

The SARS-CoV-2 Nucleocapsid protein and its interaction with Berberine and Palmatine

by

TRACEY-LEE VERMAAK

Submitted in accordance with the requirements for the degree of

MASTER OF SCIENCE

in the subject

LIFE SCIENCES

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF S GILDENHUYS

CO-SUPERVISOR: PROF S MOSEBI

October 8, 2025

Declaration

I, Tracey-Lee Vermaak, hereby declare that the dissertation, with the title: “The SARS-CoV-2 Nucleocapsid Protein and its interaction with berberine and palmitate” which I hereby submit for the degree of Master of Science in Life Sciences at the University of South Africa, is my own work and has not previously been submitted by me for a degree at this or any other institution.

I declare that the dissertation does not contain any written work presented by other persons whether written, pictures, graphs or data or any other information without acknowledging the source.

I declare that where words from a written source have been used the words have been paraphrased and referenced and where exact words from a source have been used the words have been placed inside quotation marks and referenced.

I declare that I have not copied and pasted any information from the Internet, without specifically acknowledging the source and have inserted appropriate references to these sources in the reference section of the dissertation or thesis.

I declare that during my study, I adhered to the Research Ethics Policy of the University of South Africa, received ethics approval for the duration of my study prior to the commencement of data gathering, and have not acted outside the approval conditions.

I declare that the content of my dissertation has been submitted through an electronic plagiarism detection program before the final submission for examination.

Student Signature:

Date:

Abstract

The Nucleocapsid (N) protein of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) plays various critical roles within viral replication, RNA binding and the immune response, making it a viable target for therapeutic treatments. This study investigated the biochemical interactions between the N protein and two naturally occurring isoquinoline alkaloids berberine and palmatine. The N protein was recombinantly expressed using *E. coli* BL21(DE3)-T7 express cells transformed with a pET-15b plasmid containing a consensus sequence of the protein harbouring P13L; Δ 31–33; R203K; G204R; and S413R mutations in relation to the Wuhan sequence, that was generated to represent commonly circulating variants to April 2024. Nickel Affinity Chromatography was used to purify the protein of interest, which was confirmed using Mass Spectroscopy.

Absorbance, Far Ultraviolet (UV) Circular Dichroism (CD), and Intrinsic Tryptophan Fluorescence (ITF) Spectroscopies as well as Blue Native Polyacrylamide Gel Electrophoresis (BN-PAGE) were used to investigate the native structure of the protein as well as the effects imposed by compound addition. These techniques confirmed the native protein structure, oligomerisation and nucleic acid binding capacity of the protein isolated. Far-UV-CD suggested that the binding of compounds caused slight secondary structural shifts from disorder & α -helical content toward β -sheet formation. These changes were confirmed by Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy (ATR-FTIR) and highlighted changes in the hydrophobic environment and hydrogen bonding patterns upon alkaloid binding. BN-PAGE indicated that the binding of compound does not alter the quaternary structure of the protein but allowed for partial protection of tetrameric species from urea-induced denaturation. The apparent secondary stabilisation and slight tertiary destabilisation observed by Far-UV-CD and ITF thermal denaturation likely accompanied this shift in oligomeric equilibrium, rather than being the primary drivers of stability.

In silico docking predicted various potential low affinity binding sites for both compounds, with docking scores of -30.42 kcal/mol for palmatine and -28.46 kcal/mol for berberine. Experimental validation using Isothermal Titration Calorimetry (ITC) indicates low affinity binding, between -2388.62 to -8 kcal/mol, whilst Surface Plasmon Resonance (SPR) confirmed the low affinity binding for palmatine.

Keywords

Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy (ATR-FTIR)

Berberine

Far Ultraviolet (UV) circular dichroism (CD)

Intrinsic Tryptophan Fluorescence Spectroscopy

Isothermal Titration Calorimetry (ITC)

Nucleocapsid (N) Protein

Palmitine

Protein ligand interactions

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

Surface Plasmon Resonance (SPR)

Research Outputs

1. Poster Presentation

CAES Research and Innovation Day, which was held on **May 22, 2024, at the Thamsanqa Auditorium** at the University of South Africa Florida Campus.

2. Manuscript Submissions

One manuscript is being drafted whilst another is under review.

Dedication

I dedicate this dissertation to Professor Samantha Gildenhuis. Without her unwavering support, guidance, and belief in my potential, this work would not have been possible. Her mentorship, patience, and wisdom were invaluable throughout this journey. I am deeply grateful for every opportunity she has granted me to grow both as a scientist and as an individual.

Thank you, Prof, for inspiring me to keep asking questions and to complete my thoughts as you would say.

Acknowledgements

I would like to acknowledge the profound contributions of:

My mom, Annelise, and partner, Johan. Your unwavering support and encouragement softened the hard times during this journey.

My lab mates and allies in research, Tovhowani, Ofentse, Tshepo, and Pekos. Sharing knowledge and experiences with you was invaluable.

Our lab technician Ms. Thato Maanyapelo. Your dedication and support were invaluable to this research. I would also like to thank our lab technician Mr Garland More for enabling me to perform ATR-FTIR.

The University of South Africa for hosting and funding this research.

My co-supervisor, Professor Salerwe Mosebi. Your guidance and mentorship were instrumental.

My supervisor, Professor Samantha Gildenhuis. Your contributions were not only the foundation of the journey, but also the light-house navigating us to the goal.

Abbreviations

++ssRNA	positive-sense single-stranded ribonucleic acid
ACE-2	Angiotensin-Converting Enzyme 2
ANS	8-anilinonaphthalene-1-sulfonic acid
ATR-FTIR	Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy
BN-PAGE	Blue Native Polyacrylamide Gel Electrophoresis
BSA	Bovine Serum Albumin
C-arm	C terminal arm
COVID	Coronavirus disease
CTD	C-Terminal Domain
DMSO	Dimethyl sulfoxide
DNase	Deoxyribonuclease
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylene diamine tetra-acetic acid
EU	European Union
Far-UV-CD	Far Ultra-Violet Circular Dichroism
FRET	Förster Resonance Energy Transfer
GBMV	Generalized Born with Molecular Volume
HCl	Hydrochloric acid
HV	High Voltage
IDRs	Intrinsically Disordered Regions
IEC	Ion Exchange Chromatography
IFN	Interferon
IFR	Infection Fatality Rate
IMAC	Immobilised Metal Affinity Chromatography
IPTG	isopropyl β -D-1-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
LB	Luria-Bertani
LKR	Ser/Arg-rich flexible linker
Mpro	Main Protease
MRE	Mean Residue Ellipticity

MS	Microsoft
MWM	Molecular Weight Marker
NaCl	Sodium Chloride
NH ₄ HCO ₃	Ammonium Bicarbonate
NP	Nucleocapsid protein
NSP	Non-structural Proteins
NTD	N Terminal Domain
ORFs	Open Reading Frames
PAGE	Polyacrylamide gel electrophoresis
PANGO	Phylogenetic Assignment of Named Global Outbreak
PBA/PBB/PBC	Purification Buffers A, B & C
pIFE	Primary Inner Filter Effects
PTM	Post-Translational Modification
RMSD	Root Mean Square Deviation
RNP	Ribonucleoprotein
RTC	Replication and Transcription Complex
RU	Response Units
S protein	Spike Protein
SARS-CoV-2	Severe Acute Respiratory Syndrome Corona Virus 2
SDS-PAGE	Sodium Dodecyl sulphate-polyacrylamide gel electrophoresis
sIFE	Secondary Inner filter effects
SPR	Surface Plasmon Resonance
TEMED	Tetramethylethylenediamine
TNF- α	Tumor Necrosis Factor alpha
TRIS	tris(hydroxymethyl)aminomethane
USA	United States of America
UV	Ultraviolet
VOCs	Variants of Concern
WHO	World Health Organisation
W-NP	Consensus sequence to the nucleocapsid protein
ΔH	Enthalpy
ΔS	Entropy

Standard one-letter and three-letter IUPAC codes are used for amino acid residues.

Table of Contents

Declaration	ii
Abstract	iii
Keywords	iv
Research Outputs	v
Dedication	vi
Acknowledgements	vii
Abbreviations	viii
Table of Contents	1
1. Introduction	4
1.1 Introduction to SARS-CoV-2	4
1.2 Genome Composition, Structure, and Replication Cycle of SARS-CoV-2	5
1.2.1 The Genome Composition and Structure of SARS-CoV-2	5
1.2.2 The SARS-CoV-2 Replication Cycle	7
1.3 Current SARS-CoV-2 Prevention and Treatment Strategies	10
1.4 SARS-CoV-2 Nucleocapsid Protein	11
1.5 <i>Berberis vulgaris</i> – A Herbal Therapy	16
2. Aims & Objectives	20
3. Experimental Procedures	21
3.1 Materials	21
3.2 Software	21
3.3. Methods	22
3.3.1 NP Consensus Sequence Construction	22
3.3.2 Protein Modelling	23
3.3.3 Plasmid	24
3.3.4 Cellular Competency and Transformation	25

3.3.5 Protein Expression.....	26
3.3.6 SDS-PAGE	28
3.3.7 Chromatography	29
3.3.8 Post-Purification Protein Processing.....	31
3.3.9 Mass Spectrometry	34
3.3.10 Far-UV Circular Dichroism	36
3.3.11 Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy (ATR-FTIR).....	37
3.3.12 Intrinsic Tryptophan Fluorescence	38
3.3.13 Native-PAGE	39
3.3.14 Isothermal Titration Calorimetry	40
3.3.15 Stability Studies	41
3.3.16 Surface Plasmon Resonance	41
3.3.17 Computational Docking.....	43
4. Results	46
4.1 In Silico Validation & Structural Prediction of W-NP.....	46
4.2 Expression, Purification, and Biochemical Characterisation	51
4.2.1 Optimisation of Protein Expression in <i>E.coli</i>	51
4.2.2 Protein Purification and Strategy Development	57
4.2.3 Protein Composition Analysis and Identity Confirmation.....	66
4.3. Structural Features of W-NP	69
4.3.1 Far UV CD	69
4.3.2 Fluorescence	72
4.3.3 ANS Binding Studies	75
4.3.4 BN-PAGE.....	76
4.4 Effect of Alkaloids on W-NP Structure.....	78
4.4.1 Far-UV CD	78
4.4.2 ATR-FTIR	80

4.4.3 BN-PAGE.....	82
4.4.4 Thermal Stability	83
4.5. Mechanisms of Interaction and Binding Parameters	87
4.5.1 Absorbance Profiling of Experimental Components.....	87
4.5.2 ANS Binding Studies	90
4.5.3 Fluorescence Quenching	90
4.5.4 Surface Plasmon Resonance	95
4.5.5 Isothermal Titration Calorimetry.....	99
4.5.6 Docking.....	102
5. Discussion	107
6. Conclusion	111
7. Limitations of the study.....	112
8. Future considerations	113
9. References	114
10. Appendix	158

1. Introduction

1.1 Introduction to SARS-CoV-2

In December of 2019, scientists identified a new strain of coronavirus infecting humans, and by the 11th of March 2020, this virus had spread all over the world, enabling the World Health Organization (WHO) to declare it as what we from thereon known as the COVID-19 pandemic (WHO, 2022). COVID-19 is caused by SARS-CoV-2, which belongs to the coronavirus genus, which similarly caused pandemics in the years 2003 [Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV)] and 2012 [Middle East Respiratory Syndrome (MERS)] (Abdolmaleki et al., 2022) (Coronaviruses ['corona' = crown-like spikes on the surface] (Ksiazek et al., 2003) are a family of diverse enveloped viruses that contain positive-sense single-stranded ribonucleic acid (ssRNA), which are known to infect humans and animals (Lai & Cavanagh, 1997). Sequence analysis shows that SARS-CoV-2 is a member of the *β-coronavirus* genus and *Sarbecovirus* subgenus (Gorbalenya et al., 2004).

A SARS-CoV-2 infection may be asymptomatic, or moderate to severe, where symptoms may appear after a four-to-five-day incubation and may persist for up to 10 days (Rozenchwajg et al., 2021; WHO, 2022). Infection may result in symptoms, which include fever, cough, fatigue, dyspnoea, and a loss of taste or smell, and less commonly a sore throat, headache, diarrhoea, myalgia, or rashes (WHO, 2022). A severe infection may entail respiratory failure, septic shock, organ dysfunction/failure, and could result in life-threatening respiratory pathologies or injuries (Abdolmaleki et al., 2022; V'kovski et al., 2021). The severity of infection may be enhanced with clinical comorbidities, such as hypertension, diabetes, and cardiovascular disease (Moreno et al., 2022).

The infection fatality rate (IFR) differs widely, with countries reporting average values between 0.5% to 3%, numbers which, according to WHO, may be much higher than confirmed (WHO, 2022). In addition, the IFR of SARS-CoV-2 infection is drastically influenced by age, ranging from 0.002% at age 10, to 15% at age 85 (Levin et al., 2020). Furthermore, 20% of the cases require hospitalisation, together creating a large burden on global health systems (Rozenchwajg et al., 2021).

Although the global incidence of SARS-CoV-2 infections has declined and is no longer classified as a Public Health Emergency of International Concern (PHEIC), the virus still circulates widely and it continues to evolve, giving rise to novel variants that are frequently designated as Variants of Concern (VOCs) due to the impact of mutations on transmissibility, immune escape, or disease severity (European Centre for Disease Prevention and Control, 2025). Furthermore, approximately 6 in 100 people develop post-COVID-19 condition (PCC), or Post-Acute Sequelae, also known as long COVID (Al-Aly et al., 2024; Schultheiß et al., 2022; WHO, 2025). This is defined as a continuation or development of new symptoms, of which more than 200 have been identified; where the most frequently reported symptoms include long-lasting fatigue, breathlessness, muscle or joint pain, and impaired sleep (Al-Aly et al., 2024; Ford, 2023; Hanson et al., 2022). Importantly, PCC also affects individuals who experienced mild or asymptomatic initial infections, not only those who experienced severe disease. While mortality is primarily associated with the acute phase of SARS-CoV-2 infection, PPC is characterised by persistent morbidity rather than increased direct death rates. The recognition and prevalence of PCC therefore transform the understanding of SARS-CoV-2 infection from an acute illness into a chronic health challenge for a significant portion of the population.

1.2 Genome Composition, Structure, and Replication Cycle of SARS-CoV-2

1.2.1 The Genome Composition and Structure of SARS-CoV-2

The ~29.9 kb linear single-stranded, ++ssRNA genome (Zhang, B. et al., 2022) is depicted in [Figure 1. 1.A](#). The figure shows that the largest majority of the genome is composed of two Open Reading Frames (ORFs), namely ORF1a and ORF1b, whilst the remainder of the genome consists of overlapping ORFs, namely ORF3, ORF6, ORF7a, ORF7b, ORF8, and ORF9b, all of which produce accessory proteins. Additionally, there are four structural proteins - Spike (S), Nucleocapsid (N), Membrane (M), and Envelope (E) – are encoded by the corresponding viral genes (S, N, M, and E, respectively) and are responsible for virion structure and assembly (Bai et al., 2021; Chen et al., 2020; Zhang, B. et al., 2022).

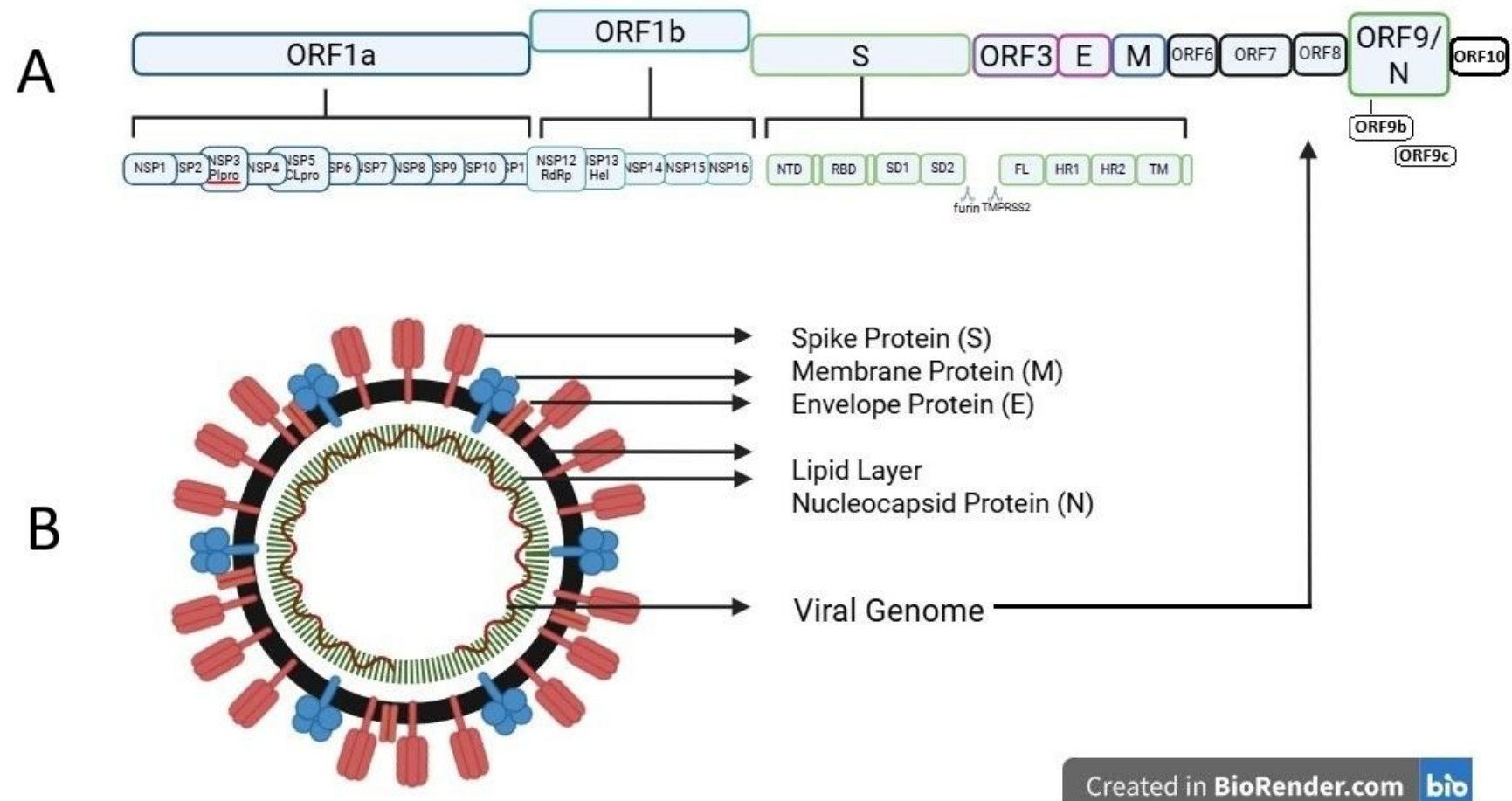


Figure 1. 1: The Structure and Genome Composition of SARS-CoV-2

Figure A depicts the genome composition of SARS-CoV-2 consisting mostly of ORF1a and ORF1b alongside the 4 structural proteins and additional ORF's that produce accessory proteins.

Figure B displays the viral structure created by the structural proteins where RNA associates with N protein and is then encapsulated with a lipid layer composed of the membrane, spike, and envelope proteins.

Adapted from Zhang et al. (2021), V'kovski et al. (2021), and Abdolmaleki et al. (2022). Created with Biorender.com.

The structure of the SARS-CoV-2 viral particle is depicted in [Figure 1. 1.B](#) demonstrating the RNA associated with the Nucleocapsid (N) protein, which in turn is associated and enclosed by a membrane alongside the Membrane (M) protein, Envelope (E) protein, and Spike (S) protein protruding from its surface (Adapted from Zhang et al. (2021), V'kovski et al. (2021), and Abdolmaleki et al. (2022)).

The N protein binds and organises viral nucleic acid within the capsid is therefore responsible for virion assembly and transcription (Chen et al., 2020; Luan et al., 2022), whilst supporting immune evasion by inhibiting the production of type I and type III interferons (IFN) (Fani et al., 2020; Kopecky-Bromberg et al., 2007; Mu et al., 2020). The M protein facilitates viral assembly and enhances the membrane curvature and structural integrity (Kaur & Gupta, 2020), the E protein participates in particle release and pathogenesis (Fani et al., 2020), and the S protein binds to host cell receptors and mediates viral entry (Millet & Whittaker, 2015; Phan, 2020; Zhou, P. et al., 2020; Zhou, R. et al., 2020). ORF1a and ORF1b are translated into two polyproteins that, following proteolytic processing, result in 16 non-structural proteins [NSPs (1 – 16)] (Zhang, Q. et al., 2021). These NSPs include the RNA-dependent RNA polymerase (RdRp) and the Replication and Transcription Complex (RTC), which mediate RNA-processing, modification, and proof-reading (V'kovski et al., 2021; Zhang, Q. et al., 2021). Together, these proteins create the required structure and functions that enable its viral life cycle.

1.2.2 The SARS-CoV-2 Replication Cycle

The steps of the viral life cycle are depicted in [Figure 1. 2](#). Viral particles are transported from an infected individual to a non-infected individual via droplets, aerosols, or fomites, and are inhaled, entering the body through the nasal-oral cavity (V'kovski et al., 2021). Virions travel to cells where their S protein binds to host Angiotensin-Converting Enzyme 2 (ACE-2) cell receptors (Moreno et al., 2022; Salamanna et al., 2020; V'kovski et al., 2021) for cellular entry. Since SARS-CoV-2 infects cells that lack ACE-2, it is believed that alternative ACE-2-independent pathways may exist. Lim et al. (2022) has identified CD147, AXL, Low density lipoprotein receptor class A 3 (LDLRAD3), TMEM30SA, KREMEN1, ASGR1 receptors, and certain C-type lectin receptors. Although the exact mechanisms behind these interactions are unknown, some may facilitate viral entry, rather than acting as

primary entry receptors, and these interactions may be context-dependent or site specific.

The binding process, membrane fusion, and entry of virion particles occur via an endosomal pathway, and this requires the pre-cleavage of the S protein at its polybasic site, located between the S1 and S2 subunits (Letko et al., 2020).

This process is followed by the uncoating and the dissociation of viral RNA from the N protein, resulting in the release of ++ssRNA into the cytoplasm, where host cell machinery directly translates it through different ribosomal frameshifts, resulting in ORF1a and ORF1b that generate the two polyproteins (pp1a and pp1ab) (V'kovski et al., 2021). These polyproteins undergo co-translational and post-translational modifications (Cheng et al., 2023; V'kovski et al., 2021) and are then cleaved by NSP3 (Papain-like protease, PL^{pro}) and NSP5 (3C-like, M^{pro}), creating the 16 non-structural proteins that constitute the viral replicase/transcriptase proteins collectively known as Non-structural Proteins (NSP 1 – 16) (V'kovski et al., 2021). NSP12 – 16 form the viral RTC (V'kovski et al., 2021). NSPs 2 – 11 perform supporting functions to the RTC and contribute to host immune evasion (V'kovski et al., 2021). Together, these NSPs both replicate and transcribe the genome, producing more genomic RNA (+) and sub-genomic (sg) mRNAs through discontinuous replication and Transcription Regulatory Sequences (V'kovski et al., 2021). The sg mRNAs then further undergo translation within the rough Endoplasmic Reticulum (ER) via a secretory pathway, ultimately creating the structural proteins of the virus (V'kovski et al., 2021). These, together with genomic RNA (+), are then packaged into vesicles through its recognition by the N protein and released as progeny viral particles through exocytosis (V'kovski et al., 2021).

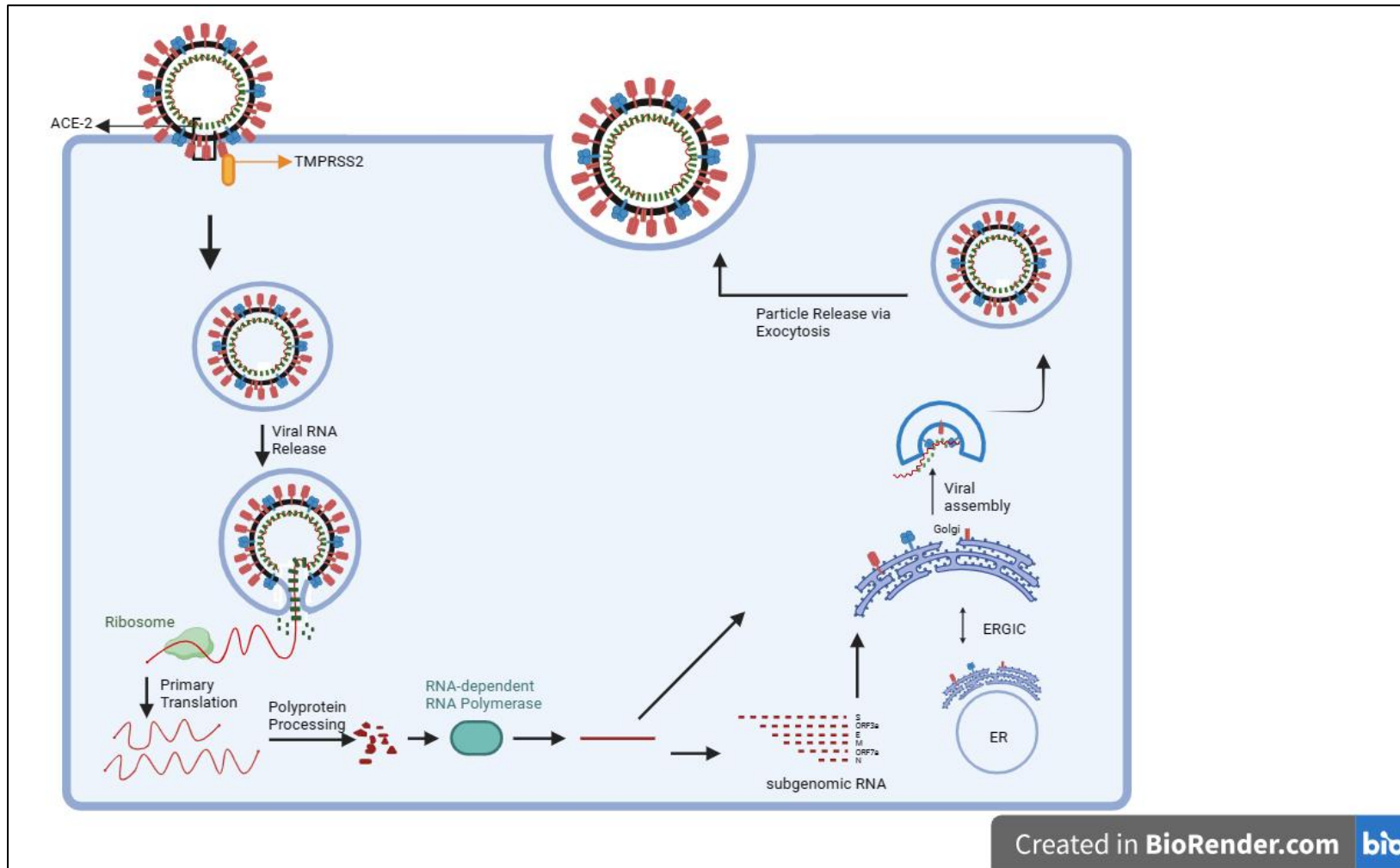


Figure 1. 2: The Replication Cycle of SARS-CoV-2

This figure schematically represents the life and replication cycle of SARS-CoV-2 starting with cellular attachment through the interaction of its S protein with ACE-2 cellular membrane receptors, followed by endosomal cellular entry. Virion-to-endosome membrane fusion allows its uncoating and the release of viral genomic RNA into the host cell cytoplasm where it undergoes primary translation by host cell machinery resulting in the formation of the viral replicase / translation complex that translate the viral genome, creating structural proteins. These along with viral RNA assemble and exit the cell via endocytosis. Polyprotein processing result in active viral replicative proteins.

1.3 Current SARS-CoV-2 Prevention and Treatment Strategies

While the initial pandemic subsided, viral spread remains prevalent, accompanied by viral mutations, including the emergence of major Variant of Concerns (VOC's) such as Alpha, Beta, Delta, and Omicron, which have been associated with altered transmissibility and immune escape, complicated by age-dependent fatalities and long-term effects, highlighting the need for advancements in the prevention and treatment of SARS-CoV-2 infection (Wu, S. et al., 2021; Zhu et al., 2021). Various prevention strategies are well known, starting with good hygiene practices that include handwashing, respiratory etiquette, surface cleaning, ventilation, and mask wearing. Prompt testing, self-isolation, as well as pre-exposure prophylaxis medications like pemivibart (Pemgarda) may further support preventative measures. Moreover, various vaccines have been well-established for the prevention of SARS-CoV-2 infection and have significantly reduced ICU admissions and mortality risks (Maleki et al., 2025), although some adverse effects, including neurological (Guillain–Barré syndrome, Bell's palsy), cardiovascular (myocarditis, pericarditis), and thrombotic events (cerebral venous sinus thrombosis) have been reported at low incidence rates (Bian et al., 2022; Faksova et al., 2024; Hosseini & Askari, 2023; Langgartner et al., 2023; Maleki et al., 2025; Meo et al., 2021).

Treatment of infection depends on the stage of infection at which the intervention takes place. During early to mid-stages of infection, replication is active, justifying the use of antiviral compounds that may prevent the infection from entering the inflammatory phase and decrease the risk of mortality (Moreno et al., 2022). During the middle to late stages of infection, inflammation is triggered, justifying the use of anti-inflammatories or immunomodulatory therapy, like dexamethasone or tocilizumab (Moreno et al., 2022).

Although this virus is novel and we have limited understanding of the mechanisms behind the infection and disease it causes a few treatments have been approved for use against its infection (FDA, 2024). During the pandemic, a variety of compounds were investigated for treatment of SARS-CoV-2 infection, such as chloroquine/hydroxychloroquine, ritonavir, interferon, ivermectin and nitazoxanide (Moreno et al., 2022; Rozencwajg et al., 2021). Unfortunately, these compounds were withdrawn after clinical trials suggested no significant benefit (Moreno et al., 2022).

Remdesivir was approved in the European Union (EU) and the United States of America (USA) for emergency use in SARS-CoV treatment, which is an adenosine analogue prodrug that inhibits viral RNA polymerases (Moreno et al., 2022). Currently, a variety of other drugs are undergoing different stages of clinical development and trials, some of which have shown to decrease hospitalisation and death in SARS-CoV-2 infections. Compounds investigated targeted various stages of the viral life cycle, for example, camostat mesylate and afamostat mesylate may inhibit TMPRSS2 during viral entry (V'kovski et al., 2021). Mastinib, an already existing drug, has been shown to inhibit 3CLpro, which is responsible for the cleavage of the polyproteins into their respective NSPs (Drayman et al., 2020).

The largest strategy of the prevention of SARS-CoV-2 infection involves the acquisition of herd immunity, during which a large enough portion of the population is protected such that transmission ceases, which for COVID-19, is believed to be at 60 to 70% (Abdolmaleki et al., 2022). Vaccination may assist in acquiring such immunity. Since the outbreak of the virus, Pfizer BNT 162b1 and Moderna mRNA-127 have been granted USFDA approval for emergency use (Meo et al., 2021). Current vaccines employ the S protein or its Receptor Binding Domain (RBD) as antigen sequences, however some studies showed reduced antibody binding and neutralisation occurring particularly in the omicron variants. The nonsynonymous mutations could potentially lead to reduced efficacy of these vaccines (Bai et al., 2021; Cao et al., 2022; Schubert et al., 2022). Similarly, a variety of vaccine compounds are undergoing development employing a variety of mechanisms including live-attenuated, viral vectors, Virus-Like Particles (VLP's), recombinant protein and nucleic acid-based vaccines, all of which carry their own pros and cons (Abdolmaleki et al., 2022; Rak et al., 2023). Such vaccines may be administered orally, intravenously, subcutaneously, or intranasally and are often used in conjunction with one another.

1.4 SARS-CoV-2 Nucleocapsid Protein

The SARS-CoV-2 Nucleocapsid (N) protein is encoded for by the ninth ORF, is 419 amino acids long, and has a theoretical molecular weight of 45,6 kDa (Lu et al., 2021). The N protein is divided into intrinsically ordered and intrinsically disordered regions (IDRs) represented in [Figure 1. 3.A](#).

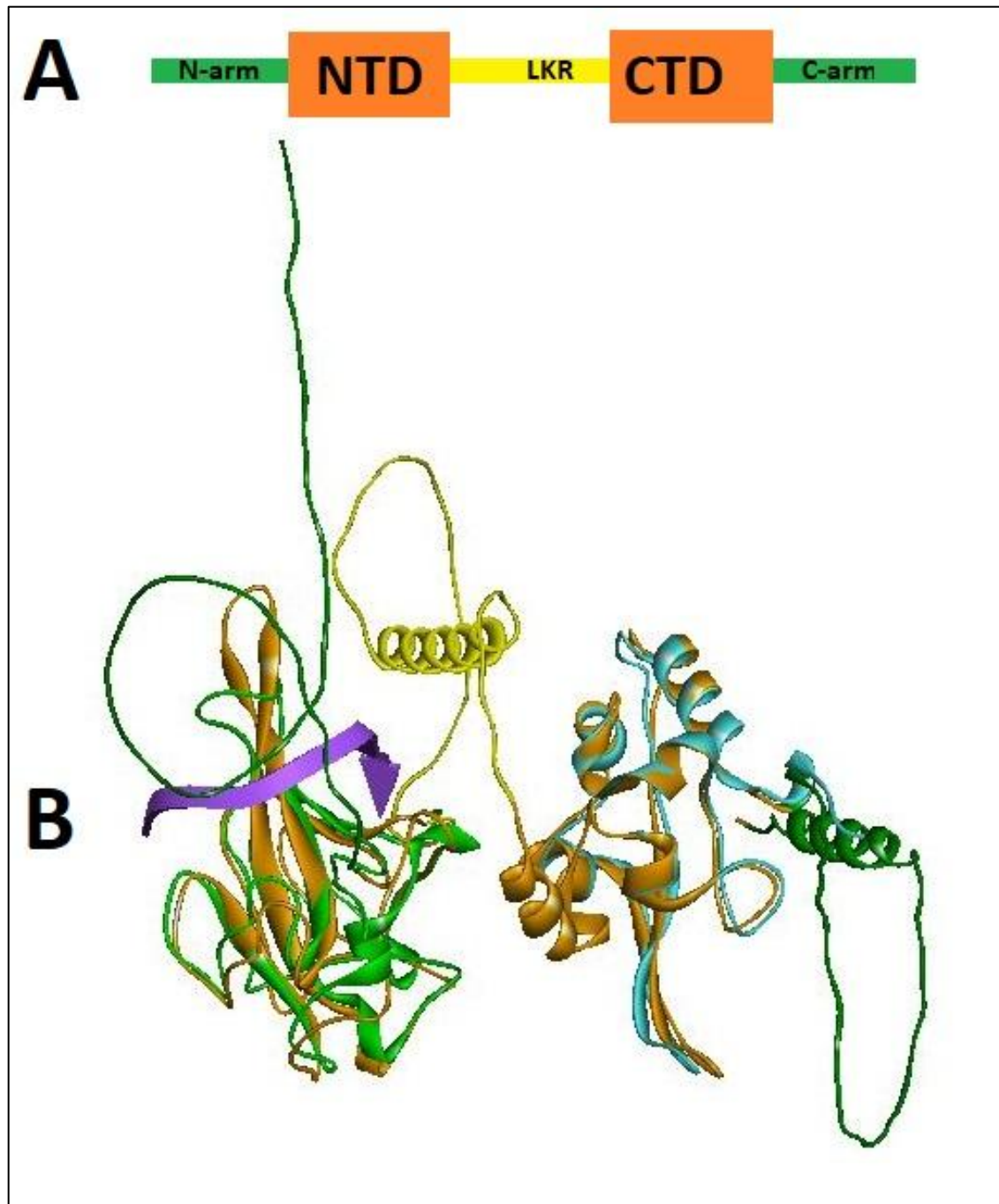


Figure 1. 3: The Primary and three-dimensional Structure of the SARS-CoV-2 Nucleocapsid Protein

Figure A display the linear arrangement of the SARS-CoV-2 nucleocapsid protein and its division into two RNA binding domains, the N-Terminal Domain (NTD) and C-Terminal Domain (CTD) that is separated by the flexible linker (LKR) and tailed by the N- and C-arms.

Figure B represents the three-dimensional organisation of the SARS-CoV-2 Nucleocapsid Protein. It displays the experimentally determined structures of the NTD (PDB: 7ACT (Dinesh et al., 2020); lime) with its RNA (purple) as well as CTD (PDB: 7XWX (Luan et al., 2022); turquoise) superimposed using Discovery Studio version 22.1.100.22290 (BIOVIA, Dassault Systemes, Discovery Studio Modeling Environment, Release 2022, Dassault Systemes: San Diego) to a full-length structure of the Wuhan sequence generated by AlphaFold. Here, the domains are in orange, the LKR in yellow, and the N- and C-arms are in green.

The ordered regions are considered to be the functional regions of the protein, consisting of two domains, the N-Terminal Domain (NTD; residues 48 - 174) and the C-Terminal Domain (CTD; residues 247 - 364) (Bai et al., 2021; Zeng, W. et al., 2020). These domains are separated by the intrinsically disordered Ser/Arg-rich flexible linker (LKR) region (residues 176 – 206) and tailed by the disordered N-arm (residues 1 – 51) and C-arm (residues 361 – 419), respectively (Bai et al., 2021; Schiavina et al., 2022; Zeng, W. et al., 2020) (**Figure 1. 3.B**).

Due to the largely disordered nature of the N protein (Schiavina et al., 2022), no full-length three-dimensional structure has been published, only that of the individual domains (PDB: 7VBD). In addition, some structures exist that contain partial coverage of oligomeric structures (PDB: 8R6E; 6WKP) or interactions with a variety of ligands (PDB: 7R98), particularly nucleic acids (PDB: 7ACS; 7ACT). The disordered nature of the protein contributes to the protein's flexibility and allosteric capability, as well as potential to bind to a wide variety of partners (Kang et al., 2020; Liu et al., 2021; Schiavina et al., 2022; Zeng, W. et al., 2020; Zhao et al., 2022). Together, these features enable its pleiotropic nature, contributing greatly to the entire viral life cycle through various functions.

The N protein is found beneath the viral envelope, tightly bound to the viral genomic RNA, forming a helically organised ribonucleoprotein (RNP) complex that is densely packed (Tarczewska et al., 2021). The RNP complex formation depends on the ordered regions, the NTD and CTD, considered to be the functional domains of the protein, that perform the RNA binding and dimerization, respectively (Chand et al., 2020; Chauhan et al., 2021; Forsythe et al., 2021; Wu, C. et al., 2021). The N protein binds nucleic acid through oligomerisation and has been found to exist in multiple states, allowing these functions to highly affect one another (Chang, Y. et al., 2021; Cubuk, Alston et al., 2024; Wu, J. et al., 2023; Zhao et al., 2021, 2024).

The NTD is shaped like a right-handed fist and consists of a core β -fold, a short helix, and β -hairpin helices (Caruso et al., 2021; Cubuk, Alston et al., 2024; Luan et al., 2022; Zeng, W. et al., 2020). The central β -core is positively charged, and, located between the β -hairpin and the β -core is a positively charged pocket (Arg R92, R107, and R149), each capable of binding RNA and neutralizing the phosphate groups

(McBride et al., 2014; Wu, W. et al., 2023). Interestingly, however, all regions are now considered to have RNA-binding capacity (Cubuk, Alston et al., 2024; Wu, W. et al., 2023).

The CTD in turn is mostly found as a diamond-shaped dimer, each monomer consisting of five α -helices, two 3_{10} helices and two β -strands. The two β -strands dimerise by inserting into one another, then binding through hydrogen and hydrophobic interactions (McBride et al., 2014; Wu, C. et al., 2021; Yang, Mei et al., 2021).

In short, the nucleic acid binding, dimerization, and ultimately RNP formation of the N protein already describes its role in genome protection, RNA chaperoning, cellular entry and uncoating, formation of RTCs, regulation of transcription and RNA processing, and ultimately genome packaging and virion assembly (Bai et al., 2021; Carlson et al., 2020; Cubuk, Alston et al., 2024; Lu et al., 2021; Surjit & Lal, 2010; Wu, J. et al., 2023; Ye et al., 2020; Zeng, W. et al., 2020). The complexity of the N protein is not limited by its IDRs and RNP formation; however, it is further extended by the multiple post-translational modification sites throughout the protein (Carlson et al., 2020; Cheng et al., 2023; Hatakeyama et al., 2021; Madahar et al., 2023; Supekar et al., 2021; Zeng, W. et al., 2020), enabling a great contribution to the viral pathogenesis and fitness through modulation of inflammation, immune evasion, and liquid-liquid phase separation (LLPS) that form liquid-like condensates (Stuwe et al., 2024).

Since the emergence of SARS-CoV-2, numerous variants have been identified. Some variants have been designated as either a Variant Under Monitoring (VUM), Variant of Interest (VOI), or Variant of Concern (VOC), some of which have been associated with mutations occurring within the nucleocapsid protein. Notably, the R203K and G204R co-occurring mutations within the SR linker region of the N protein have been associated with increased viral infectivity (Muradyan et al., 2024; Nikolakaki & Giannakouros, 2020; Shuaib et al., 2023; Wu, H., Xing, Meng, Fu, Xue, Dong, Tang et al., 2021). These mutations were prevalent in several VOCs, including Alpha, Beta, and Gamma (Rahman et al., 2021). The Delta variant displayed mutations including D63G, R203M, and D377Y, reported to contribute to enhanced viral replication (Alsuwairi et al., 2023; Hagag et al., 2022; Li, Y. et al., 2025). Mutations within the

Omicron variant included the P13L, deletion of ERS 31-33, and S413R (Alsuwairi et al., 2023; Muradyan et al., 2024; Nguyen et al., 2024; Oulas et al., 2021; Savellini et al., 2023). These may similarly influence viral behaviour.

The N protein has attracted a lot of scientific attention due to its relatively high degree of conservation across the variants, despite the occurrence of several mutation hotspots, particularly within the linker region, and is vital for immune recognition, results in higher and faster antibody response rates in patients, is used for diagnostic testing and plays a large role in the entire viral life cycle (Bai et al., 2021; Oronsky et al., 2022; Zeng, W. et al., 2020). These garner not only the necessity for continuous monitoring, but also the probable use as a target for therapeutics and vaccination that could provide immunity against novel variants.

The binding of compounds could interrupt protein function at various points, for example, compounds may block RNA binding, oligomerization, and phosphorylation of the protein that will ultimately be translated into its functions (Bouhaddou et al., 2020; Cai et al., 2021; Wu, S. et al., 2021). A variety of drugs are currently being investigated for their interaction with the N protein, some of which are from medicinal plants (Matsuo, 2021). Royster et al. (2023) demonstrated the ability of a novel compound, called K31, to bind to the N protein, inhibiting it from binding to the 5' terminus of viral genomic RNA, thereby inhibiting its replication cycle. Preliminary investigations provide promise to produce potential antiviral therapies by targeting the N protein.

However, monitoring and research of the N protein requires its expression and purification. Over time, researchers have utilized multiple expression systems, each with their own advantages, challenges, and optimisation strategies depending on the goal. Prokaryotic, *Escherichia coli* BL21(DE3) expression systems are frequently explored alongside plasmid constructs that allow for the incorporation of N-terminal Histidine tags that are used for efficient purification. Unfortunately, however, these high-level expression systems are often reported to affect the solubility of the protein and lead to the formation of inclusion bodies (Brudenell et al., 2024; Li, G. et al., 2021). These concerns are often addressed by using either lower induction temperatures (e.g., 16°C) or an unfolding-refolding strategy employing denaturants such as urea or

guanidine hydrochloride. In addition, owing to the high isoelectric point (± 10.07) as well as strong nucleic acid binding affinity, the protein is often co-purified with host cell nucleic acid (Porterfield & Zlotnick, 2010; Vidal et al., 2023), similarly addressed with an unfolding refolding strategy or high sodium chloride content. Insect or mammalian cell expression systems have been explored since they offer either eukaryotic or human-like post-translational modifications and protein folding systems, likely required for this extensively post-translationally modified protein (Brudenell et al., 2024; Vidal et al., 2023). However, the choice of the system used depends on the intended application, as well as the balancing of factors such as yield, functionality, and structural fidelity.

1.5 *Berberis vulgaris* – A Herbal Therapy

A variety of therapies have been discovered through medicinal plants, some especially for cold and flu. Since the pandemic, a variety of herbal remedies have been used for the treatment of SARS-CoV-2. To prevent and treat SARS-CoV-2 infection, Dr Chad James, a herbal medicines practitioner in South Africa, used a *Berberis vulgaris* extract, often used for gastrointestinal, urinary, and respiratory disorders (van Wyk, 2008). Analysis of the extract determined that it contains bervulcine, vulracine, berberrubine, lambertine, berbidine, obamegine, berbamine in small concentrations along with large concentrations of berberine and palmatine (Lambertine being the third highest) (Gildenhuys et al., 2023). Palmatine displayed higher concentrations in glycerol extracts, which were predominantly used for children, whereas ethanol extracts, used for adults, displayed higher concentrations of berberine.

Similar extracts, as well as the individual compounds found within them, have demonstrated anti-cancer, anti-oxidation, anti-inflammatory, anti-bacterial, antiviral, and neuroprotective properties often used to treat diarrhea, gastroenteritis, skin infection, fever, dysentery, and hepatic ailments (Ho et al., 2019; Long et al., 2019; Warowicka et al., 2020).

Berberine and palmatine, the highest concentration compounds within the extracts, are related isoquinoline alkaloids that share a tetracyclic protoberberine core structure (**Figure 1. 4**), typically occurring as chloride salts within herbal extracts. Both

compounds exhibit broad pharmacological activities, including antiviral, anti-inflammatory, and immunomodulatory effects (Jia et al., 2010; Long et al., 2019; Warowicka et al., 2020; Yan et al., 2017). Berberine is the most extensively studied owing to its great anti-inflammatory properties, characterised by reductions in IL-6, IL-8, IL-13, TNF- α , Interferon- γ , M1, and M2-mediated inflammatory actions that result in immune support and viral clearance. These actions are reported to occur through actions on mitogen-activated protein kinase extracellular signal-regulated kinase/mechanistic target of rapamycin (MEK-ERK)/mTor, the nuclear factor kappa-light-chain enhancer of activated B cells (NF- κ B), the activator protein 1 (AP-1) signalling, as well as the mitogen-activated protein kinases (P38MAPK) proinflammatory pathways (Long et al., 2019; Warowicka et al., 2020).

Palmitine has not been well-researched but has been shown to cause similar inhibition of inflammatory signalling, particularly by down-regulation of TLR4-TRIF-NF- κ B pathways and a reduction in release of TNF- α /IL-6 (Yan et al., 2017). Literature demonstrates the antiviral activity of both compounds against diverse viruses like influenza, hepatitis, and herpes, as well as anti-oxidant and anti-inflammatory actions (Yan et al., 2017), making them attractive candidates for COVID-19 therapy.

Both compounds have recently been studied in relation to SARS-CoV-2 infection directly. Berberine suppresses ACE2 and TMPRSS2 expression in epithelial cells, potentially reducing viral entry (Das et al., 2023). Berberine has demonstrated SARS-CoV-2 replication inhibition capabilities in primary human nasal and airway epithelial cells (Varghese et al., 2021). Furthermore, Berberine is believed to influence host factors like TP53 (p53), SREBF1, and TOP1, impacting antiviral immunity, lipid metabolism, and other inflammatory responses (Das et al., 2023). Both compounds are believed to have intercalation capabilities that may inhibit SARS-CoV-2 replication (He et al., 2023; Wink, 2020). Palmitine has been shown to reduce the expression of chemokine coding genes, IL6R and Vascular Endothelial Growth Factor A (VEGFA), all of which have been associated with increased severity of disease caused by the SARS-CoV-2 virus (Yan et al., 2017).

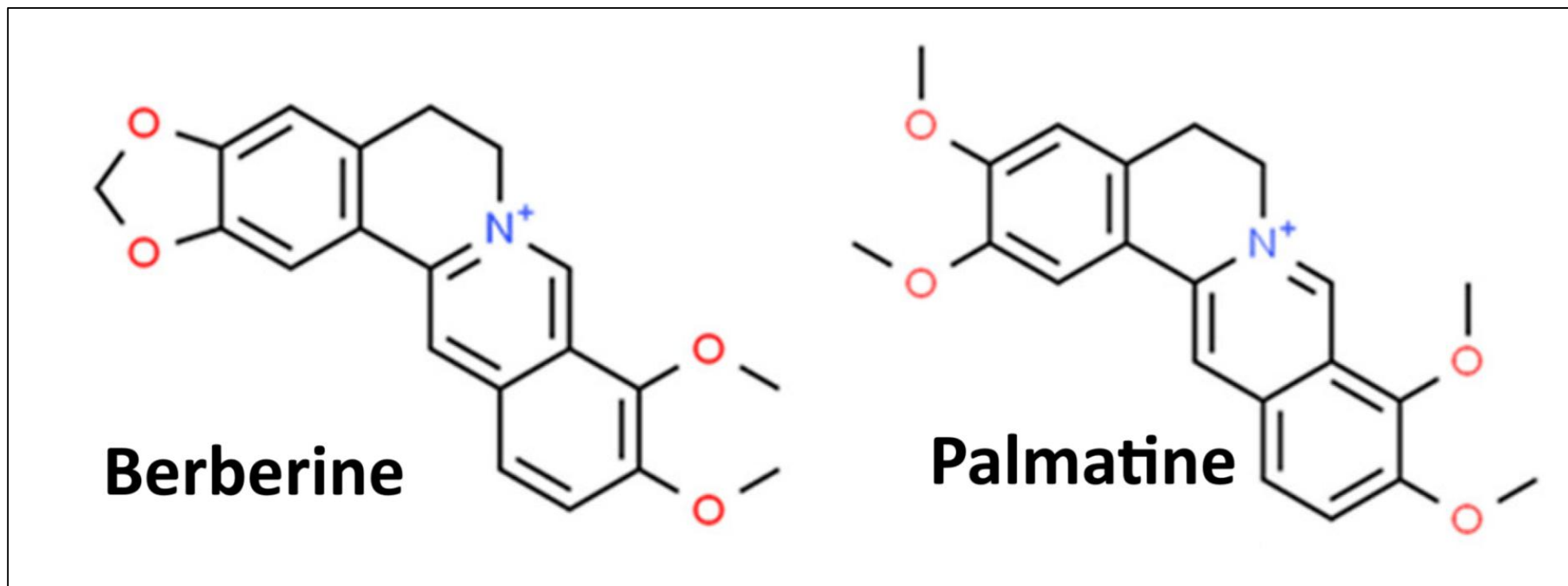


Figure 1. 4: The 2D Chemical Structures of Berberine and Palmatine

Figure A represents the chemical structure of berberine (National Center for Biotechnology Information (2023). PubChem Compound Summary for CID 2353, Berberine. Retrieved September 19, 2023 from <https://pubchem.ncbi.nlm.nih.gov/compound/Berberine>).

