UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

AYINDE KEHINDE SULAIMON

EXPRESSION, PURIFICATION, AND CHARACTERIZATION OF SARS-COV-2 ORF9b AND THE STUDY OF ITS INTERACTION WITH HUMAN HSP90 AND CYTOSOLIC TOM70.

EXPRESSÃO, PURIFICAÇÃO E CARACTERIZAÇÃO DA ORF9b DE SARS-COV-2 E ESTUDO DA SUA INTERAÇÃO COM A HSP90 HUMANA E TOM70 CITOSÓLICA

CAMPINAS

2022

AYINDE KEHINDE SULAIMON

EXPRESSION, PURIFICATION, AND CHARACTERIZATION OF SARS-COV-2 ORF9b AND THE STUDY OF ITS INTERACTION WITH HUMAN HSP90/CYTOSOLIC TOM70.

EXPRESSÃO, PURIFICAÇÃO E CARACTERIZAÇÃO DA ORF9b DE SARS-COV-2 E ESTUDO DA SUA INTERAÇÃO COM A HSP90 HUMANA E TOM70 CITOSÓLICA

Dissertation presented to the Institute of Biology of the University of Campinas in partial fulfillment of the requirements for the degree of Masters in Functional and Molecular Biology, Biochemistry track.

Dissertação apresentada ao Instituto de Biologia da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do Título de Mestre em Biologia Funcional e Molecular, na Area de Bioquimica.

Orientador: Carlos Henrique Inácio Ramos

ESTE ARQUIVO DIGITAL CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELO ALUNO AYINDE KEHINDE SULAIMON E ORIENTADA PELO PROF. DR. CARLOS HENRIQUE INÁCIO RAMOS

CAMPINAS
2022

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Mara Janaina de Oliveira - CRB 8/6972

Ay34e

Ayinde, Kehinde Sulaimon, 1994-
Expression, purification, and characterization of SARS-CoV-2 ORF9b and the study of its interaction with human HSP90 and cytosolic TOM70 / Kehinde Sulaimon Ayinde. – Campinas, SP : [s.n.], 2022.

Orientador: Carlos Henrique Inacio Ramos.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Proteínas virais reguladoras e acessórias. 2. Interação proteína-proteína. 3. SARS-CoV-2. 4. Proteínas de choque térmico HSP90. 5. Subunidade do receptor de importação mitocondrial TOM70. 6. COVID-19. I. Ramos, Carlos Henrique Inacio, 1967-. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Expressão, purificação e caracterização da ORF9b de SARS-CoV-2 e estudo da sua interação com a HSP90 humana e TOM70 citosólica
Palavras-chave em inglês:
Viral regulatory and accessory proteins
Protein-protein interactions
SARS-CoV-2
HSP90 heat shock proteins
Mitochondrial import receptor subunit TOM70
COVID-19 (Disease)
Área de concentração: Bioquímica
Titulação: Mestre em Biologia Funcional e Molecular
Banca examinadora:
Carlos Henrique Inacio Ramos [Orientador]
Leticia Maria Zanphorlin Murakami
Leandro Ramos Souza Barbosa
Data de defesa: 27-05-2022
Programa de Pós-Graduação: Biologia Funcional e Molecular

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0002-4511-2562
- Currículo Lattes do autor: http://lattes.cnpq.br/1213343472751438

Campinas, 2022.

COMISSÃO EXAMINADORA

Prof. Dr. Carlos Henrique Inacio Ramos

Dra. Leticia Maria Zanphorlin Murakami

Dr. Leandro Ramos Souza Barbosa

Prof. Dr. Daniel Martins de Souza

Dra. Adriana Franco Paes Leme

Os membros da Comissão Examinadora acima assinaram a Ata de defesa, que se encontra no processo de vida acadêmica do aluno.

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa Biologia Funcional e Molecular, na Area de Bioquimica da Unidade instituto de biologia

DEDICATION

I dedicate this work to all the lives lost to the COVID-19 pandemic.

ACKNOWLEDGEMENT

I give thanks to almighty Allah who has spared my life since the onset until this period, he is to be praised and worshiped “Al-Hamdu Lillahi”. Further, I appreciate my supervisor Professor Dr. Carlos Henrique Inacio Ramos for his wonderful mentorship, and for providing me with an excellent laboratory with state of the art equipment and techniques for my project. I am appreciating him for believing in me from the first day I got in contact with him until the end of my master’s degree. His guidance and student support made this journey a successful one.

This study was financed by the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) – Finance code 88887.657619/2021-00. This scholarship saw me through the entire length of my master’s program without any financial constraints.

This study was financed in part by the Sao Paolo Research Foundation FAPESP, grant 2017/26131-5. This funding was support was provided to the laboratory.

Similarly, I appreciate the scholarship received from SAE UNICAMP (Serviço de Apoio ao Estudante) that catered for my housing and transport allowance during my studies at UNICAMP.

I am also using the medium to appreciate both the Institute of Biology (IB) and Institute of Chemistry (IQ) for providing me with the infrastructural access including classrooms, research laboratories and other facilities that made studies easy for me throughout the period. I will not forget to appreciate the effort of UNICAMP COVID response team and the university hospital for ensuring that the COVID-19 pandemic was effectively controlled.

My reserved appreciation goes to Gluacia for her efforts in teaching me and showing me how to use most of the techniques I used for this project. I appreciate her patience and dedication to breed quality future scientists. Likewise, I express my appreciation to Annelize for her excellent support, and wonderful technical assistance throughout the entire period of this project. This also extend to Aline and other members of the laboratory; Barbara, Carolina,

Larissa, Claudia, Bianca, Caroline, Thais, Gabriel, Ana Paula, Julia, and others for their assistance one way or the other, and for giving me a wonderful and comforting environment in the laboratory.

I also appreciate my dear friends Olamide Olaoba for always being there for me throughout this period. I will not forget to appreciate the support and hospitality from the African community in Campinas (Mukaila, Abubakar, Naomi, Papa, Dr. and Drs. Alagbe, Dr. Edwin, Ibrahim, Aishat, Firdaus, and others).

Lastly, I thank my parents Mr. and Mrs. Ayinde for their moral and financial support right from my childhood until now. I really appreciate their efforts and I pray that they will grow old in good health to witness many more of my achievements in life. Also, to my family members, Taiwo, Hassan, Sofiyyah, and Khadijat, I say a very big thank you to you guys. You are the most amazing set of people in my life.

"…. and their closing prayer will be, "All praise is for Allah – Lord of the worlds!"

Qur'an 10, Verse 10.

