

https://dspace.org/

A Clot to Uncover: FOXP3 and SARS-CoV-2 Nucleocapsid Interactions and Their Effect on DNA Binding

Mcinnes, Keiran

2024
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/42068
Mcinnes, Keiran. (2024). A Clot to Uncover: FOXP3 and SARS-CoV-2 Nucleocapsid Interactions and Their Effect on DNA Binding [Master's dissertation, University of the Witwatersrand, Johannesburg]. WireDSpace.

Downloaded from a DSpace repository. https://dspace.org/



A Clot to Uncover: FOXP3 and SARS-CoV-2 Nucleocapsid Interactions and Their Effect on DNA Binding

by

Keiran Mcinnes

(1709919)

Dissertation

Submitted in fulfilment of the requirements for the degree

Master of Science

in

Molecular and Cell Biology

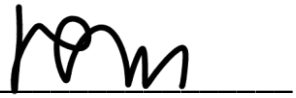
in the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa

Supervisor: Dr Sylvia Fanucchi

March 2024

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.



(Signature of candidate)

18th Day of March 2024 at Johannesburg

Research Outputs

Molecular Biosciences Research Thrust 2021

- Online GradFlash presentation (People's Choice Award)
- Silybin Binding to FOXP3 and Its Effect on DNA Binding
- Keiran Mcinnes and Sylvia Fanucchi
- University of the Witwatersrand, Johannesburg, South Africa
- 9 December, 2021

Molecular Biosciences Research Thrust 2022

- Poster presentation
- FOXP3 in COVID-19: A Clot to Uncover
- Keiran Mcinnes and Sylvia Fanucchi
- University of the Witwatersrand, Johannesburg, South Africa
- 1 December, 2021

14th Cross-Faculty Postgraduate Symposium

- GradFlash Presentation (3rd place – Masters)
- FOXP3 in COVID-19: A Clot to Uncover
- Keiran Mcinnes and Sylvia Fanucchi
- University of the Witwatersrand, Johannesburg, South Africa
- 6 – 8 September, 2023

Molecular Biosciences Research Thrust 2023

- Poster presentation
- FOXP3 in COVID-19: A Clot to Uncover
- Keiran Mcinnes and Sylvia Fanucchi
- University of the Witwatersrand, Johannesburg, South Africa
- 7 December, 2021

Abstract

During COVID-19, systemic coagulopathy, which can lead to strokes and embolisms, is often observed in COVID-19 patients and may also contribute to long COVID. This coagulopathy is the result of overactivated platelets in circulation that leads to inappropriate clot formation. FOXP3 is a transcription factor involved in platelet development and loss of FOXP3 function leads to platelets that resemble those seen during COVID-19. Thus, FOXP3 may be dysregulated in COVID-19. The SARS-CoV-2 nucleocapsid (NC) is a multifunctional protein typically associated with viral genome packaging and virion assembly. However, it is also capable of binding DNA and is potentially able to alter regulation of host protein expression. Here, potential interactions between the DNA-binding forkhead domain (FHD) of FOXP3 and the SARS-CoV-2 NC were investigated. Identification of a novel interaction between FOXP3 and SARS CoV-2 NC may provide new clues as to the pathophysiology of COVID-19. To address this aim, both proteins were overexpressed in T7 *E. coli*, purified via immobilised metal affinity chromatography, and monitored for potential interactions in the absence and presence of DNA using pull-down assays and fluorescence anisotropy. A direct interaction was identified between the two proteins in the absence of DNA. Additionally, it was found that both proteins are capable of binding to DNA at the same time, but excess NC was found to cause FHD dissociation from the FHD-NC-DNA complex. This result implicates NC in FOXP3 dysfunction which may be associated with the coagulopathy and other symptoms seen during COVID-19. Additionally, NC DNA binding does not appear to be driven by the FOXP3 consensus sequence, indicating that FOXP3 may not be the only transcription factor potentially dysregulated by NC.

Acknowledgements

The biggest thank you to my supervisor Dr Fanucchi. I am enormously grateful for allowing me to pursue this project. Thank you for the French toast, looking at my cats when I should be showing results, and manually fixing every in to in this dissertation.

Thank you to Dr Ike and Prof. Sayed for all the insights, tips, and tricks. Particular thanks to Dr Ike who helped with MD simulations. While those results didn't make it in this work, I appreciate the time you gave to show me how it's done and how to interpret the data.

Huge thank you to Heather Donald. For her mentorship in the beginning of my lab life, to her continued friendship, support, and anime recommendations.

Thank you to the Foxy group and everyone else in the PSFRU. Thank you for letting me distract you when I've just discovered a new lab party trick or something else completely mundane.

Thank you to the Wits Postgraduate Merit Award and the South African Medical Research Council for funding.

Table of Contents

Declaration 2
Research Outputs 3
Abstract 4
Acknowledgements 5
List of Figures 8
List of Tables 9
List of Equations 10
List of Abbreviations 11
Introduction 13
COVID-19 13
1.1 - Coagulopathy and Microclots 13
1.1.2 - Haemostasis 14
1.2 - The p38 MAPK Pathway 17
1.3 - FOXP3 18
1.3.1 - FOXP3 as a Transcription Factor 18
1.3.2 - The FOX Proteins and FOXP Subfamily 20
1.3.3 - Structure of FOXP3 21
1.4 - The SARS-CoV-2 Nucleocapsid 22
1.5 - Rationale 25
2 – Aims and Objectives 27
2.1 – Aim 27
2.2 – Objectives 27
3 - Methods and Materials 28
3.1 - DNA Oligonucleotides 28
3.2 - Protein Sequences and Purification 29
3.2.1 - Protein Constructs 29
3.2.2 - Transformation 29
3.2.3 - Expression 31
3.3 - Purification 32
3.3.1 - Immobilised Metal Affinity Chromatography 32
3.3.2 - Thrombin Cleavage 34
3.3.3 - Size Exclusion Chromatography 34
3.3.4 - SDS-PAGE 36
3.3.5 - Evaluation of Protein Concentration 37

3.4 - Structural Characterisation 38
3.4.1 - Circular Dichroism Spectroscopy 38
3.4.2 - Intrinsic Tryptophan Fluorescence 39
3.5 - DNA Binding and Protein-Protein Interactions 40
3.5.1 - Electrophoretic Mobility Shift Assays 40
3.5.2 - Pulldown Assays 42
3.5.3 - Protein Labelling 43
3.5.4 - Fluorescence Anisotropy 44
4 – Results 47
4.1 - Protein Purification 47
4.1.1 - FOXP3 FHD 47
4.1.2 - Nucleocapsid 49
4.1.3 - Nucleocapsid SEC 51
4.1.4 - Nucleocapsid Autolysis 53
4.1.5 - FOXP3 FHD Structural Characterisation 55
4.1.6 - Nucleocapsid Structural Characterisation 55
4.2 - Protein-Protein, Protein-DNA, and Protein-Protein-DNA Interactions 59
4.2.1 - FOXP3 FHD-NC Protein-Protein Interaction 59
4.2.2 - Protein-DNA Interactions 60
4.2.3 - Protein-Protein-DNA Interactions 63
5 - Discussion 67
5.1 - FOXP3 FHD – Nucleocapsid Interactions 67
5.2 - Protein – DNA Interactions 68
5.2.1 - FOXP3 FHD 68
5.2.2 - SARS-CoV-2 NC 69
5.3 - Protein – Protein – DNA Interactions 70
6 - Conclusion 74
7 – Supplementary Material 75
8 - References 80

List of Figures

Fig. 1: Overview of how severe SARS-CoV-2 infection leads to lung damage in the alveoli 14
Fig. 2: The coagulation cascade 15
Fig. 3: The structure of platelets in healthy individuals and those found during severe COVID-19 16
Fig. 4: The role of p38 MAPK activity in platelets 18
Fig. 5: Activated platelets in healthy vs IPEX syndrome patients on fibrinogen coated slides 19
Fig. 6: Domains of FOXA, FOXM1, FOXO, and FOXP proteins showing the diversity within the FOX protein family 20
Fig. 7: The FOXP3 FHD bound to anti-parallel DNA strands 21
Fig. 8: Model of the SARS-CoV-2 viral particle and nucleocapsid 22
Fig. 9: Structure of the SARS-CoV-2 Nucleocapsid (NC) 24
Fig. 10: Schematic of the pET-11a plasmid used to code for both the FOXP3 FHD and SARS-CoV-2 NC 29
Fig. 11: The T7 expression system 30
Fig. 12: Concept of IMAC Purification, with FOXP3 FHD (green cartoon fox) bound to the nickel charged IMAC column 32
Fig. 13: The principle of size exclusion chromatography 34
Fig. 14: Characteristic CD spectra of various protein secondary structural elements 38
Fig. 15: Electrophoretic mobility shift assays 40
Fig. 16: The concept of a pulldown assay 41
Fig. 17: Structure of NTA-ATTO 550 42
Fig. 18: Principle of fluorescence anisotropy 44
Fig. 19: FOXP3 FHD purification 47
Fig. 20: NC purification 49
Fig. 21: NC purification following thrombin cleavage 50
Fig. 22: Size exclusion chromatography of NC 51
Fig. 23: Size exclusion standards used for NC size exclusion 52
Fig. 24: NC autolysis 53
Fig. 25: FOXP3 FHD Structural Characterisation 55
Fig. 26: Size exclusion standards used for FOXP3 FHD size exclusion 56
Fig. 27: NC Structural Characterisation 57
Fig. 28: FOXP3 FHD and NC protein-protein interaction 59
Fig. 29: FOXP3 FHD DNA binding 60
Fig. 30: NC DNA binding 61
Fig. 31: Fluorescence anisotropy of the protein-protein-DNA interactions between FOXP3 FHD, NC, and the p38α promoter sequence 62
Fig. 32: Effects of increasing NC on preformed FHD-DNA complexes 64
Fig. 33: Representative docked structure between the NTD of NC (pink) and the p38α promoter sequence 64
Fig. 34: FOXP3 FHD – NC – DNA interactions 72

List of Tables

Table 1: DNA oligonucleotide sequences used in this work 27
Table 2: Protein constructs used for this study 28
Table 3: Standards used for each size exclusion column and their respective molecular masses 35
Table 4: Fluorescence anisotropy experiments performed. 45
Table 5: Overall fluorescence anisotropy results 65

List of Equations

Equation 1	$A = \varepsilon c l$	36
Equation 2	$[\theta] = (100*\theta)/(c*n*l)$	37
Equation 3	$E = hc/\lambda$	38
Equation 4	$[\text{ATTO 550}] = A_{554}/\varepsilon_{554}$	43
Equation 5	$[\text{Protein}] = ((A_{280})-0.1*(A_{554}))/\varepsilon_{280}$	43
Equation 6	Labelling Efficiency = [ATTO 550]/[Protein]	43
Equation 7	$r = (I_{\parallel}-G*I_{\perp})/(G*I_T)$	44
Equation 8	$G = I_{\perp}/I_{\parallel}$	44
Equation 9	$f(x) = (B_{max}*[\text{Ligand}])/(K_D+[\text{Ligand}])$	44

List of Abbreviations

2xYT Media	2x Yeast Tryptone Broth
AAV P5	Adeno-Associated Virus P5
ADP	Adenosine Diphosphate
APS	Ammonium Persulfate
ASK1	Apoptosis Signal-Regulating Kinase 1
A_{XXX}	Absorbance at XXX nm
CD Spectroscopy	Circular Dichroism Spectroscopy
CD-XX	Cluster of Differentiation XX (Eg: CD-25)
COVID-19	Coronavirus Disease 2019
CTD	C - Terminal Domain
DBE2	Daf-16 Family-Binding Element 2
DIC	Disseminated Intravascular Coagulopathy
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
EMSA	Electrophoretic Mobility Shift Assay
ERK	Extracellular Signal-Related Kinase
EX	Elution Fraction Number X
FHD	Forkhead Domain
FOXPX	Forkhead Box P X
FPLC	Fast Protein (Or Performance) Liquid Chromatography
FT	Flowthrough
G-Factor	Grating Factor
G3BP1/2	Ras GTPase-Activating Protein-Binding Protein 1/2
His-Tag	Hexahistidine Tag
hnRNP	Heterogeneous Nuclear Ribonucleoproteins
hsp27	Heat Shock Protein 27
$I_{\parallel}$	Intensity of emitted fluorescence parallel to the z-axis
$I_{\perp}$	Intensity of emitted fluorescence perpendicular to the z-axis
I_T	Total intensity of emitted fluorescence
IDP	Intrinsically Disordered Protein
IMAC	Immobilised Metal Affinity Chromatography
IPEX Syndrome	Immunodysregulation Polyendocrinopathy Enteropathy X-linked Syndrome
IPTG	Isopropyl β- d-1-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
JNK	c-Jun NH2-Terminal Kinase
K_D	Dissociation Constant
kDa	kilo-Daltons (1 kDa = 1000 g/mol)
LLPS	Liquid-Liquid Phase Separation
MAPK	Mitogen Activated Protein Kinase
MAP2K	Mitogen Activated Protein Kinase Kinase
MAP3K	Mitogen Activated Protein Kinase Kinase Kinase
MAS	Macrophage Activation Syndrome
MEKK1	MEK Kinase 1

MS	Mass Spectrometry
MW	Molecular Weight
NC	Nucleocapsid
NET	Neutrophil Extracellular Trap
NFAT1	Nuclear Factor of Activated T-cells 1
NHS	N-Hydroxysuccinimide
NMR	Nuclear Magnetic Resonance
NTA	Nitrilotriacetic Acid
NTD	N - Terminal Domain
OD	Optical Density
P	Pellet
PAMP	Pathogen-Associated Molecular Pattern
PMSF	Phenylmethylsulfonyl Fluoride
r	Ansiotropy
RNA	Ribonucleic Acid
ROX	Carboxyrhodamine
RPM	Rotations Per Minute
RUNX1	Runt-Related Transcription Factor 1
SARS-COV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SATB1	Special AT-Rich Sequence-Binding Protein-1
SDS-PAGE	Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
SN	Supernatant
SOC Media	Super Optimal Broth with Catabolite Repression
SSL	Single Site Long
TAE Buffer	Tris-Acetic Acid-EDTA Buffer
TAK1	Transforming Growth Factor-β-Activated Kinase 1
TBR1	T-box, Brain, 1
TBS	Tris Buffered Saline
TCEP	Tris(2-carboxyethyl)phosphine
TEMED	Tetramethylethylenediamine
TGF-β	Transforming Growth Factor-β
TIC	Transcription Initiation Complex
Treg	Regulatory T Cell
UTR	Untranslated Region
Xyyy	Residue X at position yyy of a protein (Eg: Lys262)
XyyyZ	Residue X at position yyy mutated to residue Z (Eg: Pro372Ala)
YY1	Ying Yang 1
α2AP	Alpha 2-antiplasmin
ε_{XXX}	Extinction Coefficient at XXX nm

Introduction

COVID-19

The COVID-19 pandemic officially led to the deaths of approximately 7 million people worldwide (World Health Organisation, 2023). While the worst of the pandemic is largely behind us, SARS-CoV-2, the causative agent behind COVID-19, is predicted to become endemic and possibly a seasonal occurrence like that of influenza (Telenti et al., 2021). Because of this, development of effective treatments for COVID-19 is crucial. COVID-19 is primarily a respiratory disease commonly presenting with fever, coughing, pneumonia, and loss of smell and taste (Bliddal et al., 2021; Struyf et al., 2021). However, symptoms vary widely between cases in both intensity and prevalence: from asymptomatic to life-threatening (Struyf et al., 2021). Additionally, approximately a third to one half of those who recover from COVID-19, regardless of disease severity, report ongoing complications; a state commonly referred to as "long COVID" (Carfì et al., 2020; Crook et al., 2021; Gold et al., 2021; Raveendran et al., 2021). Physical symptoms of long COVID appear to persist for weeks to months and typically include fatigue, persistent coughing, dyspnoea, muscle and joint pain, headache, and diarrhoea (Aiyegbusi et al., 2021; Carfì et al., 2020; Ho et al., 2022; Raveendran et al., 2021). Worryingly, both acute and long COVID also present with a number of neurological and cognitive symptoms; including trouble focusing, speech difficulties, memory problems, depression, and continued alterations to sense of taste and smell (Chen et al., 2021; Guo et al., 2022a, 2022b; Mao et al., 2020; Perlis et al., 2021). Furthermore, these symptoms have been linked to both physical and electrical changes in various areas of the brain caused by neurological damage during SARS-CoV-2 infection (Guo et al., 2022a, 2022b). However, the exact mechanisms behind both the physical and neurological symptoms of long COVID remain elusive. One key factor in the development of these symptoms likely lies in direct infection and subsequent damage of various tissues; such as infection of skeletal muscles leading to fatigue (Crook et al., 2021). Hyperactivation of the immune system can also contribute to symptoms of long COVID; where excessive response to the pathogen leads to lasting tissue damage (**Fig. 1**) (Crook et al., 2021; Jamal et al., 2021; Merad & Martin, 2020; Yao et al., 2021). Finally, another likely contributor in the development of long COVID is the hypercoagulable state that develops during disease onset.

1.1 - Coagulopathy and Microclots

Severe COVID-19 patients often experience disseminated intravascular coagulation (DIC), a syndrome presenting with excessive clot formation within the body (Kander, 2020; Levi et al., 2020; Yuriditsky et al., 2020). These clots can then go on to block blood flow throughout the body, resulting in widespread ischaemia leading to, among others, renal insufficiency, neurological problems, and difficulty breathing (Levi et al., 2020). The formation of larger clots can lead to embolisms in the lungs or strokes in the brains of patients, resulting in rapid deterioration and death (Klok et al., 2020). Unsurprisingly, blood drawn from COVID-19 patients has been found to possess elevated markers of coagulation (Goshua et al., 2020; Iba et al., 2020; Levi et al., 2020) as well as a greater viscosity than normal (Venter et al., 2020). Additionally, the platelets and red blood cells of COVID-19 patients show morphological abnormalities which may contribute towards excessive coagulation (Venter et al., 2020). Furthermore, these clots appear to persist following resolution of the acute SARS-CoV-2 infection in the form of microclots. These microclots have been found in the blood of patients who have otherwise completely recovered from COVID-19 (Pretorius et al., 2021) and are abnormally resistant to degradation by proteolysis and are enriched in inflammatory markers. By persisting within the circulatory systems of

patients, these microclots may be causing micro-ischaemic events through the obstruction of capillaries throughout the body. These events, while not life threatening in and of themselves, may be contributing to the common symptoms of long COVID by reducing oxygen supply to various tissues.

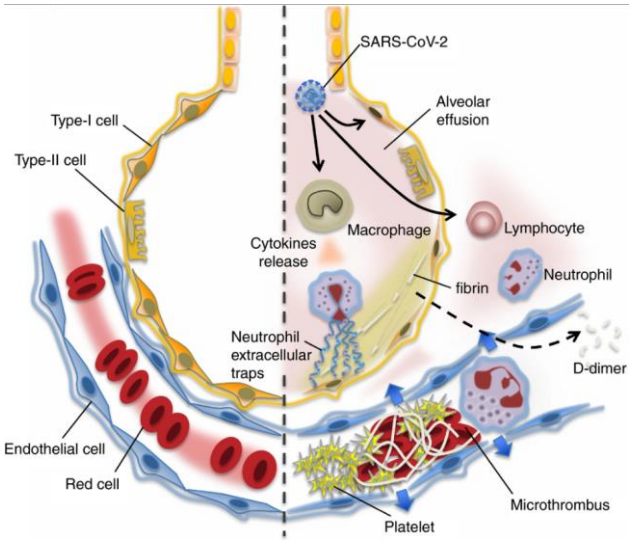


Fig. 1: Overview of how severe SARS-CoV-2 infection leads to lung damage in the alveoli. The left half of the figure depicts a healthy alveolus. The right half shows an alveolus during severe COVID-19. SARS-CoV-2 infects lung tissue, leading to inflammation and immune cell infiltration. Local inflammation leads to alveolar effusion, fluid build-up in the alveoli, preventing gas exchange. Overactive macrophages secrete large amounts of inflammatory cytokines, further worsening inflammation while also overstimulating other immune cells. The proinflammatory state promotes the formation of neutrophil extracellular traps (NETs), further obstructing gaseous exchange. Severe inflammation also leads to a hypercoagulable state, resulting in the formation of clots which only further dampens gas exchange. Taken from Iba et al, 2020.

However, the mechanisms behind the increased coagulability observed in the blood of COVID-19 patients have yet to be fully elucidated. One contributing factor appears to be the direct damage of endothelial cells (Luo et al., 2020), where damage to these cells leads to activation of clotting factors through the release of damage associated compounds. Alternatively, the release of proinflammatory cytokines leads to enhanced production of platelets as well as the upregulation, release, and activation of coagulation substrates and clotting factors (Esmon, 2005). These two factors combine to result in a hypercoagulable state; which is only worsened with overstimulation of the immune system.

1.1.2 - Haemostasis

Haemostasis is the process of blood coagulation in response to injury and is the mechanism by which blood loss is halted. Haemostasis involves an activation cascade of proteins, both freely soluble in the plasma and bound to the membranes of endothelial cells, that results in the activation of thrombin which then cleaves soluble fibrinogen to insoluble fibrin. The insoluble fibrin then forms a mesh over the injury to halt blood flow while also acting as a scaffold for wound repair (**Fig. 2**) (Dahlbäck, 2000). In COVID-19, various clotting factors and clotting associated compounds have been found to be

upregulated; such as fibrinogen, D-Dimer, thrombomodulin, and α2-antiplasmin (α2AP) (Bouck et al., 2021). α2AP in particular is a potent inhibitor of plasmin, the primary protease responsible for clot degradation, and is present in large amounts within clots produced during COVID-19 (Pretorius et al., 2021; Reed et al., 2017); explaining their abnormal resistance to degradation.

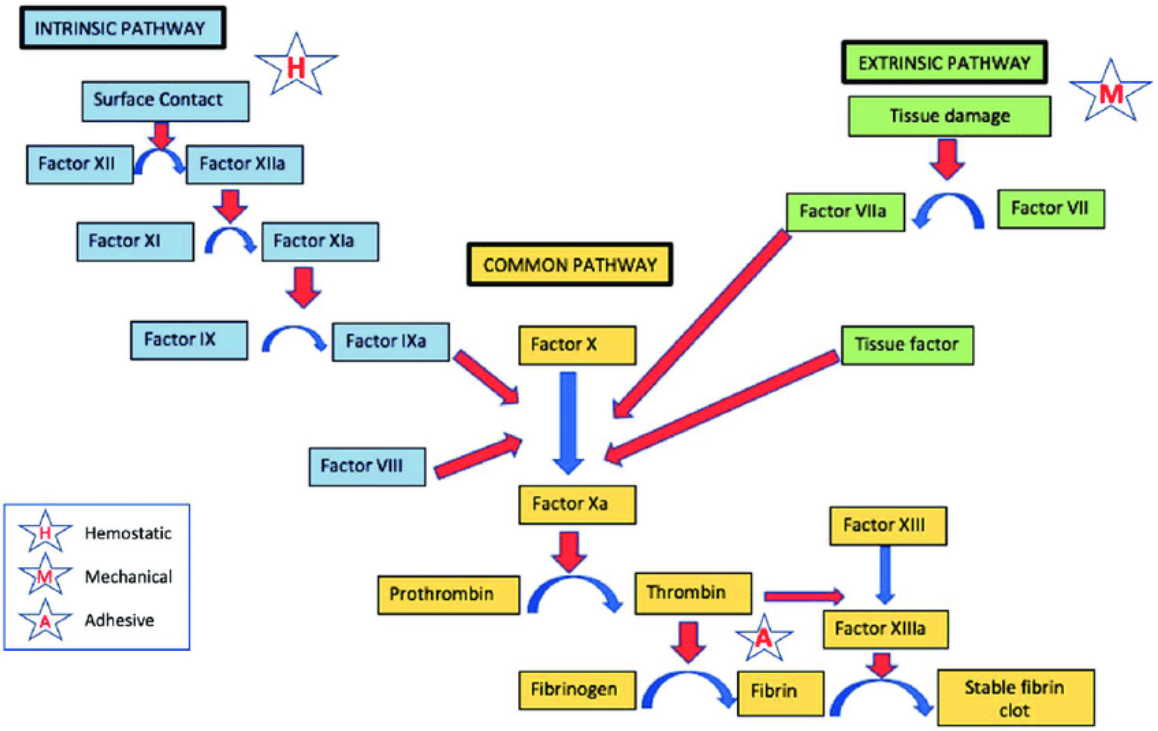


Fig. 2: The coagulation cascade. The coagulation cascade involves a series of protein activations which eventually converge on thrombin activation. Thrombin then cleaves soluble fibrinogen to insoluble fibrin. The intrinsic pathway is primarily activated through inflammation while the extrinsic pathway is activated following tissue damage. Taken from Dammann et al, 2020.

Assisting coagulation are platelets. These discoid shaped nucleus-free fragments of cytoplasm are found free floating within the plasma. During haemostasis, platelets act to adhere to the site of injury, providing a foundation upon which fibrin, coagulation factors, and additional platelets can adhere (Jurk & Kehrel, 2005).

Despite lacking a nucleus, platelets are highly functional and capable of responding to various environmental conditions. Upon adhesion, they undergo morphological changes (**Fig. 3A** and **3B**) and secrete signalling molecules such as adenosine diphosphate (ADP), serotonin, transforming growth factor – β (TGF-β), and CD-40 ligand that act to recruit more platelets and immune cells to the site of injury (Jurk & Kehrel, 2005). These morphological changes during platelet activation are the result of p38 pathway activity (Minuz et al., 2002; Saklatvala et al., 1996), an intracellular signalling cascade that will be discussed further in the following section. Following adhesion, activated p38 phosphorylates heat-shock protein 27 (hsp27). Hsp27 then binds to and promotes actin polymerisation, causing cytoskeletal rearrangement and the subsequent structural changes in the platelet (**Fig. 3B** and **3C**) (Jurk & Kehrel, 2005; Zhu et al., 1994). Interestingly, the morphological changes observed in activated platelets are reminiscent of the morphological changes observed in the platelets of COVID-19 patients (**Fig. 3D**), potentially indicating that platelets are being inappropriately activated in COVID-19. As expected, p38 activity within platelets is significantly elevated in COVID-19 with a concurrent increase in platelet aggregation (Manne et al., 2020). It is known that the Forkhead

Box Protein P3 transcription factor (FOXP3), typically associated with regulatory T cells (Tregs), is in fact required for the correct formation of platelets in megakaryocytes; the cells responsible for platelet production (Bernard et al., 2009). Loss of FOXP3 function leads to the formation of structurally and functionally abnormal platelets similar to those seen in COVID-19. Furthermore, there is tentative evidence that SARS-CoV-2 is capable of infecting both platelets and their megakaryocyte progenitors (Shen et al., 2021); albeit with limited efficiency. This suggests that the platelet pathology observed in COVID-19 may be the result of the virus infecting megakaryocytes and disrupting FOXP3 transcriptional activity, leading to the generation of malformed and functionally aberrant platelets while also potentially promoting p38 MAPK activity.

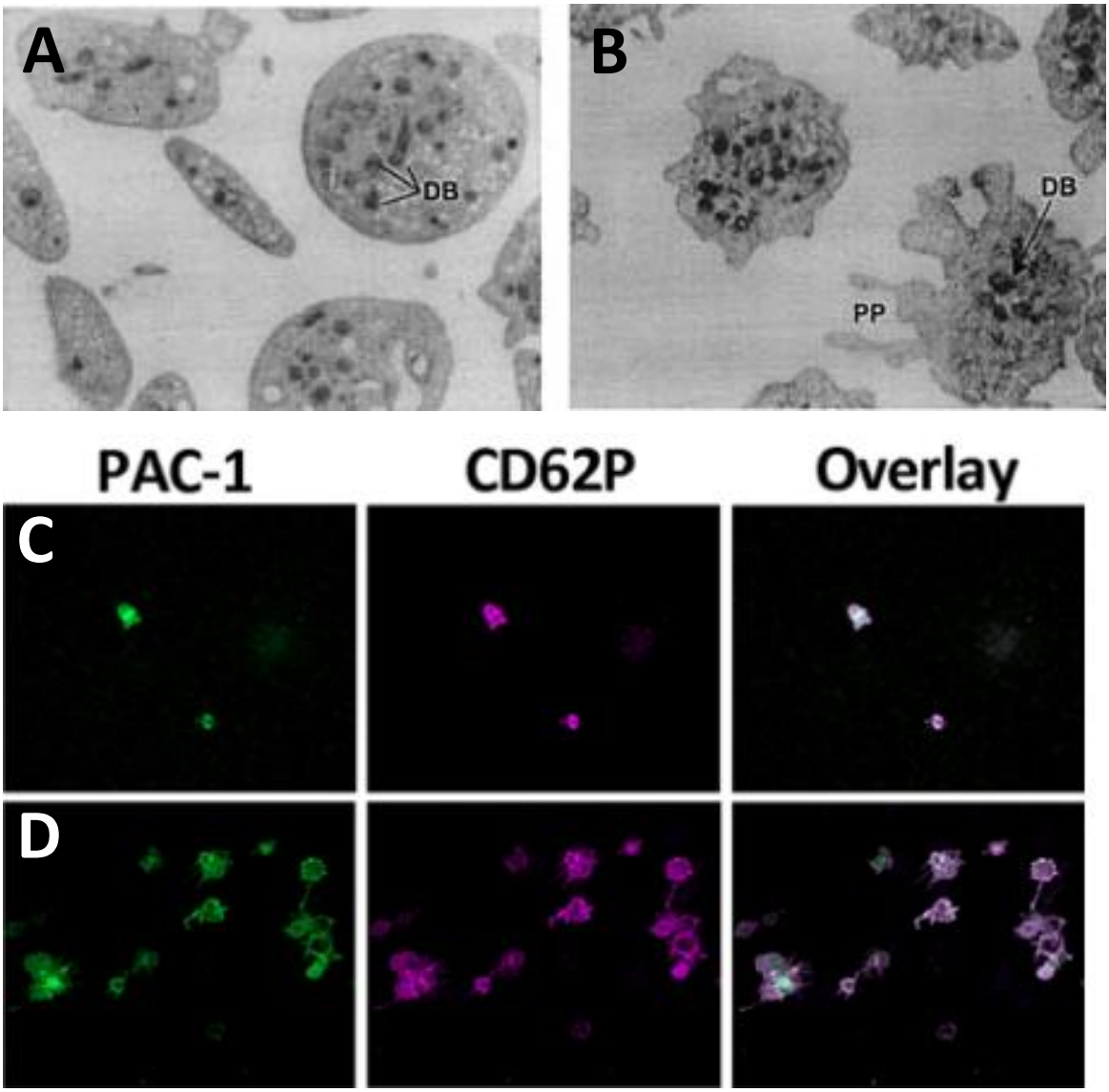


Fig. 3: The structure of platelets in healthy individuals and those found during severe COVID-19. **A -** In healthy individuals, platelets take on a rounded discoid shape prior to activation. **B** - Following activation, platelets develop pseudopodia and become flattened. **C** - Platelets stained with fluorescent antibodies that target markers of activation showing that the blood of healthy individuals possesses relatively view activated platelets. **D** - The blood of COVID-19 patients possesses a significantly elevated number of activated platelets. Additionally, the platelets possess prominent deformities such as numerous protrusions. DB – Dense bodies, also known as storage granules. PP – Pseudopodia. A and B from Jurk and Kehrel, 2005. C and D modified from Venter et al, 2020.

1.2 - The p38 MAPK Pathway

The Mitogen Activated Protein Kinase (MAPK) pathways are cellular signalling pathways involved in a variety of cellular processes such as cell proliferation, differentiation, and survival (Johnson & Lapadat, 2002). Three distinct MAPK pathways exist, based on the final integrating MAPK that the pathway ultimately activates: Extracellular signal-regulated kinase (ERK), c-Jun NH2-terminal kinase (JNK), and p38 (Johnson & Lapadat, 2002; Kim & Choi, 2010). All MAPK pathways follow a similar chain of events, where the reception of a signal leads to a phosphorylation cascade that ultimately results in the phosphorylation of certain targets resulting in a cellular response. Initially, a stimulus is received, leading to activation of the first high level MAP kinase kinase kinase (MAP3K or MAPKKK). The manner in how the stimulus is translated to MAP3K activation varies between signal type, but typically involves phosphorylation of the MAP3K by the receptor or adaptor proteins. Regardless, the now activated MAP3K phosphorylates MAP kinase kinase (MAP2K or MAPKK), the next tier in the MAPK cascade, to activate it. MAP2K then similarly phosphorylates the central MAPK to activate it. The activated MAPK then begins to phosphorylate a variety of cellular targets, such as transcription factors, in order to bring about a cellular response (Johnson & Lapadat, 2002; Kim & Choi, 2010).

The p38 pathway in particular consists of MEKK1, ASK1, and TAK1 as the MAP3K members, MKK3 and MKK6 as the MAP2K members, and p38 as the integrating MAPK (Soares-Silva et al., 2016). While four p38 isoforms have been identified; α, β, γ, and δ; the α isoform predominates and is the most thoroughly understood. While the functions and targets of each MAPK pathway are diverse, each pathway can be broadly classified according to the stimuli they respond to and the cellular response they elicit. The ERK pathways typically respond to the presence of growth factors and regulate cell division and migration. The JNK pathways respond to environmental stresses such as starvation, radiation, and oxidative stress and control the respective cellular response to these stresses. Finally, the p38 pathway responds to a number of cytokines and Pathogen-Associated Molecular Patterns (PAMPs) and is involved in a number of inflammatory responses (Johnson & Lapadat, 2002; Zhang & Liu, 2002).

The p38 MAPK pathway is heavily involved in inflammatory responses and various cellular functions. p38 activity is involved in macrophage activation (Kawasaki & Kawai, 2014), where p38 activity promotes macrophage polarisation and activation. Consequently, overactivation of the p38 pathway leads to Macrophage Activation Syndrome (MAS) that increases the risk of cardiovascular injury in severe COVID-19 (Grimes & Grimes, 2020). **Figure 4** summarises the role of p38 activity in the activation of platelets as previously mentioned (Minuz et al., 2002), as well as the role of FOXP3 as explored in the next section. Unsurprisingly, p38 activity is significantly elevated in severe COVID-19 (Manne et al., 2020). Given the role of p38 MAPK signalling in inflammation and immune responses, it is reasonable to expect pathogens to target this MAPK pathway. p38 activity has been found to be elevated in response to upregulation of the MAP3K proteins TAK1 (Moriguchi et al., 1996) and ASK1 (Ichijo et al., 1997); and thus these genes may be targeted by SARS-CoV-2 in order to upregulate p38 activity.

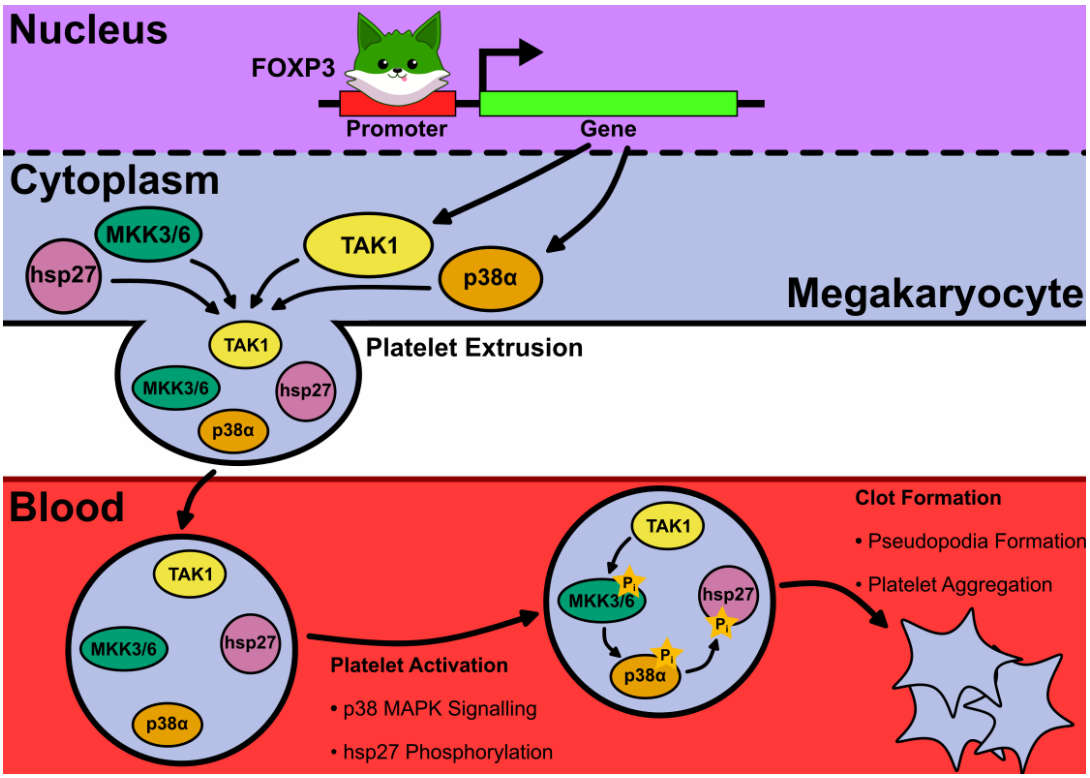


Fig. 4: The role of p38 MAPK activity in platelets. Members of the p38 MAPK signalling pathway; TAK1, MKK3/6, and p38α; are expressed in megakaryocytes and are included inside the platelet. TAK1 and p38α in particular are regulated by FOXP3 (green cartoon fox). Upon platelet activation, such as through contact with a wound or signalling by other platelets, the p38 signalling pathway undergoes a phosphorylation cascade that results in hsp27 phosphorylation. Hsp27 phosphorylation leads to the formation of pseudopodia and platelet aggregation which allows platelets to clump, seal the wound, and act as a foundation for other coagulation factors and wound healing. Figure created using Inkscape V1.2.2.

1.3 - FOXP3

1.3.1 - FOXP3 as a Transcription Factor

FOXP3 is traditionally known as the master regulator of Treg differentiation. By interacting with over 300 adaptor proteins and other transcription factors; such as RUNX1, NFAT1, EOS, and SATB1; FOXP3 is able to alter the fate of a naïve T cell; guiding it towards the Treg phenotype (Bandukwala et al., 2011; Bettelli et al., 2006; Fu et al., 2012; Ono, 2020; Sadlon et al., 2010). Tregs are immunosuppressive T cells that are critical in the prevention of autoimmune disorders. By initiating and maintaining tolerance to self-antigens, as well as tolerance to non-self but otherwise harmless antigens, Tregs prevent auto-reactive T cells from proliferating and attacking cells of the body (Bettelli et al., 2006; Dominguez-Villar & Hafler, 2018; Ono, 2020). Indeed, the presence of FOXP3 alone is capable of suppressing the cytotoxic effects of effector $CD8^+$ T cells (Hori et al., 2003; Lee & Lee, 2018; Yomogida et al., 2013). Loss of FOXP3 activity, such as through loss of function mutations, leads to the severe autoimmune disorder *immunodysregulation polyendocrinopathy enteropathy X-linked* (IPEX) *syndrome* in humans resulting in widespread inflammation and gastrointestinal problems, commonly leading to death within the first year of life (Bacchetta et al., 2018; Ono, 2020). Finally, FOXP3 has also been found to play a complex role in some cancers, where it can act as either a tumour suppressor or tumour promoter (Jia et al., 2019; Vadasz & Toubi, 2017). FOXP3 expressing Tregs often infiltrate into tumours (Gobert et al., 2009; Mizukami et al., 2008). This typically leads to tumour survival due to the immunosuppressive properties of Tregs such as in breast (Bates et al., 2006; Gobert et al., 2009),

cervical (Jordanova et al., 2008), gastric (Mizukami et al., 2008), and pancreatic cancer (Hiraoka et al., 2006). However, Treg infiltration can also lead to improved patient outcomes. For example in colorectal (Salama et al., 2009), head and neck (Badoual et al., 2006), and both B-cell and Hodgkin's lymphoma (Alvaro et al., 2005; Carreras et al., 2006). Alternatively, tumour cells themselves can express FOXP3 to varying effects (Jia et al., 2019). In breast cancer, FOXP3 expression is associated with tumour suppression through FOXP3-mediated downregulation of the oncogenes HER2 and SKP2 (Zuo, Liu, et al., 2007; Zuo, Wang, et al., 2007). In other cancers, FOXP3 expression can instead promote tumour growth. For example, in lung adenocarcinoma, FOXP3 expression leads to upregulation of cyclin D1 as well as various proteins involved in invasion and metastasis (Li et al., 2021; Li et al., 2016). The immunomodulating effects of FOXP3 make it an important player in various medical conditions; including COVID-19 (Filbin et al., 2021). However, this important transcription factor has also been found to play a role in other cells.