Figure B represents that of palmatine (National Center for Biotechnology Information (2023). PubChem Compound Summary for CID 19009, Palmatine. Retrieved September 19, 2023 from <https://pubchem.ncbi.nlm.nih.gov/compound/Palmatine>)

Palmatine has been shown to elevate the expression of the AHR gene, which is associated with the SARS-CoV-2 viral entry and disease progression and is believed to decrease the number of ACE-2 receptors (Pei et al., 2022, 2023; Zeng, J. et al., 2024). Furthermore, various computational docking (in silico) experiments predict potential binding between both palmatine and berberine (Joshi *et al.* 2021) with the SARS-CoV-2 main protease (M^{pro}), RNA-dependent RNA polymerase (RdRp) and Spike (S) glycoprotein-Human ACE2 complex (Antonopoulou et al., 2022; Joshi, A. et al., 2021; Joshi, T. et al., 2021; Majnooni et al., 2021; Narkhede et al., 2020; Oner et al., 2023), suggesting likely direct antiviral effects.

These inflammatory, immunological, and oxidative qualities displayed by both berberine and palmatine are required for the treatment of SARS-CoV-2 infections, and since these compounds could bind viral particles, proteins, and viral nucleic acids, blocking entry and replication, and modulate host cell signalling, they may provide an cheap and effective treatment to dampen overactive inflammatory responses, reducing viral replication and mitigate lung damage in COVID-19 infection.

2. Aims & Objectives

The principal aim of the study was to generate a full-length SARS-CoV-2 Nucleocapsid protein incorporating the most frequently occurring mutations in known variants to date. In addition, the study aimed to investigate the potential interactions between berberine and palmatine with the SARS-CoV-2 N protein.

To achieve these aims, the following objectives were addressed:

- Producing a consensus sequence of SARS-CoV-2 N protein containing the most frequently occurring mutations known to date.
- The production of a NP-pET-15b vector plasmid and its use for the expression and purification of the recombinant N protein using an optimised *E. coli* T7 Express based bacterial system.
- In silico screening of the binding activity of berberine and palmatine to the SARS-CoV-2 N protein through computational docking simulations.
- Identification of the isolated protein followed by characterisation of the effects of berberine and palmatine on N protein secondary, tertiary, and quaternary structure using techniques like Far-UV-CD, Intrinsic Tryptophan Fluorescence, Attenuated Total Reflectance Fourier Transformed Infrared Spectroscopy (ATR-FTIR), and Blue Native PAGE.
- Thermodynamic assessment and elucidation of binding affinities, stoichiometry, enthalpic and entropic contribution of complex formation between berberine and palmatine with the N protein using Isothermal Titration Calorimetry (ITC) and Surface Plasmon Resonance (SPR).
- Characterisation of the thermal stability of the N protein in the absence and presence of berberine and palmatine using Far-UV-CD and Intrinsic Tryptophan Fluorescence.

3. Experimental Procedures

3.1 Materials

The NP-pET-15b (LLC) was produced by Genscript® (Piscataway, NJ, USA). *E. coli* JM109, BL21(DE3)-T7 Express and NiCo21(DE3) cells were purchased from New England Biolabs (Ipswich, MA, USA). The isopropyl β -D-1-thiogalactopyranoside (IPTG) inducer, ampicillin, urea, Protease Inhibitor Cocktail VII (for HisTag Sequences), Bovine Serum Albumin (BSA), and palmatine were purchased from Melford (Ipswich, UK). Both the SDS-PAGE (#26610) and Native PAGE Molecular Weight Markers (#57030) (MWM) were purchased from Thermo Fischer Scientific (Waltham, MA, USA). Bradford's reagent was purchased from TaKaRa Bio (Kasatsu, Shiga, Japan). Berberine was purchased from Sigma-Aldrich (St Louis, MO, USA). The SP Sepharose Ion-Exchange resin was purchased from Merck (Rahway, NJ, USA), whilst the Nickel Affinity Chromatography HisTrap HP column, as well as the HiTrap Desalting column, were purchased from Cytiva (Marlborough, MA, USA). All other chemicals were of analytical grade.

3.2 Software

In this study, various specialised software tools were used in a repeating fashion to facilitate data acquisition, analysis, and the visualisation of results. Microsoft (MS) Excel (version 2504; Build 18730.20142, Microsoft, Redmond, WA, USA) was used to generate the bacterial growth curve, chromatography, and Stern-Volmer figures while or OriginLab (Origin Pro 2025 64-bit SRI Learning edition 10.2.0.196, OriginLab corporation, Northampton, MA, USA) was used to visualize all spectrophotometric results. These platforms enabled efficient data organisation and publication-quality figure creation required for this study.

The electrophoretic analyses performed in this study, SDS-PAGE and Native PAGE, were acquired using the Bio-Rad Gel Doc™ XR+ Molecular Imager along with the instrument-associated software called ImageLab (version 6, Bio-Rad Laboratories, Hercules, CA, USA).

The real-time monitoring and data acquisitions of the chromatographic techniques executed using the AKTA Prime Plus system (Cat. 11001313; GE Healthcare,

Chicago, IL, USA) and data was managed using the PrimeView and PrimeView Evaluation Software (Version 5.31, GE Healthcare, Chicago, IL, USA).

The spectroscopic measurements, more precisely the absorbance, fluorescence, and CD techniques performed on the Chirascan-plus spectrometer (Applied Photophysics Ltd., Leatherhead, UK) were managed through the Chirascan Control Panel Application Viewer (Version 4.7.0194), while data visualization and processing were conducted using the Pro-Data Viewer Application (Version 4.7.0194). Data generated here were subsequently converted for CSV export using the APL Data Converter (Version 4.5.0.11), allowing seamless integration of spectral data into downstream analytical workflows.

Additional software tools employed in this study are detailed within their respective methodological sections, corresponding to the specific experimental procedures or analyses for which they were utilized.

3.3. Methods

3.3.1 NP Consensus Sequence Construction

To ensure broad variant coverage in laboratory studies, capturing the most frequent mutations across all known variants, a consensus sequence of the nucleocapsid protein was generated using the majority rule method, whereby the most frequently occurring amino acid residue at each alignment position across all sequences was selected to produce the final consensus sequence.

A full list containing all lineages was obtained from GitHub (GitHub, 2024) and sorted alphabetically. A single amino acid sequence was obtained for each lineage in FASTA format from the NCBI Virus database (<https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/>; Last Accessed: 03/04/2024) using both a systematic and randomized approach. First, the search criteria were [refined](#) to include specifically complete sequences pertaining to taxon ID 2697049 (SARS-CoV-2) nucleocapsid phosphoprotein with a length of 415 to 425 amino acids isolated from humans. Then the distinct lineage was selected, and a single random sequence was downloaded from the results using the randomised setting on the database. The total of 2186 sequences were aligned using the MAFFT

online tool (<https://mafft.cbrc.jp/alignment/server/index.html>; Accessed 03/04/2024; (Kato et al., 2019) using default settings and a consensus sequence was subsequently generated using the EMBOSS tool (https://www.ebi.ac.uk/jdispatcher/msa/emboss_cons; Accessed 03/04/2024; (Madeira et al., 2024; Rice et al., 2000) at default plurality and setcase, with an identity of 1 and the BLOSUM62 matrix. This process was repeated to ensure reproducibility, and the resulting two sequences were visualised through alignment using the [Clustal Omega program](#) via UniProt Align ([Align | UniProt](#), Accessed 03/04/2024). This sequence was used as the final consensus sequence seen in **Figure 4. 1.A** (W-NP).

3.3.2 Protein Modelling

Three-dimensional structures for the Wuhan sequence (P0DTC9) and the W-NP consensus sequence from above were subjected to structural modelling. Domain boundaries were delineated based on established literature, with residues 48-174 defining the NTD and 247-364 defining the CTD. The sequences were extracted in FASTA format to facilitate subsequent structural modelling.

Models were generated for the NTD and CTD domains as well as full-length monomer and dimer structures for both the Wuhan sequence and W-NP using [AlphaFold Colab](#) (Jumper et al., 2021) (Accessed: 31/03/2025), an AI-driven template-less modelling tool. All monomeric and dimeric models were generated using `alphafold2_ptm` and `alphafold2_multimer_v3` models, respectively. The highest confidence (rank 1) structure, indicated by the per-residue pLDDT (predicted Local Distance Difference Test) score used to assess the reliability (Jumper et al., 2021), was used for further analysis.

Complementary to AlphaFold predictions, homology models for the NTD and CTD were constructed using the SWISS-MODEL server (Waterhouse et al., 2018) ([SWISS-MODEL; Accessed: 05/26/2025](#)) utilising the consensus-derived domain sequences. SWISS-MODEL predicts protein structures by aligning a target sequence to known homologous structures, whereas AlphaFold employs deep learning to predict structures directly from sequence data without relying on templates. Based on the Global Model Quality Estimation (GMQE) and Qualitative Model Energy Analysis

(QMEAN) quality scores Benkert et al. (2011), 9CJ6 and 6ZCO were selected as templates for the NTD and CTD, respectively. These models served to validate and complement the AlphaFold predictions, particularly in regions where AlphaFold's confidence were low, confirming the disordered content within those regions (Ruff & Pappu, 2021). The finalised models were visualized and analyzed for Solvent Accessibility Scores using Discovery Studio (Discovery Studio Visualizer, v20.1.0.19295, BIOVIA, Dassault Systèmes, San Diego, CA, USA).

3.3.3 Plasmid

A codon-optimised complementary DNA (cDNA) sequence of W-NP (the nucleocapsid protein consensus sequence as determined in [Section 3.3.1](#)) was synthesised to enhance expression efficiency within *Escherichia coli* cells by matching the codon usage preferences to those of the host (Gustafsson et al., 2004). This synthetic gene was then cloned between the NdeI and BamHI restriction sites of a pET-15b plasmid by Genscript® (Piscataway, NJ, USA). The resulting construct is referred to as NP-pET-15b from hereon.

The pET-15b plasmid contains various features that allow the facilitation of recombinant protein expression and purification (Novagen, 2002). This plasmid contains an ampicillin resistance gene, allowing for selective growth of successfully transformed bacterial cells in the presence of ampicillin. It also contains a T7 promoter that is recognised by the T7 RNA polymerase. The lacUV5 promoter controlling the chromosomally encoded T7 RNA polymerase gene may be induced using IPTG. Together, these features enable not only high-level transcription of the target gene, but also the power to control protein expression during desired growth phases, minimising the metabolic burden on host cells. In addition, it also contains a N-terminal hexahistidine (6x His) tag that enables simple and effective purification using immobilised metal affinity chromatography and its associated thrombin cleavage site that allows for the enzymatic removal of the histidine tag, together, yielding pure product close to its native state. In this study, however, the his-tag was not removed.

3.3.4 Cellular Competency and Transformation

First, *E. coli* JM109 cells, a cloning strain optimised for recombinant DNA propagation and high-yield plasmid amplification, were transformed using the synthesised plasmid termed W-NP, creating a plasmid amplification bacterial system. The plasmid DNA amplified by these cells was extracted and subsequently used to transform *E. coli* NiCo21(DE3) and *E. coli* BL21(DE3)-T7 Express, protein-expressing cells. The same methodology for the preparation of competent cells, transformation and growth curve assessment was applied across all cell lines used during this study.

To incorporate recombinant plasmid DNA, each cell line was subjected to calcium chloride chemical-based competence (Mandel & Higa, 1970) and heat shock (Chang, Y. et al., 2021). Here, the manufactured cells were grown in Luria-Bertani (LB) media (10g/l tryptone; 5g/l yeast; 5g/l NaCl) in a shaking incubator rotating at 220 RPM at 37 °C to an OD_{600 nm} of 0.5 and collected by centrifugation (3000 x *g* for 10 minutes at 4 °C). The collected cells were washed using cold 0.1 M calcium chloride and placed on ice for 30 minutes, then finally stored at –80 °C in a 0.1 M calcium chloride solution containing 30% (v/v) glycerol until needed.

The transformations were achieved using the heat-shock method described by Hanahan (1983). Competent cells were thawed on ice for 30 minutes, followed by the addition of plasmid DNA. Specifically, 4 µg/ml of both synthesised and extracted nucleocapsid phosphoprotein-containing DNA was introduced to each cell line tube. A negative and positive control was included (Lim *et al.* 2015), utilising either no DNA or the addition of 30 µg/ml of manufactured pUC-19 plasmid DNA to the competent cells. The cell and DNA mixtures were placed on ice for 30 minutes, then heated to 42 °C for 60 seconds, and placed back on ice for 2 minutes. Nine hundred microlitres of LB growth medium was added, and cells were allowed to grow for 70 minutes (220 RPM at 37 °C). From the resulting cell culture, 100 µl were plated directly as below, whilst the remainder underwent concentration via centrifugation (3000 x *g* for 10 minutes at 4 °C). Here, 800 µl of supernatant was removed and the pellet was resuspended in the remaining 100 µl of LB. The pelleted and resuspended cells were then spread on LB agar plates (1g/l Tryptone; 1g/l NaCl; 0.5g/l Yeast; 1.5g/l Agar) containing 150 µg/ml ampicillin using a flame-sterilised L-shaped spreader and then allowed to grow overnight (12-16 hours) at 37 °C. A single colony of successfully transformed cells

containing NP-pET-15b plasmid was selected from each cell line and cultured overnight (220 RPM at 37 °C) in 100 ml LB broth containing 100 µg/ml ampicillin.

A fresh 1:100 dilution culture was made from the overnight culture and incubated at the same conditions. The subculture was monitored spectrophotometrically in duplicate using an Implen P330 Nanophotometer (München, Germany) at 600 nm in 30-minute intervals for 7 hours. The early-to-mid-log phase was determined by analysing this data using MS Excel (version 2504; Build 18730.20142) during which a scatter plot was generated from all the datasets by plotting the absorbance (A) over the time (minutes).

In order to prepare glycerol stocks of each transformed cell line, cell cultures grown to early-to-mid-log phase ([Section 4.2.1](#)) were mixed at a 1:1 ratio with 30% (v/v) sterile glycerol and stored at –80 °C for future use.

E. coli JM109 cells were similarly sub-cultured and spectrophotometrically measured until an OD of 0.5, at which point, 5 ml culture was collected via centrifugation (3000 x g for 10 minutes at 4 °C). The DNA produced by these cells were extracted using the Zymo Pure Miniprep kit (Zymo Research, Irvine, CA, USA) as per instruction manual (Zymo Research, 2023). The concentration of the isolated plasmid was determined through spectrophotometric quantification using a Blue-Ray Biotech EzDrop 1000C (New Taipei City, Taiwan) during which the absorbance at 230, 260, and 280 nm in comparison to the elution buffer was measured, resulting in a somewhat pure yet high concentration plasmid.

The purified plasmid was then utilised to transform *E. coli* NiCo21(DE3) and BL21(DE3)-T7 Express cells to generate a bacterial protein expression system, following the established transformation protocol.

3.3.5 Protein Expression

3.3.5.1 Protein Expression Optimisation

The effect of temperature, IPTG concentrations, and protein expression time on the amount of recombinant nucleocapsid protein expressed by *E. coli* BL21(DE3)-T7 and

E. coli NiCo21(DE3) cells was assessed. Fifty microlitres of bacterial cell glycerol stock of the expression cell lines were added to 100 ml of LB broth supplemented with 100 µg/ml ampicillin in an Erlenmeyer flask, totalling a 1:5 culture-to-volume ratio. The cultures were grown overnight (12-16 hours) at 37 °C, shaking at 220 RPM and then sub-cultured at a 1:50 ratio in a fresh flask, similarly containing 100 µg/ml ampicillin. The sub-cultures were monitored spectrophotometrically, and protein expression was induced with either 0.1 mM, 0.4 mM, or 1 mM IPTG once the optical density reached between 0.4 and 0.6 described [Section 4.2.1](#). The cultures were allowed to express protein for 4, 6, and 8 hours as well as 12-16 hours at which point a 2 ml sample was collected via centrifugation (3000 x *g* for 10 min at 4 °C). Cultures were similarly prepared for the higher IPTG ranges (0.4 mM and 1 mM) but allowed to grow at 15 °C to which sample was collected 12 – 16 hours post induction.

The resulting cell samples were resuspended in 500 µl Resuspension Buffer (20 mM Tris-HCl, 1 M NaCl, pH = 8.00) and subjected to sonication. The resuspended cells were lysed on ice using a Q 700 sonicator (Qsonica, Newtown, CT, USA) equipped with a 220-05 conversion clamp and a 418-A probe tip. Each sample was sonicated 3 times for a total of 60 seconds, which consisted of 2-seconds on and 2-seconds off burst cycles at an amplitude of 35 kHz, releasing the cellular contents.

From the sonicated samples, 100 µl were used as crude whole cell samples. The remaining 400 µl underwent centrifugation (20 865 x *g* for 30 min at 4 °C) to facilitate separation into soluble and insoluble fractions released during sonication. The supernatant was sampled, and the insoluble fractions were resuspended in 400 µl of the same buffer. Each sample was prepared for analysis via SDS-PAGE as described in [Section 3.3.6](#).

3.3.5.2 Protein Expression for Chromatography

An *E. coli* BL21(DE3)-T7 express cell glycerol stock was prepared and sub-cultured as per [Section 3.3.5](#); however, 1.5 L of fresh LB was used for protein expression and induced at OD 0.4 - 0.6 using 0.1 mM IPTG. Cells were harvested 6 hours post-induction at 37 °C via centrifugation at 2831 x *g* at 4 °C for 20 minutes. The cell pellets were frozen at – 80 °C until needed.

When needed, the cell pellets were thawed and resuspended in 25 ml of the various Chromatography Buffer listed in [Table 3. 1](#), and an additional 10 mM magnesium chloride, DNase (4 U/mL final concentration; Cat. 10104159001; Merck; Rahway, NJ, USA), and 10 µL Protease Inhibitor Cocktail VII (for HisTag Sequences) (MEL-P51400.1) were added. The cellular mixtures were homogenized using an IKA (Staufen im Breisgau, Germany)® T25D digital ultra Turrax homogeniser at 1000 RPM for 60 seconds. The mixtures were then frozen at - 20 °C overnight (12 - 16 hours) and then later thawed on ice.

Table 3. 1: Chromatography Buffers

The contents within the various Buffer A's used in the different chromatography strategies explored.

Chromatography Buffer A	Contents
IEC Buffer A	20 mM Tris-HCl, pH 8.00
Nickel Strategy 1, 2, and 3 PBA	10 mM imidazole; 20 mM Tris-HCl; 1 M NaCl; pH 8.00
Nickel Strategy 4 PBA	45 mM imidazole; 20 mM Tris-HCl; 1 M NaCl; pH 8.00

The cells were lysed for chromatography by two rounds of sonication at 55 kHz sequenced 2-second burst and 2-second rest for 5 minutes using a QSonica Q 700 sonicator (Qsonica, Newtown, CT, USA) equipped with a 220-05 conversion clamp and a 418-A probe tip. The soluble protein was separated from the culture via centrifugation at 20 865 x g for 30 minutes at 4 °C. The supernatant was filtered using a 0.45 µm filter to remove particulates, then used to perform the Ion-Exchange and Nickel Affinity Chromatography procedures.

3.3.6 SDS-PAGE

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described by Laemmli (1970). SDS-PAGE separates proteins primarily according to molecular weight through the denaturation of proteins and uniform coating with SDS, resulting in migration through a polyacrylamide matrix under an electric field.

All samples (crude, soluble, and insoluble, as well as fractions from chromatography procedures) were mixed at a 1:1 ratio with 1x Laemmli Sample Buffer (LSB) [distilled water (50%), 0.5 M TRIS-HCl (pH = 6.8), glycerol (10% v/v), SDS (20%), β -mercaptoethanol (5%), Bromophenol Blue (2.5%)] and heated to 95 °C for 5 minutes. The sample-buffer mixtures were loaded on an SDS-PAGE prepared inhouse consisting of a 5% stacking gel [BisAcrylamide (5%), Tris-HCl (0.125 M, pH = 6.8), SDS (0.1%), APS (0.025%), TEMED (0.0004%)] and a 12% resolving gel [BisAcrylamide (12%), Tris-HCl (0.0375 M, pH = 8.8), SDS (0.1%), APS (0.05%), TEMED (0.0005%)]. Samples were run using a Running Buffer [Tris (0.3%), Glycine (1.44%), SDS (0.1%)] alongside an Unstained Protein MW Marker (ThermoFisher Scientific, Waltham, MA, Cat: 26610) containing β -galactosidase, Bovine Serum Albumin, Ovalbumin, Lactate Dehydrogenase, REase Bsp98I, β -lactoglobulin, Lysozyme. Together, it was electrophoresed until the buffer front reached the end of the gel (about 60 minutes) at a constant 140 Volts using a Bio-Rad Mini-Protean Tetra System (Hercules, CA, USA). The gels (7.2 cm x 8.6 cm) were stained for 30 minutes using InstantBlue® Coomassie Protein stain (Abcam, Cambridge, UK, cat. ISB1L, ab11921) or Coomassie stain (50% methanol, 0.05% R250 Coomassie Brilliant Blue, 10% Acetic Acid). InstantBlue-stained gels were rinsed for 10 minutes in distilled water, whilst R250 Coomassie-stained gels were then destained (5% ethanol, 7% Acetic Acid) overnight. The resulting gels were visualized using a Bio-Rad Gel Doc™ XR+ Molecular Imager along with its ImageLab (version 6, Bio-Rad) software. The images were created using a custom application that employs UV transmission illumination, a 685/55 filter, Coomassie image colour and 1x1 binning.

3.3.7 Chromatography

3.3.7.1 Ion Exchange Chromatography

Ion exchange chromatography, a technique that separates molecules based on their charge, was used to attempt to purify the W-NP protein using SP Sepharose® ion exchange resin (Cat. S1799; Merck; Rahway, NJ, USA) in conjunction with an AKTA Pure Prime plus Chromatography system (Cat. 11001313; GE Healthcare, Chicago, IL, USA). The resin used here is a strong cation exchanger that selectively binds proteins that have a net positive charge at the working pH and can later be displaced using either alternating pH or sodium (Na^+) counter ions. In this event, W-NP has a

high theoretical pI (9.950 - 10.13) with an estimated +25.776 net positive charge at pH 7.4.

The resin was equilibrated with 20% ethanol and manually packed into an XK16 column (Cat: 28988937; Cytiva, Marlborough, MA, USA) to a final bed height of 14.9 cm and total volume of 30 ml. The column was equilibrated with 10 column volumes (CV) of IEC Buffer A (20 mM Tris-HCl, pH 8.00), corresponding to a total volume of 300 ml, at a flow rate of 2 ml/min under 0.3 MPa pressure, until a stable UV baseline at 280 nm was achieved. Cell pellets were prepared as per [Section 3.3.5.2](#) with IEC Buffer A (20 mM Tris-HCl, pH 8.00), and the filtered lysate was injected onto the column using a 50 ml capacity Superloop™ (Cat: 18111382; Cytiva, Marlborough, MA, USA) at a flow rate of 0.5 ml/min. Loosely bound proteins were washed from the column in three steps, each consisting of 100 ml buffer at a flow rate of 1.5 ml/min. First, IEC Buffer A (20 mM Tris-HCl, pH 8.00), then IEC Buffer B (20 mM Tris-HCl, pH 9.00), then IEC Buffer C (20 mM Tris-HCl, pH 10.50). The column-bound proteins were eluted in 2.5 ml fractions using an 80 ml, 0% to 100%, linear gradient at 5 ml/min using IEC Buffer D (1 M NaCl; pH 8.00) and monitored at 280 nm.

Data processing followed the process described in [Section 3.2](#), and the fractions were subsequently analysed using SDS-PAGE as per [Section 3.3.6](#) to assess their protein content.

3.3.7.2 Nickel chromatography

Nickel affinity chromatography is a widely used method for purifying recombinant proteins, especially those that, like W-NP, have been incorporated into a plasmid in such a way that the synthesised protein contains a histidine (His) tag. To purify W-NP, a commercially available Nickel Chromatography Column (HisTrap HP Cat: #17524802; Cytiva, Marlborough, MA, USA) and the AKTA Pure Prime plus Chromatography system (Cat: #11001313; GE Healthcare, Chicago, IL, USA) were used. This chromatography resin and technique employs the affinity of histidine residues for Nickel ions Ni^{2+} that are coupled to nitrilotriacetic acid (NTA) agarose beads. Histidine residues bind to the column, and imidazole is used to compete for binding with the resin.

To do this, we looked at supplier-recommended protocols. A standard protocol was generated and optimised accordingly. The standard protocol followed started with the equilibration of the column using 10 column volumes of the various Purification Buffer A (20mM Tris-HCl; 1 M NaCl; pH 8.00 with either 10, 20, or 45 mM imidazole). The filtered supernatant prepared as per [Section 3.3.5.2](#) was loaded onto the column using a 50 ml capacity Superloop™ (Cat:18111382; Cytiva, Malborough, MA, USA) at 0.7 ml/min, and unbound proteins were washed from the column using 2 column volumes of the same PBA.

Non-specifically bound proteins were washed from the column using either a 25 ml or 50 ml gradient between Purification Buffer A and Purification Buffer B (20 mM Tris-HCl; 1 M NaCl; 150 mM imidazole, pH 8.00), totalling 10 column volumes at 2 ml/min. Finally, bound His-tag containing protein was eluted using a gradient between Purification Buffer B and Purification Buffer C (20 mM Tris-HCl; 1 M NaCl; pH 8.00 with either 400, 500, or 600 mM imidazole) over a 50 ml gradient at a 2 ml/min flow rate. The entire process was monitored spectrophotometrically at 280 nm whilst 5 ml and 2.5 ml fractions were collected during the wash and elution gradients, respectively. Fractions were prepared for SDS-PAGE protein content analysis as per [Section 3.3.6](#) by preparing a 1:1 sample to LSB ratio, and fractions containing high concentrations of pure protein of interest were identified.

3.3.8 Post-Purification Protein Processing

3.3.8.1 Protein concentration

Fractions collected via Nickel affinity chromatography identified through SDS-PAGE containing pure W-NP were pooled and the total volume was reduced using a stirred ultrafiltration cell (Amicon® Millipore Sigma, St. Louis, MO, USA, model 8050, cat. #5122) equipped with a 10 kDa molecular weight cut-off (MWCO) membrane (Thermo Fisher Scientific, Waltham, MA, USA, PLGC Disc 10K, 44.5 mm, PLGC04310) that selectively retains the protein, allowing the passage of lower molecular weight buffer. The ultrafiltration procedure was performed at room temperature, with the ultrafiltration cell wrapped in an ice pack to minimise heat build-up and prevent potential protein denaturation. The device was placed on a magnetic stirrer to maintain a constant 300 RPM, which will enhance the filtration efficiency and reduce membrane fouling. The

device was subjected to a nitrogen gas pressure of 0.5 mPa (or 80 psi) until the sample volume was reduced to 5 ml from the start volume.

3.3.8.2 Buffer Exchange

Purified and concentrated W-NP was prepared for quantification, storage, and downstream processing through a buffer exchange from PBC to a Storage Buffer (20 mM Tris-HCl; 1 mM ethylene diaminetetra-acetic acid (EDTA); 1 mM DTT; 10% (v/v) glycerol; pH = 8.0).

The buffer exchange was performed through size exclusion chromatography using an AKTA Pure system (Cat. 11001313; GE Healthcare, Chicago, IL, USA) and its associated software ([Section 3.2](#)), equipped with a HiTrap Desalting column (Cat. 29048684; Cytiva, Marlborough, MA, USA). This chromatography technique allows the separation of molecules based on their size, allowing small molecules (like salt) to enter the pores of the resin, slowing them down, whilst larger molecules (like proteins) pass through more quickly.

The column was first equilibrated with 10x column volumes of the same storage buffer, and the concentrated protein was applied to the column using a 50 ml capacity Superloop™ (Cat: 18111382; Cytiva, Marlborough, MA, USA) at a flow rate of 2 ml/min and subsequently eluted using the same buffer and flow rate. The entire process was monitored spectrophotometrically at 280 nm and 0.5 ml fractions were collected. The conductivity was monitored throughout to ensure buffer exchange efficiency. Fractions identified through UV 280 nm were pooled for quantification and storage.

3.3.8.3 Protein Quantification

3.3.8.3.1 Bradford Assay

The Bradford Assay (Kielkopf et al., 2020) is a colorimetric protein quantification method based on the binding of Coomassie Brilliant Blue dye to proteins, resulting in a measurable shift in absorbance proportional to protein concentration, was used to quantify the concentration of W-NP protein isolated. A stock concentration of 2 mg/ml BSA was prepared using Storage Buffer (20 mM Tris-HCl; 1 mM ethylene

diaminetetra-acetic acid (EDTA); 1 mM DTT; 10% (v/v) glycerol; pH = 8.0) and filtered using a 0.22 µm syringe filter. Serial dilutions (0.1, 0.2, 0.4, 0.6, 0.8, and 2.0 mg/ml) were prepared from the stock solution using filtered Storage Buffer. The BSA standards as well as isolated protein were prepared in duplicate in a clear micro polystyrene flat-bottom 96-well plate (Greiner Bio-One; Kremsmünster, Austria, Cat. # P1PLA024C-000096). Each sample consisted of 100 µl Bradford reagent (TaKaRa, Kusatsu, Shiga, Japan, Cat. T9310A-1), 98 µl dH₂O, and 2 µl protein. After 5 minutes of incubation (in the dark, at room temperature with shaking at 100 rpm), the plate was measured at 595 nm using a PerkinElmer EnSight Multimode plate reader (Waltham, MA, USA) and its associated software (Kaleido version 3.0.3067.117; Acapella version 5.0.1.123817). The absorbance readings were exported to CSV file format using the instrument associated software and imported into Microsoft Excel (version 2504; Build 18730.20142) where a standard curve and linear regression line equation was generated by plotting the average absorbance readings, corrected to the buffer only, over the corresponding protein concentrations. The concentration of W-NP was determined by supplementing the absorbance readings of the corresponding samples into the linear regression line equation as determined by Excel.

3.3.8.3.2 Spectroscopic Determination

The protein concentration determined by the Bradford assay was validated by spectroscopic measurement at 280 nm using a Chirascan plus V100 (Applied Photophysics, England) and its associated software ([Section 3.2](#)). Protein samples were prepared in Storage Buffer (20 mM Tris-HCl; 1 mM EDTA; 1 mM DTT; 10% (v/v) glycerol; pH = 8.0) and filtered using a 0.22 µm PVDF membrane to remove any particulates before measurement. Measurements were performed in a 1 cm quartz cuvette at 25 °C. Measurements were performed in triplicate to ensure reproducibility. Background measurements were performed using storage buffer to account for any background contributions. The data generated were exported to CSV format and analysed using MS Excel ([Section 3.2](#)). The data, along with the extinction coefficient, were applied to the Beer-Lambert Law using the following equation:

$$A = \epsilon cl$$

Equation 1

Where A is the absorbance at 280 nm determined by the Chirascan, ϵ is the computed molar extinction coefficient of the protein at 280 nm by Expasy ProtParam (Accessed: 21/05/2025) (Gasteiger et al., 2005), identified as 43 890 daltons, and l is the pathlength of the cuvette used to record the sample absorbance in cm (1 cm). This extinction coefficient was confirmed by the Edelhoch method (Pace et al., 1995):

$$\epsilon_{280} M^{-1} \cdot cm^{-1} = (5500 \times \Sigma W) + (1490 \times \Sigma Y) + (125 \times \Sigma C)$$

Equation 2

3.3.8.4 Protein Storage

To preserve protein integrity, purified, concentrated, and buffer-exchanged protein samples were frozen for long-term storage using a rate-controlled freezing method. Protein aliquots were transferred into sterile 1.5 ml tubes and sealed securely. These vials were then placed into a Mr. Frosty™ Cryo 1 °C Freezing Container (ThermoFisher Scientific, Waltham, MA, USA, Nalgene®, Cat. No. S100-0050 and 5100-0036) that was prefilled with 100% isopropanol according to the manufacturer's instructions. The container was transferred to a –80 °C freezing to achieve a uniform cooling rate of approximately –1 °C per minute. After a minimum of 12 – 16 hours (overnight), the samples were transferred to long-term storage at –80 °C until needed.

3.3.9 Mass Spectrometry

The purified W-NP protein was analysed using Mass Spectrometry (MS) to validate its identity and molecular weight. Mass spectrometry is a technique that measures the mass-to-charge ratio of ions, allowing identification and characterisation of molecules based on their mass, structure, and composition. This technique was performed by the Council for Scientific and Industrial Research (CSIR) (Pretoria, South Africa) using their instrument and associated software.

Sample preparation:

Protein purified in duplicate was applied to SDS-PAGE as per [Section 3.3.6](#) and individual bands were cut out for MS analysis. The SDS-PAGE protein gel bands were digested as per Shevchenko et al. (2006). Each gel piece was cut into small pieces

and washed in 50 mM ammonium bicarbonate (NH_4HCO_3)/50% methanol (MeOH) for 20 minutes, then centrifuged to separate supernatant and sediment. This step was repeated and then followed by a 75% acetonitrile (ACN) wash. The proteins were reduced in 10 mM DTT with 25 mM NH_4HCO_3 for 1 hour at 60 °C. Samples were spun down and the supernatant was removed. Samples were cooled to room temperature and then incubated for 10 minutes in 100% ACN. The resulting sediment was incubated in the dark for 20 minutes at room temperature in 55 mM iodoacetamide (IAA) with 25 mM NH_4HCO_3 . The remaining sediment was dehydrated for 5 minutes using 25 mM NH_4HCO_3 with 50% ACN, then centrifuged, and the supernatant was removed. The gel pieces were dried to completeness. Proteins were digested using trypsin overnight at 37 °C (Shevchenko et al., 2006). Digestion was quenched using 0.1% formic acid, and the samples were dried under vacuum. Dried samples were resuspended in 2% ACN with 0.2% formic acid, and the resulting sample was used for mass spectrometry.

Liquid Chromatography-Mass Spectrometry (LC-MS) Acquisition:

The fragmented proteins were analysed using an Evosep One LC system (Dresden, Germany) coupled to a Bruker timsTOF HT (Billerica, MA, USA) Mass Spectrometer, operated using the Bruker Compass HyStar (v6.3.1.8, Bruker, Billerica, MA, USA) acquisition platform with integrated TIMS/PASEF control modules during DDA-PASEF data acquisition. The peptides were injected and separated on a PepSep C18 column (150 μm $\times$ 8 cm, 1.5 μm particle size) and eluted using an ACN gradient over 11.5 minutes. The mass spectrometer was operated using the DDA-PASEF mode, the acquisition spectrum was set to 100-1700 m/z for the precursor MS scan with a 4 PASEF MS/MS ramps that had an intensity threshold of 2500 for scheduling MS/MS as well as a median accumulation time of 100 ms per scan, totalling cycle time of 0.53 seconds.

The resulting raw data files from the spectrum were searched using MaxQuant software using the Andromeda search algorithm and compared to a database containing sequences from SARS-CoV-2 and *E. coli* proteins. The digestion enzyme trypsin was selected, and cysteine alkylation (iodoacetamide) was chosen as a fixed modification. Methionine oxidation and N-terminal acetylation search parameters were selected as variable modifications.

3.3.10 Far-UV Circular Dichroism

The secondary structure of the isolated W-NP was investigated using Far-UV Circular Dichroism (CD), a spectroscopic technique that measures the differential absorption of left- and right-circularly polarized light between 190 to 250 nm. Protein stocks prepared as per [Sections 3.3.5.2 – 3.3.8.4](#) were diluted to a final concentration of 5 μM using Storage Buffer (20 mM Tris-HCl; 1 mM ethylene diaminetetra-acetic acid (EDTA); 1 mM DTT; 10% (v/v) glycerol; pH = 8.0). Stock solutions of compounds (palmatine and berberine) to 0.01 M concentration were prepared in storage buffer supplemented with 10% methanol. Samples subjected to CD were prepared containing 5 μM protein with or without 5 – 400 μM (5, 10, 20, 40, 50, 75, 100, 200, and 400 μM) compound or 7 M urea. Buffer only, compound only (0.001 M), and urea (7 M) only were analysed as blanks/reference measurements. All samples were filtered using a 0.45 μm syringe filter and degassed at 4 $^{\circ}\text{C}$ and 635 mmHg vacuum using a 120 RPM stirring rate for 30 minutes using a Degassing Station (TA Instruments, New Castle, DE, USA). Far-UV CD spectra were recorded using a Chirascan instrument (Applied Photophysics, Leatherhead, England) equilibrated with a Peltier temperature control unit. The instrument was prepared for acquisition using pure nitrogen. Data were collected in the far-UV range (180–280 nm, focusing on the aromatic side-chain absorption region) using a quartz cuvette with a 0.1 cm path length. Three acquisitions were recorded with a 1 nm step size, 2 nm bandwidth, 1 mm slit width, and a 0.5-second response time. The Peltier temperature controller was used to maintain the temperature at 25 $^{\circ}\text{C}$ or ramped from 20 $^{\circ}\text{C}$ to 94 $^{\circ}\text{C}$ at a step of 1 $^{\circ}\text{C}$ at a 1 $^{\circ}\text{C}/\text{min}$ rate, with spectra recorded every 5 $^{\circ}\text{C}$ with a tolerance of 0.20 $^{\circ}\text{C}$ as instructed by a T2 SHU temperature probe placed inside the cell. Instrument supplied software ([Section 3.2](#)) was used to convert raw ellipticity to molar ellipticity ($[\theta]$, $\text{deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$) using the formula:

$$[\theta] = \frac{\text{observed ellipticity (mdeg)}}{10 \cdot c \cdot l}$$

Equation 3

where l is the cuvette path length in cm, and c is the protein concentration in mol/L. BeStSel Deconvolution software (Micsonai et al., 2015) was used to estimate the secondary structural content via the single-spectrum analysis. Ellipticity changes at

characteristic wavelengths (e.g., 208 nm and 222 nm for α -helices) were analyzed to assess secondary structural alterations upon compound binding as well as thermal denaturation. All experiments were conducted in duplicate using separately purified batches of protein to ensure reproducibility.

3.3.11 Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy (ATR-FTIR)

To investigate protein conformation and interactions with berberine and palmatine, W-NP previously stored at $-80\text{ }^{\circ}\text{C}$ in Storage buffer was thawed on ice for sample preparation. Samples containing 5 μM W-NP with or without 7 M urea, or 200 μM berberine or palmatine were prepared with Storage Buffer (20 mM Tris-HCl; 1 mM ethylene diaminetetra-acetic acid (EDTA); 1 mM DTT; 10% (v/v) glycerol; pH = 8.0) and incubated at room temperature for 30 minutes.

ATR-FTIR spectra were acquired at room temperature under atmospheric pressure using a Nicolet™ iS™ 5 FTIR Spectrometer (ThermoFisher Scientific, Waltham, MA, USA) equipped with an iD7 ATR accessory, a ZnSe diamond crystal, and a deuterated triglycine sulfate (DTSG) KBr detector. The ATR crystal was cleaned with isopropanol alcohol between each measurement, after which 10 μl of the sample was applied directly onto the crystal, ensuring full contact with the surface. Instrument associated software, Omnic Paradigm (2019-2024, package version 1.9 Hot Fix 2 (Build 1), Thermofisher Scientific) was used to acquire a total of 50 scans at a resolution of 4 cm^{-1} between 525 – 4000 cm^{-1} for each sample, with air as background and absorbance as the final format. The software was used to subtract background contribution by buffer and compound only [difference in spectra (Δ Absorbance)], and the resulting spectral data were exported to TSV file format before being smoothed and visualised using OriginLab (Origin Pro 2025 64-bit SRI Learning edition 10.2.0.196, OriginLab corporation). Reproducibility was ensured between two separately purified batches of W-NP.

3.3.12 Intrinsic Tryptophan Fluorescence

Fluorescence spectroscopy leverages the intrinsic fluorescence of aromatic amino acid residues (like tryptophan) that provide insights into protein structure, dynamics, and interactions. W-NP prepared as per [Sections 3.3.5.2 – 3.3.8.4](#) was diluted to 5 μM final concentration using Storage Buffer (20 mM Tris-HCl; 1 mM ethylene diaminetetra-acetic acid (EDTA); 1 mM DTT; 10% (v/v) glycerol; pH = 8.0). In addition, 0.01 M stocks of palmatine, berberine, and 8-anilinonaphthalene-1-sulfonic acid (ANS) were prepared using 10% methanol supplemented with storage buffer. Final concentration of 5 μM protein with either 5 – 400 μM (5, 10, 20, 40, 50, 75, 100, 200, and 400 μM) compound; 5 – 250 μM (5, 10, 20, 50, 100, 250 μM) histidine; 7 M urea; 250 μM ANS were used. Buffer only, compound only (0.001 M), histidine only (0.001 M), and urea (7 M) only were analysed as blanks/reference measurements. All samples were filtered using a 0.45 μm syringe filter and degassed at 635 mmHg for 30 minutes at 4 $^{\circ}\text{C}$ using a stirring rate of 120 RPM and a Degassing Station (TA Instruments, New Castle, DE, USA).

Fluorescence measurements were performed using a Chirascan instrument (Applied Photophysics, Leatherhead, England) and its associated software ([Section 3.2](#)), which is equipped with a xenon lamp flushed and maintained with pure nitrogen gas. The Peltier temperature controller was used to regulate the temperature at 10, 15, 20, 25, 30, 35, or 40 $^{\circ}\text{C}$ or enabled to perform a stepped temperature ramp from 20 to 60 $^{\circ}\text{C}$ in 2 $^{\circ}\text{C}$ increments at a rate of 1 $^{\circ}\text{C}/\text{min}$ with a 30-second settling time. All samples were excited at 295 nm, with the ANS samples having the exception of 380 nm excitation. All emission spectra were recorded in triplicate from 250 to 550 nm with a 1 nm bandwidth within a 10 mm quartz fluorescence cuvette. Triplicates were assessed for reproducibility and averaged. Instrument control and results analysis were performed using the software listed in [Section 3.2](#), during which triplicates were assessed for reproducibility and analysis, corrections were performed, and the data were ultimately exported and visualised using MS Excel, also listed there.

Stern-Volmer analysis was performed using raw data using the following formula:

$$\frac{I_0}{I} = 1 + K_{SV} \times [Q]$$

Equation 4

Where:

I_0 is the fluorescence at the reference (protein only at 25 °C)

I is the fluorescence intensity at elevated temperature / various concentrations of quencher

KSV displayed is the Stern-Volmer quenching constant

$[Q]$ is the concentration of the quencher or the temperature

3.3.13 Native-PAGE

Native polyacrylamide gel electrophoresis (PAGE) is a method that separates proteins based on their intrinsic charge, size, and conformation, which is performed under non-reducing conditions, enabling researchers to investigate the native and functional states of proteins. The W-NP has a theoretical positive charge, requiring the use of Blue Native PAGE (BN-PAGE) for the investigation of the quaternary structure and the effect imposed by a compound. BN-PAGE often employs the use of Coomassie Brilliant Blue R250, a dye that binds non-covalently to proteins, imparting a negative charge on the protein (masks the intrinsic charge), facilitating protein migration during electrophoresis without altering the native conformation (Camacho-Carvajal et al., 2004).

Protein stocks prepared as per [Sections 3.3.5.2 – 3.3.8.4](#) were diluted to a final concentration of 5 μ M using storage buffer. Stock solutions of compounds (palmatine and berberine) to 0.01 M concentration were prepared in storage buffer supplemented with 30% methanol. Samples containing 5 μ M protein with or without 200 μ M compound and/or 7 M urea were prepared. The samples were mixed to a 1:1 ratio with BN-PAGE sample buffer (40% (v/v) glycerol) and incubated at room temperature for 30 minutes. The samples were loaded onto a BN-PAGE Gel consisting of a 5% stacking gel [BisAcrylamide (5%), Tris-HCl (0.125 M, pH = 6.8), APS (0.025%), TEMED (0.0004%)] and a 12% resolving gel [BisAcrylamide (12%), Tris-HCl (0.0375 M, pH = 8.8), APS (0.05%), TEMED (0.0005%)] alongside the NativeMark Protein Standard (Invitrogen, Thermofisher Scientific, Waltham, MA, Cat: 57030) containing Soybean Trypsin Inhibitor, Bovine Serum Albumin, Lactate Dehydrogenase, B-phycoerythrin, Apoferritin bands 1 and 2, as well as IgM Pentamer and Hexamer

proteins that range in sizes 20 to 1200 kDa. Together, they were electrophoresed on ice using a Bio-Rad Mini-Protean Tetra system (Hercules, CA, USA) at 150 Volts. The BN-PAGE was initially run using a running buffer (62 mM Tris-HCl; 480 mM glycine) supplemented with 1 mg/ml Coomassie for 20 minutes, and then later replaced with non-supplemented running buffer. The BN-PAGE was destained (5% ethanol, 7% acetic acid) overnight. The resulting gels were visualized using a Bio-Rad Gel Doc TM XR+ Molecular Imager along with its software described in [Section 3.2](#). The images were created using a custom application that employs UV transmission illumination.

3.3.14 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) experiments were conducted to assess the binding interactions between the compounds and the protein.

A c-window as per equation below was used to determine the concentration of ligand to be used for this experiment:

$$c = nK_a [\text{titrand}] \text{ or } c = n [\text{titrand}] / K_a$$

[Equation 5](#)

The ITC measurements were performed using an Affinity ITC (TA Instruments, New Castle, USA) and its related ITC Run software (TA Instruments, New Castle, USA, version 3.4.11) at a controlled temperature of 25 °C.

The protein of interest was degassed at 25 °C at 120 RPM stirring rate at 635 mmHg vacuum for 30 minutes. The 20 mM compound was titrated incrementally into the 20 μM protein solution using a stirring rate of 100 RPM to ensure thorough mixing throughout the experiment. A total of 25 injections were performed, with an injection interval set to 250 seconds. The initial injection volume was 3 μL, followed by subsequent injections of 5 μL for the remainder of the titration series.

For equilibration, default settings were utilized, with a start delay of 300 to 600 seconds. The timeout for each experiment was set to 1800 seconds to accommodate potential delays in reaction time. Baseline measurements were collected prior to the first injection, with an initial baseline period of 250 seconds to align with the injection

interval. After the final injection, the collection of data was continued for an additional 0 seconds, ensuring that any residual heat changes were accurately captured.

Data analysis was performed using NanoAnalyze (TA Instruments, New Castle, USA, version 3.7.5), allowing for the extraction of thermodynamic parameters such as binding constants, enthalpy, and entropy associated with the binding interactions.

Matching buffer to buffer, buffer to protein, and ligand to buffer titrations was used as baseline for calibration and blank controls, ensuring no signal interference. The binding affinity, binding enthalpy, and stoichiometry were determined. This technique was performed in duplicate to ensure reproducibility.

3.3.15 Stability Studies

The stability of the protein and protein-compound complexes was assessed using Far-UV-CD and intrinsic tryptophan fluorescence as per [Sections 3.3.10 and 3.3.12](#). Here, the protein unfolding and refolding were monitored using temperature ramping, recording measurements every 2 °C between 25 °C and 60 °C, with a 0.20 °C tolerance, a 30 second settling time measured for by an appropriate temperature probe, and at the end, returning to 25 °C. These measurements were taken in duplicate to ensure reproducibility. From this data, the resulting Mean Residue Ellipticity at 228 nm and 235 nm for CD measurements, and the absorbance at 335 nm for fluorescence, were normalised and plotted for comparison using MS Excel.

3.3.16 Surface Plasmon Resonance

Surface Plasmon Resonance (SPR) was employed to assess the binding interactions between W-NP and the alkaloids. This technique enables label-free, real-time evaluation of binding kinetics based on refractive index changes brought on upon binding.

The SPR assays were conducted using a Sierra SPR-24 Pro system (Bruker, Billerica, MA, USA) alongside the SPR High-Capacity Amine Sensor (Bruker Daltonics, Billerica, MA, USA, Cat # 1862614). Isolated W-NP was buffer exchanged as per

Section 3.3.8.2 to a 1x HBS Buffer (25 mM HEPES; 150 mM NaCl; 500 mM DTT; 10% (v/v) glycerol; pH 7.4), filtered using a 0.22 µm syringe filter, and diluted to a final concentration of 50 µg/ml W-NP using acetate pH 5.5 and degassed at 4 °C and 635 mmHg vacuum using a 120 RPM stirring rate for 30 minutes using a Degassing Station (TA Instruments, New Castle, DE, USA). BSA was similarly prepared only, diluted to a final concentration of 25 µg/ml using acetate pH 5.0.

W-NP was immobilized via EDC/NHS-mediated amine coupling to a density of approximately 10 000 RU on the active flow cells. Reference cells included BSA controls and activated-blocked controls.

The assays were performed at 25 °C and started with a common injection start-up phase during which PBS-T (137 mM sodium chloride; 2.7 mM Potassium Chloride; 11.9 mM Phosphate Buffer; 0.05% Tween20) was delivered 20 times at a flow rate of 25 µL/min with a 60 s contact and dissociation time as well as 10 s waiting time. The temperature data was cleared.

The kinetic assays were performed using a running buffer composed of PBS-T-DMSO [137 mM sodium chloride; 2.7 mM Potassium Chloride; 11.9 mM Phosphate Buffer; 0.05% Tween20; 5% Dimethylsulfoxide (DMSO)] delivered at a flow rate of 25 µL/min with a 30 s contact time and 0 s dissociation time after 20 startup buffer cycles.

The analytes, berberine and palmatine, were prepared to a stock concentration of 50 mM in 100% DMSO. Thereafter, the alkaloids were prepared in PBS-T-DMSO (137 mM sodium chloride; 2.7 mM Potassium Chloride; 11.9 mM Phosphate Buffer; 0.05% Tween20; 5% DMSO) as 10 serial concentrations spanning 2.5 mM to 4.88 µM.