RESUMO

O surgimento da pandemia de Covid-19, causada pelo vírus da Sindrome Respiratória Aguda Grave, em inglês *Severe Acute Respiratory Syndrome* (Sars-Cov-2) continua sendo uma ameaça significativa à saúde global, apesar do desenvolvimento precoce de vacinas. No entanto, os pesquisadores continuaram a intensificar os estudos para entender as características moleculares das proteínas virais durante a infecção, replicação e evasão do sistema imune do hospedeiro. A ORF9b, uma proteína acessória vital do vírus Sars-Cov-2, foi identificada por desempenhar um papel crítico na interação viral com o hospedeiro, visando um membro do complexo da translocase mitocondrial da membrana externa, TOM70. A ORF9b e a TOM70 interrompem a sinalização antiviral mitocondrial (MAVS), levando à evasão imune. Aqui, descrevemos a expressão, purificação e caracterização de ORF9b usando técnicas moleculares e biofísicas. Além disso, coexpressamos as duas proteínas no sistema *E coli* seguido de estudos de interação proteína-proteína. A proteína de 97 resíduos de aminoácidos foi purificada como um homodímero com uma massa molecular aproximada de 22 kDa conforme determinado por SEC-MALS. Estudos de enovelamento e conformação usando espectropolarimetria de dicroísmo circular mostraram que a proteína apresenta uma conformação randomica e é estável ao calor. O complexo ORF9b-TOM70 caracterizado por dicroísmo circular e calorimetria diferencial de varredura mostrou que o complexo está enovelado em conformação semelhante, porém mais estável termicamente quando comparado à TOM70 livre. É importante ressaltar que, usando o experimento de *pull down*, demonstramos que tanto a proteína inteira quanto o domínio C-terminal da proteína de choque térmico humana de 90 kDa (Hsp90 e C-Hsp90) não interagem com o complexo ORF9b-TOM70. Portanto, uma vez que a interação de TOM70 e Hsp90 desempenha um papel importante nas atividades antivirais mitocondrial, nossos achados mostraram que a ORF9b complexada com a TOM70,impede a interação com Hsp90, que é uma das principais explicações para a evasão de SARS-CoV-2 da imunidade inata do hospedeiro via inibição da ativação do interferon.

ABSTRACT

The emergence of the Covid-19 pandemic, caused by Severe Acute Respiratory Syndrome (Sars-Cov-2) virus remains a significant threat to the global health despite the early development of vaccines. However, researchers have continued to intensify studies into understanding the molecular characteristics of the viral proteins during infection, replication and host immune evasion. ORF9b, a vital accessory protein of the Sars-Cov-2 virus is identified to play a critical role in viral host interaction, targeting a member of the mitochondrial translocase of the outer membrane complex, TOM70. This ORF9b and TOM70 disrupts the mitochondrial antiviral signaling (MAVS), leading to immune evasion. Here, we describe the recombinant expression, purification, and characterization of ORF9b using molecular and biophysical techniques. Further, we co-expressed the two proteins in the e-coli system followed by protein-protein interaction studies. The 97 amino acid protein was purified as a homodimer with an approximate molecular mass of 22 kDa as determined by SEC-MALS. Folding and conformational studies using circular dichroism spectropolarimetry showed that the protein exhibit a random conformation and it is heat stable. ORF9b-TOM70 complex characterized by circular dichroism and differential scanning calorimetry showed that the complex is folded in similar conformation, but more thermal stable when compared to the free TOM70. Importantly, using the pull down experiment, we demonstrated that both the full length and the C-terminal domain of Human 90kDa heat shock protein (Hsp90 and C-Hsp90) does not interact with the ORF9b-TOM70 complex. Therefore, since the TOM70 and Hsp90 assembly plays an important role in the mitochondrial antivirus activities, our findings showed that ORF9b complexed with TOM70 and prevents the interaction with Hsp90, which is one major explanation for SARS-CoV-2 evasion of host innate immunity via the inhibition of interferon activation pathway.

LIST OF FIGURES

LIST OF TABLES

SUMMARY

1.0 INTRODUCTION

1.1 Cellular Protein Homeostasis

The term cellular protein homeostasis, also known as proteostasis describes the cell's operation network in preserving a correct balance between protein synthesis, protein folding/conformational maintenance, and protein degradation. Various types of cellular stresses including oxidative stress, heat shock, enzyme malfunction, and mutations can result in disturbed proteostasis (Laskowska et al., 2019). Similarly, ever-changing metabolic, environmental, physiological, and pathological conditions of cells contribute to the production and accumulation of unfolded, misfolded and/or damaged proteins which jeopardize the cell's protein homeostasis. Of no doubt, the sustainability of a cell depends significantly on a healthy proteome condition; therefore, maintaining a controlled proteomic system in the intracellular milieu is a life-long challenge to individual cells of every organism (Hoppe and Cohen, 2020). As a result, cells require a primary cellular mechanism to balance between folding and degradation of misfolded proteins, this phenomenon is referred to as the protein quality control (PQC).

Protein quality control (PQC) requires the interplay between activities of molecular chaperones, proteasome and autophagy. In the PQC system, several mechanisms are used to prevent accumulation of aggregates from incorrect protein folding, and if misfolding occur, proteins may be degraded by proteases or repaired by chaperones (Tiroli-Cepeda and Ramos, 2011). Autophagy plays a significant role in eukaryotic cells by maintaining crosstalk with the proteasome system by degrading already formed aggregates delivered to the lysosome (Ji and Kwon, 2017). Hence, failure to maintain this balance results in the accumulation of altered and misfolded proteins which are the basic causes of amyloid fibril deposition associated with several pathological disorders in humans, including cataract, Parkinson's, Alzheimer's, and prion diseases (Tiroli-Cepeda and Ramos, 2011).

1.1.1 Molecular Chaperones

According to Harlt (1996), molecular chaperones are described as proteins that assist in the proper folding, maintenance of conformational, or regulation of another protein, and by doing so do not affect the biologically active structure or become part of its final structure (Hartl, 1996). They are also said to be proteins that upon interaction with other proteins, stabilize or help the other protein gain its functionally active conformation (Borges and Ramos, 2005, Hartl and Hayer-Hartl, 2009, Wankhede et al., 2022). In addition, these chaperones recognize and interact with nonnative proteins, inhibiting the nonspecific interactions between intermediates of defective and partially folded proteins, and assisting in the proper folding process (Shiber and Ravid, 2014). (Hartl and Hayer-Hartl, 2009, Wankhede et al., 2022). In summary, they are the main players in the maintenance of a balanced proteostasis.

Molecular chaperones expression can be constitutive or induced under stressed conditions such as in cells subjected to thermal stress, therefore giving rise to them being referred to as Heat shock proteins (HSPs) or stress proteins but chaperones and HSPs are not synonymous (Ramos, 2011). HSPs are divided into Holders, Foldases, and Disaggregases based on their interaction with their client proteins. The ATP-independent holders recognize, bind and stabilize unfolded polypeptides to avoid aggregation and, upon normalized cellular conditions deliver the protein to foldases. Foldases are ATP-dependent chaperones that assist in the refolding of client proteins by the use of energy derived from ATP hydrolysis. Disaggregases on their part use ATP hydrolysis to disentangle aggregated polypeptides and deliver them for refolding into native conformation (Tiroli-Cepeda and Ramos, 2011). HSPs are classified according to their molecular weight. The most important families of HSPs that participate broadly in *de novo* protein folding and refolding are the Hsp70, Hsp60/Hsp10, Hsp90, Hsp100 and small HSPs (sHSPs). (Hartl et al., 2011, Haslbeck and Vierling, 2015).

