral particles. M proteins can form homotypic interactions, thus, creating a M-M network, which is capable of excluding some host proteins from the viral membrane. However, M proteins alone are not enough for virion formation, and they need to interact with other proteins as well. The interactions between M and S are crucial for the retention of the latter in the ER-Golgi complex, and consequently its incorporation into the viral particle. M also interacts with N to stabilize the nucleocapsid, therefore favouring the completion of viral assembly, while, as stated previously, the M-E interactions are important for the release of the viral particles (Hasöksüz zt al., 2020; Schoeman *et al*., 2022).

1.1.3.4 Nucleocapsid protein (N)

N is a multifunctional protein and is produced abundantly in the infected cells. It is composed of three highly distinct conserved domains, which are folded separately. It contains an NTD and a CTD, separated by a disordered central region called the RNA-binding domain, all of which are capable of binding to the RNA. By binding to the RNA, N is mainly implicated in RNA related processes such as transcription and translation and plays a key role in viral replication. It has also been proposed that N plays a role in viral pathogenesis of CoVs by targeting and interacting with some host proteins, possibly leading to a disturbance of the cell cycle and inhibition of interferon production (Hasöksüz *et al*., 2020; Schoeman *et al*., 2022).

1.1.3.5 Accessory proteins

In addition to structural proteins, a number of subfamily-specific accessory proteins are encoded in the genome of CoVs. They vary in number, size and location and share no similarity among subfamilies, nor to any other known viral protein. These accessory proteins are mostly dispensable for growth *in vitro* but are still maintained in the CoV genome under selective pressure, suggesting they probably confer some advantages to the virus in the host cells that they infect. Some of the roles proposed for the accessory proteins include modulation of viral pathogenicity and replication, interferon antagonists and cell death inducers, thus interfering with normal cellular processes (Oostra *et al*., 2007; McBride & Fielding, 2012).

1.1.4 Viral entry and replication

Viral entry, as mentioned in section 1.1.3.1, is mediated by the S protein, with the S1-subunit binding to the receptors of the host cell, and the S2-subunit mediating membrane fusion. The binding of the S protein to the receptor on the surface of the host cell triggers conformational changes that lead to fusion between the viral envelope and the cell membrane (Figure 2). The target receptors vary depending on the virus. SARS-CoV and SARS-CoV-2 use the angiotensin-converting enzyme 2 (ACE2), on the surface of the epithelial cells in the lungs, heart, kidney and small intestine. Other CoVs, such as the hCoV-229E, TGEV or the PEDV use the aminopeptidase N (APN) to gain entry in the host cells, while the MHV uses a biliary glycoprotein belonging to the carcinoembryonic antigen (CEA) family in the Ig superfamily (CEACAM1) (Burrell *et al*., 2017; Hasöksüz *et al*., 2020).

After binding to the receptor, the RNA genome of the CoV is transferred inside the cell, directly or via endocytosis. During the first stage of viral replication, the positive strand genomic RNA serves as the mRNA for the translation of two ORFs, 1a and 1b, each encoding a subunit of a large polyprotein, the RNA-dependent RNA polymerase (RdRp) (Schoeman *et al*., 2022). The synthesis of the 1a and 1b polyproteins involves a ribosomal frameshift during the translation of the 1a gene. The polyproteins are then cleaved by viral-encoded proteases and each polyprotein leads to between 15 and 16 nsps which assemble to form an active RNA polymerase. The newly formed polymerase then transcribes the full-length genomic RNA into a negative-sense complementary RNA, which is copied afterwards into a positive-sense sg mRNA. The positive sense sg mRNA is transcribed into full-length mRNA and also gives rise to a set of nested 3'-coterminal overlapping sg mRNAs. Each of these nested sg mRNA encodes for one structural protein downstream of the replicase. The sequence not shared among the mRNAs is translated into proteins. The full-length mRNA is then encapsidated by the N protein, which interacts with the M protein, before associating with the other structural proteins in the ER-Golgi intermediate to form new viral particles (Schoeman *et al*., 2022). After

assembly the virus exits the cell through exocytosis (Marra *et al*., 2003). The replication cycle of CoVs is illustrated in Figure 2.

Since CoVs have a relatively large genome, recombination between distinct but related viruses is frequent and is believed to be an important factor for the genetic diversity of CoVs in nature (Burrell *et al*., 2017).

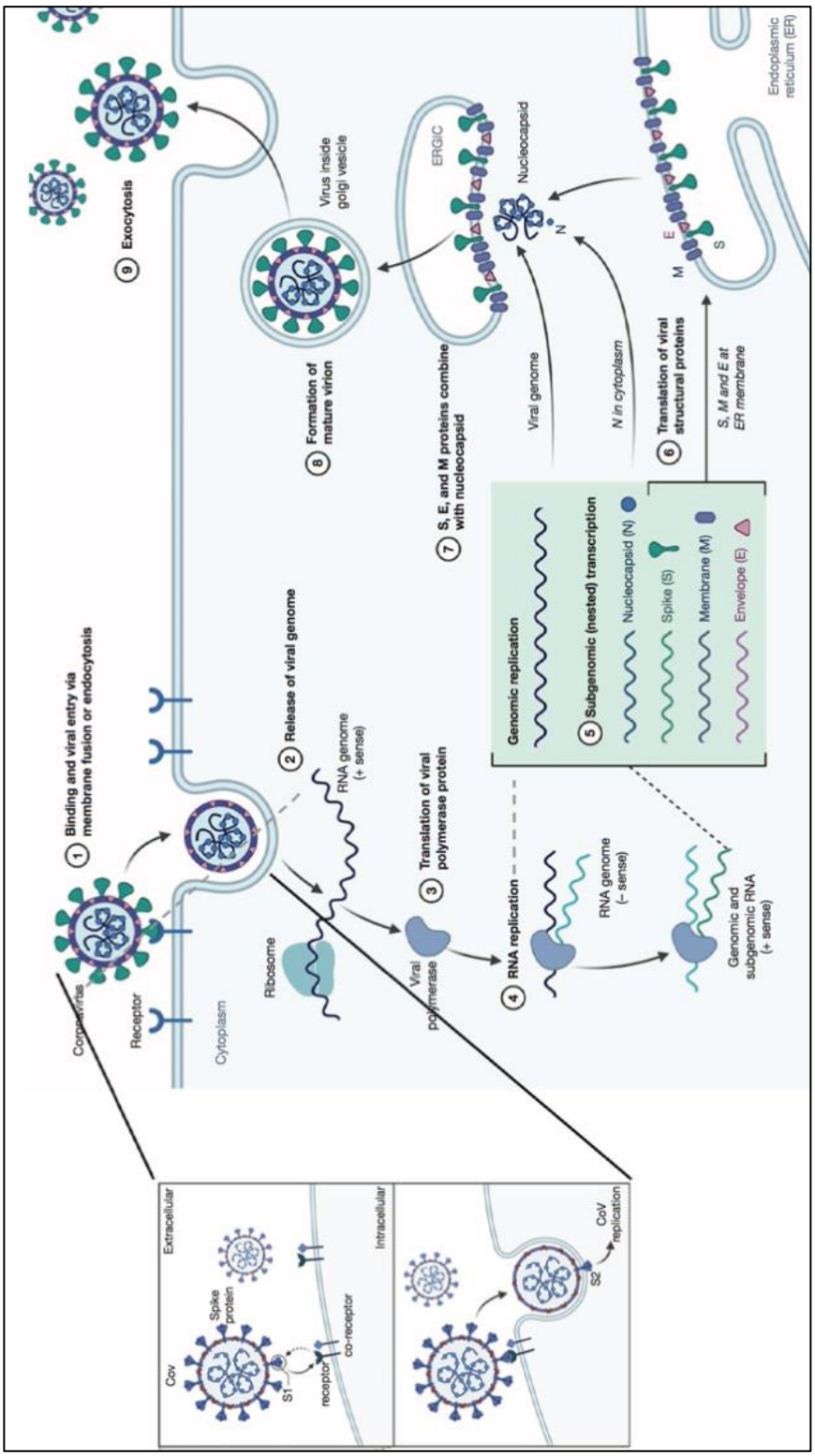

Figure 2 : Overview of the replication cycle of CoVs. The CoV enters the host cell via receptor-mediated endocytosis, releases its RNA, and uses the cell's machinery to produce viral non-structural proteins which in turn synthesise the new RNA. The new viral RNA is encapsidated and assembled with the other structural proteins, and exits the cell, forming new virions (Oostra et al*., 2007).*

1.2 Human pathogenic CoVs

The first CoVs known to infect humans, namely the hCoV-229E and hCoV-OC43 were isolated and described in the 1960s. In the majority of cases, these viruses caused relatively mild diseases of the upper respiratory tract, often considered as the "common cold" (Ludwig & Zarbock, 2020).

However, in 2003 the outbreak of a new pathogenic CoV in China, the Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV), caused a severe respiratory tract disease and raised concerns among different countries. Shortly afterwards, two new CoVs that infect humans were identified, hCoV-NL63 in the Netherlands in 2004 and hCoV-HKU1 in Hong Kong in 2005 (Van Der Hoek *et al*., 2006; Woo *et al*., 2005). These two viruses cause mild respiratory diseases and along with the hCoV-229E and the hCoV-OC43 are endemic viruses that still circulate in the human population and cause infections throughout the year (Fielding, 2011).

In September 2012, a novel CoV, even more pathogenic than SARS-CoV, was identified in a patient who was hospitalized with acute pneumonia in Saudi Arabia (Zaki *et al*., 2012). The virus was eventually named the Middle East Respiratory Syndrome Coronavirus (MERS-CoV).

Later, by the end of 2019, another novel pathogenic CoV, that caused severe respiratory tract disease emerged in China, and it was named the Severe Acute Respiratory Syndrome 2 Coronavirus (SARS-CoV-2). SARS-CoV-2 soon spread around the world, causing concern and disruption worldwide, especially due to its high transmissibility (Huang *et al*., 2020).

To highlight the potential severity CoVs, and their impact on human life, the SARS-CoV and MERS-CoV epidemics will be briefly described, before moving to SARS-CoV-2, which was the causative agent of the COVID-19 pandemic and will be the main point of this section.

1.2.1 Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV)

SARS-CoV was the first coronavirus to emerge that caused Acute Respiratory Distress Syndrome (ARDS) and a serious illness of the lower respiratory tract in humans, leading to concerns among governments and the population. The disease spread to over 30 countries and infected 8,000 people, with a mortality rate of about 10%, before eventually being controlled, owing to an unprecedented global effort. There have been no recorded cases of SARS-CoV since April 2004 (World Health Organization, 2024c).

The early prodrome includes nonspecific flu-like symptoms. The patient then gradually starts to develop more severe symptoms including fever, headache, chills, malaise and myalgias (Pormohammad *et al*., 2020). The lower respiratory tract phase of the illness starts within a week, including non-productive cough, shortness of breath and increasing respiratory distress, often requiring mechanical ventilation, whereas gastrointestinal symptoms are not very common (Coleman & Frieman, 2014).

It was initially believed that SARS-CoV originated in civet cats. However, field epidemiological and molecular screening studies have indicated that the natural reservoir of the virus was the horseshoe bat, and that the civet cat was the intermediary host to infect humans (Y. Guan *et al*., 2003). It was revealed that horseshoe bat populations host several SARS-like CoVs and also contain antibodies that target these CoVs, indicating previous infections (Lau *et al*., 2005).

Serological studies in humans have shown that SARS-CoV was indeed a novel virus that had not circulated in the human population before the 2003 outbreak in China (Y. Guan *et al*., 2003). The main ways of transmission of the virus are through close contact, via droplets or through contaminated surfaces (Low, 2004).

1.2.2 Middle East Respiratory Syndrome Coronavirus (MERS-CoV)

MERS-CoV was the next highly pathogenic coronavirus that emerged in June 2012. It was first isolated in a 60-year-old patient in Jeddah, Saudi Arabia, who presented severe respiratory illness, and died eleven days later due to renal and respiratory failure (Zaki *et al*., 2012). The virus spread to over 27 countries and is still circulating nowadays (Schoeman *et al*., 2022). As of January 2024, there have been a total of 2,609 cases and 939 reported deaths worldwide, meaning that the pathogen has a mortality rate of 36% of the infected patients (World Health Organization, 2024b). The last infection was reported in Saudi Arabia on October 26th 2023, indicating that the virus is still in circulation, causing occasional infections (European Centre for Disease Prevention and Control, 2024).

MERS-CoV causes a broad spectrum of symptoms ranging from asymptotic to mild flu-like symptoms to pneumonia and ARDS, accompanied with septic shock and multiple organ failure (Drosten *et al*., 2014). Similarly to SARS-CoV, the initial symptoms are nonspecific, and patients usually report a general malaise, sore throat, chills, low-grade fever, headache and non-productive cough, while gastrointestinal symptoms such as nausea, anorexia and abdominal pain are less frequent, but not absent (Kapoor *et al*., 2014). The lower respiratory tract is affected early in the illness and the median time from the onset of symptoms to death is about 11.5 days (Senga *et al*., 2017).

The reservoir and the source of MERS-CoV, has not been established yet. Studies suggest that bats are likely hosts of MERS-CoV because MERS-CoV-like viruses have been isolated in them (Corman *et al*., 2014). Additionally, the intermediary has not been established either, but is believed to be the dromedary camel, as MERS-CoV neutralizing antibodies were found in these animals and the incidence of human infection could be linked to them (Reusken *et al*., 2013). On the other hand, human to human transmission seems to be low and occurs only in prolonged close contact situations, as in healthcare (Schoeman *et al*., 2022).

1.2.3 Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)

SARS-CoV-2 first emerged in late December 2019 in Wuhan, China. An increase in the cases of pneumonia was initially noticed, but the cause was unknown. Later, in January 2020, the causative agent of this surge of pneumonia cases, was found to be of viral origin (Ludwig & Zarbock, 2020). The virus was given the name of SARS-CoV-2 and the disease it caused was named COVID-19. SARS-CoV-2 eventually spread worldwide, with more than 775 million cases since the outbreak, and a total of over 7 million deaths as of April 7, 2024 (World Health Organization, 2024a).

1.2.3.1 Clinical features and pathogenesis

One particularity of the COVID-19 disease, is its rather long incubation period, typically between 1 and 14 days (Bai *et al*., 2020; Lauer *et al*., 2020). This could have contributed to the spread of the virus since an infected individual could start to infect others long before experiencing any symptoms. Typically, initial symptoms include fever, dry cough, dyspnea, myalgia, fatigue and symptoms resembling pneumonia, while olfactory and taste disorders have also been reported (Chen *et al*., 2020; W. Guan *et al*., 2020). In 81% of the cases, patients only develop mild pneumonia or no symptoms at all, while 14% develop a more severe form of pneumonia accompanied by a shortness of breath and low oxygen saturation. In about 5% of the patients, the disease progresses to respiratory or multiple-organ failure (Wu & McGoogan, 2020).

The pathogenesis begins after the binding and entry of SARS-CoV-2 to the epithelial cells of the upper respiratory tract. The binding site of SARS-CoV-2 is the hACE2 receptor, found on the surface of these cells (Zhou *et al*., 2020). Once it enters the cell, the virus starts replicating and can progress into the lower respiratory tract to eventually reach the alveolar epithelial cells in the lungs (Schoeman *et al*., 2022). The progression of the virus deeper into the lungs can induce a strong immune response, leading to a cytokine storm syndrome. Similarly to SARS-CoV, this overactive immune response is believed to be the main mechanism of lung injury in patients with severe symptoms, which can lead to ARDS and eventually death (Huang *et al*., 2020).

1.2.3.2 Epidemiology

Genetic evidence suggests that SARS-CoV-2 may have originated in bats. Horseshoe bats particularly, are the source of several alpha and beta CoVs. Studies have shown that *Rhinolophus affinis* carries a CoV, denominated RaTG13, that shares 96.2% similarity with the full-length SARS-CoV-2 genome, with all the ORFs being more than 90% identical, including the hypervariable S protein (Zhou *et al*., 2020). Another bat species, *Rhinolophus malayanus* is the vector of RmYN02 that shares 93.3% similarity with the full-length genome of SARS-CoV-2 (Paraskevis *et al*., 2020). Other CoVs have been found in different species of bats with more or less genetic similarity with the SARS-CoV-2, suggesting that bats are probably the natural reservoirs of the virus (Lau *et al*., 2020). However, pangolins were initially suspected of being the source of SARS-CoV-2 as they carry CoVs. Different CoV genomes from pangolins showed between 85.5% to 92.4% similarity with the SARS-CoV-2 genome. Additionally, the RBD of the pangolin CoV from the Guangdong province in China has a high similarity with SARS-CoV-2 RBD, with only one difference in the Receptor Binding Motif (RBM), while still containing the five critical residues for hACE2 receptor binding. Nevertheless, pangolins show clinical signs and histopathological changes when they carry CoVs, suggesting that they are not the natural reservoir of the SARS-CoV-2. Bats, on the other hand, can maintain a healthy physiological state while carrying CoVs (Xiao *et al*., 2020).

1.2.3.3 SARS-CoV-2 evolution

The primary driver of SARS-CoV-2 evolution, as in all obligate parasites, is to increase its ability to transmit among its hosts. The transmission process can be divided into three main steps: the release of the virus from the infected host, its survival outside and its entry into the new host. Therefore, viral traits that enhance any of these steps could potentially be a target of natural selection (Markov *et al*., 2023). An example of the enhancement of such traits are the frequent mutations of the spike protein, to facilitate its binding to the hACE2 receptor and thereby, increase the viral entry efficiency. Such mutations are usually located in the RBD and an example is the N501Y mutation. These mutations increase transmissibility by increasing the viral load in the mucosal secretions of infected individuals since more viral particles are able to enter the host cells. Therefore, they increase infectiousness, which is defined as the capacity of an infected individual to transmit the pathogen (H. Liu *et al*., 2021; Yurkovetskiy *et al*., 2020). Other viral traits that can lead to increased transmissibility include facility of cleavage of the spike protein after binding to the host receptor (see section 1.2.3.4), broader tissue tropism, increased stability of the virion outside the host, increased duration of infectiousness and a shorter latent period (King *et al*., 2009; Lehtinen *et al*., 2021; Wallinga & Lipsitch, 2007).

RNA viruses are particularly known for their capacity of adaptive mutations in the genomic regions encoding targets of the host immune system, usually affecting the spike protein in the case of SARS-CoV-2 and often resulting in immune escape (Aleem *et al*., 2024;

Giovanetti *et al*., 2021). Immune escape is an important trait for the survival of the virus, especially in a continuously immunizing population of hosts, as in this context even a highly transmissible virus would not be able to spread. In addition, immune escape also allows for reinfection of immunized individuals, thus opening a new ecological niche for the virus. Signs of this strategy in SARS-CoV-2 were observed in the Beta and Gamma variants, which were first identified in late 2020 in South Africa and Brazil respectively. These variants showed various mutations, notably the E484K mutation that led to reduced antibody recognition and neutralization. Coincidentally, in both of these countries, epidemic data showed an increased incidence of reinfection compared to countries where other lineages were predominant (Markov *et al*., 2023). Beta and Gamma were eventually classified as Variants of Concern (VOCs). VOCs are strains of the virus that have undergone mutations with an impact on the virus's properties, such as how easily it spreads, the associated disease severity, or the performance of vaccines, therapeutic medicines, diagnostic tools, or other public health and social measures (World Health Organization, 2024d). Other variants appeared later, including the Delta and Omicron variants, also considered as VOCs. The latter accumulated an aberrant number of mutations compared to other VOCs and became the predominant variant worldwide (Markov *et al*., 2023). These variants also contain different mutations favouring immune escape.

The most frequent and impactful mutations in VOCs are found in the NTD or in the RBD of the S1-subunit of the spike protein. These domains are the main target of neutralising antibodies (nAbs) and mutations in these domains could lead to reduced effect of these antibodies against the virus (Harvey *et al*., 2021). Some of these mutations include D614G, N501Y, E484K and ΔH69/70. It is noteworthy that even though the variants emerged independently in different regions of the world, they share a set of common mutations, suggesting a possible convergent evolution (Ito *et al*., 2022).

Contrarily to the common assumption that the virulence of a pathogen will inevitably decrease over time because high virulence would kill the hosts and reduce the chances of transmission, this does not seem to be the case for SARS-CoV-2. In the spread of SARS-CoV-2, about 90% of the transmission has already been done before the average time of death, which is about 3 weeks. Hence, a high virulence is not necessarily an impediment of viability and there is no reason that it would be selected against (Markov *et al*., 2023).

Due to its role in cell entry and its high divergence, which favours the emergence of new variants and hence immune escape, the spike protein has been the subject of many studies. It has also been the subject of extensive research to develop a vaccine.

1.2.3.4 SARS-CoV-2 spike protein

The SARS-CoV-2 S protein is 1273 amino acids long (Xia, 2021a). Similarly to other CoVs, as explained in section 1.1.3.1, the S protein is divided into the S1 and S2 subunits, each playing a distinct role in the virus entry into host cells. S1 is involved in host-receptor binding, while S2 is responsible for the membrane fusion following receptor binding (Zhang *et al*., 2021).

The spike protein is composed of different domains, each having a specific function. S1 is composed of the NTD, the RBD, itself containing the Receptor Binding Motif (RBM), and two subdomains on the C-terminus, CTD1 and CTD2. The S1-subunit is preceded by a signal peptide (SP). S2 is composed of the fusion peptide (FP), the heptad repeat 1 (HR1), the central helix (CH), the connector domain (CD), the heptad repeat 2 (HR2), the transmembrane segment (TM) and the cytoplasmic tail (CT). The S1 and S2-subunits are separated by a S1/S2 cleavage site and the S2-subunit contains another cleavage site called S2' (Kovalenko *et al*., 2023).

The SARS-CoV-2 S protein forms trimers and binds to the hACE2 receptor through the RBD. The spike protein is known to adopt two main conformations, the "RBD-up" conformation, which is receptor-accessible, enabling the spike protein to bind to hACE2 receptor, and the "RBD-down" in which the RBD is inaccessible to the receptor (Figure 3). It has been shown that the S protein needs at least one RBD in the "up" conformation for successful viral infection. Cryo-EM structures of the soluble S protein and hACE2 monomer complex have shown that one S protein can bind to up to three hACE2 receptors (Kordyukova & Shanko, 2021; Zhou *et al*., 2020).

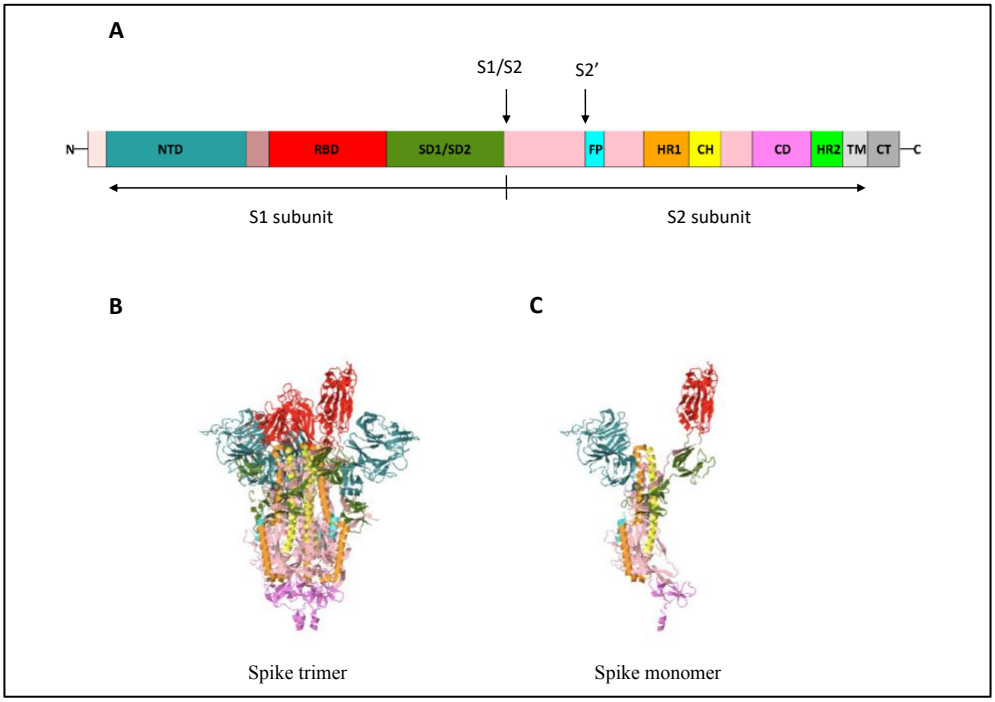

Figure 3: (A) Schematic illustration of the spike protein topology. (B) Prefusion structure of the spike trimer with one of the monomers in the "RBD-up" conformation(red). (C) Spike monomer in the "RBD-up" conformation (red) (Kovalenko et al*., 2023).*

The S1-subunit adopts a "triangular" shape in the prefusion conformation, where the NTD is found in one side and the RBD, CTD1 and CTD2 are found at the opposite side of the triangle (Zhang *et al*., 2021). In this configuration, the central helical bundle of the S2 fragment is encircled and the N-terminal end of HR1 is directed towards the viral membrane. In the post-fusion conformation, S1 is dissociated as a monomer by the action of host proteases. The S1-S2 boundary is cleaved by a furin prior to binding with the hACE2 receptor, resulting in conformational changes that expose the HR1 of S2 to the extracellular medium. In the meantime, the two subunits remain attached. A cleavage at the S2' site by serine 2 or cathepsin proteases, depending on the mechanism of entry (direct entry or endocytosis), leads to activation of the S2 (Jackson *et al*., 2022). The HR1 undergoes conformational changes, translocating its N-terminal end to project the FP towards the host cell membrane. These conformational changes bring the viral and host membranes close, eventually leading to fusion of the two, allowing the viral genome to enter in the host cytoplasm (Jackson *et al*., 2022; Zhang *et al*., 2021).

The S protein is highly glycosylated, with 22 *N*-glycosylation sites per monomer, 13 of which are found on the S1-subunit (Watanabe *et al*., 2020). As for many proteins, glycosylation

is an important post-translational modification for S, as it has been shown to play a wide range of important roles in protein folding, thermal stability, immunogenicity immune evasion, and hACE2 binding (Watanabe *et al*., 2019).

1.2.3.4.1 N-terminal domain

The NTD of the SARS-CoV-2 spike protein is 292 amino acids long, spanning from residues 14 to 306 (Cai *et al*., 2020). It contains 8 *N*-glycosylation sites, in residues 17, 61, 74, 122, 149, 165, 234 and 282 (Watanabe *et al*., 2020). Its three-dimensional structure can be divided into three regions, top, core and bottom (Figure 4A) and it essentially adopts a galectin-like fold, including ß sheets and ß strands, as well as loops which connect these structures (Cai *et al*., 2020; Yao *et al*., 2020).

The NTD function has not been conclusively elucidated, but several studies have suggested that it might play a role in the adherence to the host cell by interacting with glycans on its surface (Awasthi *et al*., 2020; Chi *et al*., 2020; Klinakis *et al*., 2021). Neutralizing antibodies directed against the NTD have also been found on the serum of infected patients (Chi *et al*., 2020). A comparison of the sequences and structures of the SARS-CoV and SARS-CoV-2 spike protein NTDs has shown that the SARS-CoV-2 NTD contains longer loops that have accumulated a large number of mutations. Some of these mutations have been associated with changes in the structure of the NTD and since the loops play a role in stabilizing the exposed tertiary structure of the NTD, structural changes may affect the recognition of this epitope by neutralizing antibodies. Epidemiological data of these mutations have shown that they have been positively selected and have been found in fast-spreading variants of SARS-CoV-2 (Klinakis *et al*., 2021).

Another study showed that the SARS-CoV-2 spike NTD contains loops that are structurally analogous to the loops of the MERS-CoV spike NTD, which is known to form sialoside-binding pockets that play an active role in the dual binding strategy of MERS-CoV, by binding to surface glycans (Awasthi *et al*., 2020; Fung & Liu, 2019).