Injections followed a multi-cycle kinetics (MCK) format, during which a single concentration of analyte was injected per cycle. Each concentration was injected for ~120 s association and ~180 s dissociation, followed by a regeneration cycle using 10 mM glycine-HCl (pH) for 30 seconds.

The data was collected using the Sierra 32 Control Panel (version 3.6.28.5) software as real-time sensogram in response units (RU) and processed using SPR Analyzer 4

(version 3.4.5) with alignment on, reference blank subtraction to the closest value, a de-spike median filter, as well as solvent correction enabled. Kinetic parameters were not derived from fitting due to results. All experiments were performed in duplicate to ensure reproducibility.

3.3.17 Computational Docking

The interaction between the nucleocapsid protein and ligands (berberine and palmatine) was assessed using molecular docking, a computational method that relies on mathematical algorithms to predict potential binding modes and energies associated with protein-ligand complexes.

A consensus sequence of the SARS-CoV-2 nucleocapsid (N) protein was utilized in this study, differing from the original Wuhan sequence. Importantly, the mutations present in the consensus sequence are located outside the N-terminal domain (NTD) and C-terminal domain (CTD), which are critical for RNA binding and dimerization, respectively. To date, no experimentally determined full-length 3D structure of the SARS-CoV-2 N protein exists, primarily due to its intrinsically disordered regions that hinder crystallization and structural resolution (Maiti et al., 2024; Schiavina et al., 2022). However, high-resolution structures of the individual NTD and CTD have been resolved using techniques such as X-ray crystallography and solid-state NMR. Comparative analysis of the AlphaFold predicted structures of the consensus sequence to the experimentally determined structure indicates that the structures closely align, such that the mutations present should not affect the secondary or tertiary structures of these domains. Consequently, for our docking studies, we employed the experimentally resolved structures of the NTD and CTD domains derived from the Wuhan sequence, as these domains are conserved and structurally representative of the consensus sequence used in our research. Unfortunately, this limits the study to functional sites only. Nonetheless, simulations were conducted using palmatine (PubChem CID: 19009) and berberine (PubChem CID: 2353), to the distinct NTD (PDB ID: 7ACT) and CTD (PDB ID: 7XWX) domains of the SARS-CoV-2 nucleocapsid protein.

Protein preparation (Creighton, 1993; Shenkin et al., 1987) and ligand preparation (Murray-Rust & Glusker, 1984) was performed using Discovery Studio (BIOVIA, Dassault Systèmes, San Diego, CA, USA). Hydrogen atoms were added (Göribitz, 1989) to the protein structures prior to preparation. Protein structures were cleaned and prepared using default parameters in the 'Prepare Protein' protocol. Pre-bound ligands and non-essential crystallographic water molecules were removed where applicable during receptor preparation using default Discovery Studio criteria. Ligands were prepared using the 'Prepare Ligands' protocol with default settings (Krammer et al., 2005). The CHARMM force field (Momany & Rone, 1992) was employed for both protein and ligand parameterisation.

Potential binding pockets on each domain were predicted using the "Define and Edit Receptor Cavities" tool in Discovery Studio. Identified cavities were manually reviewed for suitability based on structural exposure and depth. For the NTD, three potential binding sites were selected for docking, defined by the following XYZ coordinates (in Å) and radii (in Å):

NTD Site 1: X = 15.908, Y = 21.0121, Z = 29.4597, Radius = 7.2

NTD Site 2: X = 21.2023, Y = 38.767, Z = 21.7367, Radius = 5

NTD Site 3: X = 32.8974, Y = 15.6332, Z = 22.4132, Radius = 9.7

NTD Site 4: X = 22.226, Y = 43.0848, Z = 32.9299, Radius = 10.3

For the CTD, the following binding sites were selected:

CTD Site 1: X = -9.1704, Y = 99.8259, Z = -29.3606, Radius = 12.7

CTD Site 2: X = -16.0817, Y = 93.1236, Z = -26.5953, Radius = 11.4

CTD Site 3: X = -18.0768, Y = 114.279, Z = -26.8199, Radius = 9.2

Molecular docking was performed using the CDOCKER algorithm within Discovery Studio. The protocol uses a CHARMM-based molecular dynamics simulated-annealing approach to generate and refine ligand poses within a defined binding sphere. Each docking run generated five top random conformations per ligand–site combination, which were scored and ranked based on CDOCKER interaction energies (Wu, G. et al., 2003).

The Calculate Binding Energies protocol in the software was used to estimate the binding free energy and to assess conformational stability, employing *in situ* ligand minimization and Generalized Born with Molecular Volume (GBMV) implicit solvation model to approximate the effects of surrounding solvent molecules (Feig & Brooks, 2004; Tirado-Rives & Jorgensen, 2006). Poses were further evaluated by visual inspection to identify key molecular interactions such as hydrogen bonds, π - π stacking, hydrophobic contacts, and unfavourable steric clashes.

4. Results

4.1 In Silico Validation & Structural Prediction of W-NP

The Phylogenetic Assignment of Named Global Outbreak Lineages (PANGO lineage) system is a nomenclature framework used to classify and track the global evolution of the SARS-CoV-2 virus by assigning hierarchical names like B.1.1.7 to genetically distinct viral lineages based on their phylogenetic relationships and epidemiological significance (O'Toole et al., 2021; O'Toole et al., 2021; Zhu et al., 2021). Since the outbreak of the novel virus in Wuhan in December 2019, to the date of data collection in April 2024, a total of 4099 PANGO lineages had been identified. Only 2189 of these lineages had N protein sequences available for use in the generation of the consensus sequence, corresponding to 53.40% of all lineages known at the time. It is important to note that the prevalence of each lineage was not taken into account, nor were potential sampling biases occurring due to the incomplete dataset investigation, since the primary objective of this study was to identify the most frequently occurring mutations across SARS-CoV-2 lineages irrespective of the prevalence in the population. This approach allows for the generalization of the sequence, capturing a wide array of common mutations and conserved regions, ensuring that this sequence remains a pertinent tool for studying viral evolution, diagnostics, and drug development.

The consensus sequence generated in the study is shown in [Figure 4. 1.A](#). The sequence spans 416 amino acids and reveals 98.33% identity compared to the original 419 aa Wuhan sequence. This suggests that, over the 52 months, minimal deviation of the N protein has occurred, suggesting this protein is highly conserved in nature, a matter long been disputed (Gupta et al., 2020; Rahman et al., 2021; Zhang, B. et al., 2022). The consensus sequence includes amino acid substitutions at positions 203 (R to K), 204 (G to R), and 413 (S to R), along with ambiguous bases at positions 13 and 31–33. The ambiguous bases are explained by the methodology used to construct the consensus sequence in this study, where the most frequently occurring amino acid is incorporated, whilst some nucleotides may occur with similar frequencies, resulting in a high uncertainty in that position, demarked with an 'x' by the EMBOSS tool. The frequency and association of these mutations are well documented and explained; however, they exclude the frequently reported D63G mutation (Alsuwairi et al., 2023).

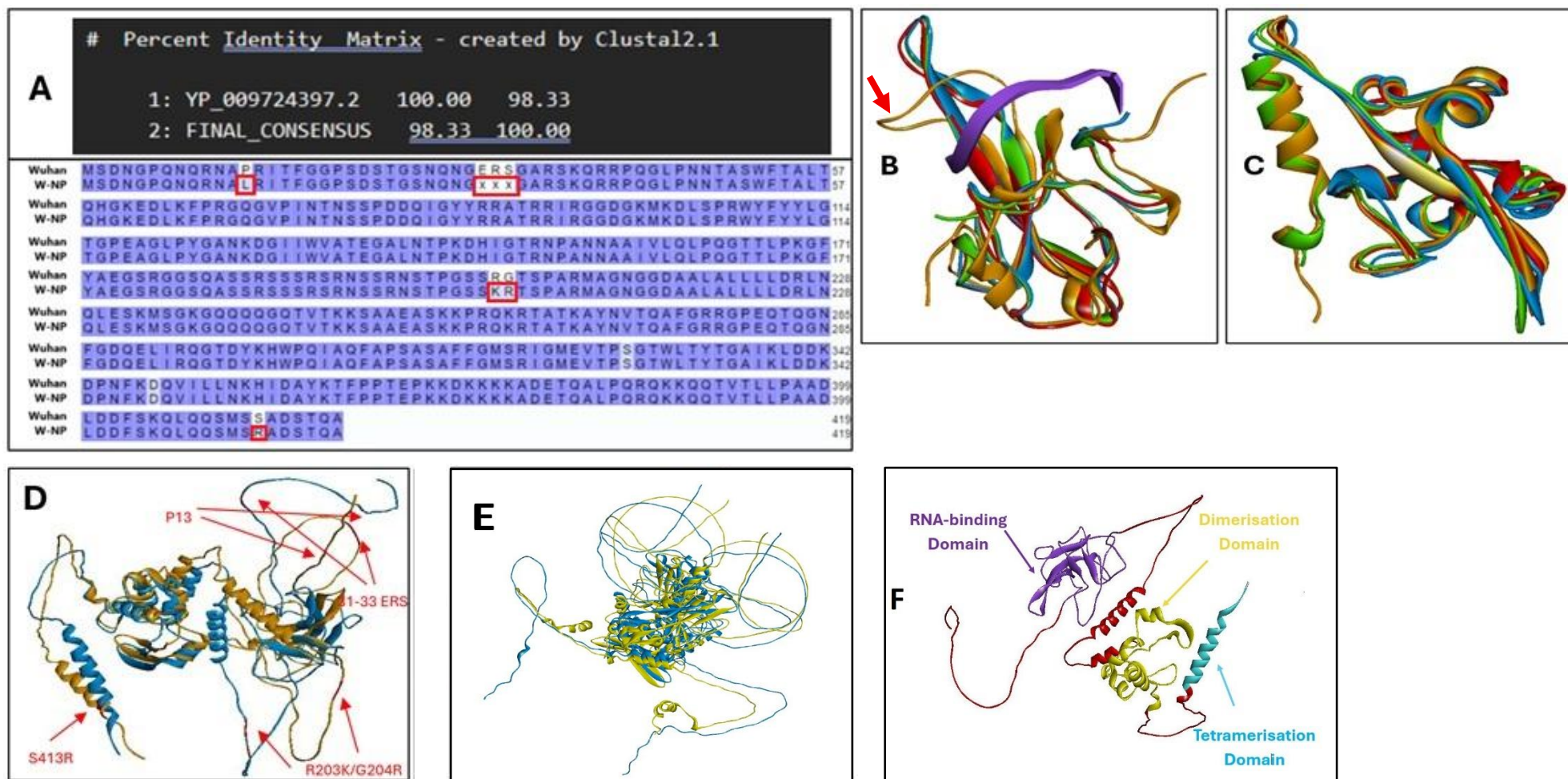


Figure 4. 1: Structural validation of W-NP via in silico prediction

A show the linear sequence alignment from UniProt, with matches in blue/purple, differences demarked in red, numerical positions indicated, and the percent identity matrix displayed at the top. Figures B-E shows the 3D structures of the Wuhan NP (green) and the W-NP (blue) predicted by AlphaFold Colab, as well as Swiss-Model Predicted W-NP (red) superimposed to the experimentally determined structures (orange). B shows the NTD (7ACT) with the nucleic acid in purple with the red arrow identifying the difference. C shows the CTD (7XWX). D shows the monomeric, and E, the dimeric forms. D highlights the location of the various mutations incorporated in red. F shows the 3D structure of the protein with the RNA-binding domain (purple), dimerization domain (yellow), tetramerization domain (cyan) and the backbone (red).

At position 13, proline (P) is substituted with another residue (x). This substitution was highly identified from the start of the pandemic, most notably P13L, and was reported to be present in 1.9% of global samples (Oulas et al., 2021; Troyano-Hernaez et al., 2021). This residue, located within the disordered N-arm region, is known to be part of a T-cell epitope that is responsible for cellular immune responses and recognition (Alsuwairi et al., 2023). This mutation is believed to disrupt recognition of CD8+ T-cell epitope, potentially enabling the virus to evade T-cell mediated immune responses (de Silva et al., 2021).

Interestingly, the P13L substitution has been associated with reduced viral transmission and death associated with infection (Oulas et al., 2021). This substitution reportedly creates a new protein-protein interaction interface (Nguyen et al., 2024) and may alter the RNA and M-protein binding affinities (Muradyan et al., 2024). The effect of the P13L mutation on protein stability, however, may be inconclusive, since some reports indicate that the mutation may enhance stability (Muradyan et al., 2024) whilst others report that it decreases stability (Oulas et al., 2021). With the frequency of natural P13L mutation occurrences, along with the reported effects on protein functionality, it was deemed appropriate to incorporate this change in the consensus to ensure accurate residue representation.

The EMBOSS tool similarly demarked positions E31, R32, and S33 in the N-arm region as ambiguous residues, however, they are frequently reported as complete deletions and were the defining mutations of the Omicron variants (Alsuwairi et al., 2023; Mak et al., 2022; Obermeyer et al., 2022; Oulas et al., 2021; Savellini et al., 2023). These deletions have been associated with a reduced risk of death, suggesting a potential link between these alterations and viral pathogenicity (Wu, S. et al., 2021).

Further investigation into the functional implications of these deletions is warranted to elucidate their precise role in modulating viral virulence and host response. It was therefore decided to incorporate these into the consensus sequence.

The R203K and G204R mutations that often occur together within the SR-rich linker region, have been found in the Alpha, Delta, and Omicron variants (Muradyan et al., 2024; Rahman et al., 2021; Savellini et al., 2023; Wu, H., Xing, Meng, Fu, Xue, Dong,

Tang et al., 2021; Wu, H., Xing, Meng, Fu, Xue, Dong, Xiao et al., 2021), and have been implicated in structural changes that affect viral pathogenesis and host interactions. These mutations are reported to reduce structural stability and flexibility, affecting monomer and dimer formation (Muradyan et al., 2024; Rahman et al., 2021) whilst enhancing RNA binding affinities and oligomerisation potential that contribute to viral replication, infectivity, and severity of infection (Guan et al., 2020; Mourier et al., 2022; Muradyan et al., 2024; Shuaib et al., 2023). This double mutation significantly impairs the virus's ability to form particles, suggesting that these changes interfere with the proper assembly of the viral structure (Nguyen et al., 2024). These mutations have been correlated with higher viral loads and reduced risk of ICU admission (Alsuwairi et al., 2023; Mourier et al., 2022), whilst no significant association has been found with mortality (Guan et al., 2020; Mourier et al., 2022; Muradyan et al., 2024; Shuaib et al., 2023). Recent studies suggest that this paradoxical effect may be due to destabilization of secondary structural elements within the disordered regions (Muradyan et al., 2024) or due to dysregulation of interferon response genes caused by hyperphosphorylation of adjacent serine residues, particularly S206 via GSK3A (Mourier et al., 2022). Furthermore, these mutations have been reported to affect rapid antigen test accuracy, potentially complicating diagnostics (Hagag et al., 2022). Collectively, the R203K/G204R mutations appear to play a dual role in SARS-CoV-2 infection, enhancing viral function while potentially modulating disease severity and immune evasion, therefore, being essential to be included in a consensus sequence.

Similarly, little is known about the S413R substitution, located within the helices of the tetramerization region (Ye et al., 2020), which is crucial for protein oligomerization and may influence viral assembly and replication, while also being part of the C-tail (Alsuwairi et al., 2023), underscoring its potential significance in viral protein function.

To date, no experimentally solved structure exists for the full-length protein, with resolved structures limited to individual domains. **Figure 4. 1.B** shows the structural superimposition of experimentally determined structures (orange) with AlphaFold2-predicted models (green and blue), while **Figure 4. 1.C** illustrate the corresponding comparison with Swiss-Model predictions (red). **Figure 4. 1.F** provides contextual mapping of the functional domains within the full 3D structure.

This comparison reveals a key difference between the experimentally determined and predicted structures of the NTD. The NTD consists of a five-stranded U-shaped right-handed antiparallel β -sheet core with a $\beta 4$ - $\beta 2$ - $\beta 3$ - $\beta 1$ - $\beta 5$ topology. From this core, $\beta 2$ - $\beta 3$ forms a long β -hairpin loop that protrudes from the β -core, consisting of basic amino acids, creating the basic finger we refer to as the palm. Electrostatic potential reveals that the putative RNA-binding site is located between this basic finger and the palm. This extending loop, however, is more structured within the AlphaFold predicted structures for both the Wuhan and W-NP sequences, as well as the W-NP structure predicted by Swiss-Model. This discrepancy could originate within the experimentally determined structure if the method could not accurately resolve the full range of conformations present (Acharya & Lloyd, 2005; Alsuwairi et al., 2023; Clore et al., 1993). However, this discrepancy could similarly be attributed to the methodology used during predictive modelling, where the training strategies rely on proteins' structure, often resulting in the over-prediction of secondary structures in IDPs (Deng et al., 2012; Ruff & Pappu, 2021), demonstrated by the reduction in pLDDT confidence scores between the 42nd and 55th residues within the various NTDs ([Section 10: Appendix 10.3](#)). However, the $\beta 2$ - $\beta 3$, along with the $\Delta 2$ - $\beta 5$ loop, is known for its “tweezer-like” interaction that performs the Watson-Crick base-pairing destabilisation, a mechanism crucial for viral genome condensation during genome packaging, transcription regulation, and phase separation (Caruso et al., 2021; Cubuk, Alston et al., 2024; Cubuk, Incicco et al., 2024; Sarkar et al., 2022), therefore, it is more likely that this discrepancy is seen due to local conformational changes that are associated with the binding of the 10mer ssRNA ligand in this case. Therefore, this strategy acknowledges that no single method can fully characterise IDPs, emphasising the importance of combining multiple approaches to understand their structure-function relationship comprehensively.

Furthermore, little to no deviation was seen within the CTD, suggesting little structural difference occurred between the Wuhan and consensus sequence, whilst highlighting the confidence of AlphaFold and Homology models (Varadi et al., 2021).

Differences emerge in the tertiary structural arrangement of the monomeric structures predicted by AlphaFold (in [Figure 4.1.D & E](#)). These differences originate at the helical region within the linker domain (in [Figure 4.1.F](#)) and are also evident in the predicted

dimeric structures. Notably, the AlphaFold-predicted dimer for the W-NP sequence exhibits a loss of secondary structural helices within the tetramerisation region, suggesting potential structural impacts on oligomerization and higher-order assembly.

In conclusion, the mutation and deletions identified through the generation of the consensus sequence of the SARS-CoV-2 nucleocapsid protein illustrate how a single genetic change can dramatically alter viral behaviour, and that the cumulative effect of these changes can affect everything from viral transmission to immune evasion, highlighting the need for a consensus sequence for laboratory use, enabling researchers to gain insights into their consequences, to better anticipate viral evolution, and to refine strategies for diagnostics, treatments, and vaccines. Ultimately, a consensus sequence provides a more comprehensive understanding of the virus, equipping us to respond more effectively to the ongoing challenges posed by SARS-CoV-2.

4.2 Expression, Purification, and Biochemical Characterisation

4.2.1 Optimisation of Protein Expression in *E.coli*

A screening study investigated the expression of recombinant N protein using (Li, G. et al., 2021; Robichon et al., 2011) and *E. coli* BL21(DE3)-T7 Express cells, superior alternatives of the *E. coli* BL21(DE3) cell line that has previously been reported to produce mostly insoluble N protein (Li, G. et al., 2021; Robichon et al., 2011).

Each of these cell lines similarly utilise the T7 system, where the target gene is cloned into an expression vector system downstream of the T7 promoter. The construct is then introduced into a T7 expression host that carries a chromosomal copy of the phage T7 RNA polymerase gene, whose expression can be induced, producing great amounts of target protein, and in addition, can be fine-tuned using the lac repressor (Studier & Moffatt, 1986). The BL21(DE3)-T7 Express cell lines are optimised to enhance the expression of toxic or insoluble proteins through the deletion of two major proteases, Lon and OmpT, that would normally degrade recombinant target protein. Furthermore, these cells allow tighter repression of T7 RNA polymerase via enhanced lacI control, reducing leaky expression.

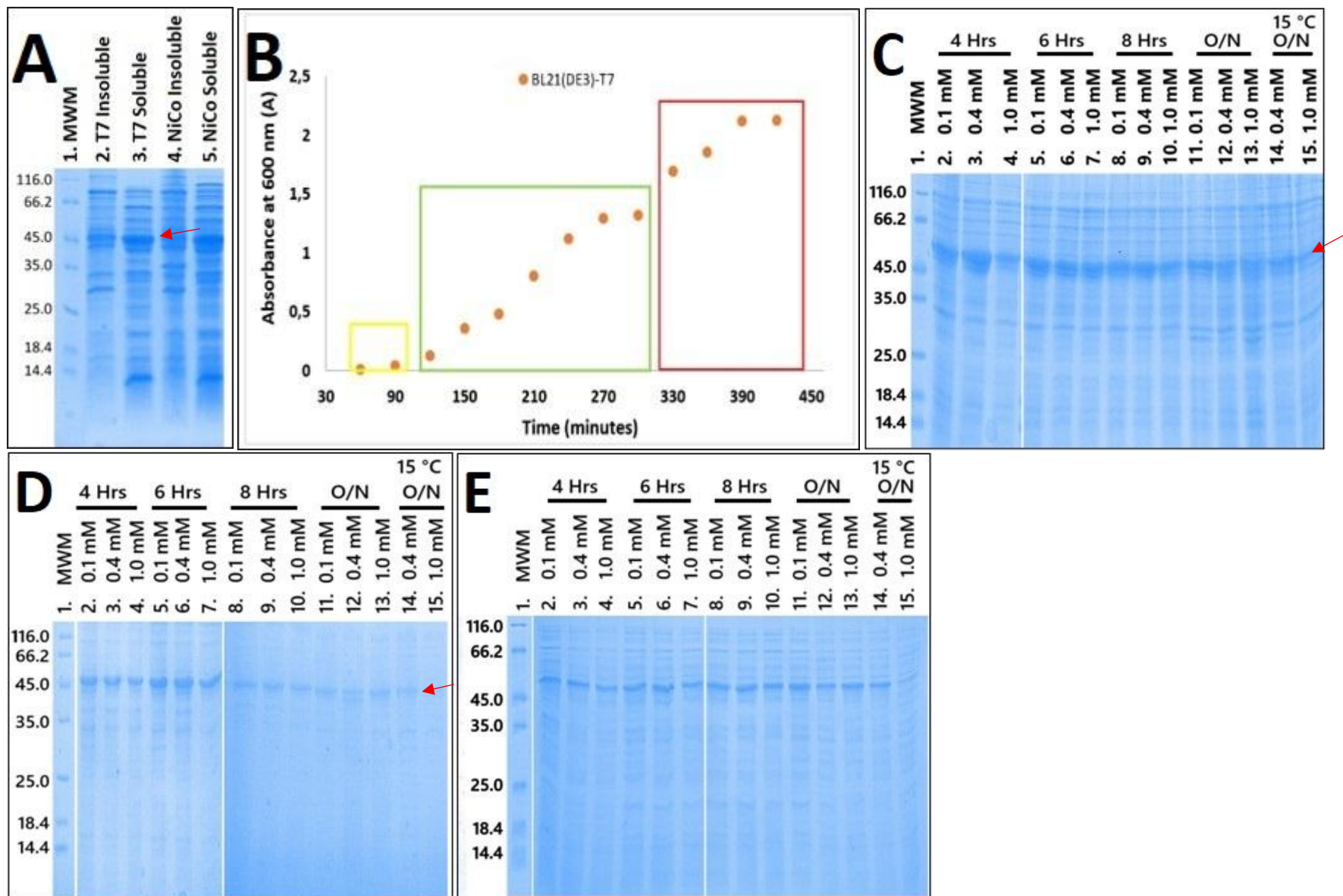


Figure 4. 2: Expression optimisation of W-NP in *E. coli*

A shows the SDS-PAGE analysis comparing soluble and insoluble protein fractions from *E. coli* BL21(DE3)-T7 and NiCo21(DE3) cultures induced with 0.1 mM IPTG at 37 °C. B shows the growth curve of *E. coli* BL21(DE3)-T7 measured spectrophotometrically, indicating lag (yellow), exponential (green), and stationary (red) phases. C-E show the SDS-PAGE analysis of crude lysate (C), insoluble (D), and soluble (E) fractions from *E. coli* BL21(DE3)-T7 cultures induced with 0.1, 0.4, or 1 mM IPTG at either 37 °C or 15 °C. All gels were run alongside a molecular weight marker (MWM) with band sizes in kilodaltons (kDa). Uninduced controls were assessed during preliminary optimisation experiments and showed minimal background expression (data not shown).

The NiCo21(DE3) cells, however, are optimised to enhance the clarity of histidine tagged proteins purified via IMAC by the engineering of three proteins that often contaminate these systems, such that they can be removed via an additional purification step (New England Biolabs., 2020).

Cells were grown to an OD between 0.4 and 0.6 and induced with the reported lower limits of IPTG (0.1 mM), then allowed to express protein for four hours. The amount of soluble and insoluble recombinant N protein expressed by the two cell lines is depicted in [Figure 4. 2.A](#).

The theoretical molecular weight of the nucleocapsid protein is reported to be 45.6 kDa (Lu et al., 2021) whilst that of the W-NP, including the his-tag, is calculated by ExPASy to be 45.41 kDa. All the results reveal an overexpressed protein band within the 45 kDa region, consistent with the theoretical values; however, calculations from the SDS-PAGE indicate a molecular weight of 49.05 kDa, significantly higher than expected. This deviation could be explained by a variety of factors extending beyond that of experimental conditions alone, such as characteristics of the N protein.

The N protein is known to possess multiple post-translational modification (PTM) sites, and these modifications influence its structural and functional properties (Cheng et al., 2023). Any such modifications may alter its migration through the gel matrix, resulting in an apparent increase in molecular weight (Parker et al., 2010). In this study, the N protein was expressed in *Escherichia coli* cells, prokaryotic hosts that lack the enzymatic machinery required for such post-translational modifications (Itakura et al., 1977), thereby making the presence of such modifications not possible. Furthermore, mass spectrometric analysis ([Section 4.2.3](#)) revealed the presence of oxidised methionine residues, a form of modification that is not enzymatically mediated but of a spontaneous process. These modifications normally do not result in detectable increases in molecular weight estimations unless they cause major conformational changes (Vogt, 1995).

Intrinsically disordered proteins often exhibit this behaviour due to a lack of stable secondary and tertiary structures that affect their interaction with SDS (Contreras-Martos et al., 2018; Schramm et al., 2019), resulting in non-uniform interaction with

SDS, creating lower charge-to-mass ratios, ultimately resulting in slower migration and an overestimation of molecular weight (Garg et al., 2026; Romero et al., 2001; Uversky et al., 2000; Uversky, 2019). The unfolded nature and the potential of incomplete denaturation further led to increased hydrodynamic radius, similarly, causing slower migration (Garg et al., 2026; Romero et al., 2001; Uversky et al., 2000; Uversky, 2019). The high proportion of charged and polar residues similarly affects the interaction with SDS and the gel matrix, further influencing the migration pattern (Maiti et al., 2024).

Nonetheless, the intense bands are believed to be W-NP, and analysis of their intensity reveals significant insights. Although the method used aimed to ensure equivalence, it is important to note that the results display the amount of protein present within a 2 ml sample of culture, and not the amount of protein produced by a particular number of cells. Analysis of the soluble and insoluble fractions isolated reveal results contrary to previous findings, indicating that the majority of the protein produced is soluble, whilst a small portion is produced as aggregation-prone N protein located within the pellet, outperforming those previously reported. Although previous studies utilised BL21(DE3) cell lines, versus the BL21(DE3) T7 Express utilised here, the observed differences in protein solubility is more likely attributed to the lower induction optical density (OD₆₀₀ 0.4 – 0.6 compared to 0.8 – 1.0), which would have reduced the metabolic burden and promoted slower, more controlled expression. When considering the intensity of background protein expression, it could be anticipated that the BL21(DE3)-T7 Express cell line produces a higher yield of recombinant N protein within both insoluble and soluble fractions.

Although both cell lines demonstrated superior expression, the T7 Express cell line was selected for subsequent work. This decision was based on its consistently higher yield of soluble recombinant protein, as well as its reportedly more robust nature, reducing the optimisation burden whilst increasing throughput potential. While the NiCo21(DE3) strain offers advantages for IMAC purification by limiting co-purification of endogenous histidine-rich proteins, this was not a significant concern in the current study, as background contaminants were either minimal or sufficiently removed using the implemented purification workflow ([Section 4.2.2](#)). Furthermore, the presence of engineered modifications in NiCo21(DE3), while useful for some applications, introduces additional genetic complexity and is less well-characterised in large-scale

protein production workflows compared to the widely adopted BL21-based T7 Express system. Therefore, due to its higher soluble expression, simplicity, and established use in high-yield expression protocols, the BL21(DE3)-T7 Express strain was deemed the most appropriate host for continued production of the SARS-CoV-2 N protein in this study.

The cellular growth of *E. coli* BL21(DE3)-T7 (orange) at the chosen conditions, as determined by spectroscopy, is depicted in [Figure 4. 2.B](#). The cells exhibited a typical bacterial growth curve with well-defined lag, exponential, and stationary growth phases. The minimal increase in OD600 between 0 – 90 minutes is indicative of the lag phase (yellow blocked data points [Figure 4. 2.B](#)) during which bacterial cells adapt and acclimate to the environment. The rapid increase in OD600 between 120 – 300 minutes represents the exponential phase (green blocked data points [Figure 4. 2.B](#)) during which active cell division occurs, demonstrating robust growth and efficient utilization of nutrients. Inducing bacterial cultures with IPTG during the exponential phase ensures optimal protein expression by maximizing metabolic activity, preventing toxicity, improving yield, and enhancing protein solubility while minimizing resource waste (Larentis et al., 2014; Sivashanmugam et al., 2009). The plateau in OD600 between 330 – 450 minutes is considered to represent the stationary phase (red blocked data [Figure 4. 2.B](#)) during which growth is limited, usually by depletion of nutrients or the formation of inhibitory products. The optimal OD (0.5) was used for plasmid extraction using the *E. coli* JM109 cells as per the manufacturer recommendation, whilst OD 0.4 - 0.6 was used for IPTG induction using the *E. coli* T7-Express & Nico21(DE3) cell lines.

The impacts of IPTG, temperature, and incubation time on the expression of crude, soluble, and insoluble W-NP protein by *E. coli* T7-Express cells is presented in [Figure 4. 2.C, D & E](#). Preliminary optimisation experiments included uninduced controls, which demonstrated minimal background expression under the conditions tested (data not shown). Interrogation of the overexpressed band reveals a trend in target protein production. Seen particularly in [Figure 4. 2.C & D](#), higher levels of IPTG induction produce lower levels of target protein, and lower levels of IPTG produce a larger amount of target protein. This phenomenon can be attributed to a variety of factors. Higher levels of IPTG increase the activity of the T7 promoter often to such an extent

that it overwhelms the cell's protein-folding machinery, resulting in the accumulation of misfolded protein that leads to the formation of inclusion bodies. This is accompanied by the growth cost due to resource diversion toward the production of unnecessary lac protein. Together, these increase the toxicity of proteins and cellular stress, that not only activate cellular proteases that breakdown overexpressed target protein, but also significantly reduces growth and viability of cells, ultimately leading to premature death, without significantly increasing target protein yield (Gomes et al., 2020; Malakar & Venkatesh, 2012; Nag et al., 2022). These factors demonstrate the upper and lower limits of IPTG concentrations seen in the results.

Extended incubation time is believed to provide more opportunity for protein to fold correctly (Farewell & Neidhardt, 1998), however, it appears that target protein production reaches a maximum after 6 hours (**Figure 4. 2.B, C & D** lanes 5 –10). This observation could be explained by a shift in cellular energy allocation as the bacterial growth reaches the stationary phase, as per **Figure 4. 2.B** growth curve. At this point, nutrients are depleted, waste accumulates, metabolic activity declines, cellular proteases and autophagy are activated (Ughy et al., 2023).

Induction at lower temperatures is believed to slow down the rate of protein synthesis such that it promotes efficient protein folding, resulting in comparable or an increased amounts of protein when compared to higher temperature counterparts (Farewell & Neidhardt, 1998; Nag et al., 2022). Analysis reveals that lower temperatures produced comparable amounts of target protein without increasing its solubility (**Figure 4. 2.C, D & E** lanes 11-15).

In totality, in these samples, the majority of the target protein remained in the soluble fraction (**Figure 4. 2.E**), following the IPTG and induction time trends. Neither condition appeared to influence the solubility of the protein produced, enabling the researcher to use a logistically simpler and resource-saving final methodology. Consequently, it was decided to incorporate the use of lower IPTG concentrations (0.1 mM IPTG) accompanied by a 6-hour incubation time at 37 °C for subsequent experiments.

4.2.2 Protein Purification and Strategy Development

Ion-Exchange Chromatography

The SARS-CoV-2 Nucleocapsid protein has previously been purified using SP cation exchange chromatography (later polished using CM cation exchange chromatography) (Hanson et al., 2022; Li, G. et al., 2021). A similar strategy was attempted here, shown in **Figure 4. 3 A & B**. Crude, insoluble and soluble samples (**Figure 4. 3.B** Lanes 2, 3 & 4) show that W-NP was successfully expressed in *E. coli* T7 cells, with the majority present within the soluble fraction. Analysis of the sample application flow through, corresponding to the major UV 280 spike seen between 20-100 minutes, clearly shows that a significant portion of the target protein did not successfully bind to the cation exchange column upon loading (**Figure 4. 3.B** Lanes 22 - 24).

The low affinity for the resin at these initial binding conditions (IEC Buffer A, 20 mM Tris, pH 8.00) could be attributed to a variety of factors.

The binding affinity depends on the pI of the protein and the working conditions. The theoretical isoelectric point (pI) of the consensus sequence (W-NP) of the nucleocapsid protein, along with its his tag, is estimated between 9.95 - 10.13 (Expasy ProtParam: (Gasteiger et al., 2005)), meaning that at pH values lower than 9.95, the protein should be well protonated, or positively charged, allowing efficient interaction with the negatively charged sulfopropyl groups within the SP Sepharose® resin (Adhikari et al., 2010). Buffer pH is recommended to be at least 0.5 - 1.5 pH units away from the pI of the protein (Cytiva, 2021).

That suggests that a protein with a pI of 9.95, at a working condition of pH 8, should theoretically have sufficient charge to facilitate the interaction between protein and resin. The binding affinity depends on the ionic strength within the system, where increased ionic strength may shield charge, reducing the binding affinity. However, the binding buffer used (IEC Buffer A, 20 mM Tris, pH 8.00) contained no salt, maintaining the low ionic strength that promotes binding (Dismer et al., 2008; Ishihara & Yamamoto, 2005).

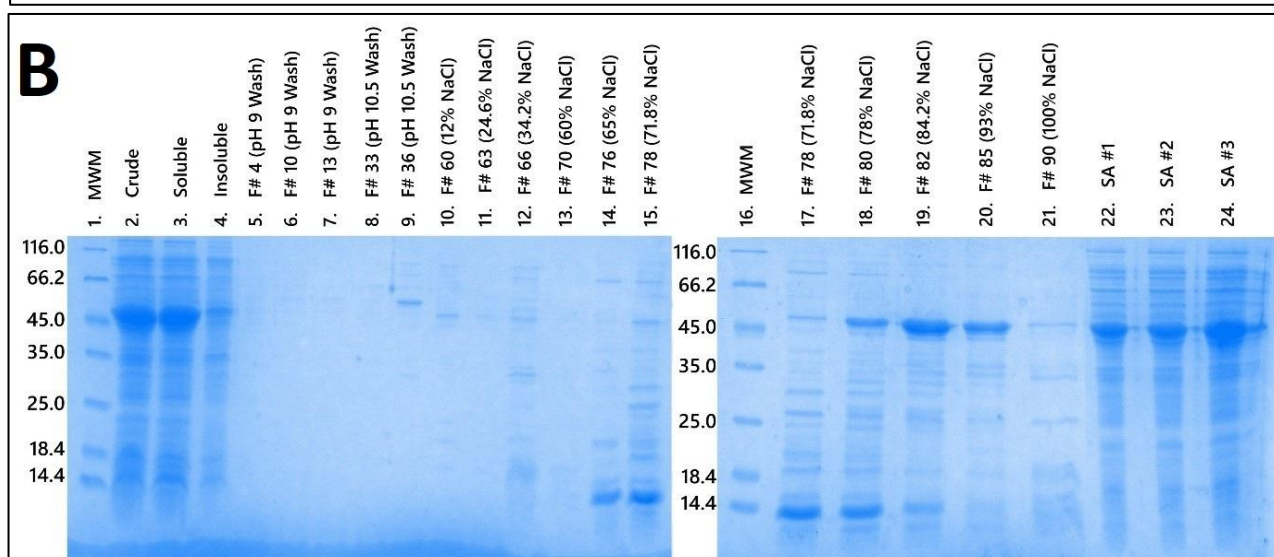
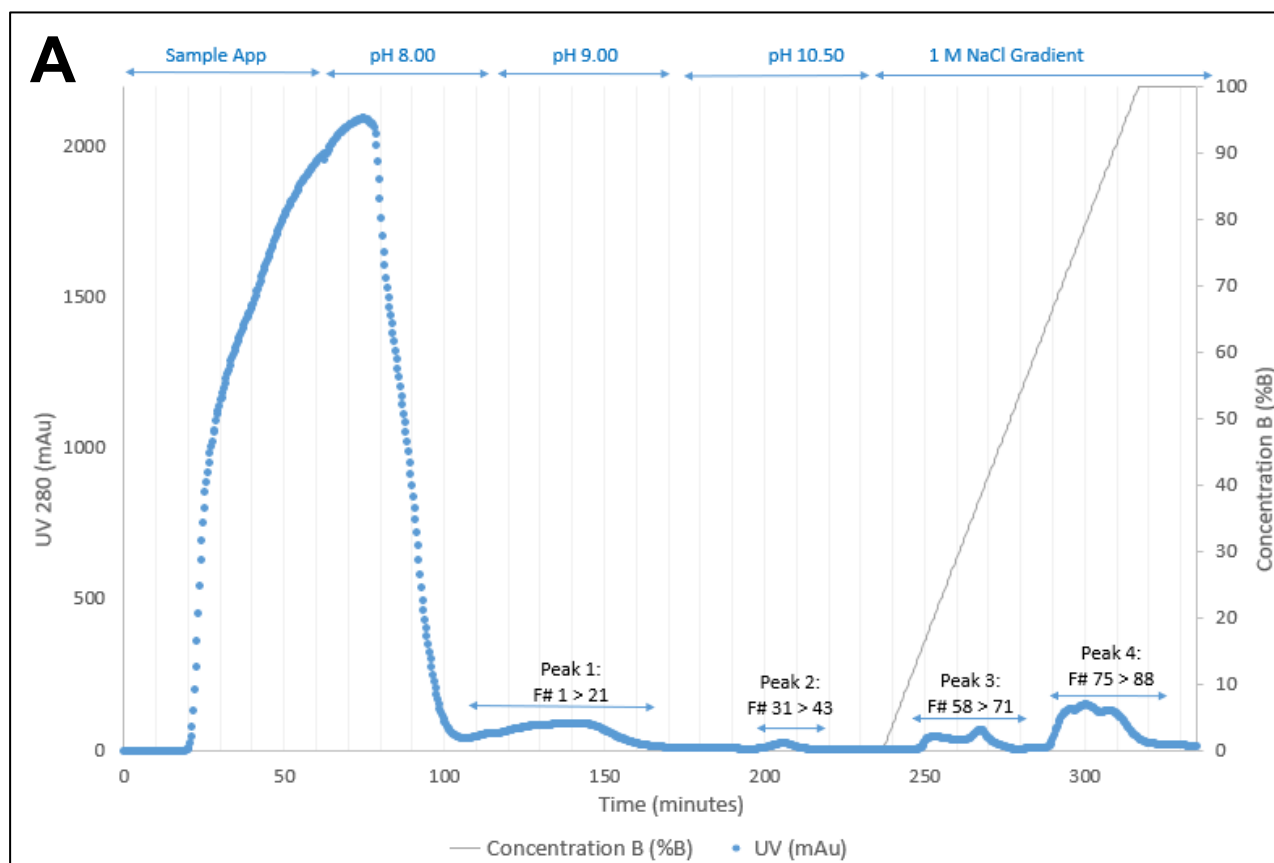


Figure 4. 3: Ion Exchange Chromatography of W-NP

A depicts the chromatogram of the IEC process displaying the UV absorbance (280 nm) in blue (protein elution), as well as the concentration of Buffer B in grey over the time in minutes. Demarked on the chromatogram is the sample application phase followed by the sequential washes using Buffers A, B, and C (20 mM Tris-HCl at pH 8.00; pH 9.00; pH 10.50 respectively) as well as the salt gradient using Buffer D (20 mM Tris-HCl, 1 M NaCl). Fractions (F#) of 2.5 mL were collected, and each peak with their corresponding fractions are demarked on the chromatogram, as well as various sample application flow through (SA). B shows the SDS-PAGE analysis of fractions collected during IEC displaying the protein content of fractions collected throughout the process against the molecular weight marker with its corresponding molecular weight (MWM) in kilodalton.

Similarly, it is unlikely that this may have been caused by the interference of the buffer with protein and resin, since the conditions chosen were within the pK_a (~ 8.1) range of the Tris buffer (Dismer et al., 2008; Ishihara & Yamamoto, 2005). The binding kinetics may also be affected by resin density, resin pore size, and poor column equilibration (Hardin et al., 2009). Although the Cytiva handbook does not directly state the pore size in nanometers or angstroms for this particular resin, it does imply a relative pore size as a 6% agarose matrix, which, according to them, is suitable for proteins up to ~ 500 kDa. However, the binding kinetics of protein to resin could most likely be attributed to the intrinsically disordered nature of the protein, as well as dimerization capabilities, that allow for flexible conformations that create transient exposure or burial of positively charged residues as well as the intrinsic heterogeneous distribution of charge that together may effectively reduce the electrostatic interactions (Arai et al., 2024; Li, W. et al., 2015; Mignon et al., 2025; Uversky, 2019).

In addition, the N protein is well known to be expressed and purified along with bacterial nucleic acid binding partners, as hypothesized to be the case here ([Section 4.2.3](#)). Interaction of the positively charged N protein with negatively charged nucleic acid could neutralize or shield the surface charge, reducing the binding affinity (Di et al., 2021; Lipfert et al., 2014). These could be improved by nucleic acid removal, often by denaturation, prior to IEC. Furthermore, if multiple proteins within the sample applied have similar isoelectric points or surface charge densities, they may compete for binding and co-elute under the various conditions.

The inefficient binding could likely be explained when considering that highly impure protein is eluted via broad peaks throughout the process ([Figure 4. 3.A & B](#)) (Wingfield, 2015). The formation of broad peaks indicates that too large a column volume was used, ultimately increasing the residency time and diluting the sample. The increased residency time could not have altered the binding affinity of the different proteins applied, but it could have changed their complete binding behaviours by moving closer to a true equilibrium by along longer association time (K_{on}), increasing the competition between the various proteins present in the sample, potentially increasing the amount of binding sites occupied by contaminants. However, this would not only change the binding profiles, but also the elution profiles.

Although the optimization of sample application as well as the NaCl gradient slope, or maintain flow rate, could enhance the binding affinity and resolution of separation, the presence of the overlapping elution profiles suggests that IEC alone may not be efficient for the purification of the N protein, and that alternative strategies like Nickel affinity chromatography may be required.

Nickel chromatography

To achieve efficient separation of W-NP target protein produced within *E. coli* BL21(DE3)-T7 Express cellular proteins via Nickel Affinity Chromatography, a standard protocol was developed and optimised from thereon (Table 4. 1). In this study, we opted to retain our 20 mM Tris pH 8.00 as per expression. Tris is used to maintain the pH of 8.00, a condition that allows histidine residues to remain deprotonated, facilitating efficient binding with the nickel ions on the affinity resin (Bornhorst & Falke, 2000; Scheich et al., 2003). In addition, Tris has minimal metal-binding properties (doesn't bind to Nickel), making it a good buffering agent that is compatible with a variety of downstream processes (Bornhorst & Falke, 2000). The study utilised 1 M NaCl concentration to enhance protein solubility, maintain ionic strength, reduce non-specific electrostatic interactions, and limit nucleic acid association (Kolesov et al., 2022; Luan et al., 2022; Zhao et al., 2021).

The standard protocol involved three purification buffers, Buffer A, Buffer B, and Buffer C, each with increasing concentrations of imidazole, each serving the purpose of a binding buffer, wash buffer, and elution buffer, respectively. In order to assess the specific conditions that allowed for the removal of contaminants whilst maintaining a high yield of pure target protein, a 50 ml linear gradient was incorporated between each buffer. This strategy relies on the association and dissociation kinetics that govern the binding and release of proteins during IMAC:

$$K_D = \frac{\kappa_{off}}{\kappa_{on}}$$

Equation 6

Where K_D is the equilibrium dissociation constant, and K_{on} and K_{off} is the rate at which proteins bind and detach, respectively. This technique results in a continuous and linear, or gradual, increase in imidazole concentration over time, allowing for

differential elution of proteins based on their affinity for the resin. The various optimisation strategies attempted are visualised in [Table 4. 1](#) and [Figure 4. 4](#).

Table 4. 1: Comparison of Ni-NTA purification strategies

This table summarises the various imidazole concentrations used in the buffers of each Nickel Affinity Chromatography strategy explored to purify W-NP.

Strategy	Binding Buffer	Wash Buffer	Elution Buffer
A) Initial Strategy	10 mM	150 mM	600 mM
B) Elution Gradient	10 mM	150 mM	400 mM
C) Wash Gradient	10 mM	150 mM but 25 ml	600 mM
D) Binding Buffer	45 mM	150 mM	600 mM

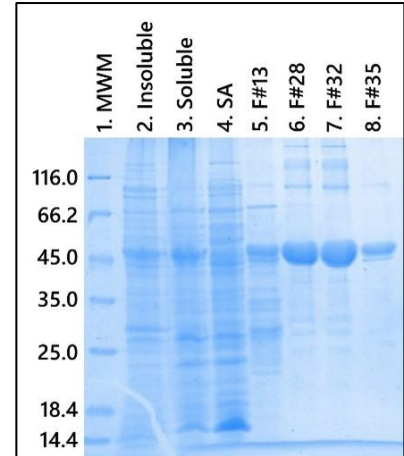
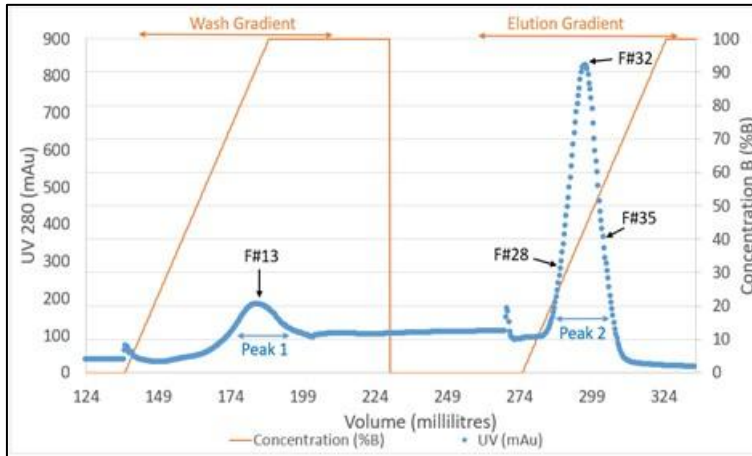
As an initial purification strategy, or strategy A ([Table 4. 1](#)), 10 mM imidazole was used in the binding buffer to prevent non-specific binding, 150 mM imidazole was used in the wash buffer to elute any non-specifically bound proteins, and 600 mM imidazole was used to elute any bound proteins.

Strategy A is displayed in [Figure 4. 4.A](#). Analysis of the protein content within fractions suggests that sufficient target protein was present within the sample that was applied (lane 3). Analysis of the sample application (SA) flow through (lane 4) indicates that the majority of the target protein successfully bound to the column, whilst the majority of contaminants did not. Ultimately, the protein eluted in two distinct UV 280 peaks, one during both the wash and the elution gradient. Analysis of the protein content eluted during those peaks (lanes 5 – 8) reveal well contaminated target protein eluting between 100 - 150 mM imidazole (peak #1; 65-100 %B in the wash gradient), and less contaminated target protein eluting between 220 - 460 mM (peak #2; 16– 70 %B in the elution gradient), concentration of imidazole. This multi-range elution profile was similarly seen and well explained during IEC ([Section: 4.2.2](#)), yet in short may be attributed to the heterogeneity in binding affinity and interactions with resins owing to the intrinsically disordered and flexible nature of the protein, complicated by its oligomeric and post-translational modification potential (although traditional PTMs are not anticipated). However, potential degradation, resin overloading, and various other experimental parameters may similarly affect the elution profile during IMAC.