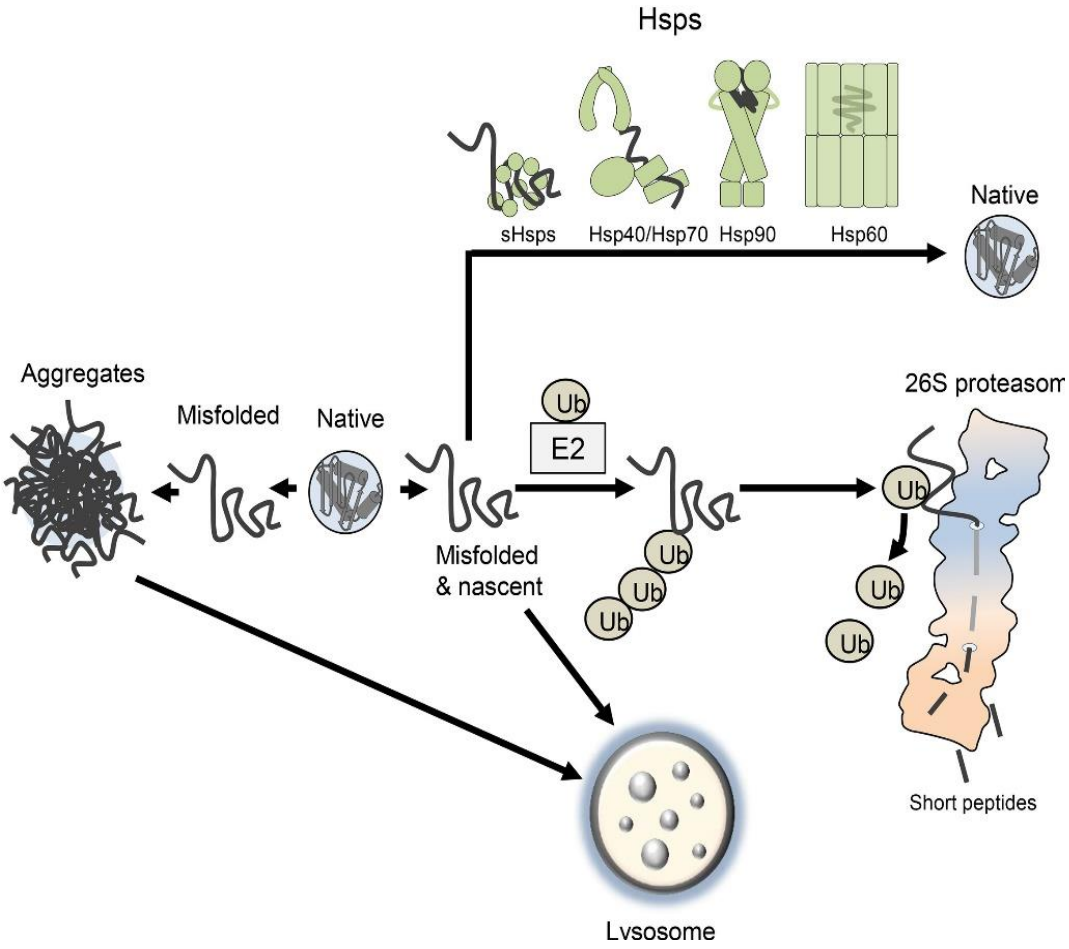


Figure 1: Cellular protein homeostasis network. In this system, the chaperones (HSPs) ensure proper protein folding and assist misfolded proteins. Aggregated proteins are labeled by ubiquitin and degraded by the proteasome system. In other cases, misfolded proteins are delivered to the lysosome and degraded by autophagy. (Adapted from Laskowska, et al., 2019).

1.1.2 Hsp90

Hsp90 is one of the main components of the PQC, and interacts with at least 10% of the proteome (da Silva and Ramos, 2012). It is a highly conserved, homodimeric, multi-domain protein present in various cellular compartments of eukaryotic cells (Ramos, 2011, Ramos and Ayinde, 2020). Hsp90 functions downstream of Hsp70 and in this regard, and does not bind nascent protein itself rather, depends on Hsp70 system to deliver protein for processing. Thereby acting as an extended substrate-binding interface to interact with unfolded, partially folded and folded polypeptides and returning them back to the Hsp70 system in a matured state ready for refolding (Young et al., 2001, Radli and Rüdiger, 2018).

The homodimeric structure of Hsp90 consists of an N-terminal domain with the ATP-binding site, a flexible domain in the middle, and C-terminal domain, all of which have specific functions. The N-terminal domain has ATPase, because it has a regulatory pocket that binds to, and hydrolyzes ATP to mediate a series of binding and unbinding of Hsp90 and client proteins (Obermann et al., 1998). The HSP90 dimer undergoes a reaction cycle characterized by changes in its conformation and ATP drives this process. Upon ATP binding, the N-terminal domains dimerizes to form the HSP90' "molecular clamp". Hartl et al, 2011 stated this ATP binding results in a closely packed arrangement of the HSP90 dimer, in a manner that individual monomers twist around each other. After hydrolysis occur, the ATPase domains in the NTD dissociate, and the Hsp90 monomers separate from the N-terminal cluster. Other co-chaperones such as p23 stabilize Hsp90 in its dimerized form before ATP hydrolysis (Figure 2) (Hartl et al., 2011).

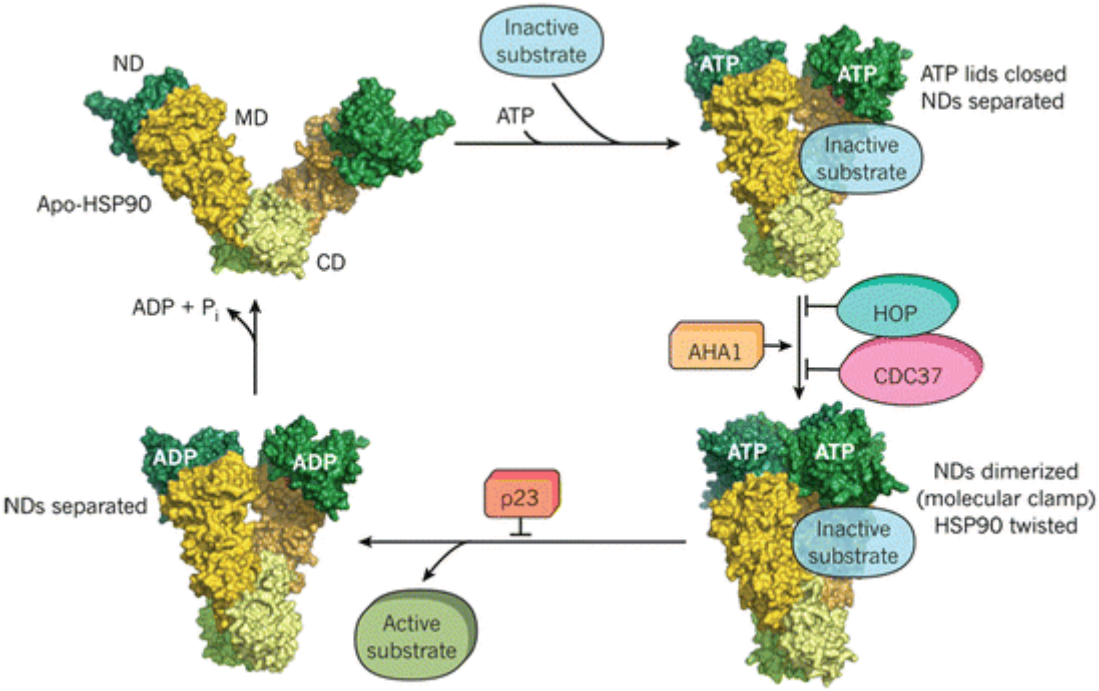


Figure 2: Hsp90 chaperone system illustrating ATP cycle. The binding of ATP to the N-terminal domain (ND) induces a conformational change and closure of the ATP lid in the ND. As a result, the ND domain is dimerized, forming a molecular clamp with twisted subunits, later dissociating upon ATP hydrolysis. (Adapted from Hartl et al., 2011).

Hsp90 cooperates with other co-chaperones for binding, many of which use tetratricopeptide repeat (TPR), a unique region of multiple repeats of 34 amino acid residue motif that are degenerate. An example is the TPR protein HOP (Hsp70-Hsp90 Organizing Protein) which serves as a link between chaperones Hsp90 and Hsp70 system for substrate transfer (Scheufler et al., 2000). A TPR motif contains two antiparallel α-helical subdomains (helix A and helix B) in the shape of a selenoid that are equivalent in length. The TPR arrangements often generates a superhelical structure in a right-handed symmetry with an amphipathic channel that can accommodate the binding regions of a target protein (Zeytuni and Zarivach, 2012, Cerveny et al., 2013) (Figure 3).