As mentioned previously, FOXP3 has been found to play an important role in the function of megakaryocytes (Bernard et al., 2009). Megakaryocytes are the cells responsible for the production of platelets (Italiano Jr. & Shivdasani, 2003) and thus play an important, if indirect, role in haemostasis. Loss of FOXP3 function leads to the impairment of platelet production, whereby the platelets produced by FOXP3 deficient megakaryocytes are abnormally structured and respond poorly following activation (**Fig. 5**) (Bernard et al., 2009). In IPEX patients, this manifests as thrombocytopaenia and occasionally gastrointestinal bleeding (Lyon et al., 1990; Wildin et al., 2002). FOXP3 has been found to regulate TAK1 and p38α of the p38 MAPK pathway (Lachmann et al., 2010; Rouillard et al., 2016). Given their role in platelet activation as discussed previously, it may be through these genes that FOXP3 dysregulation leads to platelet dysfunction in IPEX. Loss of FOXP3 function would lead to dysregulated p38α and TAK1 expression, leading to inappropriate hsp27 phosphorylation and subsequent actin polymerisation, ultimately resulting in the structural abnormalities seen in platelets during IPEX syndrome. Due to the similarities between platelets in IPEX syndrome (**Fig. 5**) and severe COVID-19 (**Fig. 3**) (Venter et al., 2020), the proliferation of abnormal platelets seen during COVID-19 may be the result of FOXP3 dysfunction in megakaryocytes caused by direct SARS-CoV-2 infection.

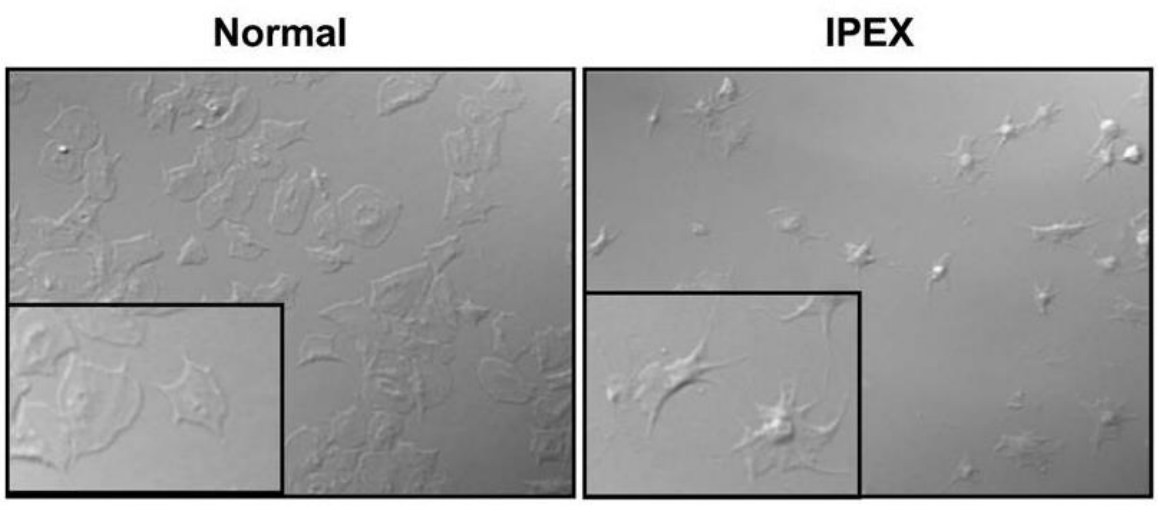


Fig. 5: Activated platelets in healthy vs IPEX syndrome patients on fibrinogen coated slides. In healthy individuals, activated platelets spread out evenly across the slide. In contrast, activated platelets from IPEX patients display a limited ability to spread. Taken from Bernard et al, 2009.

Interestingly, there is evidence that FOXP3 may be present in a small subset of macrophages. While this evidence is limited and currently controversial, the expression of FOXP3 in macrophages may lead to the development of immunosuppressive macrophages (Devaud et al., 2014; Vadasz & Toubi, 2017). It is thus interesting to envisage that similar disruption of FOXP3 by SARS-CoV-2 could lead to a population of defective $FOXP3^{+}$ immunosuppressive macrophages; contributing to the immune dysregulation seen in COVID-19.

1.3.2 - The FOX Proteins and FOXP Subfamily

FOXP3 is a member of the FOX protein family, a diverse family of transcription factors involved in development, cell proliferation, and differentiation (Lehmann et al., 2003). As transcription factors, these proteins are capable of controlling the transcription of specific genes. By binding to DNA in a sequence specific manner, they recruit or occlude the transcriptional machinery required to initiate gene transcription (Latchman, 1997). This family of proteins is highly conserved within eukaryotes, particularly in vertebrates (Hannenhalli & Kaestner, 2009). The FOX protein family consists of 20 subgroups, designated FOXA – FOXT, organised by structure (**Fig. 6**) (Laissue, 2019); with FOXT being the most recent addition to the family (Lin et al., 2021). One conserved structural element among all subgroups is the Forkhead Domain (FHD), the family namesake and the domain responsible for DNA binding. The FOXP members in particular are unique in that they possess the ability to dimerise (Chen et al., 2015). The FHD of the FOXP family was also found to form domain swapped dimers; made possible by the substitution of a proline; found in other FOX subfamilies, for an alanine in the hinge loop region of the FHD (in the case of FOXP3, Ala372) (Stroud et al., 2006). For a long time, domain swapping was thought to be important for the function of the FOXP family (Chen et al., 2015), but recent evidence suggests that domain swapping may be an artifact produced when the FHD is isolated; and indeed may be detrimental to FOXP function (Leng et al., 2022). Instead, the FOXP family may instead dimerise via the leucine zipper and the FHDs remain separate; allowing for each FHD to bind separate consensus sequences (Leng et al., 2022). Additionally, members within the FOXP subfamily are capable of forming heterodimers with each other (Li et al., 2004; Wang et al., 2003), further increasing specialisation of their functions.

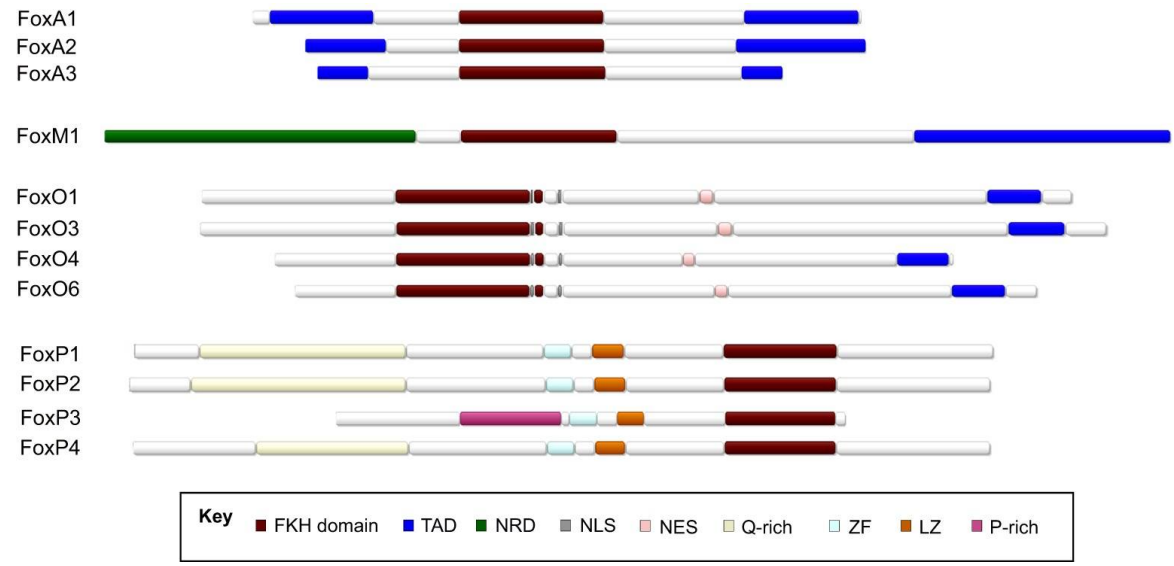


Fig. 6: Domains of FOXA, FOXM1, FOXO, and FOXP proteins showing the diversity within the FOX protein family. Despite the diversity, all FOX members possess a conserved forkhead domain responsible for DNA binding (dark red). FKH – Forkhead. TAD – Transactivating Domain. NRD – N-terminal Repressor Domain. NLS – Nuclear Localisation Signal. NES – Nuclear Export Signal. ZF – Zinc Finger. LZ – Leucine Zipper. Taken from Golson and Kaestner, 2016.

FOXP1, FOXP2, and FOXP4 are primarily involved in the development of various tissues throughout the body, with particular importance in the central nervous system (Enard et al., 2002; Lozano et al., 2021; Sin et al., 2015). FOXP3, on the other hand, is unique in that it serves as the master regulator of the Treg phenotype while also playing an important role in megakaryocytes, as previously discussed.

1.3.3 - Structure of FOXP3

FOXP3 is 47 kDa protein and possesses three domains: a leucine zipper, a zinc finger, and the C-terminal DNA binding FHD. The leucine zipper assists in protein dimerisation. Glu249, Lys250, Glu251, and Lys252 of one monomer form important hydrogen bonds with the same residues of the other monomer in order to stabilise the dimeric form (Deng et al., 2019; Kim, 2009; Li et al., 2007). The leucine zipper also plays a role in transcriptional regulation as well as association with other transcription factors (Deng et al., 2020). The function of the zinc finger is poorly understood but is thought to be similar to that of the leucine zipper (Deng et al., 2020; Kim, 2009; Song et al., 2012). The FHD is a defining feature of the FOX proteins and consists of three α-helices, three β-strands, and two loops that form the "wings". Helix 3 of the FHD inserts into the major groove of DNA (**Fig. 7**) (Bandukwala et al., 2011; Webb et al., 2017); recognising the sequence RYAAACA regardless of directionality where R = purine and Y = pyrimidine (Li et al., 2017). As mentioned previously, the FHD of the FOXP family is capable of forming domain swapped dimers in isolation. The dimeric FHD of FOXP3 is particularly stable due to a number of hydrophobic residues within the interface. This includes many residues already conserved within the FOX family, but FOXP3 possesses the additional Phe340, Leu345, Trp348, Trp366, and Met370 that further stabilise the interface (Bandukwala et al., 2011). Phe373 also plays a particularly important role in the stability of the domain swapped dimer (Perumal et al., 2015). As a result, the FHD of FOXP3 is the only FOXP FHD proposed to exist solely as a dimer under physiological conditions (Bandukwala et al., 2011); distinguishing FOXP3 among the FOXP subfamily. Due to the dimeric nature of the FHD, it is capable of binding two anti-parallel DNA strands simultaneously which allows two distant genetic loci to be brought together in close physical proximity; leading to chromatin remodelling (Chen et al., 2015; Zhang et al., 2023).

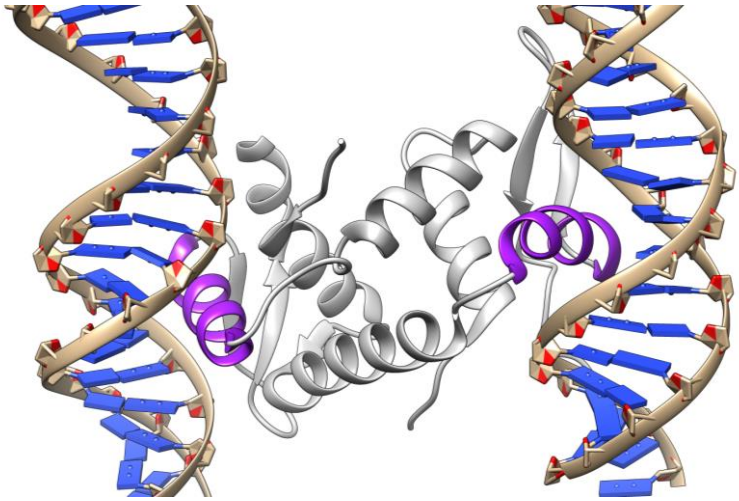


Fig. 7: The domain swapped dimeric FOXP3 FHD bound to anti-parallel DNA strands. The DNA binding helices, coloured purple, insert into the major groves of DNA to promote sequence recognition (Bandukwala et al, 2011). PDB: 3QRF. Generated using UCSF Chimera (Pettersen et al., 2004).

1.4 - The SARS-CoV-2 Nucleocapsid

The SARS-CoV-2 virion consists of four structural proteins: the spike (S), envelope (E), membrane (M), and nucleocapsid (NC) proteins (**Fig. 8A**) (Wu et al., 2020). While the S, E, and M proteins are primarily involved in the outer shell and membrane of the virus, NC (**Fig. 8B**) is involved in the encapsulation of the viral RNA genome and assists in packaging and stabilisation (Bai et al., 2021). However, the NC is multifunctional and serves additional roles. As the most prevalent viral protein within an infected cell, it assists in virion construction through interaction with the M protein (Lu et al., 2021), influences expression of viral proteins (Yang et al., 2021), and interferes with the cell's innate anti-viral responses (Lu et al., 2011; Mu et al., 2020; Zhou et al., 2008). Additionally, the NC has been found to directly contribute to the hyperinflammatory state in severe COVID-19 through the induction of Neutrophil Extracellular Trap (NET) formation (Youn et al., 2021); an inflammatory structure released by neutrophils intended to ensnare extracellular pathogens. Due to this multifunctionality, understanding the NC is critical to understanding how it is capable of subverting normal cellular processes and contributing to the hyperinflammatory state in severe COVID-19.

The SARS-CoV-2 NC is a 46 kDa protein and shares high sequence similarity to the NC of other coronaviruses (Peng et al., 2020). While the protein is primarily found as a dimer, it has been found to form higher order complexes in a concentration dependent manner (Bai et al., 2021; Lutomski et al., 2021). These higher-order complexes may be particularly relevant once bound to RNA; where local protein concentrations are elevated (Chen et al., 2007). The NC possesses two RNA binding domains linked by a largely disordered linker region (**Fig. 9A**) (Cubuk et al., 2021; Peng et al., 2020). While the structure of the full-length protein has not been solved, the structures of the individual domains have been solved. The N-terminal RNA binding domain (NTD), shown in **figure 9B**, consists of amino acids 44 – 174. The NTD resembles a fist-like structure with a prominent β-hairpin consisting of residues 90 – 107. The NTD binds RNA in the palm-like cleft; where Tyr86, Tyr109, and Tyr111 form important π-π stacking interactions and hydrogen bonds with four consecutive RNA bases. Several arginines in the β-hairpin form electrostatic interactions with the RNA backbone to further enhance RNA binding (Chang et al., 2016; Dinesh et al., 2020; Grossoehme et al., 2009; Korn et al., 2023). The C-terminal RNA binding domain (CTD) (**Fig. 9C**) consists of amino acids 255 – 364. While also participating in RNA binding, the CTD facilitates NC dimerisation. As a dimer, the CTD appears as a rectangular slab with a four-stranded β-sheet in the centre flanked by α-helices. The β-sheet is constructed when the β-hairpins from each monomer insert into the cavities of the opposite monomer, forming a continuous sheet. RNA binding is mediated by electrostatic interactions with residues Lys256, Lys257, Lys261, and Arg262 which together form a positively charged groove that runs along the back of the slab (Bai et al., 2021; Chen et al., 2007; Dinesh et al., 2020). While the CTD typically functions as a dimer when free in solution, when packing the viral genome in the cramped space of the virion, the CTD is thought to take on a helical octamer assembly which the viral RNA then wraps around (**Fig. 9D**). The phosphate backbone of the RNA is thought to become buried in the groove of the CTD. The NTD, enabled by the flexible linker region, then binds to the exposed bases (Chen et al., 2007). This packaging of the viral genome between the two NC domains both efficiently packages the RNA and protects it from degradation. It has also been found that the SARS-CoV-2 NC, like its SARS-CoV-1 predecessor, is capable of binding both single stranded and double stranded DNA in addition to RNA (Zeng et al., 2020; Zhou et al., 2020); albeit with lower affinity. The disordered regions of NC provide the protein with a high degree of flexibility, play a role in RNA binding (Wu et al., 2021), and allow the protein to perform the multitude of functions previously mentioned. Due to its high degree of flexibility, NC is considered an intrinsically disordered protein (IDP). IDPs are proteins possessing a large proportion of disordered regions with little to no permanent structure. This provides IDPs a high level of flexibility, allowing them to participate in a wide variety of typically low-affinity but high-specificity binding

events with numerous different binding partners (Uversky, 2019; Ward et al., 2004). The promiscuity of IDPs allows a single protein to participate in or perform multiple functions. This makes IDPs particularly common in viruses, as it allows for compacting of the viral genome (Xue et al., 2012). The disordered regions of NC in particular play roles in efficient packing during viral assembly, oligomerisation, RNA binding, RNA chaperoning, and template switching during replication (Chang et al., 2016; Chen et al., 2007; Wu et al., 2021).

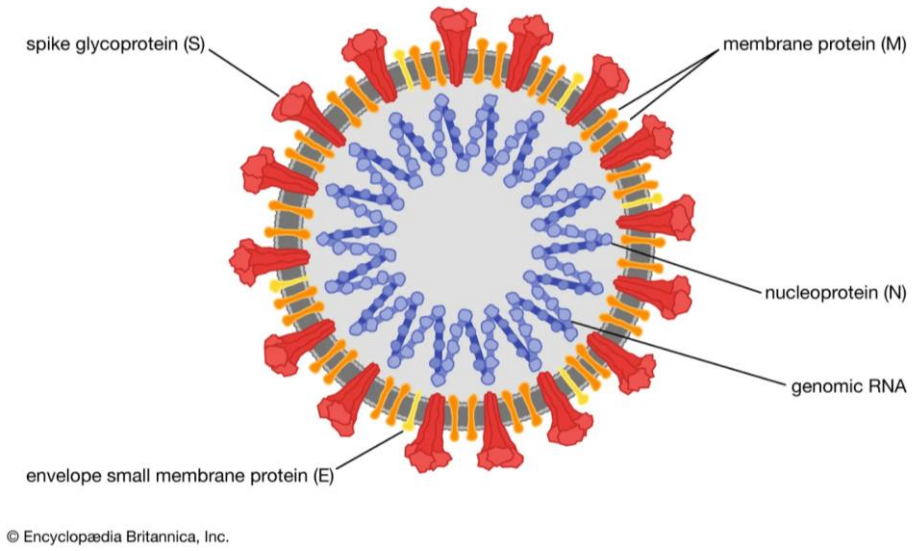


Fig. 8: Model of the SARS-CoV-2 viral particle and nucleocapsid. **A** – The virus consists of spike (S), envelope (E), membrane (M), and nucleocapsid (N) proteins. The S, E, and M proteins make up the viral coat which contains the viral genome and facilitates cellular entry. N encapsulates the viral RNA genome to allow packaging within the shell. Taken from Encyclopaedia Britannica at https://www.britannica.com/science/2019-nCov#/media/1/2158118/245650. **B** – AlphaFold prediction of the dimeric SARS-CoV-2 NC in the absence of nucleic acid, with the NTD coloured pink and CTD coloured red. Structure generated using the nucleocapsid sequence using AlphaFold V2.3.1+4 (available at: https://console.latch.bio/workflows/87049/parameters) and visualised using UCSF Chimera (Pettersen et al., 2004).

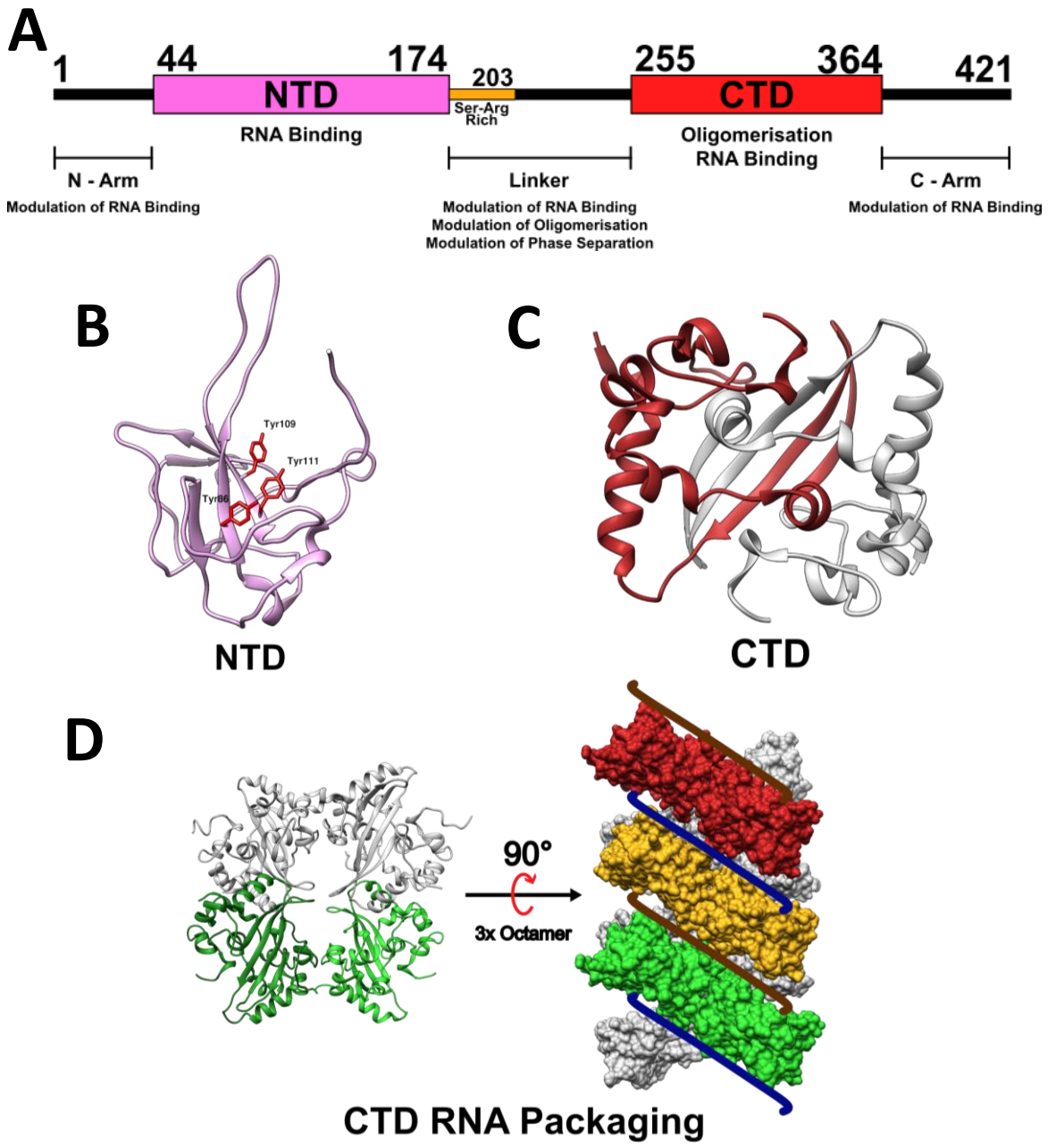


Fig. 9: Structure of the SARS-CoV-2 Nucleocapsid (NC). A – Schematic of the NC monomer, consisting of two RNA binding domains (NTD and CTD) interspersed between three disordered regions. The Ser-Arg rich region plays a role in dimerisation and is crucial for viral replication. **B** – NMR structure of the NTD (PDB: 6YI3). The NTD functions as a monomer and possesses a fist-like structure that binds RNA in the cleft visible below and to the right of the uppermost loop; analogous to the palm of the fist. The tyrosines inside the cleft, coloured red, intercalate with RNA bases to stabilise RNA binding. **C** – Crystal structure of the dimeric CTD, with each monomer coloured differently (PDB: 7CE0). The CTD domain both binds RNA and facilitates dimerisation. **D** – Proposed model of CTD nucleic acid packaging (PDB: 2CJR). Four CTD dimers form a larger CTD octamer (left) depicted as ribbons. Two dimers are coloured differently to assist visualisation. CTD octamers then stack to form a helical spindle (right) depicted as a space filling model. The corresponding green coloured dimers remain coloured green, and equivalent dimers in additional stacked octamers are coloured orange or red. The helix contains two grooves where two different RNA strands bind (blue and brown lines). CTD helix prepared by generating crystal symmetry-related copies of 2CJR. Structures visualised using UCSF Chimera (Pettersen et al., 2004).

Interestingly, the NC of SARS-CoV-1 has been found to bind to host promoter sequences to alter the transcription of host proteins directly (Yan et al., 2006). While this has yet to be shown for SARS-CoV-2, given the similarity between these proteins (Peng et al., 2020), it is likely that this ability is conserved in the SARS-CoV-2 NC. If this is indeed true, this may present a possible mechanism behind the elevated p38 pathway activity and potential FOXP3 inhibition seen in COVID-19. SARS-CoV-2 is capable of infecting megakaryocytes (Shen et al., 2021), albeit through an ACE2 independent mechanism. Upon infection, the NC may be binding to gene promoters of p38 pathway proteins, upregulating their expression while also occluding FOXP3 binding sites. Alternatively, NC has been found to interact with several human proteins such as MOV10, LARP1, and casein kinase 2, among others (Gordon et al., 2020; Wang et al., 2021). Due to this promiscuity, NC may alternatively bind to FOXP3 directly to cause dysregulation of its function.

1.5 - Rationale

Severe COVID-19 is characterised by immune system overactivation and DIC (Jamal et al., 2021; Kander, 2020; Levi et al., 2020; Merad & Martin, 2020). In the short-term, DIC can lead to life threatening strokes and embolisms. In the long-term, DIC may lead to the formation of microclots which cause capillary blockage and micro-ischaemic events. These events may contribute to long COVID; the prolonged period of symptoms following the conclusion of infection (Pretorius et al., 2021). A major facet of the immune system overactivation is the profound activation of macrophages (Giamarellos-Bourboulis et al., 2020) which is itself driven by overstimulation of the p38 MAPK signalling pathway (Grimes & Grimes, 2020). The hypercoagulable state is partially driven by the existence of overactivated and structurally deformed platelets in the plasma of patients with COVID-19 (Venter et al., 2020). Within these malformed platelets, p38 activity is similarly elevated (Manne et al., 2020), which promotes their activity and contributes to the excessive formation of clots during COVID-19. However, the mechanism behind the formation of these abnormal platelets is currently unknown. Understanding the mechanisms behind COVID-19 coagulopathies may provide opportunities for new treatments and therapies targeting the excessive formation of blood clots. One protein whose dysfunction may be related to the platelet dysfunction is FOXP3.

FOXP3 is a transcription factor traditionally associated with the Treg phenotype (Bandukwala et al., 2011; Ono, 2020) and has been found to be critically important to the immunosuppressive properties of these cells (Hori et al., 2003; Lee & Lee, 2018; Yomogida et al., 2013). Additionally, FOXP3 has been found to be important in the development of megakaryocytes and the production of platelets. Loss of FOXP3 function leads to generation of malformed and functionally aberrant platelets like those seen in severe COVID-19 (Bernard et al., 2009). Additionally, it has been found that SARS-CoV-2 is capable of infecting megakaryocytes (Shen et al., 2021), suggesting that SARS-CoV-2 may be infecting megakaryocytes and subsequently interfering with FOXP3 activity; leading to the formation of these abnormal platelets.

Given the structural and functional similarities between platelets produced by FOXP3 deficient megakaryocytes and the platelets seen in severe COVID-19 patients, it is possible that SARS-COV-2 is interfering with FOXP3 activity. Additionally, due to the importance of the p38 MAPK pathway in platelet activation, the inhibition of FOXP3 activity may manifest itself as upregulation of p38 activity. The NC of SARS-CoV-2, due to its multifunctional properties, is the most likely SARS-CoV-2 protein capable of causing this interference. The NC has been found to be capable of binding DNA (Zhou et al., 2020). Thus, while it has not been shown for the SARS-CoV-2 NC, the protein may share the ability of the highly similar SARS-CoV-1 NC to bind directly to host gene promoters and alter gene expression

(Yan et al., 2006). If this ability is conserved, the SARS-CoV-2 NC may thus be binding directly to host gene promoters which may explain the myriad of symptoms observed during COVID-19. FOXP3 has been found to regulate TAK1 and p38α (Lachmann et al., 2010; Rouillard et al., 2016). Given that upregulation of the MAP3K proteins, such as TAK1, leads to increased p38 MAPK pathway activity (Ichijo et al., 1997; Moriguchi et al., 1996), disruption of FOXP3, such as through NC DNA binding, may lead to dysregulation of this pathway and dysregulation of platelet development. By investigating the binding of FOXP3 to a short sequence in the promoter region of p38α, this project aims to determine whether the presence of SARS-CoV-2 NC alters FOXP3 DNA binding. Understanding possible interactions between FOXP3 and the SARS-CoV-2 NC will contribute to the understanding of SARS-CoV-2 pathophysiology. Additionally, this interaction may present as a potential treatment target in combatting the coagulopathy and, potentially, the immune hyperactivation of severe COVID-19.

2 – Aims and Objectives

2.1 – Aim

Determine whether the SARS-CoV-2 nucleocapsid (NC) alters the DNA binding ability of the FOXP3 FHD and whether these two proteins interact directly.

2.2 – Objectives

- Overexpress both the FOXP3 FHD and SARS-CoV-2 NC from transformed T7 *E. coli* cells and purify both proteins using immobilised metal affinity chromatography.
- Structurally characterise the secondary, tertiary, and quaternary structures of both proteins using circular dichroism spectroscopy, intrinsic tryptophan fluorescence, absorbance spectroscopy, and size exclusion chromatography.
- Assess the DNA binding ability of the FOXP3 FHD in the presence and absence of NC using electrophoretic mobility shift assays (EMSAs) and fluorescence anisotropy.
- Assess whether the NC interacts directly with the FOXP3 FHD using fluorescence anisotropy and pull-down assays.

3 - Methods and Materials

Unless otherwise stated, all experiments were performed at 20 °C and all pH values are given at 20 °C. Percentage composition values refer to w/v percentage, except for liquid/liquid solutions where v/v percentage is explicitly indicated. All reagents were of analytical grade and the majority of buffer components were purchased from Sigma-Aldrich, USA. Manufacturers of other reagents and kits are given as needed. All solutions were prepared using distilled water and filtered prior to use. All figures were created with Inkscape 1.2.2 or Microsoft PowerPoint. Protein structures were visualised using UCSF Chimera V1.14 (Pettersen et al., 2004) and PDB codes are given in respective figure legends. Initial data management was performed using Microsoft Excel and then plotted using SigmaPlot V12.2.0.45. Chemical structures were drawn using Chemdraw 19.0.0 (available at: https://chemdrawdirect.perkinelmer.cloud/js/sample/index.html).

3.1 - DNA Oligonucleotides

Both FOXP3 and NC are nucleic acid binding proteins. While NC is predominantly an RNA binding protein, it has been shown to bind DNA (Zeng et al., 2020; Zhou et al., 2020). To study protein-DNA interactions, short DNA oligonucleotides were used. FOXP3 has been found to regulate p38α (Lachmann et al., 2010; Rouillard et al., 2016) and so a sequence from the promoter of the p38α gene was used due to its role in platelet activation (Minuz et al., 2002; Saklatvala et al., 1996). The promoter sequence was accessed using the Ensembl genome browser database and scanned for FOXP3 consensus binding sites using the JASPAR database (available at https://jaspar.genereg.net/). The top result according to the scoring provided by JASPAR was used. Additional DNA sequences that do not contain the FOXP3 consensus were used to ensure specificity of FOXP3 to target sequences as well as to test specificity of NC for other DNA sequences. **Table 1** below shows the sequences of the oligonucleotides used in this work. Oligonucleotides were purchased from Integrated DNA technologies – Whitehead Scientific, South Africa.

Table 1: DNA oligonucleotide sequences used in this work. A region from the p38α gene promoter sequence was used as a FOXP3 specific sequences, while SSL and AAV P5 were used to test for specificity of FOXP3 and NC to other sequences that do not contain the FOXP3 consensus sequence. FOXP3 consensus binding site is bold and underlined. While only single stranded sequences are shown, all sequences used were double stranded oligonucleotides.

Sequence Name	Sequence (5' → 3')	Protein Specificity	Used for	Ensembl Accession
p38α	GTTATAA**TTTGTT**TATTGTCAATTG	FOXP3	EMSA Fluorescence anisotropy	ENSG00000112062
Single Site Long (SSL)	CCCAATTTCACACCTTCCTCA	TBR1	EMSA	ENSG00000158321
AAV P5	AGGGTCTCCATTTTGAAGCG	YY1	EMSA	N/A

3.2 - Protein Sequences and Purification

3.2.1 - Protein Constructs

To study the potential interaction between the FHD of FOXP3 and the SARS-CoV-2 NC, constructs of each protein were made and inserted into a vector for overexpression and subsequent purification. The isolated FOXP3 FHD was chosen due to the importance of this domain in FOXP3 function, while full-length NC was used. Both proteins possessed an N-terminal hexahistidine (His-tag) linked to the protein via a thrombin cleavage site (LVPRGS) (**Table 2**. Full protein sequences are available in **Fig. S1** and **S2**). Additionally, the FOXP3 FHD construct possessed extra residues in this linker region to improve solubility during expression.

Table 2: Protein constructs used for this study. The N-terminal regions containing the His-tag (red) and thrombin cleavage sites (purple) are shown, with the first few N-terminal residues of each respective protein shown in green. The FOXP3 FHD construct contains additional residues in the N-terminal region to assist with solubility.

Construct	Region of full-length protein	Sequence	Size (kDa)
FOXP3 FHD	336 - 417	MASMTGGQQMGM**HHHHHH**SSG**LVPRGS**MRPP…**FOXP3 FHD**	12.9 – Uncleaved 10 – Cleaved
SARS-CoV-2 NC	1 - 419	MGG**HHHHHH**SSG**LVPRGS**MSDN…**NC**	47.5 – Uncleaved 45.7 – Cleaved

Sequences coding for these proteins were codon optimised and directionally inserted into separate pET-11a plasmids using NdeI and BamHI restriction sites (**Fig. 10**) (Genscript, USA). As a selection marker, the pET-11a plasmid encodes ampicillin resistance, allowing for the selection of transformed cells through the inclusion of ampicillin in the growth medium; which otherwise inhibits the growth of untransformed cells. The protein coding sequences are under the control of a T7 promoter, which allows for induction of protein expression as explained in the next section.

3.2.2 - Transformation

For protein expression, T7 *E. coli* cells (New England Biolabs, USA) were transformed with the plasmids encoding the proteins of interest. Due to the well characterised nature of *E. coli* biology, the organism's rapid growth rate, and ease of manipulation, these bacteria serve as convenient hosts for protein expression (Rosano & Ceccarelli, 2014). Within the T7 *E. coli* genome lies a T7 RNA polymerase gene under control of a *lac* promoter. Due to the presence of an additional *lac* repressor, expression of the polymerase is prevented. By acting as a non-hydrolysable allolactose mimic, Isopropyl β-d-1-thiogalactopyranoside (IPTG) is able to suppress the effect of the *lac* repressor, thus allowing expression of the T7 polymerase. The T7 polymerase subsequently recognises and binds to the T7 promoter controlling the expression of the target protein in the plasmid, such as pET-11a, promoting

high levels of heterologous protein expression (Dubendorff & Studier, 1991) (**Fig. 11**). Due to IPTG being non-hydrolysable, IPTG need only be added once to the medium to induce continuous expression.

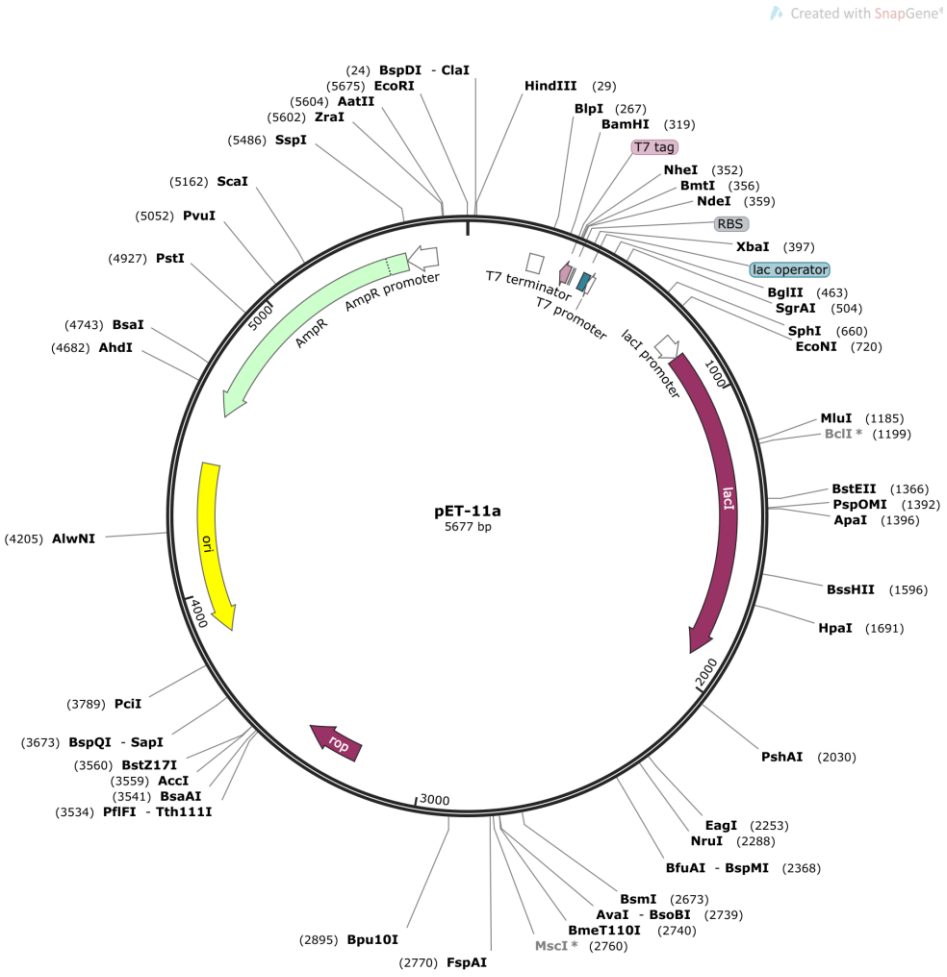


Fig. 10: Schematic of the pET-11a plasmid used to encode the gene coding for the FOXP3 FHD or SARS-CoV-2 NC. Both proteins are directionally inserted using NdeI (top right, position 359) and BamHI (top, position 319). The pET-11a plasmid encodes an ampicillin resistance gene that allows for selection of transformed cells (AmpR – top left).

Cells were transformed using the heat-shock method. Competent T7 cells were thawed on ice and 100 ng of plasmid was added. The cells were then incubated on ice for 30 minutes. The cells were subsequently heated to 42 °C for 45 seconds, and then immediately placed on ice for 5 minutes. Following this, 950 µl SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) was added to the cells and the resulting mixture incubated at 37 °C for 1 hour. Cells were then plated on 2xYT agar plates (1.6% tryptone, 1% yeast extract, 0.5% NaCl, 1.5% agarose) containing 100 µg/ml ampicillin. As described earlier, the pET-11a plasmid contains an ampicillin resistance gene, allowing for growth of transformed cells on selective media. Untransformed cells and sterile SOC medium were also plated separately as controls. Plates were incubated at 37 °C overnight. The following day, a single colony was picked and inoculated into 100 ml of 2xYT liquid medium with 100 µg/ml ampicillin and again incubated at 37 °C overnight. The

resulting culture was either used to prepare glycerol stocks or used immediately for protein expression. For the preparation of glycerol stocks, a 200 µl 50:50 mix of culture and glycerol was prepared and stored at -80 °C for future expressions.