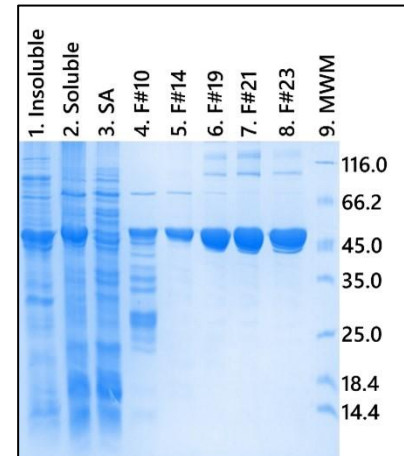
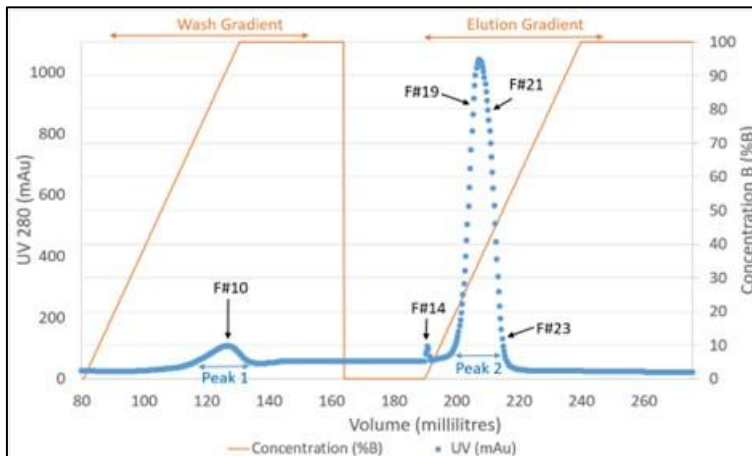
1

2

A



B



C

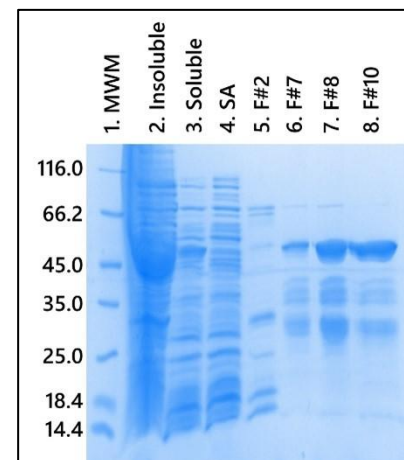
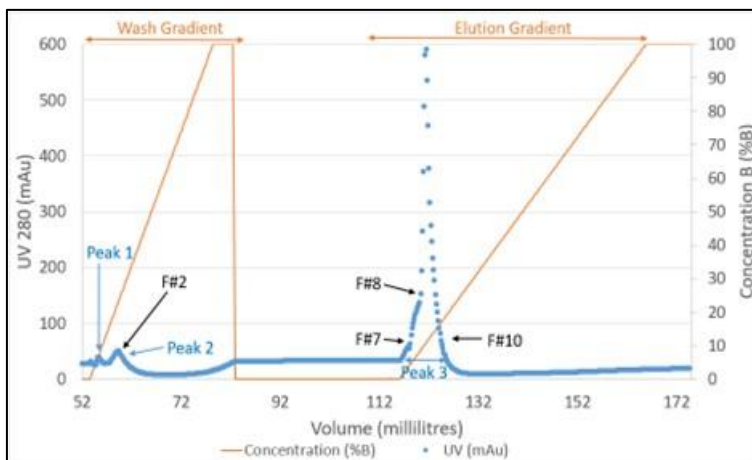


Figure 4. 4: Ni-NTA optimisation strategies for W-NP purification

This figure displays the results of the different optimization strategies used to purify W-NP via Nickel Affinity Chromatography. Panels 1 display the chromatograms illustrating the UV 280 nm absorbance (blue scatter; primary y-axis) as well as the concentration B (orange line; secondary y-axis) over the time in millilitres (x-axis) during the wash and elution gradients. Demarked on each figure is the peak numbering as well as the position and number of each fraction (F#). Panel 2 displays the corresponding SDS-PAGE result displaying the protein content of each fraction against the molecular weight marker (MWM) with its sizes marked in kilodalton (kDa). Figure A displays the initial strategy, Figure B displays the optimization of the elution gradient, and Figure C displays the optimization of the wash gradient.

These results highlighted the necessity of a systematic evaluation of the factors that influence the association and dissociation kinetics of the sample, starting with strategy B, the optimization of the elution gradient (**Figure 4. 4**), during which the imidazole concentration in Purification Buffer C (PBC) is reduced from 600 mM to 400 mM

Reducing the imidazole concentration in PBC would result in a slower increase in the concentration of imidazole during the elution gradient that would theoretically cause a slower shift in equilibrium towards the dissociation of protein, resulting in broad protein elution peaks that allow for more resolved separation of proteins (Wingfield, 2015). However, this strategy did not result in the expected behaviour. Instead, the majority of target protein eluted both earlier and faster, now between 185 mM and 285 mM imidazole (14-54 %B in the elution gradient).

This finding corroborates the hypothesis of the impact of the intrinsic nature of the N protein and nucleic acid binding capacity and behaviour on its association and dissociation kinetics (Cubuk, Alston et al., 2024), such that the equilibrium of His-tag binding to Ni^{2+} is not linear with respect to imidazole concentration, resulting in a sigmoidal dissociation process created by the shift in equilibrium induced by this change (Bornhorst & Falke, 2000; Hochuli et al., 1988; Terpe, 2003; Wingfield, 2015).

However, this strategy appeared to have resolved contaminants of smaller molecular weight than the target protein during the final elution step, yet larger molecular weight contaminants remain present, and some target protein is still lost during the wash step.

Strategy C attempted to separate contaminants earlier by optimising the wash gradient through a reduction in gradient length from 50 ml to 25 ml (**Figure 4. 4.C**). The shorter length of the gradient would cause a much faster increase between 10 mM and 150 mM imidazole. At these concentrations of imidazole, this change would theoretically shift the equilibrium such that proteins with lower affinities (somewhat higher K_d values) would elute earlier and faster, potentially minimizing their overlap with the target protein that is expected to have a much higher affinity (or lower dissociation) value. This alteration did cause earlier and faster elution of contaminants during the wash step without the loss of target protein (**Figure 4. 4.C**, lane 5), now eluting between 10-82 mM imidazole. However, this strategy allowed for smaller molecular

weight contaminants to elute alongside the target protein during the elution gradient. Together, Strategies B and C clearly demonstrate how the IMAC system does not only depend on the binding and dissociation kinetics of the target protein only but also those of the contaminants.

Another strategy frequently used to optimise the resolution and purity of an IMAC system is the optimisation of the binding buffer. The concentration of imidazole within the binding buffer plays a significant role; too little may allow a variety of contaminating proteins to bind via their intrinsic histidine residues, while too much can interfere with the target protein's binding to the column, ultimately affecting both the purity and final yield of the target protein recovered. To address this, we gradually increased the imidazole concentration in the binding buffer—first to 20 mM and then to 40 mM (results not shown)—before finalizing at 45 mM imidazole (Strategy D; [Table 4. 1](#) and [Figure 4. 5](#)).

This adjustment did not hinder W-NP from binding to the column (lane 4). In fact, it reduced the loss of target protein during the wash step when contaminants eluted between 70 – 150 mM imidazole (peak #1; lanes 5 and 6; corresponding to 26 – 100 %B in the wash step).

Notably, pure target protein eluted in two distinct peaks: one forming late in the wash step (peak #2; lanes 8-16; at 150 mM imidazole or 100 %B) and another during the elution step (peak #3; lanes 19-24; at 235 – 375 mM imidazole). This is most likely explained by the different oligomeric species seen in [Figure 4. 9](#).

These observations underscore how the starting concentration of imidazole influences contaminant binding to the column and delineate the various conditions under which the target protein elutes, supporting the idea that multiple species are present.

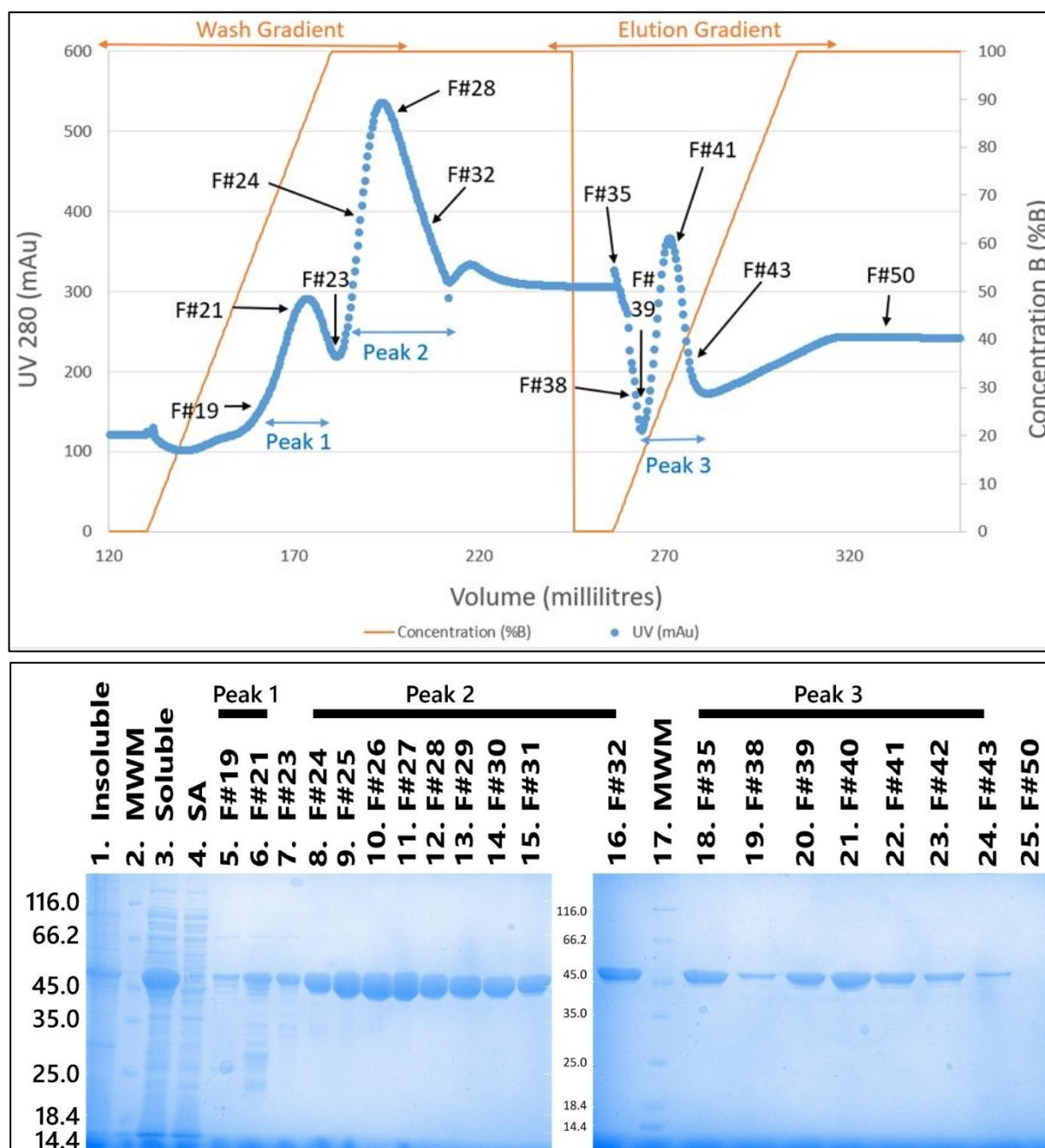


Figure 4. 5: Final Ni-NTA purification strategy

This figure displays the chromatogram (A) and SDS-PAGE (B) of the fourth and final purification strategy (D) used to purify W-NP using Nickel Affinity Chromatography. This strategy employs a 45 mM imidazole binding buffer to the initial strategy. (A) Chromatogram displaying the UV 280 nm absorbance (blue scatter; primary y-axis), as well as the concentration B (orange line; secondary y-axis), over the time in minutes (x-axis) during. The various peaks and location of the fractions (F#) are demarked. (B) SDS-PAGE results of the protein content of fractions collected using the finalized nickel affinity chromatography strategy. It also contains the name and number of each fraction (F#) and the peak to which it belongs against the molecular weight marker (MWM; lanes 2 and 17) with its sizes demarked in kilodalton (kDa), also displaying the sample application flow through (SA), insoluble and soluble fractions isolated from the cultures.

4.2.3 Protein Composition Analysis and Identity Confirmation

The quantification of purified W-NP was performed using both the Bradford assay and UV-Vis absorbance spectroscopy (**Figure 4. 6 A & B**). The Bradford assay yielded a protein concentration of 68.29 μM , whereas UV-Vis measurements indicated a slightly lower concentration of 62.15 μM .

This discrepancy is likely attributable to the presence of two distinct absorbance peaks in the UV-Vis spectrum (**Figure 4. 6 B**)—one at 280 nm, corresponding to aromatic residues (e.g., tryptophan and tyrosine), and another at 260 nm, typical of nucleic acid binding proteins like the N protein, co-purified alongside nucleic acid (Grimsley & Pace, 2004; Porterfield & Zlotnick, 2010; Vidal et al., 2023) with an A₂₆₀/280 ratio of 0.99. This spectral overlap complicates the accurate determination of protein concentration using UV-Vis spectroscopy, as both protein and nucleic acid components contribute to the overall absorbance (Porterfield & Zlotnick, 2010; Vidal et al., 2023). In addition, the binding of nucleic acids to the protein may induce conformational changes that alter hyperchromicity, further confounding spectroscopic quantification (Porterfield & Zlotnick, 2010). While several correction models have been proposed for addressing this interference, these rely on accurate knowledge of the nucleic acid extinction coefficients (Porterfield & Zlotnick, 2010), which remain undetermined in this study. Nonetheless, these findings support previous findings that suggest that these mutations introduced but do not abrogate the protein's nucleic acid binding functions (Alsuwairi et al., 2023; de Silva et al., 2021; Li, G. et al., 2021; Nguyen et al., 2024; Obermeyer et al., 2022; Oulas et al., 2021).

In turn, Coomassie G-250 used in the Bradford assay can similarly lead to an overestimation during protein quantification, since this dye is also known to interact with nucleic acids, although to a weaker extent (Wenrich & Trumbo, 2012). Depending on the downstream processing required, purified protein can be separated from nucleic acids using a variety of techniques, or the bicinchoninic acid assay could more selectively quantify the proteins with little interference from the nucleic acid (Wiechelman et al., 1988). Nevertheless, these findings suggest that the protein isolated has functional capabilities in terms of nucleic acid binding.

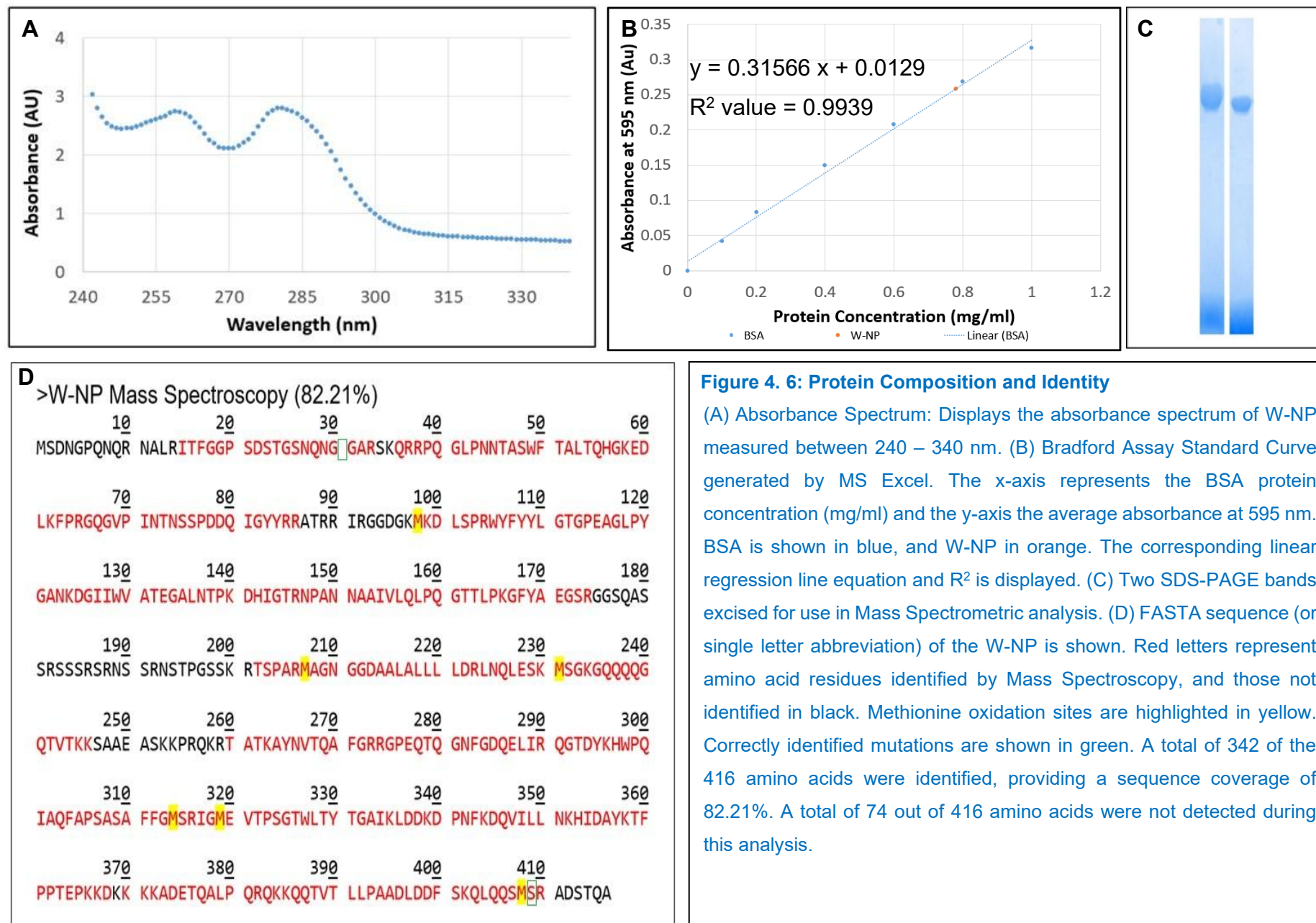


Figure 4. 6: Protein Composition and Identity

(A) Absorbance Spectrum: Displays the absorbance spectrum of W-NP measured between 240 – 340 nm. (B) Bradford Assay Standard Curve generated by MS Excel. The x-axis represents the BSA protein concentration (mg/ml) and the y-axis the average absorbance at 595 nm. BSA is shown in blue, and W-NP in orange. The corresponding linear regression line equation and R^2 is displayed. (C) Two SDS-PAGE bands excised for use in Mass Spectrometric analysis. (D) FASTA sequence (or single letter abbreviation) of the W-NP is shown. Red letters represent amino acid residues identified by Mass Spectroscopy, and those not identified in black. Methionine oxidation sites are highlighted in yellow. Correctly identified mutations are shown in green. A total of 342 of the 416 amino acids were identified, providing a sequence coverage of 82.21%. A total of 74 out of 416 amino acids were not detected during this analysis.

The mass spectrometric analysis of the duplicate gel digests obtained from separate rounds of purification (**Figure 4. 6 C & D**) confirmed that both samples analysed belong to W-NP. A total of 86 unique peptide fragments were identified between the two samples, corresponding to 82.21% of the full-length sequence. Given the incomplete coverage of the protein and PTMs identified below, an accurate size estimation cannot be reported (Beavis & Chait, 1990); however, both samples exhibited high MS scores, high intensity readings, and high MS/MS counts, which support the robustness of the identification of the mutations incorporated. Notably, shown in green (**Figure 4. 6 D**), the 31-33 ERS deletion and the S413R mutation (renumbered as position 409 in the mutated sequence) were detected, while the R203K and G204R mutations (positions 200 and 201) were not identified.

MS analysis also revealed several PTM oxidation events. Unlike traditional PTMs, amino acid residues such as methionine, cysteine, tryptophan, and tyrosine residues may undergo spontaneous oxidation (not enzyme-dependent) that can significantly influence protein function, stability, and interaction with other biomolecules (Kim et al., 2014). Some oxidation events may be required for protein function, whilst others occur when exposed to oxidative conditions, e.g., exposure to oxidizing agents like reactive oxygen species (ROS) or reactive nitrogen species (RNS). The MS analysis identified oxidized methionine (M) sites at positions 207, 231, 314, 319, 408, in both samples; W-NP sample #1 additionally showed oxidation at position 98. All the identified sites are methionine residues, frequently oxidised to methionine sulfoxide (MetO) or methionine sulfone (MetO₂) residues. The experimental protocols employed dithiothreitol (DTT) as a reducing agent to preserve protein integrity during storage and analysis. Although DTT is effective at preventing and reversing cysteine oxidation, it does not protect against the oxidation of methionine, tyrosine, or histidine residues (Kim et al., 2014; Santarino et al., 2012). The observation of methionine oxidation, despite DTT usage, could suggest that these modifications occurred before sample storage, potentially due to exposure to oxygen or other oxidative conditions during sample handling; while the mildly basic working pH (8) may further promote such oxidation, whereas these modifications are generally minimised under acidic, low-temperature, and oxygen-limited conditions (Kumar et al., 2013). Given the reversibility of methionine oxidation, these findings also highlight the potential for studying

oxidation as a regulatory mechanism in protein function, opening avenues for future research.

4.3. Structural Features of W-NP

4.3.1 Far UV CD

The Far-UV CD spectra ([Figure 4. 7.A](#)) obtained from the purified nucleocapsid protein (W-NP) exhibited a strong negative trough between 190 and 210 nm with a minimum at 200 nm accompanied by a positive band centred at 229 nm and another negative band at 235 nm. Several other research groups (Azad, 2021; Nguyen et al., 2024; Ribeiro-Filho et al., 2022; Zeng, W. et al., 2020; Zhao et al., 2021) have reported a similarly pronounced trough around 200 nm for the N protein and associated it with proteins that lack stable secondary structure or predominantly random coil proteins (Ranjbar & Gill, 2009). It is important to note that the following spectroscopic measurements were taken under low ionic strength conditions, likely influencing the conformational state and protein-RNA interactions.

Referencing [Table 4. 2](#), the feature viewer on the UniProt entry to the Wuhan SARS-CoV-2 N protein (UniProt: P0DTC9) estimates that this sequence contains 60.14% random coils, 5.97% sheets, 16.71% turns, and 16.95% helices. The AlphaFold generated structure of W-NP predicted that the incorporation of these mutations would result in reduced disordered regions alongside increased ordered regions consisting of 56.73% coils, 11.30% sheets, 12.98% turns, 19.71% helices. Deconvolution, or quantification, of the secondary structural features obtained from the Far-UV CD spectra through the BeStSel algorithm reveals that W-NP contains an estimated 7.2% helices (all of which are distorted), 20.9% antiparallel β -sheets, 14.1% turns, and 57.7% disordered (“others”).

Table 4. 2: Quantification of Nucleocapsid Protein Secondary Structure

This table quantifies the secondary structure content of the Wuhan sequence (PDB: 8FD5), the AlphaFold predicted W-NP, and BeStSel quantification of CD data in percentages.

Secondary Structures	Wuhan	W-NP	BeStSel
Helices	16.95%	19.71%	7.2%
Sheets	5.97%	11.30%	20.9%
Turns	16.71%	12.98%	14.1%
Coils	60.14%	56.73%	57.7%

The significant variation in the secondary structural content may be explained by a variety of factors. Firstly, it is important to remember the highly flexible and mostly intrinsically disordered nature of the SARS-CoV-2 nucleocapsid protein. The PDB database houses experimentally determined entries often obtained via X-ray crystallography that is limited by resolution, often resulting in missing electron densities, especially when a protein is too flexible or dynamic to adopt a single, well-defined 3D conformation, requiring inferences of the missing data (Bloomer et al., 1978; Davis et al., 2008; Kojima et al., 2024). However, regions that are highly flexible or disordered may lack observable electron density independent of overall resolution, and are therefore not modelled (Rogers et al., 2019; Uversky, 2019, 2025). This technique requires the formation of protein crystals that, in this case, may alter the arrangement of the protein, providing results that may not be accurate in terms of its native structure (Bloomer et al., 1978; Davis et al., 2008). These are some of the reasons why no full-length 3D structure exists for the nucleocapsid protein; therefore, the values of secondary structural content, quantified by UniProt, are not a true reflection of experimentally determined structures, but an overall estimate obtained by combining experimental and predictive data.

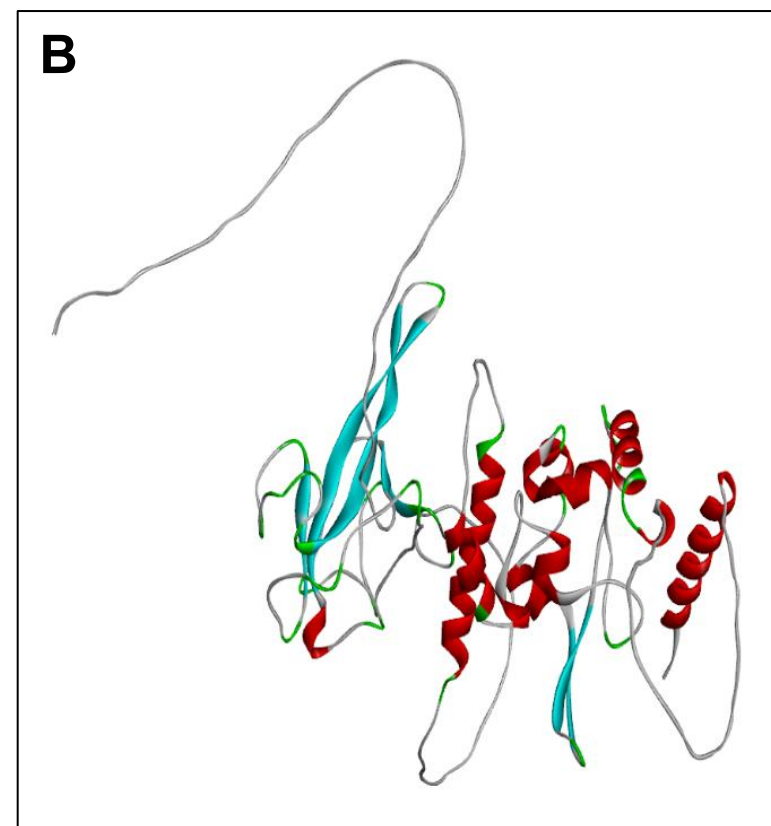
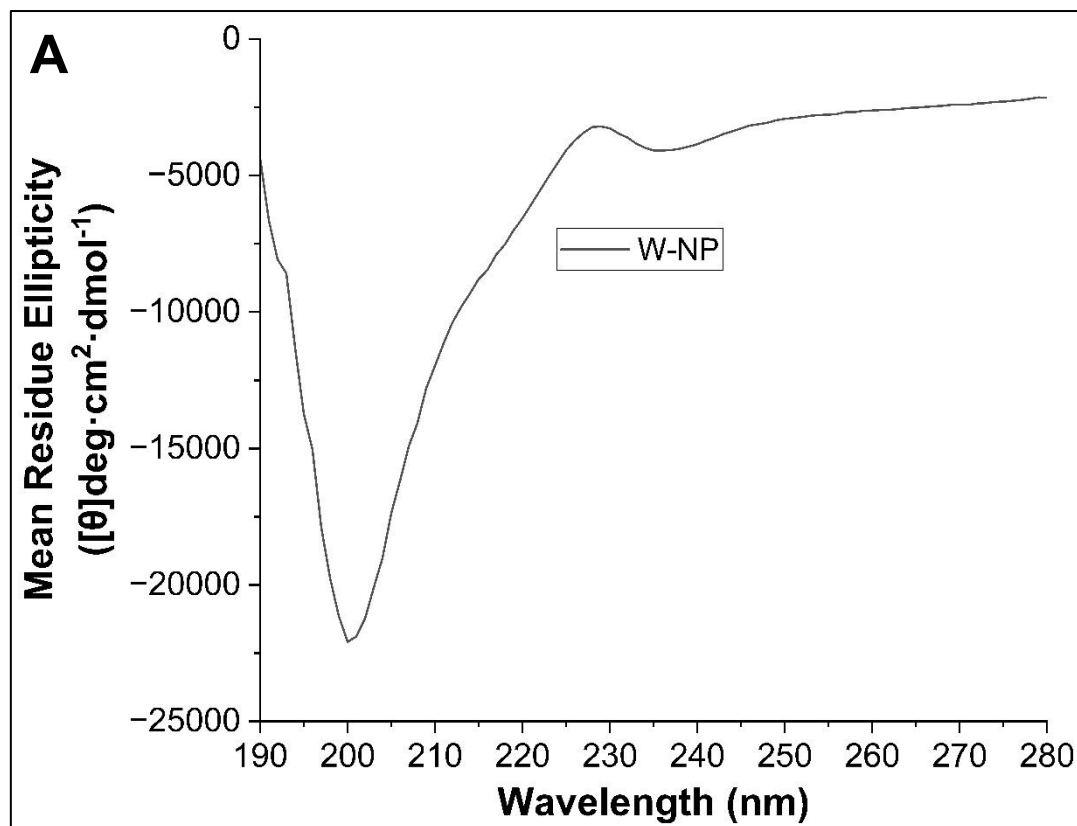


Figure 4. 7: Secondary Structure of W-NP

(A) CD spectra of 5 μ M W-NP in storage buffer (20 mM Tris-HCl; 1 mM EDTA; 1 mM DTT; 10% (v/v) glycerol; pH = 8.00), displaying wavelength on the x-axis and mean residue ellipticity on the y-axis. (B) AlphaFold predicted 3D structure of W-NP colour-coded by secondary structure showing sheet content in cyan, turns in green, and helical content in red, and disordered content in grey.

Furthermore, these conformations depend on the environmental conditions (i.e., temperature, pH, ionic strength, etc), ligand binding (i.e., nucleic acid), post-translational modifications, aggregation, and oligomeric states (Davis et al., 2008; Kojima et al., 2024; Samad et al., 2024). In addition, the secondary structure prediction algorithms are limited by training data bias, mostly relying on ordered proteins (Han et al., 2023).

Nonetheless, these observed discrepancies may be a result of the structural implications brought on by the coexisting mutations. The P13L and $\Delta 31-33$ introduce a novel oligomerization interface, promoting transient helix formation that facilitates higher-order assembly (Nguyen et al., 2024). Meanwhile, the R203K/G204R mutations within the LKR disrupt the RG motif that drives phosphorylation, resulting in enhanced liquid-liquid phase separation, RNA binding, and oligomerization (Shuaib et al., 2023; Stuwe et al., 2024; Wu, H., Xing, Meng, Fu, Xue, Dong, Tang et al., 2021; Zhu et al., 2021). All of these changes are associated with the promotion of disorder-to-order transitions characterised by increased β -sheet signals.

4.3.2 Fluorescence

When excited at 295 nm (**Figure 4. 8 A**), W-NP generates a fluorescence emission spectrum peak between 300 and 450 nm with a maximum at 336 nm. Excitation at this wavelength selectively excites tryptophan residues, and the resulting $\gamma_{\max}$ suggests that the largest contributor to this signal originates from buried tryptophan residue/s (Lakowics, 2006; Munishkina & Fink, 2007).

Referencing **Figure 4. 8 C** and **Table 4. 3** the W-NP sequence constructed here contains 29 fluorophores in total (5 tryptophan residues, 13 phenylalanine residues, and 11 tyrosine residues), constituting 6.9% of the total amino acid composition, with notable clustering in specific regions, especially within the RNA-binding domain. Three of the five tryptophan residues (Trp 49, 105, 129) are located within the NTD, whilst the remaining two (Trp 298 & 327) are located within the CTD. Analysis of Solvent Accessibility Scores (SAS) using Discovery Studio suggest that the Trp residues within the NTD are considered buried (less than 10% maximum SAS) whilst those in the CTD are considered exposed (greater than 25% maximum SAS) (**Table 4. 3**), suggesting that the signal is dominated by a Trp residue/s likely located in the NTD. Nonetheless, all the Trp residues become further buried upon dimerization (**Table 4. 3**), and potentially even more so upon trimerization.

Table 4. 3: Solvent Accessibility Scores of Monomeric & Dimeric W-NP Tryptophan Residues

This table compares the Residue Solvent Accessibility, Side Chain Solvent Accessibility, Percent Solvent Accessibility and Percent Sidechain Solvent Accessibility scores of the five tryptophan residues within the AlphaFold predicted W-NP monomeric and dimeric forms calculated using Discovery Studio.

Name Secondary	TRP49 Coil		TRP105 Sheet		TRP129 Sheet		TRP298 Helix		TRP327 Sheet	
	Monomer	Dimer	Monomer	Dimer	Monomer	Dimer	Monomer	Dimer	Monomer	Dimer
Residue Solvent Accessibility	50.30	17.60	13.08	9.56	43.09	42.54	73.79	21.88	109.03	51.78
Sidechain Solvent Accessibility	49.81	17.11	13.08	9.56	15.60	14.09	73.79	21.88	109.03	51.78
Percent Solvent Accessibility	20.77	7.27	5.40	3.95	17.79	17.57	30.47	9.04	45.03	21.38
Percent Sidechain Solvent Accessibility	26.04	8.94	6.84	5.00	8.15	7.36	38.57	11.44	56.98	27.06

Distance monitoring through the software suggests that all of these Trp residues are within close proximity (2 – 3 nm; or 20 – 30 Å), creating an environment conducive to homo-Förster Resonance Energy Transfer (FRET), where energy is transferred non-radiatively from a donor fluorophore to an acceptor fluorophore of the same species (Yang, Hui et al., 2017/03/08).

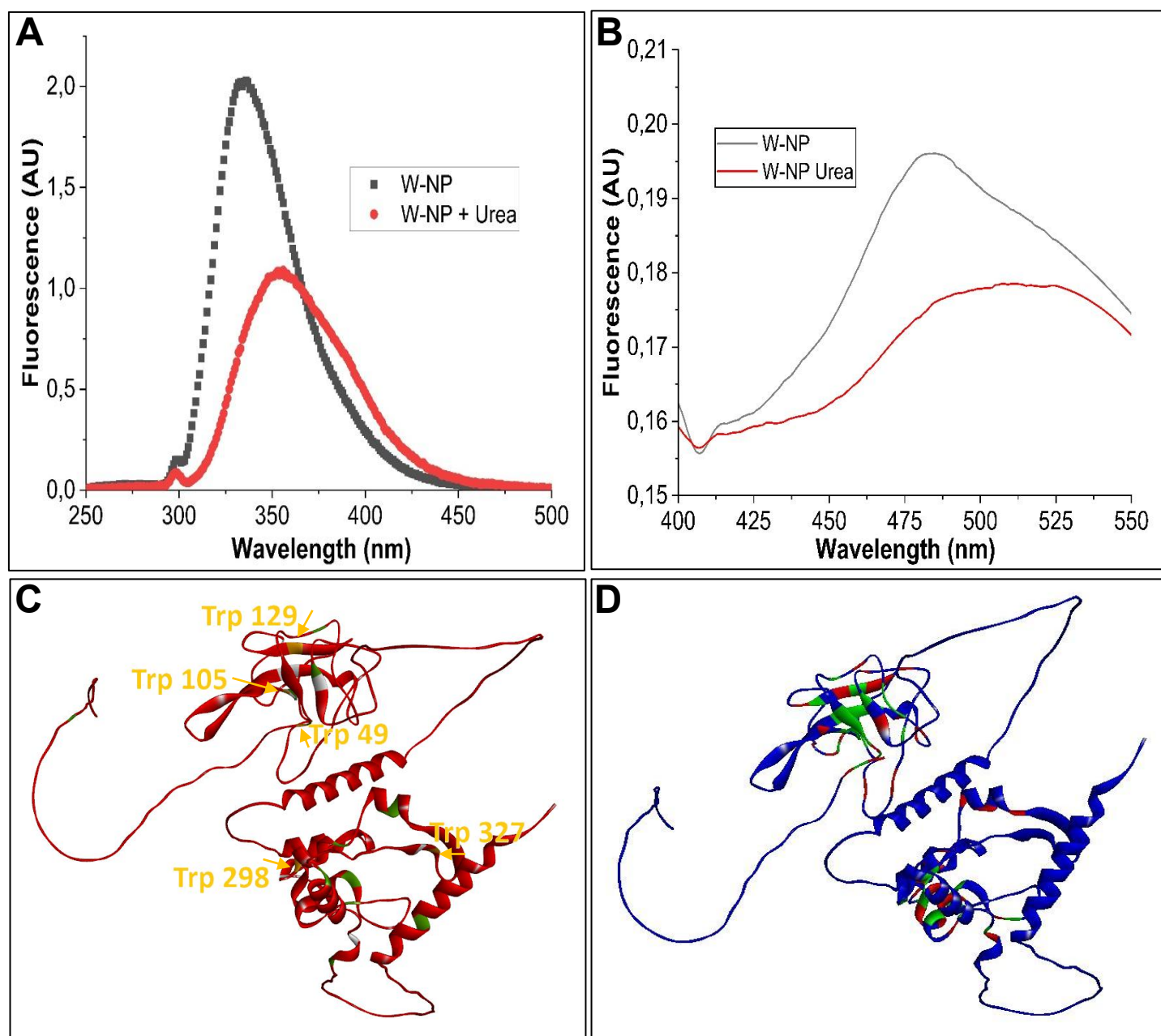


Figure 4. 8: Fluorescence of W-NP

Figures A & B display the fluorescence spectra of 5 μ M W-NP in the absence (grey) and presence (red) of 7 M urea, depicting the wavelength in nanometres (nm) on the x-axis and the fluorescence intensity in arbitrary units (AU) on the y-axis. Panel A displays the fluorescence when excited at 295 nm. Panel B displays the fluorescence intensity when in the presence of 250 μ M 8-anilinalphthalene-1-sulfonic acid (ANS) excited at 380 nm. Figures C & D display the AlphaFold predicted 3D structure of W-NP. Figure C highlights the various fluorophores with tryptophan residues in orange and demarked, phenylalanine residues in green and tyrosine residues in white. Figure D is coloured by Solvent Accessibility Scores predicted by Discovery Studio with red being the most buried, green as somewhat buried, and blue as exposed residues.

However, FRET relies not only on distance, but also on the correct orientation of residues, a factor that might change upon nucleic acid binding or dimerization. Nevertheless, such intra-protein FRET would not result in an increase in overall fluorescence intensity, but rather a mere redistribution of the excitation energy (Vivian & Callis, 2001).

The exposure of W-NP to 7 M urea causes a red shift of the emission maximum from 336 nm to 356 nm, along with a significant reduction of the emission intensity. These changes are consistent with the unfolding of protein caused by urea, leading to the exposure of tryptophan residues that now emit at longer wavelengths (red-shift of the λ_{max}) as well as their exposure to quenchers found in the environment (i.e., oxygen or buffer components) that cause the reduction of emission intensity (Lakowics, 2006).

4.3.3 ANS Binding Studies

ANS, or 8-anilinonalphthalene-1-sulfonic acid, is widely used as a fluorescent probe of hydrophobic regions within proteins, since its emission properties are altered depending on the environment it can be found (Kato & Nakai, 1980). The structure and functionality of the protein are supported through hydrophobicity studies entailing the investigation of ANS fluorescence (**Figure 4. 8 B**). Here, W-NP generates a peak between 430 – 550 nm, with a maximum at 485 nm, indicating that ANS is within a hydrophobic environment (non-polar) (Gasymov & Glasgow, 2007; Kato & Nakai, 1980). This peak is shifted to 505 nm upon the addition of 7 M urea, suggesting that ANS is in an aqueous environment (Kato & Nakai, 1980).

The W-NP consists of $\pm 25\%$ hydrophobic amino acids (I, V, L, F, C, M, A), that along with the $\pm 15\%$ positively charged amino acids (K, R, H), create both hydrophobic patches and nucleic acid binding sites. These, along with the intrinsically disordered regions and flexibility of the protein, allow the hydrophobic patches to play a key role in the stabilisation of both protein-protein and protein-nucleic acid interactions (Bai et al., 2021; Cubuk et al., 2021; Gao et al., 2021; Huang et al., 2024; Peng et al., 2020; Wu, C. et al., 2021; Yang, Minghui et al., 2024).

4.3.4 BN-PAGE

The quaternary structure of W-NP isolated at these conditions were investigated using blue native PAGE (BN-PAGE) using a 12% polyacrylamide gel presented in [Figure 4. 9](#). This gel concentration is optimal for the resolution of proteins within the 20 – 100 kDa range but is insufficient in resolving higher molecular weight species (Wittig et al., 2006), as evidenced by the lack of large molecular weight proteins within the molecular weight marker ([Figure 4. 9](#) Lane 1).

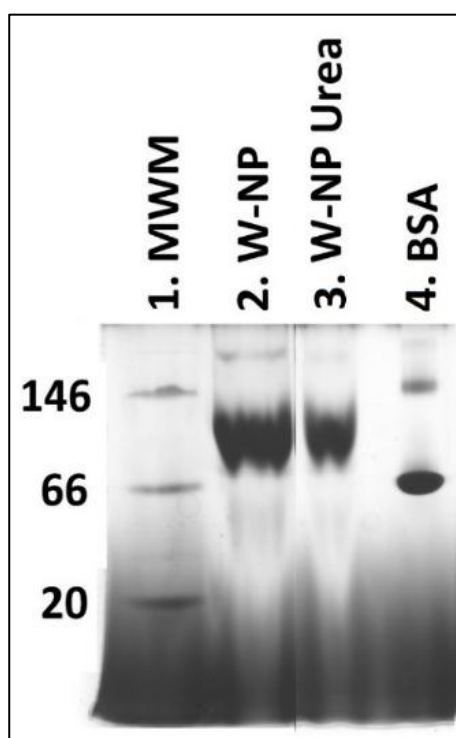


Figure 4. 9: BN-PAGE

This figure depicts the migration profile of 5 μ M W-NP in the presence/absence of 7 M urea, as well as that of BSA using a 12% Blue-Native Polyacrylamide Gel. Lane 1 displays the Molecular Weight Marker (MWM) with its corresponding sizes in kilodaltons (kDa).

BSA used as a control in this study ([Figure 4. 9](#) Lane 4), displayed a primary band within the theoretical monomeric mass size of 66 kDa, alongside two higher molecular weight species, likely representing dimeric and/or trimeric forms (~132 and 198 kDa), consistent with known BSA behaviour under the high concentrations used here (Babcock & Brancalion, 2013; Levi & Flecha, 2002).

W-NP ([Figure 4. 9](#) Lane 2) similarly generates a multi-band pattern, suggesting the presence of multiple species within the samples isolated. It consists of a faint band migrating below the 66 kDa marker corresponding to the predicted monomeric molecular mass of W-NP (~46 kDa), an intense band between the 66 and 146 kDa markers, that could represent either or dimer or trimeric forms of the protein (~90 or ~138 kDa), as well as an additional diffuse band above the 146 kDa marker, likely

representing the tetrameric form (~184 kDa). These findings align with previously reported SARS-CoV-2 N protein oligomerisation, where dimeric, trimeric and tetrameric species have been observed, and such formations are concentration and nucleic acid binding dependent (Chang, C. et al., 2014; Cubuk, Incicco et al., 2024; Sato et al., 2023; Zeng, W. et al., 2020). However, due to the limitations of the technique mentioned above, the presence of high-order oligomers usually formed when in the presence of nucleic acid (Farci et al., 2025; Zhao et al., 2024), as expected here ([Section: 4.2.3](#)), cannot be confirmed nor excluded. However, it is important to note that the protein-nucleic acid interactions affect the oligomeric state of the protein, and such interactions may not be preserved under native PAGE conditions, and if it is preserved, this may further affect the migration pattern observed (Wittig et al., 2006).

The addition of 7 M urea ([Figure 4. 9](#) Lane 3) leads to a reduction in disappearance of both monomeric and tetrameric structures. Notably, the dimeric form persists, albeit with an arguably reduced intensity. Together, these findings may be attributed to a combination of factors. The absence of higher oligomeric structures is explained by the dissociation of complexes into smaller units by the urea-induced disruption of hydrophobic interactions and hydrogen bonds. This would similarly explain the slight reduction in the dimeric form; however, under these terms it would be expected that the protein would be in monomeric form exclusively, which is not the case here. In fact, the monomeric form is also absent. The absence of the monomeric form, as well as the slight reduction in dimeric presence suggests two things. First, the urea-induced denaturation may have led to aggregation or precipitation of the protein, rendering it unable to enter the gel matrix effectively, resulting in the protein loss observed here. However, this is unlikely since the fluorescence data ([Section: 4.3](#)) displays exposed tryptophan residues under these conditions. In addition, it suggests that these oligomeric structures are more susceptible to the urea-induced dissociation, allowing the persistent presence of the dimeric/trimeric form. The selective stability of the dimeric form has previously been reported and is attributed to the structural features of the N protein's CTD, responsible for dimerization (Cubuk, Incicco et al., 2024; Luo et al., 2004; Yang, Mei et al., 2021; Zhou, R. et al., 2020). The dimerisation occurs through extensive interactions that include a hydrophobic core (Residues Ile304, Phe314, and Phe315) and a network of hydrogen bonds (created by residues Arg277

and Asn285). These interactions result in a highly stable dimer that is more resistant to denaturation by chaotropic agents like urea.

Moreover, alongside the nucleic-acid binding data detailed in [Section 4.2.3](#), these quaternary structural observations confirm that the isolated W-NP maintains its native oligomeric assembly and functional integrity.

4.4 Effect of Alkaloids on W-NP Structure

4.4.1 Far-UV CD

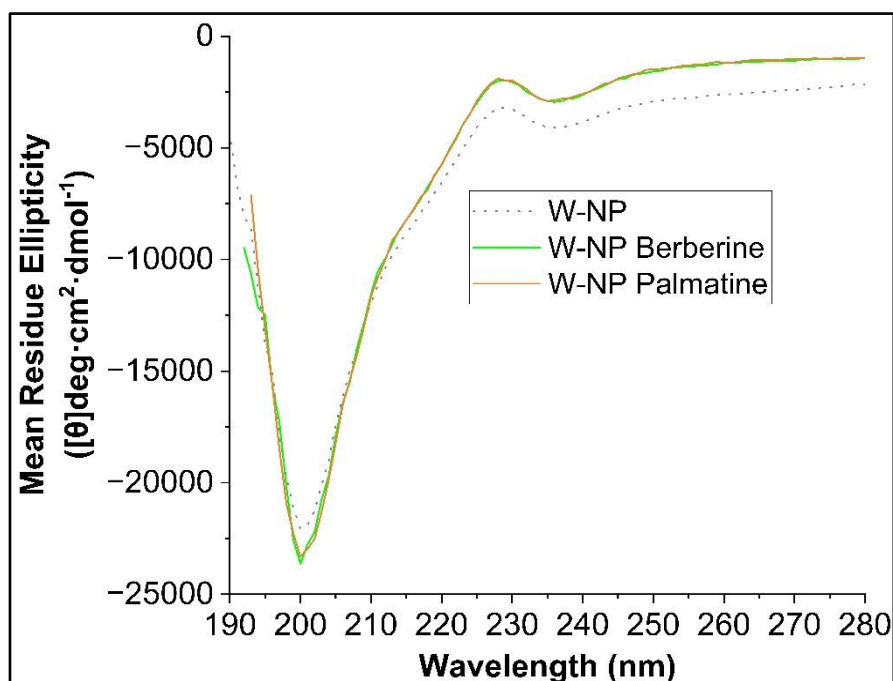
Analysis of the secondary structure of W-NP when in the presence of either compound (1:40 molar ratio) is shown in [Figure 4. 10](#). When exposed to compound, the protein exhibits similar features but with different intensities, including the strong negative trough at 200 nm alongside the positive band at ± 229 nm (229 nm for W-NP & W-NP-berberine; 228 nm for W-NP-palmatine) and trough at ± 235 nm (235 nm for W-NP and W-NP-palmatine; 236 nm for W-NP-berberine).

Deconvolution of the CD data BeStSel algorithm ([Table 4. 4](#)) reveals the binding of both compounds causes helical, turn, and disordered content transition to β -sheet content. However, subtle differences in enhanced resolution suggest different mechanisms. Specifically, berberine's interaction promotes a stronger shift from helix and disordered content towards left-handed and relaxed (slightly right-twisted) antiparallel β -sheets. In contrast, palmatine drives a more pronounced transition from turn structures into right-twisted antiparallel β -sheets, indicating a more targeted rearrangement of loop regions. These secondary structural motifs predominantly localize within the NTD, CTD, and LKR, suggesting that compound binding occurs within these functional regions or result in conformational changes in these regions and may help elucidate specific ligand interaction sites. If the alkaloid disrupts nucleic acid binding, free nucleic acid may contribute to overlapping Far-UV CD spectra, complicating the accurate interpretation of the structural changes in this case. Therefore, distinguishing direct compound-induced conformational shifts from nucleic acid-associated spectral effects requires careful control experiments and spectral deconvolution to accurately assign binding-induced alterations.

Table 4. 4: CD Deconvolution

This table compares the various secondary structural content of 5 μM W-NP in the absence and presence of 200 μM compound, measured using CD then quantified using BeStSel.

Estimated secondary structure content (%)	W-NP	W-NP Berberine	W-NP Palmatine
Helix1 ([Regular α -helix])	0	0	0
Helix2 [Distorted α -helix (2-2 residues at ends of α -helices)]	7.2	5.9	6.3
Anti1 [Antiparallel β -sheet (left-hand twisted)]	0	1.3	0.7
Anti2 [Antiparallel β -sheet (relaxed, slightly right-hand twisted)]	0.3	0.7	0.3
Anti3 [Antiparallel β -sheet (right-hand twisted)]	20.6	24	24.7
Parallel [β -sheet]	0	0	0
Turn	14.1	12.8	12.5
Others [310-helix, π -helix, β -bridge, bend, loop/irregular and invisible regions of the structure]	57.7	55.2	55.6
Helix	7.2	5.9	6.3
Antiparallel	20.9	26	25.6
Parallel	0	0	0
Turn	14.1	12.8	12.8
Others	57.7	55.2	55.6
Spectral deviation			
RMSD	0.455	0.3185	0.3058
NRMSD	0.079	0.046	0.0446

**Figure 4. 10: Effect of compounds on W-NP CD**

This figure compares the CD spectra of 5 μM W-NP (grey dot) when exposed to 200 μM berberine (Green) and palmatine (Orange), showing the wavelength in nanometres (nm) on the x-axis and the mean residue ellipticity (MRE) on the y-axis.

Fit quality metrics were excellent for all three spectra, with RMSD values of 0.4552 (W-NP), 0.3185 (W-NP-berberine) and 0.3058 (W-NP-palmatine), and normalised RMSD (NRMSD) below 0.08 in each case. The reduction in NRMSD upon ligand binding indicates that the spectra became both stronger in signal and more consistent with the BeStSel reference set, reflecting more pronounced β -fold features.

Collectively, this data demonstrates that both berberine and palmatine promote modest stabilisation and/or reorganisation antiparallel β -sheet elements within the SARS-CoV-2 nucleocapsid protein (W-NP), at the expense of distorted helical, turn, and unordered content. Such ligand-induced shifts may influence RNA-binding affinity or oligomerization and thus have potential implications for disrupting viral assembly.

4.4.2 ATR-FTIR

The ATR-FTIR spectra (4000–500 cm^{-1}) revealed subtle but reproducible differences across all conditions (**Figure 4. 11 A**), however, correction by subtraction of buffer- and compound-only spectra highlighted significant protein-associated changes in regions linked to hydration, side-chain interactions, and backbone conformation (**Figure 4. 11 B**). Proteins produce a signal in the high wavenumber region (3894–3703 cm^{-1}) associated with free terminal –OH and N–H groups (Chittur, 1998; Stuart, Barbara, 2006; Stuart, Barbara H., 1997; Surewicz et al., 1993; Van Straaten & Peppas, 1991). Changes here indicate that berberine, palmatine and urea modify or disrupt these, altering the protein's hydration shell.