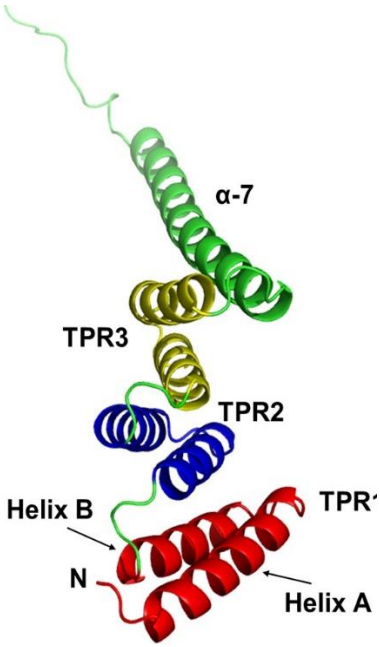


Figure 3: A typical TPR domain. Shows a structural diagram of the TPR domain of PP5. A pair of antiparallel α-helices (A and B) makes up each TPR repeat (Adapted from (Zeytuni and Zarivach, 2012)**.**

1.3 SARS-CoV-2: The new corona virus

In December 2019, the World Health Organization (WHO) named a respiratory tract disease that started in Wuhan city, China coronavirus disease 2019 (Covid19), and due to its rapid spread to other countries such as Italy and France, the WHO pronounced it a global pandemic and a public health emergency of international concern. The sequenced genome of this new virus resulted in 86.9% similarity to the genome of SARS-CoV. Hence, the new name, Corona Virus-2 of severe acute respiratory syndrome (SARS-CoV-2) was coined (Chang et al., 2020). Since then, researchers in several fields have studied and gained an excellent understanding of the genomic layout and structure of SARS-CoV-2. SARS-CoV-2 is a single-stranded positive-sense RNA virus that also shares similarities with MERS-CoV (Middle East respiratory syndrome coronavirus) (Cascella et al., 2021).

The genome of SAR-CoV-2 has about 30 kb RNA comprising 14 open reading frames (ORF). The first ORFs (ORF1a/b) are located at the 5′ end overlapping with a (-1) ribosomal frame-shift, about two-thirds of the whole genome length, and encodes polyproteins pp1a and pp1ab, respectively (Alanagreh et al., 2020, Helmy et al., 2020). The polyproteins are further processed into nonstructural proteins (NSPs), Nsp1 to Nsp16 by the viral protease (Chen et al., 2020). Other ORFs occupying the remaining one-third of the genome are located on 3′ end encodes the four major structural proteins including Spike, Envelope, Membrane, and Nucleocapsid, and some accessory proteins including Open reading frame 9 (ORF9). (Figure 4). These proteins perform various functions in the virus such as the spike protein responsible for host receptors recognition, the envelope protein responsible for virions assembly and release, and the membrane protein responsible for shaping the virion (Alanagreh et al., 2020).

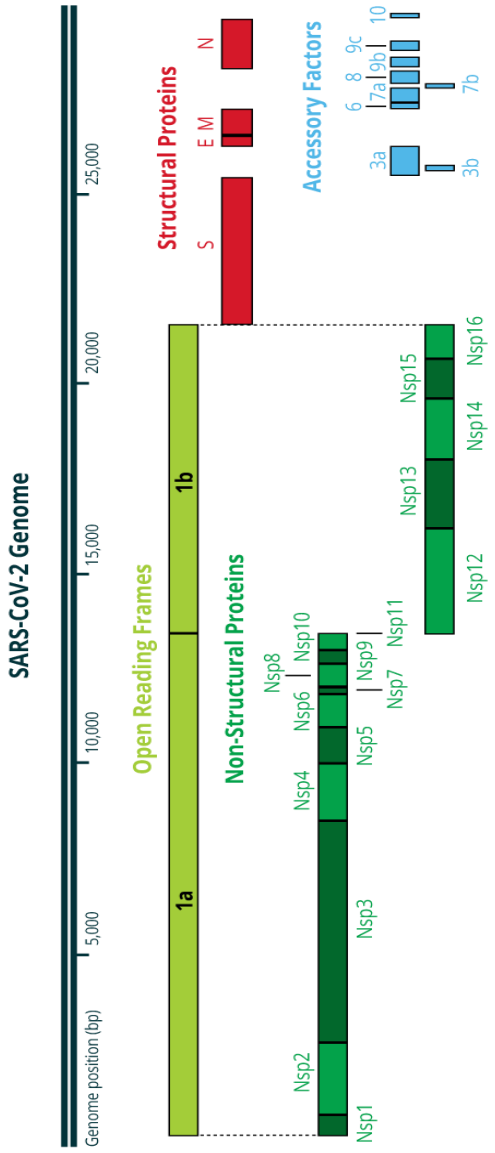


Figure 4: The SARS-CoV-2 genome layout. The viral genome encodes structural and non-structural proteins necessary for various functions in the viral recognition, entry, and infection (Adapted from (Gordon et al., 2020).

During viral infection, SARS-CoV-2 virus uses its spike to inject its genome into the host cell. The spike interact with ACE2 (human angiotensin-converting enzyme) to initiate SARS-CoV-2 attachment. Upon entry, the viral RNA is uncoated, released into the cytoplasm, and translated into respective protein products by the host ribosomes. Then, the viral proteases, PLpro and Mpro process and proteolytically cleave the translation products of the genome, for example, pp1a and pp1ab cleave into NSPs 1-16. Next, NSPs assemble into complexes and enzymes for replication and transcription, for example, the RNA-dependent RNA polymerase (RdRp), which is responsible for SARS-CoV-2 replication is formed by translated proteins NSP7, NSP8 and NSP12 (Yin et al., 2020). Then subgenomic mRNAs are translated to the structural proteins – S, M, E, N, and the accessory proteins.

1.3.1 SARS-CoV-2 and Host protein-protein interaction

In an attempt to understand the SARS-CoV-2 proteins interact with the host proteins, Gordon and his colleagues elaborated the interactions between SARS-CoV-2 proteins and human proteins associated with various cellular, molecular and physiological networks (Gordon et al., 2020). These individual interactions lead to various functions including DNA replication as seen in Nsp1, vesicle trafficking as seen in other non-structural proteins. Also, spike protein is involved in lipid modification, and Nsp8 is related to RNA processing.

Interestingly, several SARS-CoV-2 viral proteins target the major cellular proteins network that function in innate immune signaling by targeting the type I and III interferons (Park and Iwasaki, 2020). Viruses such as SARS-CoV-2 have developed unique ways to antagonize the IFN response. Previous studies of past coronaviruses have proposed that accessory proteins of SARS-CoV-1, such as ORF3b, ORF6, and ORF9b, inhibit the production of type I IFNs (Totura and Baric, 2012).

This is primarily the reason for the ability of the virus to evade host immune response. Therefore, focusing on the molecular mechanism and protein-protein interaction characteristics through which SARS-CoV-2 escape the antiviral mechanism through the interfering with the IFN production network will improve the understanding of the pathogenesis of COVID19 and provide therapeutic strategies to counteract SARS-CoV-2 infections.