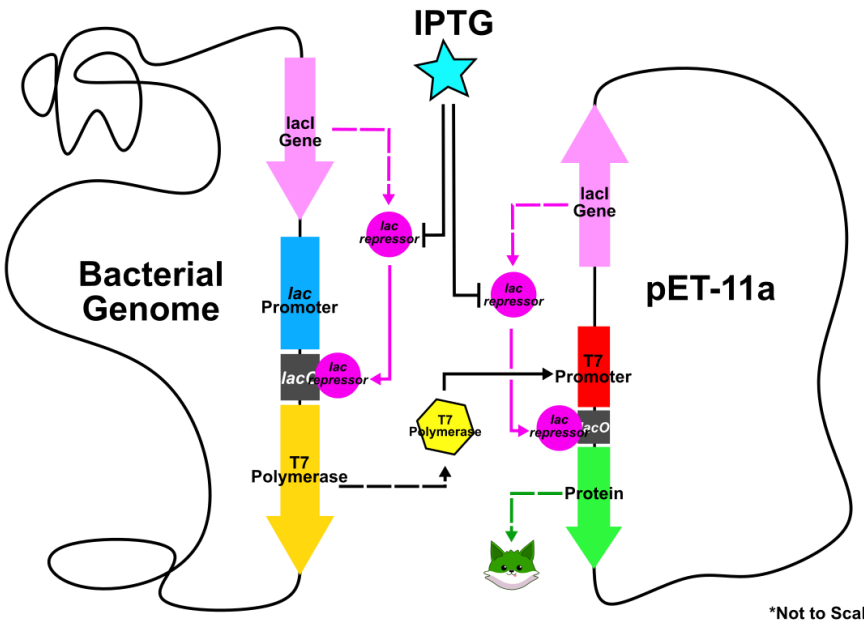


Fig. 11: The T7 expression system. Expression of the protein of interest (green, lower right) is repressed by the action of the *lac* repressor (pink circle) binding to the *lacO* regulatory sequence (grey, middle); encoded for by both the bacterial genome and plasmid (light pink arrow, top). Similarly, a T7 polymerase gene (yellow, lower left) in the bacterial genome is repressed by the same mechanism. When IPTG (light blue star, top) is added to the cell culture, the *lac* repressor dissociates from the *lacO* sequence, allowing for expression of the T7 polymerase (yellow hexagon, middle) which then readily binds to the T7 promoter (red, middle right) and transcribes the protein of interest. Dashed arrows – *produces*. Solid arrows – *binds to*. Image created using Inkscape V1.2.2.

Plasmid was extracted from cells of the overnight culture using a GeneJET Plasmid Miniprep kit (ThermoFisher, USA) and sent for sequencing (Inqaba Biotec, South Africa) to confirm plasmid integrity. Both plasmids were found to be fully intact and to code for their respective proteins without error (**Fig. S3** and **S4**).

3.2.3 - Expression

To obtain sufficient protein for subsequent studies, larger cultures of transformed T7 *E. coli* were used. The expression of both proteins had been optimised as part of previous works within our laboratory, and consequently expression trials were not performed. Both proteins were expressed using the same overarching protocol. A 100 ml liquid 2xYT culture with 100 µg/ml ampicillin overnight culture was prepared from glycerol stocks. The following day, the overnight culture was diluted 1:50 into 600 ml cultures of the same medium and incubated at 37 °C with shaking at 200 RPM. Periodically, the optical density (OD) of the culture at 600 nm was monitored using an SM1100 UV-Vis spectrophotometer

(Azzota Corporation, USA) until a desired value was reached. Once reached, IPTG was added to induce protein expression. For the FOXP3 FHD, it was found that the protein displayed significant leaky expression and the addition of IPTG did not appear to have a significant impact on final protein yields. Some human proteins have been found to induce their own expression from pET vectors in the T7 *E. coli* expression system (Briand et al., 2016). The FOXP3 FHD thus may be inducing its own expression similarly to when IPTG was added. Therefore, FOXP3 FHD was induced simply by lowering the incubation temperature to 30 °C once an OD of 1.0 was reached and incubating for a further 6 hours. For NC, once an OD of 0.6 was reached, induction was initiated by addition of IPTG to a final concentration of 0.5 mM. Flasks were then incubated at 37 °C for a further 6 hours at 180 RPM. Following incubation, cells were harvested through centrifugation at 5000 xg for 10 minutes at 4 °C. After discarding the medium, cells were resuspended in equilibration buffer (20 mM tris-HCl, 30 mM imidazole, 500 mM NaCl, pH 7.5) to a final volume of approximately 50 ml. The suspensions were stored at -20 °C until needed. Freezing of the cells both helps prevent protein degradation as well as assists in lysing of the cells through the formation of ice crystals that disrupt the cell membrane.

Prior to purification, cells were thawed at 20 °C and phenylmethylsulfonyl fluoride (PMSF) and lysozyme were added to a final concentration of 1 mM and 0.1 mg/ml, respectively. Additionally, 0.05 mg of DNase I was also added. As a serine protease inhibitor, PMSF acts to prevent protein degradation resulting from the release of proteases during the lysis process (James, 1978). Lysozyme is an enzyme that hydrolyses the β-1-4 glycosidic bonds found within the cell wall; weakening the cell's integrity and assisting lysis (Lesnierowski & Kijowski, 2007). DNase I degrades bacterial DNA for two purposes. First, degradation of DNA lessens the viscosity of the lysate, making handling and filtration easier. Second, it prevents DNA contamination; important in the purification of DNA-binding proteins (Deutscher & Burgess, 2009). For purification of NC, 0.05 mg of RNase was also added due to NC's high affinity for RNA. After addition of the above, the lysate was incubated for 30 minutes at 20 °C before sonication. Cells were lysed via 5 rounds of 30 seconds of sonication while on ice. Following sonication, the lysate was centrifuged at 18,000 xg for 25 minutes at 4 °C to separate soluble and insoluble fractions. The supernatant was used for further purification as both proteins were found to be predominantly in the soluble fraction.

3.3 - Purification

3.3.1 - Immobilised Metal Affinity Chromatography

Immobilised Metal Affinity Chromatography (IMAC) is a chromatography technique that exploits the varying affinities that different proteins have for metal ions in order to separate them. In IMAC, metal ions; commonly Ni^{2+}, Co^{2+}, or Zn^{2+}; are immobilised to a support resin via a nitrilotriacetic acid (NTA) group. As a sample of proteins is passed over this resin, proteins with an affinity for the immobilised metal become bound to the column while proteins with little to no metal affinity are eluted in the flowthrough. These bound proteins can then be subsequently eluted for collection using a high concentration of imidazole or a decrease in pH (Sulkowski, 1985). As mentioned previously, both proteins have a His-tag at their N-terminus. Histidine, through the formation of dative covalent bonds, is capable of forming complexes with metal ions. Due to this, introduction of additional histidine residues to a heterologous protein is a common technique to facilitate purification via IMAC. When passed through the metal charged IMAC column; the His-tag binds the heterologous protein to the column while other unwanted proteins elute in the flowthrough (**Fig. 12**) (Block et al., 2009; Bornhorst & Falke, 2000).

Each protein was purified on its own nickel-charged 5 ml IMAC column (HiTrap IMAC Sepharose Fast Flow, Cytiva, USA) using either an ÄKTA start or ÄKTA Prime Plus Fast Performance Liquid Chromatography (FPLC) system. Prior to purification, the column was equilibrated with 5 column volumes of equilibration buffer. The expression supernatant was filtered using a 0.45 µm filter before being passed through the column. The flowthrough was collected and the column re-equilibrated as before. The column was then washed using a high salt concentration wash (20 mM tris-HCl, 30 mM imidazole, 1.5 M NaCl, pH 7.5) for 5 column volumes. The salt wash helps remove any residual nucleic acid contamination from the lysate that may be non-specifically bound to the positively charged nickel or bound by the protein itself. It also helps remove any weakly bound proteins that have also bound the column non-specifically. In the case of FOXP3 FHD, this wash was sufficient to ensure purity from nucleic acids. For NC, which has a high affinity for RNA, the system was repeatedly paused and unpaused, allowing the salt wash to incubate on the column in approximately 1-minute intervals, as simply letting the buffer pass through was insufficient to dislodge bound RNA. This was performed until contaminants stopped eluting. Without this procedure, significant nucleic acid contamination was present in the eluted NC fractions. Following the wash, the column was again re-equilibrated. To elute the proteins, elution buffer (20 mM tris-HCl, 300 mM imidazole, 500 mM NaCl, pH 7.5) was passed through the column. Various points of the resulting elution peak were collected in fractions. These fractions, in addition to samples of the lysate supernatant, pellet, and flowthrough, were resolved using SDS-PAGE to assess protein purity; as described in **Section 3.3.4**. Proteins were stored in elution buffer at 4 °C until required.

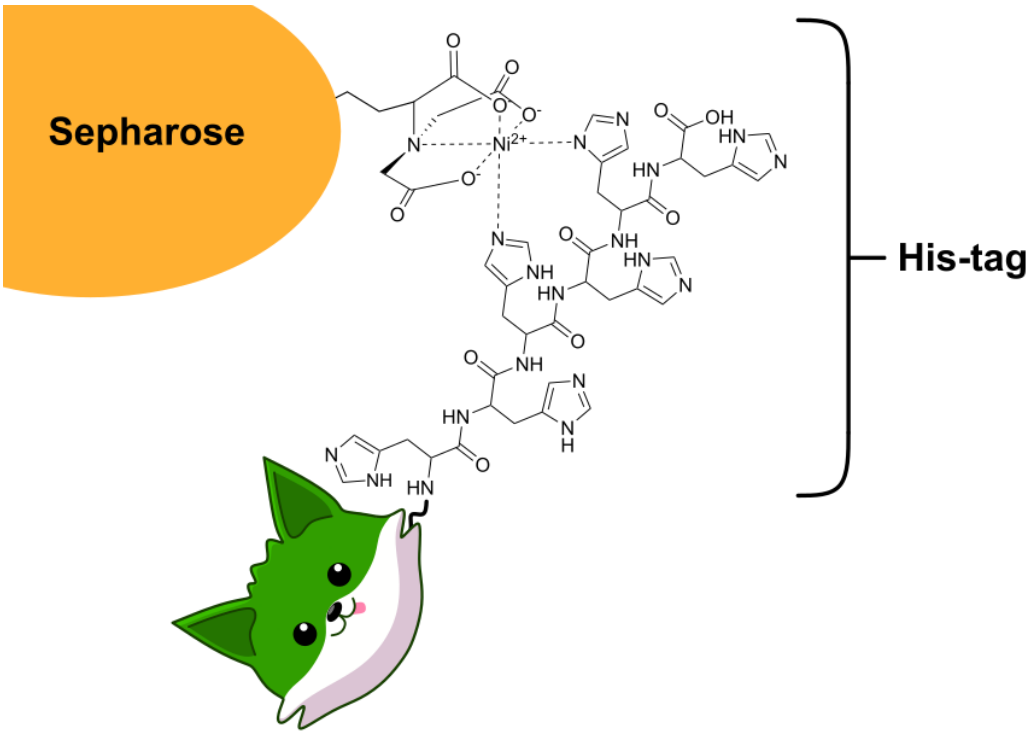


Fig. 12: Concept of IMAC Purification, with FOXP3 FHD (green cartoon fox) bound to the nickel charged IMAC column. The IMAC column consists of an NTA group chemically linked to a Sepharose matrix. The NTA group, a metal chelator, can be complexed with a metal that then acts to bind proteins of interest. When a protein mixture is passed through the column, proteins with high metal affinity; such as those with a His-tag; bind to the metal while other proteins elute out. The protein of interest can then be eluted when desired through the addition of high concentration imidazole or a decrease in pH. Image created using Inkscape V1.2.2. Chemical structures drawn using Chemdraw 19.0.0.

3.3.2 - Thrombin Cleavage

Removal of the His-tag from a protein ensures that any observed behaviour is that of the protein of interest and not that of the His-tag. Due to the inclusion of a thrombin cleavage site between the His-tag and protein, the His-tag can be removed by incubation with thrombin. Only NC underwent thrombin cleavage. Thrombin cleavage of FOXP3 FHD leads to near complete loss of protein to aggregation; possibly due to loss of the solubility enhancing residues in the N-terminal linker. However, work within our lab has found that the His-tag does not adversely affect the function of the FHD. Additionally, having one protein with a His-tag and another without is beneficial for downstream applications, as will be discussed in relevant sections. Thus, the His-tag of the FOXP3 FHD was not removed. For NC cleavage, IMAC fractions were pooled and dialysed against thrombin cleavage buffer (20 mM tris-HCl, 2 mM KCl, 500 mM NaCl, pH 8) overnight using 10K MWCO SnakeSkin dialysis tubing (ThermoFisher, USA). Following dialysis, 1 unit of human thrombin was added per 1 mg of NC. The first time thrombin cleavage was performed, hourly samples were taken and resolved using SDS-PAGE to track cleavage progress. Subsequent thrombin cleavage was performed for 6 hours. After incubation, the mixture was centrifuged at 12,000 xg for 15 minutes to remove aggregates and purified using a second round of IMAC. In addition to the 5 ml IMAC column, a 1 ml benzamidine column (HiTrap Benamadine Fast Flow, Cytiva, USA) was attached in series before the IMAC column. Both columns were equilibrated separately using 5 column volumes of equilibration buffer before being connected in series. Benzamidine, a competitive inhibitor of serine proteases such as thrombin, binds the thrombin (Tanizawa et al., 1971); immobilising it and allowing the protein of interest to elute freely. The IMAC column binds the now free His-tag and any undigested protein, while the cleaved protein of interest no longer binds and elutes from the column. Thrombin cleavage has the added benefit of removing contaminating proteins as the unwanted proteins also bind to the IMAC column like they did during the first round of IMAC purification. NC was found to elute with other contaminating proteins after the first round of IMAC, and thrombin cleavage succeeded in removing these contaminants as evidenced by SDS-PAGE analysis. Free His-tag, undigested protein, and contaminating proteins were eluted from the IMAC column as before to restore the column.

3.3.3 - Size Exclusion Chromatography

Size Exclusion Chromatography (SEC) was used both analytically and for purification. As a chromatography technique, SEC separates proteins based on their hydrodynamic volume. Within the SEC column exists beads, typically dextran or polyacrylamide, that act as the stationary phase. These beads possess pores of varying sizes which, when the sample is passed through the column, trap smaller proteins. Larger proteins are physically excluded from the beads. This results in the separation of proteins based on their hydrodynamic volume; with larger molecules eluting first and smaller molecules eluting last (**Fig. 13**). Given that a protein is suitable for the given pore size, the elution volume of a particular globular protein is inversely proportional to the log of the molecular weight (Barth et al., 1994; Hong et al., 2012; Wu, 2003).

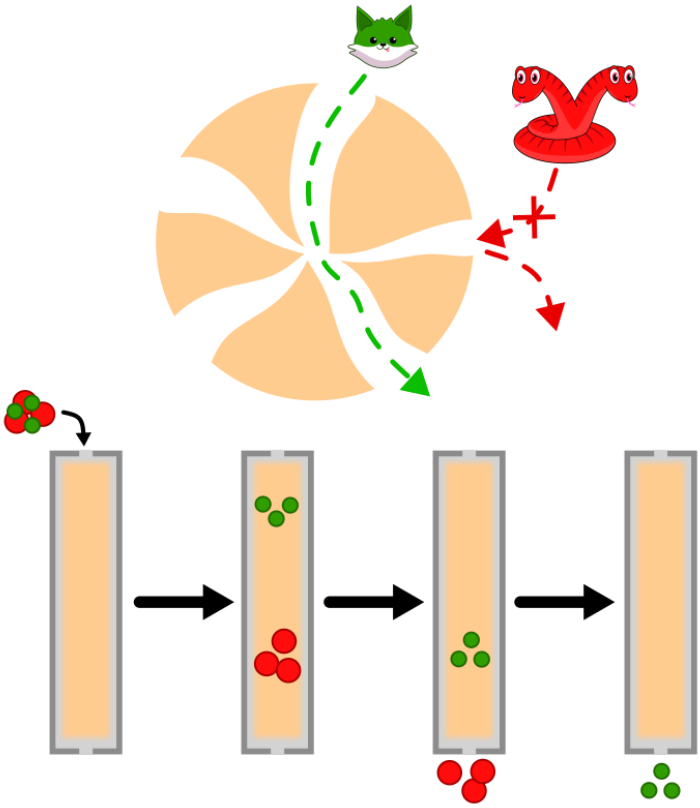


Fig. 13: The principle of size exclusion chromatography. Small proteins, in this example FOXP3 FHD (green), are able to enter the pores within the stationary phase. Larger proteins, such as NC (red), are unable to enter the pores and have access to less of the total column volume. As a result, large proteins elute first, while smaller proteins elute last. Image created using Inkscape V1.2.2.

For SEC involving NC, a Hiload 16/600 Superdex 200pg column (Cytiva, USA), suitable for separating proteins of molecular mass 10 kDa – 600 kDa, was used. For FOXP3 FHD, a HiLoad 26/600 Superdex 75pg column (Cytiva, USA), 3 kDa – 200 kDa, was used instead. An ÄKTA Prime Plus FPLC system was used for SEC. For both columns, the columns were equilibrated with 2 column volumes of equilibration buffer prior to each run. Samples were centrifuged at 12,000 xg for 12 minutes to remove aggregates prior to being placed on the column. To size each protein based off their elution time, standards of known molecular mass (Cytiva, USA) were run on each respective column; shown in **Table 3**. A standard curve was then constructed by plotting the log of the molecular mass versus the elution volume for each standard. The standard curve was then compared to the elution time of either the FHD or NC to estimate protein size.

For NC, SEC was used to purify RNA bound from unbound protein, allowing for elimination of nucleic acid contamination, as well as to remove NC degradation products. Additionally, SEC was used to determine the oligomeric state of each protein.

Table 3: Standards used for each size exclusion column and their respective molecular masses

Hiload 16/600 Superdex 200pg (NC)	
Standard	**Molecular Mass (kDa)**
Carbonic Anhydrase	29
Ovalbumin	43
Conalbumin	75
Aldolase	158
Ferritin	440
Blue Dextran	>2000

HiLoad 26/600 Superdex 75pg (FOXP3 FHD)	
Aprotinin	6.5
Ribonuclease A	13.7
Carbonic Anhydrase	29
Ovalbumin	43
Conalbumin	75
Blue Dextran	>2000

3.3.4 - SDS-PAGE

Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis (SDS-PAGE) is a technique that uses an electric field to separate proteins through a polyacrylamide gel. Due to the sieve-like nature of the polyacrylamide matrix, larger proteins migrate slower than smaller proteins, leading to separation based on size. The 3D conformation of proteins is removed as a factor due to denaturation and linearisation by SDS. Additionally, the SDS imparts an equal mass-to-charge ratio to all proteins which further eliminates protein charge as a factor affecting migration. This allows proteins to migrate based on size alone. The addition of β-mercaptoethanol reduces any disulfide bonds either within or between protein chains, allowing for complete linearisation and subunit dissociation. SDS-PAGE allows for assessment of protein purity following purification steps. Both glycine (Laemmli, 1970) and tricine (Schägger & von Jagow, 1987) SDS-PAGE are widely used; which differ in the buffering system. During initial migration through the stacking gel, glycine or tricine act as trailing ions. Due to their relatively high pKa's (9.6 and 8.15, respectively), the buffer molecules are predominantly in their zwitterionic forms and so migrate slowly behind the proteins. The separation of the buffer component from the leading chloride ions leads to a voltage gradient that helps sweep the protein samples along the gel. Once reaching the higher pH of the separating gel, the buffer molecule becomes ionised and their electrophoretic mobility sharply increases. As the buffer molecules pass the protein sample, the protein samples are left to migrate based off their charge (Schägger & von Jagow, 1987). Tricine's lower pKa allows it to migrate faster compared to glycine during the stacking phase, which better stacks smaller protein samples as they enter the separating gel. As a result, tricine gels typically possess better resolution for proteins smaller than 100 kDa, and glycine being superior for 100 – 500 kDa samples.

Here, 16% tricine gels were used for analysing both FOXP3 FHD and NC. The separating gel consisted of 16% acrylamide, 0.5% bis-acrylamide, 1 M tris-HCl, 0.1% SDS, pH 8.45, while the stacking gel consisted of 4% acrylamide, 0.12% bis-acrylamide, 0.75 M tris-HCl, 0.075% SDS, pH 8.45. Samples were prepared by adding 1 part sample to 1 part reducing sample buffer (150 mM tris-HCl, 12% SDS, 6% (v/v) β-mercaptoethanol, 30% (v/v) glycerol, 0.05% Coomassie G-250) to a volume of at least 20 µl. For particularly protein rich samples, such as lysate supernatant or IMAC flowthrough, samples were

instead made up with 1 part sample and 7 parts reducing sample buffer to dilute the protein concentration and prevent overloading the wells and creating lane distortion during sample migration. Samples were heated to 95 °C for 5 minutes to ensure full denaturation. Samples of the lysate pellet were additionally sonicated for 3 seconds to reduce viscosity and ease loading. A PageRuler Prestained Protein Ladder (ThermoFisher, USA) was used as a molecular weight marker.

The Bio-Rad Mini-PROTEAN Tetra System was used for gel casting and migration. Spacer plates with a width of 1 mm were used. While the separating gel was setting, the gel was overlaid with isopropanol to prevent atmospheric oxygen interfering with polymerisation and to ensure a flat interface between separating and stacking gel. Once polymerised, gels were assembled. The cathode buffer consisted of 0.1 M tris, 0.1 M tricine, 0.1% SDS, pH 8.25; while the anode buffer consisted of 0.1 M tris-HCl, pH 8.9. 8 µl of sample was loaded per well. Gels were initially run at 60 V until the dye front reached the interface between stacking and separating gel before the voltage was increased to 160 V and the gels electrophoresed for 1.5 – 2 hours. Gels for nucleocapsid were typically run slightly longer due to its larger size. Gels were stained using a 10% (v/v) acetic acid, 40% (v/v) methanol solution containing 0.1% (w/v) Coomassie G-250 dye. The combination of acetic acid and methanol aggregates and fixes the proteins within the gel, preventing dissociation and fading of bands. Destaining was performed overnight using a solution of the same composition of acetic acid and methanol, but without Coomassie G-250 dye. Destain solution was replaced as needed until the backgrounds were sufficiently destained. Gels were imaged using a Bio-Rad ChemiDoc XRS+ gel imaging system running ImageLab version 6.0.1.

3.3.5 - Evaluation of Protein Concentration

To determine protein concentration prior to conducting an experiment, protein was dialysed against any imidazole-free buffer suitable for the respective experiment and centrifuged at 12,000 xg for 12 minutes to remove aggregates. An absorbance spectrum from 240 – 450 nm of the protein was taken using either a Jasco V-730 absorbance spectrophotometer (JASCO Corporation, Japan) or a ThermoFisher NanoDrop One spectrophotometer. Absorbance at 280 nm was used alongside the Beer-Lambert law (**Eqn. 1**) to calculate protein concentration. Additionally, absorbance at 340 nm was subtracted from that at 280 nm to account for aggregation. There is minimal absorbance by proteins and nucleic acids at 340 nm, and so any observed absorbance at this wavelength is the result of light scattering by aggregates. By subtracting absorbance at 340 nm, the contribution to the observed absorbance by non-functional aggregated protein is removed. Thus, a more accurate concentration of soluble protein is obtained. The extinction coefficients used were obtained from ExPASy ProtParam using the respective protein sequences; available at https://web.expasy.org/protparam/ (Gasteiger et al., 2005).

Equation 1:

$$A = \varepsilon c l$$

Where A is the absorbance, ε is the molar extinction coefficient ($M^{-1} \cdot cm^{-1}$), c is the molar concentration (M), and l is the pathlength (cm). The molar extinction coefficient of the monomeric FOXP3 FHD is 24 980 $M^{-1}.cm^{-1}$. While the FOXP3 FHD is proposed to exist solely as a dimer (Bandukwala et al., 2011), for this work all concentrations reported are for the monomer. The NC possesses a monomeric molar extinction coefficient of 43 890 $M^{-1}.cm^{-1}$. Similarly, while NC primarily exists as a dimer, the monomeric concentrations are given.

3.4 - Structural Characterisation

3.4.1 - Circular Dichroism Spectroscopy

In order to determine whether each protein is correctly folded following purification, circular dichroism (CD) spectroscopy was used. CD spectroscopy is a technique used for investigating the secondary structural characteristics of a protein based on the interactions between chiral molecules and circularly polarised light. Chiral molecules will preferentially absorb either left- or right-circularly polarised light; which can be expressed as the difference in molar extinction coefficients. When the two oppositely circularly polarised light recombine, the resulting beam of light is said to become *elliptically* polarised. Thus, it is also common to instead plot the difference in absorption as the *molar ellipticity*, [θ], with the units $deg.cm^2.dmol^{-1}$ (Kelly & Price, 2000; Woody, 1995). Molar ellipticity can be normalised to the size and concentration of the protein with **Equation 2**:

Equation 2:

$$[\theta] = \frac{100 * \theta}{c * n * l}$$

Where [θ] is the normalised molar ellipticity ($deg.cm^2.dmol^{-1}$), θ is the observed ellipticity (°), c is the molar concentration (M) of protein in the sample, n is the number of residues within the protein, and l is the pathlength (cm) (Purdie, 1996; Woody, 1995).

Far-UV light, over the 180 – 240 nm range, is commonly used for CD spectroscopy. Over these wavelengths, absorption is dominated by the peptide backbone of proteins. Additionally, the extent of absorption is dependent on the bond angles between each of the many residues making up the protein. A consequence of this is that certain secondary structures possess unique and identifiable CD spectra as a result of the bond angles the peptide backbone adopts within these structures (**Fig. 14**) (Micsonai et al., 2015; Woody, 1995). Additionally, proteins with a combination of different secondary structures possess CD spectra that are intermediate to those of proteins possessing a single secondary structure type.

A Jasco J-810 CD spectropolarimeter (JASCO Corporation, Japan) was used to generate CD spectra of each protein. Protein was dialysed against 20 mM tris-HCl, 100 mM NaCl, 2.5 mM $MgCl_2$, 0.5 mM TCEP, pH 7.4. Samples were centrifuged at 12,000 xg to remove aggregates and buffer was filtered using a 0.22 µm filter. Spectra of 5 µM protein were taken from 180 – 260 nm in 0.2 nm increments at 100 nm/min using a 2 mm quartz cuvette (Hellma, Germany). A 1 second response and 5 nm bandwidth were used. Ten accumulations were taken for each sample and averaged to produce the final spectrum per sample. Data points with a tension voltage greater than 600 V were discarded prior to analysis. Three samples and a blank were taken. Finally, all samples were averaged, the blank subtracted, and the observed ellipticity converted to molar ellipticity (**Eqn. 2**). To confirm that the observed spectra were indeed the result of native folding, spectra of denatured proteins were also taken. Protein (5 µM) was denatured by incubation in the presence of 8 M urea for 30 minutes prior to having their spectra taken with the same settings as mentioned.

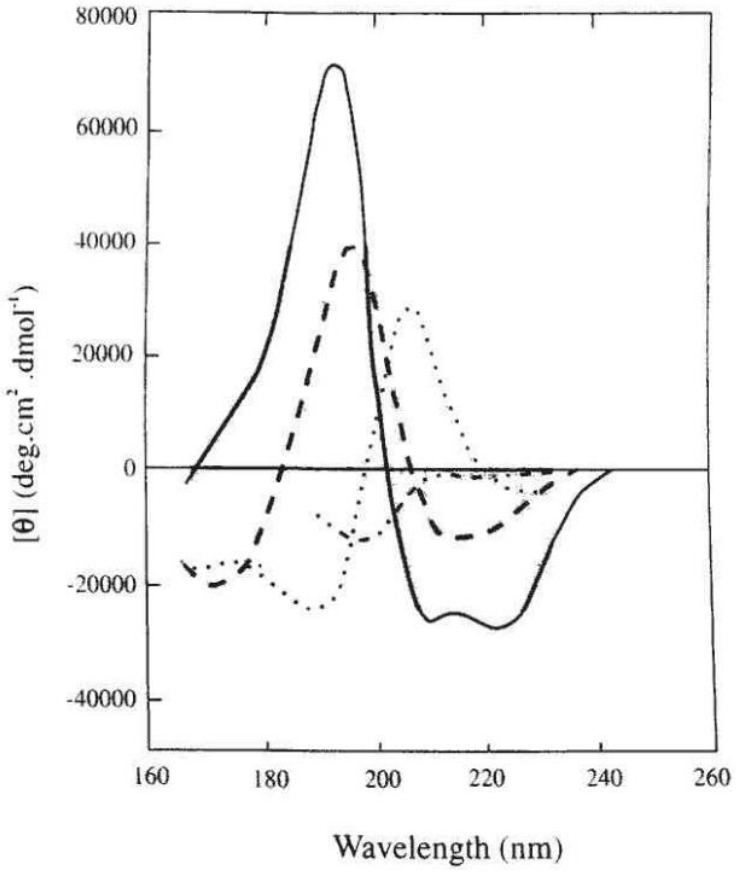


Fig. 14: Characteristic CD spectra of various protein secondary structural elements. Solid line – α-helix, long dashes - anti-parallel β-sheets, dots – β-turn, dotted dashes – random coil. Taken from (Kelly and Price, 2000).

3.4.2 - Intrinsic Tryptophan Fluorescence

Fluorescence is a subset of luminescence, the spontaneous release of light from a substance. In fluorescence, a photon of light is absorbed by an electron of a molecule, exciting the electron to a higher energy level. Following excitation, the electron near instantaneously relaxes back down to the ground state. This relaxation is accompanied by the release of a photon from the molecule; the wavelength of which is typically longer than the absorbed wavelength. Tryptophan is one of the greatest contributors to a protein's fluorescent properties. Due to its aromaticity, tryptophan possesses high absorption at 280 nm and high quantum yield upon emission. Importantly, the emission wavelength of tryptophan is influenced by its surrounding environment. The nitrogen atom of the indole ring dictates an important electronic transition that defines tryptophan's fluorescent properties, the ^1L_a transition. When in a polar environment, the formation of polar contacts with the nitrogen atom leads to a shift in the electron density around the atom which subsequently lowers the energy of the electronic transitions in the surrounding bonds, lowering the energy of the emitted photon according to **equation 3**.

Equation 3

$$E = \frac{hc}{\lambda}$$

Where E is the energy of a single photon in electron volts (eV), h is Plank's constant ($4.14*10^{-15}$ eV.s), c is the speed of light in a vacuum (299 792 458 $m.s^{-1}$), and λ is the wavelength of the photon in metres. When in a polar environment, the emission peak of tryptophan is red-shifted towards 352 nm. Conversely, when in a hydrophobic environment, there is a blue-shift towards 308 nm (Ghisaidoobe & Chung, 2014; Valeur, 2012; Vivian & Callis, 2001). Tryptophan fluorescence can thus be used to probe the tertiary structure of a protein as the spectra can be used to deduce information about the environment surrounding the tryptophan residues.

A Jasco FP-6300 spectrofluorometer (JASCO Corporation, Japan) was used to assess protein fluorescence spectra. The same samples used for CD spectroscopy could be used for fluorescence measurements. Protein was dialysed against 20 mM tris-HCl, 100 mM NaCl, 2.5 mM $MgCl_2$, 0.5 mM TCEP, pH 7.4. Samples were centrifuged at 12,000 xg to remove aggregates and buffer was filtered using a 0.22 µm filter. Protein at a concentration of 5 µM was excited using 280 nm or 295 nm excitation light and emission spectra recorded from 260 - 450 nm in 1 nm increments at 200 nm/min using a 10 mm quartz fluorescence cuvette (Hellma, Germany). A fast response with high sensitivity was used. A 5 nm slit width was used for the excitation wavelength while 2.5 nm was used for emission. Each sample was accumulated 3 times. Three samples and a blank were taken. Finally, all samples were averaged and the blank subtracted to produce the final spectrum. Like with CD spectroscopy, spectra of denatured protein were also taken.

3.5 - DNA Binding and Protein-Protein Interactions

3.5.1 - Electrophoretic Mobility Shift Assays

An Electrophoretic Mobility Shift Assay (EMSA) is a gel-based technique that can be used for the visualisation of protein-protein or protein-nucleic acid interactions. This technique is similar to SDS-PAGE in that the migration distance of a complex is a function of its size and charge. However, unlike SDS-PAGE, EMSAs are conducted using native gels, i.e. SDS is excluded and proteins are found in their native conformation. When studying protein-nucleic acid interactions, DNA or RNA is included with the protein sample. If the protein of interest is capable of binding to the nucleic acid sequence, the resulting complex possesses a larger cross-sectional area than either component alone; resulting in an upwards shift in the gel due to decreased mobility (**Fig. 15A**). This upwards shift can then be visualised using an appropriate stain, such as GelRed (**Fig. 15B**), to visually confirm the presence of nucleic acid interaction (Fried & Crothers, 1981; Garner & Revzin, 1981). While it is theoretically possible to determine dissociation constants from EMSAs, in this work they were used qualitatively to visualise interactions.

Prior to sample preparation, proteins were concentrated to an arbitrary but relatively high concentration using a 50 ml Amicon® stirred cell (MilliporeSigma, USA). Proteins were also dialysed against a low salt buffer; typically 20 mM tris-HCl, 150 mM NaCl, pH 7.5. Concentrating the proteins and use of a low salt buffer were done to reduce the effect of excessive chloride ions and other buffer components in the protein sample interfering with migration. Samples were prepared to a volume of 25 µl, with a final composition of 20 mM tris-HCl, 150 mM NaCl, 1.5 mM $MgCl_2$, 10% (v/v) glycerol, pH 7.5 at 4 °C, containing 0.5 µM DNA and increasing concentrations of protein in various DNA:protein ratios. Samples were incubated at 4 °C for a minimum of 30 minutes prior to loading onto the gel to allow for binding to occur.

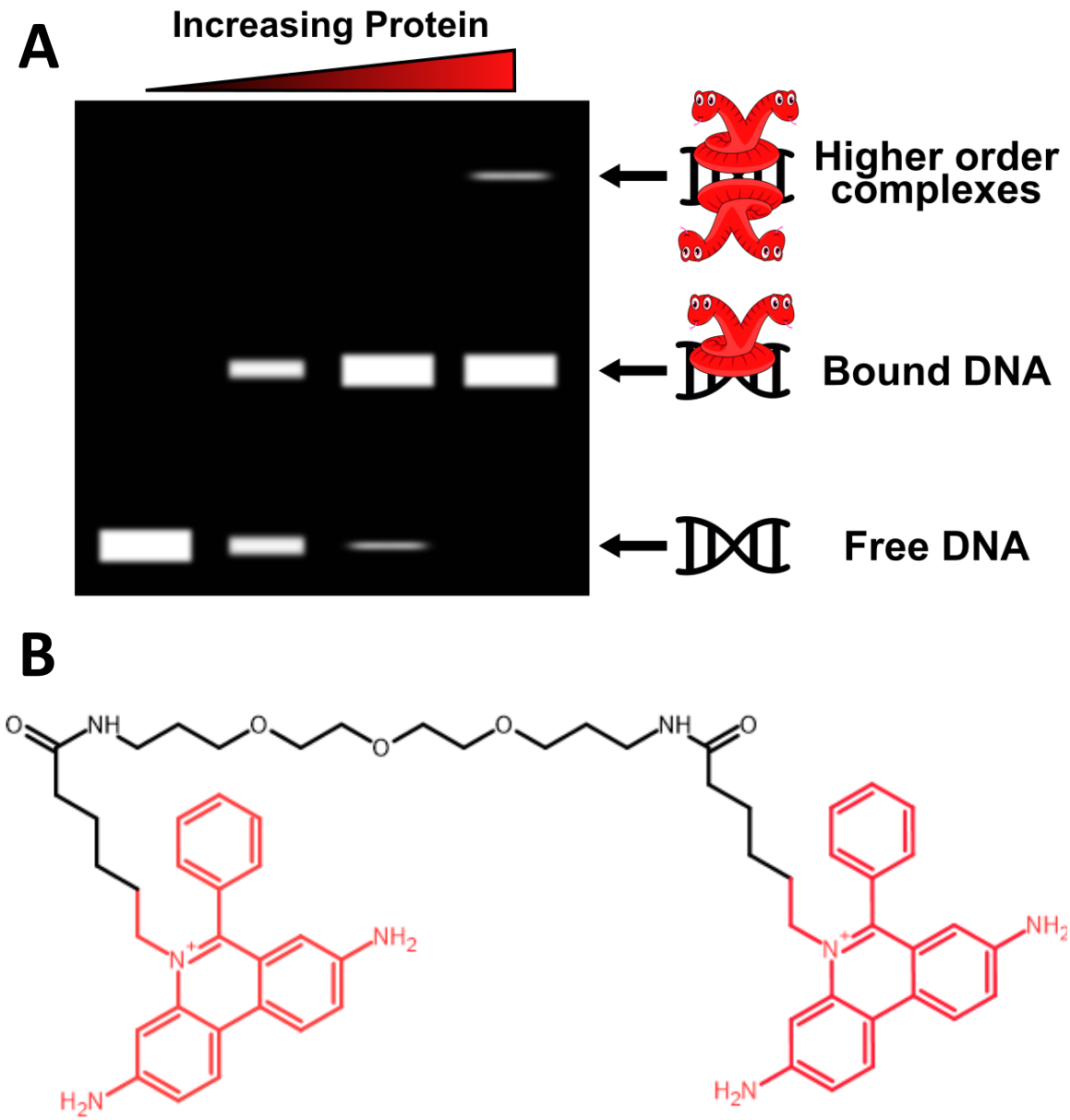


Fig. 15: Electrophoretic mobility shift assays. A - Principle of electrophoretic mobility shift assays. Free DNA migrates a certain distance based on its size (left-most lane). As more protein is included in subsequent wells, a greater proportion of DNA becomes protein-bound. The protein-DNA complex possesses a greater size than the free DNA, and so migrates slower through the gel. Higher-order complexes can also be visualised; where further protein binding to already bound DNA leads to a further shift. **B** – Proprietary chemical structure of GelRed. GelRed consists of two ethidium units (red) linked together which prevents GelRed from crossing the cell membrane. Additionally, the ethidium units are complexed with iodine instead of bromine further decreasing toxicity. The ethidium units intercalate nucleic acids, upon which their fluorescence intensity is greatly increased. Image created using Inkscape V1.2.2. Chemical structures drawn using Chemdraw 19.0.0.

EMSA gels (10% acrylamide, 0.12% bis-acrylamide, 40 mM tris-acetic acid, 2.5 mM ethylenediaminetetraacetic Acid (EDTA), 20% triethylene glycol, pH 7.8 at 4 °C) were prepared and cast in an identical manner to SDS-PAGE gels. Triethylene glycol acts as a crowding agent that helps stabilise the protein-DNA complex and improves resolution (Hellman & Fried, 2007). Gel equipment was thoroughly rinsed prior to assembly to eliminate any residual SDS. No stacking gel was used and instead the gel was one continuous 10% polyacrylamide gel. Following polymerisation, the gel and all

necessary equipment were placed at 4 °C for at least 30 minutes to equilibrate. Following equilibration of both gel and samples, 10 µl of sample was loaded per well, and the gel was run at 160 V for 1.5 hours at 4 °C using TAE (40 mM tris-acetic acid, 2.5 mM EDTA, pH 7.8 at 4 °C) as tank buffer. After migration, gels were stained using a solution of 1x GelRed nucleic acid stain (Biotium, USA) and visualised using a Bio-Rad ChemiDoc XRS+ gel imaging system running ImageLab version 6.0.1. GelRed is an intercalating nucleic acid dye that possesses two ethidium units linked together with a long linker region. Due to this, the dye has similar spectral properties to ethidium bromide while being purported as less toxic (Biotium, 2013). Upon exposure to UV light, GelRed fluoresces orange, the intensity of which greatly increases upon nucleic acid binding.