1.2.3.4.2 Receptor Binding Domain

The RBD is 223 amino acids long spanning from residues 319 to 541 (Lan *et al*., 2020). It contains 2 *N*-glycosylation sites, in residues 331 and 343 (Watanabe *et al*., 2020). The RBD can be further divided into two subdomains: a core subdomain composed of 5 ß-strands (ß1, ß2, ß3, ß4 and ß7) connected by short helices and loops, and the RBM found between ß4 and ß7, containing a long insertion with the strands ß5 and ß6 connected by short helices. The RBM spans from residues 438 to 506 and is the region of the spike protein that directly interacts with the hACE2 receptor (Figure 4B). The main types of interactions that occur between the RBM and hACE2 include hydrogen bonds, salt bridges and to a lesser extent, hydrophobic interactions (Lan *et al*., 2020). Some of the RBM residues implicated in these interactions, such as N501 or E484, have undergone mutations that are present in most fast-spreading variants. These mutations lead to increased affinity for the hACE2 receptor, increasing the fitness of the virus (H. Liu *et al*., 2021; Starr *et al*., 2020).

The RBD is the main target of nAbs, produced either after infection or vaccination. Up to 90% of the nAbs present in the serum of infected patients are directed against the RBD (Piccoli *et al*., 2020). The main neutralization mechanisms of SARS-CoV-2 by nAbs consist in preventing binding to the hACE2 receptor by binding to the RBM, or in preventing the RBD to adopt the “up” conformation (Finkelstein *et al*., 2021; Piccoli *et al*., 2020).

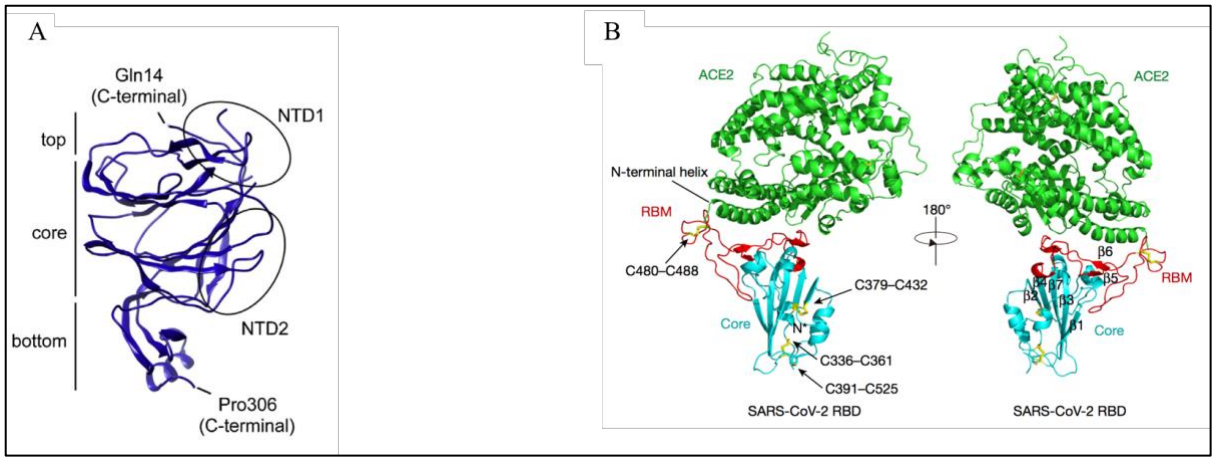

Figure 4 : (A) Close-up view of the SARS-CoV-2 spike NTD (Zhang et al.*, 2021). (B) Structure of the SARS-CoV-2 spike RBD (cyan), including the RBM (red), bound to the hACE2 receptor (green) (Lan* et al.*, 2020).*

1.2.3.5 Treatment of the SARS-CoV-2 infection

The emergence of the new SARS-CoV-2 constituted a major public health emergency, and the search for efficient treatments was and remains a high priority. Since the beginning of the pandemic, a large number of treatments have been predicted to be effective against SARS-CoV-2 and have been investigated, including antivirals, antibiotics, antimalarials and immunomodulatory drugs (Drożdżal *et al*., 2021).

One of the main treatment strategies is the use of monoclonal antibodies (mAbs). mAbs mainly act against SARS-CoV-2 by targeting the spike protein, preventing its binding to the hACE2 receptor and consequently blocking viral entry into the cell. A combination of multiple mAbs that target different epitopes on the spike protein for increased efficiency has also been proposed (AstraZeneca, 2021). The use of convalescent plasma from previously infected individuals, containing antibodies directed against SARS-CoV-2 has also been investigated (Drożdżal *et al*., 2021).

Antiviral drugs are another important strategy suggested for the treatment of COVID-19. They function mainly by targeting viral replication. Remdesivir was one of the most promising antiviral drugs that were investigated. It is an adenosine analogue that is metabolized into remdesivir triphosphate. Remdesivir triphosphate is a structural analogue of adenosine triphosphate (ATP) and competes with the natural ATP substrate to selectively inhibit the viral RdRp. This results in delayed chain termination during replication and consequently, inhibition of the viral replication (Singh *et al*., 2020).

Immunomodulating drugs have also been considered as COVID-19 treatment option. The use of immunomodulators is expected to be effective in the treatment of COVID-19 as one of the main mechanisms of organ injury caused by SARS-CoV-2 is due to an overactive immune response.

However, even though some of the above-mentioned treatments have shown interesting effects in the treatment of COVID-19, and many more are still in trials, the studies are conducted on a limited number of patients, during a limited period of time and the data obtained are insufficient in terms of clinical outcome and safety. Many others have shown very limited beneficial effects or not at all. In addition, a lot of other treatments are not in clinical trials yet due to safety concerns. Furthermore, the COVID-19 disease symptoms vary among patients and treatment needs to be tailored towards specific symptoms in order to be effective (Drożdżal *et al*., 2021).

In conclusion, there is no effective and ubiquitous treatment yet. For the vast majority of SARS-CoV-2 infections, treatment is therefore limited to supportive care. However, in the context of a highly transmissible virus such as SARS-CoV-2, still in circulation four years after the outbreak, likely to continue to spread in the future, and where the intermediary host has not been conclusively identified yet, supportive care is insufficient and prevention strategies such as vaccination must be a priority (Drożdżal *et al*., 2021; World Health Organization, 2024e).

1.2.3.6 SARS-CoV-2 vaccines

COVID-19 vaccines were rapidly developed following the SARS-CoV-2 outbreak and were approved for use in record time. Various strategies were used for the development of these vaccines, including live-attenuated or inactivated viruses, or new generation vaccines such as recombinant vaccines, virus-like particles vaccines, viral vector vaccines and in particular mRNA vaccines (Kovalenko *et al*., 2023). The latter technology had never been approved for human use previously. mRNA vaccines were among the first to be approved for emergency use by the US Food and Drug Administration (FDA), already in late 2020. Examples of the approved mRNA vaccines include Comirnaty® (Pfizer/BioNTech) and Spikevax® (Moderna), both of which containing mRNA encoding for the S protein of SARS-CoV-2, delivered in lipid nanoparticles. Some non-replicating adenovirus-based vaccines were also developed and approved for emergency use, including Vaxzevria® (Oxford/AstraZeneca), Jcovden® (Johnson&Johnson) or SputnikV® (Gamaleya Research Institute). Modified versions of the above-mentioned vaccines, to include new variants, or new vaccines were later introduced and approved by national agencies (Y.-D. Li *et al*., 2020).

In contrast to treatments, COVID-19 vaccines have been shown to be effective in the population, particularly in protecting against more severe illnesses and restricting the spread of the SARS-CoV-2 virus (Drożdżal *et al*., 2021). However, several studies have shown that the protecting effect of vaccines decreases over time and a second dose may be needed. In some cases, especially in patients with a compromised immune system, a third dose, may also be needed (Goldberg *et al*., 2021; Juno & Wheatley, 2021; Shroff *et al*., 2021). In addition, some important limitations arose regarding the vaccine distribution, particularly for mRNA vaccines. Those were initially distributed in small numbers only in highly developed countries. The limitations mainly arise from the fact that RNA is very unstable and requires very low temperatures for storage and transportation, in some cases down to -80°C (Uddin & Roni, 2021). This limits the distribution possibilities in less developed countries lacking cold chain storage infrastructure. Furthermore, the protection effect of a given vaccine is reduced against new variants (Wang *et al*., 2021).

In this context, another strategy for the development of vaccines could be of great interest. Recombinant protein vaccines are promising, because they allow to overcome some of the limitations related to other types of vaccines while showing promising results in terms of immune response elicitation. Recombinant protein vaccines are more stable, for example, than mRNA vaccines, and thus require less complex cold chain infrastructure. In addition, theoretically, they allow for the reuse of the same vector for revaccination, as opposed to viral-vector vaccines. Moreover, recombinant protein vaccines are considered among the safest in terms of side effects (Kovalenko *et al*., 2023).

1.3 Production of recombinant proteins

Recombinant proteins are produced using genetic engineering, by inserting a gene into an organism in which it is not naturally present but can be transcribed and translated into the desired protein. Recombinant proteins can be used for a wide range of applications such as structural and biophysical studies, functional assays, biomarkers, *in vitro* or *in vivo* mechanistic studies and even as therapeutics (Assenberg *et al*., 2013). Since the approval of insulin in 1982, over 430 recombinant biopharmaceutical products have been approved for human use (Walsh & Walsh, 2022). Some important recombinant biopharmaceuticals include antibodies, growth factors and a variety of vaccines (Santos *et al*., 2016; Tan *et al*., 2021).

Different platforms can be used to produce recombinant proteins. At present, the main platforms are bacterial cells, yeast cells, insect cells, mammalian cells, whole plants and plant cells. Each platform has its advantages and disadvantages in terms of cost, protein biochemistry and thus, the clinical efficiency of the product for the purpose of biopharmaceuticals. A number of constraints therefore need to be taken into account when optimizing the production of recombinant biopharmaceuticals. Furthermore, biopharmaceuticals do not necessarily have to be a perfect copy of human proteins because a small modification does not automatically alter their therapeutic value. Sometimes small modifications are deliberately introduced into recombinant proteins to enhance their pharmacokinetics or even change their pharmacodynamics (Dingermann, 2008). This opens up countless new possibilities for the development of new biopharmaceuticals.

1.3.1 Expression platforms for recombinant proteins

The first expression systems used for recombinant protein production were microbial cells. *Escherichia coli* is the most studied and its genetics is far better characterized than that of any other microorganism. The first recombinant therapeutic was insulin produced in *E. coli* (Goeddel *et al*., 1979). Since then, several others, mainly growth hormones and other relatively simple proteins, have been produced in this expression system. Microbial systems are simple to transform, cultivate and have a short generation time, making them cost-effective and industrially scalable (Chadd & Chamow, 2001; Dingermann, 2008). However, they are unable to modify proteins by glycosylation, which is the case for most human proteins. This limits their use for the production of more complex recombinant proteins (Baneyx, 1999). Proteins produced in bacterial cells usually accumulate in the form of inclusion bodies that need to be subsequently renatured (Dingermann, 2008). Another major drawback of bacterial cells is the risk of endotoxin production, which can lead to septic shock (Santos *et al*., 2016).

Yeast cells, mainly *Saccharomyces cerevisiae*, combine the fast growth and cost-effective cultivation of bacteria with the possibility of glycosylation, while producing high yields. However, the glycoproteins produced in yeast cells are rich in mannose, which are rare in natural human proteins, potentially leading to an immunogenic effect and also reduced efficiency (Dean, 1999).

Mammalian cells are currently the main platform for the production of therapeutic recombinant proteins (Lalonde & Durocher, 2017). They can produce complex proteins with glycosylation profiles that are highly similar to human proteins, since the glycosylation machinery is highly conserved among mammalian species (Dingermann, 2008). A major advantage of mammalian cells, also shared by microbial, yeast, insect and plant cells, is that they can be cultivated in contained bioreactors, which could facilitate compliance with Good Manufacturing Practices (GMPs). Another advantage of mammalian cells is that recombinant proteins can be secreted in the extracellular medium in their native form, which is not the case for *E. coli*, and there is no size limit (Gerd Gellissen *et al*., 2005). Some of the main mammalian

cell lines are the Chinese Hamster Ovary (CHO) cells, the Baby Hamster Kidney (BHK) cells or the Human Embryo Kidney 293 (HEK293) cells (Lalonde & Durocher, 2017). CHO cells are the most widely used expression system for producing biotherapeutics. This is mainly due to their glycosylation machinery, which produces proteins with very similar glycosylation profiles to human proteins, and also to their ability to achieve high production rates, up to 10 g/L, making them industrially scalable in suspension cultures (J. Y. Kim *et al*., 2012). Additionally, CHO cells are less susceptible to human viruses. Compared to other mammalian cells, many genes associated with viral entry are not expressed, even though they are present in their genome (Xu *et al*., 2011). However, this does not completely eliminate the risk of contamination with human pathogens. "As a rule of thumb", in terms of safety, it would be preferable to choose an expression system phylogenetically distant from humans, as the risk of contamination with human pathogens would be lower (Dingermann, 2008). Another drawback of CHO, or mammalian cells in general, is that they have high costs in terms of safety and cultivation, due to expensive culture media (Santos *et al*., 2016).

Transgenic animals have also been investigated and used for the production of recombinant proteins. As in the case of mammalian cells, they are capable of producing proteins with glycosylation patterns similar to native human proteins. Directing the recombinant protein production into the urine or milk of transgenic animals is an interesting approach as it would simplify their obtention and purification. This strategy was used for the production of a C1-esterase inhibitor (Ruconest®), produced in transgenic rabbit milk, which was approved by EMA and FDA in 2011 and 2014, respectively (Walsh, 2014). However, growing awareness of ethical issues related to animal experimentation has led to stricter regulations in recent years.

Insect cells are another platform for the production of recombinant proteins. They have lower cultivation costs compared to mammalian cells, yet they are still expensive compared to bacterial and yeast cells, or plants. They are not susceptible to human pathogens and are also capable of making posttranslational modifications. Insect cells are also easily industrially scalable. Some of the most used lines include Sf-9, Sf-21, Tn-368 and High-Five (Drugmand *et al*., 2012).

Finally, plants, either as whole plants or cells in suspension, are another expression system for recombinant protein production. Plants have gained in interest over the last decades due to their ability to perform higher eukaryotic posttranslational modifications, their well-defined and cost-effective cultivation medium, and their safety, as they are less susceptible to human pathogens.

1.3.2 Molecular farming

Molecular farming is the production of pharmaceutically valuable recombinant proteins in plants (Fischer & Emans, 2000). The main interest of molecular farming is that it combines the capability of higher eukaryotic posttranslational modifications, enabling to produce complex mammalian proteins, with a cost-effective and safe production process in controlled culture conditions. Molecular farming can be achieved in various settings such as in transgenic plants, transformed plant organs or transformed suspension cells.

Tobacco, cereals such as wheat or barley, fruits and vegetables including potatoes, tomatoes, strawberries and bananas are some of the used platforms for the production of biopharmaceuticals in plants. Among these platforms, tobacco is the most studied and used, due to some interesting properties, such as a relatively easy introduction of foreign genes, the possibility of stable transformation and high production of biomass (Lee *et al*., 2023). Furthermore, protein production in whole tobacco plants is mainly based on transient expression in the leaves, which eliminates the need for flowering, reducing the risk of gene

leakage in the environment (Conley *et al*., 2011). Duckweed and moss, which are also eukaryotic organisms, have also been considered for recombinant protein production. They could potentially present some advantages, mainly in terms of low cultivation cost and easy scale-up, but they are less well known (Lee *et al*., 2023).

Despite these advantages, the use of plants as a platform for producing biopharmaceuticals also presents some limitations. Firstly, even though plants are eukaryotic organisms and possess the ability to make posttranslational modifications such as glycosylation, the glycosylation profile is different to that naturally occurring in humans. Plant N-glycans often contain α1,3-fucose and ß1,2-xylose residues, which are absent in human proteins. This could be a potential problem in terms of immunogenicity and efficiency of plant-made biopharmaceuticals (Sukenik *et al*., 2018). Additionally, the production yields in plants are low, mainly due to the degradation of proteins by intracellular and extracellular proteases (Santos *et al*., 2016). Consequently, the production yields in plants are not yet comparable to those achieved in mammalian or bacterial cells.

However, efforts are being made to address these limitations. Glycoengineered plants or suspension cell lines have been obtained by knocking-out specific glycosylation enzymes responsible for the glycosylation of proteins and/or introducing enzymes enabling to produce proteins with human-like glycosylation profiles (Bosch *et al*., 2013; Castilho & Steinkellner, 2012; Mercx *et al*., 2017; Herman *et al*., 2021 and 2023). As for the degradation, multiple strategies can be used. Degradation by apoplastic and extracellular proteases can be avoided by targeting the protein in specific intracellular compartments. However, a major drawback to this strategy is that the cell would need to be lysed to recover the protein, thereby releasing various intracellular proteins, resulting in a more complex downstream processing (Santos *et al*., 2016). A better approach would be to prevent the action of proteases. This can be achieved through the inactivation of proteases by knocking out the corresponding genes, by co-expressing a protease inhibitor or by antisense mRNA (T.-G. Kim *et al*., 2008). Nonetheless, this strategy implies that the proteases degrading the proteins of interest are known, which is not always possible, as protein degradation by proteases in known to be protein specific. Other strategies include the use of protease inhibitors such as PMSF or the use of protein decoys such as gelatine or Bovine Serum Albumin (BSA) (Baur *et al*., 2005; Navarre *et al*., 2012).

Molecular farming could potentially be a valuable tool for the production of therapeutics or diagnostics in the context of an emergency situation, such as the outbreak of an infectious disease or even a bioterrorist attack (Sukenik *et al*., 2018). For example, in response to the outbreak of A/H1N1 virus in 2009, influenza virus-like particles (VLP) were produced in *Nicotiana benthamiana* leaves within three weeks of sequence availability (D'Aoust *et al*., 2010). A similar strategy was followed after the outbreak of SARS-CoV-2 in 2020 (section 1.3.4). The advantage would be that molecular farming enables rapid and scalable production, with minimal cold-chain requirements, thus promoting equitable access to therapeutics, including less developed countries (Eidenberger *et al*., 2023).

1.3.2.1 Nicotiana expression hosts

1.3.2.1.1 Nicotiana benthamiana and Nicotiana tabacum plants

Nicotiana species are the most widely used expression systems in molecular farming, mainly in the form of transgenic plants. The main species used for recombinant protein production are *N. benthamiana* and *N. tabacum* (Lee *et al*., 2023). They offer several advantages such as an easy incorporation of foreign genes, a fast growth rate and a high biomass yield, reaching up to 150 metric tons of biomass per hectare per year (S. J. Sheen, 1982). The production yields of recombinant proteins are usually higher in tobacco than in

other species, reaching grams of protein per kilogram of leaves 5 to 7 days post DNA delivery (Eidenberger *et al*., 2023). The proteins are then extracted from the leaves and further purified. An important advantage of tobacco is that it is a non-food and non-feed plant, which alleviates some of the regulatory barriers related to recombinant proteins entering the food chain. On the other hand, tobacco has the drawback of containing nicotine, even though some cultivars have lower concentrations, and higher amounts of toxic alkaloids compared to other plants (Fischer *et al*., 2004).

In a study conducted by Conley *et al.*(2011), 52 varieties of Nicotiana species were compared, in terms of growth rate, amount of biomass produced and Total Soluble Protein (TSP) concentration. The study included transgenic plants and as well as transient expression of four recombinant proteins through agroinfiltration, namely the human erythropoietin (EPO) and interleukin-10 (IL-10), an antibody against *Pseudomonas aeruginosa* (APA) and a hyperthermostable alpha-amylase from *Pyrococcus furiosus* (PFA). The study showed that *N. tabacum* produced the highest concentration of recombinant proteins, it was among the highest in terms of biomass production and presented relatively low alkaloid concentrations. Therefore, the authors concluded that *N. tabacum* was the most effective plant for large-scale transient expression of recombinant proteins, for the traits that were tested (Conley *et al*., 2011).

1.3.2.2 N. tabacum Bright Yellow 2 suspension cells

Plant cells in suspension have been attracting increasing interest for the production of complex recombinant proteins as an alternative platform to mammalian cells (Reuter *et al*., 2014). In addition to the advantages mentioned in section 1.3.2 of using plants for recombinant protein production, either as whole plants or in specific organs, cell suspension cultures share some of the advantages of mammalian or bacterial cells. While whole plants are grown in fields or greenhouses, suspension cells are cultivated in bioreactors, which leads to more control over the culture conditions and consequently more homogeneity between batches. Additionally, the use of contained bioreactors makes plant suspension cells more amenable to comply with GMPs, easing the approval of new therapeutics by regulatory agencies. Furthermore, suspension cells lack fibres, waxes or numerous secondary metabolites, which are commonly present in whole plants and could potentially interfere with the downstream purification (Reuter *et al*., 2014).

N. tabacum Bright Yellow 2 (BY-2) cells are the most commonly used suspension cells for producing recombinant proteins. The BY-2 cell line was established from callus induced on a seedling of *N. tabacum* BY-2 cultivar, at the Hantan Tobacco Experimental Station, Japan Tobacco Company by Kato *et al.* in 1972 (Nagata *et al*., 1992). Originally, the BY-2 cell line was developed at an industrial scale to investigate the possibility of using them as a raw material for cigarettes. While Nagata *et al.*(1992) referred to BY-2 as the "HeLa cell in the cell biology of higher plants", Reuter *et al.*(2014) argued that BY-2 could also be referred as the "CHO cells of molecular farming". BY-2 cells have several important advantages over other plant suspension cell lines. First, BY-2 cells have a high growth rate, with the cell density, multiplying between 80 and 100-fold over one week under optimal conditions, while having a highly synchronized growth phase and an average doubling time of approximately 16 to 24 h (Nagata *et al*., 1992). Additionally, they are readily amenable to *Agrobacterium*-mediated transformation (Nagata *et al*., 1992). According to Geelen & Inzé (2001), the use of *Agrobacterium* strain LBA4404 virG is two to five-fold more efficient in generating calli, than other *Agrobacterium* strains, and up to 500 calli can be obtained from 4 mL of BY-2 cells co-cultivated with this strain.

Despite these advantages, BY-2 cells have limitations in terms of production yields, as other plant platforms. While mammalian cells are able to routinely produce recombinant

proteins at yields up to 5 g/L, BY-2 cells can only produce 0.1 to 0.5 g/L (Walsh, 2010; Xu *et al*., 2011). As in whole plants, one of the reasons for the low yields are degradation by extracellular proteases and silencing, which is a common mechanism of defence against pathogens, present in plants (Santos *et al*., 2016).

Regardless of the yield limitations, suspension cells, in particular BY-2, are a promising alternative platform for the production of valuable recombinant proteins. In addition to the research that is being made to improve the BY-2 and other plant cell suspension cultures, lessons learned from other platforms such as mammalian or bacterial cells, which have been used for longer could be applied to accelerate the process (Santos *et al*., 2016).

1.3.2.3 Agrobacterium-mediated transformation

Agrobacterium tumefaciens is a natural plant pathogen that can infect plants through damaged tissues and cause the crown gall tumours (Gelvin, 2003). It infects a large number of plant species, mainly dicot angiosperms (Anderson, 1979; De Cleene & De Ley, 1976). Its genome is composed of a circular chromosome, a linear chromosome and two plasmids, pAtC58 and pTiC58 otherwise known as the tumour-inducing (Ti) plasmid (Goodner *et al*., 2001). The latter contains the *vir* genes (A, B, C, D, E, F), which are responsible for the virulence of *A. tumefaciens* and are the main determinant of the host range (Loper & Kado, 1979).

A. tumefaciens has been extensively used and still remains among the most commonly used methods for delivering foreign genes into plant cells, both for transient and stable transformation. It is based on its ability to transfer and integrate a portion of the Ti plasmid, called transfer DNA (T-DNA) into the plant cell genome (Gelvin, 2003). The T-DNA is flanked by specific bordering sequences of around 25 bp that are highly homogenous, and only the region between these borders is transferred to the plant cell (Yadav *et al*., 1982).

Most *vir* genes are expressed in response to plant signal compounds, typically phenolic compounds, after infection with *A. tumefaciens*. Thereby, acetosyringone is usually added during transformation experiments to enhance the efficiency (Kapila *et al*., 1997).

1.3.3 Downstream purification of recombinant proteins using affinity tags

After production, recombinant proteins must be recovered from the total intracellular soluble proteins (TSP) or from the extracellular medium in the case of secreted proteins. The protein is, therefore, often found in a complex mixture of other, host-related, proteins and requires purification. Downstream purification can be facilitated by labelling recombinant proteins with affinity tags, as they allow purification of proteins without the need of knowing their biochemical properties. Affinity tags can be defined as exogenous amino acid sequences that have high affinity for a particular chemical or biological ligand (Arnau *et al*., 2006). They need to be taken into account when designing genetic constructs encoding the recombinant protein and establishing the purification protocol.

The addition of an affinity tag often leads to unpredictable effects on the protein. They have been reported to protect the protein from proteolytic degradation and facilitate protein folding (Arnau *et al*., 2006; Mayer *et al*., 2004). On the other hand, tags have also been reported to alter the protein conformation or its biological activity in some cases (Chant *et al*., 2005; Fonda *et al*., 2002). An increase in protein yields has also been reported by some authors (Goel *et al*., 2000; Sun *et al*., 2005).

Histidine tags are the most commonly used affinity tags. They usually contain between 5 and 15 histidine amino acids and have a high affinity for chelated metallic ions, particularly nickel. Purification takes place in a chromatography column filled with a nickel resin. Tag-

fused proteins are retained due to the interaction of the imidazole side chain of the polyhistidine tag with the immobilized nickel ions on the column. Purification of His-tagged proteins from the TSP can be achieved in one step with up to 90% yield (Arnau *et al*., 2006). An interesting property of His tags is their relatively short size and small molecular weight, making them less likely to affect protein structure or function (Zhao & Huang, 2016). The V5 tag is another commonly used tag for the purification of recombinant proteins. It is an epitope tag (14 amino acid residues long) derived from the V protein of simian virus 5 (SV5) (Hanke *et al*., 1992). This epitope was recognized by a monoclonal antibody named P-k (Southern *et al*., 1991). So the purification with the V5 tag is done using an affinity matrix containing the antibody and the elution is usually done by V5 peptide or alternatively with low pH (Y. Li, 2010). Other commonly used affinity tags are based on protein-protein interactions and usually also make use of monoclonal antibodies. Some examples of such affinity tags include Myc, HA , FLAG and Streptag II tags (Arnau *et al*., 2006; Kent, 1999; Slootstra *et al*., 1997).

Finally, tag removal strategies also need to be considered as the effect of an affinity tag on the proteins is rather unpredictable. This is especially the case for biopharmaceuticals where the protein needs to be similar to those found in humans or the recipient species. Common strategies used to this end include the use of recombinant peptidases, that recognise a specific site and cleave the tag fused to the protein of interest. These peptidases are themselves tagged in order to be removed afterwards (Arnau *et al*., 2006).

1.3.4 Production of recombinant SARS-CoV-2 spike protein and its Receptor Binding Domain in plants

Among the many efforts made to mitigate the COVID-19 pandemic, the production of recombinant SARS-CoV-2 related proteins, as a means of diagnostics or therapeutics, was largely studied. Initially, the focus was on the traditional and faster producing platforms such as mammalian and microbial systems, but plant production platforms were also considered. Plant-produced SARS-CoV-2 proteins include the S, M, E and N proteins or domains such as the spike S1-subunit or the spike RBD. The production of all of these proteins was achieved in *N. benthamiana* leaves by agroinfiltration (Ceballo *et al*., 2022; Jung *et al*., 2022; Moon *et al*., 2022; Williams *et al*., 2021). Yields ranging from 4 to 25 µg of RBD/g of fresh weight have been reported in *N. benthamiana* (Diego-Martin *et al*., 2020; Siriwattananon *et al*., 2021). A plant-made vaccine was even approved for human use in Canada, but never reached the market because the company closed down. The vaccine was produced by Medicago Inc. and was named Covifenz®. It consists of SARS-CoV-2 spike VLPs co-administered with the GSK's pandemic adjuvant (GSK, 2022). However, even though transient expressions in *N. benthamiana* assures a rapid production, it requires continuous maintenance of plants at a certain developmental stage, making its upscaling more difficult, even though feasible.