1.3.2 SARS-CoV-2 ORF9b

The SARS-CoV-2 ORF9b is an alternative ORF located within the nucleocapsid (N) gene which codes for a 97 amino acid protein. The available structure of free ORF9b indicates a beta sheet-rich homodimer with a hydrophobic cavity in the center that binds lipids (Redondo et al., 2021). Interestingly, this protein has similar characteristics with its SARS-CoV homologue (ORF9b, 98aa), with about 72.4% similarities in amino acids composition. The most important finding about this protein is that it suppresses type I IFN production by binding TOM70. However, the molecular mechanism through which SARS-CoV-2 ORF9b by interacting with TOM70 inhibits the type I IFN response has not been fully elucidated (Jiang et al., 2020). TOM70 is a member of the translocase of the outer membrane complex that mediates chaperone-dependent import of preproteins in the cytosol into the mitochondria (Gava et al., 2011). TOM70 contains at least seven TPR domains in its cytosolic region. This is used to bind the Hsp90 C-terminal MEEVD motif (Young et al., 2003, Zanphorlin et al., 2016).

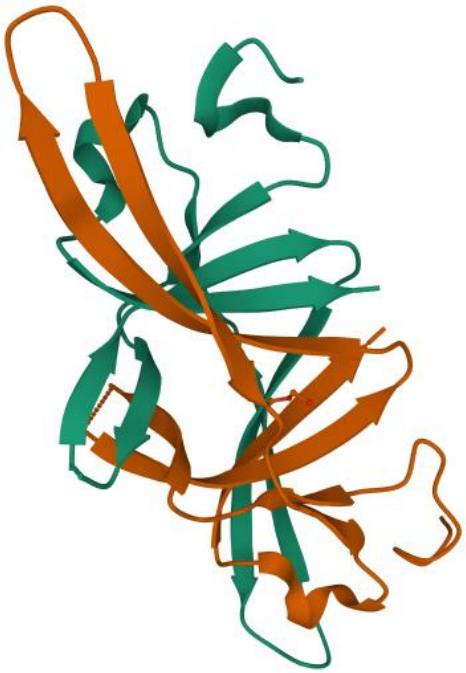


Figure 5: A X-ray Crystallographic Structure of recombinantSARS-CoV-2 ORF9b. (PDB: 6Z4U).

ORF9b is identified as a fold switcher due to the change in the orientation of the secondary structure when isolated and bound toTOM70. ORF9b alone folds into a β-sheet shape that forms a domain-swapped dimer, part of which switch into a long α-helix when bound to TOM70 (Porter, 2021).

A cryo–electron microscopy (cryo-EM) structure of the ORF9b-TOM70 complex reveals that ORF9b interacts at the TOM70 substrate binding site. ORF9b makes extensive hydrophobic interactions at the pocket on TOM70, with the interaction stabilized by four salt bridges. Furthermore, when comparing this complex with previously crystallized yeast TOM70 homologs, the interaction of ORF9b causes the helices on TOM70 to move inward and tightly wrap around ORF9b (Gordon et al., 2020). In a similar solved structure using X-ray crystallography, the binding of ORF9b induced conformational changes that are evident in the structure.

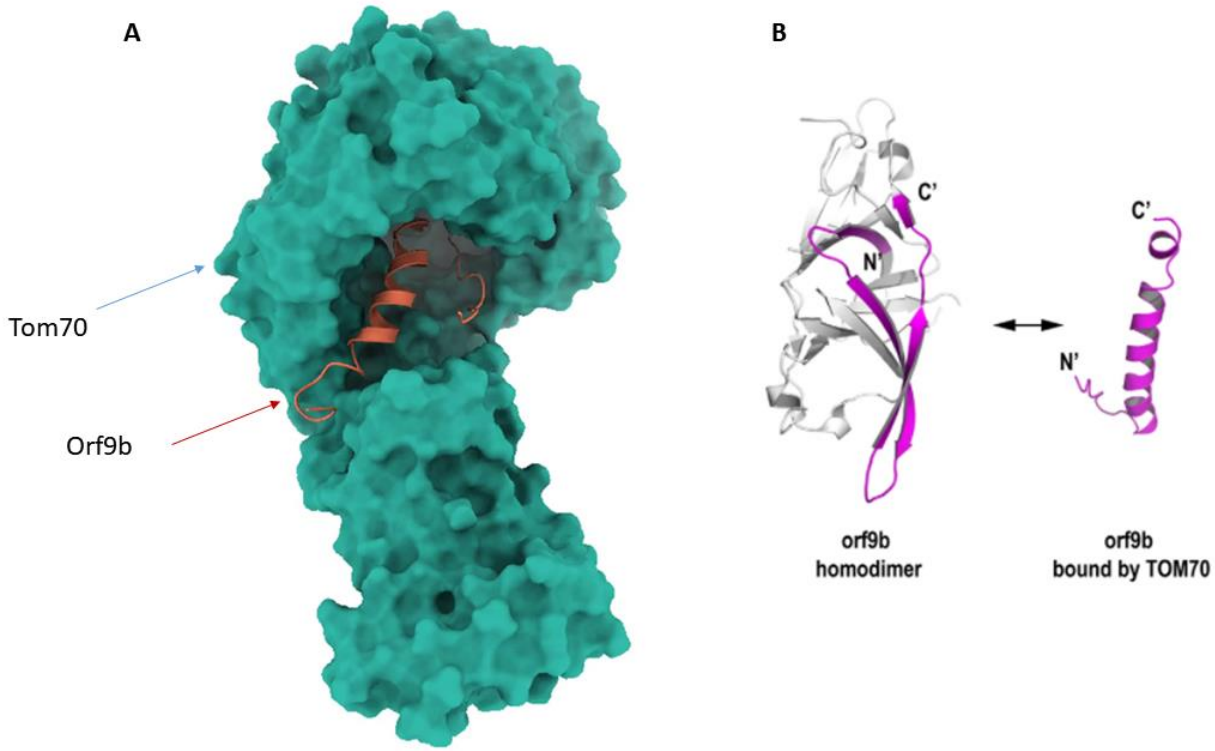


Figure 6: Surface representation of the ORF9b-TOM70 complex. (A) TOM70 and ORF9b shown in green surface and orange ribbon respectively. (B) An all-β homodimer ORF9b and an α-helix when bound to TOM70 (Adapted from (Gordon et al., 2020, Gao et al., 2021).

1.4 TOM70 and MAVS assembly

The involvement of TOM70 with the mitochondrial antiviral signaling protein (MAVS) in a cascade that produces INF-I leading to an innate immune response, which is a specific response to viral RNAs may be the most suitable explanation for the suppression of INF-I by the ORF9b-TOM70 interaction. During viral infection, the cytosolic helicase RIG-I recognizes the viral RNA which undergoes a conformational change and associates with MAVS. This triggers the association of MAVS with TOM70, which further initiates Hsp90 recruitment where the C-terminal EEVD motif of Hsp90 binds to TOM70 on the TPR domain (Liu et al., 2010). IRF3 is then phosphorylated by TBK1. This causes a dissociation of IRF3 from the MAVS complex. As a result, the IRF3 being phosphorylated translocates into the nucleus and promotes the transcription of genes encoding INF-I (F).

Conclusively, TOM70's interaction with EEVD motif of Hsp90 through the TPR domain is key to its function in the interferon pathway and induction of apoptosis in viral infection.