The addition of either berberine or palmatine alters the signal produced between 2981–2747 cm^{-1} associated with the C–H vibrations within methyl (–CH₃) and methylene (–CH₂) groups that suggest alteration in the hydrophobic environment (Chittur, 1998; Stuart, Barbara, 2006; Stuart, Barbara H., 1997; Surewicz et al., 1993), potentially due to binding interactions that result in conformational changes, particularly in regions rich in aliphatic amino acids.

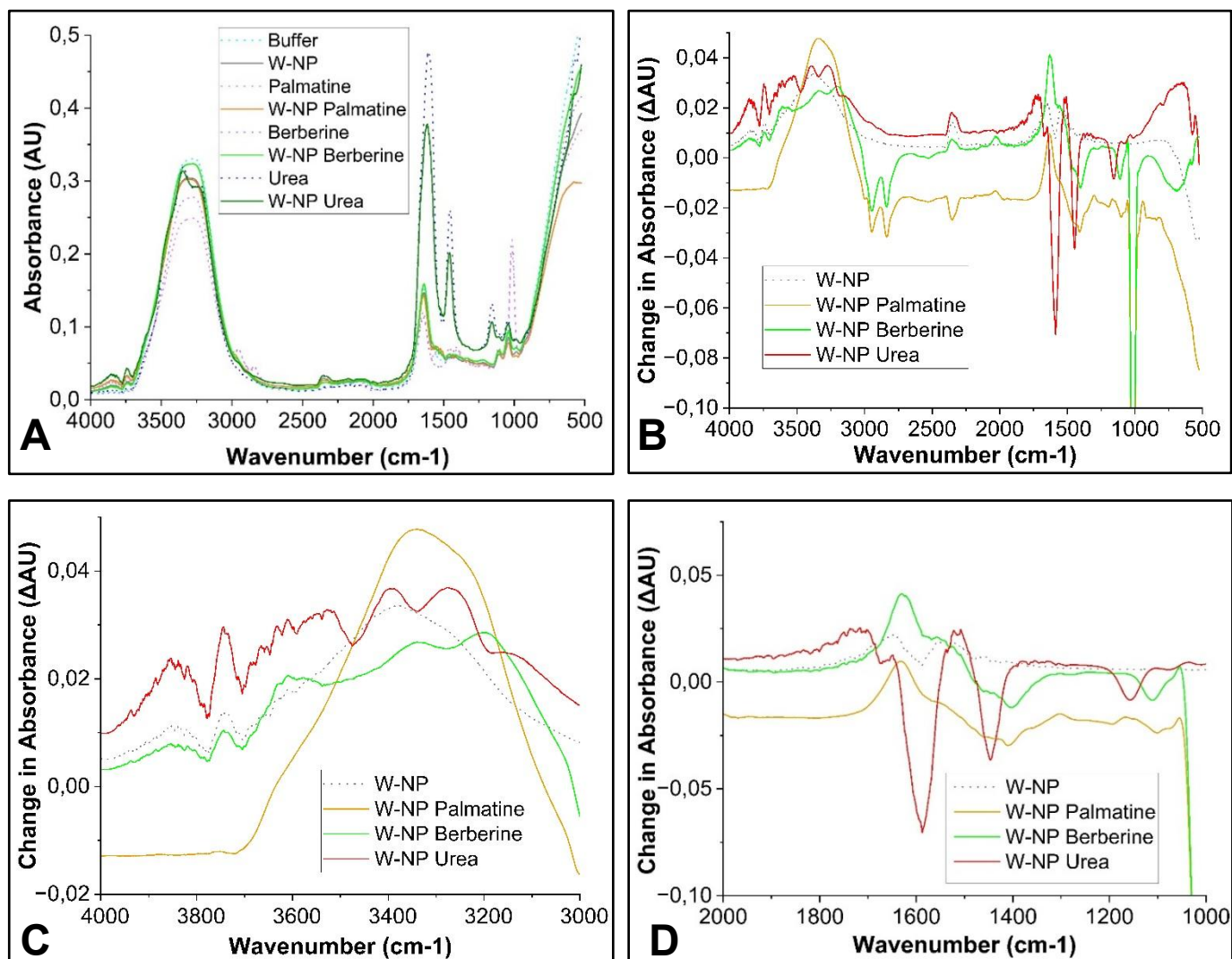


Figure 4. 11: ATR-FTIR

This figure compares the ATR-FTIR spectra of 5 μ M W-NP (grey dot) when exposed to 200 μ M berberine (Green), palmatine (Orange), or 7 M Urea (red) showing the wavenumber in reciprocal centimetres (cm⁻¹) on the x-axis and the Absorbance or Change in absorbance (A or Δ A) on the y-axis. A shows the raw spectra (including buffer blanks), B shows the corrected spectra, whilst C and D highlight specific regions.

These findings are supported by the significant variation seen in the Amide A ($\sim 3500\text{cm}^{-1}$) and Amide B ($\sim 3100\text{cm}^{-1}$) bands between $3750\text{-}2750\text{cm}^{-1}$ (Chittur, 1998; Stuart, Barbara, 2006; Stuart, Barbara H., 1997; Surewicz et al., 1993) (**Figure 4. 11 C**).

The signal here originates from O-H and N-H stretching vibrations that suggest alterations of the hydrogen bonding patterns and secondary structure upon either compound addition. Here, it would appear that berberine induces changes almost similar to urea, whilst palmatine completely alters the spectra.

Variation in signal within carbonyl, amide I and amide II regions ($1800\text{-}1300\text{ cm}^{-1}$; **Figure 4. 11 D**) is governed by stretching vibrations of C=O, C-N, C-C, and N-H bending (Chittur, 1998; Stuart, Barbara, 2006; Stuart, Barbara H., 1997; Surewicz et al., 1993), further supporting the changes brought on within the secondary structure of the protein. However, this time the two compounds alter the spectra similarly, whilst urea behaves uniquely.

Deconvolution of amide band I may reveal the exact band positions as determined by the backbone conformation and hydrogen bonding pattern (Chittur, 1998; Stuart, Barbara, 2006; Stuart, Barbara H., 1997; Surewicz et al., 1993), however, this is complicated by the intrinsically disordered nature of W-NP as well as the potential overlapping contributions in the event that nucleic acid may be present (**Section: 4.2.3**).

However, the results provide evidence for similarities and differences with regards to the binding mechanism of berberine and palmatine with the N protein.

4.4.3 BN-PAGE

The addition of either compound (**Figure 4. 12** Lanes 3 & 4) does not produce any major observable shift in oligomeric distribution, suggesting that neither compound dramatically alters the quaternary structure under the conditions investigated. However, the bands representing tetrameric structures appear to be enhanced upon addition of compounds and somewhat better retained in the presence of urea (**Figure**

4. 12 lanes 7 & 8); however, further investigations are required. It is important to note that the ligand-protein interactions may not be preserved under native PAGE conditions, since these small molecules may dissociate during the electrophoresis depending on their kinetic stability (Wittig et al., 2006). In addition, this consideration extends to those taken in [Section: 4.2.3](#), suggesting that nucleic acid is present within our sample.

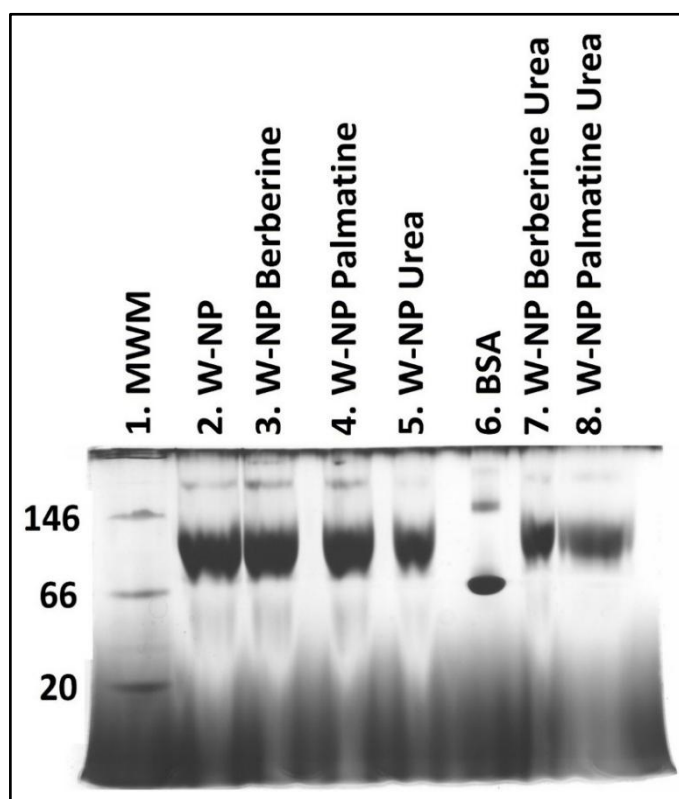


Figure 4. 12: BN-PAGE with Compound

This figure depicts the migration profile of 5 μ M W-NP in the presence/absence of 7 M urea, or 200 μ M compound (berberine or palmatine) or both, as well as that of BSA using a 12% Blue-Native Polyacrylamide Gel. Lane 1 displays the Molecular Weight Marker (MWM) with its corresponding sizes in kilodaltons (kDa).

The dissociation of nucleic acid under native PAGE conditions may alter the oligomeric state of the protein; however, its association may in turn affect the migration pattern observed.

4.4.4 Thermal Stability

The denaturation of W-NP was monitored using Far-UV-CD and intrinsic tryptophan fluorescence ([Section 10: Appendix 10. 8](#)). Analysis of emission of the Far-UV-CD at 228 nm and 235 nm, as well as intrinsic tryptophan fluorescence at 335 nm, shown in [Figure 4. 13](#), reveal insight into the secondary and tertiary structural effects of the thermal denaturation of this protein in the absence and presence of compound.

Both CD wavelengths (**Figure 4. 13 A1 & B1**) exhibit classical sigmoidal melting profiles with their largest transitions between 40 – 50 °C. Boltzmann fit using Origin Lab (**Table 4. 5**) determined that unliganded W-NP has a melting temperature (T_m) of ~47.27 °C. These values are lower than determined by Nguyen et al. (2024), who estimated the T_m to be 49.6 °C using DSF.

The same research group hypothesized that the R203K/G204R mutations increase protein stability, whilst the P13L; 31-33ERS destabilise the protein; however, no results have been reported for the combination of these mutations. The addition of either alkaloid appears to cause a small yet measurable increases in T_m (~0.08 – 0.1 °C) that suggest slight stabilization of the secondary structure of W-NP. The CD at 235 nm for berberine shows notable maintenance of structure between 35 - 45 °C (**Section 10: Appendix 10. 8**), suggesting enhanced stabilization of a certain secondary element within this temperature window, potentially attributed to stabilisation of nucleic acid.

The tryptophan fluorescence melting curve (at 335 nm) reveals a potential tertiary destabilizing effect characterized by a reduced fluorescent intensity and thermal midpoint in the presence of ligand (**Figure 4. 13 C1**). However, due to larger fitting errors, quantitative T_m values were not extracted (**Table 4. 5**).

Analysis of the reversibility of protein unfolding reveals further significant insights into protein and protein-compound behaviour. The 228 nm CD signal shows an increased refolding, or larger reversibility, after cooling with both compounds present (**Figure 4. 13 A2**); however, this effect was not observed for CD at 235 nm (**Figure 4. 13 B2 & C2**) nor fluorescence.

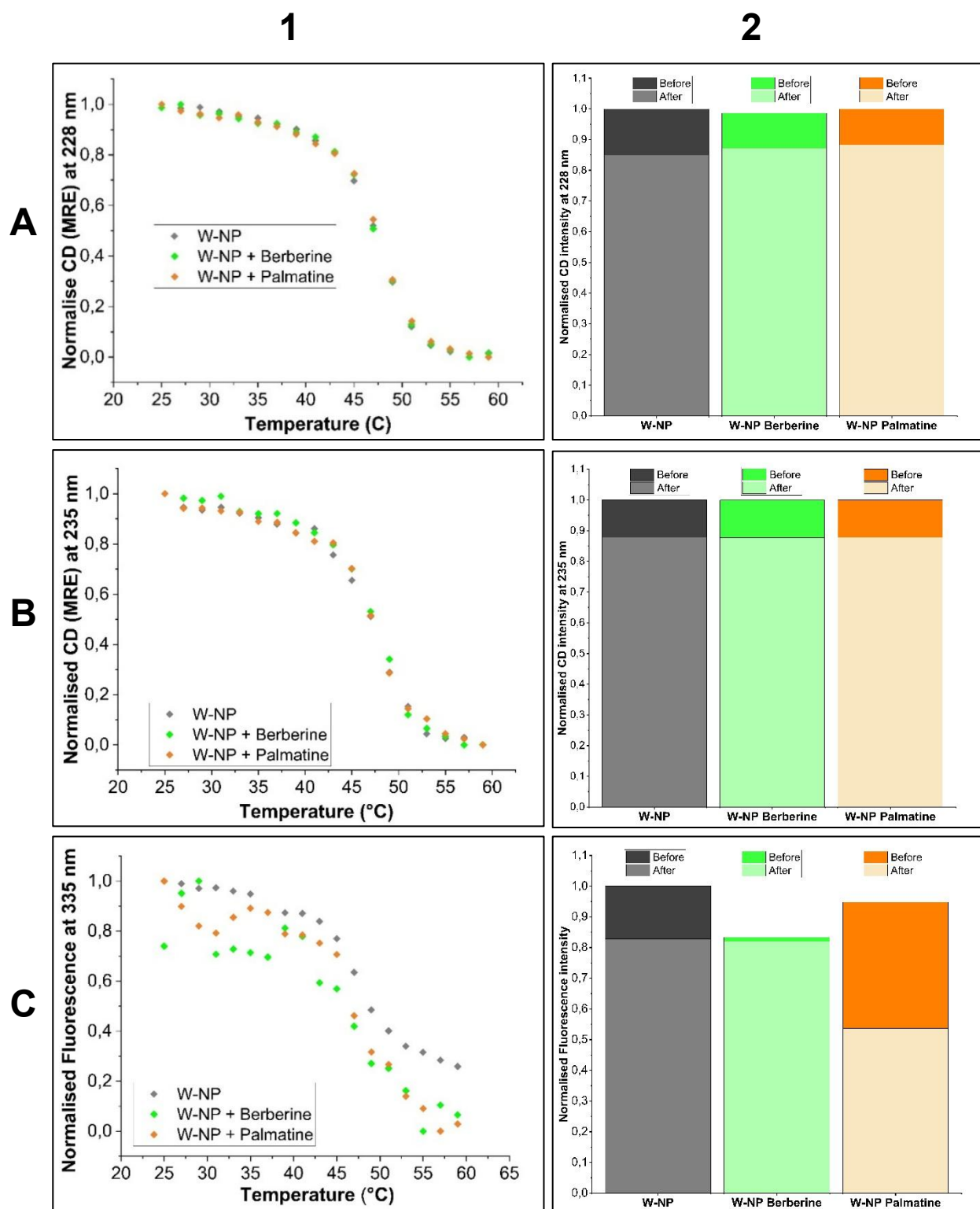


Figure 4. 13: Thermal Structural Stability

This figure depicts the thermal denaturation of W-NP with or without 200 μ M compound. Panels 1 show the normalised CD (MRE) and Fluorescence values over the temperature. A and B shows the CD values at 228 nm and 235 nm, C shows the fluorescence at 335 nm. Panel 2 shows the corresponding return after denaturation. W-NP only is shown in grey, W-NP with berberine is shown in green, and W-NP with palmatine is shown in orange.

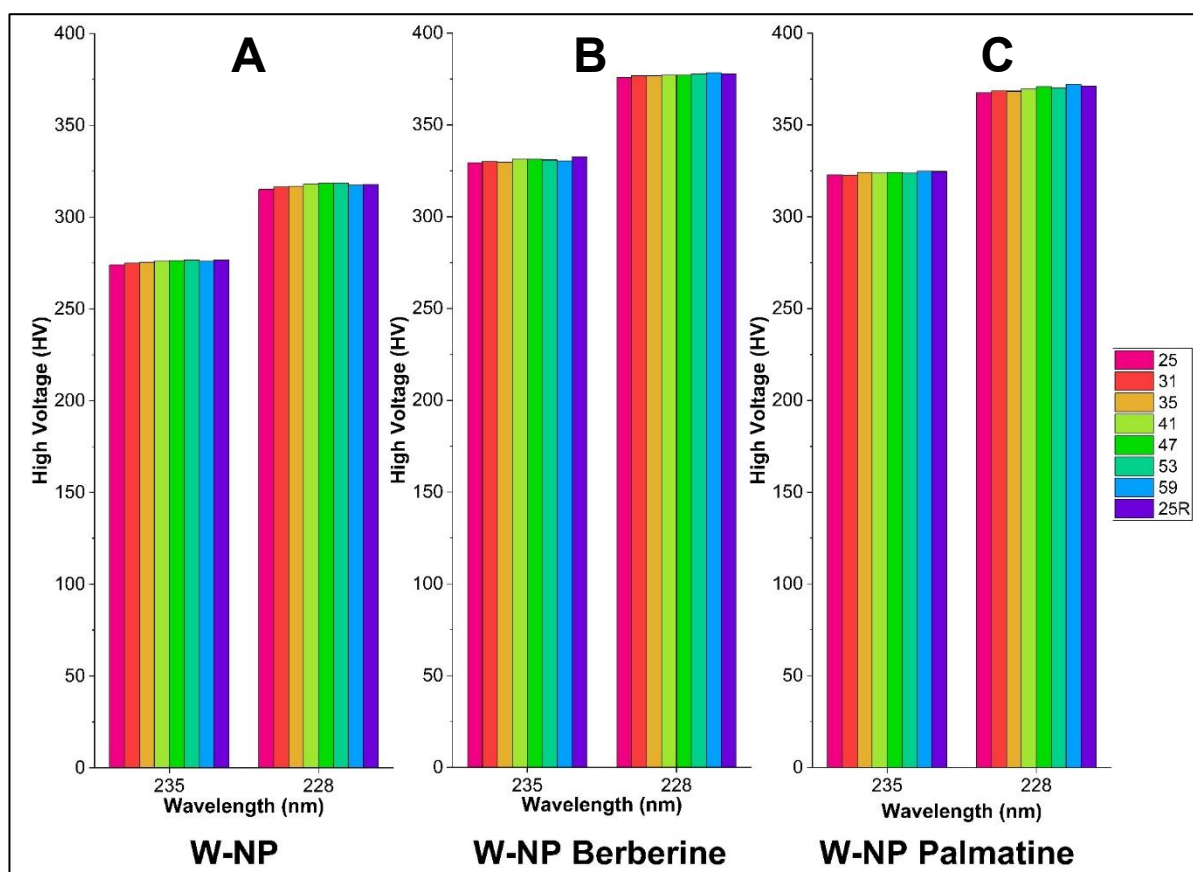


Figure 4. 14: CD High Voltage

This figure compares the CD High Voltage (HV) intensities for the various samples at the various wavelengths across the different temperatures.

Table 4. 5: The Effect of Compound on W-NP Melting Temperatures (T_m)

This table compares the effects of 200 μ M berberine and palmatine on the thermal melting temperatures extrapolated from the application of a Boltzmann fit to the CD data using OriginLab. The thermal mid point is shown in °C and the error is shown in brackets.

	W-NP	W-NP + Berberine	W-NP + Palmatine
CD 228 nm	47.19 (0.19)	47.29 (0.18)	47.49 (0.22)
CD 235 nm	47.35 (0.287)	47.45 (0.25)	47.20 (0.26)
CD Average	47.27	47.37	47.35
Fluorescence	47.24 (0.019)	47.05 (1.2)	47.92 (0.60)

Analysis of the High Voltage (HV) patterns (**Figure 4. 14**) shows that, although the HV does increase slightly with increasing temperature, it never exceeds the 600 HV threshold. This confirms that the CD data were not compromised by excessive light scattering or sample aggregation, and that the spectra can be interpreted reliably.

Such mixed structural effects have been seen in albumins and haemoglobin, where these alkaloids caused global destabilization despite binding induced rigidity (Bhadra & Kumar, 2011; Hazra et al., 2013; Islam et al., 2008; Khan & Suresh Kumar, 2015).

4.5. Mechanisms of Interaction and Binding Parameters

4.5.1 Absorbance Profiling of Experimental Components

Both berberine and palmatine exhibit characteristic UV/Vis absorbance maxima attributable to π - π electronic transitions within their extended conjugated systems. These transitions are further influenced by the presence of auxochromic substituents and heteroatoms, which modulate the electronic distribution and contribute to the observed spectral features seen in **Figure 4. 15 A** (Battu et al., 2010; Moon et al., 2021; Purwaningsih et al., 2024; Zanwar et al., 2022).

Both compounds used display varying absorbance throughout both the UV (100 – 400 nm) as well as visible range (400-700 nm), creating three distinct peaks (200-300nm; 300-385nm; 375-475nm) with maxima's at 248 nm, 331 nm (palmatine and 337 nm

(berberine) and again at 420 nm (**Figure 4. 15 A**). As expected, when in solution with protein at the same concentration, although at a much higher ratio to that of protein, whether bound or not – simply in solution, the compounds' absorbance overlaps and perturbs the absorbance of protein aromatic side-chain signals at 280 nm, causing baseline shifts and hyperchromic effects in the absorbance of protein (**Figure 4. 15 B**).

The absorbance spectra of the compounds overlap both the excitation (295 nm) and emission (300 - 400 nm) wavelengths used in the fluorescence studies of W-NP, suggesting that both primary and secondary inner filter effects (pIFE and sIFE) will be present when investigating protein compound interactions using the fluorescence technique (**Section: 4.5**) (Panigrahi & Mishra, 2019). This excessive absorbance ($A_{295} \approx 0.61$ AU; $A_{336} \approx 0.74$ AU), both pIFE (attenuation of excitation) and sIFE (reabsorption of emission) can together suppress and skew the fluorescence emission drastically.

Accurate determination of binding constants and quenching parameters requires corrections for compound absorbance and inner-filter effects, commonly applied via the Lakowics (2006) model:

$$F_{corr} = 10^{\frac{A_{ex} + A_{em}}{2}}$$

Where the fluorescence (F_{obs}) is corrected for using the excitation (A_{ex}) and emission (A_{em}). These corrections were performed to assess the fluorescence properties of protein and protein-compound interactions.

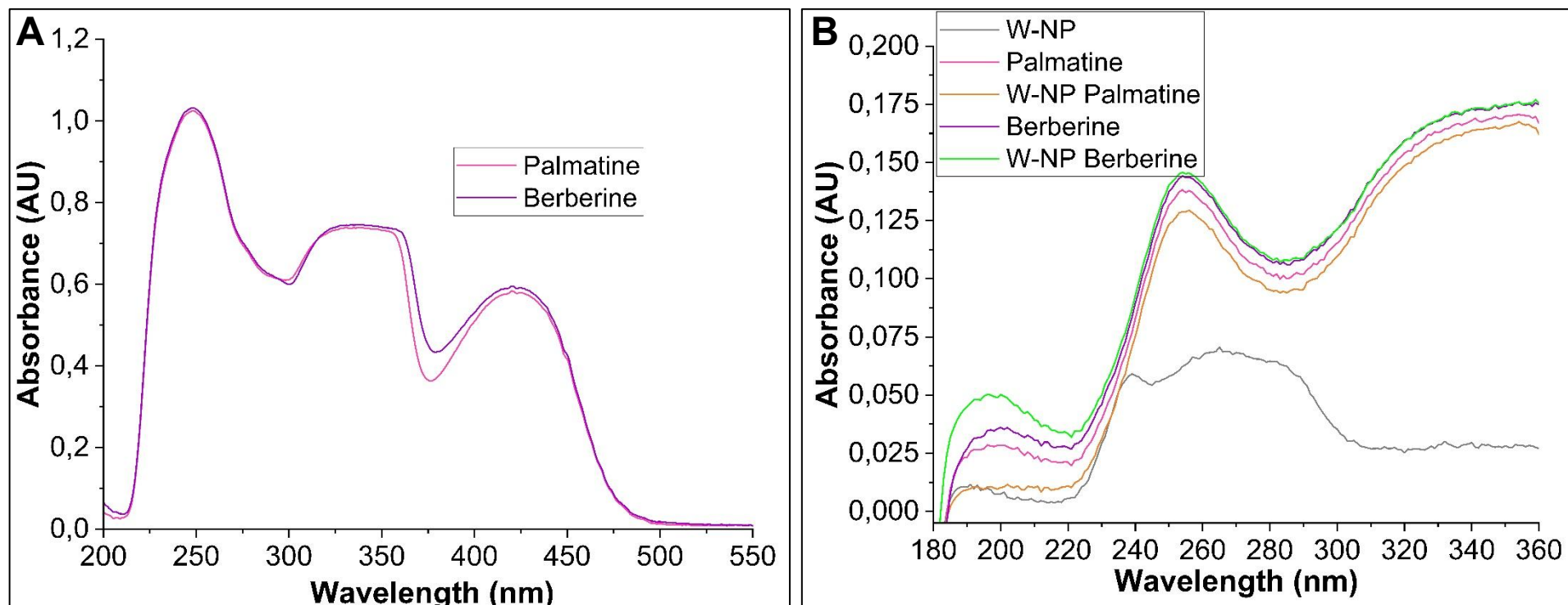


Figure 4. 15: UV-Visible Absorbance of Compounds and Complexes

These figures depict the absorbance spectra of 200 μ M compound only (A) in contrast to 5 μ M protein in the presence and absence of 200 μ M compound (B). Figure A shows a wider spectrum in comparison to Figure B. In both figures, W-NP is shown in grey, W-NP with berberine is shown in green, W-NP with palmatine is shown in orange, berberine only is shown in purple and palmatine only is shown in pink.

4.5.2 ANS Binding Studies

The hydrophobicity of W-NP, studied through the fluorescence when excited at 380 nm upon addition of ANS, was described in [Section 4.3](#). The effects that compounds have on the hydrophobicity fluorescence of the protein are displayed in [Figure 4. 16 A](#).

The addition of either compound does not alter the peak maxima (peak between 430 – 550 nm with λ_{max} 485 nm) of the signal but appears to quench the fluorescence signal. Similarly, the compound does not affect the peak maxima (λ_{max} 505 nm) in the presence of 7 M urea, but again the fluorescence is quenched. This indicates that compound binding does not alter the environment of ANS, and therefore, does not alter the hydrophobic environment of the protein, nor does it protect from unfolding. The sustained quenching is explained by the pIFE, sIFE, or dynamic quenching.

4.5.3 Fluorescence Quenching

The fluorescence spectra of W-NP were described in [Section 4.3](#), resulting in a peak between 300 – 450 nm with a λ_{max} at 336 nm, shifted to λ_{max} 356 nm upon urea addition. These spectra are significantly different in the presence of 200 μM compound ([Figure 4. 16 B & D](#)). The fluorescence spectra of compound only ([Figure 4. 16 C](#)) does not produce any signal.

The slow addition of either compound ([Figure 4. 17 A & B](#)) causes a subtle reduction in fluorescence intensity, accompanied by subtle blue shifts in emission of W-NP, whereas higher concentrations (75 – 400 μM) of either compound significantly alters the spectrum. These higher concentrations produce two distinct peaks with maximums at 332 nm as well as 360 nm (palmatine) and 328 nm and 363 nm (berberine). Of note here is the subtle difference in intensity as well as peak maxima produced by the two different compounds, supporting the hypothesis of distinct mechanisms of interaction occurring between the alkaloids ([Section 4.4](#)).

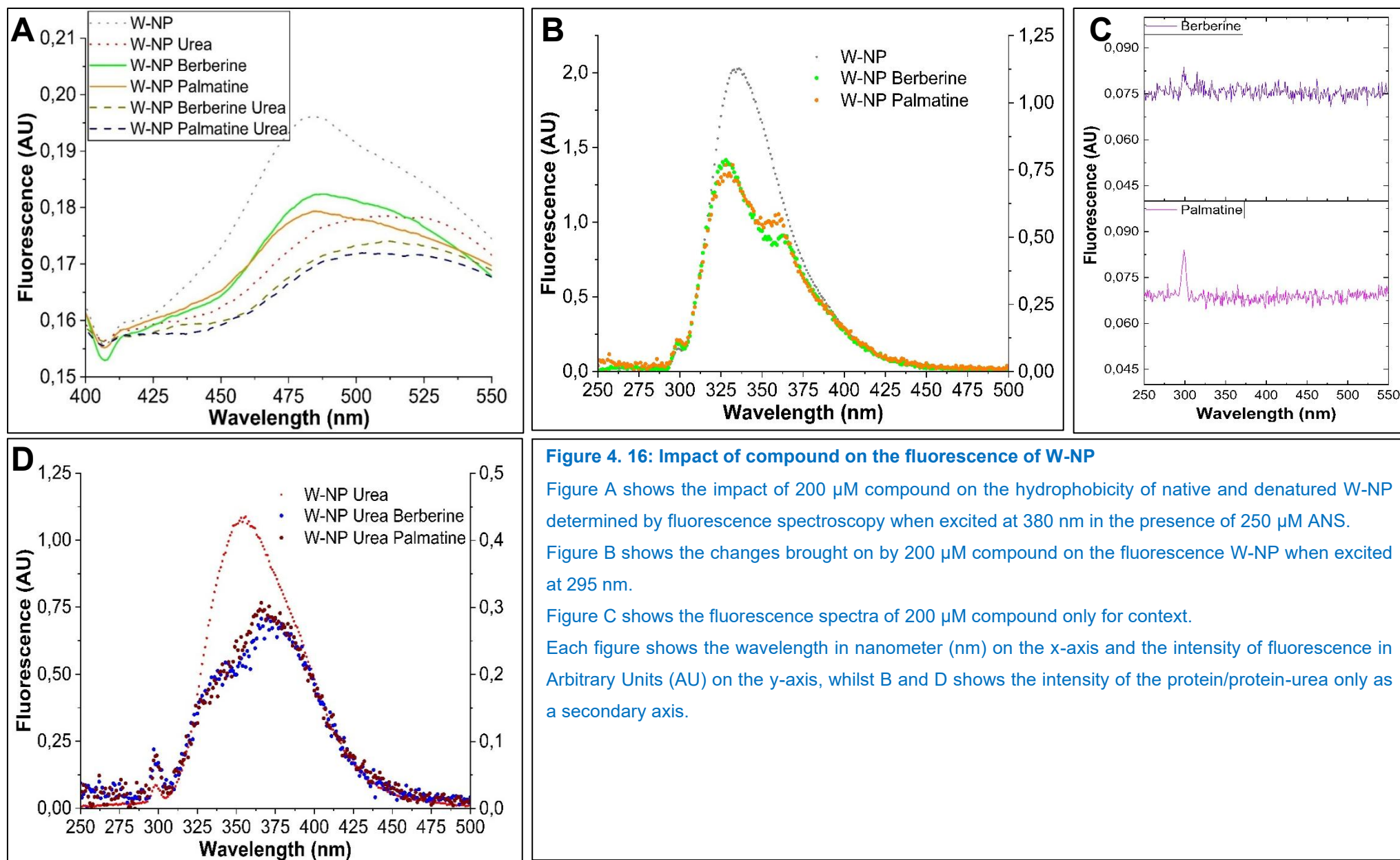


Figure 4. 16: Impact of compound on the fluorescence of W-NP

Figure A shows the impact of 200 μM compound on the hydrophobicity of native and denatured W-NP determined by fluorescence spectroscopy when excited at 380 nm in the presence of 250 μM ANS.

Figure B shows the changes brought on by 200 μM compound on the fluorescence W-NP when excited at 295 nm.

Figure C shows the fluorescence spectra of 200 μM compound only for context.

Each figure shows the wavelength in nanometer (nm) on the x-axis and the intensity of fluorescence in Arbitrary Units (AU) on the y-axis, whilst B and D shows the intensity of the protein/protein-urea only as a secondary axis.

The early changes in the emission spectra could be attributed to subtle changes occurring within the tertiary structure of the protein upon compound binding, resulting in the modulation of the efficiency of energy transfer within the protein that impacts the fluorescence intensity, during which some tryptophan residues may become more buried within the three-dimensional structure (Calhoun et al., 1986; Vivian & Callis, 2001; Yammine et al., 2019). However, these changes could also be attributed to quenching of fluorescence, or the reduction in fluorescence emission, caused by the addition of compounds, which could occur through a variety of mechanisms.

Some strategies exist to differentiate the quenching mechanism taking place (Calhoun et al., 1986; Yammine et al., 2019). Fluorescence quenching may occur due to optical artifacts, where compounds absorb either excitation and/or emission light, known as the Inner-Filter Effect (IFE) mentioned in (Section 4.5). Since the absorbance of the compound at working conditions was larger than 0.1, IFE was assumed and corrected for mathematically; however, these corrections did not significantly alter the results. Fluorescence quenching may also occur due to Förster Resonance Energy Transfer (FRET), where excited tryptophan residues transfer their energy to the compound (Calhoun et al., 1986).

FRET can be identified through fluorescent lifetime measurements whereby the fluorescence lifetime of protein decreases, such that the emission of Trp decreases whilst that of compounds increases (Lakowics, 2006).

However, the potential of FRET is dependent on both distance and orientation of residue to compound, factors that are complicated by the structural changes brought on upon nucleic acid binding and dimerization, particularly when considering the presence of multiple oligomeric species within the sample.

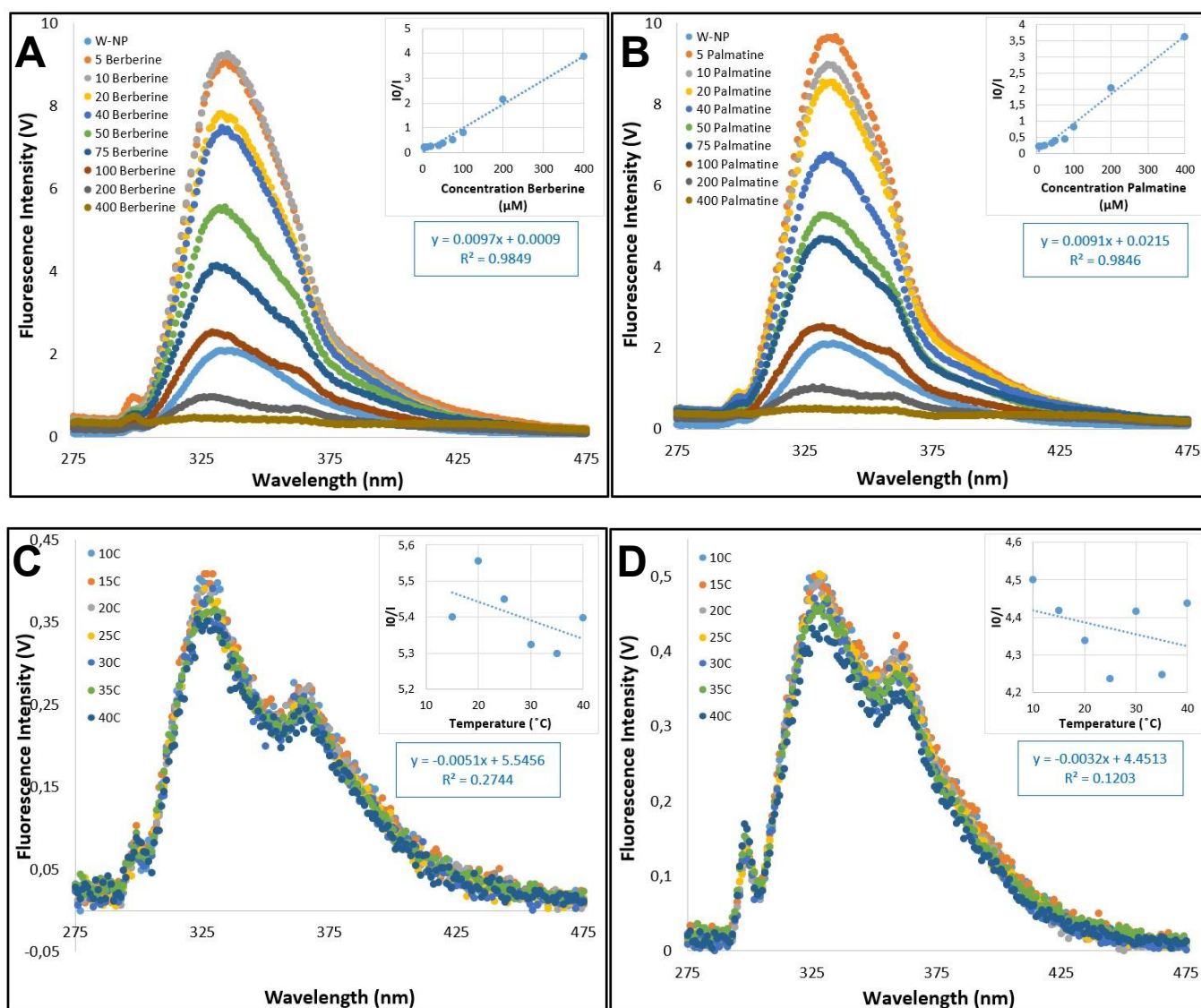


Figure 4. 17: Stern-Volmer Analysis of Fluorescence Quenching Induced by Berberine and Palmatine

These figures display the quenching effect caused by berberine (A and C) and palmatine (B and D) on the fluorescence spectra of W-NP as well as the associated Stern-Volmer response obtained from the concentration-dependent quenching (A and B) and thermal quenching (C and D) assays. Each figure displays the fluorescence intensity emitted (y-axis) per wavelength (x-axis).

Quenching of fluorescence could be attributed to either dynamic or static quenching, during which fluorescence intensity is lost due to collisions with the quencher or the formation of non-fluorescent complexes, respectively (Eftink & Ghiron, 1981). Dynamic and static quenching could be differentiated by Stern-Volmer analysis, resulting in linear and non-linear behaviour, respectively. Static quenching may, in addition, be identified through UV-Vis absorption; however, due to the absorptive nature of the compounds (Section 4.5), this was not possible. However, it is important to note that each of the quenching mechanisms may be affected by the intrinsically disordered nature of the protein, since Trp residues may be free to interact with the compound without any rigid structure to shield them, as well as the structural changes brought on upon nucleic acid binding/dimerization (Schramm et al., 2019).

Concentration-dependent Stern-Volmer analysis (Figure 4. 17 A & B) reveals non-linear behaviour for both compounds. The Stern-Volmer quenching constant (K), or slope of the line, produced by berberine ($0.0431 \mu\text{M}^{-1}$ or $43\,100 \text{ M}^{-1}$) and palmatine ($0.0469 \mu\text{M}^{-1}$ or $46\,900 \text{ M}^{-1}$) suggests that the quenching taking place is highly efficient, perhaps such that the fluorophore in question is well accessible to the quencher.

A downward deviation (toward the x-axis) of the Stern-Volmer plot could be explained by the presence of both buried and exposed tryptophan residues, such that some are more accessible to the quencher whilst others are not (Lakowics, 2006). Instead, an upward deviation is seen, suggesting that the quenching behaviour deviates from a simple dynamic mechanism, such that both dynamic and static quenching are taking place (Lakowics, 2006).

The hypothesis that dynamic quenching is taking place is corroborated by the spectra of protein-compound mixtures in the presence of urea (Figure 4. 16 D). These spectra are similarly quenched but have lost the peak at 328 nm (potentially representing buried tryptophan residues). It is further supported by the hydrophobic fluorescence studies, similarly, resulting in a quenching of fluorescence without significant alteration of the ANS environment (Figure 4. 16 A). Both observations indicate that the compounds primarily

affect the fluorescence by acting as dynamic quenchers, whilst not being able to protect the protein's native structure from urea-induced unfolding.

In turn, the negative slope obtained upon thermal quenching analysis (**Figure 4. 17 C & D**) supports the presence of static quenching, since higher temperatures reduce complex stability and therefore quenching would also decrease.

However, both scenarios are complicated by the heterogeneity of the samples analysed. Firstly, in such an event, the concentration-dependent Stern-Volmer slope then represents an average of several quenching processes, since some fluorophores are more easily quenched in one species when compared to another. In turn, the temperature-dependent Stern-Volmer slope will similarly represent an average of the different propensities, especially if some oligomeric forms are more stable than others, or some forms may maintain their quencher binding more efficiently. Therefore, the heterogeneity of the sample may mask or skew any true quenching behaviours, complicating the accurate interpretation of the quenching mechanisms from the data.

4.5.4 Surface Plasmon Resonance

Surface Plasmon Resonance (SPR) has previously been used to investigate the SARS-CoV-2 N protein. Yang et al. (2024) identified aptamers that have nanomolar affinity with SARS-CoV-2, MERS, HCoV-OC43 and -NL64 NP whilst other research groups developed a method for detection for diagnostic purposes (Anderson et al., 2021; Fan et al., 2023; Lisyte et al., 2024; Yano et al., 2022). **Figure 4. 18** displays the sensograms collected during surface plasmon resonance (SPR) experiments, investigating the interaction between W-NP and berberine (**Figure 4. 18 A**) and palmatine (**Figure 4. 18 B**).

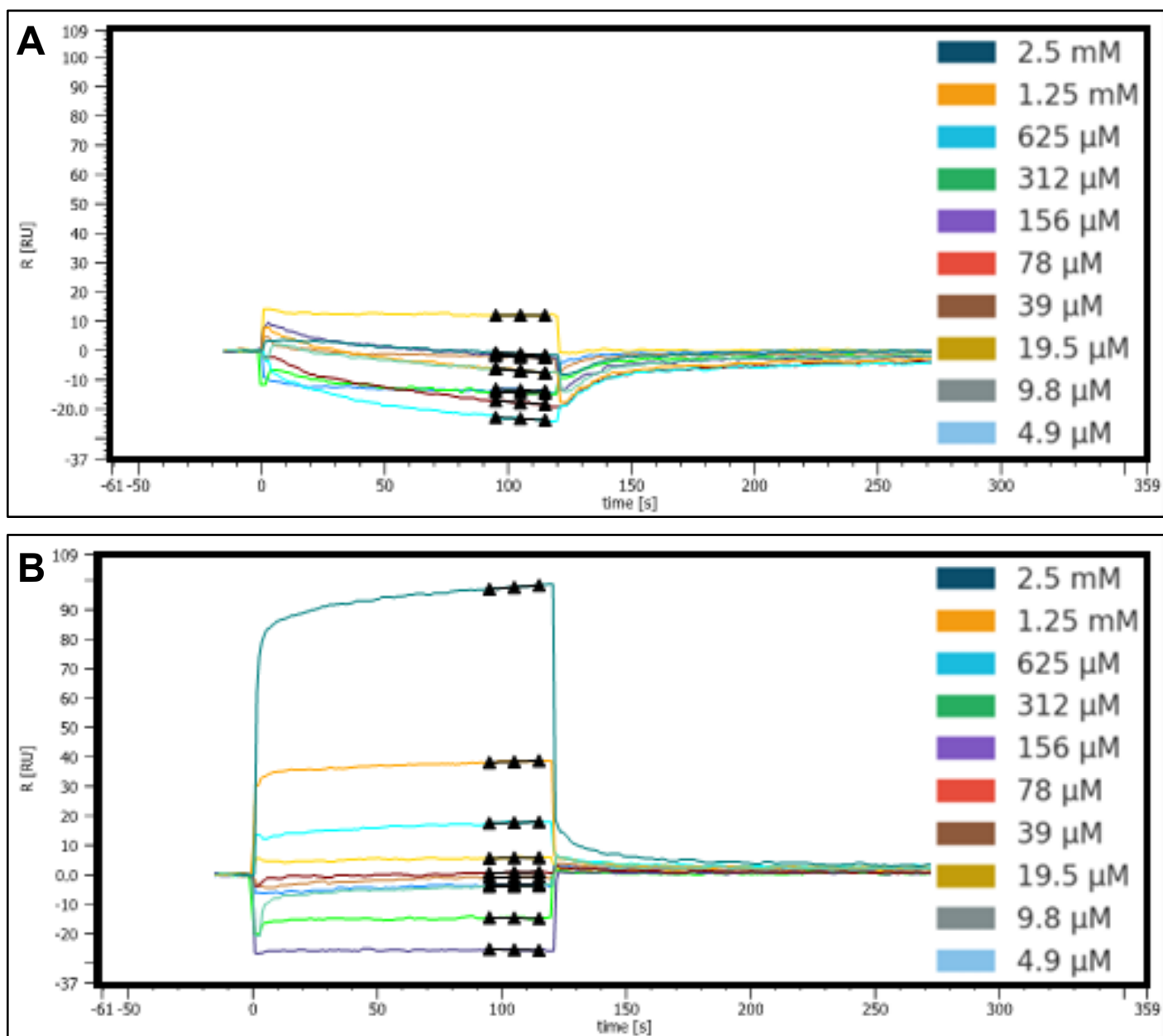


Figure 4. 18: SPR Sensograms

This figure shows the sensograms of berberine (A) and palmatine (B), overlaying the responses of analyte injection across a series of concentrations (2.5 mM to 4.88 μ M; with the highest concentration in teal). The time in seconds (s) is displayed on the x-axis and the response in Response Units (RU) on the y-axis.

For both analyses, a pre-injection baseline was observed from approximately -50 to 0 seconds, establishing a stable signal prior to analyte introduction. The injection of analyte, indicated by black triangles, occurred from 0 to approximately 100-110 seconds, marking the association phase where analyte binds to the immobilized W-NP, resulting in an increase in Response Units (RU). The rate of association is directly reflected by the steepness of the slope during this phase.

Following the cessation of analyte injection at around 100-110 seconds, the dissociation phase commenced, characterized by a decrease in RU as the buffer replaced the analyte, leading to the dissociation of bound analyte from W-NP. The rate of RU decrease during this washout period reflects the stability of the binding interaction.

The interaction between W-NP and berberine ([Figure 4. 18 A](#)) reveals minimal increase in RU, suggesting little or no response. Furthermore, the response observed lacks a clear concentration-dependent trend across the various concentrations of compound tested; specifically, the lowest concentration yielded the largest response in RU. This stochastic response suggests that, instead of specific molecular interactions occurring, the observations originate from another factor. These observations may be a result of bulk refractive index factors, during which signal changes due to differences within the refractive index between the running buffer and analyte solution (Nicoya, 2023). These observations may be due to non-specific binding, unwanted interactions between the compound and sensor surface or W-NP, often arising from hydrophobic or weak electrostatic interactions. These artifacts are avoided through surfactants within the running buffer.

Although binding was anticipated, several reasons may explain why no binding is observed between berberine and W-NP. The stock concentration of berberine was prepared at levels exceeding its solubility within DMSO; therefore, the compound may have precipitated out of solution, limiting the actual concentration of soluble compound available for binding. The method of immobilization may alter the conformation of the protein or block critical binding sites. In rare instances, some analyte-ligand interactions

result in negative RU responses; however, these instances are usually associated with significant conformational changes, which is not supported by other data in this study. Although SPR instruments typically have a lower limit of 100 Da, the low molecular weight of the compound (371.81 g/mol) may inherently render low RU responses; however, in comparison to palmatine, such changes are anticipated to still be within measurable ranges.

The interaction between W-NP and Palmatine (**Figure 4. 18 B**) yielded a low, yet observable concentration-dependent response, suggesting a specific interaction. Despite this trend, weak binding is indicated by the overall low magnitude of the RU, limiting accurate derivation of reliable affinity (K_d) values from the collected data. The slope of the association and dissociation phases suggests weak transient binding occurring between ligand and protein. A plateau, or steady-state (equilibrium), is discernible within the association phase, representing the equilibrium during which the rate of binding equals the rate of dissociation, suggesting saturation of available binding sites.

Furthermore, accurate interpretation of SPR results is limited by the multimeric species of the samples analysed, although theoretically it is more likely that the monomeric species would be immobilized to a greater extent (Frenzel et al., 2014). Furthermore, while the intrinsically disordered nature of the protein enables multifunctional interactions, it complicates the biophysical characterization of the interaction because it allows dynamic interconversion between conformations. IDPs enable binding events that do not follow the simple on/off interaction theories but may follow dynamic equilibria involving multiple bound conformations. Any disorder-to-order transitions occurring during ligand binding may contribute to the changing SPR signals. These factors lead to highly complex kinetics that deviate from the standard 1:1 Langmuir binding models often used in SPR analysis (Frenzel et al., 2014).

4.5.5 Isothermal Titration Calorimetry

Isothermal Titration Calorimetry (ITC) was used to assess the thermodynamic properties of the binding interactions between W-NP and Palmatine. Berberine was not included since the SPR data shows no binding. The N protein is known for its dynamic oligomerisation potential, as evidenced by [Figure 4. 9](#). Therefore, quantification of the proportion of each oligomeric species in solution is not possible, complicating accurate interpretation of the ITC data (Velazquez-Campoy et al., 2004; Velázquez-Campoy et al., 2004). Given this, thermodynamic parameters derived from ITC cannot be used to assign absolute binding stoichiometries; therefore, the binding parameters are reported per apparent monomeric unit, with the caveat that the true functional stoichiometry differs depending on the oligomeric state (Sparks et al., 2020). Instead, ITC serves as an indication of potential binding events, providing insights into the overall interaction dynamics rather than precise molecular ratios.

Table 4. 6: ITC experimental conditions tested

This table compares the various W-NP and ligand concentrations as well as injection volumes used during ITC, as well as the resulting thermodynamic parameters obtained from the best fit binding models.

Run	W-NP concentration (uM)	Ligand Concentration (mM)	Injection volumes	Model	Kd1 (M)	Kd2 (M)	Kd3 (M)	ΔH_1 (kJ/mol)	ΔH_2 (kJ/mol)	ΔH_3 (kJ/mol)	K_{a1} (M ⁻¹)	K_{a2} (M ⁻¹)	K_{a3} (M ⁻¹)	ΔS_1 (J/mol·K)	ΔS_2 (J/mol·K)	ΔS_3 (J/mol·K)
A	5	30	1 x 3 ul + 29 x 5 ul	Sequential Two-Site	0.0978000	0.0018250	-	- 9 994	- 5 000	-	10.22	547.90	-	-33 500.00	-16 720.00	-
B	5	20	1 x 3 ul + 29 x 7 ul	Sequential Three-Site	0.0017040	0.0999900	0.0019990	1 738	- 5 000	- 2 500	587.00	10.00	500.30	5 881.00	-16 750.00	-8 333.00
C	6.36	20	1x 5 ul + 24 x 10 ul	Sequential Two-Site	0.0465100	0.0026550	-	- 7 769	- 3 176	-	21.50	376.60	-	-25 700.00	-10 600.00	-
D	2.5	40	1 x 5 ul + 24 x 10 ul	Sequential Three-Site	0.0042210	0.0462700	0.0116600	9 855	- 5 000	- 2 499	236.90	21.61	85.76	33 100.00	-16 740.00	-8 345.00

Despite this complexity, the ITC experiments yielded consistent evidence of palmatine binding to W-NP with low affinity across all runs, although the number and nature of binding sites appeared to vary depending on the experimental conditions. Four independent runs were performed using different ligand-to-protein ratios alongside adjusted injection volumes summarised in [Table 4. 6](#). Compound-to-buffer titration is displayed in [Appendix 10. 9](#) showing little effect of compound only.