On the other hand, cell suspension cultures for recombinant protein production, cultivated in continuous bioreactors, are easier to scale up and comply with GMP practices. This could facilitate approval for human use (Santos *et al*., 2016). In this context, Rebelo *et al.*(2022) investigated the use of plant cell suspension cultures as an alternative to whole plants for the production of SARS-CoV-2 proteins. A stabilized and soluble version of the SARS-CoV-2 S protein as well as the spike RBD, as previously used by Amanat *et al.*(2020), were produced in *N. tabacum* BY-2 or *Medicago truncatula* A17 suspension cells, using either transient or stable expression. The genetic constructs contained the S protein sequence (GenBank: MN908947.3) or the RBD sequence (amino acids 319-541) optimized for expression in plant cells (Amanat *et al*., 2020; Rebelo *et al*., 2022). A polyhistidine (6xHis) tag was added on the C-terminus of the proteins for subsequent purification. The authors reported successful production and secretion of both the spike and the RBD, in both systems. However,

more degradation was observed in BY-2 cells, for both proteins. This was attributed to the action of extracellular proteases. BY-2 cells seem to have higher protease contents compared to Medicago cells, which could be the reason for the higher proteolytic activity observed (Santos *et al*., 2018).

Purification of the His-tagged RBD produced in BY-2 cells on nickel columns was attempted in the laboratory, but the protein failed to bind to the resin (Rebelo *et al*., 2022; Peeters and Navarre, unpublished data). In addition, no bands were detected in anti-His western blots performed on the His-tagged RBD produced in BY-2 cells. Together these results suggest that the polyhistidine tag might be (partially) degraded. In contrast, the histidine tag was indeed detected in the full-length S protein by performing nickel purification and anti-His western blots (Rebelo *et al*., 2022; Peeters and Navarre, unpublished data).

Importantly, the recombinant S protein produced in both systems contained the same *N*-glycosylation sites as the ones produced in mammalian HEK293 cells, but with a plant-specific glycosylation profile (Rebelo *et al*., 2022).

1.4 Context and objectives

The SARS-CoV-2 spike protein is heavily glycosylated, and glycosylation is a well-known immune escape strategy among viruses. SARS-CoV-2 is also thought to use this strategy to escape the immune system of its hosts (Markov *et al*., 2023). In addition, glycosylation plays a role, among others, in protein folding and thermal stability (Herman *et al*., 2023; Watanabe *et al*., 2020). Therefore, understanding the glycosylation of the spike protein and its impact on the protein properties could help to better understand the SARS-CoV-2 pathobiology and consequently, the development of treatments, diagnostics or vaccines.

On the other hand, plant cell suspensions have gained attention over the last decade as an alternative to mammalian or bacterial cells for producing biopharmaceuticals. However, although similar, the glycosylation machinery of plants is different from that of humans and as the glycosylation profile of the SARS-CoV-2 S protein depends on the host/production platform, the question of its impact on spike properties has been raised.

Among these different aspects, the impact of the *N*-glycosylation profile of the SARS-CoV-2 S protein on its binding to the hACE2 receptor is currently being investigated in our laboratory by Marie Peeters. So far, the S1-subunit of the SARS-CoV-2 S protein and a truncated version of 215 amino acids of its RBD, called "RBD215" have been produced by transient expression in *N. benthamiana* leaves, used for screening the different constructs. The most promising constructs were then used for the more time-consuming stable transformations of *N. tabacum* BY-2 suspension cells.

To study the interaction between the RBD and the hACE2 receptor, the protein must be purified from the extracellular medium of BY-2 cells. To this end, several purification strategies have been considered, the most convenient being the use of affinity tags such as polyhistidine or V5. In our laboratory, no RBD production was achieved in *N. benthamiana* leaves when fused to a V5 tag. Therefore, the "*RBD-V5*" construct was not tested in BY-2 cells. On the other hand, when fused to a polyhistidine tag (either 6xHis or 10xHis), significant production of the RBD was obtained in *N. benthamiana* leaves, even though the tag was not detected. The RBD constructs with a polyhistidine tag were further investigated in BY-2 cells. Indeed, when transformed with the "*RBD-His*" constructs, production of RBD was achieved in BY-2 cells as well. However, the tag could not be detected by anti-His western blots. The absence of the tag was verified by purification on a Ni^{2+} resin, in which the RBD was found in the flow through, indicating that it failed to bind to the resin, contrary to what would be expected if the tag was intact. These results with the C-terminal polyhistidine tag, are in complete accordance with results obtained independently by Rebelo *et al.* (2022), who also expressed the RBD in BY-2 cells (see section 1.3.4 of the Introduction). These results suggest that the polyhistidine tag might be damaged, probably due to host proteases. This is unfavourable because the purification of the proteins becomes more complicated. Therefore, efforts are still underway in our laboratory to add a tag to the RBD that will remain intact and ease its purification.

In this context, the objective of this work is to evaluate the impact of an N-terminal tag on the production of the SARS-CoV-2 S protein RBD in *N. tabacum* BY-2 cells. We hypothesize that the presence of an N-terminal tag would favour the production of the intact tagged RBD. This strategy was not opted initially because the RBD contains two *N*-glycosylation sites near its N-terminus and a purification tag in this position could potentially interfere with the impact of these *N*-glycans on the binding measurements between the RBD and the hACE2 receptor. For this reason, a "*10xHis-NTD-RBD*" construct was also included, in the aim of having the tag further from the *N*-glycosylation sites of the RBD. However, whether, and to which extent the tag might affect this interaction remains to be answered. For

this master thesis, the focus is confined to the production of the RBD fused to an intact/detectable purification tag on the N-terminus of the protein.

The strategy followed for testing the initial hypothesis was similar to the work carried out previously with the C-terminal tag. Firstly, genetic constructs encoding either "10xHis-RBD" or "V5-RBD" were screened by transient expression in *N. benthamiana* leaves. Additionally, the "*10xHis-NTD-RBD*" construct was also tested. Meanwhile, a less common approach of screening the constructs via transient expression in BY-2 was also tested for the RBD constructs with a 10xHis or V5 tag at either end of the protein. The construct producing the protein of interest in *N. benthamiana* was then used for stable transformation of BY-2 cells. Finally, the stably transformed BY-2 cell lines were screened by western blot to check whether the RBD was produced and, if so, the best producing lines were selected for further analyses, such as verification of the presence of the tag.

Chapter II: Results

2.1 Amplification and cloning of the *RBD* and *NTD-RBD* sequences

The sequences encoding RBD (R319-L533) or NTD-RBD (Q14-L533) of the Wuhan-hu-1 strain of SARS-CoV-2 spike protein were amplified by PCR using the template vector pUC-GW-Amp-gS, already available in our laboratory. An illustration of the pUC-GW-Amp-gS plasmid is shown in Figure 5.

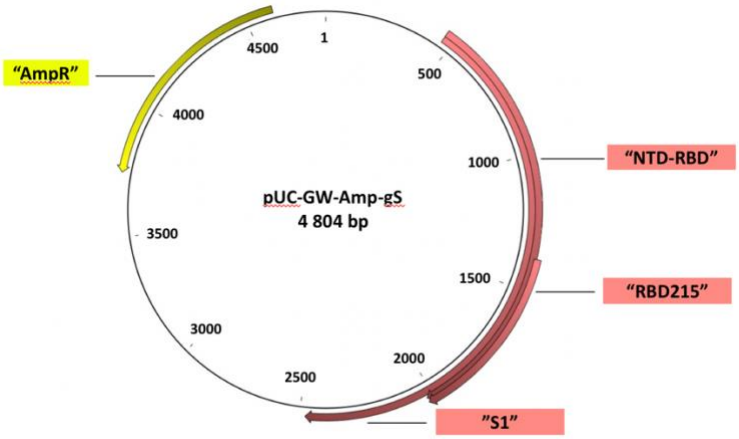

Figure 5: Illustration of the template pUC-GW-Amp-gS vector containing the sequence encoding the S1-subunit of the Wuhan-hu-1 strain of the SARS-CoV-2 spike protein, used for the amplification of the RBD and NTD-RBD sequences. "AmpR" = Ampiciline resistance gene; "S1" = S1-subunit of the SARS-CoV-2 (Wuhan-hu-1 straain); "NTD-RBD" = N-Terminal Domain (NTD) + Receptor Binding Domain (RBD) (Q14-L533); "RBD215" = truncated version of the RBD with 215 residues (R319-L533).

The template contained the sequence encoding the S1-subunit and three different primers were designed to amplify the sequence encoding both proteins (the primer sequences are given in Table 4, section 5.3.1 in Materials and Methods). The PCR products were analysed by DNA gel electrophoresis, to ensure that the amplification had been done correctly. A band at 678 bp was expected for the RBD and at 1,593 bp for the NTD-RBD. These bands were indeed observed on the DNA gel electrophoresis, indicating that the amplification was done correctly (Figure 6).

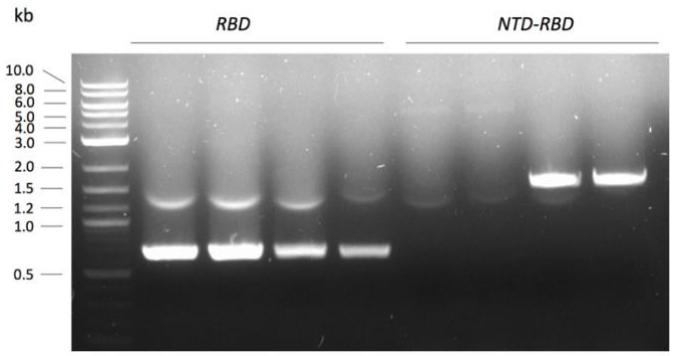

Figure 6: Agarose DNA gel electrophoresis of the PCR amplification of the RBD *and* NTD-RBD *sequences from the pUC-GW-Amp-gS template vector. Bands were expected at 678 bp and 1,593 bp for the RBD and the NTD-RBD, respectively. Ten µL of 2 log ladder and 12 µL of the PCR samples were loaded in the respective wells.*

The corresponding bands were then cut from the gel and purified before cloning them in the pGEM® -T Easy vector.

2.2 Cloning of the *RBD* and *NTD-RBD* sequences with pGEM®-T Easy

The purified PCR products were ligated into pGEM®-T Easy vectors, which contain an ampicillin resistance gene. Electrocompetent Top 10 *E. coli* cells were then transformed with the resulting plasmids. The plasmids were isolated, digested by EcoRI and analysed by DNA gel electrophoresis to verify if the sequences of the RBD or the NTD-RBD had been correctly ligated into the pGEM®-T Easy vector (Figure 7). Two bands were expected for the RBD, at 696 bp and 3,001 bp, as well for the NTD-RBD, at 1,611 bp and 3,001 bp.

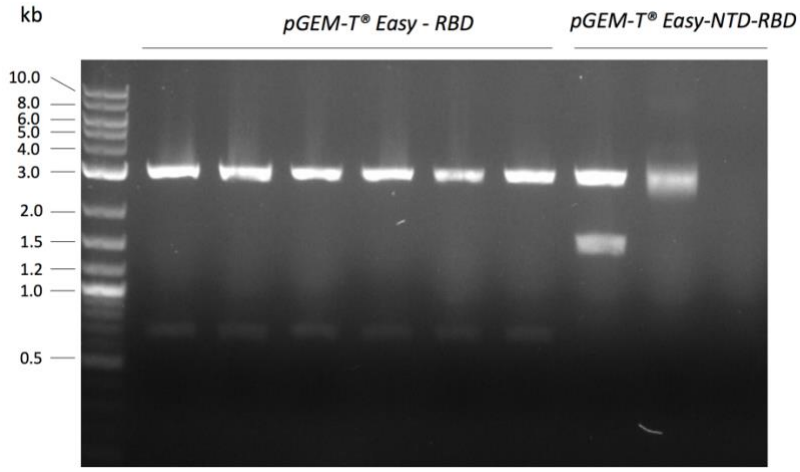

Figure 7: Agarose DNA gel electrophoresis of the control restriction of the pGEM®-T Easy-RBD and pGEM®-T Easy-NTD-RBD plasmids. The digestion was done by EcoRI and the bands were expected at 696 bp and 3,001 bp for pGEM®-T Easy-RBD, and at 1611 bp and 3,001 bp for pGEM®-T Easy-NTD-RBD. Ten µL of 2 log ladder and 12 µL of the digested samples were loaded in the respective wells.

Bands were observed at the expected size, indicating that the ligation had been done correctly in at least one sample of each ligation. Samples showing the expected restriction profile were sent for sequencing. The sequence of pGEM®-T Easy-RBD was correct while the sequence of pGEM®-T Easy-NTD-RBD was not. The ligation of the NTD-RBD PCR product into the pGEM® T-Easy vector was therefore repeated and was found to be successful (data not shown).

2.3 Obtaining of level 1 Golden Gate plasmids

The *RBD* and *NTD-RBD* sequences were cloned into pICH47742 Golden Gate acceptor vector to obtain level 1 plasmids. pICH47742 plasmid contains the *lacZ* selectable marker and an ampicillin resistance gene. The expression cassettes, inserted into the plasmid, consisted of the sequences encoding the 35S promoter (*p35S*) of the Cauliflower Mosaic Virus (CaMV), the signal peptide of the Protein Disulfide Isomerase (*spPDI*) of *Medicago Sativa*, the sequence of the purification tag (*10xHis or V5*), the sequence of the protein of interest and the terminator of the Plasma Membrane ATPase Proteolipid 1 (*tPMP1*) of *S. cerevisiae* (Figure 8). The sequences of the tags are shown in Table 1. In these constructs a double V5 tag was used.

Table 1: Amino acid sequence of the tags used in the different genetic constructs.

Tag	Sequence
10xHis	HHHHHHHHHH
V5	GKPIPNPLLGLDSTGGKPIPNPLLGLDST

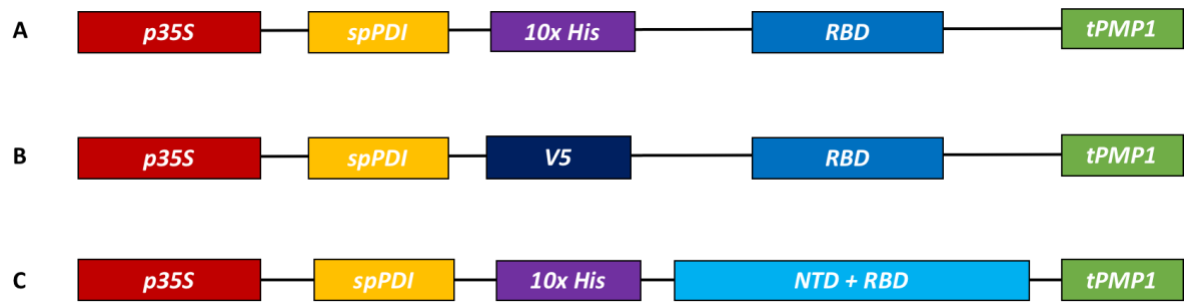

Figure 8: Schematic illustration of the expression cassettes encoding (A) the 10xHis-RBD, (B) the V5-RBD and (C) the 10xHis-NTD-RBD. p35S = *35S promoter of the Cauliflower Mosaic Virus;* spPDI = *signal peptide of the Protein Disulfide Isomerase of* Medicago Sativa; tPMP1 = *terminator of the Plasma Membrane ATPase Proteolipid 1 of* S. cerevisiae.

E. coli Top 10 cells were transformed with the level 1 plasmids. After purification, the plasmids were digested by EcoRI and analysed by DNA gel electrophoresis, in order to verify that the constructs had been ligated correctly. Plasmids with the *10xHis-RBD* expression cassette were isolated from five different transformed *E. coli* colonies. The results of the electrophoresis are shown in Figure 9.

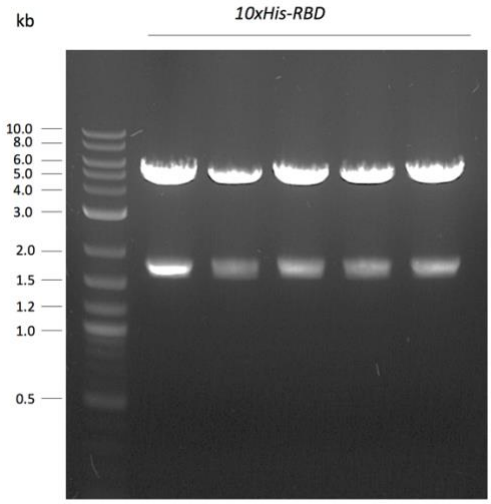

Figure 9: Agarose DNA gel electrophoresis of the control restriction of the level 1 Golden Gate plasmid containing the 10xHis-RBD *expression cassette, purified from* E. coli. *The digestion was done by EcoRI and the bands were expected at 1,720 bp and 4,857 bp. Ten µL of 2 log ladder and 12 µL of the digested samples were loaded in the respective wells.*

Two bands with the expected size of 1,720 bp and 4,857 bp were observed in the gel for all of the colonies, indicating that the ligation had been done correctly. The correct sequence of the plasmid was confirmed by sequencing.

For the *V5-RBD* expression cassette, only two transformed *E. coli* colonies were obtained. The plasmids were isolated and analysed by DNA gel electrophoresis (Figure 10).

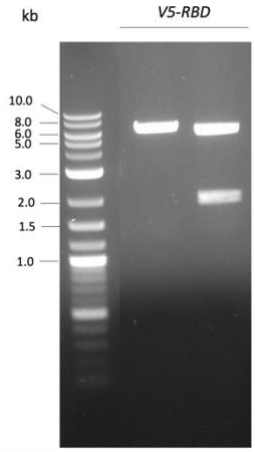

Figure 10: Agarose DNA gel electrophoresis of the control restriction of the level 1 Golden Gate plasmid containing the V5-RBD *expression cassette. The restriction was done by EcoRI and the bands were expected at 1,777 bp and 4,857 bp. Ten µL of 2 log ladder and 12 µL of the digested samples were loaded in the respective wells.*

The presence of bands at around 1,777 bp and 4,857 bp, which were the expected sizes, indicated that the ligation had been achieved correctly for the second colony. This was confirmed by sequencing.

Six *E. coli* colonies transformed with the level 1 *10xHis-NTD-RBD* plasmid were analysed. The plasmids were digested by EcoRI and analysed by DNA gel electrophoresis (Figure 11). Two bands with the expected size at 2,635 bp and 4,857 bp were observed in two of the six colonies and confirmed by sequencing.

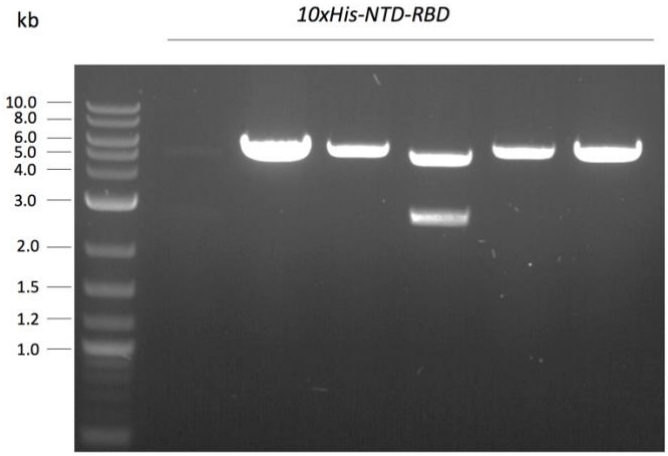

Figure 11: Agarose DNA gel electrophoresis of the control restriction of the level 1 Golden Gate plasmids containing the 10xHis-NTD-RBD *expression cassette, purified from* E. coli*. The restriction was done by EcoRI and the bands were expected at 2,635 bp and 4,857 bp. Ten µL of 2 log ladder and 12 µL of the digested samples were loaded in the respective wells.*

2.4 Obtaining of level M Golden Gate plasmids

Level M Golden Gate plasmids were obtained from level 1 plasmids. These level M plasmids were used for the transformation of the different plant expression systems and were, therefore, constructed to include the sequences of reporter genes.

The gene encoding the mCherry reporter protein was included in the constructs to confirm whether the transient transformations of the *N. benthamiana* leaves or BY-2 cells had been successful. These constructs included an expression cassette for the *mCherry* gene, which was under the control of the *p35S* promoter and the *tNos* terminator, and the expression cassette with the gene of interest. The acceptor plasmid was pAGM8031. A schematic illustration of the level M constructs is shown in Figure 12.

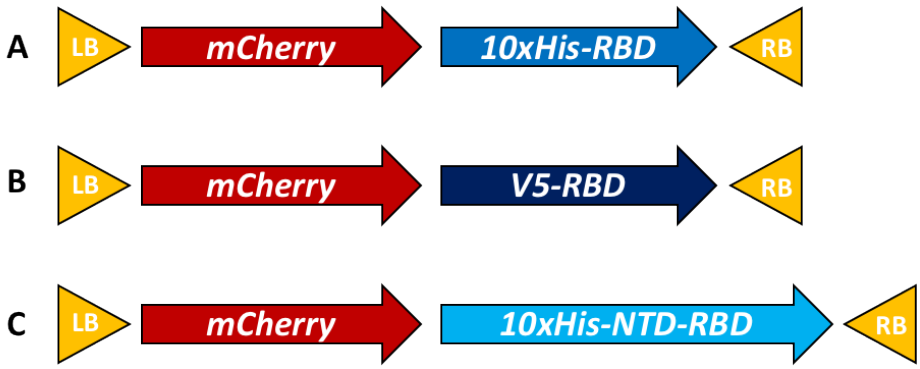

Figure 12: Schematic representation of the level M Golden Gate constructs with the expression cassettes of the mCherry *gene and the gene of interest. The* mCherry *gene is under the control of the* p35S *promoter and the* tNos *terminator. The expression cassettes are inserted into the pAGM8031 acceptor plasmid. LB = Left border; RB = Right border.*

In other level M constructs, the *nptII* gene, encoding the neomycin phosphotransferase II, which confers resistance to kanamycin and is used for the selection of the transformed BY-2 cells, was added in place of the *mCherry* sequence. These constructs consisted of the *nptII* expression cassette, the expression cassette of the gene of interest and an Rb7 Scaffold Attachment Region (SAR). The SAR has been shown experimentally in our laboratory to enable a better production of proteins of interest in BY-2 cells. The level M constructs with the *nptII* expression cassette are illustrated in Figure 13.

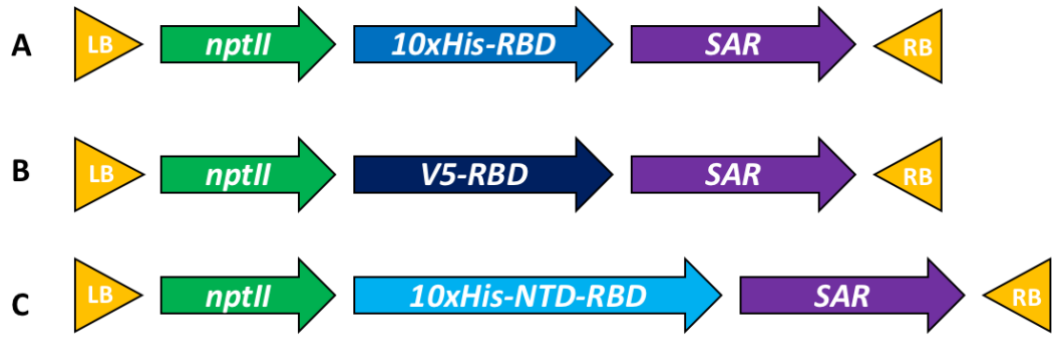

Figure 13: Schematic representation of the level M Golden Gate constructs used for the stable transformation of BY-2 cells. They contain the expression cassettes of the nptII *marker gene, the gene of interest and an Rb7* Scaffold Attachment Region *(SAR). The expression cassettes are inserted into the pAGM8031 acceptor plasmid. LB = Left border; RB = Right border.*

E. coli Top 10 cells were transformed with the level M plasmids. The purified plasmids from *E. coli* were then digested by different restriction enzymes and the restriction product was analysed by DNA gel electrophoresis. The plasmids with the *10xHis-RBD* expression cassette were digested by HindIII. Three bands were expected, namely at 449 bp and 1,157 bp for both plasmids, whereas a band at 7,100 bp was expected for the *mCherry-10xHis-RBD* plasmid and a band at 8,332 bp was expected for the *nptII-10xHis-RBD* plasmid. The DNA gel electrophoresis is shown in Figure 14.

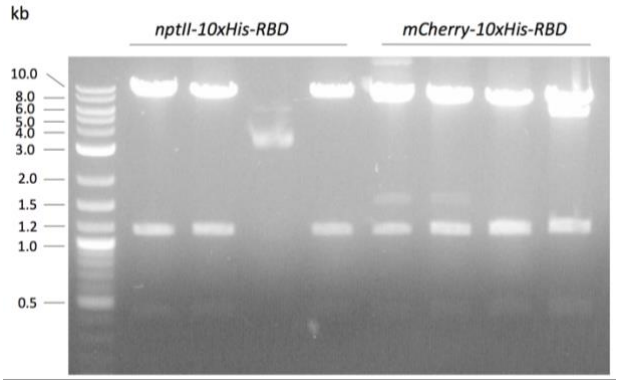

Figure 14: Agarose DNA gel electrophoresis of the control restriction of the level M Golden Gate plasmids with the 10xHis-RBD *expression cassette, purified from* E. coli. *The restriction was done by HindII and the expected bands were: 449bp, 1,157 bp and 7,100 bp for the* mCherry-10xHis-RBD *constructs and 449 bp, 1,157 bp and 8,332 bp for the* nptII-10xHis-RBD *construct. Ten µL of 2 log ladder and 12 µL of the digested samples were loaded in the respective wells.*

Samples for each plasmid were selected for the subsequent transformation of *A. tumefaciens*. The level M plasmids with the *V5-RBD* and *10xHis-NTD-RBD* expression cassettes were digested by SacI. Bands at 1,846 bp and 6,913 bp were expected for the *mCherry-V5-RBD* construct and at 1,505 bp and 8,486 bp for the *nptII-V5-RBD* construct. Bands at 1,846 bp and 7,435 bp were expected for the *mCherry-10xHis-NTD-RBD* construct and at 2,027 bp and 8,486 bp for the *nptII-10xHis-NTD-RBD* construct. The respective DNA gel electrophoreses are shown in Figures 15 and 16.

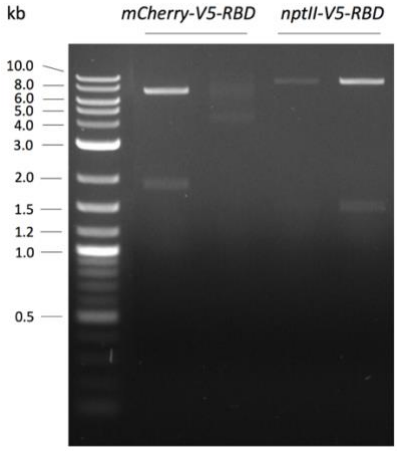

Figure 15: Agarose DNA gel electrophoresis of the control restriction of the level M Golden Gate plasmids with the V5-RBD *expression cassette, purified from* E. coli. *The restriction was done by SacI and the expected bands were: 1,846 bp and 6,913 bp for the* mCherry-V5-RBD *construct and 1,505 bp and 8,486 for the* nptII-V5-RBD *construct. Ten µL of 2 log ladder and 12 µL of the digested samples were loaded in the respective wells.*

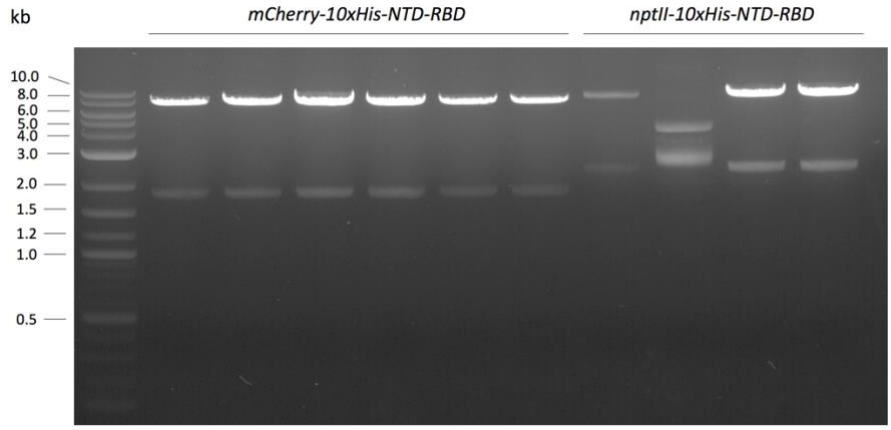

Figure 16: Agarose DNA gel electrophoresis of the control restriction of the level M Golden Gate plasmids with the 10xHis-NTD-RBD *expression cassette, purified from* E. coli*. The restriction was done by SacI and the expected bands were: 1,846 bp and 7,435 bp for the* mCherry-10xHis-NTD-RBD *construct and 2,027 bp and 8,486 bp for the* nptII-10xHis-NTD-RBD *constructs. Ten µL of 2 log ladder and 12 µL of the digested samples were loaded in the respective wells.*

Samples with the expected restriction profile were observed and were therefore selected for subsequent *A. tumefaciens* transformation.
The presence of bands at the expected size, in most samples, indicated that the level M plasmids had been ligated correctly. Therefore, samples with the correct restriction profile were selected for the subsequent transformation of *A. tumefaciens*. The results of this section are summarized in Table 2.