3.5.2 - Pulldown Assays

Pulldown assays are similar in concept to co-immunoprecipitation, but instead of using an immobilised antibody, a "bait" molecule; which can be either a purified protein or small molecule; is immobilised to a matrix in a column. Different matrices are possible, but commonly an agarose NTA-cobalt resin is used that immobilises His-tagged proteins in an identical manner to IMAC. After immobilisation, a "prey" molecule is passed through the column. If an interaction between the two occurs, the prey becomes immobilised to the matrix by the bait. The two can be eluted together and their presence measured using an appropriate technique to assess interactions (**Fig. 16**) (Wilson & Walker, 2009). While pulldown assays do not provide quantitative information on a particular interaction; like one would get from fluorescence anisotropy; they are easy to perform, quick, and provide qualitative

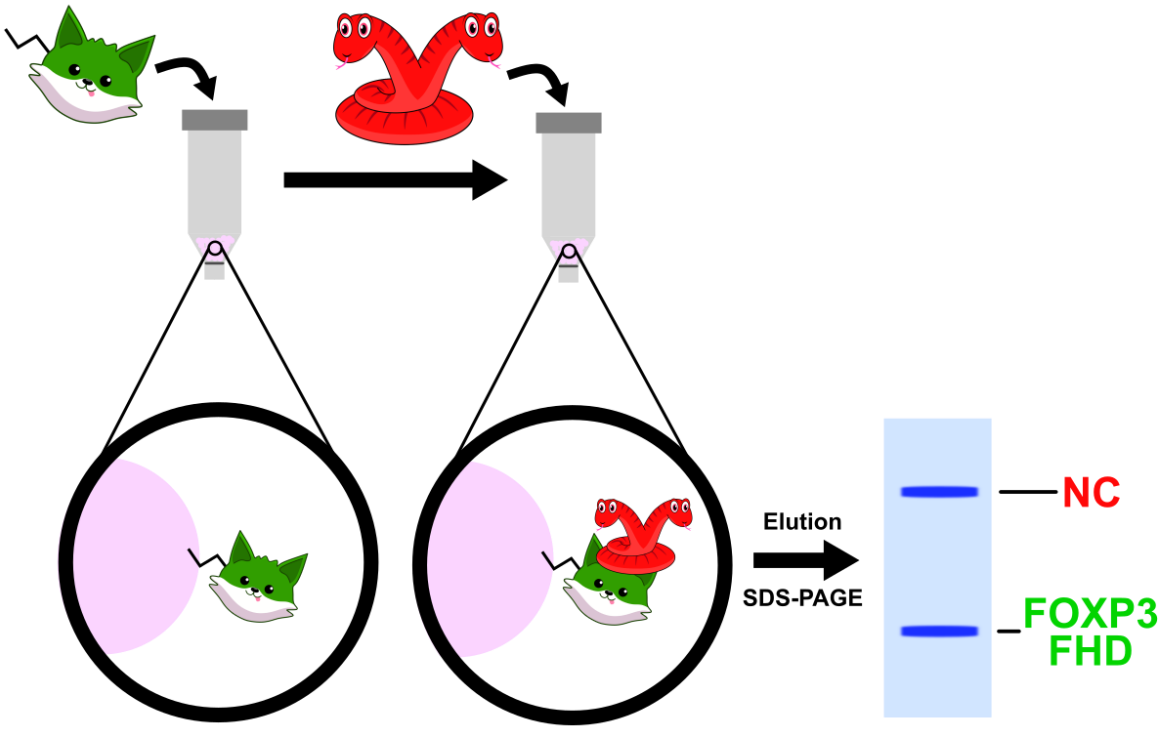


Fig. 16: The concept of a pulldown assay. A substrate, such as an NTA-cobalt resin (pink) is incubated with the *bait* protein (FOXP3 FHD). The protein attaches to the cobalt resin via His-tag complexation in an identical manner to IMAC. After, the *prey* (NC) is added. If an interaction between the two proteins occurs, NC becomes immobilised to the substrate via binding to FOXP3 FHD. Upon addition of an imidazole-rich buffer, the two proteins elute together and can be visualised using SDS-PAGE. Image created using Inkscape V1.2.2.

visual evidence of potential interactions. This makes them a good starting point to investigate potential interactions as well as to support other, more qualitative, investigations.

Pulldowns were performed using a ThermoFisher Pierce™ His Protein Interaction Pull-Down Kit (ThermoFisher, USA) and performed as per manufacturer's instructions. The FOXP3 FHD, due to its His-tag, was used as the bait. NC acted as the prey due to it not possessing a His-tag. As a control, NC was also added to a column with no bait to confirm that NC did not interact with the resin. Prior to the assay, proteins were dialysed against tris-buffered saline (TBS) (25 mM tris-HCl, 150 mM NaCl, pH 7.2). To each column, 50 µl of the provided cobalt resin slurry was added, and 150 µg of FOXP3 FHD added (or an equal volume of buffer for the control column). The columns were incubated at 4 °C for 30 minutes and the resin then washed 5 times with 400 µl 25 mM tris-HCl, 150 mM NaCl, 0.5 mM EDTA, 2.5% (v/v) glycerol, 10 mM imidazole, pH 7.2. Following washes, 150 µg NC was added and the columns incubated for 1 hour at 4 °C. The columns were then washed as previously. Proteins were eluted using a 300 mM imidazole solution. Throughout the process, various samples of the flowthroughs and resins were taken for analysis via SDS-PAGE.

3.5.3 - Protein Labelling

For some protein-DNA and protein-protein studies, it was required to label one of the proteins with a fluorescent label. Common labelling techniques include maleimide labelling of cysteine residues or complexing of an NTA-linked fluorophore to the His-tag of a protein. FOXP3 FHD possesses one cysteine residue, Cys394, but unpublished work from our lab has found that labelling of this residue leads to abrogation of DNA binding. Additionally, NC possesses no cysteine residues. Therefore, maleimide labelling was ill-suited for this work. For His-tag labelling, the His-tag of NC was removed during purification, leaving His-tag labelling of FOXP3 FHD the best choice for this study. NTA-ATTO 550 (Sigma-Aldrich, USA) was used to label FOXP3 FHD (**Fig. 17**). Protein was dialysed against 20 mM tris-HCl, 150 mM NaCl, 1 mM TCEP, 0.5 mM $NiSO_4$, pH 7.5. Following dialysis, 200 µl of 15 – 20 µM protein was incubated overnight with an equimolar amount of NTA-ATTO 550. The following day, the sample was centrifuged at 12,000 xg for 12 minutes to remove aggregates, and the reaction mixture mixed with 20 µl of dye removal resin and added to a dye removal column (ThermoFisher, USA) to remove uncomplexed dye. The column was centrifuged at 1000 xg for 30 seconds and the flowthrough

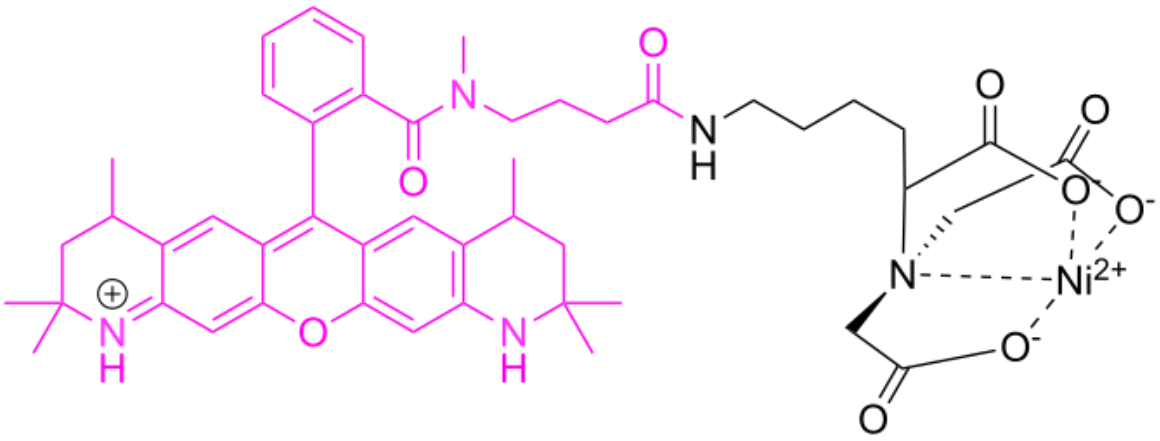


Fig. 17: Structure of NTA-ATTO 550. The ATTO 550 dye (pink, left) is linked to an NTA group (black, right) and complexes with the His-tag of the protein in a similar fashion to IMAC. Image created using Inkscape V1.2.2. Chemical structures drawn using Chemdraw 19.0.0.

collected. To determine labelling efficiency, an absorbance spectrum of the mixture was taken as previously described, and the efficiency calculated using **Equations 4**, **5**, and **6** below.

Equation 4

$$[ATTO\ 550] = \frac{A_{554}}{\varepsilon_{554}}$$

Equation 5

$$[Protein] = \frac{(A_{280}) - 0.1 * (A_{554})}{\varepsilon_{280}}$$

Equation 6

$$Labelling\ Efficiency = \frac{[ATTO\ 550]}{[Protein]}$$

Where A_{554} is the absorbance at 554 nm; ATTO 550's absorbance maximum; A_{280} is the absorbance at 280 nm, ε_{554} is the extinction coefficient of ATTO 550 at 554 nm (ε_{554} = 1.5x10^5 M^{-1}.cm^{-1}), and ε_{280} is the extinction coefficient of the protein. A correction factor of 0.1*A_{554} was applied to the protein absorbance due to absorbance at 280 nm by the label itself.

3.5.4 - Fluorescence Anisotropy

If incident light upon a fluorophore is polarised, the light it emits will remain polarised. However, due to Brownian motion, the fluorophore will continuously rotate in solution over the course of its fluorescence lifetime. Thus, as the molecule rotates, so too does the polarisation of the emitted light. Larger molecules, on account of their inherently greater mass, tumble slower than smaller molecules and thus the emitted light retains a greater degree of polarisation. Effectively, the fluorescence anisotropy of a molecule is an indirect measure of the molecule's size. This allows for assessment of binding events as these events lead to the formation of larger complexes with increasing polarisation (**Fig. 18**).

Prior to any experimental measurements, a grating factor (G-factor) is calculated by taking the ratio of the horizontal to vertical polarisation component of emitted light while the unbound sample is being analysed (**Eqn. 7**). The G-factor corrects for any inherent bias for any particular plane of polarisation caused by the optics of the polarimeter or sample. During experimental measurements, the polarimeter excites the sample using light polarised only on a single axis, typically the z-axis, and the resulting fluorescence anisotropy (*r*) of the emitted light is the ratio of the polarised component to the total light intensity with the G-factor applied (**Eqn. 8**) (Jameson & Sawyer, 1995; Steiner, 2002; Swonger & Robinson, 2018).

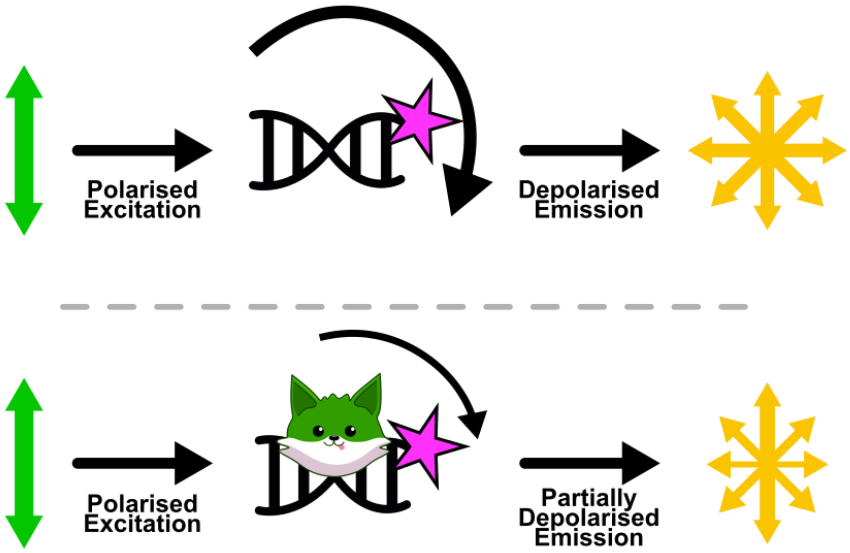


Fig. 18: Principle of fluorescence anisotropy. When excited by polarised light, the emission of a fluorophore remains polarised. However, due to Brownian motion, the emitted light can become depolarised by rotation of the fluorophore between excitation and emission. Larger molecules tumble slower in solution due to their inherent mass; thus their emitted light remains more polarised. Binding events can be assessed by monitoring the fluorescence anisotropy of a fluorescently labelled target molecule. Image created using Inkscape V1.2.2.

Equation 7

$$G = \frac{I_{\perp}}{I_{\parallel}}$$

Equation 8

$$r = \frac{I_{\parallel} - G * I_{\perp}}{G * I_T}$$

Where $I_{\parallel}$ represents the intensity of polarised light parallel to the excitation light, $I_{\perp}$ the perpendicular, I_T the total intensity, and G the grating factor. By measuring the fluorescence anisotropy of the sample at increasing concentrations of ligand, the points can be fit to a single site ligand binding curve (**Eqn. 9**).

Equation 9

$$f(x) = \frac{B_{max} * [Ligand]}{K_D + [Ligand]}$$

Where B_{max} is the maximum fluorescence anisotropy achieved at maximal binding and K_D is the dissociation constant of the reaction.

As only one species is fluorescently labelled in each fluorescence anisotropy experiment, a given binding curve only reports on the behaviour of the labelled member participating in the reaction; where the behaviour of other members can only be inferred. Thus, when investigating multi-member binding events, it is often beneficial to perform identical experiments with different members labelled

to better understand the behaviour of all participating macromolecules. For binding events involving FOXP3 FHD, NC, and DNA, either of the three may be labelled in theory. **Table 4** lists the fluorescence anisotropy experiments performed for this work. Where necessary, proteins were labelled with NTA ATTO 550 as previously described. For experiments involving labelled DNA, one strand of the oligonucleotide was labelled on the 5' end with carboxy-X-rhodamine (ROX) via an N-Hydroxysuccinimide (NHS) ester. Labelled oligonucleotides were purchased from Integrated DNA technologies – Whitehead Scientific, South Africa.

Table 4: Fluorescence anisotropy experiments performed.

Experiment	Labelled Species	Label	Titrand	[Titrand] (μM)	Titrant	Excitation (nm)	Emission (nm)
FHD – *p38α* DNA Binding	FHD	NTA ATTO 550	FHD	0.5	*p38α*	554	575
NC – *p38α* DNA binding	*p38α*	ROX (As 5' NHS ester)	*p38α*	0.2	NC	578	604
NC – FHD binding	FHD	NTA ATTO 550	FHD	0.5	NC	554	575
p38α DNA into (FHD – NC)	FHD	NTA ATTO 550	FHD NC	0.5 0.5	DNA	554	575
NC into (FHD – *p38α* DNA)	FHD	NTA ATTO 550	FHD p38α	0.5 1 or 5	NC	554	575

Fluorescence anisotropy was conducted using a Perkin Elmer LS-50B luminescence spectrophotometer. The slit width was set to 10 nm for both excitation and emission and the integration time set to 3 s. All proteins were dialysed against 20 mM tris-HCl, 150 mM NaCl, 1 mM TCEP, 0.5 mM $NiSO_4$, pH 7.5. Nickel was included in the buffer as it was found to prevent the ATTO 550 label from dissociating from the protein over the course of the experiment. Samples were incubated for 1 minute following titration to allow binding to occur. Ten readings were taken and averaged for each point. The volume increase due to titration did not exceed 10% of the starting volume.

4 – Results

Gene expression is a tightly controlled process involving a myriad of protein interactions and molecular events. The regulation of gene expression is important for correct cell functioning and overall health. The NC of SARS-CoV-2 is likely capable of disrupting this fine regulatory balance, possibly as a means of increasing viral replication. FOXP3 is a transcription factor important in immune regulation and platelet development. While NC may disrupt the function of a number of transcription factors, this work aims to study the potential interaction between NC and FOXP3 due to the latter's important role in haemostasis and immunity. To do this, the DNA binding domain of FOXP3, the FHD, and NC were overexpressed, purified, and their potential interactions studied. This was done by assessing their DNA binding ability either in isolation or together. Such an interaction could explain the widespread symptoms experienced by COVID-19 patients.

4.1 - Protein Purification

Proteins were purified to characterise potential interactions in the absence of any other confounding factors such as extraneous proteins. Proteins were expressed in T7 *E. coli* cells transformed with pET-11a plasmids. Both were purified using nickel IMAC.

4.1.1 - FOXP3 FHD

FOXP3 FHD was successfully purified using IMAC (**Fig. 19A**). FOXP3 elutes with significant tailing in the elution peak. Typically, this would necessitate a higher concentration of imidazole in the elution buffer as tailing indicates residual protein interactions with the column. However, the tailing proved to be beneficial in obtaining pure protein. The first half of the elution peak was an approximately equal mixture of contaminating proteins and FOXP3 FHD (Lane E1, **Fig. 19B**). Given the relatively small volume of this fraction (typically ≈1 ml), it was deemed impractical to purify further. The tail, meanwhile, consisted of pure FOXP3 FHD (Lanes E3 onwards) in ever decreasing concentrations. Generating a standard curve from the SDS-PAGE reveals that the FOXP3 FHD migrates with a size of 14.1 kDa (**Fig. 19C**), slightly larger than the size expected from its sequence (12.9 kDa). Given the accuracy of the standard curve (R^2 = 0.98), this discrepancy likely lies with the protein itself. Particularly hydrophobic patches of some proteins, such as the highly hydrophobic dimer interface of the FHD, are capable of binding extra SDS during denaturation, raising their apparent molecular mass (Rath et al., 2009; Tiwari et al., 2019). There does appear to be a contaminating protein at approximately 25 kDa in the elution fraction. However, this band persists following size exclusion chromatography (see **Fig. 23E** in **Section 4.1.5**) and is approximately twice the size of the monomeric FOXP3 FHD. There are cases of protein oligomers being resistant to SDS and β-mercaptoethanol denaturation to varying degrees; such as human growth hormone (Grigorian et al., 2005), human inducible nitric oxide synthase (Kolodziejski et al., 2003), human ATP synthase (Wittig et al., 2008), some lectins in plants (Renwrantz et al., 2009), and other non-human proteins (Gentile et al., 2002; Matsui et al., 2019). FOXP3 FHD has been found to be a highly stable obligate dimer held together primarily through hydrophobic interactions (Bandukwala et al., 2011; Perumal, 2018); thus this band is thought to represent dimeric FOXP3 FHD that remains associated even through SDS-PAGE denaturation.

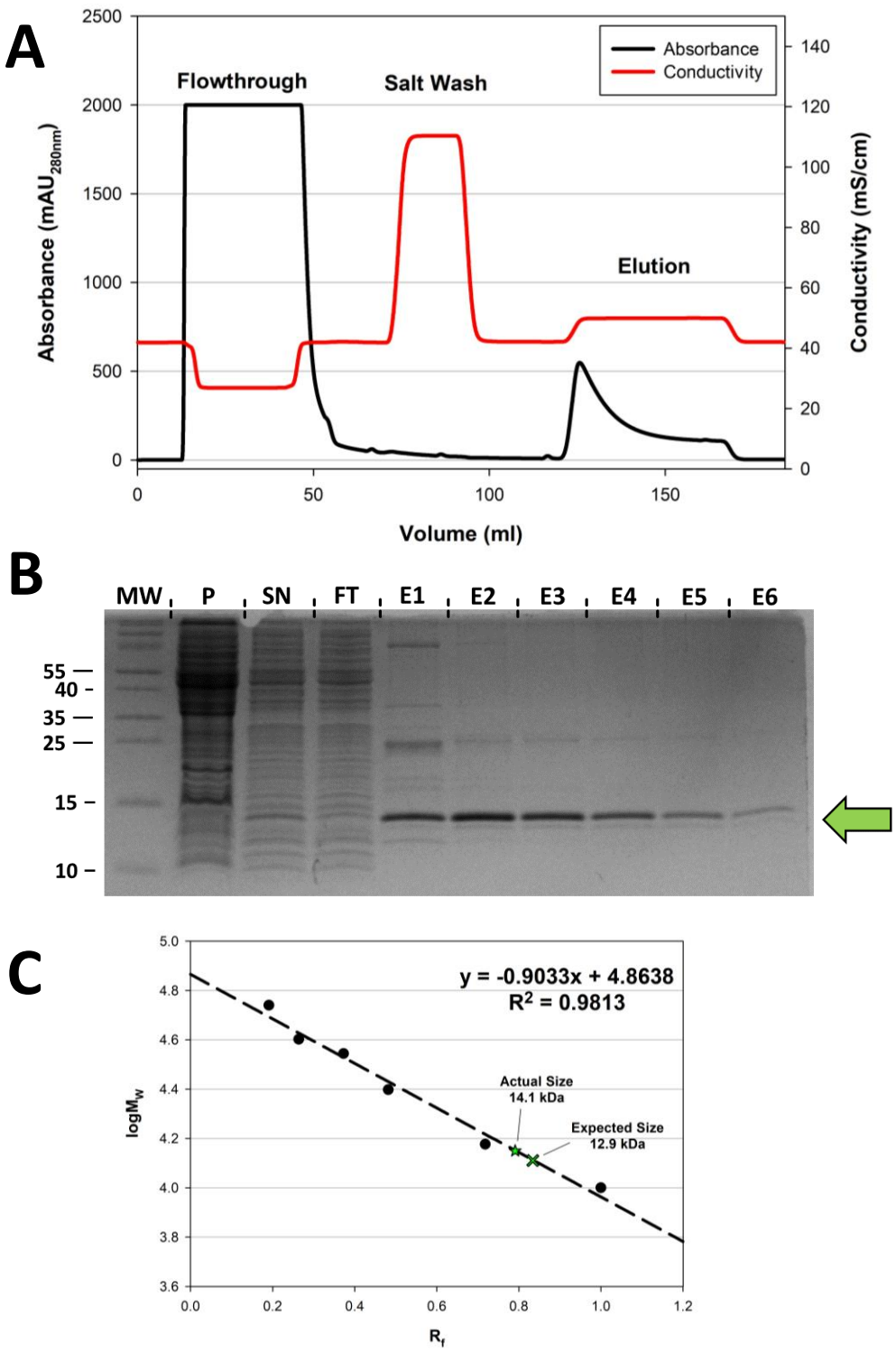


Fig. 19: FOXP3 FHD purification. A – IMAC chromatogram of FOXP3 FHD purification. As the lysate supernatant is passed through the column, unwanted bacterial proteins elute through the column (Flowthrough). Following elution of the unbound fraction, a salt wash is performed to remove loosely bound contaminating proteins and nucleic acid contamination. Finally, FOXP3 FHD is eluted using 300 mM imidazole. Conductivity is shown to monitor buffer changes. **B** – SDS-PAGE of FOXP3 FHD purification, with the position of the FHD marked with a green arrow. MW – Molecular weight marker, with sizes in kDa indicated on the left. P – Lysate pellet. SN – Lysate supernatant. FT – IMAC flowthrough, E – elution fractions 1 through 6. E1 represents the rise and top of the elution peak, with subsequent lanes representing portions of the elution tail. **C** – Standard curve generated from the SDS-PAGE to determine the size of the FOXP3 FHD. The FHD appears to possess a size of 14.1 kDa (green star), slightly larger than the 12.9 kDa expected from the sequence (green cross). Data plotted using SigmaPlot V12.

4.1.2 - Nucleocapsid

NC was similarly purified using IMAC (**Fig. 20A**). Unlike FOXP3 FHD, the salt wash during NC purification showed a significant absorbance peak. However, SDS-PAGE showed little to no protein in the corresponding fraction (Lane SW, **Fig. 20C**), suggesting that the absorbance was due to large amounts of nucleic acids within the elution. Indeed, absorbance spectroscopy of the salt wash fraction showed a spectrum typical that of nucleic acids (**Fig. 20B**) with an A_{260}/A_{280} indicating pure nucleic acid (A_{260}/A_{280} = 2.08). NC has a very high affinity for RNA and binds it non-specifically (Wu et al., 2021) and is likely binding bacterial RNA present in the lysate. Previous literature has also noted issues with nucleic acid contamination during NC purification (Luo et al., 2006; Wu et al., 2021). Addition of RNase did not completely alleviate the issue. NC does elute from the IMAC column with significant amounts of contaminating proteins (Lanes E1-3). Thrombin cleavage of the His-tag successfully removed these contaminants (**Fig. 21**). As NC no longer contained the His-tag, it was free to elute in the flowthrough during IMAC purification. Meanwhile, any contaminants, along with undigested NC and free His-tag, bound to the IMAC column as they did during initial IMAC purification.

Because the His-tag is relatively short, there is only a small difference in protein size between the uncleaved and cleaved NC (47.5 vs 45.7 kDa, respectively), resulting in only small differences seen on an SDS-PAGE gel. Interestingly, NC appears to migrate with a size of ~55 kDa during SDS-PAGE (**Fig. 20D**). Similarly to the FOXP3 FHD, the standard curve was accurate to the molecular weight marker (R^2 = 0.99) and so the discrepancy likely lies with the protein. Plasmid sequencing confirmed that the plasmid coded for the correct 47.5 kDa protein; and so the anomalous size is not the result of an incorrect protein construct. There are various reasons any given protein can migrate anomalously during SDS-PAGE such as: being a membrane protein, the prevalence of acidic residues, and disulfide bond formation, amongst others (Rath et al., 2009; Tiwari et al., 2019). The upward shift of NC on SDS-PAGE is likely caused by its intrinsically disordered structure. Firstly, IDPs, as a result of their lower hydrophobic residue content, tend to bind less SDS than a typical more structured protein; raising their mass-to-charge ratio and shifting their apparent molecular weight upwards (Rath et al., 2009; Tompa, 2002). Secondly, during denaturation of structured proteins by SDS, the polypeptide chain does not assume a completely linear coil. Instead, denatured proteins typically take on a beads-on-a-string type structure, where each “bead” is a section of structure; typically a cluster of α-helices; linked by disordered linkers. However, IDPs already possess large regions of disordered linkers in their native state, and thus, once extended by SDS denaturation, become longer than an average structured protein of similar molecular mass (Shirahama et al., 1974; Somasundaran & Hubbard, 2006) increasing their apparent molecular weight in SDS-PAGE.

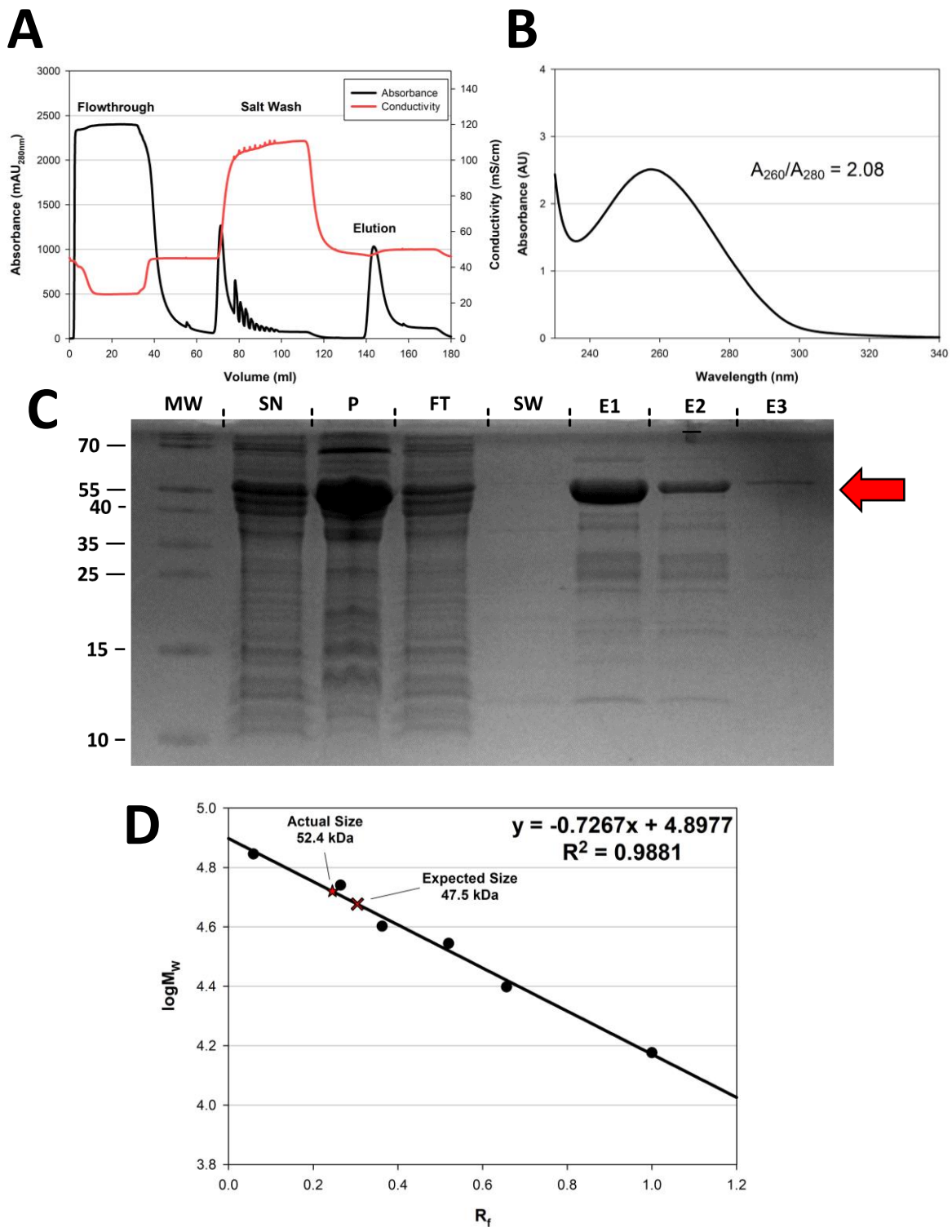


Fig. 20: NC purification. A – Initial NC IMAC purification using a nickel column. **B** - Absorbance spectrum of the NC salt wash fraction. **C** – NC purification SDS-PAGE, with the position of NC marked with a red arrow. MW – Molecular weight marker, SN – Supernatant, P – Pellet, FT – Flowthrough, SW – Salt wash, E – Elution. **D** – Standard curve generated from the SDS-PAGE to determine the size of NC. NC possess a size of 52.4 kDa (red star), slightly larger than the 47.5 kDa expected from the sequence (red cross). Data plotted using SigmaPlot V12.

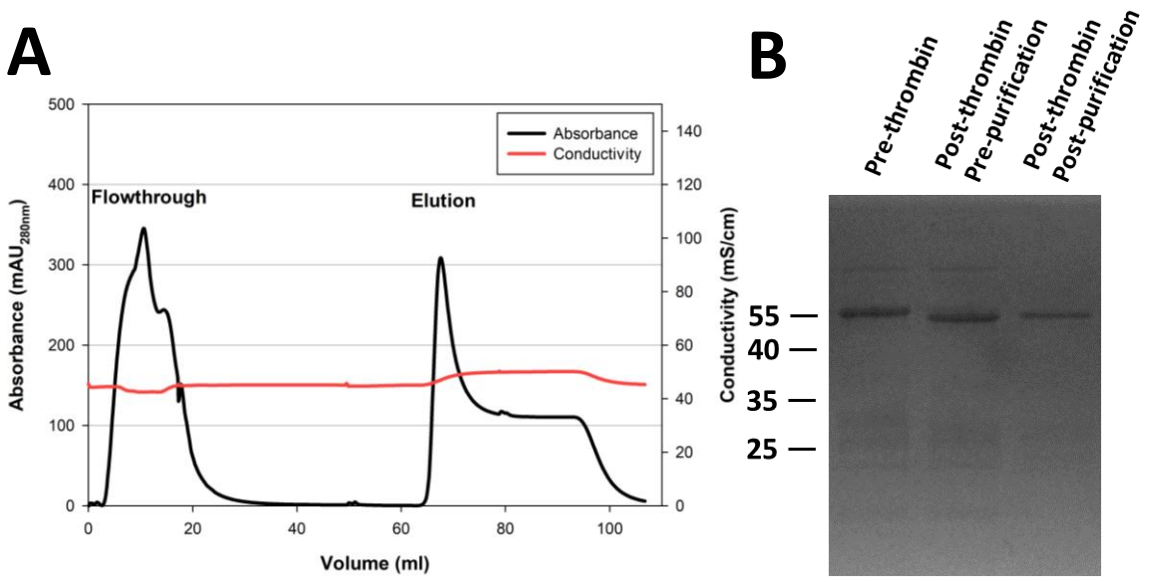


Fig. 21: NC purification following thrombin cleavage. **A** – IMAC chromatogram of cleaved NC purification using a benzamidine and nickel IMAC column in series. After cleavage, NC no longer contained the His-tag, and so elutes in the flowthrough. Uncleaved NC, free His-tag, and contaminants eluted in the elution. **B** – SDS-PAGE of NC thrombin cleavage. Pre-thrombin – NC IMAC fraction prior to addition of thrombin. Post-thrombin pre-purification – The same NC IMAC fraction but following 6 hours of incubation with thrombin. Post-thrombin post-purification – Cleaved NC following purification using a benzamidine and IMAC column. Data plotted using SigmaPlot V12.

4.1.3 - Nucleocapsid SEC

SEC was used to separate RNA-bound from RNA-unbound NC. NC elutes from the SEC column in three main peaks (**Fig. 22A**), consistent with previous literature (Wu et al., 2021). A large peak with an apparent size of ~950 kDa is seen relatively soon after the void volume (Peak 1, blue and red lines. Standards are given in **Fig. 23A** and standard curve in **23B**). However, SDS-PAGE analysis of this peak shows no detectable protein. Additionally, the A_{260nm}/A_{280nm} ration of the fraction collected for that peak was 1.68, indicating high amounts of nucleic acid content. Since the A_{260nm}/A_{280nm} ratio of pure nucleic acid is approximately 2 (Wilson & Walker, 2009), protein was likely present in low amounts; low enough to be undetectable via Coomassie staining. Instead, the peak height was likely dominated by the absorbance of the nucleic acids themselves. While absorbance by nucleic acids is greatest at 260 nm, absorbance at 280 nm is still substantial (Wilson & Walker, 2009). If, during IMAC, the salt wash is periodically paused and allowed to incubate on the column repeatedly, this absorbance peak is not present (black line); indicating that contaminating nucleic acids can be removed effectively simply using the salt wash. A second absorbance peak is seen at approximately 61 ml, corresponding to a size of ~300 kDa. NC is an intrinsically disordered protein, and the hydrodynamic volume of disordered proteins can vary up to 12x greater than a globular protein of the same molecular mass (Uversky & Dunker, 2010). Thus, sizing disordered proteins using SEC is unreliable. Given that the dimeric form of NC dominates in solution (Bai et al., 2021; Wu et al., 2021), this peak likely represents the dimeric form. This peak had an A_{260nm}/A_{280nm} of 0.63, indicating nearly nucleic acid free protein. When the proteins from this peak were separated on an SDS-PAGE gel, it revealed that it indeed consisted of full-length NC; showing that SEC can be used to purify NC of nucleic acid if it was unable to be removed during initial purification. Additionally, peak 2 is only present when performing SEC at

500 mM NaCl and disappears at 150 mM NaCl. The baseline also becomes noisier at low salt concentrations, possibly indicating a heterogenous population of RNA-bound, RNA-free, and partially RNA-bound protein. At high salt concentrations, like many nucleic acid binding proteins, salt ions outcompete NC for nucleic acid binding through electrostatic interactions which leads to dissociation of the protein from the nucleic acid (Lohman, 1986). Clear separation of RNA-bound from unbound NC via SEC was also seen by others (Luo et al., 2006; Wu et al., 2021). Finally, a third peak at 86 ml is also seen. Corresponding to a size of ~32 kDa, SDS-PAGE showed this peak consists of degraded NC (see next section), which others have also observed (Perdikari et al., 2020). Thus, SEC can also be used to purify NC from its autolysis products.

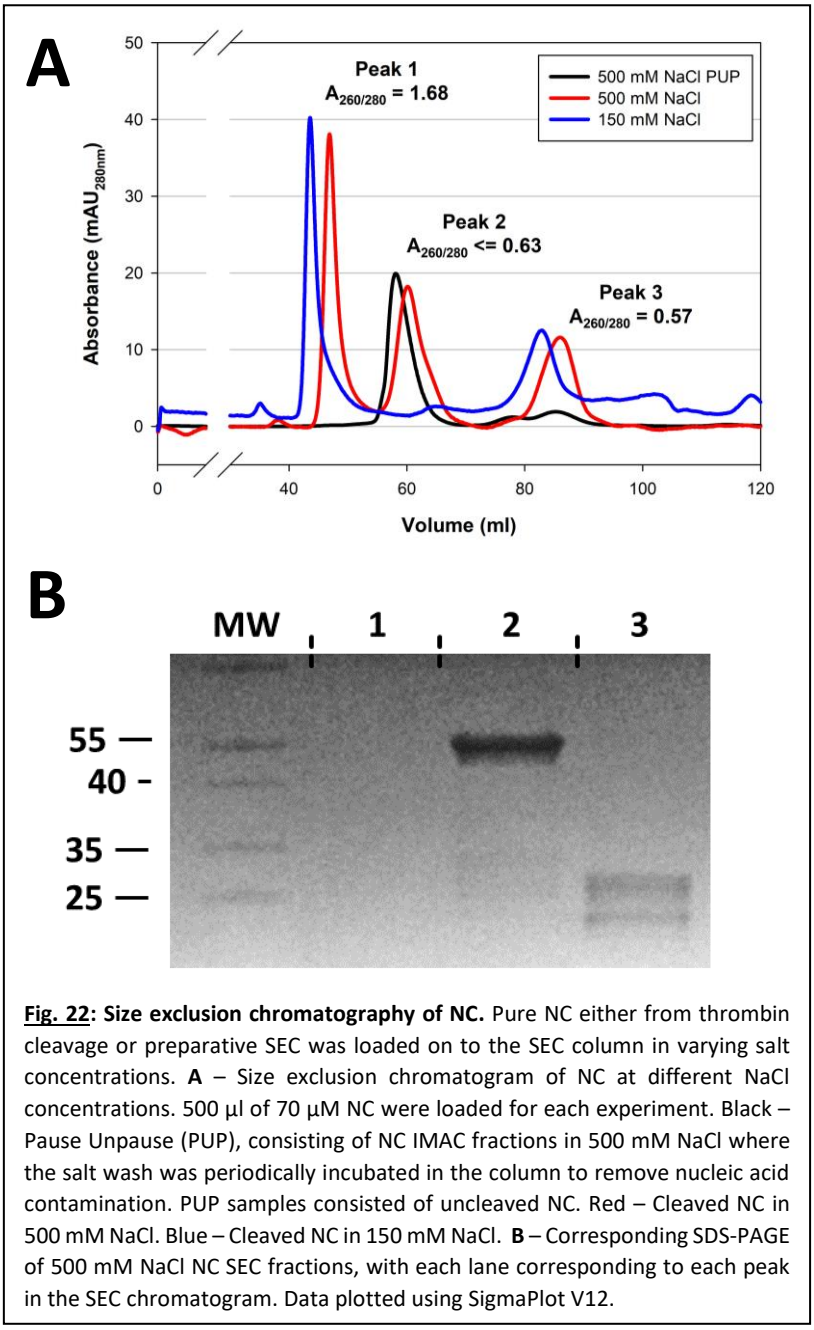


Fig. 22: Size exclusion chromatography of NC. Pure NC either from thrombin cleavage or preparative SEC was loaded on to the SEC column in varying salt concentrations. **A** – Size exclusion chromatogram of NC at different NaCl concentrations. 500 µl of 70 µM NC were loaded for each experiment. Black – Pause Unpause (PUP), consisting of NC IMAC fractions in 500 mM NaCl where the salt wash was periodically incubated in the column to remove nucleic acid contamination. PUP samples consisted of uncleaved NC. Red – Cleaved NC in 500 mM NaCl. Blue – Cleaved NC in 150 mM NaCl. **B** – Corresponding SDS-PAGE of 500 mM NaCl NC SEC fractions, with each lane corresponding to each peak in the SEC chromatogram. Data plotted using SigmaPlot V12.