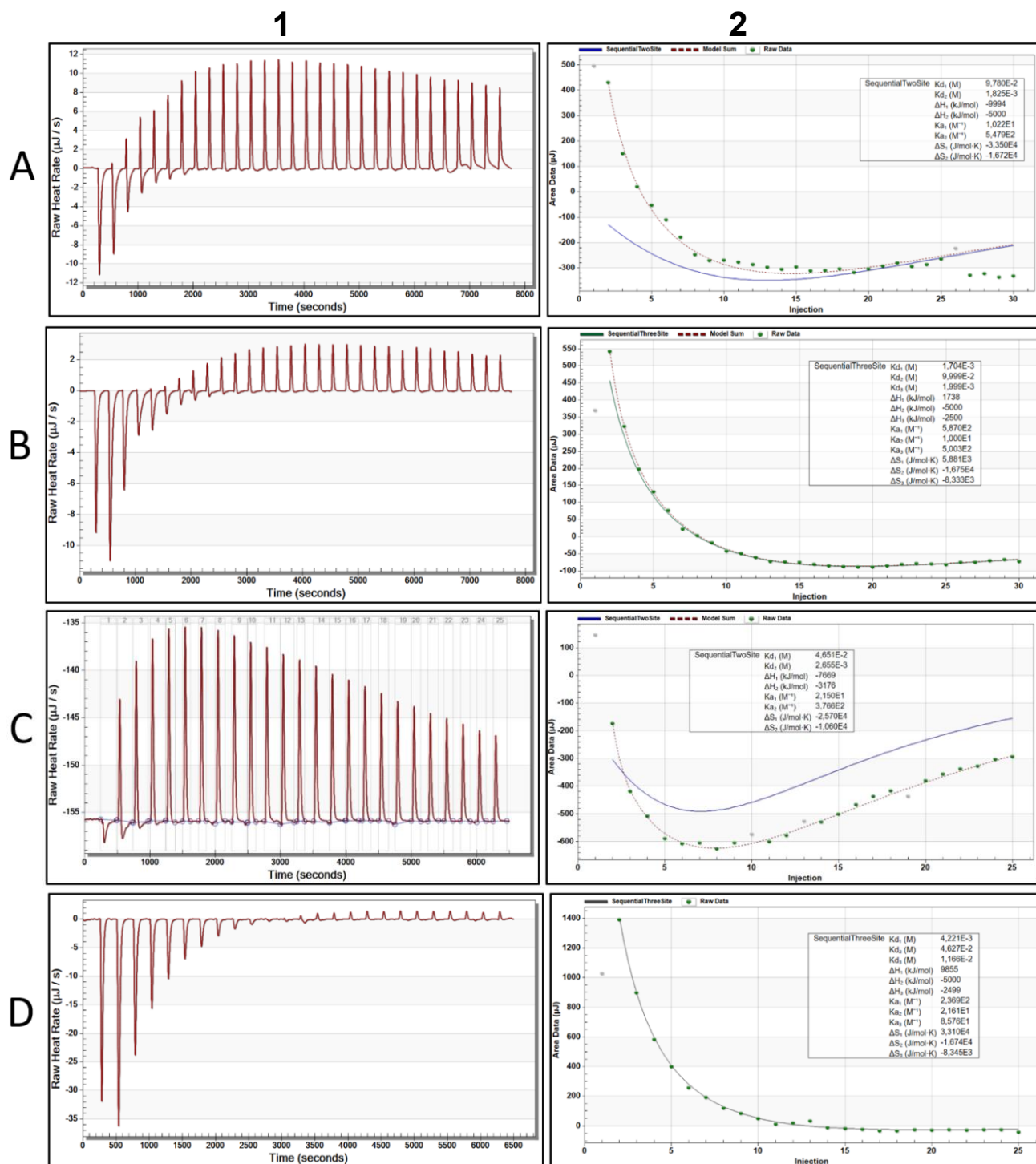


Figure 4. 19: Isothermal Titration Calorimetry Thermogram of W-NP with Palmatine

These figures show the various ITC thermograms (1) and their fits (2), depicting the interaction between W-NP and palmatine. Panels 1 show the raw thermograms, displaying the Raw Heat Rate in microjoules per second ($\mu\text{J/s}$) over time in seconds upon successive injections of compound into protein. Each downward peak corresponds to an exothermic binding event. Panels 2 show the corresponding analysed fits, displaying the integrated heat per injection (μJ) over each injection, indicative of the molar ratio of ligand to protein. Depicted are the sequential two and three site binding models and their corresponding thermodynamic parameters. Panels A, B, C, and D show the results for the various experimental conditions assessed, summarised in Table 4.5

Runs A and C, which used smaller ratios, were best fitted using a sequential two-site binding model, whereas Runs B and D, which involved higher ligand-to-protein ratios, were best described by a three-site sequential binding model. These models were selected based on statistical goodness-of-fit, but the isotherms themselves appeared relatively smooth and monophasic, without clear inflection points, an argument supported by the lack of binding within the compound to buffer titration controls ([Appendix: Appendix 10. 9](#)). This suggests that, although multi-site fitting improves the model, the underlying binding events may not represent discrete, physically distinct binding phases (Bastos & Velazquez-Campoy, 2021; Velazquez-Campoy et al., 2004; Velázquez-Campoy et al., 2004). Rather, the data may reflect overlapping, context-dependent interactions across a heterogeneous population of protein species.

The thermodynamic parameters from these models suggest that palmatine binds to W-NP with low affinity within the millimolar range, with dissociation constant (K_d) values ranging from 1.704 mM (Run B Site 1) to 99.99 mM (Run B Site 2). In Runs A and C, binding was enthalpy-driven and exothermic (e.g., Run A ΔH_1 of -9 994 kJ/mol), accompanied by negative entropy changes (e.g., Run A ΔS_1 of -33 500 J/mol·K). This typically suggests the formation of structured complexes involving hydrogen bonding, van der Waals interactions, or electrostatic contacts, and often points to specific interactions with defined binding pockets. In contrast, Runs B and D exhibited an initial binding event with a positive enthalpy (ΔH) alongside positive entropy (ΔS) values, suggesting an entropy-driven interaction like hydrophobic interactions, solvent displacement, or ligand-induced conformational changes (Lewis & Murphy, 2005; Velázquez-Campoy et al., 2004).

The presence of both binding modes may be explained by the protein's intrinsic multivalent nature, with binding characteristics that vary depending on the specific oligomeric state. However, it may suggest that palmatine engages with multiple binding surfaces or protein conformations.

Furthermore, the N protein is believed to be co-purified with nucleic acid ([Section 4.2.3](#)), which may remain associated, influencing the accessibility or conformation of palmatine binding sites, whilst indirectly contributing to the binding events by stabilising certain oligomeric forms. Nucleic acid contamination may contribute to the more structured, enthalpy-driven interactions, mimicking or modulation palmatine binding through ternary complex formation (Carlson et al., 2020; Chang, C. et al., 2014). Therefore, any variation in binding thermodynamics observed across the different runs may reflect such differences.

Overall, these results indicate that palmatine binds weakly but consistently to W-NP, with thermodynamic features suggestive of both specific and nonspecific interactions. The presence of multiple oligomeric states, potential residual RNA, and the protein's intrinsic disorder likely contribute to the apparent multiplicity of binding sites and the variability observed across different experiments. While the binding models used here provide useful approximations, they should be interpreted with caution, as the data do not support clear stoichiometric resolution. Nevertheless, the findings align with the known behaviour of coronavirus nucleocapsid proteins, which exhibit flexible, multivalent binding and context-dependent interaction profiles. These properties are likely to influence both ligand binding and the biological function of the protein, particularly in RNA packaging and viral assembly. Collectively, these results reflect the biological role of W-NP as an RNA chaperone, with regions capable of engaging in multiple, context-specific interactions that result in complex binding landscapes.

4.5.6 Docking

The potential interaction between W-NP and alkaloids berberine and palmatine was individually investigated using molecular docking simulations. Given the absence of a full-length high-resolution structure, these simulations were limited to the functional domains only, employing the NTD (PDB: 7ACT) and CTD (PDB: 7XWX). Discovery studio receptor site cavities predicted 6 potential binding sites for the NTD based on geometric and physicochemical suitability. From these, four sites were selected (sites 1 – 4), three of which extend the known RNA-binding grooves (sites 1, 2, and 4) whilst site 3 covered a

helical portion distal to the RNA-binding site (**Figure 4. 20**). The CTD analysis identified three potential pockets, with site 2 located within/near the dimerisation interface. Both berberine and palmatine were successfully docked using the CDOCKER algorithm into all selected sites across the NTD and CTD.

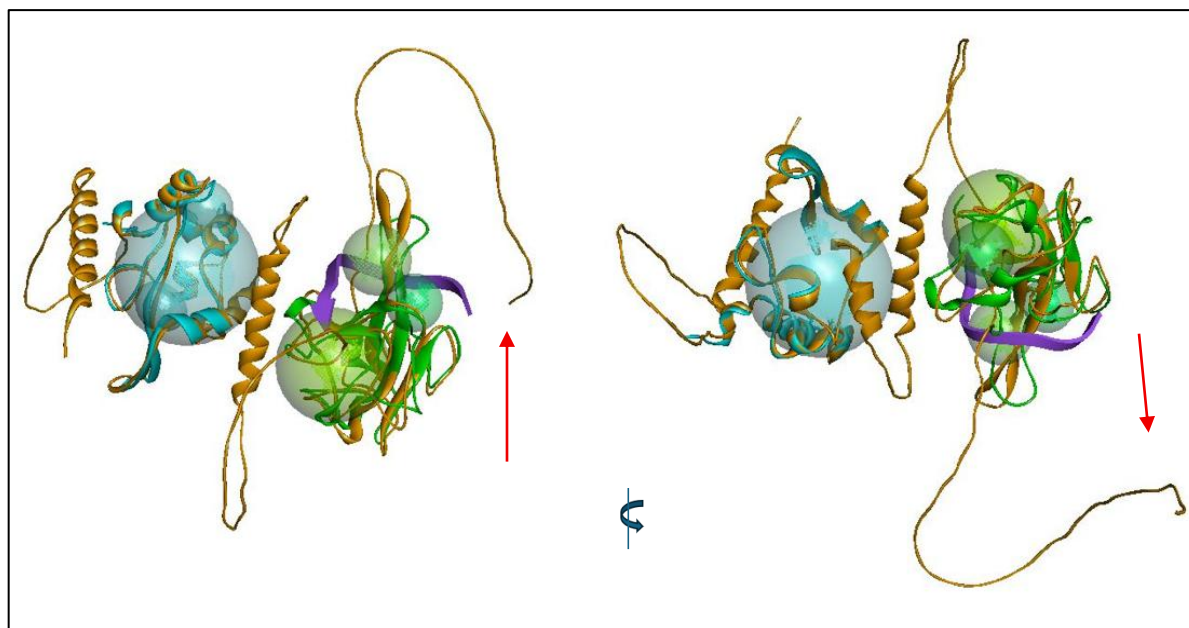


Figure 4. 20: Docking Spheres

Shows different orientations of the 3D structures of the NTD (7ACT; lime) and CTD (7XWX; blue) superimposed to the full-length structure of W-NP monomer (orange) as predicted by AlphaFold. It also shows the various spheres identified by Discovery Studio that was used for the docking simulations.

Binding energies were calculated using in situ ligand minimisation and the Generalised Born with Molecular Volume (GBMV) implicit solvent model and interactions were investigated using their 2D models

The docking results for the NTD is shown in **Figure 4. 21** whilst those for the CTD are displayed in **Figure 4. 22**. Analysis of these results revealed that multiple potential binding sites exist for both alkaloids within both domains across all sites. These interactions involved multiple non-covalent interactions between ligand and protein.

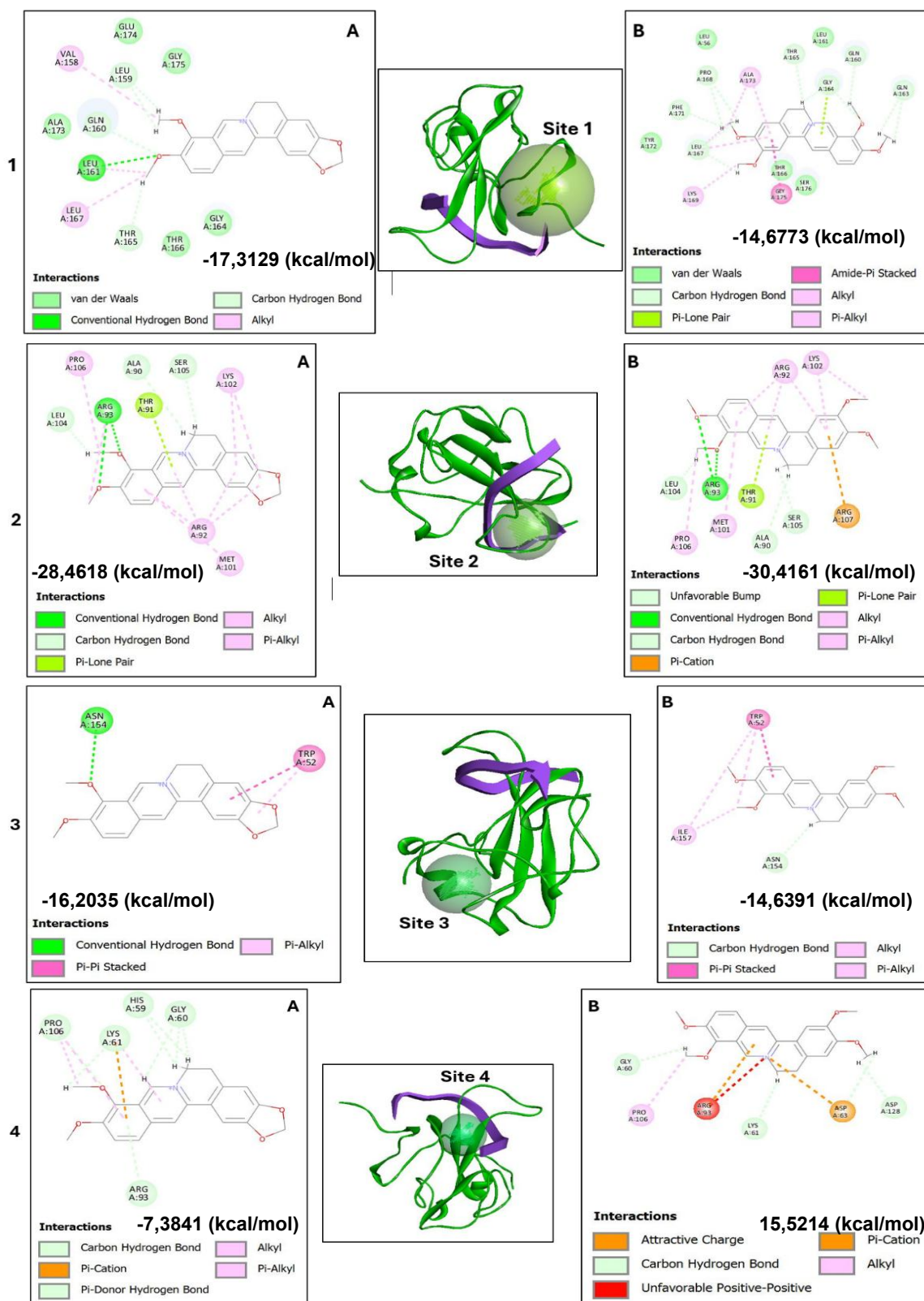


Figure 4. 21: NTD docking

These figures depict a 2D diagram showing the bonds formed during the interactions of berberine (A) and palmatine (B) with the N protein. Figures 1, 2, 3, and 4 display sites 1, 2, 3, and 4 of the NTD respectively. The middle figures depict the location of the binding site spheres within the NTD with respect to the nucleic acid binding site.

Common interactions included conventional hydrogen bonds (e.g., between polar ligand groups and NTD arginines), π – π stacking of aromatic rings (e.g., berberine and palmatine planar rings with aromatic sidechains), π –alkyl contacts (between ligand aromatics and hydrophobic residues), and various van der Waals contacts. Amide– π stacking and π –lone-pair interactions were also observed. These complex interaction networks are consistent with the polyaromatic structures of berberine and palmatine and with prior experimental findings that aromatic small molecules can engage the N-protein via stacking and electrostatic contacts (Mercaldi et al., 2022).

Docking results for the NTD showed that both alkaloids had the strongest affinity for site 2 (**Figure 4. 21**). Palmatine showed a slightly stronger affinity (-30.42 kcal/mol) for this site than berberine (-28.46 kcal/mol), although both indicate favourable interactions. The difference in the binding affinities within the same site is explained by the difference in interactions seen in the 2D diagrams resulting from the differences in alkaloid structures. The second-highest affinity site for berberine was site 3 (-17.3129 kcal/mol) in comparison to site 1 for palmatine (-14.6391 kcal/mol). Docking performed without the removal of nucleic acid co-crystalised within the NTD (not shown here) similarly resulted in positive interactions but with weaker affinities when compared to the RNA-free conditions and a shift in the most favourable site to site 4. These findings are consistent with partial competition or steric hindrances caused by the bound nucleic acid, but they also suggest simultaneous interaction with both protein and nucleic acid.

Docking performed on the three sites identified within the CTD yielded weaker affinities overall when compared to the NTD, indicating a generally more favourable target for these ligands (**Figure 4. 22**). Here, berberine bound most favourably at site 3 (-22.89 kcal/mol), whereas palmatine bound most favourably at site 1 (-28.77 kcal/mol), clearly depicting the differences in these compounds and how these differences are reflected in different binding landscapes.

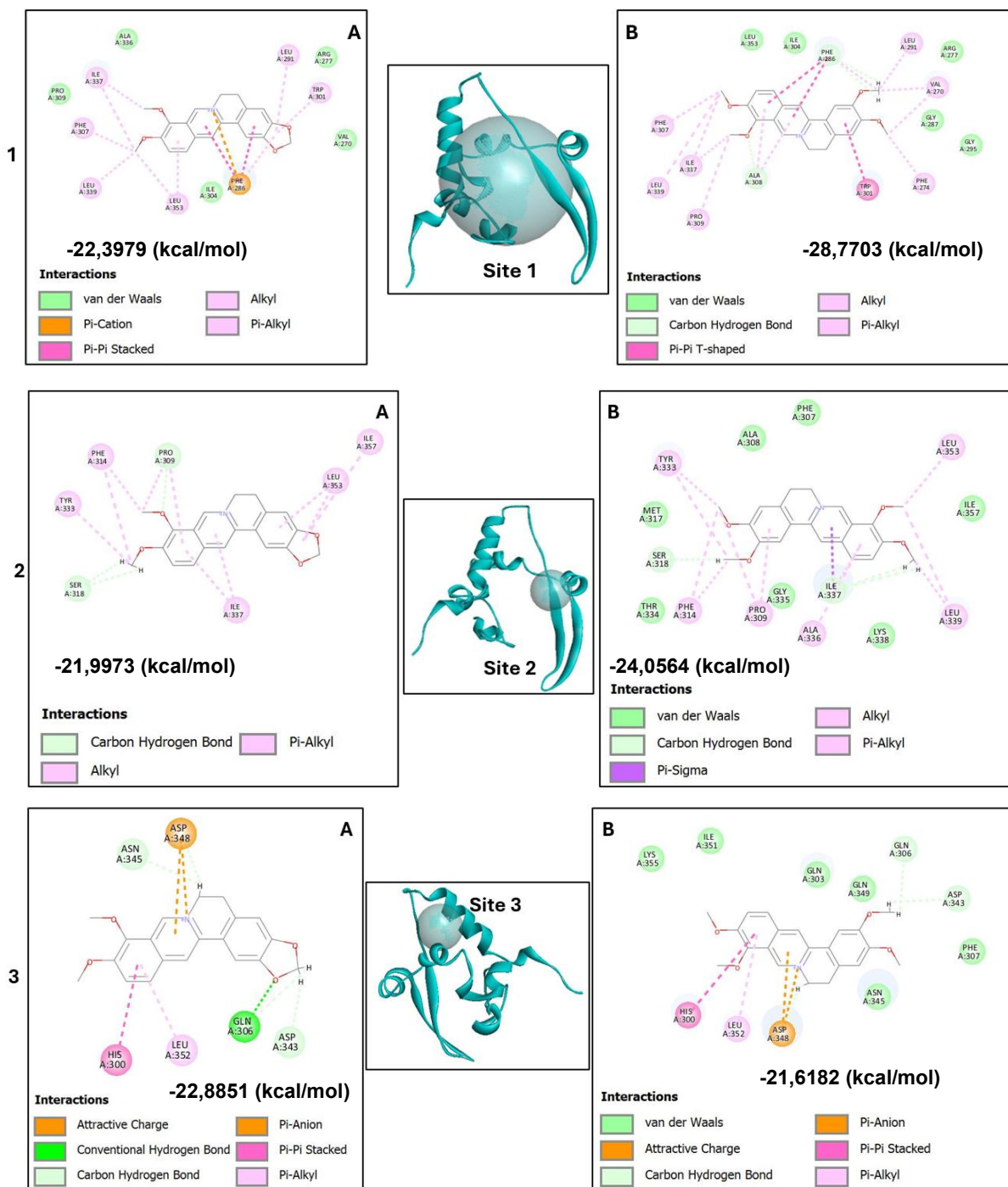


Figure 4. 22: CTD docking

Figures A and B displays the 2D diagram of the interactions of berberine and palmitine respectively. Figures 1, 2, and 3 display sites 1, 2, and 3 of the CTD respectively. Middle figures depict the location of the binding site spheres within the CTD. See comments on other figure this one also needs some optimising

The second-highest affinities for berberine were found at site 1 (-22.3979 kcal/mol) and site 2 for palmatine (-24.0564 kcal/mol).

In summary, both ligands appear to have a diverse set of non-covalent interactions across the protein with a preference towards the NTD. Although distribution across binding sites within the NTD is less evenly distributed across the sites in comparison to interactions within the CTD, indicating more site-specific binding within the NTD.

These findings support prior experimental data obtained from fluorescence spectroscopy and isothermal titration calorimetry. While binding energies were modest overall, the consistent site-specific interactions and interaction types suggest that these compounds could feasibly modulate N-protein function or interfere with RNA binding. These findings support further experimental validation and raise the possibility of allosteric or competitive inhibition mechanisms at play.

5. Discussion

The SARS-CoV-2 nucleocapsid (N) protein is an intrinsically dynamic, multifunctional protein whose RNA-binding, oligomerization and host-interaction activities are finely tuned by its folded domains (NTD & CTD) and disordered regions (N-arm, SR-rich LKR, and C-arm), that together support various functions throughout the viral lifecycle (Chang, C. et al., 2014; Favetta et al., 2024; Wu, W. et al., 2023). This study proposed a consensus to the SARS-CoV-2 N protein that displays 98.33% similarity to the Wuhan isolate, identifying P13L; Δ 31–33; R203K; G204R; and S413R as frequently existing mutations among current circulating variants ([Section 4.1: Figure 4. 1](#)). Absorbance and Circular Dichroism Spectroscopies suggested that these mutations did not abolish nucleic acid binding ([Section 4.2.3: Figure 4. 6](#)) or self-association ([Section 4.3.4: Figure 4. 9](#)) capacities. Rather, it appears that these mutations induce modest α -helix and disorder to β -sheet remodelling with the structure ([Section 4.3; Figure 4. 7](#)), underscoring the need for variant informed constructs in structural studies.

Employing the consensus sequence to the N protein presented here, this study explored the interaction of this protein (and protein-nucleic acid complexes) with two protoberberine alkaloids, berberine and palmatine, well known for their interactions with both proteins and nucleic acids through electrostatic and π - π interactions (Babalgith et al., 2022; Botova et al., 2024; Padmapriya & Barthwal, 2018).

These alkaloids appear to interact with the protein through distinct mechanisms that result in the modulation of hydration, side-chain packaging, and backbone geometries ([Section 4.4.2: Figure 4. 11](#)). These changes are characterised by secondary structural shifts from disordered content towards β -sheet enrichment ([Section 4.4.1: Table 4. 14 & Figure 4. 10](#)). These changes appear to be stabilising in nature resulting in a concurrent tertiary destabilisation ([Section 4.4.4: Figure 4. 13 & Table 4. 5](#)). Neither compound affects quaternary structure; however, it does appear to offer partial protection against urea-induced denaturation of tetrameric assemblies ([Section 4.4.3: Figure 4. 12](#)). These changes are accompanied by secondary changes, characterised by secondary stabilisation and slight tertiary destabilisation brought on by alkaloid presence ([Section 4.4.1 and Section 4.4.4](#)). These changes likely reflect the increased β -sheet content often attributed to the oligomerisation of the protein (Bai et al., 2021; Botova et al., 2024; Chang, C. et al., 2013). Furthermore, this structural rigidification could not only affect the oligomerisation of the protein, but also its RNA-binding capacity (Bai et al., 2021; Botova et al., 2024; Chang, C. et al., 2013), a crucial finding that suggests potential to inhibit N protein function.

Although the data cannot identify the exact binding site/s, it did show that these interactions could potentially take place within either the NTD, LKR, CTD or disordered regions. The data suggested that these interactions could likely occur at multiple sites ([Section 4.5.5 & Section 4.5.6](#)), but that they are extremely weak ([Sections 4.5.4 ; 4.5.5; and 4.5.6](#)) and possibly transient ([Section 4.5.4](#)). Nonetheless, these weak interactions still result distinct mechanisms of interactions that are congruent with the chemistries of the alkaloids ([Section 4.4.2](#)), where berberine's smaller, planar, rigid methylenedioxy aromatic core favours pi-pi stacking within densely packed domains (Varghese et al.,

2021; Wang et al., 2021; Warowicka et al., 2020), whilst palmatine's multi-methoxylated polar framework enables hydrogen bonding and looser binding interactions (Long et al., 2019; Padmapriya & Barthwal, 2018; Purwaningsih et al., 2024).

These interactions may allow either alkaloid to act as a molecular brace that alters the conformational flexibility of the protein, whilst maintaining the quaternary assembly. These structural changes may carry significant functional implications within this flexible multifunctional protein. Potential interactions within the NTD or CTD may significantly alter the nucleic acid binding or self-assembly of the N protein (Ye et al., 2020). Interaction within the intrinsically disordered SR-rich LKR region of the protein may exert profound allosteric effects on various key regulatory functions, including but not limited to, RNA-binding affinities, oligomerization, phase behaviour, and protein-protein or protein-nucleic interactions. Interactions here could sterically hinder access of kinases to serine residues, disrupt transient helical contacts, potentially skewing the phosphorylation balance toward unphosphorylated-like states (Botova et al., 2024; Favetta et al., 2024). Hydrogen bonding near phosphorylation motifs may alter local conformational landscapes, affecting electrostatic interactions that are critical for the dynamic assembly. Impeding or mimicking phosphorylation induced charge could suppress the LLPS formation of liquid-like condensates that promote viral replication, shifting the N protein back towards packaging-competent gel-like states (Botova et al., 2024; Favetta et al., 2024). Any such modulation may also interfere with N's recruitment of essential host or viral factors, like kinases or the M protein, ultimately impairing efficient replication and virion assembly.

Furthermore, berberine is a known stabilizer of viral RNA G-quadruplexes that influence N expression levels (Oliva et al., 2022). Dual engagement with both RNA structural motifs and N-protein surfaces may produce synergistic inhibition, potentially reducing viral gene expression and the disruption of nucleocapsid assembly. This combined action and dual targeting may allow multi-modal mechanisms and more effective antiviral effects. Beyond targeting the nucleocapsid protein, both compounds display multiple antiviral and host-modulating effects (de Silva et al., 2021; Purwaningsih et al., 2024; Wang et al., 2021; Warowicka et al., 2020) that support their therapeutic potential against SARS-CoV-2.

Although the interactions are not high-affinity or stable, even transient interaction may result in great functional outcomes. Together, the data suggest plausible mechanisms by which the alkaloids could disrupt critical N protein activities like RNA-driven oligomerization, phase separation, and post-translational regulation, establishing a solid structural foundation for continued investigation of berberine and palmatine as modulators of SARS-CoV-2 N protein functions, and offering a rationale for in-cell and in vivo exploration as mutation-resilient antivirals.

6. Conclusion

This work established a versatile, variant-informed platform for probing the SARS-CoV-2 Nucleocapsid protein. The platform incorporated the expression of an N protein consensus sequence, incorporating the persisting P13L; Δ 31–33; R203K; G204R; and S413R mutations found in various lineages, using *Escherichia coli* cells and purification via nickel-affinity chromatography. This construct retained nucleic acid binding and self-association capacity. These mutations caused increased β -sheet propensity at the expense of α -helix and disorder, underscoring the necessity of using variant-informed sequences for further studies.

Furthermore, the study used a multimodal approach that included in silico docking, SPR, ITC, ATR-FTIR, Far-UV-CD, BN-PAGE and structural spectrometric techniques demonstrating low-affinity multi-site interactions between protoberberine alkaloids berberine and palmatine with the N protein. Although neither compound disrupted native oligomeric assemblies, both induced subtle secondary structural shifts toward β -sheet enrichment and partial protection against urea-induced denaturation, effects likely arising from distinct binding footprints in folded (NTD/CTD) or intrinsically disordered (SR-rich LKR) regions. The data generated here suggest that this low-affinity binding could act as molecular braces, modulating conformational flexibility, potentially allosterically impeding key N protein activities. Previous work showing berberine's stabilisation of nucleic acid could allow a dual-target mechanism that may synergistically undermine N protein assembly and viral gene expression, together allowing these transient engagements to result in meaningful antiviral outcomes. Together, this study lays a foundation for continued exploration of the alkaloids as mutation-resilient modulators of SARS-CoV-2 N protein function as a potential treatment.

7. Limitations of the study

The research conducted here and the inferences drawn carry various limitations. Various constraints arose from the methodological design, experimental scope, and sample variability, all of which may limit the broader applicability of the observed interactions between the SARS-CoV-2 N protein and the alkaloids berberine and palmatine.

Firstly, the methodology used to obtain the consensus sequence included only mutations that are common amongst circulating variants, and therefore does not represent newly emerging, or less prevalent substitutions. Furthermore, these mutations may commonly occur as a single set, whilst a combination of these mutations may behave differently to those naturally occurring. Furthermore, this study employed *E. coli* bacteria that lack post-translational modification ability to express the recombinant N protein. These discrepancies may severely impact the various functions of the protein, its interactions and native characteristics.

In addition, the investigation of protein-ligand interactions was limited to *in vitro* binding assays. The absence of *in vivo* validation precludes firm conclusions about antiviral activity within physiological systems. Whilst these assays provide quantitative data, it does not fully mimic the complexity of the host-pathogen environments. Furthermore, whilst the study confirms binding, it does not actively quantify it, nor does it fully elucidate the binding site, mode of interactions, or whether these alkaloids interfere with functions. The structural resolutions have limitations; for example, BN-PAGE reports cannot distinguish subtle oligomeric states. In addition, due to experimental complexity, time and resource limitations, the binding experiments were conducted mostly at fixed ligand concentrations, with some error occurring in berberine preparation for SPR. This narrow concentration range restricts the ability to derive accurate thermodynamic profiles and full dose-response relationships.

Furthermore, the chemical purity and solvent interactions were controlled but not exhaustively characterized. Potential solvent effects or trace impurities might have subtly influenced the results. Nonetheless, systematic acknowledgement of the limitations

enables this study to maintain scientific rigor and transparency. While the results lay important groundwork for these interactions, further validations through physiological assays and mechanistic probing are required before consideration for therapeutic developments.

8. Future considerations

Future studies should preclude various considerations. Firstly, various physiologically relevant phosphorylation and modifications should be incorporated into the expression system, enabling the assessment of their impact on structure and alkaloid engagement. Targeted substitutions through site-directed mutagenesis in the NTD, CTD, and LKR regions may validate predicted interaction hotspots, enabling the elucidation of binding sites using fluorescence spectroscopy. Nucleic acid removal may elucidate the role of these binding partners in protein and protein-alkaloid behaviours. Furthermore, various complementary techniques may elucidate further information about the protein and protein-alkaloid complexes. Fluorescence lifetime measurements may provide knowledge on interactions. Differential Scanning Calorimetry (DSC) may further determine thermodynamic signatures. High Performance Liquid Chromatography (HPLC), Size Exclusion Chromatography (SEC), small-angle X-ray scattering (SAXS) or analytical ultracentrifugation may better characterize the quaternary/oligomeric structures within samples. Nuclear Magnetic Resonance (NMR) spectroscopy, X-ray crystallography, or cryo-EM may localise precise alkaloid binding sites within folded or disordered regions.

9. References

Abdolmaleki, G., Taheri, M. A., Paridehpour, S., Mohammadi, N. M., Tabatabaei, Y. A., Mousavi, T. & Amin, M. (2022). A comparison between SARS-CoV-1 and SARS-CoV2: an update on current COVID-19 vaccines. *DARU Journal of Pharmaceutical Sciences*, 30(2), 379–406. <https://doi.org/10.1007/s40199-022-00446-8>

Acharya, K. R. & Lloyd, M. D. (2005). The advantages and limitations of protein crystal structures. *Trends in Pharmacological Sciences*, 26(1), 10–14. <https://doi.org/10.1016/j.tips.2004.10.011>

Adhikari, S., Manthena, P. V., Sajwan, K., Kota, K. K. & Roy, R. (2010). A unified method for purification of basic proteins. *Analytical Biochemistry*, 400(2), 203–206. <https://doi.org/10.1016/j.ab.2010.01.011>

Al-Aly, Z., Davis, H., McCorkell, L., Soares, L., Wulf-Hanson, S., Iwasaki, A. & Topol, E. J. (2024). Long COVID science, research and policy. *Nature Medicine*, 30(8), 2148–2164. <https://doi.org/10.1038/s41591-024-03173-6>

Aloor, A., Aradhya, R., Venugopal, P., Gopalakrishnan Nair, B. & Suravajhala, R. (2022). Glycosylation in SARS-CoV-2 variants: A path to infection and recovery. *Biochemical Pharmacology*, 206115335. <https://doi.org/10.1016/j.bcp.2022.115335>

Alsuwairi, F. A., Alsaleh, A. N., Alsanea, M. S., Al-Qahtani, A. A., Obeid, D., Almaghrabi, R. S., Alahideb, B. M., AlAbdulkareem, M. A., Mutabagani, M. S., Althawadi, S. I., Altamimi, S. A., Alshukairi, A. N. & Alhamlan, F. S. (2023). Association of SARS-CoV-2 Nucleocapsid Protein Mutations with Patient Demographic and Clinical Characteristics during the Delta and Omicron Waves. *Microorganisms*, 11(5), 1288. <https://doi.org/10.3390/microorganisms11051288>

Anderson, G. P., Liu, J. L., Esparza, T. J., Voelker, B. T., Hofmann, E. R. & Goldman, E. R. (2021). Single-Domain Antibodies for the Detection of SARS-CoV-2 Nucleocapsid

Protein. Analytical Chemistry, 93(19), 7283–7291.
<https://doi.org/10.1021/acs.analchem.1c00677>

Antonopoulou, I., Sapountzaki, E., Rova, U. & Christakopoulos, P. (2022). Inhibition of the main protease of SARS-CoV-2 (Mpro) by repurposing/designing drug-like substances and utilizing nature's toolbox of bioactive compounds. *Computational and Structural Biotechnology Journal*, 201306–1344. <https://doi.org/10.1016/j.csbj.2022.03.009>

Arai, M., Suetaka, S. & Ooka, K. (2024). Dynamics and interactions of intrinsically disordered proteins. *Current Opinion in Structural Biology*, 84102734. <https://doi.org/10.1016/j.sbi.2023.102734>

Azad, G. K. (2021). Identification and molecular characterization of mutations in nucleocapsid phosphoprotein of SARS-CoV-2. *PeerJ*, 9e10666. <https://doi.org/10.7717/peerj.10666>

Babalgith, A. O., Al-Kuraishy, H. M., Al-Gareeb, A. I., De Waard, M., Al-Hamash, S. M., Jean-Marc, S., Negm, W. A. & Batiha, G. E. (2022). The role of berberine in Covid-19: potential adjunct therapy. *Inflammopharmacology*, 30(6), 2003–2016. <https://doi.org/10.1007/s10787-022-01080-1>

Babcock, J. J. & Brancalion, L. (2013). Bovine serum albumin oligomers in the E- and B-forms at low protein concentration and ionic strength. *International Journal of Biological Macromolecules*, 5342–53. <https://doi.org/10.1016/j.ijbiomac.2012.10.030>

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>

Bank, R. P. D.RCSB PDB Group Sequence: P0DTC9. <https://www.rcsb.org/>. Retrieved Aug 5, 2024, from https://www.rcsb.org/groups/sequence/polymer_entity/P0DTC9

Bastos, M., Abian, O., Johnson, C. M., Ferreira-da-Silva, F., Vega, S., Jimenez-Alesanco, A., Ortega-Alarcon, D. & Velazquez-Campoy, A. (2023). Isothermal titration calorimetry. *Nature Reviews Methods Primers*, 3(1), 17. <https://doi.org/10.1038/s43586-023-00199-x>

Bastos, M. & Velazquez-Campoy, A. (2021). Isothermal titration calorimetry (ITC): a standard operating procedure (SOP). *European Biophysics Journal: EBJ*, 50(3-4), 363–371. <https://doi.org/10.1007/s00249-021-01509-5>

Battu, S. K., Repka, M. A., Maddineni, S., Chittiboyina, A. G., Avery, M. A. & Majumdar, S. (2010). Physicochemical Characterization of Berberine Chloride: A Perspective in the Development of a Solution Dosage Form for Oral Delivery. *AAPS PharmSciTech*, 11(3), 1466–1475. <https://doi.org/10.1208/s12249-010-9520-y>

Beavis, R. C. & Chait, B. T. (1990). High-accuracy molecular mass determination of proteins using matrix-assisted laser desorption mass spectrometry. *Analytical Chemistry*, 62(17), 1836–1840. <https://doi.org/10.1021/ac00216a020>

Benkert, P., Biasini, M. & Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*, 27(3), 343–350. <https://doi.org/10.1093/bioinformatics/btq662>

Bhadra, K. & Kumar, G. S. (2011). Interaction of berberine, palmatine, coralyne, and sanguinarine to quadruplex DNA: A comparative spectroscopic and calorimetric study. *Biochimica Et Biophysica Acta (BBA) - General Subjects*, 1810(4), 485–496. <https://doi.org/10.1016/j.bbagen.2011.01.011>

Bian, S., Li, L., Wang, Z., Cui, L., Xu, Y., Guan, K. & Zhao, B. (2022). Allergic Reactions After the Administration of COVID-19 Vaccines. *Frontiers in Public Health*, 10878081. <https://doi.org/10.3389/fpubh.2022.878081>

Blom, N., Sicheritz-Pontén, T., Gupta, R., Gammeltoft, S. & Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics*, 4(6), 1633–1649. <https://doi.org/10.1002/pmic.200300771>

Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R. & Klug, A. (1978). Protein disk of tobacco mosaic virus at 2.8 Å resolution showing the interactions within and between subunits. *Nature*, 276(5686), 362–368. <https://doi.org/10.1038/276362a0>

Bornhorst, J. A. & Falke, J. J. (2000). [16] Purification of Proteins Using Polyhistidine Affinity Tags. *Methods in Enzymology*, 326, 245–254. [https://doi.org/10.1016/s0076-6879\(00\)26058-8](https://doi.org/10.1016/s0076-6879(00)26058-8)

Botova, M., Camacho-Zarco, A. R., Tognetti, J., Bessa, L. M., Guseva, S., Mikkola, E., Salvi, N., Maurin, D., Herrmann, T. & Blackledge, M. (2024). A specific phosphorylation-dependent conformational switch in SARS-CoV-2 nucleocapsid protein inhibits RNA binding. *Science Advances*, 10(31), eaax2323. <https://doi.org/10.1126/sciadv.aax2323>

Bouhaddou, M., Memon, D., Meyer, B., White, K. M., Rezelj, V. V., Correa Marrero, M., Polacco, B. J., Melnyk, J. E., Ulferts, S., Kaake, R. M., Batra, J., Richards, A. L., Stevenson, E., Gordon, D. E., Rojc, A., Obernier, K., Fabius, J. M., Soucheray, M., Miorin, L., . . . Krogan, N. J. (2020). The Global Phosphorylation Landscape of SARS-CoV-2 Infection. *Cell*, 182(3), 685–712.e19. <https://doi.org/10.1016/j.cell.2020.06.034>

Brudenell, E. L., Pohare, M. B., Zafred, D., Phipps, J., Hornsby, H. R., Darby, J. F., Dai, J., Liggett, E., Cain, K. M., Barran, P. E., de Silva, T. I. & Sayers, J. R. (2024). Efficient overexpression and purification of severe acute respiratory syndrome coronavirus 2 nucleocapsid proteins in *Escherichia coli*. *The Biochemical Journal*, 481(11), 669–682. <https://doi.org/10.1042/BCJ20240019>

Cai, T., Yu, Z., Wang, Z., Liang, C. & Richard, S. (2021). Arginine methylation of SARS-Cov-2 nucleocapsid protein regulates RNA binding, its ability to suppress stress granule

formation, and viral replication. *The Journal of Biological Chemistry*, 297(1), 100821. <https://doi.org/10.1016/j.jbc.2021.100821>

Calhoun, D. B., Vanderkooi, J. M., Holtom, G. R. & Englander, S. W. (1986). Protein fluorescence quenching by small molecules: protein penetration versus solvent exposure. *Proteins*, 1(2), 109–115. <https://doi.org/10.1002/prot.340010202>

Camacho-Carvajal, M. M., Wollscheid, B., Aebersold, R., Steimle, V. & Schamel, W. W. A. (2004). Two-dimensional Blue native/SDS gel electrophoresis of multi-protein complexes from whole cellular lysates: a proteomics approach. *Molecular & Cellular Proteomics: MCP*, 3(2), 176–182. <https://doi.org/10.1074/mcp.T300010-MCP200>

Cao, Y., Wang, J., Jian, F., Xiao, T., Song, W., Yisimayi, A., Huang, W., Li, Q., Wang, P., An, R., Wang, J., Wang, Y., Niu, X., Yang, S., Liang, H., Sun, H., Li, T., Yu, Y., Cui, Q., . . . Xie, X. S. (2022). Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature*, 602(7898), 657–663. <https://doi.org/10.1038/s41586-021-04385-3>

Carlson, C. R., Asfaha, J. B., Ghent, C. M., Howard, C. J., Hartooni, N., Safari, M., Frankel, A. D. & Morgan, D. O. (2020). Phosphoregulation of Phase Separation by the SARS-CoV-2 N Protein Suggests a Biophysical Basis for its Dual Functions. *Molecular Cell*, 80(6), 1092–1103.e4. <https://doi.org/10.1016/j.molcel.2020.11.025>

Caruso, Í P., Sanches, K., Poian, A. T. D., Pinheiro, A. S. & Almeida, F. C. L. (2021). Dynamics of the SARS-CoV-2 nucleoprotein N-terminal domain triggers RNA duplex destabilization. *Biophysical Journal*, 120(14), 2814–2827. <https://doi.org/10.1016/j.bpj.2021.06.003>

Chand, G. B., Banerjee, A. & Azad, G. K. (2020). Identification of novel mutations in RNA-dependent RNA polymerases of SARS-CoV-2 and their implications on its protein structure. *PeerJ*, 8e9492. <https://doi.org/10.7717/peerj.9492>

Chang, C., Chen, C. M., Chiang, M., Hsu, Y. & Huang, T. (2013). Transient Oligomerization of the SARS-CoV N Protein – Implication for Virus Ribonucleoprotein Packaging. *PLoS ONE*, 8(5), e65045. <https://doi.org/10.1371/journal.pone.0065045>

Chang, C., Hou, M., Chang, C., Hsiao, C. & Huang, T. (2014). The SARS coronavirus nucleocapsid protein--forms and functions. *Antiviral Research*, 10339–50. <https://doi.org/10.1016/j.antiviral.2013.12.009>

Chang, Y., Wang, Y., Zhang, X., Iqbal, J., Lu, M. & Du, Y. (2021). Transcriptional regulation of small heat shock protein genes by heat shock factor 1 (HSF1) in *Liriomyza trifolii* under heat stress. *Cell Stress and Chaperones*, 26(5), 835–843. <https://doi.org/10.1007/s12192-021-01224-2>

Chauhan, A., Avti, P., Shekhar, N., Prajapat, M., Sarma, P., Bhattacharyya, A., Kumar, S., Kaur, H., Prakash, A. & Medhi, B. (2021). Structural and conformational analysis of SARS CoV 2 N-CTD revealing monomeric and dimeric active sites during the RNA-binding and stabilization: Insights towards potential inhibitors for N-CTD. *Computers in Biology and Medicine*, 134104495. <https://doi.org/10.1016/j.combiomed.2021.104495>

Chen, R., Kan, L., Duan, F., He, L., Wang, M., Cui, J., Zhang, Z. & Zhang, Z. (2021). Surface plasmon resonance aptasensor based on niobium carbide MXene quantum dots for nucleocapsid of SARS-CoV-2 detection. *Microchimica Acta*, 188(10), 316. <https://doi.org/10.1007/s00604-021-04974-z>

Chen, Y., Liu, Q. & Guo, D. (2020). Emerging coronaviruses: Genome structure, replication, and pathogenesis. *Journal of Medical Virology*, 92(4), 418–423. <https://doi.org/10.1002/jmv.25681>

Cheng, N., Liu, M., Li, W., Sun, B., Liu, D., Wang, G., Shi, J. & Li, L. (2023). Protein post-translational modification in SARS-CoV-2 and host interaction. *Frontiers in Immunology*, 131068449. <https://doi.org/10.3389/fimmu.2022.1068449>

Chittur, K. K. (1998). FTIR/ATR for protein adsorption to biomaterial surfaces. *Biomaterials*, 19(4), 357–369. [https://doi.org/10.1016/S0142-9612\(97\)00223-8](https://doi.org/10.1016/S0142-9612(97)00223-8)

Clore, G. M., Robien, M. A. & Gronenborn, A. M. (1993). Exploring the Limits of Precision and Accuracy of Protein Structures Determined by Nuclear Magnetic Resonance Spectroscopy. *Journal of Molecular Biology*, 231(1), 82–102. <https://doi.org/10.1006/jmbi.1993.1259>

Contreras-Martos, S., Nguyen, H. H., Nguyen, P. N., Hristozova, N., Macossay-Castillo, M., Kovacs, D., Bekesi, A., Oemig, J. S., Maes, D., Pauwels, K., Tompa, P. & Lebrun, P. (2018). Quantification of Intrinsically Disordered Proteins: A Problem Not Fully Appreciated. *Frontiers in Molecular Biosciences*, 5 <https://doi.org/10.3389/fmolb.2018.00083>

Creighton, T. E. (1993). *Proteins: Structures and Molecular Properties*. Macmillan.