Table 2: Summary of the level M Golden Gate constructs, restriction enzymes used for the control restrictions and the expected bands in DNA gel electrophoresis of digested samples purified from E. coli*.*

Level M Golden Gate construct	Restriction enzyme	Expected bands (bp)
mCherry-10xHis-RBD	HindIII	449, 1,157, 7100
nptII-10xHis-RBD	HindIII	449, 1157, 8332
mCherry-10xHis-RBD	SacI	1846, 6856
nptII-10xHis-RBD	SacI	1448, 8486
mCherry-V5-RBD	SacI	1846, 6913
nptII-V5-RBD	SacI	1505, 8486
mCherry-10xHis-NTD-RBD	SacI	1846, 7435
nptII-10xHis-NTD-RBD	SacI	2027, 8486

2.5 Obtaining of transformed *A. tumefaciens* cell lines

LBA4404 VirG *A. tumefaciens* cells were transformed by electroporation with the different level M Golden Gate plasmids. The plasmids were then purified, digested by SacI, and the restriction products were analysed by DNA gel electrophoresis. The sizes of the expected bands for the *mCherry-10xHis-RBD* and *nptII-10xHis-RBD* constructs, as well as for the *mCherry-V5-RBD, mCherry-10xHis-NTD-RBD, nptII-V5-RBD* and *nptII-10xHis-NTD-RBD* constructs, digested by SacI are shown in Table 2. The respective DNA gel electrophoreses are shown in Figures 17, 18 and 19.

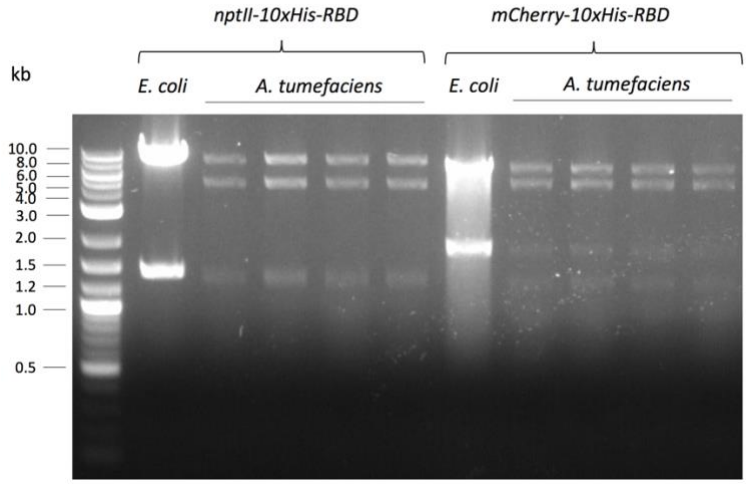

Figure 17: Agarose DNA gel electrophoresis of the control restriction of the level M Golden Gate plasmids with the 10xHis-RBD *expression cassette purified from* A. tumefaciens*. The level M plasmids used for the transformation of* A. tumefaciens *cells, purified from* E. coli *are also presented. The restriction was done by SacI and the expected bands were: 1,448 bp and 8,486 bp for the* nptII-10xHis-RBD *construct and 1,846 bp and 6,856 for the* mCherry-10xHis-RBD *construct. Ten µL of 2 log ladder and 12 µL of the digested samples were loaded in the respective wells.*

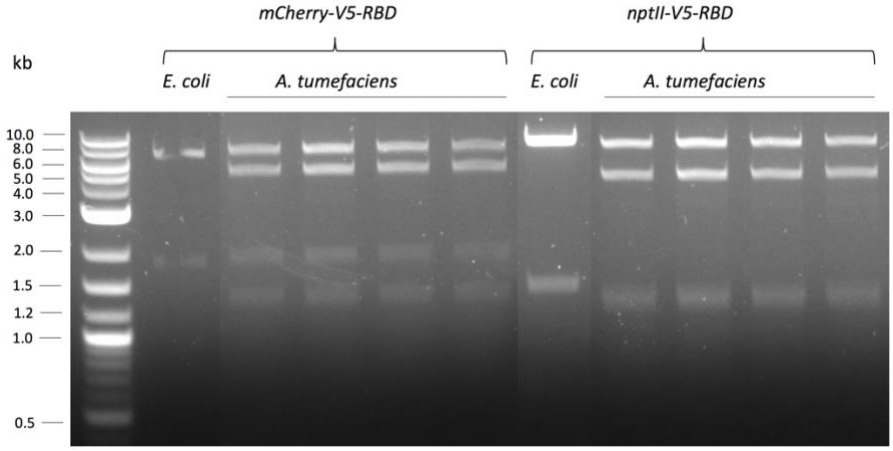

Figure 18: Agarose DNA gel electrophoresis of the control restriction of the level M Golden Gate plasmids with the V5-RBD *expression cassette purified from* A. tumefaciens*. The level M plasmids used for the transformation of* A. tumefaciens *cells, purified from* E. coli *are also presented. The restriction was done by SacI and the expected bands were: 1,846 bp and 6,913 bp for the* mCherry-V5-RBD *construct and 1,505 bp and 8,486 the* nptII-V5-RBD *construct. Ten µL of 2 log ladder and 12 µL of the digested samples were loaded in the respective wells.*

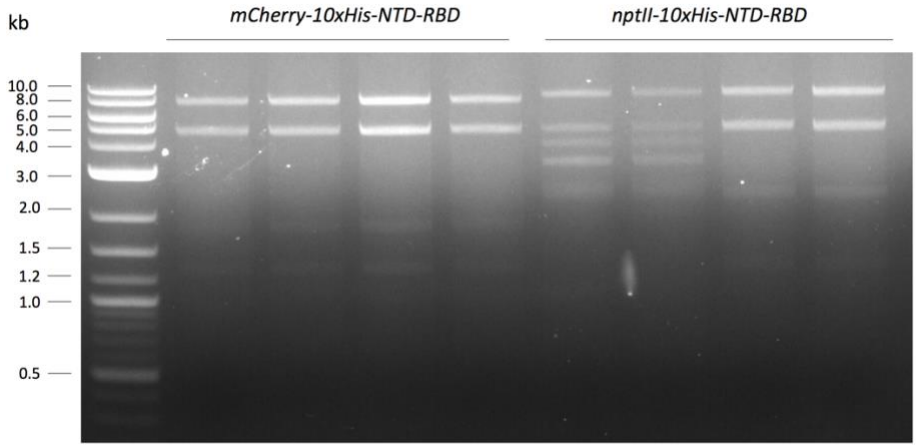

Figure 19: Agarose DNA gel electrophoresis of the control restriction of the level M Golden Gate plasmids with the 10xHis-NTD-RBD *expression cassette purified from* A. tumefaciens*. The restriction was done by SacI and the expected bands were: 1,846 bp and 7,435 bp for the mCherry-10xHis-NTD-RBD construct and 2,027 bp and 8,486 for the nptII-10xHis-NTD-RBD construct.*

Bands at the expected sizes were observed in all samples, with the exception of the first and second samples of the *nptII-10xHis-NTD-RBD* construct, where two additional bands were observed. The bands can probably be due to contamination with nearby colonies at the moment of picking the transformed *A. tumefaciens* colonies to start liquid cultures. A summary of the results of this section and an overview of the different plant transformations are given in Table 3.

Table 3: Overview of the level M Golden Gate constructs used to transform A. tumefaciens*, restriction enzymes, the expected bands in DNA gel electrophoresis of digested samples and the plant transformations performed in this work ("✓" : the transformation was performed ; N/A = the constructs are not typically used for the given transformation; "\" = the transformation with the given construct was not performed in this work).*

Level M Golden Gate construct	Restriction enzyme	Expected bands (bp)	Plant system transformations		
			N. benthamiana	Transient BY-2	Stable BY-2
mCherry-10xHis-RBD	SacI	1846, 6856	✓	✓	N/A
nptII-10xHis-RBD	SacI	1448, 8486	N/A	N/A	✓
mCherry-V5-RBD	SacI	1846, 6913	✓	✓	N/A
nptII-V5-RBD	SacI	1505, 8486	N/A	N/A	\
mCherry-10xHis-NTD-RBD	SacI	1846, 7435	✓	\	N/A
nptII-10xHis-NTD-RBD	SacI	2027, 8486	N/A	N/A	\

2.6 Transient expression in *N. benthamiana* leaves

2.6.1 Protein production

2.6.1.1 The RBD constructs

Transient expressions of the *mCherry-10xHis-RBD*, *mCherry-V5-RBD* and *mCherry-10xHis-NTD-RBD* from the level M plasmids were performed by agroinfiltration of leaves of 4-week-old *N. benthamiana* plants. The infiltrations were done in triplicates for each construct and non-infiltrated leaves were used as negative control. Total Soluble Leaf Proteins (TSLP) were extracted from leaf discs 5 days after infiltration. TSLP were then analysed by western blot using two different primary antibodies: an anti-mCherry antibody to verify whether the transformation had been successful, and an anti-RBD antibody to check for production of RBD or NTD-RBD. TSLP from leaves transformed with the *mCherry-RBD-10xHis* level M plasmid, *i.e.* a plasmid encoding the RBD with a C-terminal 10xHis tag, which was available in our laboratory, was used as positive control.

The results of the western blots performed with the anti-mCherry primary antibody are shown in Figure 20.

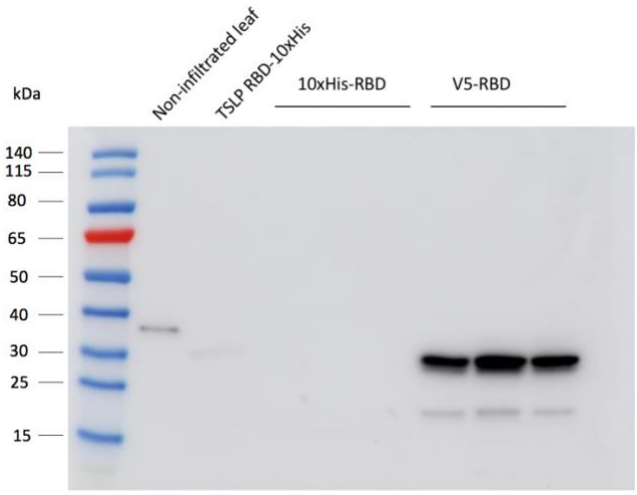

Figure 20: Anti-mCherry western blot of the TSLP from N. benthamiana *leaves transformed with the* mCherry-10xHis-RBD *or* mCherry-V5-RBD *constructs. Each well was loaded with 26 µg of TSLP. Rabbit anti-mCherry and anti-rabbit antibodies conjugated to the horseradish peroxidase (HRP) were used as primary and secondary antibodies, respectively. Revelation was done using a chemiluminescence substrate. TSLP from non-infiltrated* N. benthamiana *leaves and TSLP from* N. benthamiana *leaves transformed with the* mCherry-RBD-10xHis *construct were used as negative and positive controls, respectively.*

The presence of the mCherry signals at the expected molecular mass (30 kDa) as well as the fainter bands at a lower molecular mass, which is a typical western blot pattern of this protein, indicated that the leaf transformation had been successful for the *mCherry-V5-RBD* construct. A nonspecific, unexplained band, with a higher molecular mass than the expected one was observed in the negative control, and a very faint band at the expected size, and thereby corresponding to mCherry, was observed in the positive control. In the case of the *mCherry-10xHis-RBD* construct, the absence of mCherry bands was attributed to the fact that the leaves were not transformed with the appropriate line of *A. tumefaciens*. Instead, one of the *A. tumefaciens* lines containing the *nptII-10xHis-RBD* plasmid was used for the transformation,

due to a mislabelling of the tubes. However, detection of RBD signals using the anti-RBD antibodies strongly suggested that the transformation had indeed been successful (Figure 21).

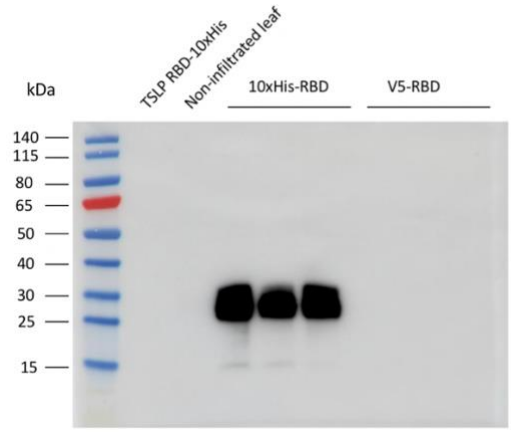

Figure 21: Anti-RBD western blots of the TSLP from N. benthamiana *leaves transformed with the* mCherry-10xHis-RBD *or* mCherry-V5-RBD *constructs. Each well was loaded with 26 µg of TSLP. Rabbit anti-RBD and anti-rabbit antibodies conjugated to the horseradish peroxidase (HRP) were used as primary and secondary antibodies, respectively. The revelation was done using a chemiluminescence substrate. TSLP from* N. benthamiana *leaves transformed with the* mCherry-RBD-10xHis *construct and TSLP from non-infiltrated* N. benthamiana *leaves were used as positive and negative controls, respectively.*

The theoretical molecular mass of the RBD is 28 kDa. This western blot indicated that there was a production of RBD with the *10xHis-RBD* expression cassette, but not with the *V5-RBD*. Some degradation bands were also observed at lower molecular weights. These results were in accordance with previous results of transient expression in *N. benthamiana* leaves of RBD constructs with a 10xHis or a V5 tag on the C-terminus, obtained in our laboratory.

2.6.1.2 The NTD-RBD construct

Transient expression of the *mCherry-10xHis-NTD-RBD* construct was also performed in *N. benthamiana* leaves. The western blot with an anti-mCherry primary antibodies is shown in Figure 22.

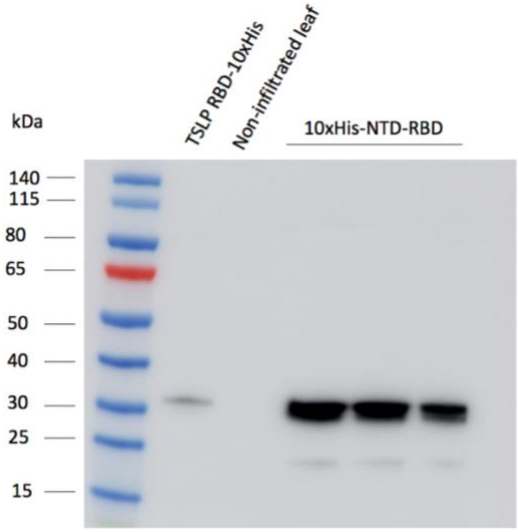

Figure 22: Anti-mCherry western blot of the TSLP from N. benthamiana *leaves transformed with the* mCherry-10xHis-NTD-RBD *construct. Each well was loaded with 26 µg of TSLP. Rabbit anti-mCherry and anti-rabbit antibodies conjugated to the horseradish peroxidase (HRP) were used as primary and secondary antibodies, respectively. The revelation was done using a chemiluminescence substrate. TSLP from* N. benthamiana *leaves transformed with the* mCherry-RBD-10xHis *construct and TSLP from non-infiltrated* N. benthamiana *leaves were used as positive and negative controls, respectively.*

The presence of intense bands at around 30 kDa and a fainter band at lower molecular mass, which is a typical pattern of mCherry, indicated that the transformation had been successful. Results of an anti-RBD western blot are shown in Figure 23.

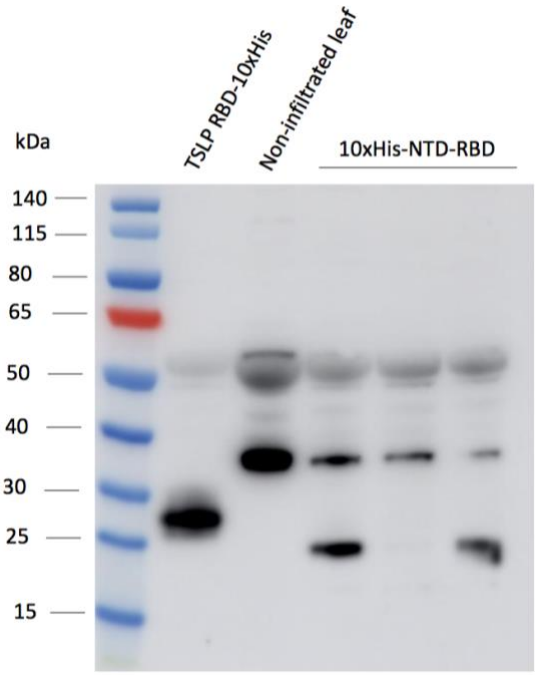

Figure 23: Anti-RBD western Blot of the TSLP from N. benthamiana *leaves transformed with the* mCherry-10xHis-NTD-RBD *construct. Each well was loaded with 26 µg of TSLP. Rabbit anti-RBD and anti-rabbit antibodies conjugated to the horseradish peroxidase (HRP) were used as primary and secondary antibodies, respectively. The revelation was done using a chemiluminescence substrate. TSLP from* N. benthamiana *leaves transformed with the* mCherry-RBD-10xHis *construct and TSLP from non-infiltrated* N. benthamiana *leaves were used as positive and negative controls, respectively.*

The theoretical molecular mass of the NTD-RBD, without glycosylation, is 68 kDa. No band with this mass was observed, indicating the absence of intact NTD-RBD. However, one additional band at about 25 kDa was observed in the samples of transformed leaves, while not observed in non-infiltrated leaves. This band could be a degradation product, suggesting that NTD-RBD was produced but degraded by plant proteases.

2.6.2 Tag intactness

The TSLP from the leaves that had produced RBD with the *mCherry-10xHis-RBD* construct were further analysed to check whether the 10xHis tag was still present. This was done by a western blot with anti-His primary antibodies, followed by a purification assay using a Ni^{2+}-NTA resin. The positive control was the SARS-CoV-2 S1-subunit with a 10xHis tag on its C-terminus, which had been produced previously in our laboratory and was confirmed to have a detectable tag. The anti-His western blot is shown in Figure 24.

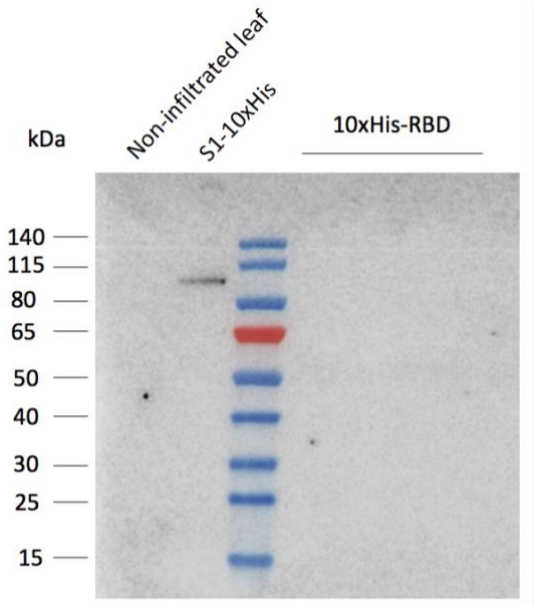

Figure 24: Anti-His western blot of the TSLP from N. benthamiana *leaves transformed with the* mCherry-10xHis-RBD *construct. Each well was loaded with 26 µg of TSLP. Anti-rabbit antibodies conjugated to the horseradish peroxidase (HRP) were used as secondary antibodies and the revelation was done using a chemiluminescence substrate. TSLP from* N. benthamiana *leaves transformed with the SARS-CoV-2* S1-10xHis *expression cassette were used as positive control and TSLP from non-infiltrated* N. benthamiana *leaves were used as negative control.*

A signal at 28 kDa (the theoretical size of the RBD) was expected for this western blot, but it was not observed, indicating that the tag was probably (partially) degraded. The purification assay was performed to confirm that the tag was indeed undetectable and to avoid a false negative result, as anti-His antibodies are known to be not very sensitive. The TSLP sample was mixed with the equilibrated Ni^{2+}-NTA resin. The washing and elution were done using increasing concentrations of imidazole (from 20 to 300 mM). The different fractions were then analysed by western blot using an anti-RBD primary antibody (Figure 25).

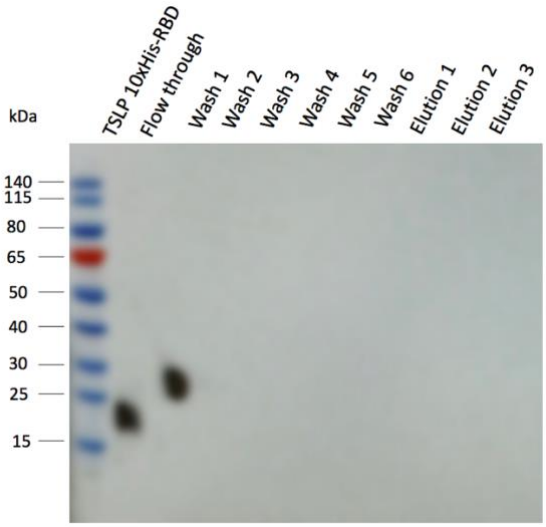

Figure 25: Anti-RBD western blot of the Ni^{2+}-NTA purification fractions of the TSLP from N. benthamiana *leaves transformed with the* mCherry-10xHis-RBD *construct. Each well was loaded with 25 µg of TSLP. Rabbit anti-RBD and anti-rabbit antibodies conjugated to the horseradish peroxidase (HRP) were used as primary and secondary antibodies, respectively. The revelation was done using a chemiluminescence substrate. TSLP from* N. benthamiana *leaves transformed with the* mCherry-10xHis-RBD *construct (the same sample used for the purification) were used as positive control.*

The presence of a band corresponding to the mass of RBD in the elution fractions would indicate that the 10xHis tag was still fused to the protein. However, a band at the expected size could only be found in the flow through fraction, confirming that the tag was probably degraded. The lower molecular mass of the band observed in the positive control could probably be due to degradation by proteases, since the sample had been used for several manipulations. This degradation could probably have occurred during the western blot manipulations, after the purification fractions had been obtained. Otherwise, there would be no reason for such a difference in size between the RBD bands observed in the positive control and in the flow through fraction, given that the RBD in both cases was derived from the same TSLP sample.

The purification results and those of the anti-His western blot confirmed that the tag might indeed be lost during protein production. These results are in accordance with previous results from transient expression of RBD with a C-terminal 10xHis tag in *N. benthamiana* leaves, performed in our laboratory.

2.7 Transient expression in BY-2 cells

The *mCherry-10xHis-RBD* and *mCherry-V5-RBD* constructs were then transiently expressed in BY-2 cells. Transient expression in BY-2 cells may be of interest as it can provide beforehand a more accurate idea than transient expression in *N. benthamiana* of what can be expected from stable expressions, as the latter usually takes up to 3 months.

In this experiment, the BY-2 cells were co-cultivated with *A. tumefaciens* cell lines containing the appropriate plasmids, and the expression of proteins was achieved in 3 days. The experiment also included RBD constructs with the 10xHis or V5 tag on the C-terminus, with the aim of comparing the expression of RBD with the 10xHis tag on either end of the protein. *A. tumefaciens* cell lines containing plasmids with the *mCherry-RBD-10xHis* and the *mCherry-RBD-V5* constructs had been obtained previously and were supplied by Marie Peeters. After 3 days, the Total Soluble Proteins (TSP) were extracted from the transformed BY-2 cells and analysed by western blot, using two different primary antibodies: anti-mCherry and anti-RBD antibodies (Figures 26 and 27, respectively).

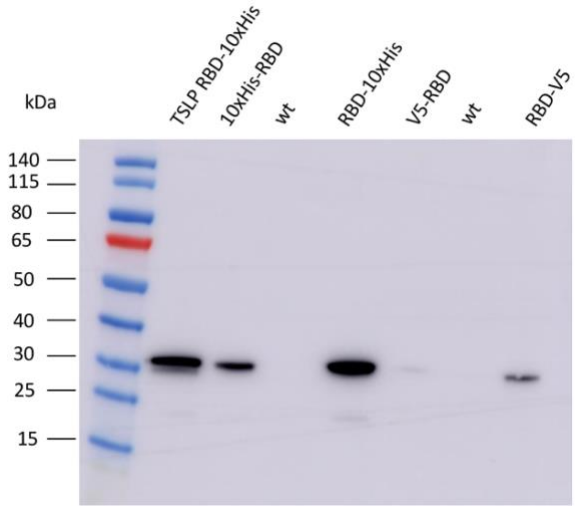

Figure 26: Anti-mCherry western blot of the TSP from BY-2 cells transiently transformed with different RBD constructs. Each well was loaded with 26 µg of TSLP or TSP. Rabbit anti-mCherry and anti-rabbit antibodies conjugated to the horseradish peroxidase (HRP) were used as primary and secondary antibodies, respectively. The revelation was done using a chemiluminescence substrate. TSLP from N. benthamiana *leaves transformed with the* mCherry-RBD-10xHis *construct were used as positive control and TSP from the same BY-2 cultures transiently transformed, but which had not been co-cultivated with* A. tumefaciens *were used as negative control.*

The band at 30 kDa which was present in the samples of transformed cells but not in the negative control had the expected mass of mCherry, indicating that the cells had been successfully transformed. As a positive control TSLP from *N. benthamiana* leaves transformed with the *mCherry-RBD-10xHis* construct were used since constructs with mCherry had never been transformed into BY-2 cells.

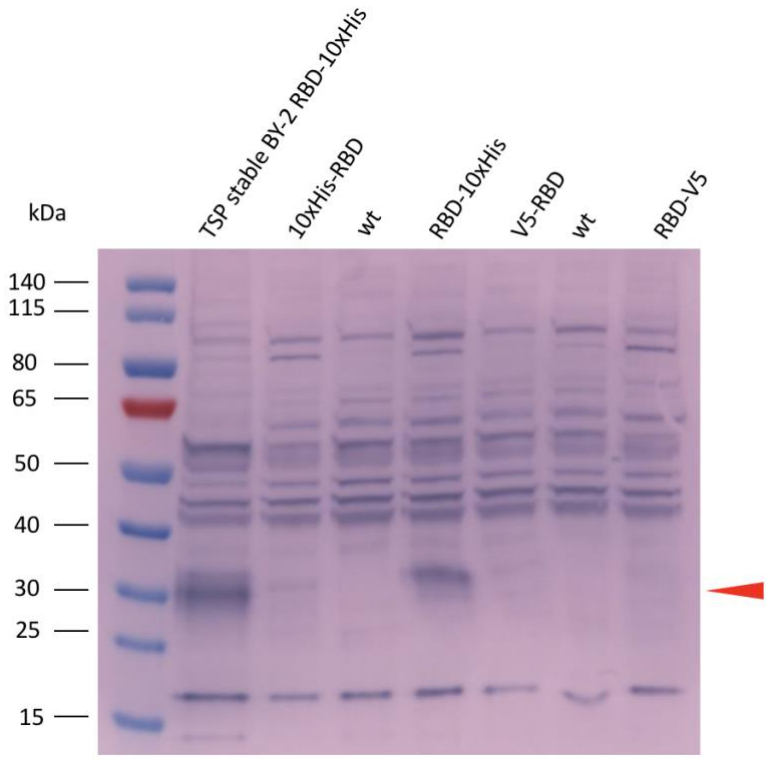

Figure 27: Anti-RBD western blot of the TSP from BY-2 cells transiently transformed with different RBD constructs. Each well was loaded with 26 µg of TSP. Rabbit anti-RBD and anti-rabbit antibodies conjugated to an alkaline phosphatase were used as primary and secondary antibodies, respectively. The revelation was done using a precipitating substrate. TSP from BY-2 cells stably transformed with the nptII-RBD-10xHis *construct was used as positive control and TSP from the same BY-2 cultures transiently transformed, but which had not been co-cultivated with* A. tumefaciens *were used as negative control.*

On the western blot revealed by the RBD antibodies (Figure 27), faint bands were observed around 30 kDa in the TSP samples of transformed cells, corresponding to the band observed in the positive control and thus, to the RBD. The intensity of the RBD band in the RBD-10xHis sample and the positive control was quite similar, while its intensity was fainter in the other samples, but the band was still distinguishable. Two negative controls were used because the transformations with each tag were performed at two different moments.