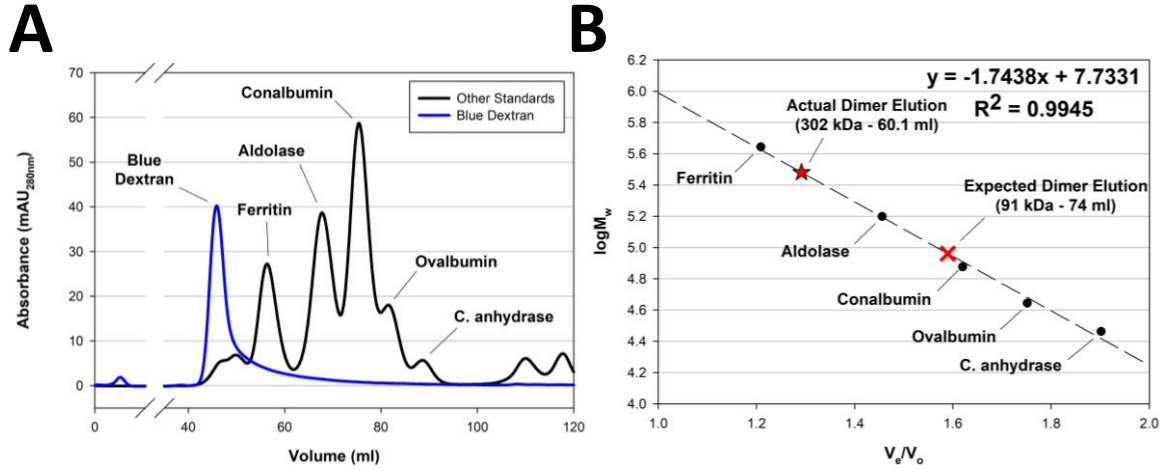


Fig. 23: Size exclusion standards used for NC size exclusion. A – Original chromatograms of the standards. Blue dextran was used to determine the void volume of the column while other standards were used to construct the standard curve. **B** – Corresponding standard curve generated from the chromatogram. The log of the molecular weight of each standard was plotted against its elution volume with respect to the void volume. Also shown are the expected and actual elution times of dimeric NC (red cross and red star, respectively). Data plotted using SigmaPlot V12.

4.1.4 - Nucleocapsid Autolysis

While working with purified NC, it was noticed that extraneous bands in SDS-PAGE would appear over time with a concomitant decrease in the intensity of the main NC band at ~55 kDa. Initially, these bands migrate to ~35 kDa, but with time additional bands at ~25 kDa and finally ~16 kDa appear (**Fig. 24A**). Previous literature with both the SARS-CoV-1 and SARS-CoV-2 NC have found similar results (Lutomski et al., 2021; Mark et al., 2008; Perdikari et al., 2020). Initially, research in cell culture suggested this degradation was due to intracellular proteases. However, this behaviour is also observed with purified NC. Mark et al. (2008) found that the cleavage does not appear to be the result of contaminating proteases as no cleavage of fluorescently labelled casein was observed when the casein was added to the purified NC sample. Similarly, no cleavage of ovalbumin or RNase, which both contain the putative SR motif that is frequently cleaved in NC autolysis, was observed. Instead, autolysis appears to be an innate ability of NC. However, the mechanism remains unknown. Interestingly, peptides corresponding to the regions of NC that are commonly cleaved are not cleaved in the presence of full-length NC (Mark et al., 2008). It has been suggested that autolysis of NC allows for unpacking of the RNA genome upon infection, immune evasion, and separation of the two domains which independently carry out particular functions throughout the infected cell (Lutomski et al., 2021). Indeed, previous mass spectrometry (MS) analysis has found that the majority of NC autolysis products consist of intact NTD and CTD domains (Lutomski et al., 2021). These two domains are separated by an Ser-Arg rich region, which is more prone to degradation than other dipeptide repeats (Guruprasad et al., 1990).

To confirm whether the extraneous bands seen during SDS-PAGE were indeed NC autolysis products, bands were excised and sent for MS sequencing. Bands corresponding to the main ~55 kDa NC band, a ~25 kDa band, an ~18 kDa band, and a ~16 kDa band were excised. Unexpectedly, all three autolysis product bands produced peptides that cover the full-length protein (**Fig. 24B**), suggesting these bands do not represent any particular region of protein. Rather, the bands appear to represent heterogenous mixtures of random autolysis products. Given that a heterogenous mixture is expected to produce a

distribution of differently sized products, one would not expect a distinct SDS-PAGE band. These results can also not be reconciled with previous work that found specific sizes to represent particular proteoforms (Lutomski et al., 2021). Regardless, there does appear to be slight enrichment of peptides originating from either the NTD or CTD domains, possibly indicating that autolysis products do indeed possess intact domains. Given that each domain can function independently, the separated domains are likely still functional despite protein degradation.

SEC was used to successfully purify full-length NC from its autolysis products as previously discussed. Visible autolysis is apparent after only one day following purification, but full-length NC remains predominant until approximately three days post-purification. Freshly purified NC was used for experiments as soon as possible, typically within a day of purification. However, purification of NC from autolysis products immediately prior to each experiment was impractical. Because of this, some degree of autolysis was unavoidable.

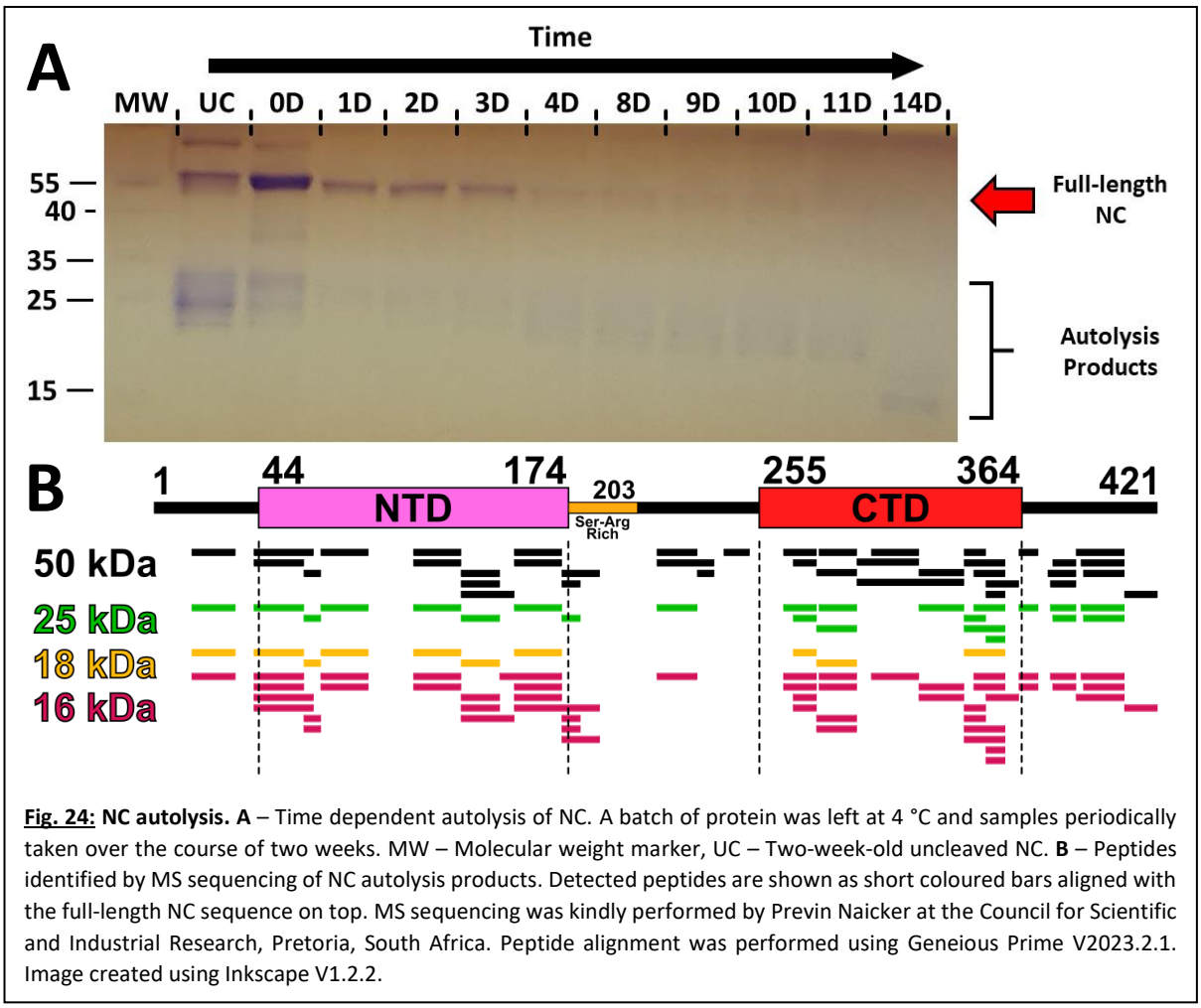


Fig. 24: NC autolysis. A – Time dependent autolysis of NC. A batch of protein was left at 4 °C and samples periodically taken over the course of two weeks. MW – Molecular weight marker, UC – Two-week-old uncleaved NC. **B** – Peptides identified by MS sequencing of NC autolysis products. Detected peptides are shown as short coloured bars aligned with the full-length NC sequence on top. MS sequencing was kindly performed by Previn Naicker at the Council for Scientific and Industrial Research, Pretoria, South Africa. Peptide alignment was performed using Geneious Prime V2023.2.1. Image created using Inkscape V1.2.2.

4.1.5 - FOXP3 FHD Structural Characterisation

To confirm whether the FOXP3 FHD was correctly folded following purification, various techniques were used to probe its structure (**Fig. 25**). CD spectroscopy, used to assess secondary structure, showed a predominantly α-helical spectrum (**Fig. 25A**) corresponding to the secondary structure visible in the crystal structure (**Fig. 25F**) (Bandukwala et al., 2011). The α-helical spectrum is lost upon urea denaturation, instead showing a portion of a disordered spectrum, indicating that the original spectrum represents native folding. Intrinsic tryptophan fluorescence of the FHD, a technique for probing the local environment surrounding tryptophan residues, showed a prominent peak at 342 nm with a shoulder at approximately 330 nm (**Fig. 25B**). There is a small peak at 295 nm indicating scattering of excitation light by aggregation. Upon denaturation, the shoulder is lost and a bathochromic shift is observed as all the tryptophans becomes solvent exposed; indicating the protein is folded and is not in a denatured state following purification. FOXP3 FHD has four tryptophan residues: Trp348, Trp366, Trp381, and Trp407. Trp348 and Trp366 are directly involved in the highly hydrophobic dimer interface (**Fig. 25F**) (Bandukwala et al., 2011); thus the shoulder at 330 nm in the non-denatured spectrum is likely due to these residues.

The absorbance spectrum of FOXP3 FHD is characteristic of tryptophan absorbance; with a peak at 280 nm and a shoulder at 290 nm (**Fig. 25C**) (Edelhoch, 1967). As mentioned, the FHD possesses four tryptophans, along with only two tyrosines. Tryptophan possesses a larger extinction coefficient when compared to tyrosine (Edelhoch, 1967), so the FHD's absorbance spectrum is dominated by its tryptophan residues.

As mentioned previously, the FHD has been found to exist as an obligate dimer. This was confirmed using SEC where the FHD eluted with an apparent size of ~28 kDa (**Fig. 25D**. Standards are given in **Fig. 26A** and standard curve in **26B**), with no visible peak where the monomeric FHD was expected to elute. When fractions from SEC were visualised using SDS-PAGE, the FHD migrated with a size of approximately 14 kDa (**Fig. 25E**). This confirms that the FHD is indeed exclusively a dimer and corresponds with previous literature (Bandukwala et al., 2011). Thus, the secondary, tertiary, and quaternary structure of the FOXP3 FHD was as expected for the native state following purification.

4.1.6 - Nucleocapsid Structural Characterisation

Similarly, NC structural characterisation was performed to confirm correct folding. The NC CD spectrum is characteristic of predominantly disordered proteins (**Fig. 27A**); agreeing with previously published findings (Zeng et al., 2020). Due to absorbance by the urea itself, the urea denatured spectrum reaches the signal noise cutoff at 218 nm, preventing meaningful comparison between the native and denatured spectra. Intrinsic tryptophan fluorescence of NC shows a prominent double peak at 330 nm and 338 nm (**Fig. 27B**). NC has five tryptophan residues; Trp52, Trp108, Trp132, Trp301, and Trp330; all of which are found in either the relatively structured NTD (Trp52, Trp108, Trp132) (**Fig. 27D**) or the CTD (Trp301, Trp330) (**Fig. 27E**). Trp52 and Trp330 are relatively more solvent exposed than the others in the solved structures, possibly contributing to the peak at 338 nm; while the rest contribute to the peak at 330 nm. Upon denaturation with 8 M urea, the double peak is lost and a single peak at 350 nm is seen. Additionally, the intensity of this peak is reduced compared to the native spectrum; potentially indicating quenching upon solvent exposure. This indicates that the protein is folded following purification. Given these data, along with SEC as previously shown, purified NC was considered correctly folded. This was backed up by the observation that NC binds significant bacterial nucleic acids during purification, as previously described, showing NC is still functional at the time of purification.

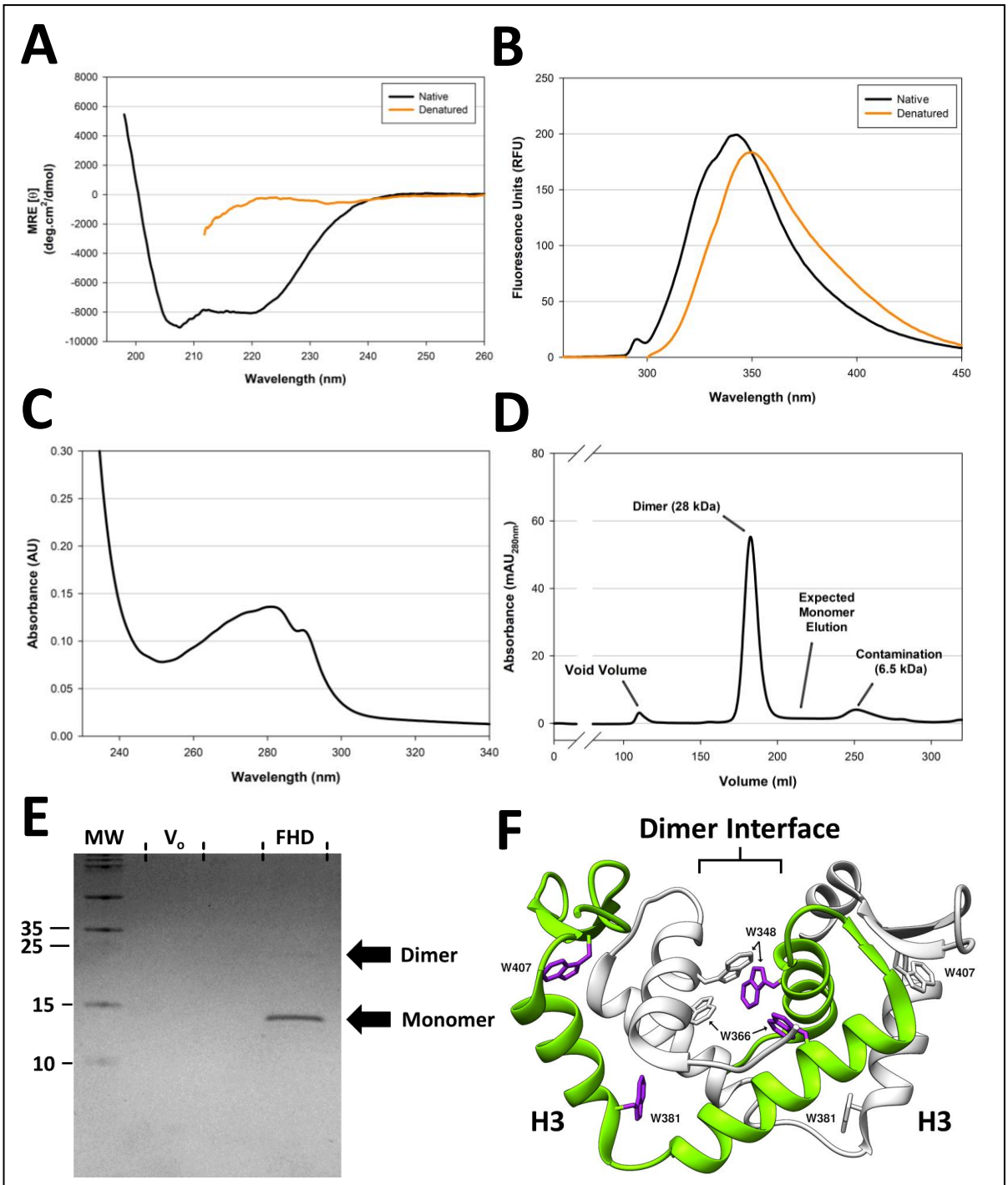


Fig. 25: FOXP3 FHD Structural Characterisation. To confirm that the FOXP3 FHD was folded following purification, various structural characteristics where measured. **A** – Circular dichroism spectrum of native (black) and 8 M urea denatured (orange) FOXP3 FHD. **B** – Intrinsic tryptophan fluorescence of native (black) and 8 M urea denatured (orange) FOXP3 FHD. **C** – Absorbance spectrum of FOXP3 FHD. **D** – SEC of purified FOXP3 FHD. **E** – Corresponding SDS-PAGE of FOXP3 FHD SEC. MW – Molecular weight marker, V_o – Void volume. **F** – FOXP3 FHD crystal structure of the domain swapped dimer (PDB: 3QRF) showing the tryptophan residues present in the FHD. Monomers are coloured green and white. The tryptophans of one monomer are coloured purple while the other white. Trp348 and Trp366 are directly involved in the hydrophobic dimer interface. The DNA binding helix 3 (H3) are also indicated. Data plotted using SigmaPlot V12. Structure visualised using UCSF Chimera (Pettersen et al., 2004).

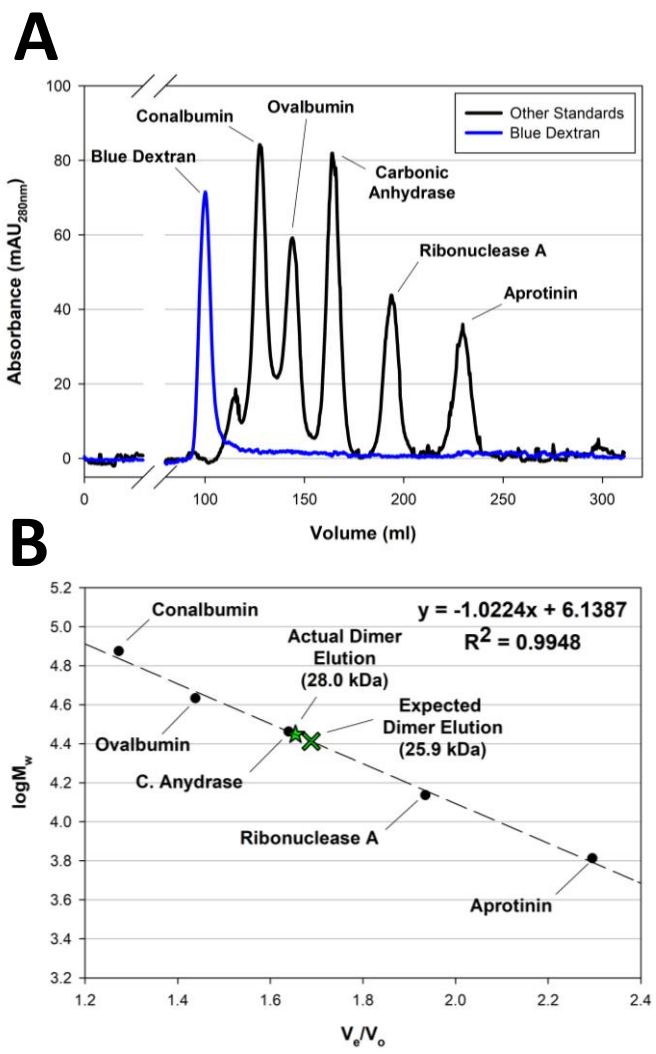


Fig. 26: Size exclusion standards used for FOXP3 FHD size exclusion. **A** – Original chromatograms of the calibration proteins using a HiLoad 26/600 Superdex 75pg column. 1 ml of 1 mg/ml blue dextran (blue) was used to determine the void volume of the column while 1 ml containing 3 – 4 mg/ml of the other standards (black) were used. **B** – Standard curve generated from the chromatogram. The log of the molecular weight of each standard was plotted against its elution volume with respect to the void volume. Linear regression of the resulting plot allowed for determination of the FHD size from its elution volume. Also shown are the expected (green cross) and actual (green star) elution times of the dimeric FOXP3 FHD. Data plotted using SigmaPlot V12.

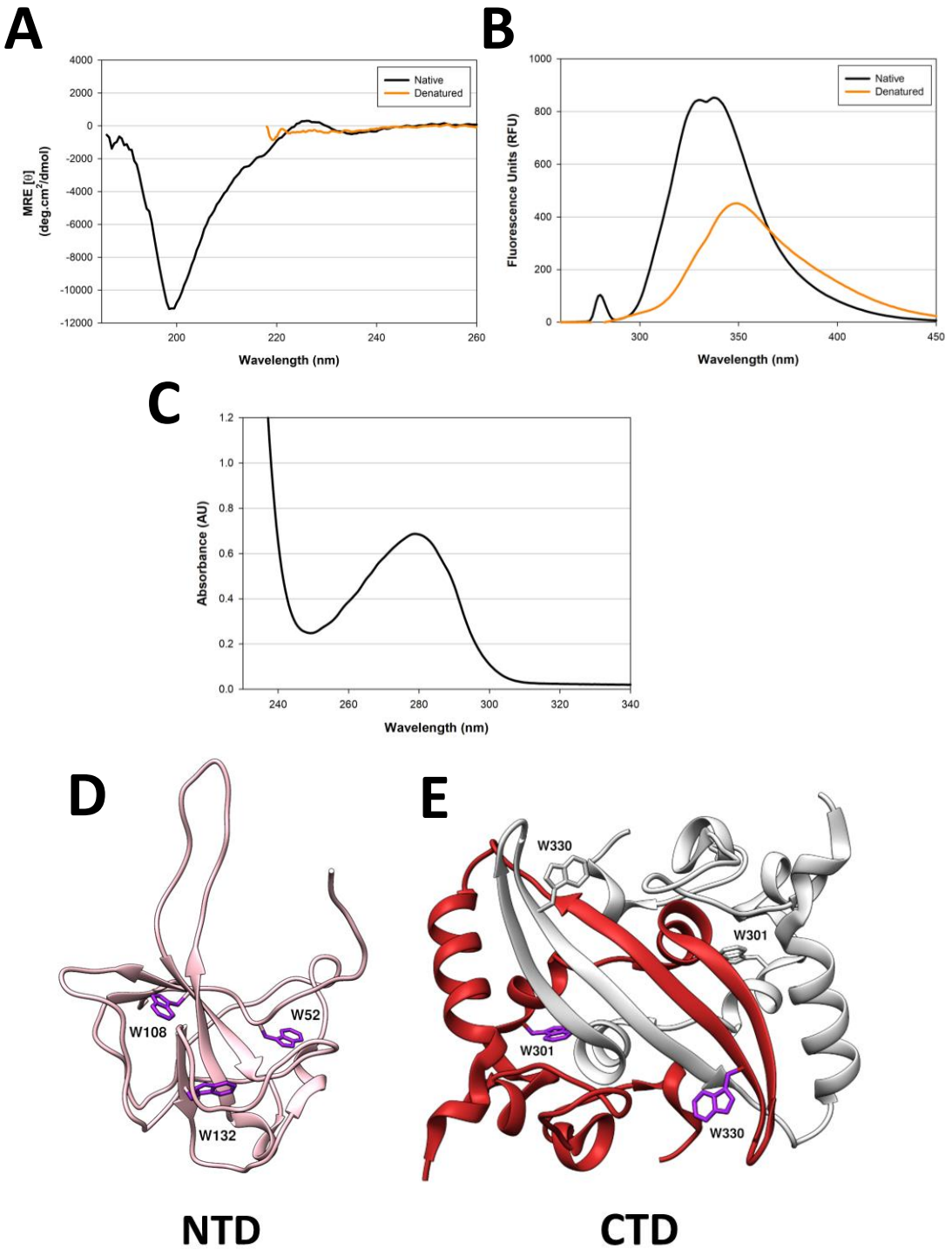


Fig. 27: NC Structural Characterisation. A – Circular dichroism spectrum of native (black) and 8 M urea denatured (orange) NC. **B** – Intrinsic tryptophan fluorescence of native (black) and 8 M urea denatured (orange) NC. **C** – Absorbance spectrum of NC. **D** and **E** – Nuclear Magnetic Resonance (NMR) structure of NTD (left, pink, PDB: 6YI3) and crystal structure of CTD (right, red/white, PDB: 7CE0) of NC showing tryptophans (purple and white). Data plotted using SigmaPlot V12. Structures visualised using UCSF Chimera (Pettersen et al., 2004).

4.2 - Protein-Protein, Protein-DNA, and Protein-Protein-DNA Interactions

Nucleocapsid, in addition to packaging the viral genome, plays an important role in viral replication. By interfering with host proteins, NC interferes with the cell's innate antiviral response to promote replication (Gordon et al., 2020; Lu et al., 2021; Mu et al., 2020; Wang et al., 2021; Zhou et al., 2008). Overactivated platelets are a prominent feature during COVID-19 (Pretorius et al., 2021; Venter et al., 2020), and FOXP3 is important in platelet generation (Bernard et al., 2009). Thus, NC may also be interacting with FOXP3 to bring about these symptoms. NC may interact with the FHD of FOXP3 through direct protein-protein interactions, or interactions may instead be mediated through the presence of DNA. Thus, to understand whether the FHD and NC interact, it is necessary to investigate their interactions both in the absence and presence of DNA. By studying their interactions in the absence of DNA, its possible to determine whether the two proteins interact directly, which may cause abrogation of FHD DNA binding. Studies of their interaction in the presence of DNA can be used to determine whether interactions exist when one or both proteins are DNA bound, and whether these interactions modify the DNA binding behaviour of the FHD.

4.2.1 - FOXP3 FHD-NC Protein-Protein Interaction

Transcription factors rely heavily on protein-protein interactions to exert transcriptional control over the genes they regulate. As such, direct interaction between the FOXP3 FHD and NC was investigated using pulldown assays and fluorescence anisotropy (**Fig. 28**). Pulldown assays revealed a weak interaction between the two proteins (**Fig. 28A**). Despite a large concentration of His-tagged FHD (the bait) bound to the column during the pulldown assay, the majority of untagged NC (the prey) placed on the column did not bind and instead flowed through. A small amount of NC remained on the column and eluted with FOXP3 FHD in the elution. Despite the small amount of NC bound to the FOXP3 FHD, this binding is a true interaction as NC did not bind to the column in the absence of the FOXP3 FHD; as evidenced by the lack of NC bands in the control lanes.

Fluorescence anisotropy revealed similar results, with a steady increase in the anisotropy occurring with increasing NC concentrations (**Fig. 28B**). While the isotherm could be fit (K_D = 25.1 ± 29.73 μM, R^2 = 0.8733), the curve does not reach saturation and thus does not accurately quantify the true binding affinity between these two proteins. Higher concentration titrations, to sample points closer to saturation, were not performed due to the very high concentration of protein required. Regardless, this, along with the pulldown assay, shows that FOXP3 FHD and NC do interact directly in the absence of DNA.

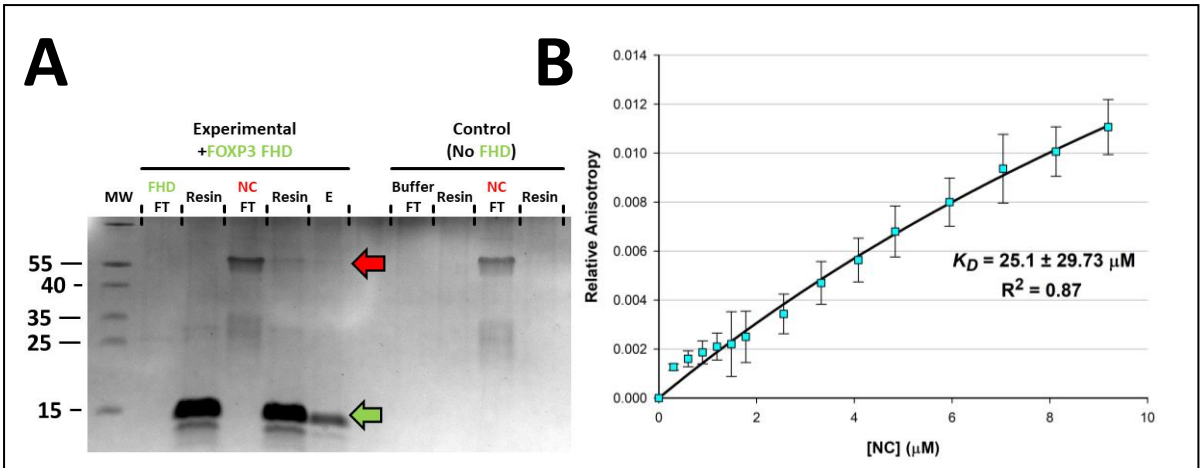


Fig. 28: FOXP3 FHD and NC protein-protein interaction. FOXP3 FHD and NC show weak direct interaction. **A** – Pulldown assay with FOXP3 FHD (green arrow) as the bait and NC (red arrow) as prey. 150 µg of the His-tagged FOXP3 FHD was placed on the column containing a cobalt charged NTA resin. After washing off unbound FHD, 150 µg of NC was added and the column incubated for 1 hour. After washing off unbound NC, proteins were eluted using a 300 mM imidazole solution and analysed using SDS-PAGE. The control was performed similarly but with the omission of the FOXP3 FHD. MW – Molecular weight marker, FT – Flowthrough of the FHD (bait) or NC (prey), Resin – Sample of the column resin following addition of the FHD or NC at respective steps. E - Elution. **B** – Fluorescence anisotropy showing a relatively weak interaction between FOXP3 FHD and NC (K_D > 25 µM). FOXP3 FHD labelled with NTA ATTO 550 was titrated with high concentration NC and the resulting anisotropy measured after each titration. Data fit using a single site ligand binding curve using SigmaPlot V12. Error bars represent standard deviation of triplicate measurements.

4.2.2 - Protein-DNA Interactions

Both FOXP3 and NC bind nucleic acids to perform their respective functions. Thus, it is important to study their nucleic acid binding in isolation to fully understand potential interactions between them. FOXP3 FHD was found to bind the p38α promoter sequence (**Fig. 29**). EMSAs showed that at a DNA:Protein ratio of 1:8 (equivalent to a protein concentration of 4 µM), the majority of DNA is protein bound; with no free DNA remaining at a ratio of 1:20 (**Fig. 29A**). Additionally, EMSAs using different DNA sequences specific for other transcription factors (AAV P5, specific for YY1, and SSL DNA, specific for TBR1), showed that, unlike the p38α sequence, FOXP3 binds only very weakly and non-specifically to these sequences as evidenced by the slight smearing and lack of a distinctly shifted DNA-bound band in the EMSA gel (**Fig. 29B**). The FOXP3 FHD – p38α promoter sequence interaction was also confirmed using fluorescence anisotropy which showed that the FHD binds to the sequence with K_D = 0.14 ± 0.03 µM (R^2 = 0.96) (**Fig. 29C**).

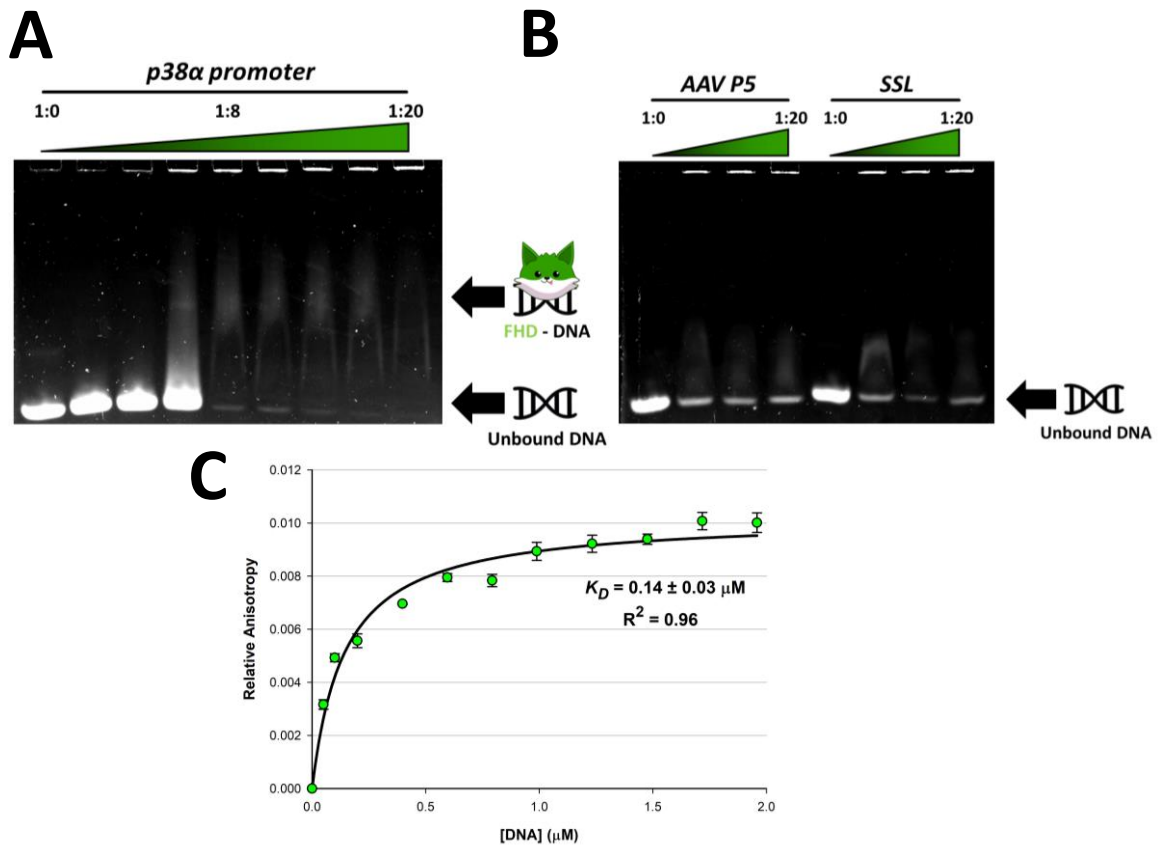


Fig. 29: FOXP3 FHD DNA binding. The FHD binds specifically to the p38α promoter sequence. **A** – EMSA of FOXP3 FHD binding to the p38α promoter sequence. Samples containing 0.5 μM p38α promoter sequence was incubated with increasing FHD concentrations and then electrophoresed on a native polyacrylamide gel. **B** – EMSA of the FHD with DNA sequences not specific for FOXP3, showing non-specific binding. **C** – Fluorescence anisotropy binding isotherm of FOXP3 FHD to the p38α promoter sequence. FOXP3 FHD labelled with NTA ATTO 550 was titrated with high concentration DNA and the resulting anisotropy measured after each titration. Data fit using a single site ligand binding curve using SigmaPlot V12. Error bars represent standard deviation of triplicate measurements.

Surprisingly, NC was also found to bind the p38α promoter sequence using both EMSA and fluorescence anisotropy (**Fig. 30**). It is known that NC is capable of binding both RNA and DNA non-specifically (Zeng et al., 2020; Zhou et al., 2020). This was indeed observed, as evidenced by the smeared "columns" that increase with protein concentration (**Fig. 30A**). However, in addition to non-specific DNA binding, NC also appears to bind to the p38α sequence specifically; leading to the presence of a distinct upward shifted DNA band. Compared to the FOXP3 FHD, the NC – DNA binding interaction appears to be lower affinity as the DNA does not become fully bound even at a DNA:Protein ratio of 1:40. This band is only weakly present with the AAV P5 and SSL sequences; while the non-specific binding remains unchanged (**Fig. 30B** and **30C**, respectively). Fluorescence anisotropy using ROX labelled p38α promoter shows that NC binds the sequence with moderate affinity (K_D = 1.84 ± 0.22 μM, R^2 = 0.98) (**Fig. 30D**).

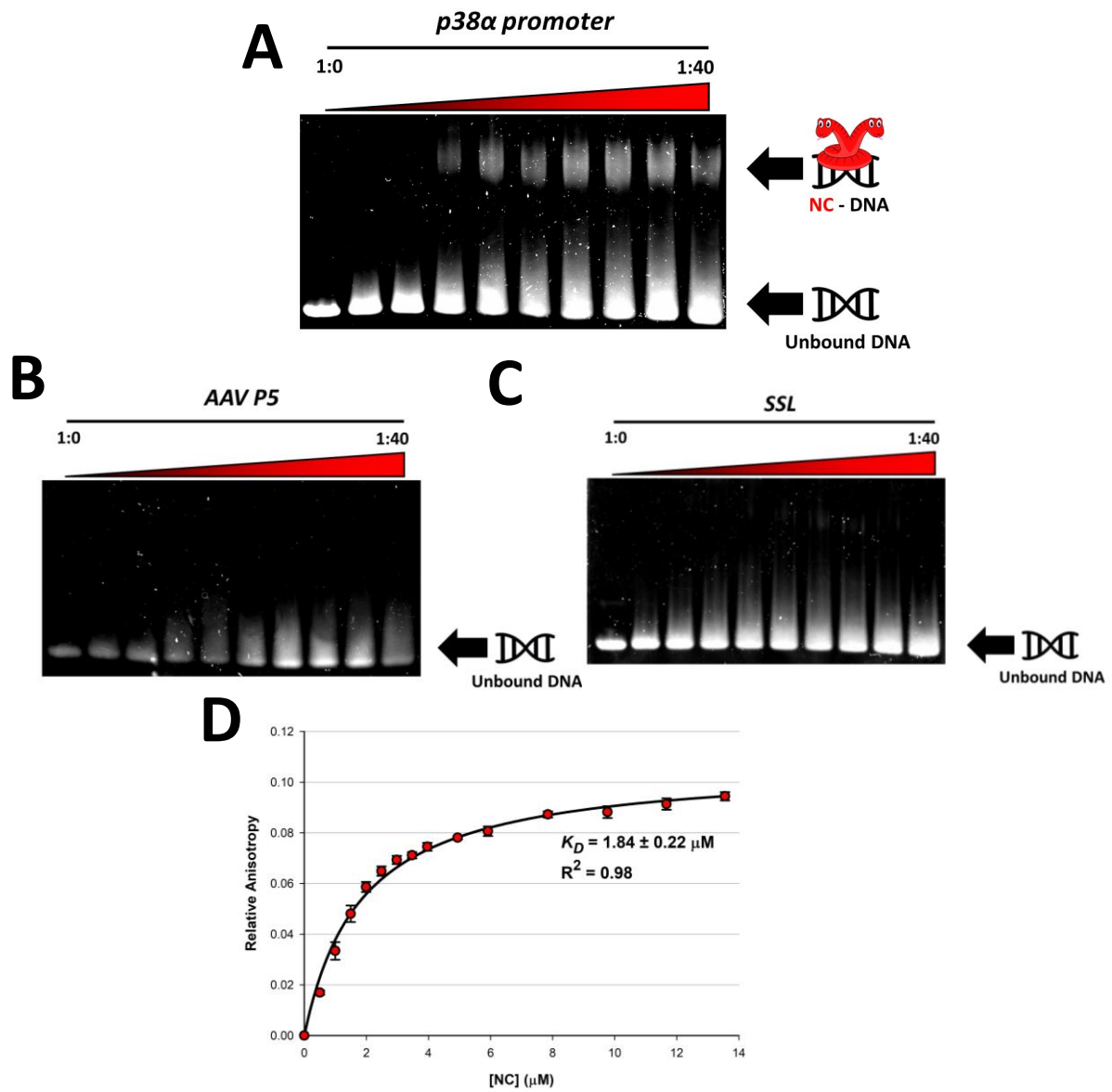


Fig. 30: NC DNA binding. NC binds DNA non-specifically, but also shows specific binding in the case of the p38α promoter sequence. **A** – EMSA showing NC binding to the p38α promoter sequence. Samples containing 0.5 μM p38α promoter sequence was incubated with increasing NC concentrations and then electrophoresed on a native polyacrylamide gel. **B** and **C** – NC EMSAs to non-FOXP3 specific sequences, AAV P5 and SSL, showing only non-specific smearing with no specific binding. **D** – Fluorescence anisotropy of NC binding to the p38α promoter sequence, showing that NC binds the sequence with moderate affinity. ROX labelled p38α promoter was titrated with high concentration NC and the resulting anisotropy measured after each titration. Data fit using a single site ligand binding curve using SigmaPlot V12. Error bars represent standard deviation of triplicate measurements.