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), 1936. <https://doi.org/10.1038/s41467-021-21953-3>

Cubuk, J., Alston, J., Incicco, J. J., Holehouse, A., Hall, K., Stuchell-Brereton, M. & Soranno, A. (2024). The disordered N-terminal tail of SARS-CoV-2 Nucleocapsid protein forms a dynamic complex with RNA. *Nucleic Acids Research*, 52(5), 2609–2624. <https://doi.org/10.1093/nar/gkad1215>

Cubuk, J., Incicco, J. J., Hall, K. B., Holehouse, A. S., Stuchell-Brereton, M. D. & Soranno, A. (2024). The dimerization domain of SARS CoV 2 Nucleocapsid protein is partially disordered as a monomer and forms a high affinity dynamic complex. *bioRxiv: The*

Cytiva. Ion Exchange Chromatography & Chromatofocusing, Principles and Methods Handbook (2021). Cytiva. <https://cdn.cytivalifesciences.com/api/public/content/digi-13101-pdf>

Das, A., Khan, S., Roy, S. & Das, S. (2023). Phytochemicals for mitigating the COVID-19 crisis: evidence from pre-clinical and clinical studies. *Exploration of Drug Science*, 1(5), 336–376. <https://doi.org/10.37349/eds.2023.00024>

Davis, A. M., St-Gallay, S. A. & Kleywegt, G. J. (2008). Limitations and lessons in the use of X-ray structural information in drug design. *Drug Discovery Today*, 13(19), 831–841. <https://doi.org/10.1016/j.drudis.2008.06.006>

de Silva, T. I., Liu, G., Lindsey, B. B., Dong, D., Moore, S. C., Hsu, N. S., Shah, D., Wellington, D., Mentzer, A. J., Angyal, A., Brown, R., Parker, M. D., Ying, Z., Yao, X., Turtle, L., Dunachie, S., Aanensen, D. M., Abudahab, K., Adams, H., . . . Dong, T. (2021). The impact of viral mutations on recognition by SARS-CoV-2 specific T cells. *iScience*, 24(11), 103353. <https://doi.org/10.1016/j.isci.2021.103353>

Deng, X., Eickholt, J. & Cheng, J. (2012). A comprehensive overview of computational protein disorder prediction methods. *Molecular bioSystems*, 8(1), 114–121. <https://doi.org/10.1039/c1mb05207a>

Di, D., Dileepan, M., Ahmed, S., Liang, Y. & Ly, H. (2021). Recombinant SARS-CoV-2 Nucleocapsid Protein: Expression, Purification, and Its Biochemical Characterization and Utility in Serological Assay Development to Assess Immunological Responses to SARS-CoV-2 Infection. *Pathogens*, 10(8), 1039. <https://doi.org/10.3390/pathogens10081039>

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>

Dismer, F., Petzold, M. & Hubbuch, J. (2008). Effects of ionic strength and mobile phase pH on the binding orientation of lysozyme on different ion-exchange adsorbents. *Journal of Chromatography. A*, 1194(1), 11–21. <https://doi.org/10.1016/j.chroma.2007.12.085>

Drayman, N., Jones, K. A., Azizi, S., Froggatt, H. M., Tan, K., Maltseva, N. I., Chen, S., Nicolaescu, V., Dvorkin, S., Furlong, K., Kathayat, R. S., Firpo, M. R., Mastrodomenico, V., Bruce, E. A., Schmidt, M. M., Jedrzejczak, R., Muñoz-Alía, M. Á, Schuster, B., Nair, V., . . . Tay, S. (2020). Drug repurposing screen identifies masitinib as a 3CLpro inhibitor that blocks replication of SARS-CoV-2 in vitro. *bioRxiv: The Preprint Server for Biology*, 2020.08.31.274639. <https://doi.org/10.1101/2020.08.31.274639>

E. coli Cell Culture Concentration from OD600 Calculator <https://www.agilent.com/>. https://www.agilent.com/store/biocalculators/calcODBacterial.jsp?_requestid=4918775

Eftink, M. R. & Ghiron, C. A. (1981). Fluorescence quenching studies with proteins. *Analytical Biochemistry*, 114(2), 199–227. [https://doi.org/10.1016/0003-2697\(81\)90474-7](https://doi.org/10.1016/0003-2697(81)90474-7)

European Centre for Disease Prevention and Control. SARS-CoV-2 variants of concern as of 27 June 2025. <https://www.ecdc.europa.eu/en/covid-19/variants-concern>. Retrieved 15/08/2025, from <https://www.ecdc.europa.eu/en/covid-19/variants-concern>

Experts in Protein Expression and Purification | NEB Retrieved May 24, 2026, from <https://www.neb.com/en/protein-expression/experts-in-protein-expression-and-purification>

Faksova, K., Walsh, D., Jiang, Y., Griffin, J., Phillips, A., Gentile, A., Kwong, J. C., Macartney, K., Naus, M., Grange, Z., Escolano, S., Sepulveda, G., Shetty, A., Pillsbury, A., Sullivan, C., Naveed, Z., Janjua, N. Z., Giglio, N., Perälä, J., . . . Hviid, A. (2024). COVID-19 vaccines and adverse events of special interest: A multinational Global Vaccine Data Network (GVDN) cohort study of 99 million vaccinated individuals. *Vaccine*, 42(9), 2200–2211. <https://doi.org/10.1016/j.vaccine.2024.01.100>

Fan, L., Du, B., Pei, F., Hu, W., Feng, S., Liu, B., Tong, Z., Tan, W. & Mu, X. (2023). A Novel SPR Immunosensor Based on Dual Signal Amplification Strategy for Detection of SARS-CoV-2 Nucleocapsid Protein. *Biosensors*, 13(5), 549. <https://doi.org/10.3390/bios13050549>

Fani, M., Teimoori, A. & Ghafari, S. (2020). Comparison of the COVID-2019 (SARS-CoV-2) pathogenesis with SARS-CoV and MERS-CoV infections. *Future Virology*, 10.2217/fvl-0050. <https://doi.org/10.2217/fvl-2020-0050>

Farci, D., Graça, A. T., Hall, M., Haniewicz, P., Kereiche, S., Faull, P., Kirkpatrick, J., Tramontano, E., Schröder, W. P. & Piano, D. (2025). Characterization of SARS-CoV-2 nucleocapsid protein oligomers. *Journal of Structural Biology*, 217(1), 108162. <https://doi.org/10.1016/j.jsb.2024.108162>

Farewell, A. & Neidhardt, F. C. (1998). Effect of temperature on in vivo protein synthetic capacity in *Escherichia coli*. *Journal of Bacteriology*, 180(17), 4704–4710. <https://doi.org/10.1128/JB.180.17.4704-4710.1998>

Favetta, B., Wang, H., Cubuk, J., Barai, M., Ramirez, C., Gormley, A. J., Murthy, S., Soranno, A., Shi, Z. & Schuster, B. S. (2024). Phosphorylation Toggles the SARS-CoV-2 Nucleocapsid Protein Between Two Membrane-Associated Condensate States. *bioRxiv*, 2024.10.17.618867. <https://doi.org/10.1101/2024.10.17.618867>

FDA. Coronavirus (COVID-19) | Drugs. <https://www.fda.gov/drugs/emergency-preparedness-drugs/coronavirus-covid-19-drugs>. Retrieved 15/05/2025, from <https://www.fda.gov/drugs/emergency-preparedness-drugs/coronavirus-covid-19-drugs>

Feig, M. & Brooks, C. L. (2004). Recent advances in the development and application of implicit solvent models in biomolecule simulations. *Current Opinion in Structural Biology*, 14(2), 217–224. <https://doi.org/10.1016/j.sbi.2004.03.009>

Ford, N. D. (2023). Long COVID and Significant Activity Limitation Among Adults, by Age — United States, June 1–13, 2022, to June 7–19, 2023. *MMWR. Morbidity and Mortality Weekly Report*, 72 <https://doi.org/10.15585/mmwr.mm7232a3>

Forsythe, H. M., Rodriguez Galvan, J., Yu, Z., Pinckney, S., Reardon, P., Cooley, R. B., Zhu, P., Rolland, A. D., Prell, J. S. & Barbar, E. (2021). Multivalent binding of the partially disordered SARS-CoV-2 nucleocapsid phosphoprotein dimer to RNA. *Biophysical Journal*, 120(14), 2890–2901. <https://doi.org/10.1016/j.bpj.2021.03.023>

Frenzel, D., Glück, J. M., Brener, O., Oesterhelt, F., Nagel-Steger, L. & Willbold, D. (2014). Immobilization of Homogeneous Monomeric, Oligomeric and Fibrillar A β Species for Reliable SPR Measurements. *PLoS ONE*, 9(3), e89490. <https://doi.org/10.1371/journal.pone.0089490>

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>

Garg, A., Gielnik, M. B., & Kjaergaard, M. (2026). Sequence determinants of the hypomobility of intrinsically disordered proteins in SDS-PAGE. <https://www.biorxiv.org/content/10.64898/2026.03.24.714011v1>

Gasteiger, E.(et al.). (2005). Protein Identification and Analysis Tools on the ExPASy Server. In J. M. Walker (Ed.), *The Proteomics Protocols Handbook* (pp. 571–607). Humana Press.

Gasymov, O. K. & Glasgow, B. J. (2007). ANS fluorescence: Potential to augment the identification of the external binding sites of proteins. *Biochimica Et Biophysica Acta (BBA) - Proteins and Proteomics*, 1774(3), 403–411. <https://doi.org/10.1016/j.bbapap.2007.01.002>

Gibbs, E. B. & Showalter, S. A. (2015). Quantitative Biophysical Characterization of Intrinsically Disordered Proteins. *Biochemistry*, 54(6), 1314–1326. <https://doi.org/10.1021/bi501460a>

Gildenhuys, S., Steenkamp, P. A., Sekele, L., Mosebi, S. & James, C. C. (2023). An Analysis in Identifying the Mechanism of Action of the Alkaloids Found in *Berberis vulgaris* on the Treatment and Prevention of SARS-CoV-2. *Natural Product Communications*, 18(12), 1934578X231212835. <https://doi.org/10.1177/1934578X231212835>

GitHub. CoV-lineages (2024). <https://github.com/cov-lineages/lineages>

Gomes, L., Monteiro, G. & Mergulhão, F. (2020). The Impact of IPTG Induction on Plasmid Stability and Heterologous Protein Expression by *Escherichia coli* Biofilms. *International Journal of Molecular Sciences*, 21(2), 576. <https://doi.org/10.3390/ijms21020576>

Gorbalenya, A. E., Snijder, E. J. & Spaan, W. J. M. (2004). Severe acute respiratory syndrome coronavirus phylogeny: toward consensus. *Journal of Virology*, 78(15), 7863–7866. <https://doi.org/10.1128/JVI.78.15.7863-7866.2004>

Görbitz, C. H. (1989). Hydrogen-bond distances and angles in the structures of amino acids and peptides. *Acta Crystallographica Section B: Structural Science*, 45(4), 390–395. <https://doi.org/10.1107/S0108768189003666>

Grimsley, G. R. & Pace, C. N. (2004). Spectrophotometric determination of protein concentration. *Current Protocols in Protein Science*, <https://doi.org/10.1002/0471140864.ps0301s33>

Guan, Q., Sadykov, M., Mfarrej, S., Hala, S., Naeem, R., Nugmanova, R., Al-Omari, A., Salih, S., Mutair, A. A., Carr, M. J., Hall, W. W., Arold, S. T. & Pain, A. (2020). A genetic barcode of SARS-CoV-2 for monitoring global distribution of different clades during the COVID-19 pandemic. *International Journal of Infectious Diseases*, 100216–223. <https://doi.org/10.1016/j.ijid.2020.08.052>

Gupta, R., Charron, J., Stenger, C. L., Painter, J., Steward, H., Cook, T. W., Faber, W., Frisch, A., Lind, E., Bauss, J., Li, X., Sirpilla, O., Soehnen, X., Underwood, A., Hinds, D., Morris, M., Lamb, N., Carcillo, J. A., Bupp, C., . . . Prokop, J. W. (2020). SARS-CoV-2 (COVID-19) structural and evolutionary dynamicome: Insights into functional evolution and human genomics. *Journal of Biological Chemistry*, 295(33), 11742–11753. <https://doi.org/10.1074/jbc.RA120.014873>

Gustafsson, C., Govindarajan, S. & Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends in Biotechnology*, 22(7), 346–353. <https://doi.org/10.1016/j.tibtech.2004.04.006>

Hagag, I. T., Pyrc, K., Weber, S., Balkema-Buschmann, A., Groschup, M. H. & Keller, M. (2022). Mutations in SARS-CoV-2 nucleocapsid in variants of concern impair the sensitivity of SARS-CoV-2 detection by rapid antigen tests. *Frontiers in Virology*, 2 <https://doi.org/10.3389/fviro.2022.971862>

Han, B., Ren, C., Wang, W., Li, J. & Gong, X. (2023). Computational Prediction of Protein Intrinsically Disordered Region Related Interactions and Functions. *Genes*, 14(2), 432. <https://doi.org/10.3390/genes14020432>

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*, 166(4), 557–580. [https://doi.org/10.1016/S0022-2836\(83\)80284-8](https://doi.org/10.1016/S0022-2836(83)80284-8)

Hanson, S. W., Abbafati, C., Aerts, J. G., Al-Aly, Z., Ashbaugh, C., Ballouz, T., Blyuss, O., Bobkova, P., Bonsel, G., Borzakova, S., Buonsenso, D., Butnaru, D., Carter, A., Chu, H., De Rose, C., Diab, M. M., Ekbom, E., El Tantawi, M., Fomin, V., . . . Vos, T. (2022). Estimated Global Proportions of Individuals With Persistent Fatigue, Cognitive, and Respiratory Symptom Clusters Following Symptomatic COVID-19 in 2020 and 2021. *JAMA*, 328(16), 1604–1615. <https://doi.org/10.1001/jama.2022.18931>

Hardin, A. M., Harinarayan, C., Malmquist, G., Axén, A. & van Reis, R. (2009). Ion exchange chromatography of monoclonal antibodies: Effect of resin ligand density on dynamic binding capacity. *Journal of Chromatography. A*, 1216(20), 4366–4371. <https://doi.org/10.1016/j.chroma.2008.08.047>

Hatakeyama, D., Masuda, T., Miki, R., Ohtsuki, S. & Kuzuhara, T. (2021). In-vitro acetylation of SARS-CoV and SARS-CoV-2 nucleocapsid proteins by human PCAF and GCN5. *Biochemical and Biophysical Research Communications*, 557273–279. <https://doi.org/10.1016/j.bbrc.2021.03.173>

Hazra, S., Hossain, M. & Kumar, G. S. (2013). Binding of isoquinoline alkaloids berberine, palmatine and coralyne to hemoglobin : structural and thermodynamic characterization studies. *Molecular BioSystems*, 9(1), 143–153. <https://doi.org/10.1039/C2MB25345C>

He, Z., Yuan, J., Zhang, Y., Li, R., Mo, M., Wang, Y. & Ti, H. (2023). Recent advances towards natural plants as potential inhibitors of SARS-Cov-2 targets. *Pharmaceutical*

Biology, 61(1), 1186–1210.
<https://doi.org/https://doi.org/10.1080/13880209.2023.2241518>

Ho, Y., Lu, J., Huang, Y. & Lai, Z. (2019). Palmatine inhibits Zika virus infection by disrupting virus binding, entry, and stability. *Biochemical and Biophysical Research Communications*, 518(4), 732–738. <https://doi.org/10.1016/j.bbrc.2019.08.120>

Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R. & Stüber, D. (1988). Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent. *Bio/Technology*, 6(11), 1321–1325. <https://doi.org/10.1038/nbt1188-1321>

Hosseini, R. & Askari, N. (2023). A review of neurological side effects of COVID-19 vaccination. *European Journal of Medical Research*, 28(1), 102. <https://doi.org/10.1186/s40001-023-00992-0>

Huang, Y., Chen, J., Chen, S., Huang, C., Li, B., Li, J., Jin, Z., Zhang, Q., Pan, P., Du, W., Liu, L. & Liu, Z. (2024). Molecular characterization of SARS-CoV-2 nucleocapsid protein. *Frontiers in Cellular and Infection Microbiology*, 14 <https://doi.org/10.3389/fcimb.2024.1415885>

Image Lab Software | Bio-Rad Retrieved Apr 8, 2024, from <https://www.bio-rad.com/en-za/product/image-lab-software?ID=KRE6P5E8Z>

Ishihara, T. & Yamamoto, S. (2005). Optimization of monoclonal antibody purification by ion-exchange chromatography. Application of simple methods with linear gradient elution experimental data. *Journal of Chromatography. A*, 1069(1), 99–106. <https://doi.org/10.1016/j.chroma.2004.10.040>

Islam, M. M., Pandya, P., Chowdhury, S. R., Kumar, S. & Kumar, G. S. (2008). Binding of DNA-binding alkaloids berberine and palmatine to tRNA and comparison to ethidium:

Spectroscopic and molecular modeling studies. *Journal of Molecular Structure*, 891(1), 498–507. <https://doi.org/10.1016/j.molstruc.2008.04.043>

Itakura, K., Hirose, T., Crea, R., Riggs, A. D., Heyneker, H. L., Bolivar, F. & Boyer, H. W. (1977). Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science* (New York, N.Y.), 198(4321), 1056–1063. <https://doi.org/10.1126/science.412251>

Jia, F., Zou, G., Fan, J. & Yuan, Z. (2010). Identification of palmatine as an inhibitor of West Nile virus. *Archives of Virology*, 155(8), 1325–1329. <https://doi.org/10.1007/s00705-010-0702-4>

Joshi, A., Kaur, M., Kaur, R., Grover, A., Nash, D. & El-Mohandes, A. (2021). Predictors of COVID-19 Vaccine Acceptance, Intention, and Hesitancy: A Scoping Review. *Frontiers in Public Health*, 9698111. <https://doi.org/10.3389/fpubh.2021.698111>

Joshi, T., Joshi, T., Pundir, H., Sharma, P., Mathpal, S. & and Chandra, S. (2021). Predictive modeling by deep learning, virtual screening and molecular dynamics study of natural compounds against SARS-CoV-2 main protease. *Journal of Biomolecular Structure and Dynamics*, 39(17), 6728–6746. <https://doi.org/10.1080/07391102.2020.1802341>

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>

Kang, S., Yang, M., Hong, Z., Zhang, L., Huang, Z., Chen, X., He, S., Zhou, Z., Zhou, Z., Chen, Q., Yan, Y., Zhang, C., Shan, H. & Chen, S. (2020). Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting

sites. *Acta Pharmaceutica Sinica. B*, 10(7), 1228–1238.
<https://doi.org/10.1016/j.apsb.2020.04.009>

Kato, A. & Nakai, S. (1980). Hydrophobicity determined by a fluorescence probe method and its correlation with surface properties of proteins. *Biochimica Et Biophysica Acta*, 624(1), 13–20. [https://doi.org/10.1016/0005-2795\(80\)90220-2](https://doi.org/10.1016/0005-2795(80)90220-2)

Katoh, K., Rozewicki, J. & Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, 20(4), 1160–1166. <https://doi.org/10.1093/bib/bbx108>

Kaur, S. P. & Gupta, V. (2020). COVID-19 Vaccine: A comprehensive status report. *Virus Research*, 288198114. <https://doi.org/10.1016/j.virusres.2020.198114>

Khan, A. Y., Hossain, M. & Kumar, G. S. (2013). Binding of plant alkaloids berberine and palmatine to serum albumins: a thermodynamic investigation. *Molecular Biology Reports*, 40(1), 553–566. <https://doi.org/10.1007/s11033-012-2092-z>

Khan, A. Y. & Suresh Kumar, G. (2015). Natural isoquinoline alkaloids: binding aspects to functional proteins, serum albumins, hemoglobin, and lysozyme. *Biophysical Reviews*, 7(4), 407–420. <https://doi.org/10.1007/s12551-015-0183-5>

Kielkopf, C. L., Bauer, W. & Urbatsch, I. L. (2020a). Bradford Assay for Determining Protein Concentration. *Cold Spring Harbor Protocols*, 2020(4), 102269. <https://doi.org/10.1101/pdb.prot102269>

Kielkopf, C. L., Bauer, W. & Urbatsch, I. L. (2020b). Methods for Measuring the Concentrations of Proteins. *Cold Spring Harbor Protocols*, 2020(4), 102277. <https://doi.org/10.1101/pdb.top102277>

Kim, G., Weiss, S. J. & Levine, R. L. (2014). Methionine Oxidation and Reduction in Proteins. *Biochimica Et Biophysica Acta*, 1840(2), 10.1016/j.bbagen.2013.04.038. <https://doi.org/10.1016/j.bbagen.2013.04.038>

Kojima, M., Abe, S., Furuta, T., Hirata, K., Yao, X., Kobayashi, A., Kobayashi, R. & Ueno, T. (2024). High-throughput structure determination of an intrinsically disordered protein using cell-free protein crystallization. *Proceedings of the National Academy of Sciences of the United States of America*, 121(25), e2322452121. <https://doi.org/10.1073/pnas.2322452121>

Kolesov, D. E., Sinegubova, M. V., Safenkova, I. V., Vorobiev, I. I. & Orlova, N. A. (2022). Antigenic properties of the SARS-CoV-2 nucleoprotein are altered by the RNA admixture. *PeerJ*, 10e12751. <https://doi.org/10.7717/peerj.12751>

Kopecky-Bromberg, S. A., Martínez-Sobrido, L., Frieman, M., Baric, R. A. & Palese, P. (2007). Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists. *Journal of Virology*, 81(2), 548–557. <https://doi.org/10.1128/JVI.01782-06>

Krammer, A., Kirchhoff, P. D., Jiang, X., Venkatachalam, C. M. & Waldman, M. (2005). LigScore: a novel scoring function for predicting binding affinities. *Journal of Molecular Graphics & Modelling*, 23(5), 395–407. <https://doi.org/10.1016/j.jmgm.2004.11.007>

Ksiazek, T. G., Erdman, D., Goldsmith, C. S., Zaki, S. R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J. A., Lim, W., Rollin, P. E., Dowell, S. F., Ling, A., Humphrey, C. D., Shieh, W., Guarner, J., Paddock, C. D., Rota, P., Fields, B., . . . Anderson, L. J. (2003). A novel coronavirus associated with severe acute respiratory syndrome. *The New England Journal of Medicine*, 348(20), 1953–1966. <https://doi.org/10.1056/NEJMoa030781>

Kumar, N., Ippel, H., Weber, C., Hackeng, T. & Mayo, K. H. (2013). Protein lysine-N ζ alkylation and O-phosphorylation mediated by DTT-generated reactive oxygen species. *Protein Science: A Publication of the Protein Society*, 22(3), 327–346. <https://doi.org/10.1002/pro.2214>

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>

Lai, M. M. & Cavanagh, D. (1997). The molecular biology of coronaviruses. *Advances in Virus Research*, 481–100. [https://doi.org/10.1016/S0065-3527\(08\)60286-9](https://doi.org/10.1016/S0065-3527(08)60286-9)

Lakowics, J. R. (2006). Protein Fluorescence. In J. R. Lakowicz (Ed.), *Principles of Fluorescence Spectroscopy* (3rd ed., pp. 529–575). Springer US.

Langgartner, D., Winkler, R., Brunner-Weisser, J., Rohleder, N., Jarczok, M. N., Gündel, H., Weimer, K. & Reber, S. O. (2023). COVID-19 vaccination exacerbates ex vivo IL-6 release from isolated PBMCs. *Scientific Reports*, 139496. <https://doi.org/10.1038/s41598-023-35731-2>

Larentis, A. L., Nicolau, J. F. M. Q., Esteves, G. d. S., Vareschini, D. T., de Almeida, F. V. R., dos Reis, M. G., Galler, R. & Medeiros, M. A. (2014). Evaluation of pre-induction temperature, cell growth at induction and IPTG concentration on the expression of a leptospiral protein in *E. coli* using shaking flasks and microbioreactor. *BMC Research Notes*, 7671. <https://doi.org/10.1186/1756-0500-7-671>

Letko, M., Marzi, A. & Munster, V. (2020). Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nature Microbiology*, 5(4), 562–569. <https://doi.org/10.1038/s41564-020-0688-y>

Levi, V. & Flecha, F. L. G. (2002). Reversible fast-dimerization of bovine serum albumin detected by fluorescence resonance energy transfer. *Biochimica Et Biophysica Acta*, 1599(1-2), 141–148. [https://doi.org/10.1016/s1570-9639\(02\)00414-4](https://doi.org/10.1016/s1570-9639(02)00414-4)

Levin, A. T., Hanage, W. P., Owusu-Boaitey, N., Cochran, K. B., Walsh, S. P. & Meyerowitz-Katz, G. (2020). Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. *European Journal of Epidemiology*, 35(12), 1123–1138. <https://doi.org/10.1007/s10654-020-00698-1>

Lewis, E. A., & Murphy, K. P. (2005). Isothermal Titration Calorimetry. In G. Ulrich Nienhaus (Ed.), *Protein-Ligand Interactions: Methods and Applications* (pp. 1–15). Humana Press. <https://doi.org/10.1385/1-59259-912-5:001>

Li, F. Q., Xiao, H., Tam, J. P. & Liu, D. X. (2005). Sumoylation of the nucleocapsid protein of severe acute respiratory syndrome coronavirus. *FEBS Letters*, 579(11), 2387–2396. <https://doi.org/10.1016/j.febslet.2005.03.039>

Li, G., Li, W., Fang, X., Song, X., Teng, S., Ren, Z., Hu, D., Zhou, S., Wu, G. & Li, K. (2021). Expression and purification of recombinant SARS-CoV-2 nucleocapsid protein in inclusion bodies and its application in serological detection. *Protein Expression and Purification*, 186105908. <https://doi.org/10.1016/j.pep.2021.105908>

Li, W., Persson, B. A., Morin, M., Behrens, M. A., Lund, M. & Zackrisson Oskolkova, M. (2015). Charge-Induced Patchy Attractions between Proteins. *The Journal of Physical Chemistry B*, 119(2), 503–508. <https://doi.org/10.1021/jp512027j>

Li, Y., Li, M., Xiao, H., Liao, F., Shen, M., Ge, W., Ou, J., Liu, Y., Chen, L., Zhao, Y., Wan, P., Liu, J., Chen, J., Lan, X., Wu, S., Ding, Q., Li, G., Zhang, Q. & Pan, P. (2025). The R203M and D377Y mutations of the nucleocapsid protein promote SARS-CoV-2

infectivity by impairing RIG-I-mediated antiviral signaling. *PLOS Pathogens*, 21(1), e1012886. <https://doi.org/10.1371/journal.ppat.1012886>

Lim, S., Zhang, M. & Chang, T. L. (2022). ACE2-Independent Alternative Receptors for SARS-CoV-2. *Viruses*, 14(11), 2535. <https://doi.org/10.3390/v14112535>

Lipfert, J., Doniach, S., Das, R. & Herschlag, D. (2014). Understanding Nucleic Acid–Ion Interactions. *Annual Review of Biochemistry*, 83(Volume 83, 2014), 813–841. <https://doi.org/10.1146/annurev-biochem-060409-092720>

Lisyte, V., Kausaite-Minkstimiene, A., Brasiunas, B., Popov, A. & Ramanaviciene, A. (2024). Surface Plasmon Resonance Immunosensor for Direct Detection of Antibodies against SARS-CoV-2 Nucleocapsid Protein. *International Journal of Molecular Sciences*, 25(16), 8574. <https://doi.org/10.3390/ijms25168574>

Little, D. R., MacLachlan, B. J., Watson, G. M., Vivian, J. P. & Gully, B. S. (2020). A pocket guide on how to structure SARS-CoV-2 drugs and therapies. *Biochemical Society Transactions*, 48(6), 2625–2641. <https://doi.org/10.1042/BST20200396>

Liu, X., Verma, A., Garcia, G. J., Ramage, H., Lucas, A., Myers, R. L., Michaelson, J. J., Coryell, W., Kumar, A., Charney, A. W., Kazanietz, M. G., Rader, D. J., Ritchie, M. D., Berrettini, W. H., Schultz, D. C., Cherry, S., Damoiseaux, R., Arumugaswami, V. & Klein, P. S. (2021). Targeting the coronavirus nucleocapsid protein through GSK-3 inhibition. *Proceedings of the National Academy of Sciences of the United States of America*, 118(42), e2113401118. doi: 10.1073/pnas.2113401118. Epub 2021 Sep 30. <https://doi.org/10.1073/pnas.2113401118>

LLC, G. S.pET-15b Sequence and Map. [https://www.snapgene.com/plasmids/pet_and_duet_vectors_\(novagen\)/pET-15b](https://www.snapgene.com/plasmids/pet_and_duet_vectors_(novagen)/pET-15b). Retrieved May 11, 2026, from [https://www.snapgene.com/plasmids/pet_and_duet_vectors_\(novagen\)/pET-15b](https://www.snapgene.com/plasmids/pet_and_duet_vectors_(novagen)/pET-15b)

Long, J., Song, J., Zhong, L., Liao, Y., Liu, L. & Li, X. (2019). Palmatine: A review of its pharmacology, toxicity and pharmacokinetics. *Biochimie*, 162176–184. <https://doi.org/10.1016/j.biochi.2019.04.008>

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), 502. <https://doi.org/10.1038/s41467-020-20768-y>

Luan, X., Li, X., Li, Y., Su, G., Yin, W., Jiang, Y., Xu, N., Wang, F., Cheng, W., Jin, Y., Zhang, L., Xu, H. E., Xue, Y. & Zhang, S. (2022). Antiviral drug design based on structural insights into the N-terminal domain and C-terminal domain of the SARS-CoV-2 nucleocapsid protein. *Science Bulletin*, 67(22), 2327–2335. <https://doi.org/10.1016/j.scib.2022.10.021>

Luo, C., Luo, H., Zheng, S., Gui, C., Yue, L., Yu, C., Sun, T., He, P., Chen, J., Shen, J., Luo, X., Li, Y., Liu, H., Bai, D., Shen, J., Yang, Y., Li, F., Zuo, J., Hilgenfeld, R., . . . Jiang, H. (2004). Nucleocapsid protein of SARS coronavirus tightly binds to human cyclophilin A. *Biochemical and Biophysical Research Communications*, 321(3), 557–565. <https://doi.org/10.1016/j.bbrc.2004.07.003>

Madahar, V., Dang, R., Zhang, Q., Liu, C., Rodgers, V. G. J. & Liao, J. (2023). Human Post-Translational SUMOylation Modification of SARS-CoV-2 Nucleocapsid Protein Enhances Its Interaction Affinity with Itself and Plays a Critical Role in Its Nuclear Translocation. *Viruses*, 15(7), 1600. <https://doi.org/10.3390/v15071600>

Madeira, F., Madhusoodanan, N., Lee, J., Eusebi, A., Niewielska, A., Tivey, A. R. N., Lopez, R. & Butcher, S. (2024). The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Research*, 52(W1), W521–W525. <https://doi.org/10.1093/nar/gkae241>

Maiti, S., Singh, A., Maji, T., Saibo, N. V. & De, S. (2024). Experimental methods to study the structure and dynamics of intrinsically disordered regions in proteins. *Current Research in Structural Biology*, 7100138. <https://doi.org/10.1016/j.crstbi.2024.100138>

Majnooni, M. B., Fakhri, S., Bahrami, G., Naseri, M., Farzaei, M. H. & Echeverría, J. (2021). Alkaloids as Potential Phytochemicals against SARS-CoV-2: Approaches to the Associated Pivotal Mechanisms. *Evidence-Based Complementary and Alternative Medicine: eCAM*, 20216632623. <https://doi.org/10.1155/2021/6632623>

Mak, G. C. K., Lau, S. S. Y., Wong, K. K. Y., Lau, C. S., Lam, E. T. K., Ng, K. H. L. & Chan, R. C. W. (2022). Evaluation of Rapid Antigen Detection Kits for Detection of SARS-CoV-2 B.1.1.529. *Dr. Sulaiman Al Habib Medical Journal*, 4(2), 99–102. <https://doi.org/10.1007/s44229-022-00009-8>

Malakar, P. & Venkatesh, K. V. (2012). Effect of substrate and IPTG concentrations on the burden to growth of *Escherichia coli* on glycerol due to the expression of Lac proteins. *Applied Microbiology and Biotechnology*, 93(6), 2543–2549. <https://doi.org/10.1007/s00253-011-3642-3>

Maleki, B., Sadeghian, A. M. & Ranjbar, M. (2025). Impact of vaccination against SARS-CoV-2 on mortality risk, ICU admission rate, and hospitalization length in hospitalized COVID-19 patients: a cross-sectional study. *BMC Infectious Diseases*, 25(1), 144. <https://doi.org/10.1186/s12879-025-10530-4>

Mandel, M. & Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology*, 53(1), 159–162. [https://doi.org/10.1016/0022-2836\(70\)90051-3](https://doi.org/10.1016/0022-2836(70)90051-3)

Matsuo, T. (2021). Viewing SARS-CoV-2 Nucleocapsid Protein in Terms of Molecular Flexibility. *Biology*, 10(6), 454. <https://doi.org/10.3390/biology10060454>

McBride, R., van Zyl, M. & Fielding, B. C. (2014). The coronavirus nucleocapsid is a multifunctional protein. *Viruses*, 6(8), 2991–3018. <https://doi.org/10.3390/v6082991>

Meo, S. A., Bukhari, I. A., Akram, J., Meo, A. S. & Klonoff, D. C. (2021). COVID-19 vaccines: comparison of biological, pharmacological characteristics and adverse effects of Pfizer/BioNTech and Moderna Vaccines. *European Review for Medical and Pharmacological Sciences*, 25(3), 1663–1669. https://doi.org/10.26355/eurev_202102_24877

Mercaldi, G. F., Bezerra, E. H. S., Batista, F. A. H., Tonoli, C. C. C., Soprano, A. S., Shimizu, J. F., Nagai, A., da Silva, J. C., Filho, H. V. R., do Nascimento Faria, J., da Cunha, M. G., Zeri, A. C. M., Nascimento, A. F. Z., Proenca-Modena, J. L., Bajgelman, M. C., Rocco, S. A., Lopes-de-Oliveira, P. S., Cordeiro, A. T., Bruder, M., . . . Trivella, D. B. B. (2022). Discovery and structural characterization of chicoric acid as a SARS-CoV-2 nucleocapsid protein ligand and RNA binding disruptor. *Scientific Reports*, 12(1), 18500. <https://doi.org/10.1038/s41598-022-22576-4>

Micsonai, A., Wien, F., Kernya, L., Lee, Y., 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 of the United States of America*, 112(24), 3095. <https://doi.org/10.1073/pnas.1500851112>

Mignon, K., Galle, M., Van der Eecken, R., Haesaerts, S., Demulder, M., De Greve, H., De Veylder, L. & Loris, R. (2025). Purification and characterization of the intrinsically disordered *Arabidopsis thaliana* protein SOG1. *Protein Expression and Purification*, 229106678. <https://doi.org/10.1016/j.pep.2025.106678>

Millet, J. K. & Whittaker, G. R. (2015). Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. *Virus Research*, 202120–134. <https://doi.org/10.1016/j.virusres.2014.11.021>

Momany, F. A. & Rone, R. (1992). Validation of the general purpose QUANTA ®3.2/CHARMm® force field. *Journal of Computational Chemistry*, 13(7), 888–900. <https://doi.org/rg/10.1002/jcc.540130714>

Moon, J. M., Ratliff, K. M., Hagele, A. M., Stecker, R. A., Mumford, P. W. & Kerksick, C. M. (2021). Absorption Kinetics of Berberine and Dihydroberberine and Their Impact on Glycemia: A Randomized, Controlled, Crossover Pilot Trial. *Nutrients*, 14(1), 124. <https://doi.org/10.3390/nu14010124>

Moreno, S., Alcázar, B., Dueñas, C., González Del Castillo, J., Olalla, J. & Antela, A. (2022). Use of Antivirals in SARS-CoV-2 Infection. Critical Review of the Role of Remdesivir. *Drug Design, Development and Therapy*, 16827–841. <https://doi.org/10.2147/DDDT.S356951>

Mourier, T., Shuaib, M., Hala, S., Mfarrej, S., Alofi, F., Naeem, R., Alsomali, A., Jorgensen, D., Subudhi, A. K., Ben Rached, F., Guan, Q., Salunke, R. P., Ooi, A., Esau, L., Douvropoulou, O., Nugmanova, R., Perumal, S., Zhang, H., Rajan, I., . . . Pain, A. (2022). SARS-CoV-2 genomes from Saudi Arabia implicate nucleocapsid mutations in host response and increased viral load. *Nature Communications*, 13(1), 601. <https://doi.org/10.1038/s41467-022-28287-8>

Mu, J., Fang, Y., Yang, Q., Shu, T., Wang, A., Huang, M., Jin, L., Deng, F., Qiu, Y. & Zhou, X. (2020). SARS-CoV-2 N protein antagonizes type I interferon signaling by suppressing phosphorylation and nuclear translocation of STAT1 and STAT2. *Cell Discovery*, 6(1), 1–4. <https://doi.org/10.1038/s41421-020-00208-3>

Muegge, I. & Martin, Y. C. (1999). A general and fast scoring function for protein-ligand interactions: a simplified potential approach. *Journal of Medicinal Chemistry*, 42(5), 791–804. <https://doi.org/10.1021/jm980536j>

Munishkina, L. A. & Fink, A. L. (2007). Fluorescence as a method to reveal structures and membrane-interactions of amyloidogenic proteins. *Biochimica Et Biophysica Acta (BBA) - Biomembranes*, 1768(8), 1862–1885. <https://doi.org/10.1016/j.bbamem.2007.03.015>

Muradyan, N., Arakelov, V., Sargsyan, A., Paronyan, A., Arakelov, G. & Nazaryan, K. (2024). Impact of mutations on the stability of SARS-CoV-2 nucleocapsid protein structure. *Scientific Reports*, 14(1), 5870. <https://doi.org/10.1038/s41598-024-55157-8>

Murray-Rust, P. & Glusker, J. P. (1984). Directional hydrogen bonding to sp²- and sp³-hybridized oxygen atoms and its relevance to ligand-macromolecule interactions. *Journal of the American Chemical Society*, 106(4), 1018–1025. <https://doi.org/10.1021/ja00316a034>

Nag, N.(et al.). (2022). Strategies to improve the expression and solubility of recombinant proteins in *E. coli*. In T. Tripathi& V. K. Dubey (Eds.), *Advances in Protein Molecular and Structural Biology Methods* (pp. 1–12). Academic Press. <https://doi.org/10.1016/B978-0-323-90264-9.00001-5>

Narkhede, R. R., Cheke, R. S., Ambhore, J. P. & Shinde, S. D. (2020). The Molecular Docking Study of Potential Drug Candidates Showing Anti-COVID-19 Activity by Exploring of Therapeutic Targets of SARS-CoV-2. *Eurasian Journal of Medicine and Oncology*, 4(3), 185–195. <https://doi.org/10.14744/ejmo.2020.31503>

NCBI.NCBI Virus. <https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/>. Retrieved Apr 21, 2026, from <https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/>

Nguyen, A., Zhao, H., Myagmarsuren, D., Srinivasan, S., Wu, D., Chen, J., Piszczek, G. & Schuck, P. (2024). Modulation of Biophysical Properties of Nucleocapsid Protein in the Mutant Spectrum of SARS-CoV-2. *eLife*, 13 <https://doi.org/10.7554/eLife.94836.2>

Nicoya. (2023,). Become an SPR expert with our 10 tips for success.
<https://nicoyalife.com/blog/spr-tips/top-10-tips-for-high-quality-spr-data/>

Nikolakaki, E. & Giannakouros, T. (2020). SR/RS Motifs as Critical Determinants of Coronavirus Life Cycle. *Frontiers in Molecular Biosciences*, 7219.
<https://doi.org/10.3389/fmolb.2020.00219>

Novagen. Novagen pET System Manual.
https://kirschner.med.harvard.edu/files/protocols/Novagen_petsystem.pdf. Retrieved
May 16, 2025, from
https://kirschner.med.harvard.edu/files/protocols/Novagen_petsystem.pdf

O'Toole, Á, Scher, E., Underwood, A., Jackson, B., Hill, V., McCrone, J. T., Colquhoun, R., Ruis, C., Abu-Dahab, K., Taylor, B., Yeats, C., du Plessis, L., Maloney, D., Medd, N., Attwood, S. W., Aanensen, D. M., Holmes, E. C., Pybus, O. G. & Rambaut, A. (2021). Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. *Virus Evolution*, 7(2), veab064. <https://doi.org/10.1093/ve/veab064>

Obermeyer, F., Jankowiak, M., Barkas, N., Schaffner, S. F., Pyle, J. D., Yurkovetskiy, L., Bosso, M., Park, D. J., Babadi, M., MacInnis, B. L., Luban, J., Sabeti, P. C. & Lemieux, J. E. (2022). Analysis of 6.4 million SARS-CoV-2 genomes identifies mutations associated with fitness. *Science (New York, N.Y.)*, 376(6599), 1327–1332.
<https://doi.org/10.1126/science.abm1208>

Oliva, R., Mukherjee, S., Manisegaran, M., Campanile, M., Del Vecchio, P., Petraccone, L. & Winter, R. (2022). Binding Properties of RNA Quadruplex of SARS-CoV-2 to Berberine Compared to Telomeric DNA Quadruplex. *International Journal of Molecular Sciences*, 23(10), 5690. <https://doi.org/10.3390/ijms23105690>

Oner, E., Al-Khafaji, K., Mezher, M. H., Demirhan, I., Suhail Wadi, J., Belge Kurutas, E., Yalin, S. & Choowongkamon, K. (2023). Investigation of berberine and its derivatives in

Sars Cov-2 main protease structure by molecular docking, PROTOX-II and ADMET methods: in machine learning and in silico study. *Journal of Biomolecular Structure & Dynamics*, 41(19), 9366–9381. <https://doi.org/10.1080/07391102.2022.2142848>

Oronsky, B., Larson, C., Caroen, S., Hedjran, F., Sanchez, A., Prokopenko, E. & Reid, T. (2022). Nucleocapsid as a next-generation COVID-19 vaccine candidate. *International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases*, 122529–530. <https://doi.org/10.1016/j.ijid.2022.06.046>

O'Toole, Á, Hill, V., Pybus, O. G., Watts, A., Bogoch, I. I., Khan, K., Messina, J. P., Tegally, H., Lessells, R. R., Giandhari, J., Pillay, S., Tumedi, K. A., Nyepetsi, G., Kebabonye, M., Matsheka, M., Mine, M., Tokajian, S., Hassan, H., Salloum, T., . . . Kraemer, M. U. G. (2021). Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2 with grinch. *Wellcome Open Research*, 6121. <https://doi.org/10.12688/wellcomeopenres.16661.2>

Oulas, A., Zanti, M., Tomazou, M., Zachariou, M., Minadakis, G., Bourdakou, M. M., Pavlidis, P. & Spyrou, G. M. (2021). Generalized linear models provide a measure of virulence for specific mutations in SARS-CoV-2 strains. *PloS One*, 16(1), e0238665. <https://doi.org/10.1371/journal.pone.0238665>

Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Science: A Publication of the Protein Society*, 4(11), 2411–2423. <https://doi.org/10.1002/pro.5560041120>

Padmapriya, K. & Barthwal, R. (2018). Structural and biophysical insight into dual site binding of the protoberberine alkaloid palmatine to parallel G-quadruplex DNA using NMR, fluorescence and Circular Dichroism spectroscopy. *Biochimie*, 147153–169. <https://doi.org/10.1016/j.biochi.2018.02.002>

Panigrahi, S. K. & Mishra, A. K. (2019). Inner filter effect in fluorescence spectroscopy: As a problem and as a solution. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews*, 41100318. <https://doi.org/10.1016/j.jphotochemrev.2019.100318>

Parker, C. E.(et al.). (2010). *Mass Spectrometry for Post-Translational Modifications*. In O. Alzate (Ed.), *Neuroproteomics* . CRC Press/Taylor & Francis.

Pei, H., Zeng, J., Chen, W., He, Z. & Du, R. (2022). Network pharmacology and molecular docking integrated strategy to investigate the pharmacological mechanism of palmatine in Alzheimer's disease. *Journal of Biochemical and Molecular Toxicology*, 36(11), e23200. <https://doi.org/10.1002/jbt.23200>

Pei, H., Zeng, J., He, Z., Zong, Y., Zhao, Y., Li, J., Chen, W. & Du, R. (2023). Palmatine ameliorates LPS-induced HT-22 cells and mouse models of depression by regulating apoptosis and oxidative stress. *Journal of Biochemical and Molecular Toxicology*, 37(1), e23225. <https://doi.org/10.1002/jbt.23225>

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/emboj.2020105938>

Phan, T. (2020). Novel coronavirus: From discovery to clinical diagnostics. *Infection, Genetics and Evolution*, 79104211. <https://doi.org/10.1016/j.meegid.2020.104211>

Porterfield, J. Z. & Zlotnick, A. (2010). A simple and general method for determining the protein and nucleic acid content of viruses by UV absorbance. *Virology*, 407(2), 281–288. <https://doi.org/10.1016/j.virol.2010.08.015>

Purwaningsih, I., Maksum, I. P., Sumiarsa, D. & Sriwidodo, S. (2024). Isolation and Quantification of Palmatine from *Fibraurea tinctoria* Lour: In Vitro Antioxidant Activity and

In Silico Antidiabetic Activity Evaluation. *Drug Design, Development and Therapy*, 183443–3459. <https://doi.org/10.2147/DDDT.S454091>

Rahman, M. S., Islam, M. R., Alam, A. S. M. R. U., Islam, I., Hoque, M. N., Akter, S., Rahaman, M. M., Sultana, M. & Hossain, M. A. (2021). Evolutionary dynamics of SARS-CoV-2 nucleocapsid protein and its consequences. *Journal of Medical Virology*, 93(4), 2177–2195. <https://doi.org/10.1002/jmv.26626>

Rak, A., Isakova-Sivak, I. & Rudenko, L. (2023). Overview of Nucleocapsid-Targeting Vaccines against COVID-19. *Vaccines*, 11(12), 1810. <https://doi.org/10.3390/vaccines11121810>

Ranjbar, B. & Gill, P. (2009). Circular dichroism techniques: biomolecular and nanostructural analyses- a review. *Chemical Biology & Drug Design*, 74(2), 101–120. <https://doi.org/10.1111/j.1747-0285.2009.00847.x>

Ribeiro-Filho, H. V., Jara, G. E., Batista, F. A. H., Schleder, G. R., Tonoli, C. C. C., Soprano, A. S., Guimarães, S. L., Borges, A. C., Cassago, A., Bajgelman, M. C., Marques, R. E., Trivella, D. B. B., Franchini, K. G., Figueira, A. C. M., Benedetti, C. E. & Lopes-de-Oliveira, P. S. (2022). Structural dynamics of SARS-CoV-2 nucleocapsid protein induced by RNA binding. *PLOS Computational Biology*, 18(5), e1010121. <https://doi.org/10.1371/journal.pcbi.1010121>

Rice, P., Longden, I. & Bleasby, A. (2000). EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics*, 16(6), 276–277. [https://doi.org/10.1016/S0168-9525\(00\)02024-2](https://doi.org/10.1016/S0168-9525(00)02024-2)

Robichon, C., Luo, J., Causey, T. B., Benner, J. S. & Samuelson, J. C. (2011). Engineering Escherichia coli BL21(DE3) Derivative Strains To Minimize E. coli Protein Contamination after Purification by Immobilized Metal Affinity Chromatography. *Applied and Environmental Microbiology*, 77(13), <https://doi.org/10.1128/AEM.00119-11>

Rogers, D. M., Jasim, S. B., Dyer, N. T., Auvray, F., Réfrégiers, M. & Hirst, J. D. (2019). Electronic Circular Dichroism Spectroscopy of Proteins. *Chem*, 5(11), 2751–2774. <https://doi.org/10.1016/j.chempr.2019.07.008>

Romero, P., Obradovic, Z., Li, X., Garner, E. C., Brown, C. J. & Dunker, A. K. (2001). Sequence complexity of disordered protein. *Proteins*, 42(1), 38–48. [https://doi.org/10.1002/1097-0134\(20010101\)42:1<38::aid-prot50>3.0.co;2-3](https://doi.org/10.1002/1097-0134(20010101)42:1<38::aid-prot50>3.0.co;2-3)

Royster, A., Ren, S., Ma, Y., Pintado, M., Kahng, E., Rowan, S., Mir, S. & Mir, M. (2023). SARS-CoV-2 Nucleocapsid Protein Is a Potential Therapeutic Target for Anticoronavirus Drug Discovery. *Microbiology Spectrum*, 11(3), e0118623. <https://doi.org/10.1128/spectrum.01186-23>

Rozenchwajg, S., Blet, A., Lamer, A., Boisson, M., Clavier, T. & Abou-Arab, O. (2021). SARS-CoV-2 vaccination efficacy on hospitalisation and variants. *Anaesthesia, Critical Care & Pain Medicine*, 100867. <https://doi.org/10.1016/j.accpm.2021.100867>

Ruff, K. M. & Pappu, R. V. (2021). AlphaFold and Implications for Intrinsically Disordered Proteins. *Journal of Molecular Biology*, 433(20), 167208. <https://doi.org/10.1016/j.jmb.2021.167208>

Ruiz, N., Jarosova, P., Taborsky, P. & Gargallo, R. (2022). Study of the interaction of the palmatine alkaloid with hybrid G-quadruplex/duplex and i-motif/duplex DNA structures. *Biophysical Chemistry*, 281106715. <https://doi.org/10.1016/j.bpc.2021.106715>

Salamanna, F., Maglio, M., Landini, M. P. & Fini, M. (2020). Body Localization of ACE-2: On the Trail of the Keyhole of SARS-CoV-2. *Frontiers in Medicine*, 7594495. <https://doi.org/10.3389/fmed.2020.594495>

Samad, A., Hamza, A., Imam, M. A., Ahmad Chaudhary, A., Alawam, A. S., Abdullah Almuqri, E., Islam, A. & Parveen, S. (2024). pH induced structural and conformational changes in nucleocapsid protein leads to intermediate like conformation: a biophysical and computational approach. *Journal of Biomolecular Structure and Dynamics*, 0(0), 1–12. <https://doi.org/10.1080/07391102.2024.2442791>

Santarino, I. B., Oliveira, S. C. B. & Oliveira-Brett, A. M. (2012). Protein reducing agents dithiothreitol and tris(2-carboxyethyl)phosphine anodic oxidation. *Electrochemistry Communications*, 23114–117. <https://doi.org/10.1016/j.elecom.2012.06.027>

Sarkar, M. & Madabhavi, I. (2022). SARS-CoV-2 variants of concern: a review. *Monaldi Archives for Chest Disease = Archivio Monaldi Per Le Malattie Del Torace*, 93(3), <https://doi.org/10.4081/monaldi.2022.2337>

Sarkar, S., Runge, B., Russell, R. W., Movellan, K. T., Calero, D., Zeinalilathori, S., Quinn, C. M., Lu, M., Calero, G., Gronenborn, A. M. & Polenova, T. (2022). Atomic-Resolution Structure of SARS-CoV-2 Nucleocapsid Protein N-Terminal Domain. *Journal of the American Chemical Society*, 144(23), 10543–10555. <https://doi.org/10.1021/jacs.2c03320>

SARS-CoV-2 variants of concern as of 25 April 2025 <https://www.ecdc.europa.eu>. Retrieved May 15, 2025, from <https://www.ecdc.europa.eu/en/covid-19/variants-concern>

Sato, R., Tomioka, Y., Sakuma, C., Nakagawa, M., Kurosawa, Y., Shiba, K., Arakawa, T. & Akuta, T. (2023). Detection of concentration-dependent conformational changes in SARS-CoV-2 nucleoprotein by agarose native gel electrophoresis. *Analytical Biochemistry*, 662114995. <https://doi.org/10.1016/j.ab.2022.114995>

Savellini, G. G., Anichini, G. & Cusi, M. G. (2023). SARS-CoV-2 omicron sub-lineages differentially modulate interferon response in human lung epithelial cells. *Virus Research*, 332199134. <https://doi.org/10.1016/j.virusres.2023.199134>

Scheich, C., Sievert, V. & Büssow, K. (2003). An automated method for high-throughput protein purification applied to a comparison of His-tag and GST-tag affinity chromatography. *BMC Biotechnology*, 312. <https://doi.org/10.1186/1472-6750-3-12>

Schiavina, M., Pontoriero, L., Tagliaferro, G., Pierattelli, R. & Felli, I. C. (2022). The Role of Disordered Regions in Orchestrating the Properties of Multidomain Proteins: The SARS-CoV-2 Nucleocapsid Protein and Its Interaction with Enoxaparin. *Biomolecules*, 12(9), 1302. <https://doi.org/10.3390/biom12091302>

Schramm, A., Bignon, C., Brocca, S., Grandori, R., Santambrogio, C. & Longhi, S. (2019). An arsenal of methods for the experimental characterization of intrinsically disordered proteins – How to choose and combine them? *Archives of Biochemistry and Biophysics*, 676108055. <https://doi.org/10.1016/j.abb.2019.07.020>

Schubert, M., Bertoglio, F., Steinke, S., Heine, P. A., Ynga-Durand, M. A., Maass, H., Sammartino, J. C., Cassaniti, I., Zuo, F., Du, L., Korn, J., Milošević, M., Wenzel, E. V., Krstanović, F., Polten, S., Pribanić-Matešić, M., Brizić, I., Baldanti, F., Hammarström, L., . . . Hust, M. (2022). Human serum from SARS-CoV-2-vaccinated and COVID-19 patients shows reduced binding to the RBD of SARS-CoV-2 Omicron variant. *BMC Medicine*, 20(1), 102. <https://doi.org/10.1186/s12916-022-02312-5>

Schultheiß, C., Willscher, E., Paschold, L., Gottschick, C., Klee, B., Henkes, S., Bosurgi, L., Dutzmann, J., Sedding, D., Frese, T., Girndt, M., Höll, J., Gekle, M., Mikolajczyk, R. & Binder, M. (2022). The IL-1 β , IL-6, and TNF cytokine triad is associated with post-acute sequelae of COVID-19 - PMC. *Cell Reports Medicine*, 3(6), <https://doi.org/10.1016%2Fj.xcrm.2022.100663>

Shenkin, P. S., Yarmush, D. L., Fine, R. M., Wang, H. J. & Levinthal, C. (1987). Predicting antibody hypervariable loop conformation. I. Ensembles of random conformations for

ringlike structures. Biopolymers, 26(12), 2053–2085.
<https://doi.org/10.1002/bip.360261207>

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, 1(6), 2856–2860. <https://doi.org/10.1038/nprot.2006.468>

Shuaib, M., Adroub, S., Mourier, T., Mfarrej, S., Zhang, H., Esau, L., Alsomali, A., Alofi, F. S., Ahmad, A. N., Shamsan, A., Khogeer, A., Hashem, A. M., Almontashiri, N. A. M., Hala, S. & Pain, A. (2023). Impact of the SARS-CoV-2 nucleocapsid 203K/204R mutations on the inflammatory immune response in COVID-19 severity. *Genome Medicine*, 15(1), 54. <https://doi.org/10.1186/s13073-023-01208-0>

Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J. & Li, Q. (2009). Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Science: A Publication of the Protein Society*, 18(5), 936–948. <https://doi.org/10.1002/pro.102>

Sparks, S., Hayama, R., Rout, M. P. & Cowburn, D. (2020). Analysis of Multivalent IDP Interactions: Stoichiometry, Affinity, and Local Concentration Effect Measurements. *Methods in Molecular Biology* (Clifton, N.J.), 2141463–475. https://doi.org/10.1007/978-1-0716-0524-0_23

Stuart, B. (2006). FTIR of Biomolecules. *Reviews in Cell Biology and Molecular Medicine*, <https://doi.org/10.1002/3527600906.mcb.200300059>

Stuart, B. H. (1997). In Ando D. J. (Ed.), *Biological Applications of Infrared Spectroscopy*. John Wiley & Sons.