Revelation with the alkaline phosphatase precipitating substrate is more sensitive than that with HRP chemiluminescence and faint bands indicated that the protein concentrations were very low. However, it is noteworthy that the bands in the anti-mCherry western blot were not very intense either, at least compared to the positive control, except for the RBD-10xHis sample. Thus, the low concentration of RBD in the TSP could be partially attributed to the transformation efficiency, rather than entirely to the level of protein synthesis.

Finally, it is worth mentioning that the results of transient expressions in BY-2 cells and *N. benthamiana* leaves showed differences between the two systems, as observed in the expression of *mCherry-RBD-V5* and *mCherry-V5-RBD* constructs.

2.8 Stable expression in BY-2 cells

2.8.1 Protein production

The stable production of the RBD with an N-terminal tag in BY-2 cells was the main objective of this work. The construct encoding 10xHis-RBD seemed to be the most promising as compared to the V5-RBD one. Thereby, only this construct was tested in stable expression. Due to time constraints, the construct encoding 10xHis-NTD-RBD was not tested either.

BY-2 cells were co-cultivated with an *A. tumefaciens* cell line transformed the *nptII-10xHis-RBD* plasmid. The co-cultivation cultures were then spread on solid MSCCK and incubated until calli appeared. After selection, calli were maintained on MSK plates. Calli before the second passage on solid MSCCK are shown in Figure 28.

Figure 28: BY-2 calli transformed with the nptII-10xHis-RBD construct, three weeks after the first passage onto solid MSCCK (© 2024 Joel Skënderaj).

After selection, 18 liquid cultures were started from independent calli. The liquid cultures were diluted twice before collecting the cells and the extracellular medium (EM), and extracting the TSP.

TSP and EM from 11 liquid cultures were analysed by western blot using anti-RBD primary antibodies, while the other seven liquid cultures did not grow as expected. Since the constructs contained a sequence encoding a signal peptide that targets the protein in the secretory pathway, the majority of the RBD was expected to be found in the EM. TSP were revealed using the alkaline phosphatase precipitating substrate since it is more sensitive, and previous experiments in our laboratory had shown a very efficient secretion of the RBD into the EM, and thereby less RBD in the TSP fractions. EM fractions were revealed by chemiluminescence with the HRP. The western blot results from the TSP and EM samples are shown in Figure 29 and Figure 30, respectively.

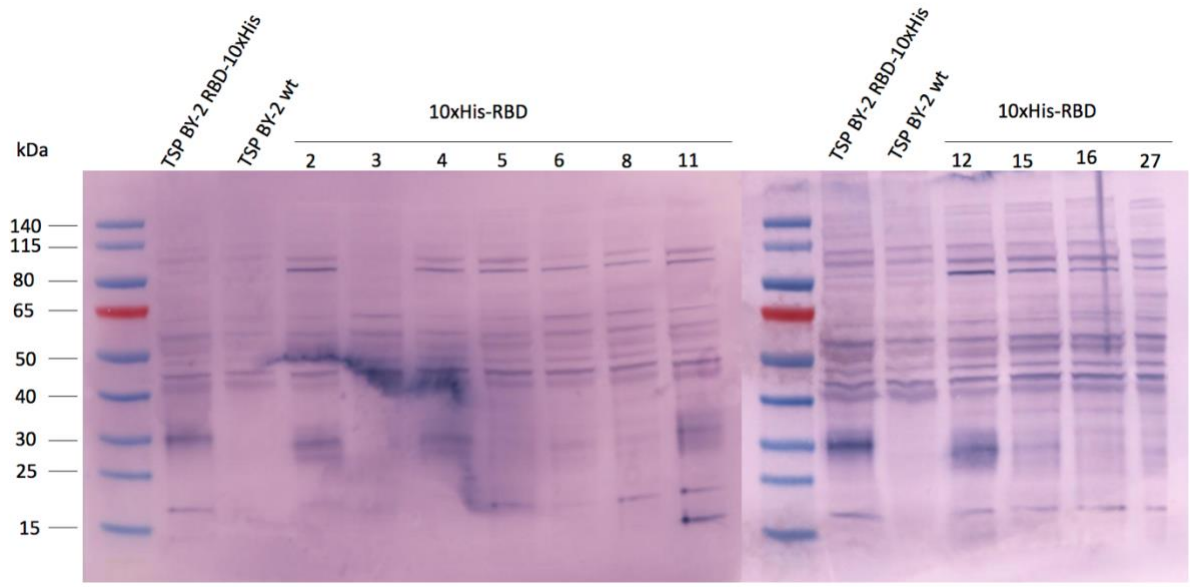

Figure 29: Anti-RBD western blot of the TSP from stably transformed BY-2 cells with the nptII-10xHis-RBD *construct. Each well was loaded with 25 µg of TSLP. Rabbit anti-RBD and Anti-rabbit antibodies conjugated to alkaline phosphatase were used as primary and secondary antibodies, respectively. The revelation was done using a precipitating substrate. TSP from BY-2 cells stably transformed with the* nptII-RBD-10xHis *construct were used as positive control and TSP from wt BY-2 cells were used as negative control.*

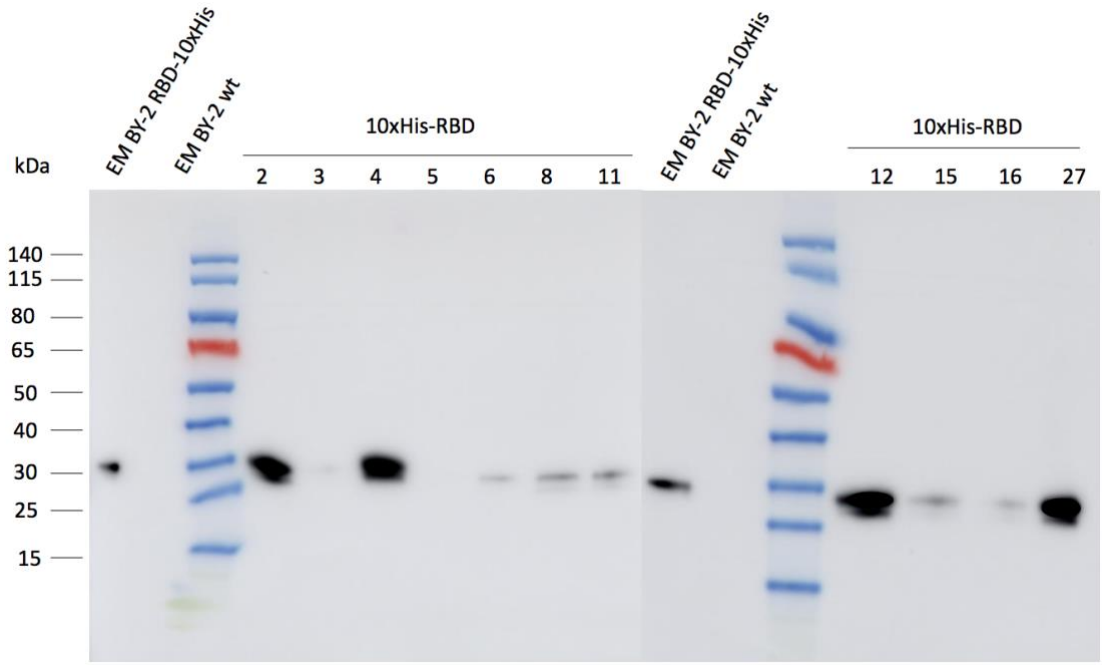

Figure 30: Anti-RBD western blot of the EM from stably transformed BY-2 cells with the nptII-10xHis-RBD *construct. Forty µL of EM were loaded in each well. Rabbit anti-RBD and anti-rabbit antibodies conjugated to the horseradish peroxidase (HRP) were used as primary and secondary antibodies, respectively. The revelation was done using a chemiluminescence substrate. The EM from BY-2 cells stably transformed with the* nptII-RBD-10xHis *construct was used as positive control and the EM from wt BY-2 cells were used as negative control.*

A band at 30 kDa, about the same size as the RBD in the positive control, was observed in most BY-2 cell lines, in both the TSP and the EM, with the exception of lines 3 and 5. This indicates that the RBD was indeed produced in most cell lines. As expected, the majority of the RBD was found in the EM, which as can be inferred from the intensities of the bands.

A noticeable difference in terms of band intensities was also observed among the different cell lines in either the TSP or the EM fractions. A given cell line had the same relative band intensities in both fractions, compared with the positive control, indicating that differences in intensities were due to differences in protein production rather than protein secretion. Differences in the levels of protein production were hypothesized to be due to differences in the number of copies of integrated T-DNA in the cell genome.

2.8.2 Tag intactness

The RBD in the TSP or EM of the best producing BY-2 lines was further analysed for the presence of the 10xHis tag. The lines analysed were the 2, 4, 12 and 27. His-tagged ULP1 produced in *E. coli* and purified on a nickel resin was used as a positive control, while TSP and EM from wt BY-2 cells were used as negative controls. The detection was done by His Probe^{TM}-HRP. The results are shown in Figure 31.

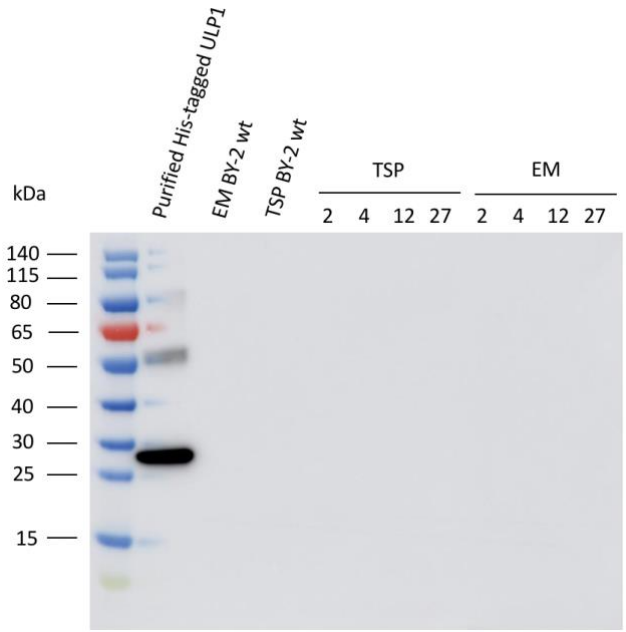

Figure 31: His Probe^{TM}-HRP western blot of the TSP and EM from stably transformed BY-2 cells with the nptII-10xHis-RBD *construct. Forty µL of EM and 25 µg of TSP were loaded in each well. The revelation was done using a chemiluminescence substrate. Twenty µL of purified ULP1 fused to a His tag, produced in* E. coli *was used as a positive control and 25.2 µL of TSP as well as 40 µL of EM from wt BY-2 cells were used as negative controls.*

No bands were detected at the expected size, neither in the TSP or the EM, indicating that the 10xHis tag could have been (partially) degraded. The result is similar to what was observed in *N. benthamiana*.

Chapter III: Discussion

3.1 Protein production in *N. benthamiana* leaves

3.1.1 RBD

In this work, RBD and NTD-RBD fused to an N-terminal affinity tag were produced in different plant systems. In *N. benthamiana* leaves, differences in the yield of the RBD were observed, depending on the type of N-terminal tag fused to the protein. RBD production was analysed by western blot, using anti-RBD antibodies. In the TSLPs of *N. benthamiana* leaves transformed with the constructs encoding 10xHis-RBD, bands were observed at the expected size, showing that the RBD was indeed produced and present in its intact from. In contrast, no bands were detected in the western blots of TSLPs from leaves transformed with the constructs encoding V5-RBD, indicating that no intact RBD was present (Figure 21). Similar results had been observed previously in our laboratory, in transient expression of constructs encoding RBD with a C-terminal 6xHis, 10xHis or V5 tag in *N. benthamiana* leaves. Indeed, anti-RBD western blots had showed bands corresponding to the expected size of the RBD in samples with the 6xHis and 10xHis tags but not with the V5 tag.

No conclusive answer has yet been found to explain the differences in protein expression and/or degradation. The most likely scenario to explain the absence of RBD signals in the V5-RBD samples is that the protein was produced but degraded by endogenous proteases. RBD can be expected to be produced, since the promoter and terminator are the same as in the *10xHis-RBD* expression cassettes.

Affinity tags, on the other hand, have shown unpredictable effects on the production of recombinant proteins in the past, including protein overexpression. Indeed, it has been reported that polyhistidine tags can lead to overexpression in *E. coli*, namely via uneven subcellular distribution (Park *et al*., 2015). Uneven subcellular distribution refers to a phenomenon, common in eukaryotic cells, in which different molecules, such as mRNA or tRNA, are localized in specific regions of the cell. This phenomenon could make these subcellular regions more favourable for protein translation, leading to increased protein production. The authors reported that a polyhistidine tag could serve as a signal for the subcellular uneven distribution of a recombinant protein, leading to its overexpression. They also reported that this effect is more pronounced when the tag is added on the N-terminus of the protein than to the C-terminus, but the mechanism is not yet clear. In the work of this master thesis, however, the signal peptide included in the expression cassettes directed the protein to the secretory pathway, and the 35S promoter already led to overexpression. Although the situation was not the same as in the study conducted by Park *et al.* (2015), an effect due to the tag may still exist. Even though Park *et al.* results were based on polyhistidine tags, which are the most commonly used affinity tags, the same phenomenon could potentially occur when a V5 tag is added to the protein.

Overexpression can lead to protein accumulation that, in turn, can result in protein misfolding. Misfolding can occur if the folding machinery, including chaperones, is overwhelmed and protein accumulation exceeds its folding capacity. Misfolded proteins tend to aggregate in order to minimize the contact of their hydrophobic regions with the aqueous, intracellular environment. Protein aggregates, in turn, could be toxic for the cells, thus misfolded proteins are often more susceptible to degradation by the action of proteases, as a defence mechanism (J. Liu & Howell, 2016). The degradation could therefore be a potential reason for not observing bands in western blots. In this scenario, however, degradation bands at lower molecular weights could still have been visible. Nevertheless, a complete absence of bands is also possible, as it could be due to an enhanced degradation of the RBD. In this case,

the masses of the degradation products would be below the lowest molecular mass that can be detected by the gel electrophoresis. An alternative for the absence of signals could also be a degradation of the epitopes recognized by the used antibodies.

Another potential issue with this hypothesis is that the synthesis and degradation of proteins is a continuous process. So, even if the protein is quickly degraded, faint bands corresponding to the intact protein or degradation fragments should still be detectable in the western blot. It is possible, however, that at the moment of the extraction, the expression of the *V5-RBD* cassette has significantly decreased since the plasmids are only transiently expressed. Therefore, the bands corresponding to the intact protein or degradation fragments might not be detectable in the western blot. The hypothesis of correct production and enhanced degradation afterwards is therefore plausible.

A noticeable difference between the V5 and 10xHis tags is their size. The V5 tag used in this work is about three times longer than the 10xHis tag. A large tag, added to the protein, could negatively impact its proper folding, thus leading to misfolding. In addition, the presence of hydrophobic residues in the V5 tag, such as leucine and isoleucine, could enhance its negative effect on protein folding through unwanted hydrophobic interactions. The outcome in this scenario would be the same as in the case of accumulation due to overexpression, *i.e.* misfolding and consequently, degradation by endogenous proteases.

3.1.2 NTD-RBD

Regarding the production of NTD-RBD fused to an N-terminal 10xHis tag, the TSLP samples from infiltrated leaves showed a band around 25 kDa in addition to the bands also detected in TSLP from non-infiltrated leaves (Figure 23). This 25 kDa band was observed at a lower molecular mass than the expected mass of the RBD alone. This lower band could correspond to a degradation product, indicating that the 10xHis-NTD-RBD was produced but was not intact, probably due to degradation by endogenous proteases. The production of NTD-RBD was attempted for the first time in our laboratory and no reports were found in the literature, so no previous results could be used for comparison.

The detection of bands in the negative control in the western blot suggests that they were due to non-specific interactions with the antibodies.

3.2 Transient expression in BY-2 cells

BY-2 cells were transiently transformed with *A. tumefaciens* containing an *mCherry* expression cassette and a cassette expressing the RBD fused to a 10xHis or V5 tag at either end of the protein. The successful transformation of the BY-2 cells was confirmed by western blots with anti-mCherry primary antibodies. Different intensities of the mCherry bands were observed in the TSP from cells transformed with the different constructs (Figure 26), suggesting a difference in the concentration of mCherry produced in the transformed cells.

The *mCherry* expression cassettes were identical in the different constructs. In addition, the BY-2 cells were from the same batch cultures in the transformations with constructs with the same type of tag. On the other hand, transient transformations with constructs containing different tags, 10xHis or V5, were performed at two different moments, so the BY-2 cells between these transformations are from different batch cultures. Considering that the experimental conditions were the same, the only difference in in this experiment is the *A. tumefaciens* cells, which originate from different batches. There was maybe an intrinsic variability among the *A. tumefaciens* cells containing the different genetic constructs, that made them more or less efficient in the transformation of BY-2 cells. Potentially, the handling and manipulation of the *A. tumefaciens* cells in various experiments in our laboratory could also

have had an impact on their transformation capacity. These presumed differences in transformation capacity could result in differences in the number of BY-2 cells transformed. Delivery of T-DNA by *A. tumefaciens* indeed occurs in a random manner and the number of copies of the constructs that are delivered could vary among the plant cells. Even though, in this case, a large number of transformed cells were analysed, instead of cells coming from a single cell, as in stable transformations, differences in the number of collected BY-2 cells that had been transformed, could impact the results of the western blots. This could be a potential reason for the differences in band intensities observed in the western blot, and thereby in mCherry concentrations.

In terms of RBD production, faint bands corresponding to the expected size were observed in all TSP samples (Figure 27). Even though faint, different band intensities could be distinguished. These bands are absent in the TSP from wt BY-2 cells, used here as negative controls, confirming that RBD is indeed produced. The relative intensities of the RBD bands in the different samples follow the same pattern as the intensities of the mCherry bands, meaning that if a given sample shows a relatively intense mCherry band, it also shows a relatively intense RBD band. However, the ratio of intensities between the RBD and mCherry bands does not seem to be the same in all samples. For example, when comparing the production of 10xHis-RBD and RBD-10xHis, the difference in the intensity of the mCherry bands (Figure 26) is less clear than the difference of intensity of the RBD bands (Figure 27). These observations suggest that the concentration of RBD produced in BY-2 cells transformed with different constructs is partially due to the number of transformed BY-2 cells but might also be partially due to the efficiency of production and/or due to degradation by endogenous proteases. This could indicate that the effect of the production efficiency or susceptibility to degradation depending on the tag position is more important than the number of transformed BY-2 cells in the differences of RBD concentration observed. It seems that the RBD is produced less efficiently or is more susceptible to degradation when the tag is added on the N-terminus, contrary to what has been reported by Park *et al.* (2015).

BY-2 cells and *N. benthamiana* leaves seem to show a difference as expression systems for recombinant proteins, as noticed with the production of RBD with a V5 tag. Even though at very low concentrations, the RBD seems to be present in the TSP from BY-2 cells transiently transformed with the *V5-RBD* and *RBD-V5* expression cassettes, as opposed to the TSLP of *N. benthamiana* leaves transformed with the same cassettes. This indicates that transient expression in *N. benthamiana* leaves might not be completely accurate for predicting stable expression in BY-2 cells. This could be expected, as transient expression in *N. benthamiana* leaves is, in the frame of this work, only a model, used as a faster screening method when producing recombinant proteins in plant systems that require longer periods for the production, such as BY-2 cells. However, differing results could lead to biased interpretations. The model should therefore be carefully considered.

3.3 Stable expression in BY-2 cells

The same *A. tumefaciens* strains used for the transient expression of 10xHis-RBD in *N. benthamiana* leaves were also used for the stable transformation of BY-2 cells. RBD was produced in most of the cell lines, as demonstrated by anti-RBD western blots, with the exception of lines 3 and 5 (Figures 29 and 30). The absence of RBD in these lines was most likely due to lack of production, as a result of silencing, degradation by proteases or insertion of an incomplete T-DNA sequence, rather than to an unsuccessful transformation. The BY-2 transformants had indeed been cultivated in media supplemented with kanamycin before protein extraction. The fact that the cells were able to grow in these media indicates that they had integrated the *nptII* gene, and thus, that the transformation had been successful.

As expected, most of the RBD was found in the EM (Figure 30), since a signal peptide was included in the genetic construct (Figure 8A). This indicates that the *spPDI* worked well for the secretion of the RBD. Nevertheless, a small quantity of RBD was still detected in the TSP (Figure 29). This was expected, since the production is continuous. At any given moment, RBD should therefore be present inside the cell, in the secretory pathway. The same result had been achieved by Rebelo *et al.*(2022). The RBD was produced in stably transformed BY-2 cells. The genetic constructs used included an *RBD-6xHis* expression cassette under the control of the CaMV 35S promoter and 35S terminator, as well as an expression cassette for the kanamycin resistance gene (Rebelo *et al*., 2022). The authors reported the production and secretion of the RBD in the EM, but the tag was not detected either.

Moreover, heterogeneity in the concentrations of RBD could be observed among the different cell lines, as indicated by the differences in RBD band intensities in the western blots. Heterogeneity was expected, because the quantity of RBD produced depends, among others, on the location and the number of copies of the genetic construct that have been inserted by *A. tumefaciens* in the BY-2 cell genome during their co-cultivation.

Finally, the bands corresponding to the RBD in the TSP samples of cells transformed with the *10xHis-RBD* expression cassette were found at a slightly lower molecular mass than the band corresponding to the RBD when the tag was added at the C-terminus, used here as positive control. The RBD band in the positive control had the expected mass. This could indicate that more degradation occurred in the samples of RBD with an N-terminal tag. RBD contains its two *N*-glycosylation sites near its N-terminus, precisely at residues 331 and 343 (Watanabe *et al*., 2020). Another hypothesis could therefore be that the tag interferes with the *N*-glycosylation machinery, for example through steric hindrance, thus altering the glycosylation profile of the RBD. This could result in a slightly different molecular mass in the gel electrophoresis.

3.4 Tag intactness

The RBD produced in *N. benthamiana* leaves transformed with the *10xHis-RBD* construct was tested for the presence of the tag by two different methods, namely by anti-His western blot or purification on a nickel resin. In either case, the tag was not detected. Therefore, it was hypothesised that the tag might not be present on the protein, or at least, partially degraded. The same result had previously been observed in our laboratory in transient expressions of RBD with a C-terminal 6xHis and 10xHis tags in *N. benthamiana* leaves, as well as in stable expressions in BY-2 cells. The latter was also investigated by Rebelo *et al.* (2022). The authors reported that a part of the C-terminal region, including the tag, was degraded by proteases. The tag could, therefore, not be detected by anti-His western blots, indicating that it was indeed degraded. The most probable reason for the absence of the tag, in

the case of this work as well, is its degradation by endogenous proteases. This would occur after protein synthesis, since RBD has been shown to be produced correctly.

Curiously, the tag was detected by anti-His western blot performed on the S1-subunit fused to a C-terminal 10xHis tag, expressed in *N. benthamiana* leaves and in stably transformed BY-2 cells in our laboratory. The protein was purified from the BY-2 TSP on a nickel resin. The purified protein showed a high-mannose *N*-glycosylation profile, which is typical of proteins found in the ER. This suggests that the intact tag is found in the early steps of the secretory pathway, where the protein has not yet been subject to proteases. This was not observed in RBD expressions. The reason could be related to the size difference between RBD and the S1-subunit. A larger protein might take more time to be processed in the ER, increasing the chances to have higher quantities of the protein from the early stages of the secretory pathway in the TSP at the moment of protein extraction. The S1-subunit fused to a 10xHis tag also lost its tag in the later stages of the secretory pathway, as confirmed by anti-His western blots of the BY-2 EM, probably also due to degradation by proteases. Rebelo *et al*. (2022) also reported that the His tag was maintained when the whole spike protein (S1+S2) was expressed in BY-2 cells. A purification assay was the carried out on TSPs from the transformed cells using a His-trap column resulting in a partially purified sample.

Concerning the BY-2 cells transformed with the *10xHis-RBD* construct in this work, the 10xHis tag was not detected on the RBD, neither in the fraction secreted to the EM nor in the TSP. This indicates that the tag might have been (partially) degraded by the action of proteases. These observations are similar to what had been observed previously, with His-tagged RBD produced in *N. benthamiana* leaves or BY-2 cells, showcasing the effect of endogenous proteases in the degradation of recombinant proteins.

Chapter IV: Conclusion and perspectives

4.1 Conclusion

The objective of this work was to evaluate the effect of an N-terminal tag on the production of the SARS-CoV-2 spike protein RBD in BY-2 cells. To do so, different level M Golden Gate plasmids were obtained, containing *10xHis-RBD, 10xHis-NTD-RBD* or *V5-RBD* expression cassettes, as well as selection markers.

The genetic constructs were first evaluated through transient expression in *N. benthamiana* leaves. RBD was successfully produced in leaves transformed with the *10xHis-RBD* expression cassette but not with the *V5-RBD*. The use of *N. benthamiana* leaves as an expression system is very well established and is the most widely used plant system to produce recombinant biopharmaceuticals. However, in this work, it mainly served as an intermediary to predict the possible outcome of stable expressions in BY-2 cells. Two different analyses, namely anti-His western blot and purification on a nickel resin, revealed that the 10xHis tag was not detected on the RBD when it was added to the N-terminus of the protein, suggesting that the tag might be degraded. *10xHis-NTD-RBD* expression was also achieved in *N. benthamiana* leaves, but the protein was not found in its intact form in the TSLP.

Transient expression experiments were also performed in BY-2 cells, with the expression cassettes encoding the RBD fused to a 10xHis or V5 tag at either end of the protein. All the constructs led to RBD production, but with very low yields. Variability in the RBD yields was, however, observed, indicating that some constructs might lead to a better RBD production than others. Particularly, the *RBD-10xHis* construct led to the highest RBD yields, followed by *10xHis-RBD*. It is also important to note that the RBD was produced with constructs with a V5 tag in BY-2 cells, which was not the case in *N. benthamiana* leaves, indicating that there are differences between the two expression systems when producing recombinant proteins.

Finally, RBD production was achieved in stably transformed BY-2 suspension cells, which was the main objective of this work. The genetic constructs used for the transformation contained a *10xHis-RBD* expression cassette including a signal peptide. The signal peptide led to the successful secretion of the RBD in the extracellular medium, easing its downstream purification. The 10xHis tag was, however, not detected by anti-His western blots, indicating that it might have been degraded. Nevertheless, the extracellular medium has a lower content in other proteins compared to the TSP, which simplifies the purification of the protein of interest. Secretion of the RBD would therefore be particularly interesting if the BY-2 production system is to be upscaled.