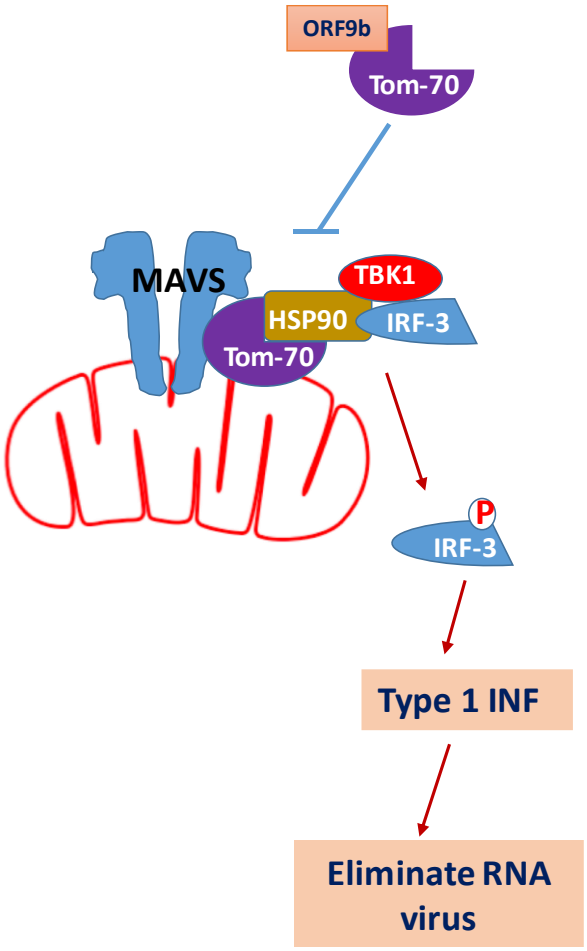


Figure 7: MAVS cascade. MAVS bind to TOM70 which subsequently leads to binding of the Hsp90/TBK1/IRF3 complex. TBK1 phosphorylates IRF3 which subsequently dissociates from the complex, promoting transcription of genes encoding IFN-I causing viral elimination. The binding of ORF9b to TOM70 disrupt this process

3.0 AIMS AND OBJECTIVES OF THIS STUDY

The following are the aims and objective of this study

I. Clone, express and purify ORF9b and produce a pure form of the recombinant ORF9bprotein, and characterize its structural and biophysical characteristics.
II. Co-express the ORF9b-TOM70 complex, determine its stability and characterize the biophysical characteristics of the complex.
III. Express and purify TOM70, Hsp90 and C-Hsp90, for protein-protein interaction studies with ORF9b

3.0 MATERIALS AND METHODS

3.1 Materials

All experiments were carried out under the guidance of prof. Dr. Carlos Ramos, and all materials were made available by his laboratory at the Institute of Chemistry, University of Campinas (UNICAMP), unless stated otherwise.

3.2 Methods

3.2.1 Plasmids and heat shock transformation

A codon-optimized SARS-CoV-2 ORF9b cloned into a pET28a with the insertion of a polyhistidine tag (His-tag) was purchased from GenScript (USA). The original 97 amino-acid sequence expressed to include an insertion of 20 amino acid residues containing 6 concomitant histidine residues and cleavage site for TEV protease. Other plasmids used to express the remaining proteins including (i) full-length human Hsp90 (UniProtKB - P07900), (ii) C-terminal domain of human Hsp90 (residues 566-732), and (iii) Cytosolic fragment of TOM70 (UniProtKB - O94826, residues 111-608) used in this project were obtained from already characterized plasmids stored in a -80 ^{o}C bio-freezer in the laboratory. The respective plasmids that encode other proteins are listed in the table below

Table 1: Plasmids information

cDNA	Vector	Organism	Antibiotics
ORF9b	pET28a	SAR-CoV-2	Kanamycin
TOM70	pProEx	Human	Ampicillin
Hsp90	pProEx	Human	Ampicillin
C-Hsp90	pProEx	Human	Ampicillin

All plasmids containing respective genes of interest were transformed into *E. coli* BL21(DE3). For transformation, BL21 competent cells were thaw in an ice bath at 4 ^{o}C and was incubated for 30 min after adding 1 µL of plasmid solution. After incubation cells were subjected to heat shock at 42 ^{o}C for 90 s and rapidly re-incubated on ice for 2 min. Then 800 µL of SOC medium (20 g/L peptone; 5 g/L yeast extract; 500 mg/L NaCl; 2.5 mM KCl; 10 µM $MgCl_2$; 20 µM glucose) was added and incubated at 37 ^{o}C for 60 min. In the case where we performed co-expression of ORF9b and TOM70, 1 µL of the respective plasmids were added into a single competent cell batch. The incubation at 37 ^{o}C was assisted with mild shaking for 2 hrs. After a successful transformation, cells were centrifuged at 8000 rpm (EppendORF-5418 centrifuge) for 30 s and the supernatant was discarded leaving about 10 µL. Cells were then plated on solid LB medium (10 g/L peptone; 5 g/L yeast extract; 10 g/L NaCl; 1.5% Agar) prepared in respective antibiotics, and cultivated at 37 ^{o}C in the incubator overnight. Cell growth was observed on plates the following morning, and plates were stored at 4 ^{o}C in the refrigerator.

3.2.2 Expression of recombinant protein

Pilot expression tests were performed for ORF9b to determine the best conditions for its expression in *E. coli* system. To prepare the pre-inoculum, three colonies were picked from the plate and dropped into a tube containing 10 mL LB medium (10 g/L peptone; 5 g/L yeast extract; 10 g/L NaCl) with 10 µL of 30 µg/mL kanamycin. This was placed in the shaker rotating at 200 rpm at 37 ^{o}C overnight. The following morning, 2.5 mL of the pre-inoculum was added to two different flasks containing 50 mL of LB medium. This flask was placed in a shaker at 37 ^{o}C/200 rpm and optical density (OD) at 600 nm was checked at intervals. At an OD of 0.6 – 0.8, 0.4 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) was added to induce protein expression. The two flasks were expressed separately at 20 ^{o}C and 37 ^{o}C respectively. Aliquots of 1.5 mL were taken from the medium after 1 h, 2 h, 3 h, 4 h and overnight. Aliquots were then centrifuged and 8000 rpm (EppendORF-5418 centrifuge) for 1 min, supernatant was discarded, the pellet was lysed, and the expressed protein was detected by SDS-PAGE.

For large-scale expression, pre-inoculum was prepared in 100 mL of liquid LB medium containing respective antibiotics for each plasmid. This was placed in the shaker rotating at 200 rpm at 37 ^{o}C overnight. Then, the following morning, 25 mL of the pre-inoculum was added into an Erlynmeyer's flask containing liquid LB medium and the same antibiotics then cultivated at 37 ^{o}C until the OD is between 0.6 and 0.8. At this point, a suitable concentration of IPTG was added and proteins were expressed at their respective best temperature and duration. ORF9b was expressed with 0.5 mM IPTG at 37 ^{o}C/200 rpm for 4 hrs. Other proteins were expressed according to the established laboratory protocol (Table 2). After protein expression, samples were centrifuged at 35000 rpm for 15 min at 4 ^{o}C. The pellets were kept in the freezer at -20 ^{o}C

Table 2: Conditions used for protein expression in *E. coli* BL21 (DE3)

Protein	IPTG conc. (mM)	Temperature (^{o}C)	Duration (hrs)
ORF9b	0.5	37	4
TOM70	0.8	37	4
ORF9b+TOM70	1.0	37	4
Hsp90	1.0	18	16
C-Hsp90	0.8	37	4

3.2.3 Bacteria cell lysis

Frozen pellets were re-suspended in 20 mL of lysis buffer composed of Na_2HPO_4/NaH_2PO_4 pH 7.5; 250 mM NaCl; 20 mM Imidazole; 1 mM PMSF; 30 µg/mL lysozyme; 5 units of *RNase-free* DNase which was incubated for 30 min on ice. Following this, cells were lysed in a sonicator (*Misonix Ultrasonic Liquid Processor*), at 10 pulse, 12 s intervals, 1.5 min, and at an amplitude of 40 W. The suspension was then centrifuged at 14000*g* for 20 min at 4 ^{o}C, supernatant was filtered with a 0.45 µM (Millex® GV) filter and ready to be applied for purification processes.