4.2.3 - Protein-Protein-DNA Interactions

Protein-protein-DNA interactions play a major role in transcriptional regulation. The FOXP3 FHD and NC have been shown to interact directly, and can specifically bind the p38α promoter independently. However, the presence of protein-protein interactions may influence protein-DNA interactions, and vice-versa. As such, formation of a protein-protein-DNA ternary complex was assessed using fluorescence anisotropy. p38α promoter was titrated into 0.5 µM NTA ATTO 550 labelled FOXP3 FHD in the presence of 0.5 µM NC to assess whether the two proteins could bind the DNA simultaneously. Since the concentration of protein is low, only a small proportion of protein is expected to exist as an FHD-NC complex while the majority remains as free protein. Titration resulted in a two-phase curve (**Fig. 31**). Initially, there is a sharp increase in the anisotropy that appears to follow the same binding pattern as NC-DNA. The large initial increase in the anisotropy indicates that both the FHD and NC are binding to the DNA simultaneously as the relatively small FHD becomes part of a much larger FHD-NC-DNA complex. However, at DNA concentrations above 0.5 µM there is a decrease in the anisotropy that eventually more closely resembles the binding behaviour of FHD-DNA. This indicates that once there is excess DNA present in the cuvette compared to the proteins, the FHD goes from participating in a large FHD-NC-DNA complex to a smaller FHD-DNA complex as the proteins dissociate and bind separate DNA oligonucleotides. Since it is the FHD that is labelled, the binding curve then resembles that of FHD-DNA binding.

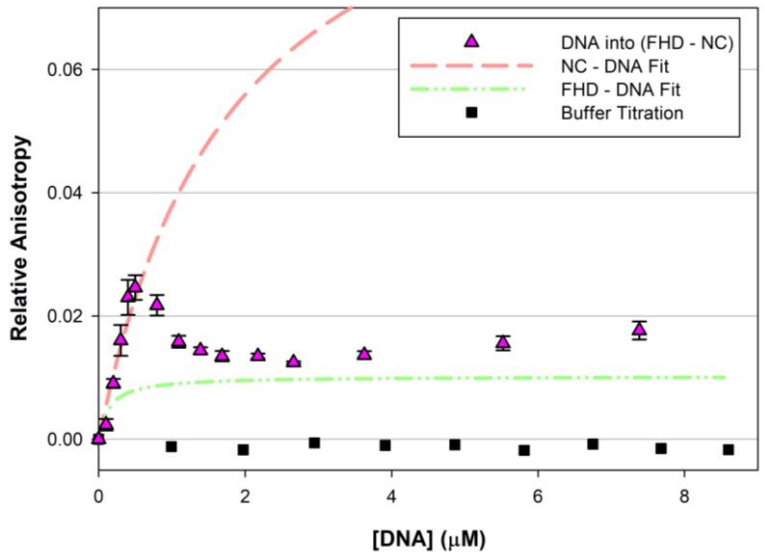


Fig. 31: Fluorescence anisotropy of the protein-protein-DNA interactions between FOXP3 FHD, NC, and the p38α promoter sequence. NTA ATTO 550 labelled FHD in the presence of 0.5 µM NC was titrated with p38α promoter sequence (pink triangles). The FHD and NC can bind to the p38α promoter simultaneously when protein concentrations are greater than the DNA concentration ([DNA] < 0.5 µM). Once the concentration of DNA increases beyond the protein concentration, each protein binds separate oligonucleotides. Also shown are the previous NC-DNA (red line) and FHD-DNA (green line) fits for comparison. A buffer titration (black squares) shows that the decreasing anisotropy values are not due to label dissociation. Data fit using a single site ligand binding curve using SigmaPlot V12. Error bars indicate standard deviation of triplicate measurements.

To study the effect of increasing NC on already DNA bound FHD, NC was titrated into NTA ATTO 550 labelled FHD in the presence of saturating (1 µM or 5 µM) p38α promoter DNA. Concentrations were chosen to ensure saturation based on the FHD-DNA K_D shown previously (**Fig. 29**) (K_D = 0.14 ± 0.03 µM). The FHD-DNA mixture was allowed to incubate for one hour at room temperature prior to measurements to allow for binding to occur. When in the presence of 5 µM DNA, a clear binding isotherm was produced (**Fig. 32**, purple circles), as NC binds to the FHD-DNA complex to form a larger FHD-NC-DNA complex. Additionally, fitting this curve reveals that NC binds to the FHD-DNA complex with approximately the same K_D compared to when no FHD is present (K_D = 1.92 ± 0.49 µM vs K_D = 1.84 ± 0.22 µM, respectively). This indicates that not only can NC bind to the sequence in both the presence and absence of FHD, but that the FHD does not significantly alter the affinity of NC for the DNA. In the presence of lower DNA concentrations (1 µM), there is an initial increase in the anisotropy prior to ~1.2 µM NC. The initial increase approximately follows that of the 5 µM DNA binding isotherm, indicating similar binding behaviour as a larger FHD-NC-DNA complex is formed. However, following titration of 1.2 µM NC, there is a decrease in the anisotropy which indicates FHD dissociates from the FHD-NC-DNA complex. These results suggest that when NC is not in excess, FHD and NC can bind to the DNA simultaneously; in agreement with the results shown in **Fig. 31**. However, once NC is in excess, NC successfully outcompetes FHD for DNA binding, leading to FHD dissociation from the DNA. The FHD then possibly binds free NC, forming FHD-NC complexes. Given the weak association between the FHD and NC (**Fig. 28**), these complexes are not in abundance at the concentrations used in the experiment and so do not greatly raise the anisotropy. Interestingly, FHD dissociation from the DNA does not appear to begin immediately once the NC concentration matches that of the DNA. Instead, dissociation of the FHD in the 1 µM curve only begins at approximately 1.2 µM NC; suggesting a slight excess of NC is required to compete with the FHD for DNA binding. Indeed, when in the presence of 5 µM DNA, FHD dissociation is not readily apparent in the concentrations tested. Regardless, points lying beyond a concentration of 5 µM NC in the 5 µM DNA plot were excluded during curve fitting, as FHD dissociation may be present.

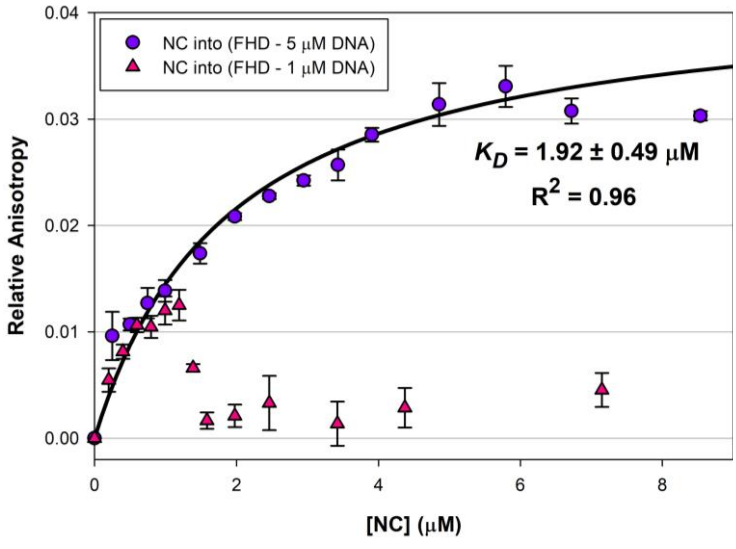


Fig. 32: Effects of increasing NC on preformed FHD-DNA complexes. NTA ATTO 550 labelled FHD in the presence of either 5 µM (circles, purple) or 1 µM (triangles, pink) p38α DNA was titrated with NC. The last three points of the 5 µM DNA plot were excluded from data fitting. Data fit using a single site ligand binding curve using SigmaPlot V12.Error bars indicate standard deviation of triplicate measurements.

The NTD of NC is thought to provide some level of binding specificity (Iserman et al., 2020; Redzic et al., 2021). Thus, protein-DNA computational docking was performed between the NTD of NC and the p38α promoter oligonucleotide using either PyDockDNA (Rodríguez-Lumbreras et al., 2022), MDockPP (Huang & Zou, 2010), or HDOCK (Yan et al., 2020) to see if it could shed light on how NC binds to the sequence (**Fig. 33**). The packaging signal, a short sequence present in the p38α promoter that is discussed in-depth later, is separate from the FOXP3 consensus sequence, and various docking poses showed the NTD binding on or near the packaging signal. In many poses, this left the FOXP3 consensus site free. This could be the mechanism that allows both the FHD and NC to bind to the sequence simultaneously, as NC binds to the packaging signal and the FHD binds to the FOXP3 consensus site. However, many other poses also showed the NTD binding to other regions of the sequence. Additionally, the docking results do not fully replicate how NTD is expected to bind nucleic acids, where Tyr86, Tyr109, and Tyr111 are thought to intercalate with the bases via π-π stacking (Chang et al., 2016). Unfortunately no structure of the NTD bound to double stranded nucleic acid has been solved to allow comparison. Thus, these docking results only suggest, rather than prove, that the NTD may bind to the packaging signal present in the p38α promoter sequence.

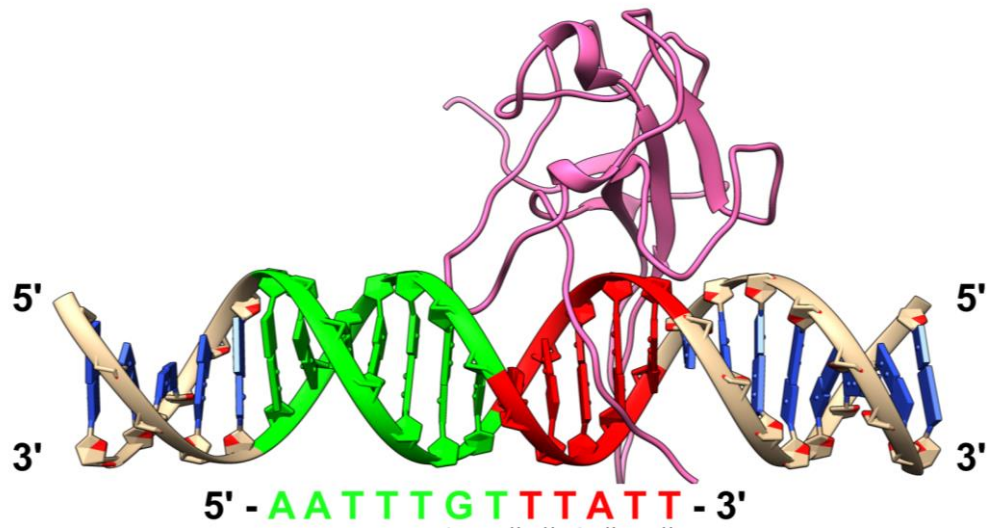


Fig. 33: Representative docked structure between the NTD of NC (pink) and the p38α promoter sequence. The NMR structure of the NTD (PDB: 6YI3) was docked with p38α promoter sequence (created using PyMOL Builder Tool) using either PyDockDNA, MDockPP, or HDOCK. In many poses, the NTD docks on or near the packaging signal (red), which leaves the FOXP3 consensus (green) open. Visualised using UCSF Chimera (Pettersen et al., 2004).

The fluorescence anisotropy results, relevant K_D and B_{max} values, and significant findings are summarised in **Table 5**. Both the FOXP3 FHD and NC can bind to the p38α promoter sequence independently. Additionally, both proteins can bind to the DNA sequence at the same time, and NC binds to the sequence with the same affinity regardless of whether the FHD is present or not. However, excess NC leads to dissociation of the FHD. If this interaction is present in the body, it may lead to FOXP3 dysregulation as FOXP3 is prevented from reaching its target sites in the genome, leading to transcriptional dysregulation.

Table 5: Overall fluorescence anisotropy results. K_D, B_{max}, R^2, and relevant findings/notes are given where applicable.

Experiment	K_D (µM)	B_{max}	R^2	Findings/notes
FOXP3 FHD – *p38α*	0.14 ± 0.03	0.0102 ± 0.0004	0.96	-
NC – *p38α*	1.84 ± 0.22	0.1073 ± 0.004	0.98	-
FOXP3 FHD – NC	> 25	N/A	N/A	Inaccurate fit due to weak binding
p38α DNA into (FHD – NC) complex	N/A	N/A	N/A	Both proteins bind when DNA is limiting. Bind separately when DNA in excess
NC into (FHD – *p38α* DNA) complex				Both bind when NC is limiting, but excess NC leads to competition and FHD dissociation
- 5 µM *p38α*	1.92 ± 0.49	0.0424 ± 0.005	0.96	
- 1 µM *p38α*	N/A	N/A	N/A	

5 - Discussion

While the worst of the COVID-19 pandemic is behind us, SARS-CoV-2 is predicted to become endemic and circulate seasonally much like influenza (Telenti et al., 2021). Thus, understanding how COVID-19 sometimes leads to severe disease is crucial for managing future SARS-CoV-2 outbreaks.

One of the hallmarks of COVID-19 is Disseminated Intravascular Coagulopathy (DIC) (Kander, 2020; Levi et al., 2020; Yuriditsky et al., 2020). In severe COVID-19, DIC can lead to large thrombi that may lead to strokes and embolisms. Furthermore, even in minor or asymptomatic COVID-19, the formation of small but long-lasting clots is thought to be a contributing factor to long-COVID (Pretorius et al., 2021). A common finding in COVID-19 patients is the presence of structurally and functionally abnormal platelets that resemble platelets seen in those suffering from IPEX; an autoimmune condition arising from loss of function mutations in the FOXP3 transcription factor (Bernard et al., 2009; Venter et al., 2020). Thus, COVID-19 may be causing FOXP3 dysfunction that ultimately results in platelet dysregulation and resulting DIC.

SARS-CoV-2 possesses the multifunctional nucleocapsid protein (NC) which, alongside its ability to bind and package the viral genome, performs various roles within the cell to facilitate successful viral replication. One such function may be a direct alteration of host expression caused by NC binding to host promoter regions (Yan et al., 2006). Due to NC's multifunctional abilities, and the ability of the highly similar SARS-CoV-1 NC to directly alter host expression, it would likely be this protein that would alter FOXP3 expression. By investigating if and how NC directly interacts with FOXP3, a better understanding of COVID-19 pathophysiology can be attained which can be used to guide future COVID-19 therapeutics. Additionally, it will provide potential insight into FOXP3 regulation of platelet development and possibly FOXP3 function in general.

5.1 - FOXP3 FHD – Nucleocapsid Interactions

Protein-protein interactions are critically important in gene regulation; where the formation of a complex, even in the absence of DNA, may promote or abolish expression of a particular gene (Voet et al., 2016). Because of this, it is important to determine whether FOXP3 FHD and NC interact directly in the absence of DNA. NC has been found to bind a number of host proteins such as hnRNP C, G3BP1, G3BP2, and other RNA processing proteins (Cascarina & Ross, 2020; Gordon et al., 2020; Luo et al., 2005; Savastano et al., 2020), which it does to promote replication of the viral genome. Other SARS-CoV-2 proteins also target other cellular proteins for similar reasons (Gordon et al., 2020; Kee et al., 2022; Zhou et al., 2021). Thus, there is precedent for NC to bind host proteins.

In this work, pulldown assays and fluorescence anisotropy revealed a weak (K_D > 25 μM) interaction between FOXP3 FHD and NC (**Fig. 28**). Weak transient protein-protein interactions, referred to as the quinary structure of proteins (McConkey, 1982), form a critically important part in cellular processes such as metabolism, signalling pathways, and expression. While weak, these interactions are often numerous and therefore exert a profound effect on cellular processes (Nooren & Thornton, 2003; Wirth & Gruebele, 2013). Weak protein-protein interactions are also common during assembly of some viral capsids (Tellinghuisen et al., 1999) where, due to hysteresis, the resulting protein-protein complexes are more stable than one would predict from the initial interactions alone. They are a central part of liquid-liquid phase separation (LLPS) which itself is involved in a variety of cellular

processes (Uversky, 2019; Wirth & Gruebele, 2013). NC has been found to undergo LLPS in a concentration and RNA dependent manner (Cascarina & Ross, 2020; Iserman et al., 2020; Perdikari et al., 2020; Redzic et al., 2021; Savastano et al., 2020). In the cell, the resulting liquid droplets locally concentrate RNA and RNA processing proteins, such as RNA dependent RNA polymerase, to promote replication of the viral genome (Savastano et al., 2020). Thus, the presence of a weak interaction between FOXP3 FHD and NC may be indicative of potential interactions *in vivo*, possibly involving NC's ability to phase separate and locally concentrate particular proteins. While simple assays for studying LLPS are available (Perdikari et al., 2020), they could not be performed as part of this work due to time constraints. These assays, as part of future work, may provide valuable insight into whether the FHD of FOXP3 becomes incorporated into NC phase-separated droplets.

Additionally, the crowded environment of the cell typically results in stronger interactions than those seen in relatively dilute *in vitro* conditions (Wirth & Gruebele, 2013). Thus, an interaction between NC and FOXP3 *in vivo* may be stronger than that seen here. Typically, NC is largely excluded from the nucleus of infected cells (Iserman et al., 2020). However, FOXP3 is only localised to the nucleus following translation in the cytoplasm (Douglass et al., 2012; Hancock & Özkaynak, 2009), a process which is regulated by post-translational ubiquitination and by FOXP3's nuclear localisation signal (Ni et al., 2019). Thus, there is potential for FOXP3 and NC to interact *in vivo*; particularly during pro-inflammatory states where FOXP3 nuclear localisation is downregulated (Deng et al., 2019). Future *in vivo* protein localisation assays would similarly provide insight into potential direct protein-protein interactions.

5.2 - Protein – DNA Interactions

5.2.1 - FOXP3 FHD

Understanding the DNA binding ability of either protein in isolation is important to allow for comparison when both proteins are studied together. As a transcription factor, FOXP3 binds DNA to regulate the expression of hundreds of genes involved in a wide range of processes such as signal transduction, expression, and RNA processing; including some MAPK pathway proteins (Marson et al., 2007; Zheng et al., 2007). FOXP3 has been found to regulate p38α as well as the upstream MAP3K TAK1 (Lachmann et al., 2010; Rouillard et al., 2016). p38 MAPK signalling plays an important role in a variety of processes such as macrophage activation (Grimes & Grimes, 2020), T cell and Treg function (Huang et al., 2013; Sadeghifard et al., 2017), and platelet activity (Minuz et al., 2002), and is significantly elevated in severe COVID-19 (Manne et al., 2020). Thus, understanding FOXP3 DNA binding, and how it may be disrupted, is important in understanding how these processes are regulated.

In this work, the FOXP3 FHD was found to bind the p38α promoter sequence with moderate-to-high affinity (**Fig. 29**, K_D = 0.14 ± 0.03 μM). FOXP3 literature is scarce in quantitative FHD – DNA binding affinity; rather primarily focusing on relative changes in DNA binding caused by respective experimental conditions. The only example of quantitative FOXP3 FHD-DNA binding found in literature reported a K_D of 4.7 μM between the FHD and the Daf-16 family-binding element 2 (DBE2), a 21-mer oligonucleotide commonly used to study FOX transcription factors (Furuyama et al., 2000), using isothermal titration calorimetry (Li et al., 2017). The discrepancy between the affinity obtained in this work (K_D = 0.14 ± 0.03) and between the FHD and DBE2 may lie in differing conditions; where their

binding studies were performed at 250 mM salt while here they were performed at 150 mM salt. Increasing salt concentrations typically leads to decreased affinity between protein and DNA as the salt begins to outcompete the protein for DNA binding through electrostatic interactions (Lohman, 1986). Other differences may be caused by the differing flanking regions of the oligonucleotide used. The wing region of the FHD augments DNA binding affinity through interaction with flanking bases outside of the consensus sequence (Dai et al., 2021). Additionally, particular DNA sequences may induce DNA conformational changes that also alter the affinity of the FHD for DNA (Dai et al., 2021). The non-specific binding of the FHD to non-consensus sequences, such as SSL and AAV P5, is likely the result of non-specific electrostatic interactions between the positively charged FHD (theoretical pI = 10.42) and the negatively charged DNA backbone. Indeed, transcription factors are believed to use these weak non-specific electrostatic interactions to quickly scan the genome for their consensus binding sites (Suter, 2020).

5.2.2 - SARS-CoV-2 NC

NC is primarily an RNA binding protein (Bai et al., 2021; Chen et al., 2007), but has also been found to bind DNA (Zeng et al., 2020; Zhou et al., 2020). These works have shown that NC binds DNA non-specifically; driven primarily through electrostatic interactions. Non-specific binding was indeed prominent in NC EMSAs (**Fig. 30**), visible as a smeared tail above the unbound DNA band. This non-specific binding appears uniform across all tested DNA sequences, in accordance with NC's ability to non-specifically interact with DNA. Unexpectedly, NC was found to bind the p38α promoter sequence specifically in addition to non-specifically, as evidenced by the shifted DNA bands closer to the top of the gel. While NC predominantly binds nucleic acids non-specifically, it does show a particular preference for the RNA sequence 5'-AAUAA-3' (Redzic et al., 2021), which the 5' untranslated region (UTR) of the viral genome is particularly rich in. This sequence acts as a packaging signal that, along with similar packaging signals in the 3' UTR, bind NC to drive packaging of the viral genome during virion assembly. The p38α promoter sequence possesses the DNA equivalent of this sequence (5'-GTTATAATTTGT**TTATT**GTCAATTG-3', where the packaging signal is bolded and the FOXP3 consensus underlined) while the other tested sequences do not; which may explain why NC displays specific binding for this sequence. This sequence also partially overlaps that of the FOXP3 consensus sequence. Initially, this was thought to indicate that NC and FOXP3 FHD may compete for DNA binding. However, as discussed in **Section 5.3**, this is not the case.

Interestingly, the work that found that the SARS-CoV-1 NC directly binds to DNA and alters host expression also observed specific binding during EMSAs (Yan et al., 2006). However, the oligonucleotides in their study possessed neither a packaging signal nor any similarity to the p38α promoter sequence as determined by BLAST analysis. This suggests that there may be different sequences that NC may differentially recognise and bind to.

To further characterise the interaction between NC and the p38α DNA, fluorescence anisotropy was performed. NC binds to the p38α promoter sequence with moderate affinity (K_D = 1.84 ± 0.22 μM). Unfortunately, comparison with previous literature is difficult, as previously published NC affinity values vary over several orders of magnitude. Some previous works have reported affinities in the nanomolar range (Wu et al., 2021; Zeng et al., 2020), others in the tens of micromolar (Chang et al., 2009; Dinesh et al., 2020; Korn et al., 2023; Redzic et al., 2021), and yet more in the hundreds of micromolar up into the millimolar range (Redzic et al., 2021). Such differences may partly be the result of differing techniques used to obtain K_D values. These papers used a combination of EMSAs, NMR, ITC, and fluorescence anisotropy to quantify NC nucleic acid binding. Another large factor is the nucleic

acid itself. The affinity of NC for a particular nucleic acid sequence appears to be heavily influenced by a number of factors such as its sequence, the type of nucleic acid, whether it possesses secondary structure, and its rigidity (Redzic et al., 2021). Sequences containing packaging signals and stem-loops are typically bound by NC with greater affinity than those without (Chen et al., 2007; Iserman et al., 2020; Redzic et al., 2021). The p38α promoter oligonucleotide is not predicted to form any particular secondary structures, and thus it may be the packaging signal mentioned previously which drives specific binding. The p38α promoter sequence used in this work was taken directly from the promoter of the p38α gene, and since NC can bind to the sequence *in vitro*, it may bind the sequence *in vivo* and potentially disrupt transcriptional regulation. Additionally, if NC is recognising the 5' – AATAA – 3' packaging signal, FOXP3 may not be the only transcription factor dysregulated by NC. NC may bind to any sequence containing a packaging signal and potentially disrupt any nearby transcription factors. As mentioned previously, the SARS-CoV-1 NC, which shares 90% sequence identity with the SARS-CoV-2 NC (Peng et al., 2020), has been found to directly bind host promoters and alter gene expression (Yan et al., 2006). Thus, this function is likely preserved in the SARS-CoV-2 NC. Future work in cell culture could be used to verify if the SARS-CoV-2 NC does indeed bind to genomic DNA in a more physiologically relevant environment. Determining which sequence(s) the NC preferentially binds to would also help determine which transcription factors are disrupted during COVID-19.

5.3 - Protein – Protein – DNA Interactions

The FOXP3 FHD and SARS-CoV-2 NC interact directly with each other. Additionally, they both bind to the same DNA oligonucleotide separately, raising the question of whether they could bind simultaneously or compete for binding. Interestingly, both the FHD and NC simultaneously bind to the DNA to form an FHD-NC-DNA complex (**Fig. 31**). At the protein concentrations used (0.5 μM for both proteins), the proteins are expected to be predominantly unbound from each other prior to the addition of DNA given the K_D of the FHD-NC interaction (K_D > 25 μM, **Fig. 28**). The large increase in the anisotropy compared to that of FHD-DNA binding (**Fig. 29**) indicates that binding is not the result of only FHD binding to the DNA. Additionally, this increase only occurs after the addition of DNA, indicating this interaction is mediated by the DNA itself. Then, as more DNA is added, the FHD dissociates from the FHD-NC-DNA complex, leading to a decrease in the anisotropy. The binding curve then more closely follows that of the FHD-DNA binding curve, indicating that the FHD is now part of an FHD-DNA complex. Given the excess of DNA in solution after dissociation, NC is likely bound to its own DNA as it was shown to be able to bind the p38α promoter DNA (**Fig. 30**). The packaging signal that NC may be recognising does partially overlap the FOXP3 consensus site. Since the two proteins are, at least initially, binding the DNA simultaneously, this indicates that either the NC binds to the packaging signal in the minor groove of the DNA; which remains free as the FHD binds the major groove (Bandukwala et al., 2011); or NC is recognising some other sequence. Computational docking between the NTD of NC and the p38α promoter sequence did have poses where the NTD bound near or on the packaging signal (**Fig. 33**), possibly supporting the idea that the FOXP3 consensus remains free upon NC binding. Unfortunately, no structures of NC binding to double stranded nucleic acid are available to compare with to confirm whether NC does indeed bind to the minor groove. The dissociation of the two proteins in excess DNA may be the result of the highly dynamic nature of IDP interactions, where many transient interactions allow the protein to quickly change binding partners, leading to quick dissociation and separate binding once excess DNA is available (Cubuk et al., 2021; Uversky, 2019).

Competition for DNA binding is readily apparent when NC is instead titrated into a preformed FHD-DNA complex. In these experiments, DNA acts as a limited resource and so the proteins are more likely

to either form FHD-NC-DNA complexes or compete with each other rather than bind DNA separately. When only 1 μM DNA is present and NC is titrated in, both proteins initially bind to the DNA to form an FHD-NC-DNA complex (**Fig. 32**). Once the concentration of NC exceeds the concentration of DNA, however, the decrease in the anisotropy reveals that the FHD begins to dissociate; indicating that NC only begins to outcompete the FHD once free DNA binding sites are depleted. Indeed, when the concentration of DNA is increased to 5 μM, dissociation is not seen in the concentrations tested; as there are more free DNA binding sites for NC to bind to. Additionally, NC binds to the DNA with similar affinity whether the FHD is bound or not (K_D = 1.92 ± 0.49 and 1.84 ± 0.22, respectively). Once the FHD has dissociated, it likely forms FHD-NC complexes with free NC like that seen in **Fig. 28**, resulting in the anisotropy remaining positive despite dissociation of the larger FHD-DNA complex. While NC is primarily localised to the cytoplasm (Iserman et al., 2020; Surjit et al., 2005), NC is present in smaller amounts in the nucleus (Gao et al., 2021; Kumar et al., 2020; Surjit et al., 2005); potentially also entering the nucleolus (Wurm et al., 2001). Thus, NC is present in the nucleus to potentially compete with FOXP3. NC is the most abundant viral protein present in an infected cell (Lu et al., 2021) and may be present in sufficient amounts to compete with FOXP3 for DNA binding; causing dysregulation of FOXP3 target genes. Alternatively, even with low amounts of NC, NC may disrupt FOXP3 transcriptional control. During the initiation of transcription, transcription factors often work together to bring about assembly of the transcription initiator complex (TIC). Multiple transcription factors bound to a promoter, even with no direct interaction between them, may be required to act as a foundation for TIC assembly (Brown, 2002; Voet et al., 2016). Thus, NC binding may physically block interactions necessary for TIC assembly; either by blocking FOXP3 binding partners or by blocking the binding of other transcription factors to the DNA.

FOXP3 plays several important roles. In haemostasis, FOXP3 is important for the proper development of platelets. Platelets produced in patients with loss of function FOXP3 mutations are morphologically and functionally similar to those seen during COVID-19 (Bernard et al., 2009; Venter et al., 2020). NC interference with FOXP3 activity may be the cause of these abnormally activated platelets; which in turn may be contributing to the DIC seen in COVID-19. NC forms direct protein-protein interactions with the FOXP3 FHD (**Fig. 28**). In the cell, these interactions may sequester FOXP3 in the cytoplasm, where NC concentrations are greatest, preventing FOXP3 from entering the nucleus. Alternatively, NC present in the nucleus may bind directly to FOXP3, altering its transcriptional regulating abilities. NC in the nucleus may otherwise bind near FOXP3 on the DNA (**Fig. 30**, **31**, and **32**). This may disrupt FOXP3 function by sterically hindering FOXP3 binding partners, preventing formation of the TIC and subsequent protein expression. Alternatively, if sufficient in great enough quantities in the nucleus, NC may compete with FOXP3 for DNA binding (**Fig. 32**), leading to complete FOXP3 dissociation from the DNA and loss of transcriptional regulation. FOXP3 regulates the expression of TAK1 and p38α (Latchman, 1997; Rouillard et al., 2016), which are important members of the p38 MAPK signalling pathway involved in platelet activation (Minuz et al., 2002). This work in particular used a sequence from the p38α promoter itself to show that the NC of SARS-CoV-2 is capable of competing with the FHD of FOXP3 for DNA binding *in vitro*. Like much of the FOX family, FOXP3 is largely a repressor of transcription (Dai et al., 2021; Ono, 2020). Disruption of FOXP3 would therefore likely lead to increased expression of these proteins. Increased expression of TAK1 in particular has been shown to increase p38 MAPK activity (Moriguchi et al., 1996). Dysregulation of FOXP3 by NC through any of the above mechanisms may therefore lead to increased MAPK activity through elevated TAK1 and p38α expression, resulting in excess platelet activation. The overly stimulated platelets may then lead to greater clot formation (Venter et al., 2020), increasing the risk of strokes and embolisms during COVID-19. Thus, through NC, both via direct protein-protein and protein-protein-DNA interactions, SARS-CoV-2 may be dysregulating FOXP3, leading to the coagulopathy seen during COVID-19.

FOXP3 is also important in other processes and hence disruption of FOXP3 may have wider reaching effects during COVID-19. Notably, FOXP3 is best-known for its role in immune regulation, where it controls Treg differentiation (Bettelli et al., 2006; Ono, 2020). It has been found that SARS-CoV-2 can infect T cells via an ACE-2 independent mechanism, with a slight bias for $CD4^+$ T cells (Shen et al., 2022), which Tregs do express (Ono, 2020). A second prominent hallmark of COVID-19 is severe systemic inflammation. In particular, in severe COVID-19, there is depletion of circulating Tregs which ultimately leads to severe inflammation by both activated and autoreactive T cells (Wang et al., 2020). While this depletion is thought to be caused in part by secretion of pro-inflammatory cytokines that suppress FOXP3 function (Kalfaoglu et al., 2021), NC dysregulation of FOXP3, through the previously mentioned mechanisms, in SARS-CoV-2 infected Tregs may be a confounding factor that further promotes inflammation during COVID-19. FOXP3 is also expressed in a number of epithelial cells, where it primarily acts as a tumour suppressor (Chen et al., 2008; Karanikas et al., 2008; Wang et al., 2009; Zuo et al., 2007), where its dysfunction could contribute to other symptoms.

Disruption of transcription factor function by NC may not be limited to FOXP3. As previously mentioned, NC may be recognising a sequence outside of the FOXP3 consensus: either the packaging signal or some other region in the flanking regions of the oligonucleotide used in this study. Additionally, Yan et al. (2006), who found that the NC of SARS-CoV-1 can alter host expression, used DNA oligonucleotides that did not possess a packaging signal nor any similarity to the p38α promoter sequence used here; yet specific binding by NC was still seen during an EMSA. Thus, NC can likely recognise and bind different nucleic acid sequences. These sites may be spread throughout the human genome and thus allow NC to bind sporadically throughout various regions. If NC does bind to promoter regions that contain packaging signals or other sequences it recognises, it has the potential to disrupt the activity of any transcription factors that may bind in proximity to that site. Indeed, if true, this may contribute to the widespread and sometimes seemingly unrelated symptoms that present during COVID-19 (Struyf et al., 2021).

The question remains of sequence specificity of incoming NC once the FHD has dissociated from the DNA. While there is some sequence outside of the FOXP3 consensus that allows NC to bind specifically, there is evidently some additional factor that allows excess NC to cause FHD dissociation and further NC binding. NC has been found to bind nucleic acids cooperatively (Chang et al., 2009; Redzic et al., 2021; Savastano et al., 2020). In particular, it is thought that the NTD acts as the “seed” for cooperative binding, where the NTD binds to nucleic acids with some level of specificity, and then acts to recruit more NC which subsequently binds with less specificity (Iserman et al., 2020; Redzic et al., 2021). A possible chain of events for FHD – NC – DNA interactions is that the NTD of a single NC initially binds to the p38α promoter sequence specifically. The CTD may either remain free in solution or bind to the DNA non-specifically through electrostatic interactions. When NC is limiting, there is insufficient free NC to promote cooperative binding as free NC quickly binds unassociated DNA. Then, once no more free DNA binding sites are present, DNA-bound NC cooperatively assists free NC in non-specific DNA binding; successfully dislodging the FHD (**Fig. 34**). To confirm this, fluorescence anisotropy with labelled NC would need to be performed which, due to time constraints, was unfortunately not done. Additionally, whether the FHD can bind to a preformed NC – DNA complex; where the NC may have the opportunity to completely bind the DNA and prevent FHD binding; is still unanswered; which similarly requires labelled NC fluorescence anisotropy experiments.

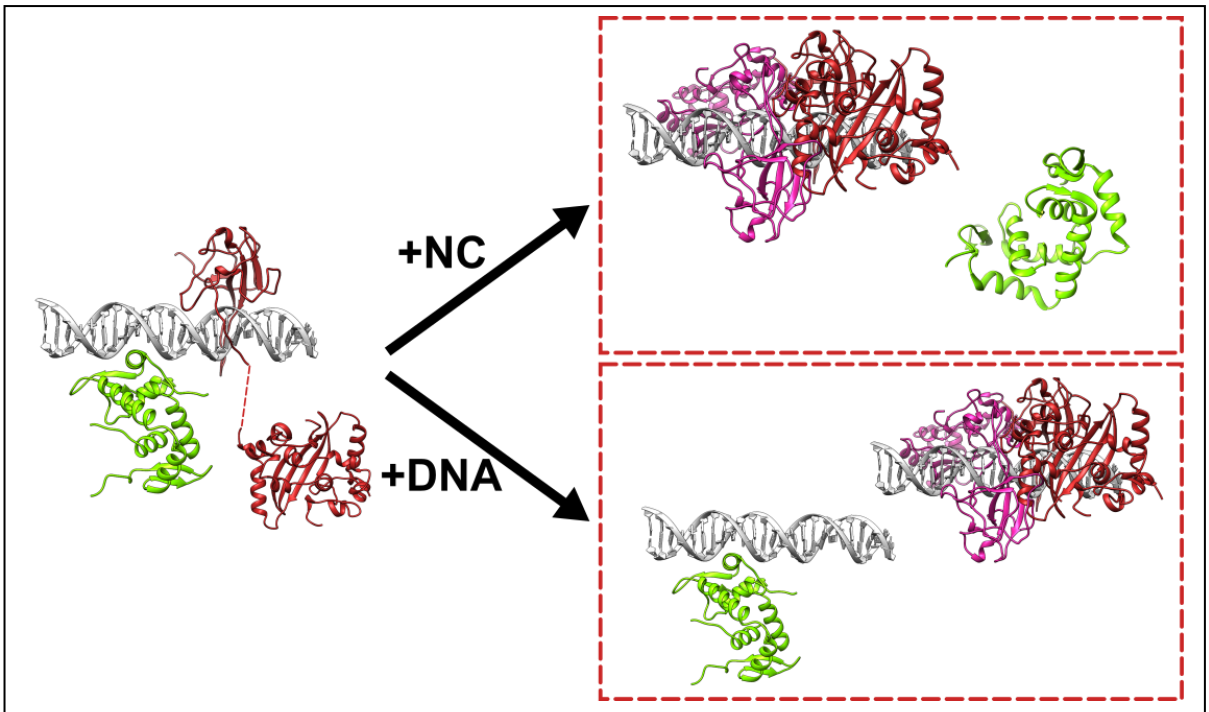


Fig. 34: FOXP3 FHD – NC – DNA interactions. When neither DNA nor NC is in excess, both the FHD (green) and NC (red) can bind to DNA (left). Since the NTD is proposed to provide sequence specificity, it may be that initially only the NTD of NC binds while the CTD remains free. If, from this initial state, additional NC is added, FHD dissociation occurs; possibly brought about through NC cooperativity (top right). If instead DNA is added, the two proteins dissociate and bind separate DNA molecules (bottom right). PDB codes: 3QRF (FHD), 6YI3 (NTD), 7CE0 (CTD), DNA structure generated using PyMol Builder tool. Visualised using UCSF Chimera (Pettersen et al., 2004). Image created using Inkscape V1.2.2.

6 - Conclusion

Here, interactions between the FHD of FOXP3, the SARS-CoV-2 nucleocapsid, and DNA have been investigated. The FHD and NC can both bind to a FOXP3 specific DNA oligonucleotide from the p38α promoter sequence at the same time when no single component is in excess. However, once NC is in excess, FHD dissociates from the DNA. Given that NC is the most abundant viral protein in an infected cell, and given FOXP3's important role in platelet generation, this may explain the widespread coagulopathy seen during COVID-19. Additionally, FOXP3 has important roles in other processes, particularly in immune regulation. SARS-CoV-2 has been found to infect some of the tissues that FOXP3 is expressed in and, if NC similarly interferes with FOXP3 in these tissues, it may explain the other widespread and seemingly unrelated symptoms seen during COVID-19. Additionally, the binding of NC to the oligonucleotide was found to not be driven by the FOXP3 consensus sequence; indicating that FOXP3 may not be the only transcription factor affected by NC DNA binding.