Overall, the issues encountered, such as the absence of RBD, the absence of the tag or the lower molecular mass than expected of the detected bands, were hypothesized to be due to the action of endogenous proteases, which might be enhanced by the misfolding of the proteins. Degradation of several recombinant proteins produced in different plant systems has been frequently observed in our laboratory, affecting either the tags or the entire proteins. The results obtained in this work are complementary to what had been achieved previously, either in our laboratory or by other authors, in either *N. benthamiana* leaves or BY-2 cells. These results show that the RBD can also be successfully produced and secreted in BY-2 cells when an N-terminal 10xHis tag is added.

4.2 Future perspectives

The ultimate objective of the broader ongoing study in our laboratory regarding the RBD is to study the impact of its *N*-glycosylation profile on the interaction with the hACE2 receptor. To do so, the RBD must be produced in sufficient quantities and purified for the subsequent binding measurements. The next step is, therefore, to compare the RBD production in the different BY-2 cell lines, including the lines obtained previously with constructs with the 10xHis tag on the C-terminus, in order to select the best producing lines. The comparison could be done, for example, by the Enzyme-Linked Immunosorbent Assay (ELISA) method. In this case, purification using affinity resins would not be possible as the 10xHis tag has not been detected on the RBD in any of the BY-2 cells TSPs. Other purification protocols would therefore be needed and are currently being set up in our laboratory.

Alternatively, since the His tag appears to be degraded, the stable expression of RBD with a V5 tag could also be considered, due to the indications that it can be produced in BY-2 cells, as observed in transient expression experiments. However, the V5 tag is not guaranteed to be intact. As it was done with the *V5-RBD* or *RBD-V5* expression cassettes, the *10xHis-NTD-RBD* expression cassette could also be tested through transient expressions in BY-2 cells to determine whether a difference is observed compared with *N. benthamiana* leaves. If a difference is observed and the protein is produced correctly, the construct could be used for stable transformations of BY-2 cells. NTD-RBD could then potentially be used to complement the binding measurements of RBD alone with the hACE2 receptor.

As degradation often seems to occur at the extremities of the protein, leading to tag degradation, it would be interesting to test an *NTD-10xHis-RBD* expression cassette. Including an affinity tag in the middle of a recombinant protein is not a commonly used strategy, as the tag might interfere with the proper folding of the protein. However, NTD and RBD are two distinct domains of the S1-subunit, which tend to fold independently (Xia, 2021b), so the presence of a tag between the two might not be as disruptive as in the middle of another protein sequence. Another potential problem with this strategy is that the tag might be hidden in the 3D structure of the protein and not be accessible to the resin. This question would need to be tested in silico with structure prediction programs such as AlphaFold (https://alphafold.ebi.ac.uk/), but also experimentally. Perhaps in this case, if the protein is expressed, the tag would be less susceptible to degradation by proteases and would still be present on the protein and accessible to the affinity resin, allowing its purification.

In a broader perspective, it could be interesting to further investigate transient expression in BY-2 cells, as this is a more accurate model for predicting the outcome in stable expression in BY-2 cells compared to *N. benthamiana* leaves. So far, the problem with transient expression in BY-2 cells is that the production yields are very low. The protocol used in our laboratory could be further optimised experimentally, by varying different parameters such as the co-cultivation time, the culture medium or the different volumes of cell cultures and/or medium used.

Finally, although currently insufficient, protein production yields in BY-2 cells are expected to increase in the longer term. The RBD, or a modified version, produced in BY-2 cells could then be investigated as a potential alternative vaccine against another SARS-CoV-2 strain that may emerge in the future. In the same order of ideas, BY-2 cells could be used as a platform to produce other recombinant vaccines or biopharmaceuticals.

Chapter V: Materials and methods

5.1 Culture media

5.1.1 Lysogeny Broth medium

The Lysogeny Broth (LB) medium was used for the cultivation of *E. coli*. It contains 10g/L tryptone, 5 g/L yeast extract and 10 g/L NaCl. Solid medium was obtained by adding 15 g/L agar.

For the selection of *E. coli* transformed with pGEM®-T Easy and level 1 Golden Gate plasmids, 100 µg/mL ampicillin was added (LBA). *E. coli* transformed with *L*evel M Golden Gate plasmids were selected by adding100 µg/mL spectinomycin to the solid LB medium (LBS). A mixture of 100 µL of 100 mM Isopropyl ß-D-1-thiogalactopyranoside (IPTG) and 20 µL 50 mM 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) was spread on the plates for blue-white screening.

5.1.2 2x Yeast Extract and Tryptone medium supplemented with glucose and magnesium

The 2x Yeast Extract and Tryptone medium (2YT), supplemented with glucose and magnesium, was used for the cultivation of *A. tumefaciens*. It contains 16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl, 1 g/L glucose and 0.2 g/L $MgSO_4$. Solid medium was obtained by adding 13 g/L of agar. For selection, 20 µg/mL rifampicin, 40 µg/mL gentamicin and 100 µg/mL spectinomycin were added (2YT RGS).

5.1.3 Murashige and Skoog medium

The Murashige and Skoog (MSI) medium was used for the cultivation of BY-2 suspension cells. It contains 30 g/L saccharose, 4.4 g/L Murashige and Skoog basal salts mixture supplemented with vitamins (MP Biomedicals, Solon, OH), 0.2 g/L KH_2PO_4, 2.5 mg/L thiamine, 50 mg/L myo-inositol and 0.2 mg/L 2,4- dichlorophenoxyacetid acid (2,4-D). The pH was adjusted to 5.8 by addition of KOH. Solid medium was obtained by adding 8 g/L agar. Selection was done by adding antibiotics. MSK medium contains 100 mg/mL kanamycin. MSCCK medium contains 400 mg/L carbenicillin, 500 mg/L cefotaxime and 100 mg/L kanamycin.

MSII medium, used for co-cultivation of BY-2 cells with *A. tumefaciens*, was obtained in the same manner as MSI but with an additional 1.8 g/L glucose and the pH adjusted to 5.3.

5.1.4 Paul medium

Paul medium was used for the cultivation of BY-2 suspension cells in the case of a transient transformation. It contains 10 g/L saccharose and 4.3 g/L Murashige and Skoog basal salts mixture supplemented with vitamins (MP Biomedicals, Solon, OH). The pH was adjusted to 5.8 with KOH. Solid medium was obtained by adding 5 g/L agar.

5.2 Cellular strains and plants

5.2.1 Escherichia coli

Electrocompetent *E. coli* TOP10 cells were used for the selection and amplification of plasmids. The cells were spread on LBA or LBS, depending on the resistance genes, and incubated for 16 h at 37°C in the dark. Liquid cultures of 5 mL LBA or LBS were then grown in the dark for 16h at 37°C on a rotary shaker at 120 rpm.

5.2.2 Agrobacterium tumefaciens

Electrocompetent *A. tumefaciens* LBA4404VirG cells were used for the transformation of *N. benthamiana* leaves and BY-2 cells. *A. tumefaciens* cells were spread or streaked on 2YT RGS plates or grown in 10 mL of liquid 2YT RGS, in the dark, at 28°C on a rotary shaker at 120 rpm.

5.2.3 *Nicotiana tabacum* Bright Yellow 2 suspension cells

N. tabacum BY-2 suspension cells were grown in liquid MSI medium in the dark, at 25°C on a rotary shaker at 110 rpm. The cultures were grown in either 4 mL MSI using 6-well CellStar culture plates (Greiner Bio-One) or in 50 mL MSI using 250 mL Erlenmeyer flasks. In the case of 50-mL cultures, an 8% (v/v) inoculum was transferred weekly into fresh medium. BY-2 calli were grown on MSCCK agar plates for nine weeks and transferred to fresh plates every three weeks. They were then transferred to MSK agar plates and then transferred to fresh plates every three weeks.

5.2.4 *Nicotiana benthamiana* plants

N. benthamiana wild type seeds were germinated on potting soil, in transparent boxes kept at 25°C with a 16h-light/8h-dark photoperiod for two weeks. The seedlings were then transferred to Jiffy® pellets and cultivated for two additional weeks under the same conditions. Afterwards, the plants were transferred into plastic pots filled with potting soil and grown at 25°C and 80% humidity. After a period of two weeks, the plants were ready for agroinfiltration.

5.3 DNA manipulations

5.3.1 DNA amplification by Polymerase Chain Reaction (PCR)

The amplification of DNA by PCR was performed with the Q5® High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA). The reaction mix consisted of 10 µL 5X Q5 Reaction Buffer, 1 µL 10 mM dNTPs, 2.5 µL 10 µM Forward Primer, 2.5 µL 10 µM Reverse Primer, 1 µL Template DNA (pUC-GW-Amp-gS) diluted 40x or 200x, 0.5 µL Q5 High-Fidelity DNA Polymerase and 32.5 µL milli-Q water to reach a final volume of 50 µL. The primers used are shown in Table 4.

Table 4: Primers used for the PCR amplification of the sequences encoding the RBD and the NTD-RBD.

Primer	Sequence 5'-3'
RBD Fwd	GGTCTCAAGGGTGCAACCTACAGAA
RBD Rev	TGAGACCAAGCTTACAAGTTTGTTGACTTCTTTGG
NTD-RBD Fwd	GGTCTCACAGTGCGTGAACCTTACT
NTD-RBD Rev	TGAGACCAAGCTTACAAGTTTGTTGACTTCTTTGG

The PCR amplification reaction was performed following the conditions shown in Table 5.

Table 5: Conditions for the PCR amplification of the sequences encoding the RBD and the NTD-RBD.

Step		Temperature	Time
Initial denaturation		98 °C	10 s
25 cycles	Denaturation	98 °C	10 s
	Annealing	64[a]/63[b] °C	30 s
	Extension	72 °C	20[a]/40[b] s
Initial extension		72 °C	2 min

a: RBD/ b: NTD-RBD

After the initial extension, a final extension was performed with the addition of 1 µL dATPs 1 mM and 1 µL GoTaq® polymerase, for the subsequent ligation of the PCR fragment in pGEM®-T Easy vector. This extension was performed at 72°C for 10 min. The temperatures used for the annealing step were calculated using the NEB Tm Calculator (https://tmcalculator.neb.com/#!/main).

5.3.2 Agarose gel electrophoresis

The agarose gel was prepared by mixing 1% w/v agarose and TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8) and dissolving by heating. After cooling down, 0.005% (v/v) Midori Green Advance (Nippon Genetics EUROPE, Germany) was added.

The gel was cast in an appropriate mold with a comb and left to polymerize. DNA was mixed with 1/6 (v/v) Gel Loading Dye, Purple (New England Biolabs, Ipswich, MA) (2.5% Ficoll® 400, 10 mM EDTA, 3.3 mM Tris-HCl, 0.08% SDS, 0.02% Dye 1, 0.0008% Dye 2, pH 8). The gel was placed in the electrophoresis apparatus in TAE buffer and loaded with 10 µL 2-log DNA ladder at 0.1 µg/µL (New England Biolabs, Ipswich, MA) and with 12 µL of samples. The migration was set at 100 V for the appropriate time. The gel was then revealed under UV light (265 nm) (Molecular Imager Gel DocTM XR+, Bio-Rad).

5.3.3 Gel purification

After migration on agarose gel, the desired band was carefully cut and transferred into a 1.5 mL Eppendorf tube. The gel purification was carried out using the innuPrep DoublePure kit (Analytikjena, Jena, Germany). Initially, 650 µL of Gel Solubilizer were added to the Eppendorf tube and the mixture was incubated at 50°C for 10 minutes with regular mixing to ensure dissolution of the gel. After incubation, 50 µL of Binding Optimizer were added and the mixture was vortexed. The mixture was then transferred onto a column and incubated for 5 minutes at room temperature (RT). The column was then centrifuged at 11,000 g (Centrifuge 5425, Eppendorf AG, Hamburg, Germany) for 1 minute and the flow-through was discarded.

Two washings with 700 µL of Washing Solution each followed by centrifugation at 11,000 g for 1 minute were performed. After the second wash, the column was centrifuged an additional time at 11,000 g for 2 minutes. The column was then transferred onto an elution tube and 30 µL of milli-Q water were added. The column was incubated for 5 minutes at RT. After incubation, the column was centrifuged at 11,000 g for 1 minute and the final product was retrieved into the elution tube and stored at -20°C.

5.3.4 Ligation of DNA inserts into pGEM®- T Easy vectors

The ligation reaction was carried out by mixing 5 µL 2x Rapid Ligation Buffer, 1 µL the pGEM®-T Easy vector, 3 µL DNA insert and 1 µL T4 DNA ligase (New England Biolabs, Ipswich, MA) into a 0.5 mL tube. The mixture was mixed by pipetting and incubated at 4°C overnight (O/N).

5.3.5 Transformation of electrocompetent *E. coli* cells by electroporation

Prior to electroporation, desalting of the ligation mix is necessary to prevent arcing. To this end, a 6-well CellStar culture plate (Greiner Bio-One) was filled with 15 mL milli-Q water and a 0.025 µm nitrocellulose membrane filter (Millipore) was placed on the surface. The ligation mix was pipetted on the membrane and left for 10 min. In the meantime, a 45 µL aliquot of electrocompetent *E. coli* TOP 10 cells were thawed on ice for 10 min. After desalting, 5 µL of the ligation mix were added to the E. *coli* cells and the mixture was transferred into a 0.1 cm wide electroporation cuvette (Bio-Rad, Hercules, USA). The electroporation was carried out with the following parameters: C = 25 µF, R = 200 Ω and V = 1.75 kV. The electroporation time constant should be between 3.5 and 4.5 ms. The cells were then quickly recovered into 1 mL LB medium and incubated at 37°C for 1 h. After incubation, 300 µL of cells were spread on LBA plates containing 100 µL 100 mg/ml IPTG and 20 µL 50 mg/mL X-Gal and incubated at 37°C for 16 h.

5.3.6 *E. coli* DNA plasmid isolation (Miniprep)

Prior to the Miniprep, 5-mL of LBA or LBS bacterial culture were incubated O/N at 37°C, with shaking at 120 rpm. The liquid culture was transferred into a 2-mL Eppendorf tube and centrifuged at 4,000 g (Centrifuge 5425, Eppendorf AG, Hamburg, Germany) for 1 min. The supernatant was discarded, and the cells were resuspended in 300 µL resuspension buffer S1 (50 mM Tris-HCl, 10 mM EDTA, 100 µg/mL RNase A, pH 8) at 4°C by vortexing. Three hundred µL of lysis buffer S2 (200 mM NaOH and 1% SDS) were then added, followed by gentle mixing. The mix was incubated for maximum 3 min before the addition of 300 µL of neutralization buffer S3 (2.8 M potassium acetate, pH 5.1) at 4°C. The mix samples were gently mixed and incubated on ice for 10 min. After incubation, the cells were centrifuged at 18,213 g for 5 min (Centrifuge 5430R, Eppendorf AG, Hamburg, Germany) at 4°C and the supernatant was transferred into a 1.5-mL Eppendorf tube. The samples were centrifuged again for 2 min at 20,238 g (Centrifuge 5425, Eppendorf AG, Hamburg, Germany) at RT. The supernatant was transferred into a new 1.5-mL Eppendorf tube and 600 µL isopropanol were added under a hood, followed by gentle mixing. The samples were then centrifuged at maximum speed (18,213 g) for 15 min at 4°C. The supernatant was discarded and 500 µL 70% (v/v) ethanol were added. The samples were then vortexed and centrifuged at maximum speed (18,213 g) for 5 min at 4°C. The supernatant was discarded, and the pellet was dried for about 15 min in a vacuum centrifuge (Speed vac). After drying, the pellet was dissolved in 30 µL milli-Q water and the samples were stored at -20°C.

5.3.7 Control digestion by restriction enzymes

A control digestion was performed by mixing 2 µL of miniprep DNA, 1 µL of the adequate restriction enzyme, 2.5 µL Cut Smart buffer (New England Biolabs, Ipswich, MA) and 19.5 µL of milli-Q water. The mix was incubated for 1 h at 37°C. The digested samples were then analysed by agarose gel electrophoresis.

5.3.8 DNA sequencing

DNA plasmid sequencing was carried out by an external firm, Microsynth Seqlab, using the Sanger method. The reaction mixture consisted of 1 µL purified DNA, 3 µL of the corresponding primer and 8 µL milli-Q water. The samples were labelled and sent to the sequencing facility.

5.3.9 SmartPure Miniprep of *E. coli* plasmids for Golden Gate Assembly

Another method for the isolation of plasmids from *E. coli* cells, using the SmartPure Miniprep kit, was necessary, as using the miniprep protocol described in section 5.3.6 does not lead to successful ligations when using the Golden Gate technique.

The SmartPure Miniprep was done using the SMARTPURE SK-PLPU-100 plasmid kit (Eurogentec, Seraing, Belgium). Between 1- and 1.5-mL O/N LBA or LBS liquid bacteria culture were transferred into a 1.5 mL Eppendorf tube, centrifuged at 10,000 g (Centrifuge 5425, Eppendorf AG, Hamburg, Germany) for 30 s and the supernatant was discarded. The cells were then resuspended in 250 µL Resuspension Buffer (50 mM Tris-HCl, 10 mM EDTA, 100 µg/mL RNAse A, pH 8) at 4°C by vortexing. Next, 250 µL Lysis Buffer (200 mM NaOH, 1% SDS) were added and the tubes were gently inverted several times. This step was allowed to continue for 5 min, before the addition of 350 µL Neutralization Buffer (2.8 M potassium acetate, pH 5.1), followed by gentle mixing. The samples were then centrifuged at 16,000 g (Centrifuge 5425, Eppendorf AG, Hamburg, Germany) for 10 min at RT. The supernatant was transferred into a SmartPure column placed in a collection tube, which was then centrifuged at 6,000 g (Centrifuge 5425, Eppendorf AG, Hamburg, Germany) for 1 min at RT. The flow-through was discarded. The column was then washed by adding 650 µL of Wash Buffer and centrifuged at 12,000 g (Centrifuge 5425, Eppendorf AG, Hamburg, Germany) for 1 min at RT. The wash step was repeated once, followed by an additional centrifugation in the same conditions, to remove any residual liquid. Subsequently, the SmartPure column, containing the isolated DNA, was placed into a 1.5-mL Eppendorf tube and the elution was done by adding 50 µL milli-Q water. The column was left to stand for 1 minute at RT and then centrifuged at 12,000 g for 1 min at RT. The purified DNA was stored at -20°C.

5.3.10 DNA quantification

DNA concentration and purity were measured using the NanoDrop microvolume spectrophotometer (Thermo Fischer Scientific, Waltham, MA).

A blank was taken with 1 µL of milli-Q water. One µL of purified DNA was set onto the NanoDrop pedestal. The purity was determined as the 260/280 and 260/230 ratio, where the purity was considered satisfying when the ratios were under 2.2 and between 1.8 to 2.0 respectively. The concentration was given in ng/µL.

5.3.11 Golden Gate cloning

The cloning of the different inserts into the acceptor plasmid, according to the Golden Gate method, was done using the MoClo kit (Weber *et al*., 2011). After quantification, 20 fmol of each insert and of the acceptor plasmid were mixed in a final volume of 25 µL. The volume of each component was calculated as follows:

$$Volume\ (\mu L) = \frac{20\ (fmol) \times size\ (bp)}{concentration\ (ng/\mu L) \times 1520\ ((fmol \times bp)/ng)}$$

The ligation reaction mix also included 1.5 µL T4 DNA ligase (New England Biolabs), 1 µL of the adequate restriction enzyme and 2.5 µL 10X ligation buffer. The final volume of 25 µL was reached by adding milli-Q water. The reaction mix was then incubated in a thermal cycler in the conditions shown in Table 6.

Table 6: Golden Gate thermal cycle for the restriction-ligation reaction.

Temperature		Time
50 cycles	37 °C	2 min
	16 °C	5 min
BsaI	50 °C	5 min
BpiI	37 °C	
80 °C		10 min
16 °C		indefinitely

5.3.12 Transformation of electrocompetent *A. tumefaciens* cells by electroporation

A 45 µL electrocompetent aliquot of *A. tumefaciens* LBA4404VirG cells was thawed on ice and diluted 4 times with 10% (w/v) glycerol at 4°C. Two µL of plasmid were then added and mixed with the cells. The mix was transferred into a 0.1 cm wide electroporation cuvette (Bio-Rad, Hercules, USA). The electroporation was done using the following parameters: C = 25 µF, R = 400 Ω and V = 12.5 kV (Gene Pulser™, Bio-Rad). The time constant should be between 8 and 12 ms. The cells were then quickly recovered into 1 mL of 2YT medium without antibiotics and incubated for 4 h at 28°C with shaking. A culture volume of 250 µL was then spread on 2YT RGS plates and incubated for 3 days at 28°C.

5.3.13 SmartPure Miniprep of *A. tumefaciens* plasmids

The SmartPure purification of *A. tumefaciens* plasmids was done in a similar way to that of *E. coli* plasmids, using the same kit (SMARTPURE SK-PLPU-100 plasmid kit (Eurogentec, Seraing, Belgium)), with the difference being the preparation of samples. In the case of *A. tumefaciens*, between 7 and 8 mL of O/N liquid LBS culture were centrifuged at 4,000 g (Centrifuge 5804 R, Eppendorf AG, Germany) for 5 min at 28°C, before removing the supernatant. The pellet was resuspended in 250 µL Resuspension Buffer and transferred into a 1.5 mL Eppendorf tube. The rest of the protocol is identical to that of the SmartPure Miniprep of *E. coli* plasmids (section 5.3.9).

5.4 *N. benthamiana* leaves and BY-2 cells transformation

5.4.1 Agroinfiltration of *N. benthamiana* leaves

Three days before the infiltration, *A. tumefaciens* cells containing the construct of interest were streaked on 2YT RGS plates. The day before, 10 mL 2YT RGS were inoculated with a bacterial streak, suspended by vortexing and incubated O/N at 28°C with shaking, to obtain an OD_{600} (Optical Density at 600 nm) of about 3. The day of the agroinfiltration, the OD_{600} was measured, and the appropriate volume of culture was added to leaf infiltration medium (10 mM MES, 10 mM $MgCl_2$, adjusted with KOH to pH 5.6) to obtain an OD_{600} of 0.8 in a final volume of 10 mL. The cultures were centrifuged at 4,200 g (Centrifuge 5804 R, Eppendorf AG, Hamburg, Germany) for 10 min, the supernatant was discarded, and the pellet was resuspended in 10 mL leaf infiltration medium by pipetting. The leaves of 4-week-old *N. benthamiana* plants were then infiltrated on the abaxial side with a 1 mL needless syringe.

5.4.2 Stable transformation of BY-2 cells

Every step was performed under laminar flow and with sterile material to avoid contamination.

5.4.2.1 BY-2 cells and A. tumefaciens cultivation

Five days before co-cultivation, 4 mL of a 7-day-old BY-2 cultivation were transferred into a 250 mL Erlenmeyer flask containing 50 mL MSI medium. Four days before co-cultivation, *A. tumefaciens* LBA4404VirG cells containing the genetic constructs of interest were streaked on 2YT RGS plates. The day before co-cultivation, 10 mL 2YT RGS were inoculated with a bacterial streak, suspended by vortexing and incubated O/N at 28°C with shaking.

5.4.2.2 *Preparation of A. tumefaciens cells*

The OD_{600} was measured by diluting 100 µL of culture in 900 µL 2YT RGS. The appropriate volumes of *A. tumefaciens* cultures were used to inoculate two 2YT RGS cultures to obtain an OD_{600} of 0.5 or 0.7 in a final volume of 10 mL. The cultures were then incubated at 28°C with shaking for 4-5 h to reach an OD_{600} between 1 and 1.5. After incubation, the culture volume required to obtain an OD_{600} of 1.5 in 1 mL was transferred into a 1.5 mL Eppendorf tube. The culture was then centrifuged at 8,000 g (Centrifuge 5425, Eppendorf AG, Germany) for 2 min and the supernatant was discarded. The pellet was resuspended in 1 mL MSII medium by pipetting and the culture was centrifuged again, at 8,000 g for 2 min. The supernatant was discarded, and the pellet was resuspended in 1 mL MSII medium by pipetting.

5.4.2.3 Preparation of BY-2 cells

For each transformation, 25 mL of 5-day-old BY-2 cultures were transferred into a 50 mL Falcon tube and centrifuged at 2,000 g for 5 min (Rotafix 32A, Hettich). The supernatant was discarded by vacuum pump and the pellet was resuspended in 30-40 mL MSII medium by gently by pipetting. This centrifugation step was repeated once, and the supernatant was discarded again. The pellet was suspended in a final volume of 25 mL MSII medium.

5.4.2.4 Co-cultivation of BY-2 cells with A. tumefaciens

A volume of 12.5 µL freshly prepared 100 mM acetosyringone (19.6 mg/mL in 100% ethanol) was added in each Falcon tube containing BY-2 cells for the transformation. In a 6-well CellStar culture plate (Greiner Bio-One), 4 mL BY-2 cell culture were mixed with different volumes of *A. tumefaciens* cultures: 0, 10, 20, 50, 100 and 200 µL. The plate was gently mixed by rotating and placed in the dark at 24°C without agitation for 48 h.

5.4.2.5 Cultivation of transformed BY-2 cells on solid medium

After 2 days of co-cultivation, the co-cultures were transferred in 15 mL Falcon tubes and the wells were washed with 8 mL MSI. The tubes were centrifuged at 2,000 g for 5 min (Rotafix 32A, Hettich) and the supernatant was removed with a vacuum pump. The cells were then resuspended in 10 mL MSI. The tubes were centrifuged, and the supernatant was removed again. The pellet was resuspended in 10 mL MSCCK by pipetting. The tubes were then centrifuged at 2,000 g for 5 min and the supernatant was removed until a cell volume between 1.5 and 1.8 mL was reached. The cells were gently mixed by pipetting and spread on MSCCK plates. The plates were incubated at 24°C in the dark for three weeks until the apparition of calli. When the calli were big enough, 27 of them were transferred on new MSCCK plates by using a scalpel. The calli were then transferred on new MSCCK plates every 3 weeks, again under the same incubation conditions. After two passages, the calli were transferred on MSK plates.

5.4.3 Transient transformation of BY-2 cells

The protocol of transient transformation of BY-2 cells is similar to that of the stable transformation. The main differences are in the medium used for cultivation of BY-2 cells and the incubation period.

5.4.3.1 BY-2 cells and A. tumefaciens cultivations

Five days before co-cultivation, 4 mL of a 7-day-old BY-2 cultivation were transferred into a 250 mL Erlenmeyer flask containing 50 mL MSI medium. Four days before co-cultivation, *A. tumefaciens* LBA4404VirG cells containing the genetic constructs of interest were streaked on 2YT RGS plates. The day before co-cultivation, 10 mL 2YT RGS were inoculated with a bacterial streak, suspended by vortexing and incubated O/N at 28°C with shaking.

5.4.3.2 Preparation of A. tumefaciens cells

The OD_{600} was measured by diluting 100 µL of culture in 900 µL2YT RGS. The appropriate volume of *A. tumefaciens* cultures was inoculated into 2YT RGS to obtain an OD_{600} of 0.45 in a final volume of 10 mL. The culture was then incubated at 28°C with shaking for 4-5 h to reach an OD_{600} between 0.8 and 1. Eight mL of cells were then transferred in a sterile Falcon tube and centrifuged at 4,000 g (Centrifuge 5804 R, Eppendorf AG, Hamburg, Germany) for 5 min. The supernatant was discarded, and the cells were resuspended in 1 mL Paul medium, transferred in a sterile 1.5 mL Eppendorf tube and centrifuged at 9,300 g (Centrifuge 5425, Eppendorf AG, Hamburg, Germany) for 3 min. The supernatant was discarded again, and the pellet was resuspended in 240 µL Paul medium (30 µL Paul medium per mL of culture pelleted).

5.4.3.3 Preparation of BY-2 cells

The BY-2 culture was transferred to a 50 mL Falcon tube and the cells were centrifuged at 700 g (Rotafix 32A, Hettich) for 5 min. The supernatant was removed to reach a cell volume of around 9 mL. The cells were resuspended in 40 mL Paul medium by gently inverting the tube. This step was repeated once. A final centrifugation was performed at 700 g for 5 min and the supernatant was removed to a final volume of 10 mL.