3.2.4 Protein purification

a. *Nickel affinity chromatography*

The first step for protein purification used is the nickel affinity chromatography because of the presence of His-tag in all the protein used in this study. Affinity chromatography is a powerful method for separating desired protein from complex mixtures (Porath, 1992). This involve the localization an interacting molecule referred to as affinity ligand onto a solid matrix to create a stationary phase, while the target molecule is in the mobile phase. In this case, electron donors present at the surface of the protein such as the thiol group of cysteine residues, indole groups of tryptophan, or imidazole group of histidine binds with coordinated metal on the solid phase (Spriestersbach et al., 2015).

Therefore, in this study, 5 mL HisTrap™ HP (GE Healthcare) pre-packed with nickel ions (Ni^{2+}), which can interact with the imidazole group present at the histidine residues of the poly-his tagged proteins was used. After column binding and washing steps, contaminants were removed, including untagged proteins and bacteria extracts. The bound protein was then eluted by competitive affinity using imidazole.

ORF9b was purified by coupling 5 mL HisTrap™ HP (GE Healthcare) with FPLC (Fast Protein Liquid Chromatography) AKTA system (Amersham Pharmacia Biotech). First, the filtered lysed protein (~10 mL) was loaded onto a column equilibrated with buffer A (25 mM Tris HCl pH 8.0; 200 mM NaCl; 20 mM imidazole), and 10 column volume (CV) of Buffer A was passed to remove protein contaminants that did not bind as flow through. After that, 30 CV linear percentage increase (0-100%) of buffer B containing 500 mM imidazole was passed through the system to elute the protein. The protein concentration peak was monitored on the chromatogram, and the fractions corresponding with the protein of interest was collected. The FPLC system was kept at a flow of 3 mL/min at a maximum pressure of 0.3 Mpa. The protein was dialyzed by loading into a 3,000 Da (Spectra/por-1) membrane placed in a beaker containing 2 L of buffer A, overnight under mild agitation.

However, other proteins were purified by affinity chromatography using the manual method. In this method, the FPLC system was excluded. His-tagged proteins were loaded onto the HisTrap™ HP (GE Healthcare) using a sterile syringe. Thereafter, column was washed by passing 3 CV Buffer A to remove contaminants. This was followed by preparing ascending percentage of 500 mM imidazole Buffer B solution (10%, 20%, 40%, 50%, 60%, 80%, and 100%). 3 CV of each was passed through the column slowly to elute the protein, and collected in falcon tubes. SDS-PAGE was performed to determine the fraction(s) that elutes the pure protein sample. See Table 3 for the buffers used for each protein.

b. Size exclusion chromatography.

After purification by affinity, a second step purification method known as size exclusion or gel filtration chromatography was used to achieve the purest form of the protein as possible. This method separated proteins according to their size and molecular mass. In addition to separating different proteins of varying size, one may resolve oligomeric forms of a particular protein, and it is generally used as a final purification step after at least one other purification step. Furthermore, it can be used to exchange the buffer of a sample for a different one (Duong-Ly and Gabelli, 2014).

The column used for size exclusion chromatography in this study was HiLoad Superdex 200 26/60 (GE healthcare), coupled with FPLC (Fast Protein Liquid Chromatography) AKTA system (Amersham Pharmacia Biotech). Proteins samples from the affinity chromatography step were injected into a pre-equilibrated column with suitable Gel filtration buffer (GFB) (Table 3). After injecting the protein, the column was eluted with a CV of GFB, at a flow rate of 2.5 mL/min and maximum pressure of 0.5 Mpa. Eluents were collected by the FPLC collection system and Absorbance at 280 nm was monitored to determine the elution volume of the proteins.

Table 3: Buffers used for affinity and gel filtration chromatography

Protein	Buffers
ORF9b	**A** 25 mM Tris HCl pH 8.0; 200 mM NaCl; 20 mM imidazole **B** 25 mM Tris HCl pH 8.0; 200 mM NaCl; 500 mM imidazole **GF** 25 mM Tris HCl pH 8.0; 200 mM NaCl.
TOM70	**A** 20 mM Tris HCl pH 7.4; 500 mM NaCl; 20 mM imidazole **B** 20 mM Tris HCl pH 7.4; 500 mM NaCl; 500 mM imidazole **GF** 25 mM Tris HCl pH 7.4; 200 mM NaCl.
ORF9b-TOM70	**A** 50 mM Tris HCl pH 8.0; 200 mM NaCl; 20 mM imidazole **B** 50 mM Tris HCl pH 8.0; 200 mM NaCl; 500 mM imidazole **GF** 25 mM Tris HCl pH 8.0; 200 mM NaCl.
Hsp90	**A** 20 mM Na_2HPO_4/NaH_2PO_4 pH 7.4; 500 mM NaCl **B** 20 mM Na_2HPO_4/NaH_2PO_4 pH 7.4; 500 mM NaCl; 500 mM Imidazole **GF** 20 mM Tris HCl pH 7.4; 150 mM NaCl
C-Hsp90	**A** 20 mM Na_2HPO_4/NaH_2PO_4 pH 7.4; 500 mM NaCl **B** 20 mM Na_2HPO_4/NaH_2PO_4 pH 7.4; 500 mM NaCl; 500 mM Imidazole **GF** 20 mM Tris HCl pH 8.0; 150 mM NaCl

c. *SDS-PAGE*

During all purification processes, fractions collected at each chromatography stage were analyzed by SDS-PAGE. The samples were mixed in a 1:1 ratio with the electrophoresis loading buffer (Tris-HCl 50 mM pH 6.8; DTT 100 mM; SDS 2%; bromophenol blue 0.1%; glycerol 10%). The samples and the molecular mass (MM) marker (*Broad Range Promega*) were applied in a 12% SDS polyacrylamide gel (SDS-PAGE), followed by electrophoresis at 40 mA, 300 V for 1 h. After electrophoresis, gels were colored with *Coomassie Brilliant Blue*

and discolored with a solution containing ethanol:acetic acid:H_2O 3:2:35. Bands corresponding to a protein of interest were pictured and analyzed by the *ImageJ* program (Schneider et al., 2012).

d. *TEV Protease*

A stable variant of tobacco etch virus (TEV) protease was used for the cleavage of the His-tagged proteins; ORF9b and Hsp90. The ENLYFQG sequence present at the poly-his tag serves as the site for TEV protease. TEV protease recognizes the amino acid sequence ENLYFQ/G with high efficiency (Raran-Kurussi et al., 2017). Each protein was mixed with TEV protease in the ratio 1:10 mg/mL (TEV protease:Protein). The mixture, kept in a membrane (Spectra/por-1), was placed in a beaker containing GF buffer under mild agitation for 4 hrs. Following this, the mixture was passed through another affinity chromatography so that the cleaved protein (with no His-tag) flows through the nickel column, while the poly histidine residues interact and bind with the column.