However, the *in vitro* conditions used here differ significantly from physiological conditions, so whether these observations are present *in vivo* remains unanswered. This is particularly important when considering the disordered nature of NC, a property which often causes different protein behaviour when in the crowded environment of the cell. Additionally, how these interactions are maintained when longer DNA sequences are used still requires investigation. Future cell culture work, such as luciferase assays, would help confirm whether this interaction is present in a physiologically relevant context. Additionally, further investigations into whether NC can similarly interfere with other transcription factors would help unravel the molecular mechanisms behind COVID-19 pathogenesis. FOXP3 FHD and NC show a weak but measurable direct interaction in the absence of DNA. NC has been shown to undergo liquid-liquid phase separation; wherein it locally recruits and concentrates certain proteins to assist during viral replication and assembly. LLPS is driven by weak interactions, and the weak interactions between NC and the FHD may be indicative of potential FOXP3 sequestering *in vivo* via LLPS. Whether this occurs and whether this process is specific or purely incidental requires further investigation. These results point toward a mechanism behind COVID-19 pathophysiology.

7 – Supplementary Material

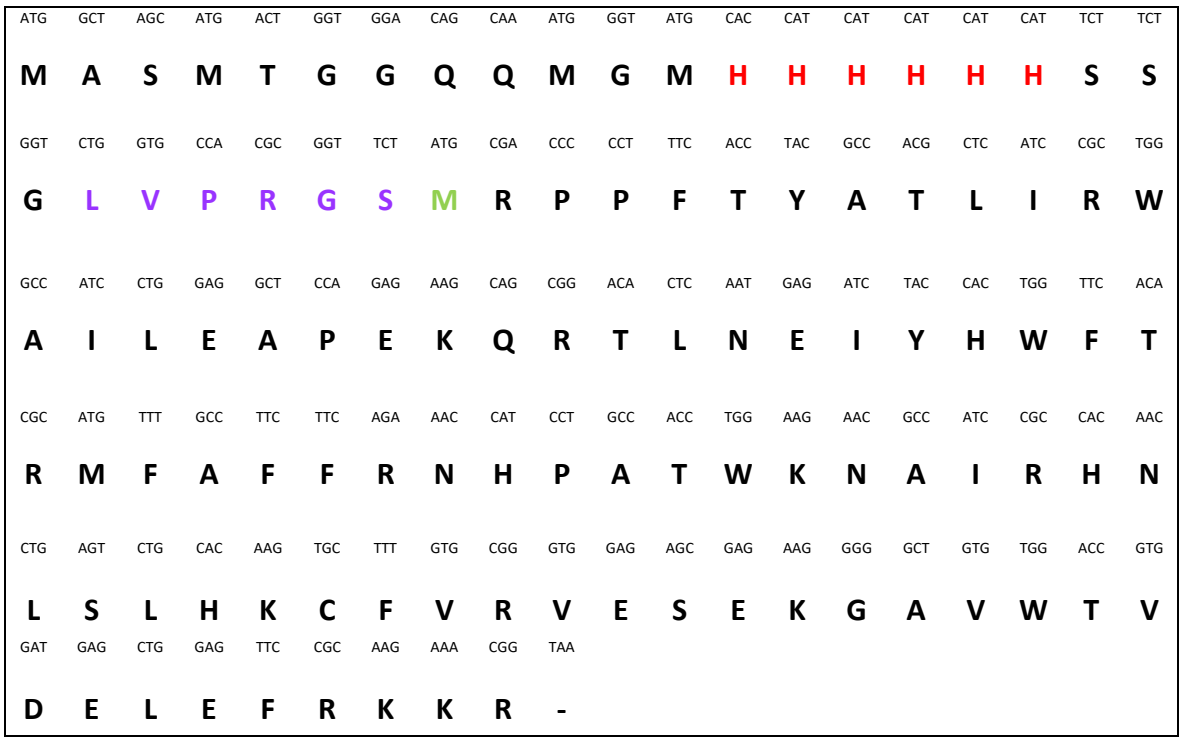
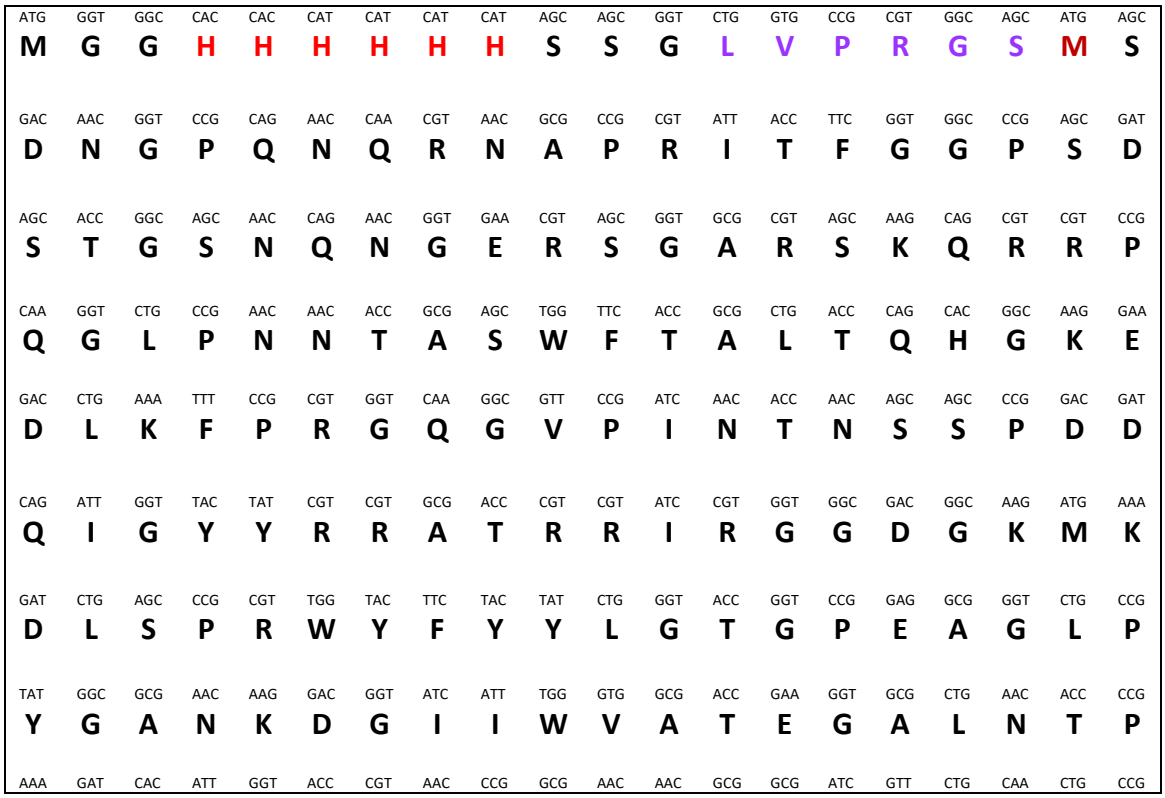


Fig. S1: FOXP3 FHD coding and protein sequence. The his-tag is shown in red, thrombin cleavage site in purple, and the first residue of the FHD in green.



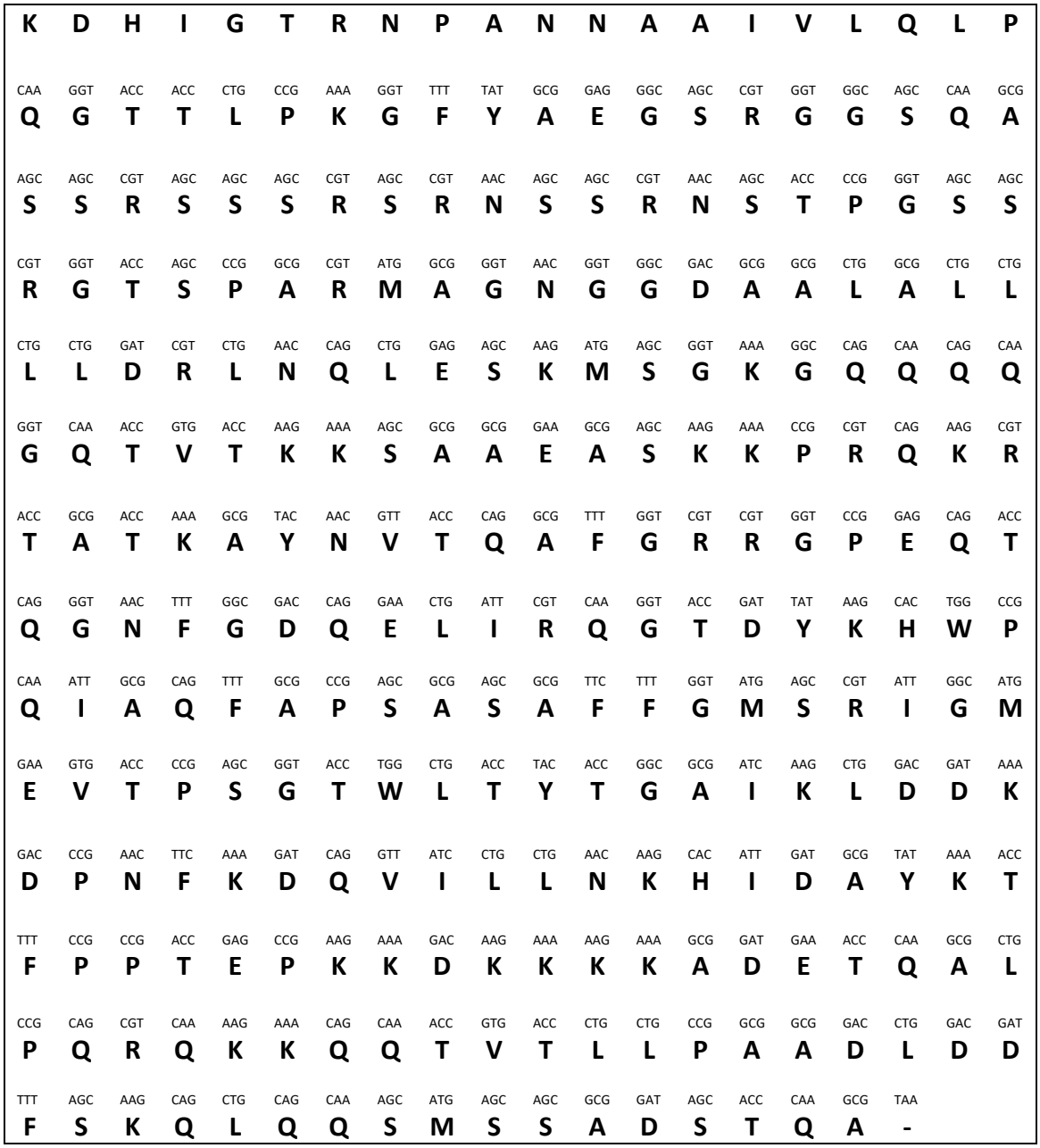


Fig. S2: SARS-CoV-2 coding and protein sequence. The his-tag is shown in red, thrombin cleavage site in purple, and the first residue of NC in maroon.

```
Score:230 bits(586), Expect:2e-85,
Method:Compositional matrix adjust.,
Identities:109/109(100%), Positives:109/109(100%), Gaps:0/109(0%)

Query  1    MASMTGGQQMGMHHHHHHSSGLVPRGSMRPPFTYATLIRWAILEAPEKQRTLNEIYHWFT  60
            MASMTGGQQMGMHHHHHHSSGLVPRGSMRPPFTYATLIRWAILEAPEKQRTLNEIYHWFT
Sbjct  1    MASMTGGQQMGMHHHHHHSSGLVPRGSMRPPFTYATLIRWAILEAPEKQRTLNEIYHWFT  60

Query  61   RMFAFFRNHPATWKNAIRHNLSLHKCFVRVESEKGAVWTVDELEFRKKR  109
            RMFAFFRNHPATWKNAIRHNLSLHKCFVRVESEKGAVWTVDELEFRKKR
Sbjct  61   RMFAFFRNHPATWKNAIRHNLSLHKCFVRVESEKGAVWTVDELEFRKKR  109
```

Fig. S3: Plasmid sequencing confirmed that the pET-11a plasmid contained the gene correctly coding for the FOXP3 FHD. The expected protein sequence (Query) was aligned with the translated gene sequence obtained via Sanger sequencing (Sbjct). Sequencing was performed by Inqaba Biotec, South Africa.

```
Score:896 bits(2316), Expect:0.0,
Method:Compositional matrix adjust.,
Identities:437/437(100%), Positives:437/437(100%), Gaps:0/437(0%)

Query  1    MGGHHHHHHSSGLVPRGSMSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRP  60
            MGGHHHHHHSSGLVPRGSMSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRP
Sbjct  1    MGGHHHHHHSSGLVPRGSMSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRP  60

Query  61   QGLPNNTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMK  120
            QGLPNNTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMK
Sbjct  61   QGLPNNTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMK  120

Query  121  DLSPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLP  180
            DLSPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLP
Sbjct  121  DLSPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLP  180

Query  181  QGTTLPKGFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAGNGGDAALALL  240
            QGTTLPKGFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAGNGGDAALALL
Sbjct  181  QGTTLPKGFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAGNGGDAALALL  240

Query  241  LLDRLNQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQT  300
            LLDRLNQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQT
Sbjct  241  LLDRLNQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQT  300

Query  301  QGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDDK  360
            QGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDDK
Sbjct  301  QGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDDK  360

Query  361  DPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTVTLLPAADLDD  420
            DPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTVTLLPAADLDD
Sbjct  361  DPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTVTLLPAADLDD  420

Query  421  FSKQLQQSMSSADSTQA  437
            FSKQLQQSMSSADSTQA
Sbjct  421  FSKQLQQSMSSADSTQA  437
```

Fig. S4: Plasmid sequencing confirmed that the pET-11a plasmid contained the gene correctly coding for the SARS-CoV-2 NC. The expected protein sequence (Query) was aligned with the translated gene sequence obtained via Sanger sequencing (Sbjct). Sequencing was performed by Inqaba Biotec, South Africa.

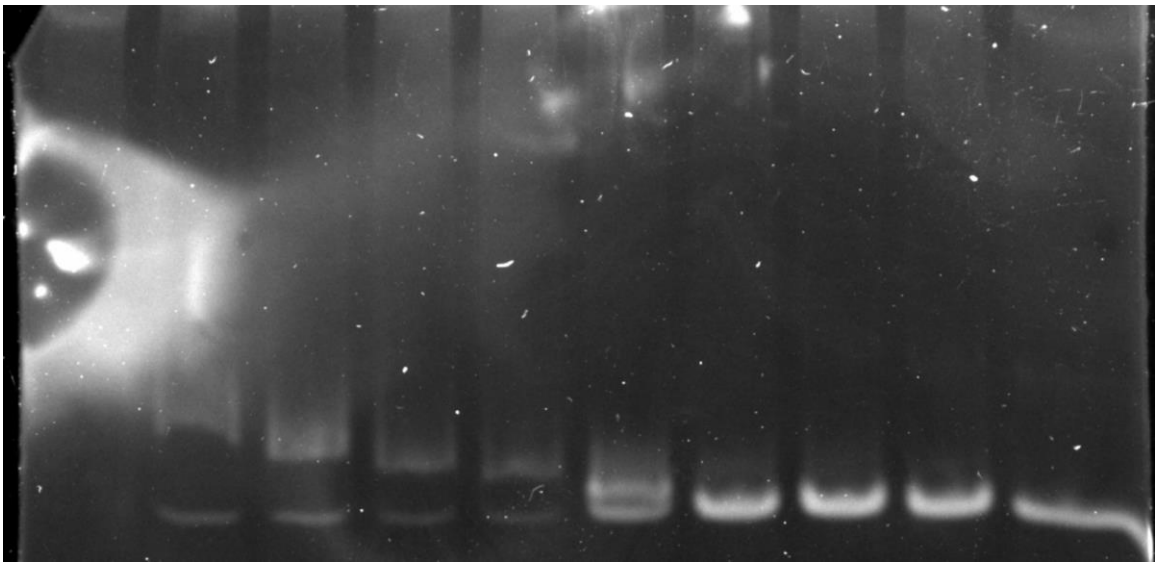


Fig. Spooky: Spencer the Spooky Specter, the sociable science spirit who spectated my studies and shielded against sinister sensations.

8 - References

Aiyegbusi, O. L., Hughes, S. E., Turner, G., Rivera, S. C., McMullan, C., Chandan, J. S., Haroon, S., Price, G., Davies, E. H., Nirantharakumar, K., Sapey, E., & Calvert, M. J. (2021). Symptoms, complications and management of long COVID: A review. *Journal of the Royal Society of Medicine*, *114*(9), 428–442. https://doi.org/10.1177/01410768211032850

Alvaro, T., Lejeune, M., Salvadó, M. T., Bosch, R., García, J. F., Jaén, J., Banham, A. H., Roncador, G., Montalbán, C., & Piris, M. A. (2005). Outcome in Hodgkin's lymphoma can be predicted from the presence of accompanying cytotoxic and regulatory T cells. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *11*(4), 1467–1473. https://doi.org/10.1158/1078-0432.CCR-04-1869

Bacchetta, R., Barzaghi, F., & Roncarolo, M.-G. (2018). From IPEX syndrome to FOXP3 mutation: A lesson on immune dysregulation. *Annals of the New York Academy of Sciences*, *1417*(1), 5–22. https://doi.org/10.1111/nyas.13011

Badoual, C., Hans, S., Rodriguez, J., Peyrard, S., Klein, C., Agueznay, N. E. H., Mosseri, V., Laccourreye, O., Bruneval, P., Fridman, W. H., Brasnu, D. F., & Tartour, E. (2006). Prognostic value of tumor-infiltrating CD4+ T-cell subpopulations in head and neck cancers. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *12*(2), 465–472. https://doi.org/10.1158/1078-0432.CCR-05-1886

Bai, Z., Cao, Y., Liu, W., & Li, J. (2021). The SARS-CoV-2 nucleocapsid protein and its role in viral structure, biological functions, and a potential target for drug or vaccine mitigation. *Viruses*, *13*(6), 1115. https://doi.org/10.3390/v13061115

Bandukwala, H. S., Wu, Y., Feuerer, M., Chen, Y., Barboza, B., Ghosh, S., Stroud, J. C., Benoist, C., Mathis, D., Rao, A., & Chen, L. (2011). Structure of a domain-swapped FOXP3 Dimer on DNA and its function in regulatory T cells. *Immunity*, *34*(4), 479–491. https://doi.org/10.1016/j.immuni.2011.02.017

Barth, H. G., Jackson, C., & Boyes, B. E. (1994). Size Exclusion Chromatography. *Analytical Chemistry*, *66*(12), 595–620. https://doi.org/10.1021/ac00084a022

Bates, G. J., Fox, S. B., Han, C., Leek, R. D., Garcia, J. F., Harris, A. L., & Banham, A. H. (2006). Quantification of regulatory T cells enables the identification of high-risk breast cancer patients and those at risk of late relapse. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *24*(34), 5373–5380. https://doi.org/10.1200/JCO.2006.05.9584

Bernard, J. J., Seweryniak, K. E., Koniski, A. D., Spinelli, S. L., Blumberg, N., Francis, C. W., Taubman, M. B., Palis, J., & Phipps, R. P. (2009). Foxp3 regulates megakaryopoiesis and platelet function. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *29*(11), 1874–1882. https://doi.org/10.1161/ATVBAHA.109.193805

Bettelli, E., Carrier, Y., Gao, W., Korn, T., Strom, T. B., Oukka, M., Weiner, H. L., & Kuchroo, V. K. (2006). Reciprocal developmental pathways for the generation of pathogenic effector T H 17 and regulatory T cells. *Nature*, *441*(7090), Article 7090. https://doi.org/10.1038/nature04753

Biotium. (2013). *Safety Report for GelRed® and GelGreen®*. Biotium. https://biotium.com/wp-content/uploads/2013/07/GelRed-and-GelGreen-Safety-Report.pdf

Bliddal, S., Banasik, K., Pedersen, O. B., Nissen, J., Cantwell, L., Schwinn, M., Tulstrup, M., Westergaard, D., Ullum, H., Brunak, S., Tommerup, N., Feenstra, B., Geller, F., Ostrowski, S. R., Grønbæk, K., Nielsen, C. H., Nielsen, S. D., & Feldt-Rasmussen, U. (2021). Acute and persistent symptoms in non-hospitalized PCR-confirmed COVID-19 patients. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-92045-x

Bouck, E. G., Denorme, F., Holle, L. A., Middelton, E. A., Blair, A. M., de Laat, B., Schiffman, J. D., Yost, C. C., Rondina, M. T., Wolberg, A. S., & Campbell, R. A. (2021). COVID-19 and sepsis are associated with different abnormalities in plasma procoagulant and fibrinolytic activity. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *41*(1), 401–414. https://doi.org/10.1161/ATVBAHA.120.315338

Briand, L., Marcion, G., Kriznik, A., Heydel, J. M., Artur, Y., Garrido, C., Seigneuric, R., & Neiers, F. (2016). A self-inducible heterologous protein expression system in *Escherichia coli*. *Scientific Reports*, *6*(1), Article 1. https://doi.org/10.1038/srep33037

Brown, T. A. (2002). Genomes. In *Genomes. 2nd edition* (2nd ed.). Wiley-Liss. https://www.ncbi.nlm.nih.gov/books/NBK21115/

Carfì, A., Bernabei, R., Landi, F., & for the Gemelli Against COVID-19 Post-Acute Care Study Group. (2020). Persistent symptoms in patients after acute COVID-19. *JAMA*, *324*(6), 603–605. https://doi.org/10.1001/jama.2020.12603

Carreras, J., Lopez-Guillermo, A., Fox, B. C., Colomo, L., Martinez, A., Roncador, G., Montserrat, E., Campo, E., & Banham, A. H. (2006). High numbers of tumor-infiltrating FOXP3-positive regulatory T cells are associated with improved overall survival in follicular lymphoma. *Blood*, *108*(9), 2957–2964. https://doi.org/10.1182/blood-2006-04-018218

Cascarina, S. M., & Ross, E. D. (2020). A proposed role for the SARS-CoV-2 nucleocapsid protein in the formation and regulation of biomolecular condensates. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, *34*(8), 9832–9842. https://doi.org/10.1096/fj.202001351

Chang, C., Lo, S.-C., Wang, Y.-S., & Hou, M.-H. (2016). Recent insights into the development of therapeutics against coronavirus diseases by targeting N protein. *Drug Discovery Today*, *21*(4), 562–572. https://doi.org/10.1016/j.drudis.2015.11.015

Chang, C.-K., Hsu, Y.-L., Chang, Y.-H., Chao, F.-A., Wu, M.-C., Huang, Y.-S., Hu, C.-K., & Huang, T.-H. (2009). Multiple nucleic acid binding sites and intrinsic disorder of severe acute respiratory syndrome coronavirus nucleocapsid protein: Implications for ribonucleocapsid protein packaging. *Journal of Virology*, *83*(5), 2255–2264. https://doi.org/10.1128/JVI.02001-08

Chen, C.-Y., Chang, C., Chang, Y.-W., Sue, S.-C., Bai, H.-I., Riang, L., Hsiao, C.-D., & Huang, T. (2007). Structure of the SARS Coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. *Journal of Molecular Biology*, *368*(4), 1075–1086. https://doi.org/10.1016/j.jmb.2007.02.069

Chen, G.-Y., Chen, C., Wang, L., Chang, X., Zheng, P., & Liu, Y. (2008). Cutting edge: Broad expression of the FoxP3 locus in epithelial cells: a caution against early interpretation of fatal inflammatory diseases following in vivo depletion of FoxP3-expressing cells. *Journal of Immunology (Baltimore, Md.: 1950)*, *180*(8), 5163–5166. https://doi.org/10.4049/jimmunol.180.8.5163

Chen, X., Laurent, S., Onur, O. A., Kleineberg, N. N., Fink, G. R., Schweitzer, F., & Warnke, C. (2021). A systematic review of neurological symptoms and complications of COVID-19. *Journal of Neurology*, *268*(2), 392–402. https://doi.org/10.1007/s00415-020-10067-3

Chen, Y., Chen, C., Zhang, Z., Liu, C.-C., Johnson, M. E., Espinoza, C. A., Edsall, L. E., Ren, B., Zhou, X. J., Grant, S. F. A., Wells, A. D., & Chen, L. (2015). DNA binding by FOXP3 domain-swapped dimer suggests mechanisms of long-range chromosomal interactions. *Nucleic Acids Research*, *43*(2), 1268–1282. https://doi.org/10.1093/nar/gku1373

Crook, H., Raza, S., Nowell, J., Young, M., & Edison, P. (2021). Long covid—Mechanisms, risk factors, and management. *BMJ*, *374*, n1648. https://doi.org/10.1136/bmj.n1648

Cubuk, J., Alston, J. J., Incicco, J. J., Singh, S., Stuchell-Brereton, M. D., Ward, M. D., Zimmerman, M. I., Vithani, N., Griffith, D., Wagoner, J. A., Bowman, G. R., Hall, K. B., Soranno, A., & Holehouse, A. S. (2021). The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-21953-3

Dahlbäck, B. (2000). Blood coagulation. *The Lancet*, *355*(9215), 1627–1632. https://doi.org/10.1016/S0140-6736(00)02225-X

Dai, S., Qu, L., Li, J., & Chen, Y. (2021). Toward a mechanistic understanding of DNA binding by forkhead transcription factors and its perturbation by pathogenic mutations. *Nucleic Acids Research*, *49*(18), 10235–10249. https://doi.org/10.1093/nar/gkab807

Deng, G., Song, X., Fujimoto, S., Piccirillo, C. A., Nagai, Y., & Greene, M. I. (2019). Foxp3 Post-translational Modifications and Treg Suppressive Activity. *Frontiers in Immunology*, *10*. https://www.frontiersin.org/articles/10.3389/fimmu.2019.02486

Deng, G., Song, X., & Greene, M. I. (2020). FoxP3 in Treg cell biology: A molecular and structural perspective. *Clinical & Experimental Immunology*, *199*(3), 255–262. https://doi.org/10.1111/cei.13357

Deutscher, M., P., & Burgess, R. (2009). *Guide to Protein Purification* (2nd ed., Vol. 463). Academic Press.

Devaud, C., Yong, C. S. M., John, L. B., Westwood, J. A., Duong, C. P. M., House, C. M., Denoyer, D., Li, J., Darcy, P. K., & Kershaw, M. H. (2014). Foxp3 expression in macrophages associated with RENCA tumors in mice. *PLOS ONE*, *9*(9), e108670. https://doi.org/10.1371/journal.pone.0108670

Dinesh, D. C., Chalupska, D., Silhan, J., Koutna, E., Nencka, R., Veverka, V., & Boura, E. (2020). Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein. *PLoS Pathogens*, *16*(12), e1009100. https://doi.org/10.1371/journal.ppat.1009100

Dominguez-Villar, M., & Hafler, D. A. (2018). Regulatory T cells in autoimmune disease. *Nature Immunology*, *19*(7), Article 7. https://doi.org/10.1038/s41590-018-0120-4

Douglass, S., Ali, S., Meeson, A. P., Browell, D., & Kirby, J. A. (2012). The role of FOXP3 in the development and metastatic spread of breast cancer. *Cancer and Metastasis Reviews*, *31*(3), 843–854. https://doi.org/10.1007/s10555-012-9395-3

Dubendorff, J. W., & Studier, F. W. (1991). Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *Journal of Molecular Biology*, *219*(1), 45–59. https://doi.org/10.1016/0022-2836(91)90856-2

Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins *. *Biochemistry*, *6*(7), 1948–1954. https://doi.org/10.1021/bi00859a010

Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S. L., Wiebe, V., Kitano, T., Monaco, A. P., & Pääbo, S. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*, *418*(6900), 869–872. https://doi.org/10.1038/nature01025

Esmon, C. T. (2005). The interactions between inflammation and coagulation. *British Journal of Haematology*, *131*(4), 417–430. https://doi.org/10.1111/j.1365-2141.2005.05753.x

Filbin, M. R., Mehta, A., Schneider, A. M., Kays, K. R., Guess, J. R., Gentili, M., Fenyves, B. G., Charland, N. C., Gonye, A. L. K., Gushterova, I., Khanna, H. K., LaSalle, T. J., Lavin-Parsons, K. M., Lilley, B. M., Lodenstein, C. L., Manakongtreecheep, K., Margolin, J. D., McKaig, B. N., Rojas-Lopez, M., … Goldberg, M. B. (2021). Longitudinal proteomic analysis of severe COVID-19 reveals survival-associated signatures, tissue-specific cell death, and cell-cell interactions. *Cell Reports Medicine*, *2*(5). https://doi.org/10.1016/j.xcrm.2021.100287

Fried, M., & Crothers, D. M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Research*, *9*(23), 6505–6525.

Fu, W., Ergun, A., Lu, T., Hill, J. A., Haxhinasto, S., Fassett, M. S., Gazit, R., Adoro, S., Glimcher, L., Chan, S., Kastner, P., Rossi, D., Collins, J. J., Mathis, D., & Benoist, C. (2012). A multiple redundant genetic switch locks in the transcriptional signature of T regulatory cells. *Nature Immunology*, *13*(10), 972–980. https://doi.org/10.1038/ni.2420

Furuyama, T., Nakazawa, T., Nakano, I., & Mori, N. (2000). Identification of the differential distribution patterns of mRNAs and consensus binding sequences for mouse DAF-16 homologues. *Biochemical Journal*, *349*(Pt 2), 629–634.

Gao, T., Gao, Y., Liu, X., Nie, Z., Sun, H., Lin, K., Peng, H., & Wang, S. (2021). Identification and functional analysis of the SARS-COV-2 nucleocapsid protein. *BMC Microbiology*, *21*(1), 58. https://doi.org/10.1186/s12866-021-02107-3

Garner, M. M., & Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: Application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Research*, *9*(13), 3047–3060.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In J. M. Walker (Ed.), *The Proteomics Protocols Handbook* (pp. 571–607). Humana Press. https://doi.org/10.1385/1-59259-890-0:571

Gentile, F., Amodeo, P., Febbraio, F., Picaro, F., Motta, A., Formisano, S., & Nucci, R. (2002). SDS-resistant active and thermostable dimers are obtained from the dissociation of homotetrameric β-glycosidase from hyperthermophilic sulfolobus solfataricus in SDS: Stabilizing role of the A-C intermonomeric interface. *Journal of Biological Chemistry*, *277*(46), 44050–44060. https://doi.org/10.1074/jbc.M206761200

Ghisaidoobe, A. B. T., & Chung, S. J. (2014). Intrinsic tryptophan fluorescence in the detection and analysis of proteins: A focus on Förster resonance energy transfer techniques. *International Journal of Molecular Sciences*, *15*(12), 22518–22538. https://doi.org/10.3390/ijms151222518

Giamarellos-Bourboulis, E. J., Netea, M. G., Rovina, N., Akinosoglou, K., Antoniadou, A., Antonakos, N., Damoraki, G., Gkavogianni, T., Adami, M.-E., Katsaounou, P., Ntaganou, M., Kyriakopoulou, M., Dimopoulos, G., Koutsodimitropoulos, I., Velissaris, D., Koufargyris, P., Karageorgos, A., Katrini, K., Lekakis, V., … Koutsoukou, A. (2020). Complex immune dysregulation in covid-19 patients with severe respiratory failure. *Cell Host & Microbe*, *27*(6), 992-1000.e3. https://doi.org/10.1016/j.chom.2020.04.009

Gobert, M., Treilleux, I., Bendriss-Vermare, N., Bachelot, T., Goddard-Leon, S., Arfi, V., Biota, C., Doffin, A. C., Durand, I., Olive, D., Perez, S., Pasqual, N., Faure, C., Ray-Coquard, I., Puisieux, A., Caux, C., Blay, J.-Y., & Ménétrier-Caux, C. (2009). Regulatory T cells recruited through

CCL22/CCR4 are selectively activated in lymphoid infiltrates surrounding primary breast tumors and lead to an adverse clinical outcome. *Cancer Research*, *69*(5), 2000–2009. https://doi.org/10.1158/0008-5472.CAN-08-2360

Gold, J. E., Okyay, R. A., Licht, W. E., & Hurley, D. J. (2021). Investigation of Long COVID prevalence and its relationship to Epstein-Barr Virus reactivation. *Pathogens*, *10*(6), Article 6. https://doi.org/10.3390/pathogens10060763

Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O'Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., Tummino, T. A., Hüttenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., Foussard, H., Batra, J., Haas, K., Modak, M., … Krogan, N. J. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, *583*(7816), Article 7816. https://doi.org/10.1038/s41586-020-2286-9

Goshua, G., Pine, A. B., Meizlish, M. L., Chang, C.-H., Zhang, H., Bahel, P., Baluha, A., Bar, N., Bona, R. D., Burns, A. J., Cruz, C. S. D., Dumont, A., Halene, S., Hwa, J., Koff, J., Menninger, H., Neparidze, N., Price, C., Siner, J. M., … Lee, A. I. (2020). Endotheliopathy in COVID-19-associated coagulopathy: Evidence from a single-centre, cross-sectional study. *The Lancet Haematology*, *7*(8), e575–e582. https://doi.org/10.1016/S2352-3026(20)30216-7

Grigorian, A. L., Bustamante, J. J., Hernandez, P., Martinez, A. O., & Haro, L. S. (2005). Extraordinarily stable disulfide-linked homodimer of human growth hormone. *Protein Science : A Publication of the Protein Society*, *14*(4), 902–913. https://doi.org/10.1110/ps.041048805

Grimes, J. M., & Grimes, K. V. (2020). p38 MAPK inhibition: A promising therapeutic approach for COVID-19. *Journal of Molecular and Cellular Cardiology*, *144*, 63–65. https://doi.org/10.1016/j.yjmcc.2020.05.007

Grossoehme, N. E., Li, L., Keane, S. C., Liu, P., Dann, C. E., Leibowitz, J. L., & Giedroc, D. P. (2009). Coronavirus N Protein N-Terminal Domain (NTD) specifically binds the Transcriptional Regulatory Sequence (TRS) and melts TRS-cTRS RNA duplexes. *Journal of Molecular Biology*, *394*(3), 544–557. https://doi.org/10.1016/j.jmb.2009.09.040

Guo, P., Benito Ballesteros, A., Yeung, S. P., Liu, R., Saha, A., Curtis, L., Kaser, M., Haggard, M. P., & Cheke, L. G. (2022a). COVCOG 1: Factors predicting physical, neurological and cognitive symptoms in long COVID in a community sample. a first publication from the COVID and cognition study. *Frontiers in Aging Neuroscience*, *14*. https://www.frontiersin.org/article/10.3389/fnagi.2022.804922

Guo, P., Benito Ballesteros, A., Yeung, S. P., Liu, R., Saha, A., Curtis, L., Kaser, M., Haggard, M. P., & Cheke, L. G. (2022b). COVCOG 2: Cognitive and memory deficits in long COVID: A second publication from the COVID and cognition study. *Frontiers in Aging Neuroscience*, *14*. https://www.frontiersin.org/article/10.3389/fnagi.2022.804937

Guruprasad, K., Reddy, B. V., & Pandit, M. W. (1990). Correlation between stability of a protein and its dipeptide composition: A novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Engineering*, *4*(2), 155–161. https://doi.org/10.1093/protein/4.2.155

Hancock, W. W., & Özkaynak, E. (2009). Three distinct domains contribute to nuclear transport of murine Foxp3. *PLOS ONE*, *4*(11), e7890. https://doi.org/10.1371/journal.pone.0007890

Hannenhalli, S., & Kaestner, K. H. (2009). The evolution of Fox genes and their role in development and disease. *Nature Reviews. Genetics*, *10*(4), 233–240. https://doi.org/10.1038/nrg2523

Hellman, L. M., & Fried, M. G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. *Nature Protocols*, *2*(8), 1849–1861. https://doi.org/10.1038/nprot.2007.249

Hiraoka, N., Onozato, K., Kosuge, T., & Hirohashi, S. (2006). Prevalence of FOXP3+ regulatory T cells increases during the progression of pancreatic ductal adenocarcinoma and its premalignant lesions. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *12*(18), 5423–5434. https://doi.org/10.1158/1078-0432.CCR-06-0369

Ho, C.-Y., Salimian, M., Hegert, J., O'Brien, J., Choi, S. G., Ames, H., Morris, M., Papadimitriou, J. C., Mininni, J., Niehaus, P., Burke, A., Canbeldek, L., Jacobs, J., LaRocque, A., Patel, K., Rice, K., Li, L., Johnson, R., LeFevre, A., … Drachenberg, C. (2022). Postmortem assessment of olfactory tissue degeneration and microvasculopathy in patients with COVID-19. *JAMA Neurology*. https://doi.org/10.1001/jamaneurol.2022.0154

Hong, P., Koza, S., & Bouvier, E. S. P. (2012). Size-Exclusion Chromatography for the analysis of protein biotherapeutics and their aggregates. *Journal of Liquid Chromatography & Related Technologies*, *35*(20), 2923–2950. https://doi.org/10.1080/10826076.2012.743724

Hori, S., Nomura, T., & Sakaguchi, S. (2003). Control of regulatory T cell development by the transcription factor Foxp3. *Science (New York, N.Y.)*, *299*(5609), 1057–1061. https://doi.org/10.1126/science.1079490

Huang, G., Wang, Y., & Chi, H. (2013). Control of T cell fates and immune tolerance by p38α signaling in mucosal CD103+ dendritic cells. *The Journal of Immunology*, *191*(2), 650–659. https://doi.org/10.4049/jimmunol.1300398

Iba, T., Levy, J. H., Levi, M., & Thachil, J. (2020). Coagulopathy in COVID-19. *Journal of Thrombosis and Haemostasis*, *18*(9), 2103–2109. https://doi.org/10.1111/jth.14975

Ichijo, H., Nishida, E., Irie, K., ten Dijke, P., Saitoh, M., Moriguchi, T., Takagi, M., Matsumoto, K., Miyazono, K., & Gotoh, Y. (1997). Induction of apoptosis by ASK1, a mammalian MAPKKK that activates SAPK/JNK and p38 signaling pathways. *Science (New York, N.Y.)*, *275*(5296), 90–94. https://doi.org/10.1126/science.275.5296.90

Iserman, C., Roden, C. A., Boerneke, M. A., Sealfon, R. S. G., McLaughlin, G. A., Jungreis, I., Fritch, E. J., Hou, Y. J., Ekena, J., Weidmann, C. A., Theesfeld, C. L., Kellis, M., Troyanskaya, O. G., Baric, R. S., Sheahan, T. P., Weeks, K. M., & Gladfelter, A. S. (2020). Genomic RNA Elements Drive Phase separation of the SARS-CoV-2 nucleocapsid. *Molecular Cell*, *80*(6), 1078-1091.e6. https://doi.org/10.1016/j.molcel.2020.11.041

Italiano Jr., J. E., & Shivdasani, R. A. (2003). Megakaryocytes and beyond: The birth of platelets. *Journal of Thrombosis and Haemostasis*, *1*(6), 1174–1182. https://doi.org/10.1046/j.1538-7836.2003.00290.x

Jamal, M., Bangash, H. I., Habiba, M., Lei, Y., Xie, T., Sun, J., Wei, Z., Hong, Z., Shao, L., & Zhang, Q. (2021). Immune dysregulation and system pathology in COVID-19. *Virulence*, *12*(1), 918–936. https://doi.org/10.1080/21505594.2021.1898790

James, G. T. (1978). Inactivation of the protease inhibitor phenylmethylsulfonyl fluoride in buffers. *Analytical Biochemistry*, *86*(2), 574–579. https://doi.org/10.1016/0003-2697(78)90784-4

Jameson, D. M., & Sawyer, W. H. (1995). [12] Fluorescence anisotropy applied to biomolecular interactions. In *Methods in Enzymology* (Vol. 246, pp. 283–300). Academic Press. https://doi.org/10.1016/0076-6879(95)46014-4

Jia, H., Qi, H., Gong, Z., Yang, S., Ren, J., Liu, Y., Li, M.-Y., & Chen, G. G. (2019). The expression of FOXP3 and its role in human cancers. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, *1871*(1), 170–178. https://doi.org/10.1016/j.bbcan.2018.12.004

Johnson, G. L., & Lapadat, R. (2002). Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases. *Science (New York, N.Y.)*, *298*(5600), 1911–1912. https://doi.org/10.1126/science.1072682

Jordanova, E. S., Gorter, A., Ayachi, O., Prins, F., Durrant, L. G., Kenter, G. G., van der Burg, S. H., & Fleuren, G. J. (2008). Human leukocyte antigen class I, MHC class I chain-related molecule A, and CD8+/regulatory T-cell ratio: Which variable determines survival of cervical cancer patients? *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *14*(7), 2028–2035. https://doi.org/10.1158/1078-0432.CCR-07-4554

Jurk, K., & Kehrel, B. E. (2005). Platelets: Physiology and biochemistry. *Seminars in Thrombosis and Hemostasis*, *31*(4), 381–392. https://doi.org/10.1055/s-2005-916671

Kalfaoglu, B., Almeida-Santos, J., Tye, C. A., Satou, Y., & Ono, M. (2021). T-cell dysregulation in COVID-19. *Biochemical and Biophysical Research Communications*, *538*, 204–210. https://doi.org/10.1016/j.bbrc.2020.10.079

Kander, T. (2020). Coagulation disorder in COVID-19. *The Lancet Haematology*, *7*(9), e630–e632. https://doi.org/10.1016/S2352-3026(20)30218-0

Karanikas, V., Speletas, M., Zamanakou, M., Kalala, F., Loules, G., Kerenidi, T., Barda, A. K., Gourgoulianis, K. I., & Germenis, A. E. (2008). Foxp3 expression in human cancer cells. *Journal of Translational Medicine*, *6*, 19. https://doi.org/10.1186/1479-5876-6-19

Kawasaki, T., & Kawai, T. (2014). Toll-Like Receptor signaling pathways. *Frontiers in Immunology*, *5*. https://www.frontiersin.org/article/10.3389/fimmu.2014.00461

Kee, J., Thudium, S., Renner, D. M., Glastad, K., Palozola, K., Zhang, Z., Li, Y., Lan, Y., Cesare, J., Poleshko, A., Kiseleva, A. A., Truitt, R., Cardenas-Diaz, F. L., Zhang, X., Xie, X., Kotton, D. N., Alysandratos, K. D., Epstein, J. A., Shi, P.-Y., … Korb, E. (2022). SARS-CoV-2 disrupts host epigenetic regulation via histone mimicry. *Nature*, *610*(7931), Article 7931. https://doi.org/10.1038/s41586-022-05282-z

Kelly, S. M., & Price, N. C. (2000). The use of circular dichroism in the investigation of protein structure and function. *Current Protein & Peptide Science*, *1*(4), 349–384. https://doi.org/10.2174/1389203003381315

Kim, C. H. (2009). FOXP3 and its role in the immune system. *Advances in Experimental Medicine and Biology*, *665*, 17–29. https://doi.org/10.1007/978-1-4419-1599-3_2

Kim, E. K., & Choi, E.-J. (2010). Pathological roles of MAPK signaling pathways in human diseases. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1802*(4), 396–405. https://doi.org/10.1016/j.bbadis.2009.12.009

Klok, F. A., Kruip, M. J. H. A., van der Meer, N. J. M., Arbous, M. S., Gommers, D., Kant, K. M., Kaptein, F. H. J., van Paassen, J., Stals, M. a. M., Huisman, M. V., & Endeman, H. (2020). Confirmation of the high cumulative incidence of thrombotic complications in critically ill ICU patients with COVID-19: An updated analysis. *Thrombosis Research*, *191*, 148–150. https://doi.org/10.1016/j.thromres.2020.04.041

Kolodziejski, P. J., Rashid, M. B., & Eissa, N. T. (2003). Intracellular formation of ”undisruptable” dimers of inducible nitric oxide synthase. *Proceedings of the National Academy of Sciences*, *100*(24), 14263–14268. https://doi.org/10.1073/pnas.2435290100

Korn, S. M., Dhamotharan, K., Jeffries, C. M., & Schlundt, A. (2023). The preference signature of the SARS-CoV-2 Nucleocapsid NTD for its 5'-genomic RNA elements. *Nature Communications*, *14*(1), Article 1. https://doi.org/10.1038/s41467-023-38882-y

Kumar, A., Parveen, A., Kumar, N., Bairy, S., Kaushik, V., Chandola, C., Sharma, J., Sharma, P., Agarwal, A., Pandey, A., Goyal, P., & Neerathilingam, M. (2020). *Characterization of nucleocapsid (N) protein from novel coronavirus SARS-CoV-2*.