5.4.3.4 *Co-cultivation of BY-2 cells with A. tumefaciens*

Before the co-cultivation, 5 µL of freshly prepared 100 mM acetosyringone were added to the 10 mL BY-2 cells. TheBY-2 cells were then mixed with the *A. tumefaciens* cultures in 15-mL Falcon tubes at a ratio of 60 µL of *A. tumefaciens* cells per 2 mL of BY-2 culture. The negative control consisted of BY-2 cells without *A. tumefaciens* co-cultivation. The tubes were then put on an agitator at 100 rpm for 5 min for homogenisation. Subsequently, between 7 and 9 co-culture cell drops of 140 µL were placed on solid Paul medium plates using a pipet with cut tips. The plates were then covered with Parafilm and incubated in the dark at 25°C for 3 days.

5.4.3.5 Harvest of the BY-2 cells

After the incubation, the cell drops were retrieved with a scalpel into a 50 mL Falcon tube and washed with Paul medium, at a ratio of 450 µL of medium per drop. The cells were transferred in a filtration tube lined with 4-ply Miracloth layers (Calbiochem) and centrifuged at 4,000 g (Rotafix 32A, Hettich) for 5 min. The extracellular medium was discarded, while the cells were transferred in 2 mL screw caped Precellys tubes containing glass beads and frozen in liquid nitrogen.

5.5 Protein manipulations

5.5.1 Protein extraction from *N. benthamiana* leaves

Three leaf discs (60 mg) were sampled from agroinfiltrated leaves 5 days post infiltration and frozen in liquid nitrogen. A negative control was also included, by sampling non-infiltrated leaves in the same way. The leaf discs were then transferred into a glass tissue grinder size 22 (DWK life sciences, Mainz, Germany) and 700 µL extraction medium (60 mM Tris, 250 mM Sorbitol, 2 mM EDTA, pH 8) supplemented with 0.6% Polyclar AT, 10 mM DTT, 1 mM phenylmethylsulfonylfluoride (PMSF) and 1 µg/mL of protease inhibitors (leupeptin, aprotinin, antipain, pepstain and chymostain). The discs were grinded on ice until a homogenous solution was reached. The extracts were centrifuged at 5,000 g (Centrifuge 5430R, Eppendorf AG, Hamburg, Germany) for 5 min at 4°C and the supernatant was recovered in Beckman tubes (Beckman Coulter® Centrifuge Tubes). The supernatant was then centrifuged at 54,000 g for 15 min at 4°C (Optima™ MAX Ultracentrifuge, Beckman Coulter). The supernatant containing the Total Soluble Leaf Proteins (TSLP) was then transferred into a 1.5-mL Eppendorf tube and frozen on liquid nitrogen before storage at -20°C.

5.5.2 Protein extraction from BY-2 cells

After three passages on MSCCK plates, the calli were resuspended in 4 mL MSK in 6-well CellStar culture plates (Greiner Bio-One) and diluted twice in 10 days. Between 2 and 4 mL of BY-2 liquid culture was transferred into a filtration tube lined with 4-ply Miracloth layers (Calbiochem) and centrifuged at 4,000 g for 5 min (Rotafix 32A, Hettich). The extracellular medium (EM) was obtained in the flow through and the cells were contained on the filter. The cells were subsequently transferred in 2 mL screw caped Precellys tubes containing glass beads. The EM and the cells were then frozen in liquid nitrogen. For Total Soluble Protein (TSP) extraction, cell samples were unfrozen and resuspended in 700 µL extraction medium (60 mM Tris, 250 mM Sorbitol, 2 mM EDTA, pH 8) supplemented with 1 µg/mL protease inhibitors (leupeptin, aprotinin, antipain, pepstain and chymostain) and 1 mM PMSF at 4°C. The samples were grinded three times at 5,000 rpm for 40 sec (Precellys 24 lysis and homogenization, Bertin Technologies) and cooled on ice between cycles. The rest of the protocol was done in a cold room at 4°C and on ice. The samples were centrifuged at 5,000 rpm for 5 min (Centrifuge 5430R, Eppendorf AG, Hamburg, Germany) and the supernatant was transferred into a new 1.5-mL Eppendorf tube. If the supernatant was still turbid, this step was repeated. The samples were then centrifuged at 9,400 g (Centrifuge 5430R, Eppendorf AG, Hamburg, Germany) for 7 min. The supernatant was transferred into a Beckman tube (Beckman Coulter® Centrifuge Tubes) and centrifuged at 54,000 g (OptimaTM MAX Ultracentrifuge, Beckman Coulter) for 15 min. The supernatant, containing the TSP, was then transferred to a new 1.5-mL Eppendorf tube and frozen in liquid nitrogen.

5.5.3 TSP quantification by the Bradford method

Three µL of TSP or 4 µL of TSLP sample were mixed with 97 µL 0.1 M NaOH and prepared in triplicates. The BSA standards were prepared in double, at increasing concentrations: 0, 2, 4, 6, 8 and 10 µg/mL, to obtain the calibration curves. One mL Bradford reagent (Coomassie brilliant blue G-250 100 mg/L, 5% (v/v) ethanol, 8.5% (v/v) H_3PO_4) was added to all the samples and standards. The tubes were vortexed and left for 15 min at RT. After 15 min, 220 µL of each sample were transferred into a 96-well plate (Greiner Bio-One) and the absorbance was measured at 595 nm (SPECTROstar Nano BMG Labtech).

5.5.4 Protein gel electrophoresis and western blot

5.5.4.1 Sample preparation

For the TSP, after the concentration was determined (section 5.5.3), the appropriate volume of sample was diluted in extraction medium to have 30 µg of protein in a total volume of 50 µL. For the EM, 40 µL of sample were used without any dilution. The samples were mixed with 1/6 (v/v) Laemmli buffer (0.5 M TrisHCl, pH 6.8, 100 g/L SDS, 30% (v/v) glycerol, 0.12 g/L bromophenol blue) supplemented with 0.6 M DTT, 1 mM PMSF and a protease inhibitor cocktail (leupeptin, aprotinin, antipain, pepstain and chymostain, each at 1 µg/mL), and vortexed. The samples were put into boiling water for 5 min and centrifuged at 20,238 g (Centrifuge 5425, Eppendorf AG, Hamburg, Germany) for 1 min. If a sample changed colour from blue to green/yellow, the pH was adjusted with 1 µL of 1 M Tris pH 9.

5.5.4.2 Gel migration

A 4-20% polyacrylamide gel (SurePAGE™, GenScript) was placed into an electrophoresis cell filled with MOPS buffer (6.06 g/L Tris, 10.46 g/L MOPS, 1 g/L SDS, 0.3 g/L Na_2EDTA). The gel was loaded with 10 µL PageRuler Prestained Protein Ladder (Thermofisher) and either 42 (12 wells gel) or 44 µL (10 wells gel) of each sample, depending on the number of wells in the gel. The migration was set at a constant amperage of 50 mA/gel. After migration, the gel was placed into transfer buffer (6 g/L Tris, 3 g/L glycine, 0.2 g/L SDS and 20% (v/v) methanol) for at least 15 min under agitation.

5.5.4.3 Transfer

The transfer was done using the Trans-Blot Turbo kit (BioRad). The gel and the membrane were placed between blotting papers and air bubbles were removed using a roller previously immersed in transfer buffer. The transfer was set at a constant amperage of 1.5 A for 5 min for 1 gel, or 2.5 A for 10 min for 2 gels. After transfer, the membranes were placed in blocking buffer (20% (v/v) 5x Tris Buffered Saline (TBS) (15 g/L Tris, 40 g/L NaCl and 1 g/L KCl, pH 7.6), 30 g/L milk powder, 0.5% (v/v) Tween20) between 0.5 and 4h with agitation at RT, or O/N at 4°C.

5.5.4.4 Preparation of western Blot

The membranes were washed in washing buffer (20% (v/v) 5x TBS, 2.78 g/L milk powder, 1% (v/v) Tween 20) for 10 min with agitation. They were transferred afterwards in 50 mL Falcon tubes containing 5 mL washing buffer and the appropriate quantity of the primary antibody. The membranes were incubated with rotation for 1h at RT or alternatively O/N at 4°C. After incubation, the membranes were washed 3 times with washing buffer for 5 minutes with agitation at RT. After the third wash, the membranes were transferred in a new 50 mL Falcon tube containing 5 mL of wash buffer and the appropriate quantity of secondary antibody, and they were incubated again for 1 h at RT with rotation. After incubation with the secondary antibody, the membranes were washed again 3 times for 5 min with agitation in washing buffer. The primary and secondary antibodies used are listed in Table 7.

Table 7 : List of primary and secondary antibodies used for western blots.

Function	Antibody	Dilution
Primary antibodies	Rabbit polyclonal anti-RBD antibody, SinoBiological #40592-T62	1:2000
	Rabbit polyclonal anti-mCherry antibody, Rockland #600-401-P16	1:3000
	Penta-His Tag Monoclonal Antibody, ThermoScientific, # P-21315	1:1000
Probes	HisProbe™ – HRP, ThermoScientific, #RJ240374	1:1000
Secondary antibodies	Alkaline phosphatase-conjugated goat anti-rabbit IgG, Sigma, #A3687	1:10000
	HRP-conjugated rat anti-rabbit IgG, SynAbs, #4500058305	1:10000

5.5.4.5 Western blot revelation

For revelation with the Horseradish Peroxidase (HRP), after the washings, 990 µL luminol (BM Chemiluminescence Western Blotting Substrate, Roche) were mixed with 10 µL enhancer (BM Chemiluminescence Western Blotting Substrate, Roche) and spread on the membrane. Chemiluminescence was then detected using the Amersham™ Imager 600 (GE Healthcare).

For revelation with the alkaline phosphatase, the membrane was rinsed with milli-Q water and incubated in 5 mL BM Purple (Roche) in the dark, for 20-45 minutes or until the bands appeared. When the band were visible, the reaction was stopped by washing the membrane with milli-Q water and a picture of the membrane was taken in the Amersham™ Imager 600 (GE Healthcare).

Bibliography

Aleem, A., Akbar Samad, A. B., & Vaqar, S. (2024). Emerging Variants of SARS-CoV-2 and Novel Therapeutics Against Coronavirus (COVID-19). In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK570580/

Amanat, F., Stadlbauer, D., Strohmeier, S., Nguyen, T. H. O., Chromikova, V., McMahon, M., Jiang, K., Arunkumar, G. A., Jurczyszak, D., Polanco, J., Bermudez-Gonzalez, M., Kleiner, G., Aydillo, T., Miorin, L., Fierer, D. S., Lugo, L. A., Kojic, E. M., Stoever, J., Liu, S. T. H., … Krammer, F. (2020). A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nature Medicine*, *26*(7), 1033–1036. https://doi.org/10.1038/s41591-020-0913-5

Anderson, A. R. (1979). Host Specificity in the Genus *Agrobacterium*. *Phytopathology*, *69*(4), 320. https://doi.org/10.1094/Phyto-69-320

Arnau, J., Lauritzen, C., Petersen, G. E., & Pedersen, J. (2006). Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expression and Purification*, *48*(1), 1–13. https://doi.org/10.1016/j.pep.2005.12.002

Assenberg, R., Wan, P. T., Geisse, S., & Mayr, L. M. (2013). Advances in recombinant protein expression for use in pharmaceutical research. *Current Opinion in Structural Biology*, *23*(3), 393–402. https://doi.org/10.1016/j.sbi.2013.03.008

AstraZeneca. (2021, October 11). *AZD7442 reduced risk of developing severe COVID-19 or death in TACKLE Phase III outpatient treatment trial*. https://www.astrazeneca.com/media-centre/press-releases/2021/azd7442-phiii-trial-positive-in-covid-outpatients.html#!

Awasthi, M., Gulati, S., Sarkar, D. P., Tiwari, S., Kateriya, S., Ranjan, P., & Verma, S. K. (2020). The Sialoside-Binding Pocket of SARS-CoV-2 Spike Glycoprotein

Structurally Resembles MERS-CoV. *Viruses*, *12*(9), 909. https://doi.org/10.3390/v12090909

Bai, Y., Yao, L., Wei, T., Tian, F., Jin, D.-Y., Chen, L., & Wang, M. (2020). Presumed Asymptomatic Carrier Transmission of COVID-19. *JAMA*, *323*(14), 1406–1407. https://doi.org/10.1001/jama.2020.2565

Baneyx, F. (1999). Recombinant protein expression in Escherichia coli. *Current Opinion in Biotechnology*, *10*(5), 411–421. https://doi.org/10.1016/s0958-1669(99)00003-8

Baur, A., Reski, R., & Gorr, G. (2005). Enhanced recovery of a secreted recombinant human growth factor using stabilizing additives and by co-expression of human serum albumin in the moss Physcomitrella patens. *Plant Biotechnology Journal*, *3*(3), 331–340. https://doi.org/10.1111/j.1467-7652.2005.00127.x

Bosch, D., Castilho, A., Loos, A., Schots, A., & Steinkellner, H. (2013). N-glycosylation of plant-produced recombinant proteins. *Current Pharmaceutical Design*, *19*(31), 5503–5512. https://doi.org/10.2174/1381612811319310006

Burrell, C. J., Howard, C. R., & Murphy, F. A. (2017). Chapter 31—Coronaviruses. In C. J. Burrell, C. R. Howard, & F. A. Murphy (Eds.), *Fenner and White's Medical Virology (Fifth Edition)* (pp. 437–446). Academic Press. https://doi.org/10.1016/B978-0-12-375156-0.00031-X

Cai, Y., Zhang, J., Xiao, T., Peng, H., Sterling, S. M., Walsh, R. M., Rawson, S., Rits-Volloch, S., & Chen, B. (2020). Distinct conformational states of SARS-CoV-2 spike protein. *Science*, *369*(6511), 1586–1592. https://doi.org/10.1126/science.abd4251

Castilho, A., & Steinkellner, H. (2012). Glyco-engineering in plants to produce human-like N-glycan structures. *Biotechnology Journal*, *7*(9), 1088–1098. https://doi.org/10.1002/biot.201200032

Ceballo, Y., López, A., González, C. E., Ramos, O., Andújar, I., Martínez, R. U., & Hernández, A. (2022). Transient production of receptor-binding domain of SARS-CoV-2 in Nicotiana benthamiana plants induces specific antibodies in immunized mice. *Molecular Biology Reports*, *49*(7), 6113–6123. https://doi.org/10.1007/s11033-022-07402-4

Chadd, H. E., & Chamow, S. M. (2001). Therapeutic antibody expression technology. *Current Opinion in Biotechnology*, *12*(2), 188–194. https://doi.org/10.1016/S0958-1669(00)00198-1

Chant, A., Kraemer-Pecore, C. M., Watkin, R., & Kneale, G. G. (2005). Attachment of a histidine tag to the minimal zinc finger protein of the Aspergillus nidulans gene regulatory protein AreA causes a conformational change at the DNA-binding site. *Protein Expression and Purification*, *39*(2), 152–159. https://doi.org/10.1016/j.pep.2004.10.017

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., Xia, J., Yu, T., Zhang, X., & Zhang, L. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study. *The Lancet*, *395*(10223), 507–513. https://doi.org/10.1016/S0140-6736(20)30211-7

Chi, X., Yan, R., Zhang, J., Zhang, G., Zhang, Y., Hao, M., Zhang, Z., Fan, P., Dong, Y., Yang, Y., Chen, Z., Guo, Y., Zhang, J., Li, Y., Song, X., Chen, Y., Xia, L., Fu, L., Hou, L., … Chen, W. (2020). A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science*, *369*(6504), 650–655. https://doi.org/10.1126/science.abc6952

Coleman, C. M., & Frieman, M. B. (2014). Coronaviruses: Important Emerging Human Pathogens. *Journal of Virology*, *88*(10), 5209–5212. https://doi.org/10.1128/JVI.03488-13

Conley, A. J., Zhu, H., Le, L. C., Jevnikar, A. M., Lee, B. H., Brandle, J. E., & Menassa, R. (2011). Recombinant protein production in a variety of Nicotiana hosts: A comparative analysis. *Plant Biotechnology Journal*, *9*(4), 434–444. https://doi.org/10.1111/j.1467-7652.2010.00563.x

Corman, V. M., Ithete, N. L., Richards, L. R., Schoeman, M. C., Preiser, W., Drosten, C., & Drexler, J. F. (2014). Rooting the Phylogenetic Tree of Middle East Respiratory Syndrome Coronavirus by Characterization of a Conspecific Virus from an African Bat. *Journal of Virology*, *88*(19), 11297–11303. https://doi.org/10.1128/JVI.01498-14

D'Aoust, M., Couture, M. M. -J., Charland, N., Trépanier, S., Landry, N., Ors, F., & Vézina, L. (2010). The production of hemagglutinin-based virus-like particles in plants: A rapid, efficient and safe response to pandemic influenza. *Plant Biotechnology Journal*, *8*(5), 607–619. https://doi.org/10.1111/j.1467-7652.2009.00496.x

De Cleene, M., & De Ley, J. (1976). The host range of crown gall. *The Botanical Review*, *42*(4), 389–466. https://doi.org/10.1007/BF02860827

Dean, N. (1999). Asparagine-linked glycosylation in the yeast Golgi. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1426*(2), 309–322. https://doi.org/10.1016/S0304-4165(98)00132-9

Diego-Martin, B., González, B., Vazquez-Vilar, M., Selma, S., Mateos-Fernández, R., Gianoglio, S., Fernández-del-Carmen, A., & Orzáez, D. (2020). Pilot Production of SARS-CoV-2 Related Proteins in Plants: A Proof of Concept for Rapid Repurposing of Indoor Farms Into Biomanufacturing Facilities. *Frontiers in Plant Science*, *11*. https://doi.org/10.3389/fpls.2020.612781

Dingermann, T. (2008). Recombinant therapeutic proteins: Production platforms and challenges. *Biotechnology Journal*, *3*(1), 90–97. https://doi.org/10.1002/biot.200700214

Drosten, C., Meyer, B., Müller, M. A., Corman, V. M., Al-Masri, M., Hossain, R., Madani, H., Sieberg, A., Bosch, B. J., Lattwein, E., Alhakeem, R. F., Assiri, A. M., Hajomar, W., Albarrak, A. M., Al-Tawfiq, J. A., Zumla, A. I., & Memish, Z. A. (2014). Transmission of MERS-coronavirus in household contacts. *The New England Journal of Medicine*, *371*(9), 828–835. https://doi.org/10.1056/NEJMoa1405858

Drożdżal, S., Rosik, J., Lechowicz, K., Machaj, F., Szostak, B., Przybyciński, J., Lorzadeh, S., Kotfis, K., Ghavami, S., & Łos, M. J. (2021). An update on drugs with therapeutic potential for SARS-CoV-2 (COVID-19) treatment. *Drug Resistance Updates: Reviews and Commentaries in Antimicrobial and Anticancer Chemotherapy*, *59*, 100794. https://doi.org/10.1016/j.drup.2021.100794

Drugmand, J.-C., Schneider, Y.-J., & Agathos, S. N. (2012). Insect cells as factories for biomanufacturing. *Biotechnology Advances*, *30*(5), 1140–1157. https://doi.org/10.1016/j.biotechadv.2011.09.014

Eidenberger, L., Kogelmann, B., & Steinkellner, H. (2023). Plant-based biopharmaceutical engineering. *Nature Reviews Bioengineering*, *1*(6), 426–439. https://doi.org/10.1038/s44222-023-00044-6

European Centre for Disease Prevention and Control. (2024). *MERS-CoV worldwide overview*. https://www.ecdc.europa.eu/en/middle-east-respiratory-syndrome-coronavirus-mers-cov-situation-update

Fehr, A. R., & Perlman, S. (2015). Coronaviruses: An Overview of Their Replication and Pathogenesis. In H. J. Maier, E. Bickerton, & P. Britton (Eds.), *Coronaviruses:*

Methods and Protocols (pp. 1–23). Springer. https://doi.org/10.1007/978-1-4939-2438-7_1

Fielding, B. C. (2011). Human coronavirus NL63: A clinically important virus? *Future Microbiology*, *6*(2), 153–159. https://doi.org/10.2217/fmb.10.166

Finkelstein, M. T., Mermelstein, A. G., Parker Miller, E., Seth, P. C., Stancofski, E.-S. D., & Fera, D. (2021). Structural Analysis of Neutralizing Epitopes of the SARS-CoV-2 Spike to Guide Therapy and Vaccine Design Strategies. *Viruses*, *13*(1), 134. https://doi.org/10.3390/v13010134

Fischer, R., & Emans, N. (2000). Molecular farming of pharmaceutical proteins. *Transgenic Research*, *9*(4), 279–299. https://doi.org/10.1023/A:1008975123362

Fischer, R., Stoger, E., Schillberg, S., Christou, P., & Twyman, R. M. (2004). Plant-based production of biopharmaceuticals. *Current Opinion in Plant Biology*, *7*(2), 152–158. https://doi.org/10.1016/j.pbi.2004.01.007

Fonda, I., Kenig, M., Gaberc-Porekar, V., Pristovaek, P., & Menart, V. (2002). Attachment of histidine tags to recombinant tumor necrosis factor-alpha drastically changes its properties. *TheScientificWorldJournal*, *2*, 1312–1325. https://doi.org/10.1100/tsw.2002.215

Fung, T. S., & Liu, D. X. (2019). Human Coronavirus: Host-Pathogen Interaction. *Annual Review of Microbiology*, *73*(1), 529–557. https://doi.org/10.1146/annurev-micro-020518-115759

Geelen, D. N., & Inzé, D. G. (2001). A bright future for the bright yellow-2 cell culture. *Plant Physiology*, *127*(4), 1375–1379.

Gelvin, S. B. (2003). Agrobacterium-Mediated Plant Transformation: The Biology behind the "Gene-Jockeying" Tool. *Microbiology and Molecular Biology Reviews*, *67*(1), 16–37. https://doi.org/10.1128/MMBR.67.1.16-37.2003

Gerd Gellissen, Alexander W.M. Strasser, & Manfred Suckow. (2005). Key and Criteria to the Selection of an Expression Platform. *Production of Recombinant Proteins. Novel Microbial and Eucaryotic Expression Systems*, 3.

Giovanetti, M., Benedetti, F., Campisi, G., Ciccozzi, A., Fabris, S., Ceccarelli, G., Tambone, V., Caruso, A., Angeletti, S., Zella, D., & Ciccozzi, M. (2021). Evolution patterns of SARS-CoV-2: Snapshot on its genome variants. *Biochemical and Biophysical Research Communications*, *538*, 88–91. https://doi.org/10.1016/j.bbrc.2020.10.102

Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K., & Riggs, A. D. (1979). Expression in Escherichia coli of chemically synthesized genes for human insulin. *Proceedings of the National Academy of Sciences*, *76*(1), 106–110. https://doi.org/10.1073/pnas.76.1.106

Goel, A., Colcher, D., Koo, J. S., Booth, B. J., Pavlinkova, G., & Batra, S. K. (2000). Relative position of the hexahistidine tag effects binding properties of a tumor-associated single-chain Fv construct. *Biochimica Et Biophysica Acta*, *1523*(1), 13–20. https://doi.org/10.1016/s0304-4165(00)00086-6

Goldberg, Y., Mandel, M., Bar-On, Y. M., Bodenheimer, O., Freedman, L., Haas, E. J., Milo, R., Alroy-Preis, S., Ash, N., & Huppert, A. (2021). Waning Immunity after the BNT162b2 Vaccine in Israel. *New England Journal of Medicine*, *385*(24). https://doi.org/10.1056/NEJMoa2114228

Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., Goldman, B. S., Cao, Y., Askenazi, M., Halling, C., Mullin, L., Houmiel, K., Gordon, J., Vaudin, M., Iartchouk, O., Epp, A., Liu, F., Wollam, C., Allinger, M., … Slater, S. (2001). Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58. *Science (New York, N.Y.)*, *294*(5550), 2323–2328. https://doi.org/10.1126/science.1066803

GSK. (2022, February 24). *Medicago and GSK announce the approval by Health Canada of COVIFENZ®, an adjuvanted plant-based COVID-19 vaccine*. https://www.gsk.com/en-gb/media/press-releases/medicago-and-gsk-announce-the-approval-by-health-canada-of-covifenz/

Guan, W., Ni, Z., Hu, Y., Liang, W., Ou, C., He, J., Liu, L., Shan, H., Lei, C., Hui, D. S. C., Du, B., Li, L., Zeng, G., Yuen, K.-Y., Chen, R., Tang, C., Wang, T., Chen, P., Xiang, J., … Zhong, N. (2020). Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine*, *382*(18), 1708–1720. https://doi.org/10.1056/NEJMoa2002032

Guan, Y., Zheng, B. J., He, Y. Q., Liu, X. L., Zhuang, Z. X., Cheung, C. L., Luo, S. W., Li, P. H., Zhang, L. J., Guan, Y. J., Butt, K. M., Wong, K. L., Chan, K. W., Lim, W., Shortridge, K. F., Yuen, K. Y., Peiris, J. S. M., & Poon, L. L. M. (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science (New York, N.Y.)*, *302*(5643), 276–278. https://doi.org/10.1126/science.1087139

Hanke, T., Szawlowski, P., & Randall, R. E. (1992). Construction of solid matrix-antibody-antigen complexes containing simian immunodeficiency virus p27 using tag-specific monoclonal antibody and tag-linked antigen. *The Journal of General Virology*, *73 (Pt 3)*, 653–660. https://doi.org/10.1099/0022-1317-73-3-653

Harvey, W. T., Carabelli, A. M., Jackson, B., Gupta, R. K., Thomson, E. C., Harrison, E. M., Ludden, C., Reeve, R., Rambaut, A., COVID-19 Genomics UK (COG-UK) Consortium, Peacock, S. J., & Robertson, D. L. (2021). SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews. Microbiology*, *19*(7), 409–424. https://doi.org/10.1038/s41579-021-00573-0

HASÖKSÜZ, M., KILIÇ, S., & SARAÇ, F. (2020). Coronaviruses and SARS-COV-2. *Turkish Journal of Medical Sciences*, *50*(9), 549–556. https://doi.org/10.3906/sag-2004-127

Herman, X., Far, J., Peeters, M., Quinton, L., Chaumont, F., & Navarre, C. (2023). In vivo deglycosylation of recombinant glycoproteins in tobacco BY-2 cells. *Plant Biotechnology Journal*, *21*(9), 1773–1784. https://doi.org/10.1111/pbi.14074

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., … Cao, B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, *395*(10223), 497–506. https://doi.org/10.1016/S0140-6736(20)30183-5

Ito, J., Suzuki, R., Uriu, K., Itakura, Y., Zahradnik, J., Deguchi, S., Wang, L., Lytras, S., Tamura, T., Kida, I., Nasser, H., Shofa, M., Begum, M. M., Tsuda, M., Oda, Y., Fujita, S., Yoshimatsu, K., Ito, H., Nao, N., … Sato, K. (2022). *Convergent evolution of the SARS-CoV-2 Omicron subvariants leading to the emergence of BQ.1.1 variant*. bioRxiv. https://doi.org/10.1101/2022.12.05.519085

Jackson, C. B., Farzan, M., Chen, B., & Choe, H. (2022). Mechanisms of SARS-CoV-2 entry into cells. *Nature Reviews Molecular Cell Biology*, *23*(1), 3–20. https://doi.org/10.1038/s41580-021-00418-x

Jung, J.-W., Zahmanova, G., Minkov, I., & Lomonossoff, G. P. (2022). Plant-based expression and characterization of SARS-CoV-2 virus-like particles presenting a native spike protein. *Plant Biotechnology Journal*, *20*(7), 1363–1372. https://doi.org/10.1111/pbi.13813

Juno, J. A., & Wheatley, A. K. (2021). Boosting immunity to COVID-19 vaccines. *Nature Medicine*, *27*(11), 1874–1875. https://doi.org/10.1038/s41591-021-01560-x

Kapila, J., De Rycke, R., Van Montagu, M., & Angenon, G. (1997). An *Agrobacterium*-mediated transient gene expression system for intact leaves. *Plant Science*, *122*(1), 101–108. https://doi.org/10.1016/S0168-9452(96)04541-4