Studier, F. W. & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology*, 189(1), 113–130. [https://doi.org/10.1016/0022-2836\(86\)90385-2](https://doi.org/10.1016/0022-2836(86)90385-2)

Stuwe, H., Reardon, P. N., Yu, Z., Shah, S., Hughes, K. & Barbar, E. J. (2024). Phosphorylation in the Ser/Arg-rich region of the nucleocapsid of SARS-CoV-2 regulates phase separation by inhibiting self-association of a distant helix. *The Journal of Biological Chemistry*, 300(6), 107354. <https://doi.org/10.1016/j.jbc.2024.107354>

Sun, Z., Zheng, X., Ji, F., Zhou, M., Su, X., Ren, K. & Li, L. (2021). Mass Spectrometry Analysis of SARS-CoV-2 Nucleocapsid Protein Reveals Camouflaging Glycans and Unique Post-Translational Modifications. *Infectious Microbes & Diseases*, 3(3), 149. <https://doi.org/10.1097/IM9.0000000000000071>

Supekar, N. T., Shajahan, A., Gleinich, A. S., Rouhani, D. S., Heiss, C., Chapla, D. G., Moremen, K. W. & Azadi, P. (2021). Variable posttranslational modifications of severe acute respiratory syndrome coronavirus 2 nucleocapsid protein. *Glycobiology*, 31(9), 1080–1092. <https://doi.org/10.1093/glycob/cwab044>

Surewicz, W. K., Mantsch, H. H. & Chapman, D. (1993). Determination of protein secondary structure by Fourier transform infrared spectroscopy: a critical assessment. *Biochemistry*, 32(2), 389–394. <https://doi.org/10.1021/bi00053a001>

Surjit, M., & Lal, S. K. (2010). The Nucleocapsid Protein of the SARS Coronavirus: Structure, Function and Therapeutic Potential. In S. K. Lal (Ed.), *Molecular Biology of the SARS-Coronavirus* (pp. 129–151). Springer. https://doi.org/10.1007/978-3-642-03683-5_9

Tarczewska, A., Kolonko-Adamska, M., Zarębski, M., Dobrucki, J., Ożyhar, A. & Greb-Markiewicz, B. (2021). The method utilized to purify the SARS-CoV-2 N protein can affect

its molecular properties. *International Journal of Biological Macromolecules*, 188391–403. <https://doi.org/10.1016/j.ijbiomac.2021.08.026>

Terpe, K. (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*, 60(5), 523–533. <https://doi.org/10.1007/s00253-002-1158-6>

Tirado-Rives, J. & Jorgensen, W. L. (2006). Contribution of conformer focusing to the uncertainty in predicting free energies for protein-ligand binding. *Journal of Medicinal Chemistry*, 49(20), 5880–5884. <https://doi.org/10.1021/jm060763i>

Troyano-Hernaez, P., Reinosa, R. & Holguin, A. (2021). Evolution of SARS-CoV-2 Envelope, Membrane, Nucleocapsid, and Spike Structural Proteins from the Beginning of the Pandemic to September 2020: A Global and Regional Approach by Epidemiological Week. *Viruses*, 13(2), 243. <https://doi.org/10.1002/jmv.26626>

Ughy, B., Nagyapati, S., Lajko, D. B., Letoha, T., Prohaszka, A., Deeb, D., Der, A., Pettko-Szandtner, A. & Szilak, L. (2023). Reconsidering Dogmas about the Growth of Bacterial Populations. *Cells*, 12(10), 1430. <https://doi.org/10.3390/cells12101430>

Uversky. (2019). Intrinsically Disordered Proteins and Their “Mysterious” (Meta)Physics. *Frontiers in Physics*, 7 <https://doi.org/10.3389/fphy.2019.00010>

Uversky. (2025). A blurry view of fuzzy objects: on the roles of low-resolution structural techniques in discovery and early characterization of intrinsically disordered proteins. *Frontiers in Biophysics*, 3 <https://doi.org/10.3389/frbis.2025.1693360>

Uversky, Gillespie, J. R. & Fink, A. L. (2000). Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins*, 41(3), 415–427. [https://doi.org/10.1002/1097-0134\(20001115\)41:3<415::aid-prot130>3.0.co;2-7](https://doi.org/10.1002/1097-0134(20001115)41:3<415::aid-prot130>3.0.co;2-7)

Van Straaten, J. & Peppas, N. A. (1991). ATR-FTIR analysis of protein adsorption on polymeric surfaces. *Journal of Biomaterials Science, Polymer Edition*, 2(2), 113–121. <https://doi.org/10.1163/156856291X00106>

van Wyk, B. -. (2008). A broad review of commercially important southern African medicinal plants. *Journal of Ethnopharmacology*, 119(3), 342–355. <https://doi.org/10.1016/j.jep.2008.05.029>

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Žídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., . . . Velankar, S. (2021). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, 50(D1), D439–D444. <https://doi.org/10.1093/nar/gkab1061>

Varghese, F. S., van Woudenberg, E., Overheul, G. J., Eleveld, M. J., Kurver, L., van Heerbeek, N., van Laarhoven, A., Miesen, P., den Hartog, G., de Jonge, M. I. & van Rij, R. P. (2021). Berberine and Obatoclax Inhibit SARS-Cov-2 Replication in Primary Human Nasal Epithelial Cells In Vitro. *Viruses*, 13(2), 282. <https://doi.org/10.3390/v13020282>

Velazquez-Campoy, A., Leavitt, S. A. & Freire, E. (2004). Characterization of protein-protein interactions by isothermal titration calorimetry. *Methods in Molecular Biology* (Clifton, N.J.), 26135–54. <https://doi.org/10.1385/1-59259-762-9:035>

Velázquez-Campoy, A., Ohtaka, H., Nezami, A., Muzammil, S. & Freire, E. (2004). Isothermal titration calorimetry. *Current Protocols in Cell Biology*, Chapter 17Unit 17.8. <https://doi.org/10.1002/0471143030.cb1708s23>

Vidal, L. E. L., Figueira-Mansur, J., Jurgilas, P. B., Argondizzo, A. P. C., Pestana, C. P., Martins, F. O., da Silva Junior, H. C., Miguez, M., Loureiro, B. O., Marques, C. d. F. S., Trinta, K. S., da Silva, L. B. R., de Mello, M. B., da Silva, E. D., Bastos, R. C. & Esteves,

G. (2023). Process development and characterization of recombinant nucleocapsid protein for its application on COVID-19 diagnosis. *Protein Expression and Purification*, 207106263. <https://doi.org/10.1016/j.pep.2023.106263>

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](https://doi.org/10.1016/S0006-3495(01)76183-8)

V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. (2021). Coronavirus biology and replication: implications for SARS-CoV-2. *Nature Reviews. Microbiology*, 19(3), 155–170. <https://doi.org/10.1038/s41579-020-00468-6>

Vogt, W. (1995). Oxidation of methionyl residues in proteins: tools, targets, and reversal. *Free Radical Biology & Medicine*, 18(1), 93–105. [https://doi.org/10.1016/0891-5849\(94\)00158-g](https://doi.org/10.1016/0891-5849(94)00158-g)

Wang, Z., Li, K., Maskey, A. R., Huang, W., Toutov, A. A., Yang, N., Srivastava, K., Geliebter, J., Tiwari, R., Miao, M. & Li, X. (2021). A small molecule compound berberine as an orally active therapeutic candidate against COVID-19 and SARS: A computational and mechanistic study. *The FASEB Journal*, 35(4), e21360. <https://doi.org/10.1096/fj.202001792R>

Warowicka, A., Nawrot, R. & Goździcka-Józefiak, A. (2020). Antiviral activity of berberine. *Archives of Virology*, 165(9), 1935–1945. <https://doi.org/10.1007/s00705-020-04706-3>

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R. & Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46(W1), W296–W303. <https://doi.org/10.1093/nar/gky427>

Wenrich, B. R. & Trumbo, T. A. (2012). Interaction of nucleic acids with Coomassie Blue G-250 in the Bradford assay. *Analytical Biochemistry*, 428(2), 93–95. <https://doi.org/10.1016/j.ab.2012.06.014>

WHO. (2022). Coronavirus disease (COVID-19). https://www.who.int/health-topics/coronavirus#tab=tab_1. Retrieved 26 June 2023, from https://www.who.int/health-topics/coronavirus#tab=tab_1

WHO. Post COVID-19 condition (long COVID). [https://www.who.int/news-room/fact-sheets/detail/post-covid-19-condition-\(long-covid\)](https://www.who.int/news-room/fact-sheets/detail/post-covid-19-condition-(long-covid)). [https://www.who.int/news-room/fact-sheets/detail/post-covid-19-condition-\(long-covid\)](https://www.who.int/news-room/fact-sheets/detail/post-covid-19-condition-(long-covid))

Wiechelman, K. J., Braun, R. D. & Fitzpatrick, J. D. (1988). Investigation of the bicinchoninic acid protein assay: Identification of the groups responsible for color formation. *Analytical Biochemistry*, 175(1), 231–237. [https://doi.org/10.1016/0003-2697\(88\)90383-1](https://doi.org/10.1016/0003-2697(88)90383-1)

Wingfield, P. T. (2015). Overview of the purification of recombinant proteins. *Current Protocols in Protein Science*, 806.1.1–6.1.35. <https://doi.org/10.1002/0471140864.ps0601s80>

Wink, M. (2020). Potential of DNA Intercalating Alkaloids and Other Plant Secondary Metabolites against SARS-CoV-2 Causing COVID-19. *Diversity*, 12(5), 175. <https://doi.org/10.3390/d12050175>

Wittig, I., Braun, H. & Schägger, H. (2006). Blue native PAGE. *Nature Protocols*, 1(1), 418–428. <https://doi.org/10.1038/nprot.2006.62>

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), 102681. <https://doi.org/10.1016/j.isci.2021.102681>

Wu, G., Robertson, D. H., Brooks, C. L. & Vieth, M. (2003). Detailed analysis of grid-based molecular docking: A case study of CDOCKER-A CHARMM-based MD docking algorithm. *Journal of Computational Chemistry*, 24(13), 1549–1562. <https://doi.org/10.1002/jcc.10306>

Wu, H., Xing, N., Meng, K., Fu, B., Xue, W., Dong, P., Tang, W., Xiao, Y., Liu, G., Luo, H., Zhu, W., Lin, X., Meng, G. & Zhu, Z. (2021). Nucleocapsid mutations R203K/G204R increase the infectivity, fitness, and virulence of SARS-CoV-2. *Cell Host & Microbe*, 29(12), 1788–1801.e6. <https://doi.org/10.1016/j.chom.2021.11.005>

Wu, H., Xing, N., Meng, K., Fu, B., Xue, W., Dong, P., Xiao, Y., Liu, G., Luo, H., Zhu, W., Lin, X., Meng, G., & Zhu, Z. (2021). Nucleocapsid mutation R203K/G204R increases the infectivity, fitness and virulence of SARS-CoV-2. <https://www.biorxiv.org/content/10.1101/2021.05.24.445386v1>

Wu, J., Kuan, I. -, Guo, J., Hsu, W., Tang, W., Chan, H., Chen, Y., Chen, B., Wu, H. & Liao, J. C. (2023). SARS-CoV-2 N protein mediates intercellular nucleic acid dispersion, a feature reduced in Omicron. *iScience*, 26(2), <https://doi.org/10.1016/j.isci.2023.105995>

Wu, S., Tian, C., Liu, P., Guo, D., Zheng, W., Huang, X., Zhang, Y. & Liu, L. (2021). Effects of SARS-CoV-2 mutations on protein structures and intraviral protein–protein interactions. *Journal of Medical Virology*, 93(4), 2132–2140. <https://doi.org/10.1002/jmv.26597>

Wu, W., Cheng, Y., Zhou, H., Sun, C. & Zhang, S. (2023). The SARS-CoV-2 nucleocapsid protein: its role in the viral life cycle, structure and functions, and use as a potential target

in the development of vaccines and diagnostics. *Virology Journal*, 20(1), 6. <https://doi.org/10.1186/s12985-023-01968-6>

Yammine, A., Gao, J. & Kwan, A. H. (2019). Tryptophan Fluorescence Quenching Assays for Measuring Protein-ligand Binding Affinities: Principles and a Practical Guide. *Bio-Protocol*, 9(11), e3253. <https://doi.org/10.21769/BioProtoc.3253>

Yan, B., Wang, D., Dong, S., Cheng, Z., Na, L., Sang, M., Yang, H., Yang, Z., Zhang, S. & Yan, Z. (2017). Palmatine inhibits TRIF-dependent NF- κ B pathway against inflammation induced by LPS in goat endometrial epithelial cells. *International Immunopharmacology*, 45194–200. <https://doi.org/10.1016/j.intimp.2017.02.004>

Yang, H., Xiao, X., Zhao, X., & Wu, Y. Intrinsic fluorescence spectra of tryptophan, tyrosine and phenylalanine. , 10255 1199–1206. <https://doi.org/10.1117/12.2268397>. <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10255/102554M/Intrinsic-fluorescence-spectra-of-tryptophan-tyrosine-and-phenylalanine/10.1117/12.2268397.full>.

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://doi.org/10.3389/fchem.2020.624765>

Yang, M., Li, C., Ye, G., Shen, C., Shi, H., Zhong, L., Tian, Y., Zhao, M., Wu, P., Hussain, A., Zhang, T., Yang, H., Yang, J., Weng, Y., Liu, X., Wang, Z., Gan, L., Zhang, Q., Liu, Y., . . . Zhao, Y. (2024). Aptamers targeting SARS-CoV-2 nucleocapsid protein exhibit potential anti pan-coronavirus activity. *Signal Transduction and Targeted Therapy*, 9(1), 40. <https://doi.org/10.1038/s41392-024-01748-w>

Yano, T., Kajisa, T., Ono, M., Miyasaka, Y., Hasegawa, Y., Saito, A., Otsuka, K., Sakane, A., Sasaki, T., Yasutomo, K., Hamajima, R., Kanai, Y., Kobayashi, T., Matsuura, Y.,

Itonaga, M. & Yasui, T. (2022). Ultrasensitive detection of SARS-CoV-2 nucleocapsid protein using large gold nanoparticle-enhanced surface plasmon resonance. *Scientific Reports*, 12(1), 1060. <https://doi.org/10.1038/s41598-022-05036-x>

Ye, Q., West, A. M. V., Silletti, S. & Corbett, K. D. (2020). Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. *Protein Science: A Publication of the Protein Society*, 29(9), 1890–1901. <https://doi.org/10.1002/pro.3909>

Zanwar, A., Sen, D., Memom, M., Sanathra, R., Patel, D., Maheshwari, R. & Sen, A. (2022). View of Spectrophotometric Methods for the Analysis of Berberine Hydrochloride and Eugenol in Formulated Emulgel. *Journal of Natural Remedies*, 22(3), <https://doi.org/10.18311/jnr/2022/30166>

Zeng, J., Pei, H., Wu, H., Chen, W., Du, R. & He, Z. (2024). Palmatine attenuates LPS-induced neuroinflammation through the PI3K/Akt/NF- κ B pathway. *Journal of Biochemical and Molecular Toxicology*, 38(1), e23544. <https://doi.org/10.1002/jbt.23544>

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. <https://doi.org/10.1016/j.bbrc.2020.04.136>

Zhang, B., Tian, J., Zhang, Q., Xie, Y., Wang, K., Qiu, S., Lu, K. & Liu, Y. (2022). Comparing the Nucleocapsid Proteins of Human Coronaviruses: Structure, Immunoregulation, Vaccine, and Targeted Drug. *Frontiers in Molecular Biosciences*, 9761173. <https://doi.org/10.3389/fmolb.2022.761173>

Zhang, Q., Xiang, R., Huo, S., Zhou, Y., Jiang, S., Wang, Q. & Yu, F. (2021). Molecular mechanism of interaction between SARS-CoV-2 and host cells and interventional therapy. *Signal Transduction and Targeted Therapy*, 6(1), 233. <https://doi.org/10.1038/s41392-021-00653-w>

Zhao, H., Nguyen, A., Wu, D., Li, Y., Hassan, S. A., Chen, J., Shroff, H., Piszczek, G. & Schuck, P. (2022). Plasticity in structure and assembly of SARS-CoV-2 nucleocapsid protein. *bioRxiv: The Preprint Server for Biology*, 2022.02.08.479556. <https://doi.org/10.1101/2022.02.08.479556>

Zhao, H., Syed, A. M., Khalid, M. M., Nguyen, A., Ciling, A., Wu, D., Yau, W., Srinivasan, S., Esposito, D., Doudna, J. A., Piszczek, G., Ott, M. & Schuck, P. (2024). Assembly of SARS-CoV-2 nucleocapsid protein with nucleic acid. *Nucleic Acids Research*, 52(11), 6647–6661. <https://doi.org/10.1093/nar/gkae256>

Zhao, H., Wu, D., Nguyen, A., Li, Y., Adão, R. C., Valkov, E., Patterson, G. H., Piszczek, G. & Schuck, P. (2021). Energetic and structural features of SARS-CoV-2 N-protein co-assemblies with nucleic acids. *iScience*, 24(6), 102523. <https://doi.org/10.1016/j.isci.2021.102523>

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

Zhou, R., Zeng, R., von Brunn, A. & Lei, J. (2020a). Structural characterization of the C-terminal domain of SARS-CoV-2 nucleocapsid protein. *Molecular Biomedicine*, 12. <https://doi.org/10.1186/s43556-020-00001-4>

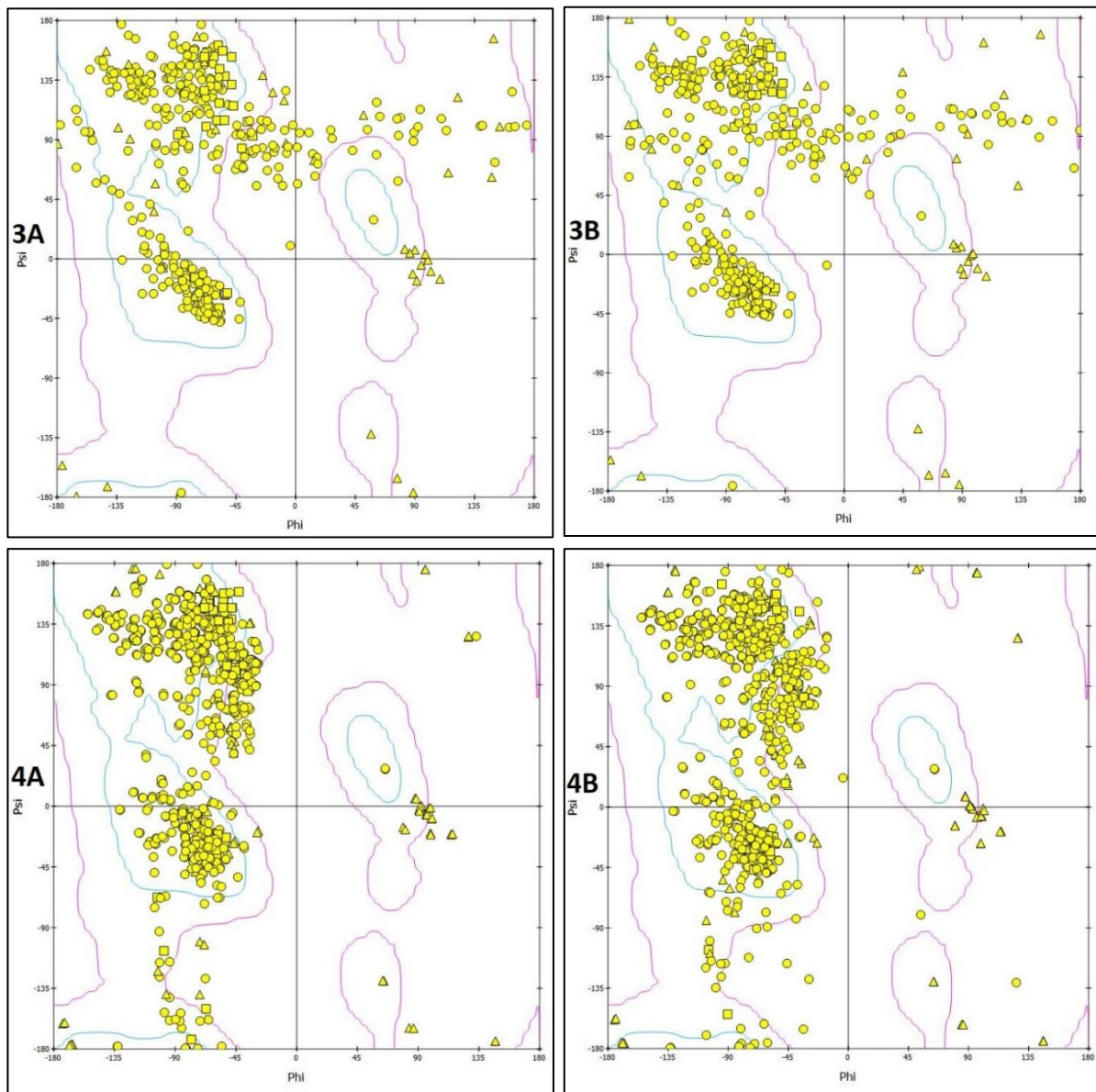
Zhou, R., Zeng, R., von Brunn, A. & Lei, J. (2020b). Structural characterization of the C-terminal domain of SARS-CoV-2 nucleocapsid protein. *Molecular Biomedicine*, 1(1), 2. <https://doi.org/10.1186/s43556-020-00001-4>

Zhu, Z., Liu, G., Meng, K., Yang, L., Liu, D. & Meng, G. (2021). Rapid Spread of Mutant Alleles in Worldwide SARS-CoV-2 Strains Revealed by Genome-Wide Single Nucleotide

Polymorphism and Variation Analysis. *Genome Biology and Evolution*, 13(2), evab015.
<https://doi.org/10.1093/gbe/evab015>

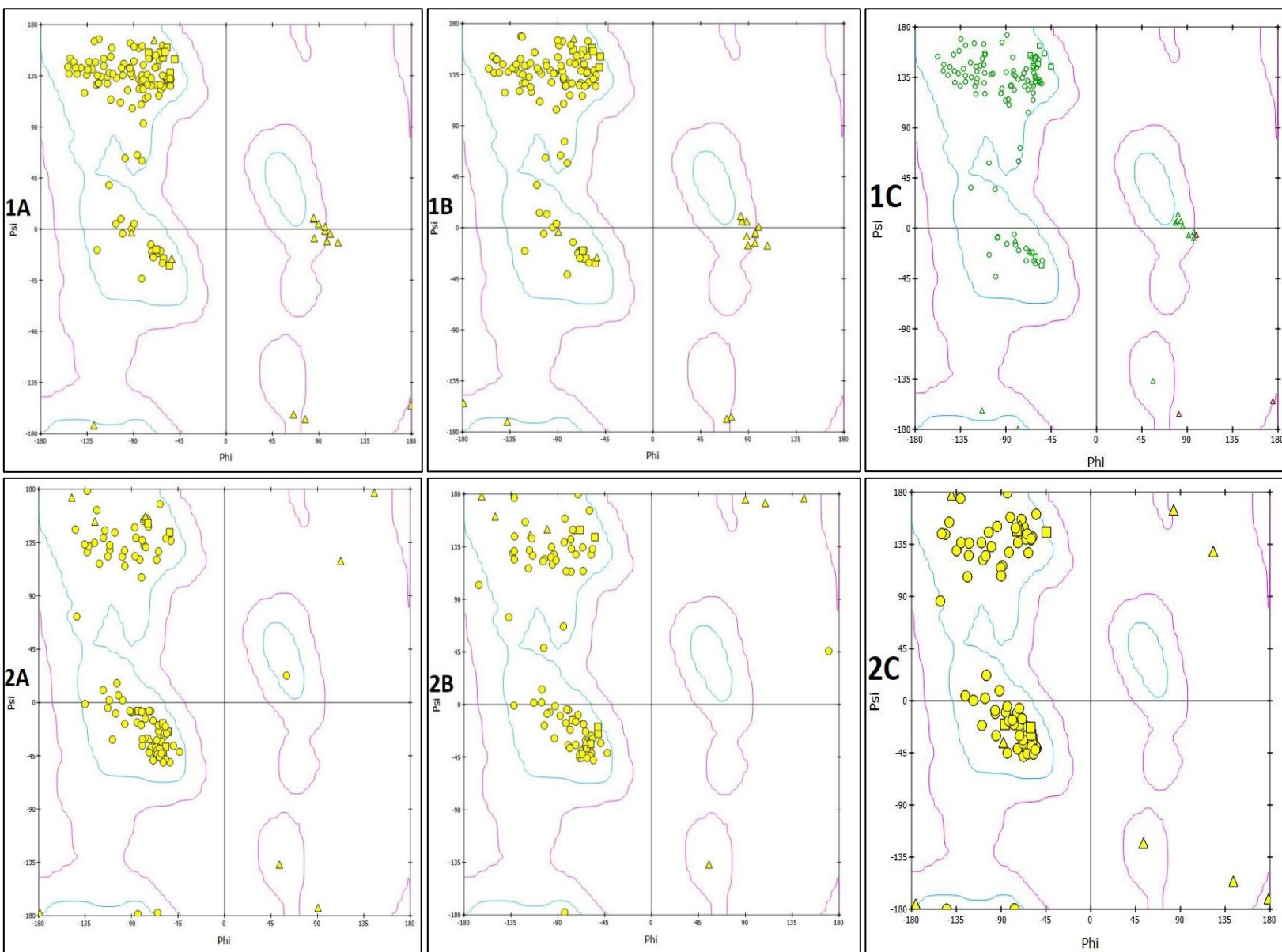
Zymo Research. (2023). ZymoPURE™ Plasmid Miniprep Kit Instruction Manual [computer software]. Taiwan:

10. Appendix



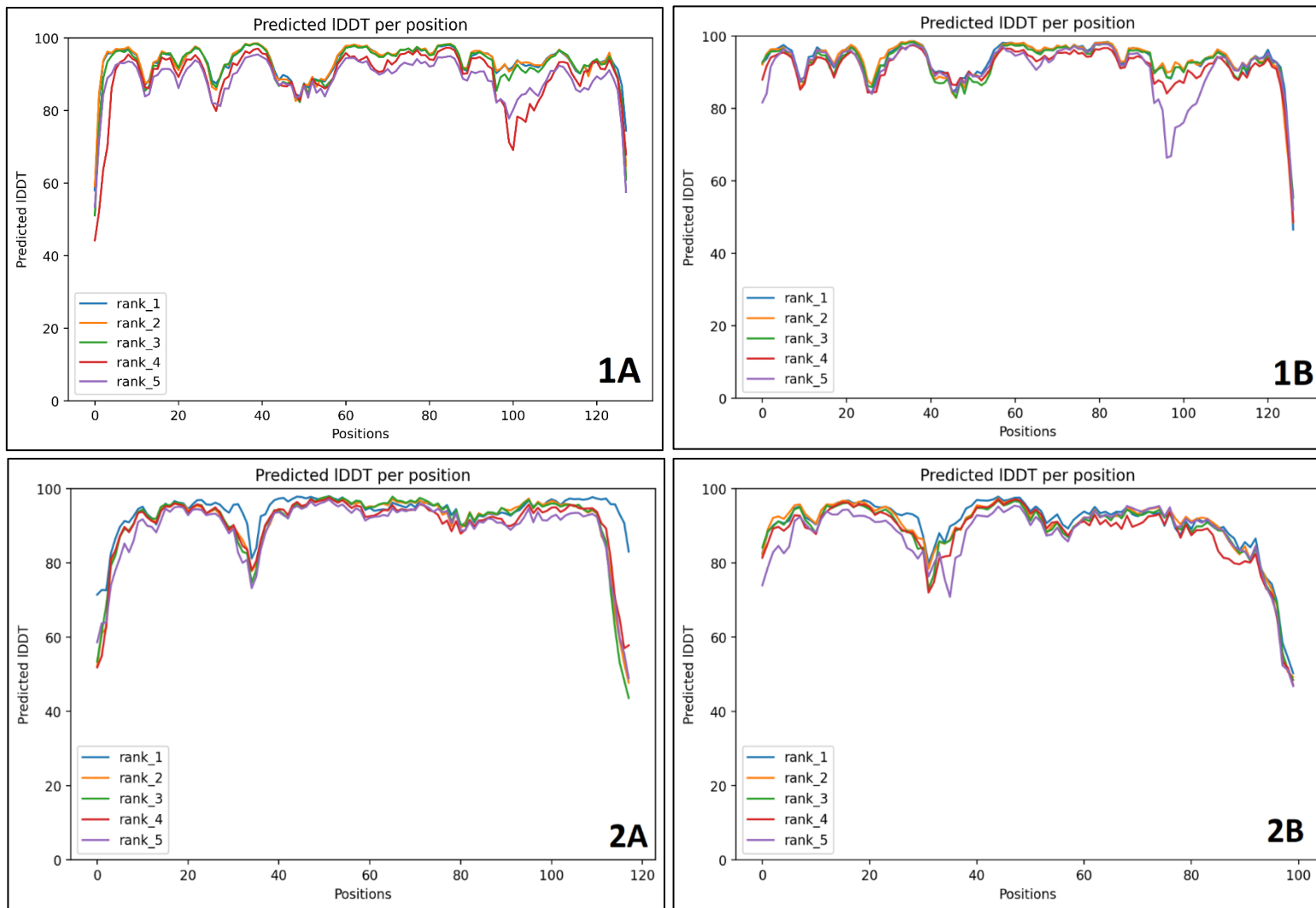
Appendix 10. 1: Ramachandran Plots of the predicted structure for monomeric and dimeric N protein

Ramachandran Plots for the Monomeric (3) and Dimeric (4) structures predicted by AlphaFold for the Wuhan (A) and W-NP (B) sequences.



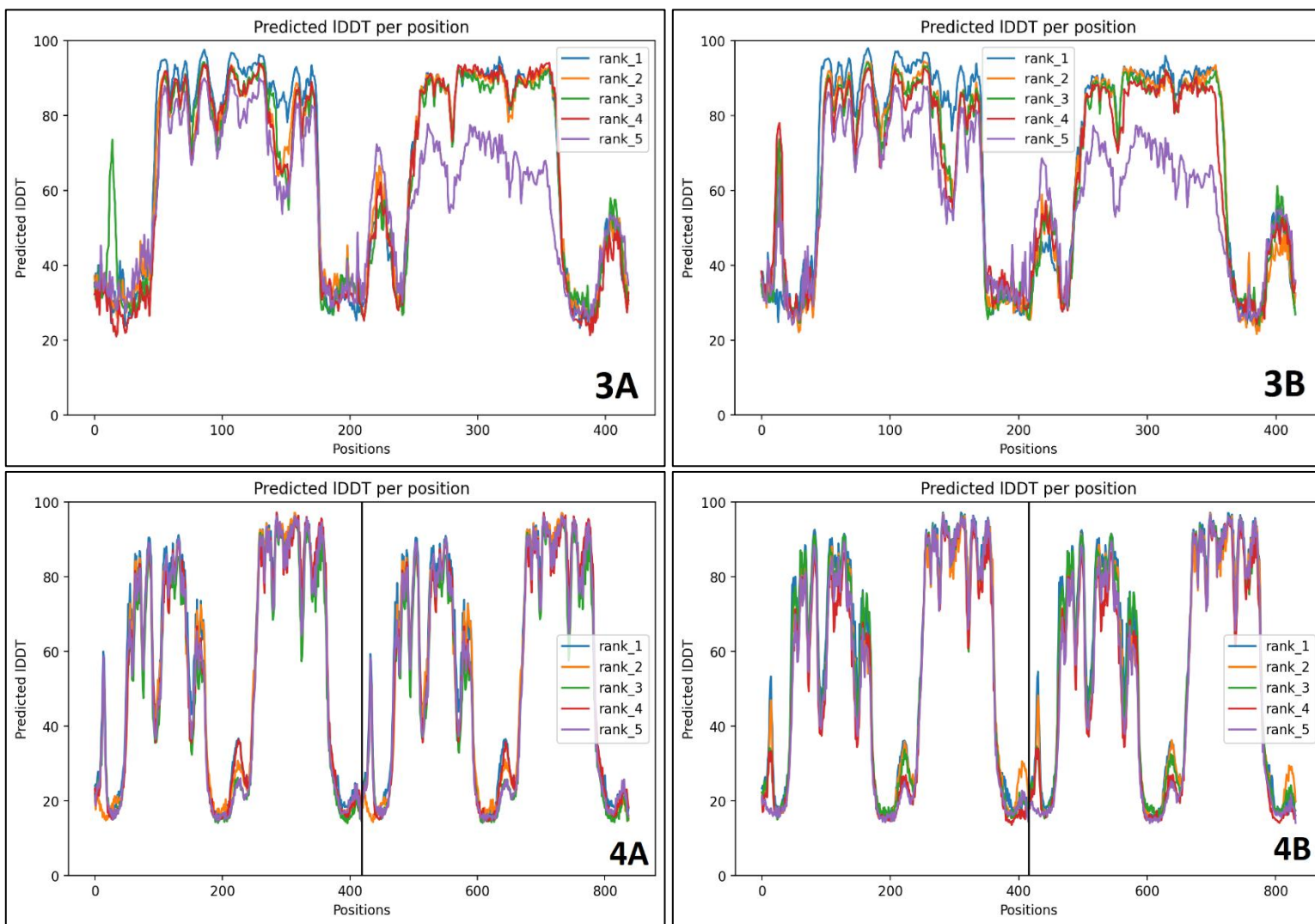
Appendix 10. 2: Ramachandran Plots for the predicted structure for N protein domains

This figure displays the Ramachandran plot illustrating the distribution of backbone dihedral angles (ϕ and ψ) for predicted structures of NP NTD (1) & CTD (2) for the Wuhan (A), W-NP (B) and Swiss-Model (C) Predicted Structures



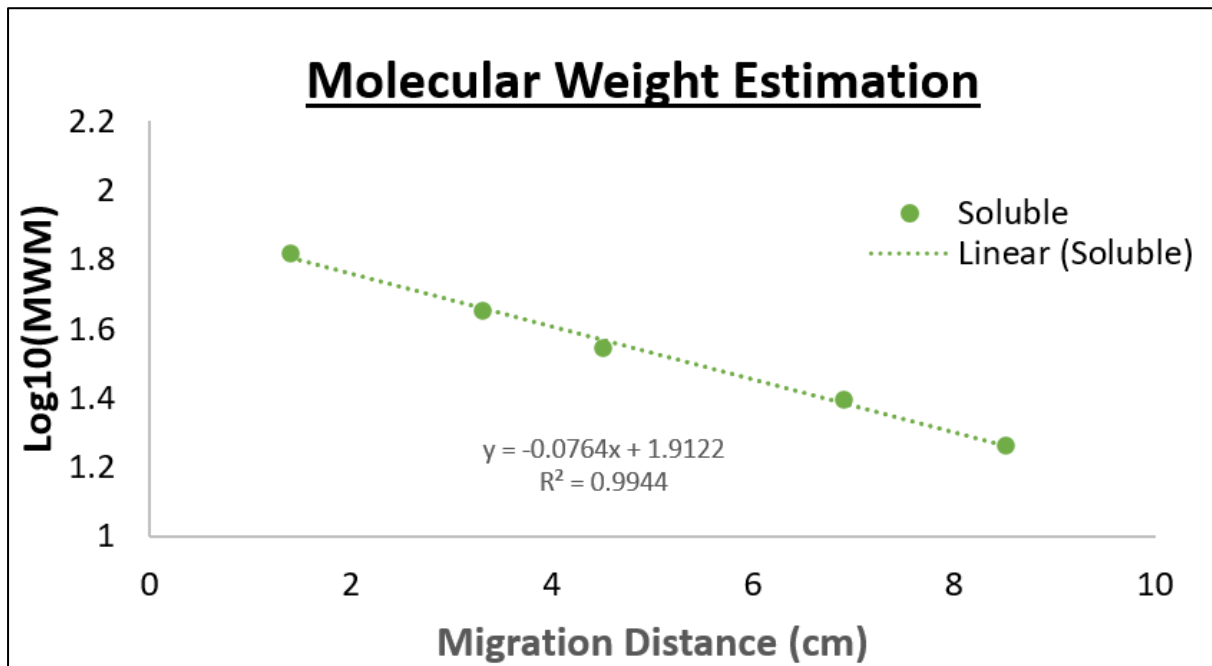
Appendix 10. 3: pLDDT scores for the predicted structure for N protein domains

This plot displays the per-residue confidence (pLDDT) scores for NTD (1) & CTD (2) for the Wuhan (A) and W-NP (B) Predicted Structures. Low confidence suggests disordered or unreliable.



Appendix 10. 4: pLDDT scores for the predicted structure for monomeric and dimeric N protein

This plot displays the per-residue confidence (pLDDT) scores for the Monomeric (3) and Dimeric (4) structures predicted by AlphaFold for the Wuhan (A) and W-NP (B) sequences.



$$y = -0.0764x + 1.9122$$

$$y = -0.0764 \times 2.9 + 1.9122$$

$$\text{Unknown log10} = 1.69064$$

$$\text{MW} = 10^{1.69064}$$

$$\text{MW} = 49.05 \text{ kDa}$$

Appendix 10. 5: Molecular Weight Estimation

Figure A displays the linear regression line used to estimate the molecular weight of the over expressed protein, anticipated to be W-NP, from the BL21(DE3)-T7 express cell line soluble fractions. The migration of the proteins within the molecular weight marker (MWM) in centimeters (cm) is shown on the x-axis and the log of the molecular weight of those proteins are shown on the y-axis. Demarked is the linear regression line calculated by MS Excel. B shows the mathematics solving the line equation using the migration profile of W-NP.

$$A = \epsilon cl$$

$$c = \frac{A}{\epsilon l}$$

$$c = \frac{2.73}{(43890)(1)}$$

$$c = 0.00006215 \text{ M}$$

$$c = 62.115 \mu\text{M}$$

$$y = 0,3156 x + 0,0129$$

$$x = (y - 0,0129) / 0,3156$$

$$x = (0,258 - 0,0129) / 0,3156$$

$$x = 0,78 \text{ mg/ml}$$

$$0,78 \text{ mg/ml} \times 4 = 3,11 \text{ mg/ml}$$

$$3,11 \text{ mg/ml} \times 5 \text{ ml} = 15,53 \text{ mg}$$

$$15,53 \text{ mg} / 1000 = 0,01553 \text{ g}$$

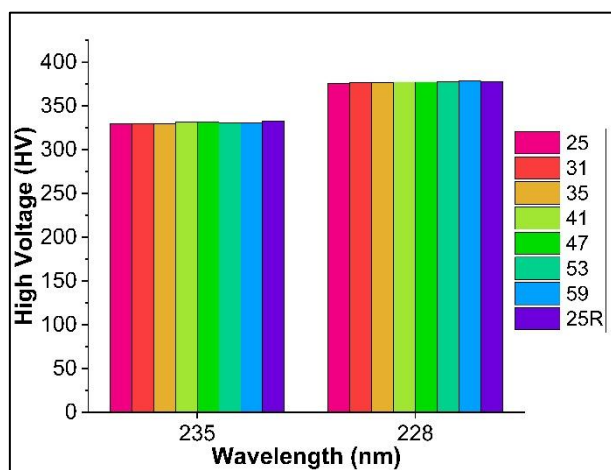
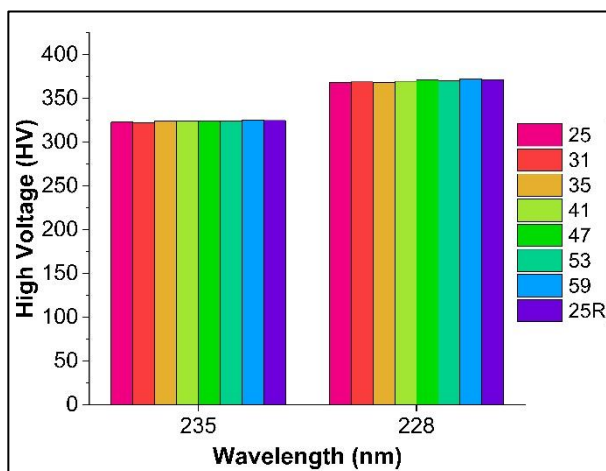
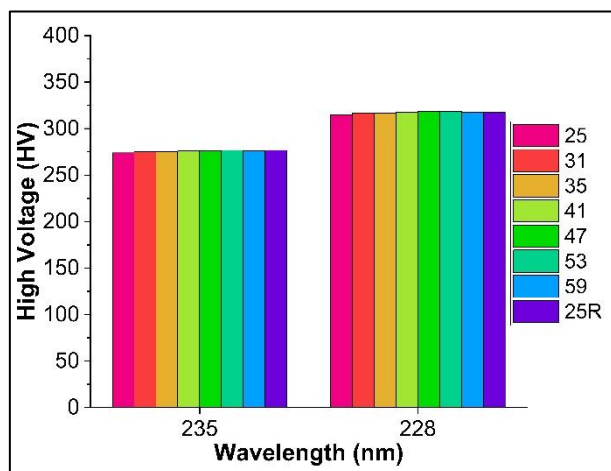
$$0,01553 \text{ g} / 45\,490.59 \text{ g/mol} = 0,0000003414 \text{ mol}$$

$$0,0000003414 \text{ mol} / 0,005 \text{ L} = 0,00006829 \text{ M}$$

$$0,00006829 \text{ M} \times 1000 \times 1000 = 68,29 \mu\text{M}$$

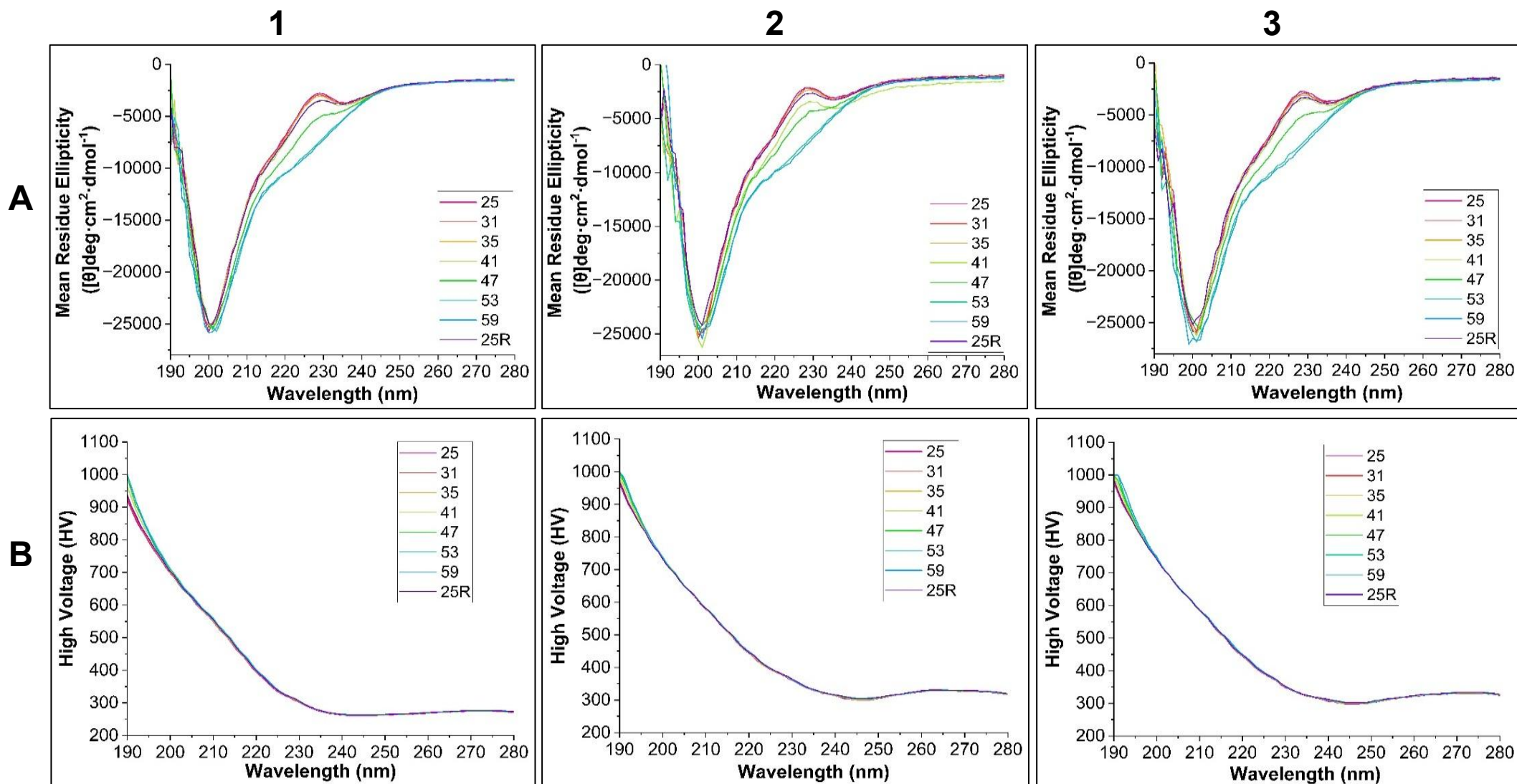
Appendix 10. 7: Concentration Determination of purified W-NP

A shows the estimation of concentration of W-NP isolated using absorbance spectroscopy described in **Sections 3.3.8.3.2 & 4.2.3** .B shows the estimation of the concentration of W-NP isolated using the Bradford assay described in **Sections 3.3.8.3.1 & 4.2.3**



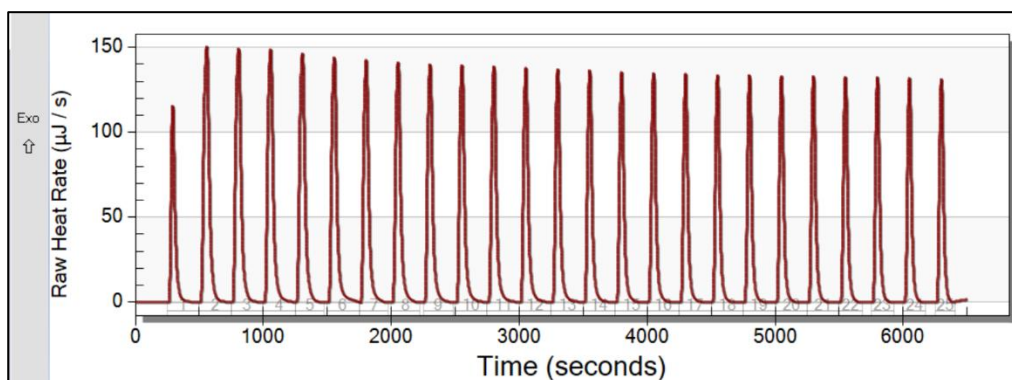
Appendix 10. 6: HV of thermal CD at 235 nm and 228 nm

This figure displays the High Voltage (HV) intensities obtained during the 235 and 228 nanometres (nm) CD thermal ramp for W-NP in the absence (A) and presence of 200 μM berberine (B) and palmatine (C).



Appendix 10. 8: Impact of Compounds on W- NP CD Thermal Structural Stability

This figure displays the CD Spectra (A) and High Voltage (B) intensities obtained during the thermal ramp for W-NP in the absence (1) and presence of 200 μM berberine (2) and palmitate (3).



Appendix 10. 9: ITC of Compound Titrated into Buffer

This figure depicts the raw ITC thermogram obtained from titrating compound into buffer only. The raw thermogram displays the Raw Heat Rate in microjoules per second ($\mu\text{J}/\text{s}$) over time in seconds upon successive injections of palmatine into storage buffer (20 mM Tris-HCl; 1 mM ethylene diaminetetra-acetic acid (EDTA); 1 mM DTT; 10% (v/v) glycerol; pH = 8.0) only.