Lachmann, A., Xu, H., Krishnan, J., Berger, S. I., Mazloom, A. R., & Ma'ayan, A. (2010). ChEA: Transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics (Oxford, England)*, *26*(19), 2438–2444. https://doi.org/10.1093/bioinformatics/btq466

Laemmli, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, *227*(5259), 680–685. https://doi.org/10.1038/227680a0

Laissue, P. (2019). The forkhead-box family of transcription factors: Key molecular players in colorectal cancer pathogenesis. *Molecular Cancer*, *18*(1), 5. https://doi.org/10.1186/s12943-019-0938-x

Latchman, D. S. (1997). Transcription factors: An overview. *The International Journal of Biochemistry & Cell Biology*, *29*(12), 1305–1312. https://doi.org/10.1016/S1357-2725(97)00085-X

Lee, W., & Lee, G. R. (2018). Transcriptional regulation and development of regulatory T cells. *Experimental & Molecular Medicine*, *50*(3), e456. https://doi.org/10.1038/emm.2017.313

Lehmann, O. J., Sowden, J. C., Carlsson, P., Jordan, T., & Bhattacharya, S. S. (2003). Fox's in development and disease. *Trends in Genetics*, *19*(6), 339–344. https://doi.org/10.1016/S0168-9525(03)00111-2

Leng, F., Zhang, W., Ramirez, R. N., Leon, J., Zhong, Y., Hou, L., Yuki, K., Veeken, J. van der, Rudensky, A. Y., Benoist, C., & Hur, S. (2022). The transcription factor FoxP3 can fold into two dimerization states with divergent implications for regulatory T cell function and immune homeostasis. *Immunity*, *55*(8), 1354-1369.e8. https://doi.org/10.1016/j.immuni.2022.07.002

Lesnierowski, G., & Kijowski, J. (2007). Lysozyme. In R. Huopalahti, R. López-Fandiño, M. Anton, & R. Schade (Eds.), *Bioactive Egg Compounds* (pp. 33–42). Springer. https://doi.org/10.1007/978-3-540-37885-3_6

Levi, M., Thachil, J., Iba, T., & Levy, J. H. (2020). Coagulation abnormalities and thrombosis in patients with COVID-19. *The Lancet Haematology*, *7*(6), e438–e440. https://doi.org/10.1016/S2352-3026(20)30145-9

Li, B., Samanta, A., Song, X., Iacono, K. T., Brennan, P., Chatila, T. A., Roncador, G., Banham, A. H., Riley, J. L., Wang, Q., Shen, Y., Saouaf, S. J., & Greene, M. I. (2007). FOXP3 is a homo-oligomer and a component of a supramolecular regulatory complex disabled in the human XLAAD/IPEX autoimmune disease. *International Immunology*, *19*(7), 825–835. https://doi.org/10.1093/intimm/dxm043

Li, C., Wang, H., Fang, H., He, C., Pei, Y., & Gai, X. (2021). FOXP3 facilitates the invasion and metastasis of non-small cell lung cancer cells through regulating VEGF, EMT and the Notch1/Hes1 pathway. *Experimental and Therapeutic Medicine*, *22*(3), 958. https://doi.org/10.3892/etm.2021.10390

Li, J., Jiang, L., Liang, X., Qu, L., Wu, D., Chen, X., Guo, M., Chen, Z., Chen, L., & Chen, Y. (2017). DNA-binding properties of FOXP3 transcription factor. *Acta Biochimica et Biophysica Sinica*, *49*(9), 792–799. https://doi.org/10.1093/abbs/gmx079

Li, S., Weidenfeld, J., & Morrisey, E. E. (2004). Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. *Molecular and Cellular Biology*, *24*(2), 809–822. https://doi.org/10.1128/MCB.24.2.809-822.2004

Li, Y., Li, D., Yang, W., Fu, H., Liu, Y., & Li, Y. (2016). Overexpression of the transcription factor FOXP3 in lung adenocarcinoma sustains malignant character by promoting G1/S transition gene CCND1. *Tumour Biology: The Journal of the International Society for Oncodevelopmental Biology and Medicine*, *37*(6), 7395–7404. https://doi.org/10.1007/s13277-015-4616-3

Lin, H.-Y., Zhu, C.-Q., Zhang, H.-H., Shen, Z.-C., Zhang, C.-X., & Ye, Y.-X. (2021). The genetic network of Forkhead gene family in development of brown planthoppers. *Biology*, *10*(9), Article 9. https://doi.org/10.3390/biology10090867

Lohman, T. M. (1986). Kinetics of protein-nucleic acid interactions: Use of salt effects to probe mechanisms of interaction. *CRC Critical Reviews in Biochemistry*, *19*(3), 191–245. https://doi.org/10.3109/10409238609084656

Lozano, R., Gbekie, C., Siper, P. M., Srivastava, S., Saland, J. M., Sethuram, S., Tang, L., Drapeau, E., Frank, Y., Buxbaum, J. D., & Kolevzon, A. (2021). FOXP1 syndrome: A review of the literature and practice parameters for medical assessment and monitoring. *Journal of Neurodevelopmental Disorders*, *13*(1), 18. https://doi.org/10.1186/s11689-021-09358-1

Lu, S., Ye, Q., Singh, D., Cao, Y., Diedrich, J. K., Yates, J. R., Villa, E., Cleveland, D. W., & Corbett, K. D. (2021). The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-020-20768-y

Lu, X., Pan, J., Tao, J., & Guo, D. (2011). SARS-CoV nucleocapsid protein antagonizes IFN-β response by targeting initial step of IFN-β induction pathway, and its C-terminal region is critical for the antagonism. *Virus Genes*, *42*(1), 37–45. https://doi.org/10.1007/s11262-010-0544-x

Luo, H., Chen, J., Chen, K., Shen, X., & Jiang, H. (2006). Carboxyl terminus of Severe Acute Respiratory Syndrome Coronavirus nucleocapsid protein: self-association analysis and nucleic acid binding characterization. *Biochemistry*, *45*(39), 11827–11835. https://doi.org/10.1021/bi0609319

Luo, H., Chen, Q., Chen, J., Chen, K., Shen, X., & Jiang, H. (2005). The nucleocapsid protein of SARS coronavirus has a high binding affinity to the human cellular heterogeneous nuclear ribonucleoprotein A1. *FEBS Letters*, *579*(12), 2623–2628. https://doi.org/10.1016/j.febslet.2005.03.080

Luo, W., Yu, H., Gou, J., Li, X., Sun, Y., Li, J., & Liu, L. (2020). *Clinical pathology of critical patient with novel Coronavirus pneumonia (COVID-19)*. https://www.preprints.org/manuscript/202002.0407/v1

Lutomski, C. A., El-Baba, T. J., Bolla, J. R., & Robinson, C. V. (2021). Multiple roles of SARS-CoV-2 N protein facilitated by proteoform-specific interactions with RNA, host proteins, and convalescent antibodies. *JACS Au*, *1*(8), 1147–1157. https://doi.org/10.1021/jacsau.1c00139

Lyon, M. F., Peters, J., Glenister, P. H., Ball, S., & Wright, E. (1990). The scurfy mouse mutant has previously unrecognized hematological abnormalities and resembles Wiskott-Aldrich syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(7), 2433–2437.

Manne, B. K., Denorme, F., Middleton, E. A., Portier, I., Rowley, J. W., Stubben, C., Petrey, A. C., Tolley, N. D., Guo, L., Cody, M., Weyrich, A. S., Yost, C. C., Rondina, M. T., & Campbell, R. A.

(2020). Platelet gene expression and function in patients with COVID-19. *Blood*, *136*(11), 1317–1329. https://doi.org/10.1182/blood.2020007214

Mao, L., Jin, H., Wang, M., Hu, Y., Chen, S., He, Q., Chang, J., Hong, C., Zhou, Y., Wang, D., Miao, X., Li, Y., & Hu, B. (2020). Neurologic Manifestations of Hospitalized Patients With Coronavirus Disease 2019 in Wuhan, China. JAMA Neurology, 77(6), 683–690. https://doi.org/10.1001/jamaneurol.2020.1127

Mark, J., Li, X., Cyr, T., Fournier, S., Jaentschke, B., & Hefford, M. A. (2008). SARS coronavirus: Unusual lability of the nucleocapsid protein. *Biochemical and Biophysical Research Communications*, *377*(2), 429–433. https://doi.org/10.1016/j.bbrc.2008.09.153

Marson, A., Kretschmer, K., Frampton, G. M., Jacobsen, E. S., Polansky, J. K., MacIsaac, K. D., Levine, S. S., Fraenkel, E., von Boehmer, H., & Young, R. A. (2007). Foxp3 occupancy and regulation of key target genes during T-cell stimulation. *Nature*, *445*(7130), 931–935. https://doi.org/10.1038/nature05478

Matsui, T., Kamata, S., Ishii, K., Maruno, T., Ghanem, N., Uchiyama, S., Kato, K., Suzuki, A., Oda-Ueda, N., Ogawa, T., & Tanaka, Y. (2019). SDS-induced oligomerization of Lys49-phospholipase A2 from snake venom. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-38861-8

McConkey, E. H. (1982). Molecular evolution, intracellular organization, and the quinary structure of proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *79*(10), 3236–3240. https://doi.org/10.1073/pnas.79.10.3236

Merad, M., & Martin, J. C. (2020). Pathological inflammation in patients with COVID-19: A key role for monocytes and macrophages. *Nature Reviews Immunology*, *20*(6), Article 6. https://doi.org/10.1038/s41577-020-0331-4

Micsonai, A., Wien, F., Kernya, L., Lee, Y.-H., Goto, Y., Réfrégiers, M., & Kardos, J. (2015). Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proceedings of the National Academy of Sciences*, *112*(24), E3095–E3103. https://doi.org/10.1073/pnas.1500851112

Minuz, P., Gaino, S., Zuliani, V., Tommasoli, R. M., Benati, D., Ortolani, R., Zancanaro, C., Berton, G., & Santonastaso, C. L. (2002). Functional role of p38 mitogen activated protein kinase in platelet activation induced by a thromboxane A2 analogue and by 8-iso-prostaglandin F2alpha. *Thrombosis and Haemostasis*, *87*(5), 888–898.

Mizukami, Y., Kono, K., Kawaguchi, Y., Akaike, H., Kamimura, K., Sugai, H., & Fujii, H. (2008). Localisation pattern of Foxp3+ regulatory T cells is associated with clinical behaviour in gastric cancer. *British Journal of Cancer*, *98*(1), 148–153. https://doi.org/10.1038/sj.bjc.6604149

Moriguchi, T., Kuroyanagi, N., Yamaguchi, K., Gotoh, Y., Irie, K., Kano, T., Shirakabe, K., Muro, Y., Shibuya, H., Matsumoto, K., Nishida, E., & Hagiwara, M. (1996). A novel kinase cascade mediated by mitogen-activated protein kinase kinase 6 and MKK3. *The Journal of Biological Chemistry*, *271*(23), 13675–13679. https://doi.org/10.1074/jbc.271.23.13675

Mu, J., Xu, J., Zhang, L., Shu, T., Wu, D., Huang, M., Ren, Y., Li, X., Geng, Q., Xu, Y., Qiu, Y., & Zhou, X. (2020). SARS-CoV-2-encoded nucleocapsid protein acts as a viral suppressor of RNA interference in cells. *Science China. Life Sciences*, *63*(9), 1413–1416. https://doi.org/10.1007/s11427-020-1692-1

Ni, X., Kou, W., Gu, J., Wei, P., Wu, X., Peng, H., Tao, J., Yan, W., Yang, X., Lebid, A., Park, B. V., Chen, Z., Tian, Y., Fu, J., Newman, S., Wang, X., Shen, H., Li, B., Blazar, B. R., … Lu, L. (2019). TRAF6

directs FOXP3 localization and facilitates regulatory T-cell function through K63-linked ubiquitination. *The EMBO Journal*, *38*(9), e99766. https://doi.org/10.15252/embj.201899766

Nooren, I. M. A., & Thornton, J. M. (2003). Diversity of protein-protein interactions. *The EMBO Journal*, *22*(14), 3486–3492. https://doi.org/10.1093/emboj/cdg359

Ono, M. (2020). Control of regulatory T-cell differentiation and function by T-cell receptor signalling and Foxp3 transcription factor complexes. *Immunology*, *160*(1), 24–37. https://doi.org/10.1111/imm.13178

Peng, Y., Du, N., Lei, Y., Dorje, S., Qi, J., Luo, T., Gao, G. F., & Song, H. (2020). Structures of the SARS-CoV-2 nucleocapsid and their perspectives for drug design. *The EMBO Journal*, *39*(20), e105938. https://doi.org/10.15252/embj.2020105938

Perdikari, T. M., Murthy, A. C., Ryan, V. H., Watters, S., Naik, M. T., & Fawzi, N. L. (2020). SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs. *The EMBO Journal*, *39*(24), e106478. https://doi.org/10.15252/embj.2020106478

Perlis, R. H., Ognyanova, K., Santillana, M., Baum, M. A., Lazer, D., Druckman, J., & Della Volpe, J. (2021). Association of acute symptoms of COVID-19 and symptoms of depression in adults. *JAMA Network Open*, *4*(3), e213223. https://doi.org/10.1001/jamanetworkopen.2021.3223

Perumal, K. (2018). *Stability and kinetics of folding of the domain-swapped dimer of the FOXP3 forkhead domain* [PhD Thesis]. University of Witwatersrand.

Perumal, K., Dirr, H. W., & Fanucchi, S. (2015). A single amino acid in the hinge loop region of the FOXP Forkhead domain is significant for dimerisation. *The Protein Journal*, *34*(2), 111–121. https://doi.org/10.1007/s10930-015-9603-4

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Pretorius, E., Vlok, M., Venter, C., Bezuidenhout, J. A., Laubscher, G. J., Steenkamp, J., & Kell, D. B. (2021). Persistent clotting protein pathology in Long COVID/Post-Acute Sequelae of COVID-19 (PASC) is accompanied by increased levels of antiplasmin. *Cardiovascular Diabetology*, *20*, 172. https://doi.org/10.1186/s12933-021-01359-7

Purdie, N. (1996). Circular dichroism and the conformational analysis of biomolecules. *Journal of the American Chemical Society*, *118*(50), 12871–12871. https://doi.org/10.1021/ja965689f

Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G., & Deber, C. M. (2009). Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proceedings of the National Academy of Sciences*, *106*(6), 1760–1765.

Raveendran, A. V., Jayadevan, R., & Sashidharan, S. (2021). Long COVID: An overview. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, *15*(3), 869–875. https://doi.org/10.1016/j.dsx.2021.04.007

Redzic, J. S., Lee, E., Born, A., Issaian, A., Henen, M. A., Nichols, P. J., Blue, A., Hansen, K. C., D'Alessandro, A., Vögeli, B., & Eisenmesser, E. Z. (2021). The inherent dynamics and interaction sites of the SARS-CoV-2 nucleocapsid N-terminal region. *Journal of Molecular Biology*, *433*(15), 167108. https://doi.org/10.1016/j.jmb.2021.167108

Reed, G. L., Houng, A. K., Singh, S., & Wang, D. (2017). α2-Antiplasmin: New Insights and Opportunities for Ischemic Stroke. *Seminars in Thrombosis and Hemostasis*, *43*(2), 191–199. https://doi.org/10.1055/s-0036-1585077

Renwrantz, L., Marquart, A., Möck, A., & Richards, E. (2009). Molecular properties of the partially SDS-resistant lectin from the albumen gland of *Helix pomatia* and demonstration of lectin-

related molecules on the surface of *H. pomatia* haemocytes. *Journal of Molluscan Studies*, *75*(1), 41–49. https://doi.org/10.1093/mollus/eyn037

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Frontiers in Microbiology*, *5*. https://doi.org/10.3389/fmicb.2014.00172

Rouillard, A. D., Gundersen, G. W., Fernandez, N. F., Wang, Z., Monteiro, C. D., McDermott, M. G., & Ma'ayan, A. (2016). The harmonizome: A collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database*, *2016*, baw100. https://doi.org/10.1093/database/baw100

Sadeghifard, N., Talaei, M., Nasser, A., & Maleki, F., farajallah, farajollah, F,. (2017). T regulatory cell and body tolerance. *Der Pharmacia Lettre*, *9*, 23–28.

Sadlon, T. J., Wilkinson, B. G., Pederson, S., Brown, C. Y., Bresatz, S., Gargett, T., Melville, E. L., Peng, K., D'Andrea, R. J., Glonek, G. G., Goodall, G. J., Zola, H., Shannon, M. F., & Barry, S. C. (2010). Genome-wide identification of human FOXP3 target genes in natural regulatory T cells. *The Journal of Immunology*, *185*(2), 1071–1081. https://doi.org/10.4049/jimmunol.1000082

Saklatvala, J., Rawlinson, L., Waller, R. J., Sarsfield, S., Lee, J. C., Morton, L. F., Barnes, M. J., & Farndale, R. W. (1996). Role for p38 mitogen-activated protein kinase in platelet aggregation caused by collagen or a thromboxane analogue. *The Journal of Biological Chemistry*, *271*(12), 6586–6589. https://doi.org/10.1074/jbc.271.12.6586

Salama, P., Phillips, M., Grieu, F., Morris, M., Zeps, N., Joseph, D., Platell, C., & Iacopetta, B. (2009). Tumor-infiltrating FOXP3+ T regulatory cells show strong prognostic significance in colorectal cancer. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *27*(2), 186–192. https://doi.org/10.1200/JCO.2008.18.7229

Savastano, A., Ibáñez de Opakua, A., Rankovic, M., & Zweckstetter, M. (2020). Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates. *Nature Communications*, *11*(1), 6041. https://doi.org/10.1038/s41467-020-19843-1

Schägger, H., & von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry*, *166*(2), 368–379. https://doi.org/10.1016/0003-2697(87)90587-2

Shen, S., Zhang, J., Fang, Y., Lu, S., Wu, J., Zheng, X., & Deng, F. (2021). SARS-CoV-2 interacts with platelets and megakaryocytes via ACE2-independent mechanism. *Journal of Hematology & Oncology*, *14*(1), 72. https://doi.org/10.1186/s13045-021-01082-6

Shen, X.-R., Geng, R., Li, Q., Chen, Y., Li, S.-F., Wang, Q., Min, J., Yang, Y., Li, B., Jiang, R.-D., Wang, X., Zheng, X.-S., Zhu, Y., Jia, J.-K., Yang, X.-L., Liu, M.-Q., Gong, Q.-C., Zhang, Y.-L., Guan, Z.-Q., … Zhou, P. (2022). ACE2-independent infection of T lymphocytes by SARS-CoV-2. *Signal Transduction and Targeted Therapy*, *7*(1), Article 1. https://doi.org/10.1038/s41392-022-00919-x

Shirahama, K., Tsujii, K., & Takagi, T. (1974). Free-boundary electrophoresis of sodium dodecyl sulfate-protein polypeptide complexes with special reference to SDS-polyacrylamide gel electrophoresis. *Journal of Biochemistry*, *75*(2), 309–319. https://doi.org/10.1093/oxfordjournals.jbchem.a130398

Sin, C., Li, H., & Crawford, D. A. (2015). Transcriptional regulation by FOXP1, FOXP2, and FOXP4 dimerization. *Journal of Molecular Neuroscience*, *55*(2), 437–448. https://doi.org/10.1007/s12031-014-0359-7

Soares-Silva, M., Diniz, F. F., Gomes, G. N., & Bahia, D. (2016). The Mitogen-Activated Protein Kinase (MAPK) pathway: role in immune evasion by trypanosomatids. *Frontiers in Microbiology*, *7*. https://www.frontiersin.org/article/10.3389/fmicb.2016.00183

Somasundaran, P., & Hubbard, A. T. (2006). *Encyclopedia of surface and colloid science* (2nd ed). Taylor & Francis. http://www.informaworld.com/smpp/title~content=t713172975

Song, X., Li, B., Xiao, Y., Chen, C., Wang, Q., Liu, Y., Berezov, A., Xu, C., Gao, Y., Wu, S.-L., Zhang, H., Karger, B. L., Hancock, W. W., Wells, A. D., Zhou, Z., & Greene, M. I. (2012). Structural and biological features of FOXP3 dimerization relevant to regulatory T cell function. *Cell Reports*, *1*(6), 665–675. https://doi.org/10.1016/j.celrep.2012.04.012

Steiner, R. F. (2002). Fluorescence Anisotropy: Theory and Applications. In J. R. Lakowicz (Ed.), *Topics in fluorescence spectroscopy: Principles* (pp. 1–52). Springer US. https://doi.org/10.1007/0-306-47058-6_1

Stroud, J. C., Wu, Y., Bates, D. L., Han, A., Nowick, K., Paabo, S., Tong, H., & Chen, L. (2006). Structure of the forkhead domain of FOXP2 bound to DNA. *Structure (London, England: 1993)*, *14*(1), 159–166. https://doi.org/10.1016/j.str.2005.10.005

Struyf, T., Deeks, J. J., Dinnes, J., Takwoingi, Y., Davenport, C., Leeflang, M. M., Spijker, R., Hooft, L., Emperador, D., Domen, J., Horn, S. R. A., Bruel, A. V. den, & Group, C. C.-19 D. T. A. (2021). Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19. *Cochrane Database of Systematic Reviews*, *2*. https://doi.org/10.1002/14651858.CD013665.pub2

Surjit, M., Kumar, R., Mishra, R. N., Reddy, M. K., Chow, V. T. K., & Lal, S. K. (2005). The Severe Acute Respiratory Syndrome Coronavirus nucleocapsid protein is phosphorylated and localizes in the cytoplasm by 14-3-3-mediated translocation. *Journal of Virology*, *79*(17), 11476–11486. https://doi.org/10.1128/jvi.79.17.11476-11486.2005

Suter, D. M. (2020). Transcription Factors and DNA Play Hide and Seek. *Trends in Cell Biology*, *30*(6), 491–500. https://doi.org/10.1016/j.tcb.2020.03.003

Swonger, K. N., & Robinson, A. S. (2018). Using fluorescence anisotropy for ligand binding kinetics of membrane proteins. *Current Protocols in Protein Science*, *93*(1), e63. https://doi.org/10.1002/cpps.63

Tanizawa, K., Ishii, S., Hamaguchi, K., & Kanaoka, Y. (1971). Aromatic amidines as competitive inhibitors of trypsin. *J. Biochem.*, *69*(5), 893–899. https://doi.org/10.1093%2Foxfordjournals.jbchem.a129540

Telenti, A., Arvin, A., Corey, L., Corti, D., Diamond, M. S., García-Sastre, A., Garry, R. F., Holmes, E. C., Pang, P. S., & Virgin, H. W. (2021). After the pandemic: Perspectives on the future trajectory of COVID-19. *Nature*, *596*(7873), Article 7873. https://doi.org/10.1038/s41586-021-03792-w

Tellinghuisen, T. L., Hamburger, A. E., Fisher, B. R., Ostendorp, R., & Kuhn, R. J. (1999). In vitro assembly of Alphavirus cores by using nucleocapsid protein expressed in *Escherichia coli*. *Journal of Virology*, *73*(7), 5309–5319.

Tiwari, P., Kaila, P., & Guptasarma, P. (2019). Understanding anomalous mobility of proteins on SDS-PAGE with special reference to the highly acidic extracellular domains of human E- and N-cadherins. *ELECTROPHORESIS*, *40*(9), 1273–1281. https://doi.org/10.1002/elps.201800219

Tompa, P. (2002). Intrinsically unstructured proteins. *Trends in Biochemical Sciences*, *27*(10), 527–533. https://doi.org/10.1016/s0968-0004(02)02169-2

Uversky, V. N. (2019). Intrinsically disordered proteins and their “mysterious” (meta)physics. *Frontiers in Physics*, *7*. https://www.frontiersin.org/articles/10.3389/fphy.2019.00010

Uversky, V. N., & Dunker, A. K. (2010). Understanding protein non-folding. *Biochimica Et Biophysica Acta*, *1804*(6), 1231–1264. https://doi.org/10.1016/j.bbapap.2010.01.017

Vadasz, Z., & Toubi, E. (2017). FoxP3 expression in macrophages, cancer, and B cells-is it real? *Clinical Reviews in Allergy & Immunology*, *52*(3), 364–372. https://doi.org/10.1007/s12016-016-8572-5

Valeur, B. (2012). *Molecular fluorescence: principles and applications* (2nd ed.). Wiley-VCH. https://onlinelibrary.wiley.com/doi/book/10.1002/3527600248

Venter, C., Bezuidenhout, J. A., Laubscher, G. J., Lourens, P. J., Steenkamp, J., Kell, D. B., & Pretorius, E. (2020). Erythrocyte, platelet, serum ferritin, and P-Selectin pathophysiology implicated in severe hypercoagulation and vascular complications in COVID-19. *International Journal of Molecular Sciences*, *21*(21), E8234. https://doi.org/10.3390/ijms21218234

Vivian, J. T., & Callis, P. R. (2001). Mechanisms of tryptophan fluorescence shifts in proteins. *Biophysical Journal*, *80*(5), 2093–2109. https://doi.org/10.1016/S0006-3495(01)76183-8

Voet, D., Voet, J., & Pratt, C. (2016). *Fundamentals of Biochemistry* (5th ed.). Wiley-VCH.

Wang, B., Lin, D., Li, C., & Tucker, P. (2003). Multiple domains define the expression and regulatory properties of Foxp1 forkhead transcriptional repressors. *The Journal of Biological Chemistry*, *278*(27), 24259–24268. https://doi.org/10.1074/jbc.M207174200

Wang, L., Liu, R., Li, W., Chen, C., Katoh, H., Chen, G.-Y., McNally, B., Lin, L., Zhou, P., Zuo, T., Cooney, K. A., Liu, Y., & Zheng, P. (2009). Somatic single-hits inactivate the X-linked tumor suppressor FOXP3 in the prostate. *Cancer Cell*, *16*(4), 336–346. https://doi.org/10.1016/j.ccr.2009.08.016

Wang, L., Sola, I., Enjuanes, L., & Zuñiga, S. (2021). MOV10 helicase interacts with Coronavirus nucleocapsid protein and has antiviral activity. *mBio*, *12*(5), 10.1128/mbio.01316-21. https://doi.org/10.1128/mbio.01316-21

Wang, W., Su, B., Pang, L., Qiao, L., Feng, Y., Ouyang, Y., Guo, X., Shi, H., Wei, F., Su, X., Yin, J., Jin, R., & Chen, D. (2020). High-dimensional immune profiling by mass cytometry revealed immunosuppression and dysfunction of immunity in COVID-19 patients. *Cellular & Molecular Immunology*, *17*(6), 650–652. https://doi.org/10.1038/s41423-020-0447-2

Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., & Jones, D. T. (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *Journal of Molecular Biology*, *337*(3), 635–645. https://doi.org/10.1016/j.jmb.2004.02.002

Webb, H., Steeb, O., Blane, A., Rotherham, L., Aron, S., Machanick, P., Dirr, H., & Fanucchi, S. (2017). The FOXP2 forkhead domain binds to a variety of DNA sequences with different rates and affinities. *The Journal of Biochemistry*, *162*(1), 45–54. https://doi.org/10.1093/jb/mvx003

Wildin, R. S., Smyk-Pearson, S., & Filipovich, A. H. (2002). Clinical and molecular features of the immunodysregulation, polyendocrinopathy, enteropathy, X linked (IPEX) syndrome. *Journal of Medical Genetics*, *39*(8), 537–545. https://doi.org/10.1136/jmg.39.8.537

Wilson, K., & Walker, J. M. (Eds.). (2009). *Principles and techniques of biochemistry and molecular biology* (7th ed). Cambridge University Press.

Wirth, A. J., & Gruebele, M. (2013). Quinary protein structure and the consequences of crowding in living cells: Leaving the test-tube behind. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *35*(11), 984–993. https://doi.org/10.1002/bies.201300080

Wittig, I., Velours, J., Stuart, R., & Schägger, H. (2008). Characterization of domain interfaces in monomeric and dimeric ATP synthase. *Molecular & Cellular Proteomics: MCP*, *7*(5), 995–1004. https://doi.org/10.1074/mcp.M700465-MCP200

Woody, R. W. (1995). Circular dichroism. *Methods in Enzymology*, *246*, 34–71. https://doi.org/10.1016/0076-6879(95)46006-3

World Health Organisation. (2023). *COVID-19 deaths | WHO COVID-19 dashboard*. WHO Coronavirus (COVID-19) Dashboard. https://data.who.int/dashboards/covid19/deaths

Wu, C., Qavi, A. J., Hachim, A., Kavian, N., Cole, A. R., Moyle, A. B., Wagner, N. D., Sweeney-Gibbons, J., Rohrs, H. W., Gross, M. L., Peiris, J. S. M., Basler, C. F., Farnsworth, C. W., Valkenburg, S. A., Amarasinghe, G. K., & Leung, D. W. (2021). Characterization of SARS-CoV-2 nucleocapsid protein reveals multiple functional consequences of the C-terminal domain. *iScience*, *24*(6). https://doi.org/10.1016/j.isci.2021.102681

Wu, C.-S. (Ed.). (2003). *Handbook of size exclusion chromatography and related techniques: Revised and expanded* (2nd ed.). CRC Press. https://doi.org/10.1201/9780203913321

Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., Hu, Y., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., Yuan, M.-L., Zhang, Y.-L., Dai, F.-H., Liu, Y., Wang, Q.-M., Zheng, J.-J., Xu, L., Holmes, E. C., & Zhang, Y.-Z. (2020). A new coronavirus associated with human respiratory disease in China. *Nature*, *579*(7798), Article 7798. https://doi.org/10.1038/s41586-020-2008-3

Wurm, T., Chen, H., Hodgson, T., Britton, P., Brooks, G., & Hiscox, J. A. (2001). Localization to the nucleolus is a common feature of coronavirus nucleoproteins, and the protein may disrupt host cell division. *Journal of Virology*, *75*(19), 9345–9356. https://doi.org/10.1128/JVI.75.19.9345-9356.2001

Xue, B., Dunker, A. K., & Uversky, V. N. (2012). Orderly order in protein intrinsic disorder distribution: Disorder in 3500 proteomes from viruses and the three domains of life. *Journal of Biomolecular Structure & Dynamics*, *30*(2), 137–149. https://doi.org/10.1080/07391102.2012.675145

Yan, X., Hao, Q., Mu, Y., Timani, K. A., Ye, L., Zhu, Y., & Wu, J. (2006). Nucleocapsid protein of SARS-CoV activates the expression of cyclooxygenase-2 by binding directly to regulatory elements for nuclear factor-kappa B and CCAAT/enhancer binding protein. *The International Journal of Biochemistry & Cell Biology*, *38*(8), 1417–1428. https://doi.org/10.1016/j.biocel.2006.02.003

Yang, M., He, S., Chen, X., Huang, Z., Zhou, Z., Zhou, Z., Chen, Q., Chen, S., & Kang, S. (2021). Structural insight into the SARS-CoV-2 nucleocapsid protein C-Terminal domain reveals a novel recognition mechanism for viral transcriptional regulatory sequences. *Frontiers in Chemistry*, *8*. https://www.frontiersin.org/article/10.3389/fchem.2020.624765

Yao, C., Bora, S. A., Parimon, T., Zaman, T., Friedman, O. A., Palatinus, J. A., Surapaneni, N. S., Matusov, Y. P., Cerro Chiang, G., Kassar, A. G., Patel, N., Green, C. E. R., Aziz, A. W., Suri, H., Suda, J., Lopez, A. A., Martins, G. A., Stripp, B. R., Gharib, S. A., … Chen, P. (2021). Cell-type-specific immune dysregulation in severely ill COVID-19 patients. *Cell Reports*, *34*(1), 108590. https://doi.org/10.1016/j.celrep.2020.108590

Yomogida, K., Wu, S., Baravati, B., Avendano, C., Caldwell, T., Maniaci, B., Zhu, Y., & Chu, C.-Q. (2013). Cell penetrating recombinant Foxp3 protein enhances Treg function and ameliorates arthritis. *Biochemical and Biophysical Research Communications*, *434*(2), 263–267. https://doi.org/10.1016/j.bbrc.2013.02.114

Youn, Y.-J., Lee, Y.-B., Kim, S.-H., Jin, H. K., Bae, J., & Hong, C.-W. (2021). Nucleocapsid and spike proteins of SARS-CoV-2 drive neutrophil extracellular trap formation. *Immune Network*, *21*(2), e16. https://doi.org/10.4110/in.2021.21.e16

Yuriditsky, E., Horowitz, J. M., Merchan, C., Ahuja, T., Brosnahan, S. B., McVoy, L., & Berger, J. S. (2020). Thromboelastography profiles of critically ill patients with coronavirus disease 2019. *Critical Care Medicine*, *48*(9), 1319–1326. https://doi.org/10.1097/CCM.0000000000004471

Zeng, W., Liu, G., Ma, H., Zhao, D., Yang, Y., Liu, M., Mohammed, A., Zhao, C., Yang, Y., Xie, J., Ding, C., Ma, X., Weng, J., Gao, Y., He, H., & Jin, T. (2020). Biochemical characterization of SARS-CoV-2 nucleocapsid protein. *Biochemical and Biophysical Research Communications*, *527*(3), 618–623. https://doi.org/10.1016/j.bbrc.2020.04.136

Zhang, W., Leng, F., Wang, X., Ramirez, R. N., Park, J., Benoist, C., & Hur, S. (2023). FOXP3 recognizes microsatellites and bridges DNA through multimerization. *Nature*, 1–9. https://doi.org/10.1038/s41586-023-06793-z

Zhang, W., & Liu, H. T. (2002). MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Research*, *12*(1), Article 1. https://doi.org/10.1038/sj.cr.7290105

Zheng, Y., Josefowicz, S. Z., Kas, A., Chu, T.-T., Gavin, M. A., & Rudensky, A. Y. (2007). Genome-wide analysis of Foxp3 target genes in developing and mature regulatory T cells. *Nature*, *445*(7130), 936–940. https://doi.org/10.1038/nature05563

Zhou, B., Liu, J., Wang, Q., Liu, X., Li, X., Li, P., Ma, Q., & Cao, C. (2008). The nucleocapsid protein of severe acute respiratory syndrome coronavirus inhibits cell cytokinesis and proliferation by interacting with translation elongation factor 1alpha. *Journal of Virology*, *82*(14), 6962–6971. https://doi.org/10.1128/JVI.00133-08

Zhou, F., Wan, Q., Chen, S., Chen, Y., Wang, P.-H., Yao, X., & He, M. (2021). Attenuating innate immunity and facilitating β-coronavirus infection by NSP1 of SARS-CoV-2 through specific redistributing hnRNP A2/B1 cellular localization. *Signal Transduction and Targeted Therapy*, *6*(1), Article 1. https://doi.org/10.1038/s41392-021-00786-y

Zhou, R., Zeng, R., von Brunn, A., & Lei, J. (2020). Structural characterization of the C-terminal domain of SARS-CoV-2 nucleocapsid protein. *Molecular Biomedicine*, *1*(1), 2. https://doi.org/10.1186/s43556-020-00001-4

Zhu, Y., O'Neill, S., Saklatvala, J., Tassi, L., & Mendelsohn, M. (1994). Phosphorylated HSP27 associates with the activation-dependent cytoskeleton in human platelets. *Blood*, *84*(11), 3715–3723. https://doi.org/10.1182/blood.V84.11.3715.bloodjournal84113715

Zuo, T., Liu, R., Zhang, H., Chang, X., Liu, Y., Wang, L., Zheng, P., & Liu, Y. (2007). FOXP3 is a novel transcriptional repressor for the breast cancer oncogene SKP2. *The Journal of Clinical Investigation*, *117*(12), 3765–3773. https://doi.org/10.1172/JCI32538

Zuo, T., Wang, L., Morrison, C., Chang, X., Zhang, H., Li, W., Liu, Y., Wang, Y., Liu, X., Chan, M. W. Y., Liu, J.-Q., Love, R., Liu, C.-G., Godfrey, V., Shen, R., Huang, T. H.-M., Yang, T., Park, B. K., Wang, C.-Y., … Liu, Y. (2007). FOXP3 is an X-linked breast cancer suppressor gene and an important repressor of the HER-2/ErbB2 oncogene. *Cell*, *129*(7), 1275–1286. https://doi.org/10.1016/j.cell.2007.04.034