Kapoor, M., Pringle, K., Kumar, A., Dearth, S., Liu, L., Lovchik, J., Perez, O., Pontones, P., Richards, S., Yeadon-Fagbohun, J., Breakwell, L., Chea, N., Cohen, N. J., Schneider, E., Erdman, D., Haynes, L., Pallansch, M., Tao, Y., Tong, S., … Feikin, D. R. (2014). Clinical and laboratory findings of the first imported case of Middle East respiratory syndrome coronavirus to the United States. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *59*(11), 1511–1518. https://doi.org/10.1093/cid/ciu635

Kent, U. M. (1999). Purification of antibodies using protein A-sepharose and FPLC. *Methods in Molecular Biology (Clifton, N.J.)*, *115*, 29–33. https://doi.org/10.1385/1-59259-213-9:29

Kim, J. Y., Kim, Y.-G., & Lee, G. M. (2012). CHO cells in biotechnology for production of recombinant proteins: Current state and further potential. *Applied Microbiology and Biotechnology*, *93*(3), 917–930. https://doi.org/10.1007/s00253-011-3758-5

Kim, T.-G., Lee, H.-J., Jang, Y.-S., Shin, Y.-J., Kwon, T.-H., & Yang, M.-S. (2008). Co-expression of proteinase inhibitor enhances recombinant human granulocyte-macrophage colony stimulating factor production in transgenic rice cell suspension culture. *Protein Expression and Purification*, *61*(2), 117–121. https://doi.org/10.1016/j.pep.2008.06.005

King, A. A., Shrestha, S., Harvill, E. T., & Bjørnstad, O. N. (2009). Evolution of Acute Infections and the Invasion-Persistence Trade-Off. *The American Naturalist*, *173*(4), 446–455. https://doi.org/10.1086/597217

Klinakis, A., Cournia, Z., & Rampias, T. (2021). N-terminal domain mutations of the spike protein are structurally implicated in epitope recognition in emerging SARS-CoV-2 strains. *Computational and Structural Biotechnology Journal*, *19*, 5556–5567. https://doi.org/10.1016/j.csbj.2021.10.004

Kordyukova, L. V., & Shanko, A. V. (2021). COVID-19: Myths and Reality. *Biochemistry. Biokhimiia*, *86*(7), 800–817. https://doi.org/10.1134/S0006297921070026

Kovalenko, A., Ryabchevskaya, E., Evtushenko, E., Nikitin, N., & Karpova, O. (2023). Recombinant Protein Vaccines against Human Betacoronaviruses: Strategies, Approaches and Progress. *International Journal of Molecular Sciences*, *24*(2), 1701. https://doi.org/10.3390/ijms24021701

Lalonde, M.-E., & Durocher, Y. (2017). Therapeutic glycoprotein production in mammalian cells. *Journal of Biotechnology*, *251*, 128–140. https://doi.org/10.1016/j.jbiotec.2017.04.028

Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., & Wang, X. (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, *581*(7807), 215–220. https://doi.org/10.1038/s41586-020-2180-5

Lau, S. K. P., Luk, H. K. H., Wong, A. C. P., Li, K. S. M., Zhu, L., He, Z., Fung, J., Chan, T. T. Y., Fung, K. S. C., & Woo, P. C. Y. (2020). *Possible Bat Origin of Severe Acute Respiratory Syndrome Coronavirus 2—Volume 26, Number 7—July 2020—Emerging Infectious Diseases journal—CDC*. *26*(7). https://doi.org/10.3201/eid2607.200092

Lau, S. K. P., Woo, P. C. Y., Li, K. S. M., Huang, Y., Tsoi, H.-W., Wong, B. H. L., Wong, S. S. Y., Leung, S.-Y., Chan, K.-H., & Yuen, K.-Y. (2005). Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proceedings of the*

National Academy of Sciences, *102*(39), 14040–14045. https://doi.org/10.1073/pnas.0506735102

Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., Azman, A. S., Reich, N. G., & Lessler, J. (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine*, *172*(9), 577–582. https://doi.org/10.7326/M20-0504

Lee, J., Lee, S.-K., Park, J.-S., & Lee, K.-R. (2023). Plant-made pharmaceuticals: Exploring studies for the production of recombinant protein in plants and assessing challenges ahead. *Plant Biotechnology Reports*, *17*(1), 53–65. https://doi.org/10.1007/s11816-023-00821-0

Lehtinen, S., Ashcroft, P., & Bonhoeffer, S. (2021). On the relationship between serial interval, infectiousness profile and generation time. *Journal of the Royal Society, Interface*, *18*(174), 20200756. https://doi.org/10.1098/rsif.2020.0756

Li, Y. (2010). Commonly used tag combinations for tandem affinity purification. *Biotechnology and Applied Biochemistry*, *55*(2), 73–83. https://doi.org/10.1042/BA20090273

Li, Y.-D., Chi, W.-Y., Su, J.-H., Ferrall, L., Hung, C.-F., & Wu, T.-C. (2020). Coronavirus vaccine development: From SARS and MERS to COVID-19. *Journal of Biomedical Science*, *27*(1), 104. https://doi.org/10.1186/s12929-020-00695-2

Liu, H., Zhang, Q., Wei, P., Chen, Z., Aviszus, K., Yang, J., Downing, W., Jiang, C., Liang, B., Reynoso, L., Downey, G. P., Frankel, S. K., Kappler, J., Marrack, P., & Zhang, G. (2021). The basis of a more contagious 501Y.V1 variant of SARS-CoV-2. *Cell Research*, *31*(6), 720–722. https://doi.org/10.1038/s41422-021-00496-8

Liu, J., & Howell, S. H. (2016). Managing the protein folding demands in the endoplasmic reticulum of plants. *New Phytologist*, *211*(2), 418–428. https://doi.org/10.1111/nph.13915

Loper, J. E., & Kado, C. I. (1979). Host range conferred by the virulence-specifying plasmid of Agrobacterium tumefaciens. *Journal of Bacteriology*, *139*(2), 591–596. https://doi.org/10.1128/jb.139.2.591-596.1979

Low, D. E. (2004). Why SARS will not return: A polemic. *CMAJ : Canadian Medical Association Journal*, *170*(1), 68–69.

Ludwig, S., & Zarbock, A. (2020). Coronaviruses and SARS-CoV-2: A Brief Overview. *Anesthesia and Analgesia*, 10.1213/ANE.0000000000004845. https://doi.org/10.1213/ANE.0000000000004845

Markov, P. V., Ghafari, M., Beer, M., Lythgoe, K., Simmonds, P., Stilianakis, N. I., & Katzourakis, A. (2023). The evolution of SARS-CoV-2. *Nature Reviews Microbiology*, *21*(6), 361–379. https://doi.org/10.1038/s41579-023-00878-2

Marra, M. A., Jones, S. J. M., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S. N., Khattra, J., Asano, J. K., Barber, S. A., Chan, S. Y., Cloutier, A., Coughlin, S. M., Freeman, D., Girn, N., Griffith, O. L., Leach, S. R., Mayo, M., McDonald, H., Montgomery, S. B., … Roper, R. L. (2003). The Genome Sequence of the SARS-Associated Coronavirus. *Science*, *300*(5624), 1399–1404. https://doi.org/10.1126/science.1085953

Mayer, A., Sharma, S. K., Tolner, B., Minton, N. P., Purdy, D., Amlot, P., Tharakan, G., Begent, R. H. J., & Chester, K. A. (2004). Modifying an immunogenic epitope on a therapeutic protein: A step towards an improved system for antibody-directed enzyme prodrug therapy (ADEPT). *British Journal of Cancer*, *90*(12), 2402–2410. https://doi.org/10.1038/sj.bjc.6601888

McBride, R., & Fielding, B. C. (2012). The Role of Severe Acute Respiratory Syndrome (SARS)-Coronavirus Accessory Proteins in Virus Pathogenesis. *Viruses*, *4*(11), 2902–2923. https://doi.org/10.3390/v4112902

Moon, K.-B., Jeon, J.-H., Choi, H., Park, J.-S., Park, S.-J., Lee, H.-J., Park, J. M., Cho, H. S., Moon, J. S., Oh, H., Kang, S., Mason, H. S., Kwon, S.-Y., & Kim, H.-S. (2022). Construction of SARS-CoV-2 virus-like particles in plant. *Scientific Reports*, *12*(1), 1005. https://doi.org/10.1038/s41598-022-04883-y

Nagata, T., Nemoto, Y., & Hasezawa, S. (1992). Tobacco BY-2 Cell Line as the "HeLa" Cell in the Cell Biology of Higher Plants. In K. W. Jeon & M. Friedlander (Eds.), *International Review of Cytology* (Vol. 132, pp. 1–30). Academic Press. https://doi.org/10.1016/S0074-7696(08)62452-3

Navarre, C., De Muynck, B., Alves, G., Vertommen, D., Magy, B., & Boutry, M. (2012). Identification, gene cloning and expression of serine proteases in the extracellular medium of Nicotiana tabacum cells. *Plant Cell Reports*, *31*(10), 1959–1968. https://doi.org/10.1007/s00299-012-1308-y

Neuman, B. W., Kiss, G., Kunding, A. H., Bhella, D., Baksh, M. F., Connelly, S., Droese, B., Klaus, J. P., Makino, S., Sawicki, S. G., Siddell, S. G., Stamou, D. G., Wilson, I. A., Kuhn, P., & Buchmeier, M. J. (2011). A structural analysis of M protein in coronavirus assembly and morphology. *Journal of Structural Biology*, *174*(1), 11–22. https://doi.org/10.1016/j.jsb.2010.11.021

Oostra, M., de Haan, C. A. M., & Rottier, P. J. M. (2007). The 29-nucleotide deletion present in human but not in animal severe acute respiratory syndrome coronaviruses disrupts the functional expression of open reading frame 8. *Journal of Virology*, *81*(24), 13876–13888. https://doi.org/10.1128/JVI.01631-07

Paraskevis, D., Kostaki, E. G., Magiorkinis, G., Panayiotakopoulos, G., Sourvinos, G., & Tsiodras, S. (2020). Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event. *Infection, Genetics and Evolution*, *79*, 104212. https://doi.org/10.1016/j.meegid.2020.104212

Park, W.-J., You, S.-H., Choi, H.-A., Chu, Y.-J., & Kim, G.-J. (2015). Over-expression of recombinant proteins with N-terminal His-tag via subcellular uneven distribution in Escherichia coli. *Acta Biochimica Et Biophysica Sinica*, *47*(7), 488–495. https://doi.org/10.1093/abbs/gmv036

Piccoli, L., Park, Y.-J., Tortorici, M. A., Czudnochowski, N., Walls, A. C., Beltramello, M., Silacci-Fregni, C., Pinto, D., Rosen, L. E., Bowen, J. E., Acton, O. J., Jaconi, S., Guarino, B., Minola, A., Zatta, F., Sprugasci, N., Bassi, J., Peter, A., De Marco, A., … Veesler, D. (2020). Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell*, *183*(4), 1024-1042.e21. https://doi.org/10.1016/j.cell.2020.09.037

Pormohammad, A., Ghorbani, S., Khatami, A., Farzi, R., Baradaran, B., Turner, D. L., Turner, R. J., Bahr, N. C., & Idrovo, J.-P. (2020). Comparison of confirmed COVID-19 with SARS and MERS cases - Clinical characteristics, laboratory findings, radiographic signs and outcomes: A systematic review and meta-analysis. *Reviews in Medical Virology*, *30*(4), e2112. https://doi.org/10.1002/rmv.2112

Rebelo, B. A., Folgado, A., Ferreira, A. C., & Abranches, R. (2022). Production of the SARS-CoV-2 Spike protein and its Receptor Binding Domain in plant cell suspension cultures. *Frontiers in Plant Science*, *13*. https://doi.org/10.3389/fpls.2022.995429

Reusken, C. B., Haagmans, B. L., Müller, M. A., Gutierrez, C., Godeke, G.-J., Meyer, B., Muth, D., Raj, V. S., Vries, L. S.-D., Corman, V. M., Drexler, J.-F., Smits, S. L., El

Tahir, Y. E., De Sousa, R., van Beek, J., Nowotny, N., van Maanen, K., Hidalgo-Hermoso, E., Bosch, B.-J., … Koopmans, M. P. (2013). Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: A comparative serological study. *The Lancet Infectious Diseases*, *13*(10), 859–866. https://doi.org/10.1016/S1473-3099(13)70164-6

Reuter, L. J., Bailey, M. J., Joensuu, J. J., & Ritala, A. (2014). Scale-up of hydrophobin-assisted recombinant protein production in tobacco BY-2 suspension cells. *Plant Biotechnology Journal*, *12*(4), 402–410. https://doi.org/10.1111/pbi.12147

S. J. Sheen. (1982). Biomass and chemical composition of tobacco plants under high density growth. *35th Tobacco Chemists' Research Conference, Winston-Salem, N.C. 1981*.

Santos, R. B., Abranches, R., Fischer, R., Sack, M., & Holland, T. (2016). Putting the Spotlight Back on Plant Suspension Cultures. *Frontiers in Plant Science*, *7*. https://doi.org/10.3389/fpls.2016.00297

Santos, R. B., Chandrasekar, B., Mandal, M. K., Kaschani, F., Kaiser, M., Both, L., van der Hoorn, R. A. L., Schiermeyer, A., & Abranches, R. (2018). Low Protease Content in Medicago truncatula Cell Cultures Facilitates Recombinant Protein Production. *Biotechnology Journal*, *13*(7), e1800050. https://doi.org/10.1002/biot.201800050

Schoeman, D., Gordon, B., & Fielding, B. C. (2022). Coronaviruses. *Encyclopedia of Infection and Immunity*, 241–258. https://doi.org/10.1016/B978-0-12-818731-9.00052-5

Senga, M., Arabi, Y. M., & Fowler, R. A. (2017). Clinical spectrum of the Middle East respiratory syndrome coronavirus (MERS-CoV). *Journal of Infection and Public Health*, *10*(2), 191–194. https://doi.org/10.1016/j.jiph.2016.04.008

Shroff, R. T., Chalasani, P., Wei, R., Pennington, D., Quirk, G., Schoenle, M. V., Peyton, K. L., Uhrlaub, J. L., Ripperger, T. J., Jergović, M., Dalgai, S., Wolf, A., Whitmer, R.,

Hammad, H., Carrier, A., Scott, A. J., Nikolich-Žugich, J., Worobey, M., Sprissler, R., … Bhattacharya, D. (2021). Immune responses to two and three doses of the BNT162b2 mRNA vaccine in adults with solid tumors. *Nature Medicine*, *27*(11), 2002–2011. https://doi.org/10.1038/s41591-021-01542-z

Singh, A. K., Singh, A., Singh, R., & Misra, A. (2020). Remdesivir in COVID-19: A critical review of pharmacology, pre-clinical and clinical studies. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, *14*(4), 641–648. https://doi.org/10.1016/j.dsx.2020.05.018

Siriwattananon, K., Manopwisedjaroen, S., Shanmugaraj, B., Rattanapisit, K., Phumiamorn, S., Sapsutthipas, S., Trisiriwanich, S., Prompetchara, E., Ketloy, C., Buranapraditkun, S., Wijagkanalan, W., Tharakhet, K., Kaewpang, P., Leetanasaksakul, K., Kemthong, T., Suttisan, N., Malaivijitnond, S., Ruxrungtham, K., Thitithanyanont, A., & Phoolcharoen, W. (2021). Plant-Produced Receptor-Binding Domain of SARS-CoV-2 Elicits Potent Neutralizing Responses in Mice and Non-human Primates. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.682953

Slootstra, J. W., Kuperus, D., Plückthun, A., & Meloen, R. H. (1997). Identification of new tag sequences with differential and selective recognition properties for the anti-FLAG monoclonal antibodies M1, M2 and M5. *Molecular Diversity*, *2*(3), 156–164. https://doi.org/10.1007/BF01682203

Southern, J. A., Young, D. F., Heaney, F., Baumgärtner, W. K., & Randall, R. E. (1991). Identification of an epitope on the P and V proteins of simian virus 5 that distinguishes between two isolates with different biological characteristics. *The Journal of General Virology*, *72 (Pt 7)*, 1551–1557. https://doi.org/10.1099/0022-1317-72-7-1551

Starr, T. N., Greaney, A. J., Hilton, S. K., Ellis, D., Crawford, K. H. D., Dingens, A. S., Navarro, M. J., Bowen, J. E., Tortorici, M. A., Walls, A. C., King, N. P., Veesler, D., & Bloom, J. D. (2020). Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell*, *182*(5), 1295-1310.e20. https://doi.org/10.1016/j.cell.2020.08.012

Sukenik, S. C., Karuppanan, K., Li, Q., Lebrilla, C. B., Nandi, S., & McDonald, K. A. (2018). Transient Recombinant Protein Production in Glycoengineered Nicotiana benthamiana Cell Suspension Culture. *International Journal of Molecular Sciences*, *19*(4), 1205. https://doi.org/10.3390/ijms19041205

Sun, Q.-M., Chen, L.-L., Cao, L., Fang, L., Chen, C., & Hua, Z.-C. (2005). An improved strategy for high-level production of human vasostatin120-180. *Biotechnology Progress*, *21*(4), 1048–1052. https://doi.org/10.1021/bp049583x

Tan, E., Chin, C. S. H., Lim, Z. F. S., & Ng, S. K. (2021). HEK293 Cell Line as a Platform to Produce Recombinant Proteins and Viral Vectors. *Frontiers in Bioengineering and Biotechnology*, *9*. https://doi.org/10.3389/fbioe.2021.796991

Uddin, M. N., & Roni, M. A. (2021). Challenges of Storage and Stability of mRNA-Based COVID-19 Vaccines. *Vaccines*, *9*(9), 1033. https://doi.org/10.3390/vaccines9091033

Van Der Hoek, L., Pyrc, K., & Berkhout, B. (2006). Human coronavirus NL63, a new respiratory virus. *FEMS Microbiology Reviews*, *30*(5), 760–773. https://doi.org/10.1111/j.1574-6976.2006.00032.x

Wallinga, J., & Lipsitch, M. (2007). How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings. Biological Sciences*, *274*(1609), 599–604. https://doi.org/10.1098/rspb.2006.3754

Walsh, G. (2010). Biopharmaceutical benchmarks 2010. *Nature Biotechnology*, *28*(9), 917–924. https://doi.org/10.1038/nbt0910-917

Walsh, G. (2014). Biopharmaceutical benchmarks 2014. *Nature Biotechnology*, *32*(10), 992–1000. https://doi.org/10.1038/nbt.3040

Walsh, G., & Walsh, E. (2022). Biopharmaceutical benchmarks 2022. *Nature Biotechnology*, *40*(12), 1722–1760. https://doi.org/10.1038/s41587-022-01582-x

Wang, Z., Schmidt, F., Weisblum, Y., Muecksch, F., Barnes, C. O., Finkin, S., Schaefer-Babajew, D., Cipolla, M., Gaebler, C., Lieberman, J. A., Oliveira, T. Y., Yang, Z., Abernathy, M. E., Huey-Tubman, K. E., Hurley, A., Turroja, M., West, K. A., Gordon, K., Millard, K. G., … Nussenzweig, M. C. (2021). mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature*, *592*(7855), 616–622. https://doi.org/10.1038/s41586-021-03324-6

Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S., & Crispin, M. (2020). Site-specific glycan analysis of the SARS-CoV-2 spike. *Science (New York, N.Y.)*, *369*(6501), 330–333. https://doi.org/10.1126/science.abb9983

Watanabe, Y., Bowden, T. A., Wilson, I. A., & Crispin, M. (2019). Exploitation of glycosylation in enveloped virus pathobiology. *Biochimica Et Biophysica Acta. General Subjects*, *1863*(10), 1480–1497. https://doi.org/10.1016/j.bbagen.2019.05.012

Weber, E., Engler, C., Gruetzner, R., Werner, S., & Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. *PloS One*, *6*(2), e16765. https://doi.org/10.1371/journal.pone.0016765

Weiss, S. R., & Navas-Martin, S. (2005). Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiology and Molecular Biology Reviews: MMBR*, *69*(4), 635–664. https://doi.org/10.1128/MMBR.69.4.635-664.2005

Williams, L., Jurado, S., Llorente, F., Romualdo, A., González, S., Saconne, A., Bronchalo, I., Martínez-Cortes, M., Pérez-Gómez, B., Ponz, F., Jiménez-Clavero, M. Á., & Lunello, P. (2021). The C-Terminal Half of SARS-CoV-2 Nucleocapsid Protein, Industrially Produced in Plants, Is Valid as Antigen in COVID-19 Serological Tests. *Frontiers in Plant Science*, *12*, 699665. https://doi.org/10.3389/fpls.2021.699665

Woo, P. C. Y., Lau, S. K. P., Chu, C., Chan, K., Tsoi, H., Huang, Y., Wong, B. H. L., Poon, R. W. S., Cai, J. J., Luk, W., Poon, L. L. M., Wong, S. S. Y., Guan, Y., Peiris, J. S. M., & Yuen, K. (2005). Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *Journal of Virology*, *79*(2), 884–895. https://doi.org/10.1128/JVI.79.2.884-895.2005

World Health Organization. (2024a). *Coronavirus disease (COVID-19): SARS-CoV-2 variant classifications and definitions.* https://www.who.int/health-topics/coronavirus#tab=tab_1

World Health Organization. (2024b). *Middle East respiratory syndrome coronavirus (MERS-CoV).* https://www.who.int/health-topics/middle-east-respiratory-syndrome-coronavirus-mers

World Health Organization. (2024c). *Severe acute respiratory syndrome (SARS).* https://www.who.int/health-topics/severe-acute-respiratory-syndrome

World Health Organization. (2024d). *Tracking SARS-CoV-2 Variants*. https://www.who.int/activities/tracking-SARS-CoV-2-variants

World Health Organization. (2024e). *Vaccines and immunization*. https://www.who.int/health-topics/vaccines-and-immunization#tab=tab_1

Wu, Z., & McGoogan, J. M. (2020). Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of

72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA*, *323*(13), 1239–1242. https://doi.org/10.1001/jama.2020.2648

Xia, X. (2021a). Domains and Functions of Spike Protein in Sars-Cov-2 in the Context of Vaccine Design. *Viruses*, *13*(1), 109. https://doi.org/10.3390/v13010109

Xia, X. (2021b). Domains and Functions of Spike Protein in Sars-Cov-2 in the Context of Vaccine Design. *Viruses*, *13*(1), 109. https://doi.org/10.3390/v13010109

Xiao, K., Zhai, J., Feng, Y., Zhou, N., Zhang, X., Zou, J.-J., Li, N., Guo, Y., Li, X., Shen, X., Zhang, Z., Shu, F., Huang, W., Li, Y., Zhang, Z., Chen, R.-A., Wu, Y.-J., Peng, S.-M., Huang, M., … Shen, Y. (2020). Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature*, *583*(7815), 286–289. https://doi.org/10.1038/s41586-020-2313-x

Xu, X., Nagarajan, H., Lewis, N. E., Pan, S., Cai, Z., Liu, X., Chen, W., Xie, M., Wang, W., Hammond, S., Andersen, M. R., Neff, N., Passarelli, B., Koh, W., Fan, H. C., Wang, J., Gui, Y., Lee, K. H., Betenbaugh, M. J., … Wang, J. (2011). The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nature Biotechnology*, *29*(8), 735–741. https://doi.org/10.1038/nbt.1932

Yadav, N. S., Vanderleyden, J., Bennett, D. R., Barnes, W. M., & Chilton, M. D. (1982). Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proceedings of the National Academy of Sciences of the United States of America*, *79*(20), 6322–6326. https://doi.org/10.1073/pnas.79.20.6322

Yao, H., Song, Y., Chen, Y., Wu, N., Xu, J., Sun, C., Zhang, J., Weng, T., Zhang, Z., Wu, Z., Cheng, L., Shi, D., Lu, X., Lei, J., Crispin, M., Shi, Y., Li, L., & Li, S. (2020). Molecular Architecture of the SARS-CoV-2 Virus. *Cell*, *183*(3), 730-738.e13. https://doi.org/10.1016/j.cell.2020.09.018

Yurkovetskiy, L., Wang, X., Pascal, K. E., Tomkins-Tinch, C., Nyalile, T. P., Wang, Y., Baum, A., Diehl, W. E., Dauphin, A., Carbone, C., Veinotte, K., Egri, S. B., Schaffner, S. F., Lemieux, J. E., Munro, J. B., Rafique, A., Barve, A., Sabeti, P. C., Kyratsous, C. A., … Luban, J. (2020). Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell*, *183*(3), 739-751.e8. https://doi.org/10.1016/j.cell.2020.09.032

Zaki, A. M., Van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. M. E., & Fouchier, R. A. M. (2012). Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. *New England Journal of Medicine*, *367*(19), 1814–1820. https://doi.org/10.1056/NEJMoa1211721

Zhang, J., Xiao, T., Cai, Y., & Chen, B. (2021). Structure of SARS-CoV-2 spike protein. *Current Opinion in Virology*, *50*, 173–182. https://doi.org/10.1016/j.coviro.2021.08.010

Zhao, D., & Huang, Z. (2016). Effect of His-Tag on Expression, Purification, and Structure of Zinc Finger Protein, ZNF191(243-368). *Bioinorganic Chemistry and Applications*, *2016*, 8206854. https://doi.org/10.1155/2016/8206854

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., Chen, H.-D., Chen, J., Luo, Y., Guo, H., Jiang, R.-D., Liu, M.-Q., Chen, Y., Shen, X.-R., Wang, X., … Shi, Z.-L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, *579*(7798), 270–273. https://doi.org/10.1038/s41586-020-2012-7

Effect of an N-terminal tag on the production of SARS-CoV-2 spike protein receptor binding domain in *Nicotiana tabacum* Bright Yellow 2 cells.

Joel Skënderaj

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) is the causative agent of the Coronavirus Disease 19 (COVID-19), which started a pandemic in 2019. The SARS-CoV-2 virion is composed of four structural proteins, namely the Spike (S), Membrane (M), Envelope (E) and Nucleocapsid (N) proteins. The S protein is a heavily glycosylated trimer, with 66 *N*-glycosylation sites, that binds to the human angiotensin-converting enzyme 2 (hACE2) receptor at the host cell surface and enables viral entry by acting as a class I membrane fusion protein. More specifically, the binding occurs through the Receptor Binding Domain (RBD). The S protein, in particular the RBD, is the main target of host neutralizing antibodies. As a result, it has been extensively studied to better understand its properties and develop vaccines.

Recombinant vaccines, focused on the S protein or its RBD, were among the various strategies considered for vaccine development. Among the different recombinant protein production platforms, plants have gained increasing interest in recent years. They offer advantages compared to mammalian cells, such as low production costs and low risk of contamination with human pathogens. Plant suspension cells, such as the *Nicotiana tabacum* Bright Yellow 2 (BY-2) cells, offer additional advantages, as they are easily amenable to comply with good manufacturing practices and easily scalable, making them particularly interesting for the production of recombinant proteins of pharmaceutical value. However, issues related to these platforms, such as low yields due to degradation by endogenous proteases, have hindered their commercialisation.

In this work, two recombinant protein production platforms were used, namely *Nicotiana benthamiana* leaves and BY-2 suspension cells. Genetic constructs encoding the RBD with a 10xHis or V5 N-terminal tag were transiently expressed in both platforms, leading to the successful production of the protein. An affinity tag was included in the genetic constructs to enable a simple and efficient downstream purification of the protein. The tag was added on the N-terminus of the protein, as previous results with a C-terminal tag had shown tag degradation. However, in this case as well, the N-terminal tag was not detected by western blots, indicating that it might have been degraded by proteases. Another construct, encoding the N-terminal domain (NTD) and the RBD of the SARS-CoV-2 S protein, namely *10xHis-NTD-RBD*, was also transiently expressed in *N. benthamiana* leaves for the first time. Finally, the *10xHis-RBD* construct was stably expressed in BY-2 cells. While the tag was not detected by western blot, the protein was successfully produced and secreted in the extracellular medium, which still makes its purification easier.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté des bioingénieurs
Croix du Sud, 2 bte L7.05.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/agro**