e. *Western blot*

In order to confirm a clean cleavage of the his-tag, western blot was performed against anti-his antibody, in this case, Mouse Anti-His antibody (GE Healthcare). Protein bands were transferred from SDS-PAGE gel into nitrocellulose membrane in transfer buffer (Tris base 3.0g; glycine 14.4g; SDS 0.1%; Methanol 20%) using a western blot chamber, at 80 mA, 300 V for 1 h. The membrane was then blocked with 5% BSA for 1 h on a shaker, then stored overnight at 4 ^{o}C. On the following day, the primary antibody, mouse anti-his antibody (1:3000) was added to the membrane, and incubated for 1 h. This was followed by the applying a secondary antibody, goat polyclonal IgG (HRP) incubated for 1 h. Thereafter, the membrane was treated with ECL™ western blotting reagents for detection with a chemiluminescence detector device (Amersham imager 600). All antibodies used were prepared in TBST (10% TBS; 0.1% Tween20).

f. *Determination of protein concentration*

Protein concentration is measured using the Edelhock method (Edelhoch, 1967), in which absorbance of the protein sample was measured at 280 nm, assuming all tryptophan, tyrosine, and cysteine residues are completely exposed in the presence of 6M guanidine chloride (GdmCl) in 25 mM Na_2HPO_4/NaH_2PO_4, pH 6.5. Prior to measurement, protein was incubated with the buffer for 10 min at room temperature, then absorbance was measured using a UV-Vis spectrophotometer device. The absorbance value was then converted to concentration using the Beer-Lambert equation as follows:

$$A_{280nm} = \varepsilon.l.C$$

Where A is the absorbance, l is the optical path length (cm), Ɛ is the molar absorption coefficient, and C the molar concentration. The molar absorption coefficient is a sample dependent property, and is calculated taking into consideration the molar absorbance coefficient of tryptophan, tyrosine and cysteine residue (Pace et al., 1995). Therefore, it is calculated according to the equation below:

$$\varepsilon = (n^oTrp*5500) + (n^oTyr*1490) + (n^oCys*125)$$

3.2.5 Circular dichroism (CD) spectropolarimetry

Circular dichroism spectropolarimetry (CD) is an efficient technique for rapidly evaluating the folding, stability, and binding properties of proteins. Using this technique, we can get a very good estimation of the fraction of the residues involved in α-helix, β-sheet or disorderly formation (random coil) (Correa and Ramos, 2009). For example, proteins rich in α-helical have negative bands at 222 nm and 208 nm and a positive band at 193 nm, β-pleated sheets (β-helices) rich proteins have negative bands at 218 nm and positive bands at 195 nm. In comparison, disordered proteins have very low ellipticity above 210 nm and negative bands near 195 nm (Greenfield, 2006); Figure 9).

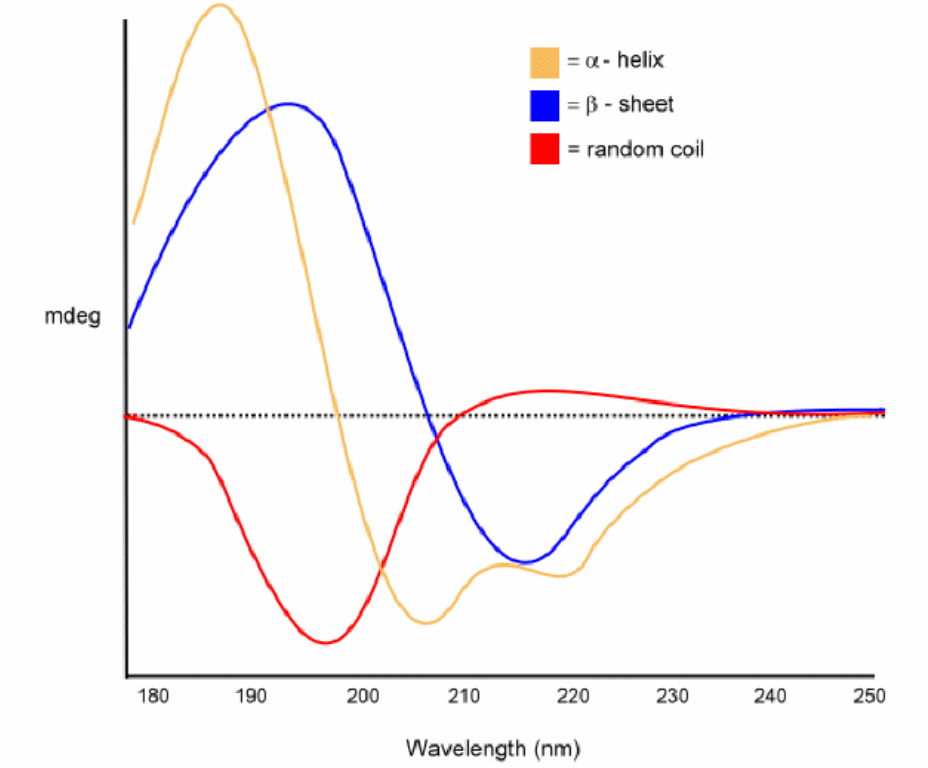


Figure 8: Circular dichroism profile of protein secondary structures. Showing CD spectra of helix, sheet and random coil. (Adapted from (Correa and Ramos, 2009).

CD is also used to study the stability of a protein by increasing denaturant conditions such as temperature, chemicals (usually urea and guanidinium chloride) and extremes of pH (usually acidic). As the denaturant condition increases the stability of the protein decreases and then unfolds. When proteins are exposed to denaturant condition, the spectra of folded and unfolded differ from each other. The unfolding transition can be easily determined by choosing a wavelength where the difference in signal for folded and unfolded protein is large (e.g. 222 nm in α-helical proteins) (Correa and Ramos, 2009).

In this study, each CD spectrum was read using spectropolarimeter J-720 (JASCO) device under constant flow of 10 L/min N_2. 5 µM of protein samples were pipetted into a 2 mm optical length quartz and placed in the device. Each final spectrum was theaverage of 16 different readings at 200 to 260 nm wavelength measured at 20 nm/min with 1 s response time. All reading was maintained at 25 ^{o}C with a coupled temperature controller (*Peltier Type Control System PFD 4255-jasco*). Results are expressed as mean residue molar ellipticity, [θ].

$$[\theta] = \frac{\Theta \,.\, MM \,.\, 100}{C \,.\, l \,.\, n}$$

Where θ is the ellipticity in degrees, l is the optical path in cm, C is the concentration in mg/mL, M is the molecular mass and n is the number of residues in the protein. The mean residue molar ellipticity [θ] is given as deg.cm^{2}.dmol^{-1}.

Also, for ORF9b, samples were prepared with varying concentrations of trifluoroethanol (TFE; 1 to 20%). Low TFE concentrations is often used to increase the propensity of certain polypeptides/proteins to form secondary structures, especially helical conformation (Correa and Ramos, 2009). The temperature-induced unfolding (20-90 ^{o}C) profiles of proteins (TOM70 and ORF9b-TOM70) were determined by following the CD signal at 222 nm using 5 µM protein samples in a 2 mm quartz cuvette. Then, the temperature at the middle of the transition (Tm) was determined.

3.2.6 Fluorescence

The intrinsic tryptophan fluorescence spectrometry allows the use of tryptophan residues to obtain information about local protein conformation. When the protein, or the region around the residue, is well folded, tryptophan is not exposed but remain in a non-polar environment, emitting fluorescence at a maximum intensity of less than about 340 nm and can be as low as 310 nm. However, when the protein, or the region around the residue, is unfolded or not properly folded, the tryptophan residue is exposed to the polar environment and fluoresce at wavelengths higher than 340.

In this study, 10 µM of TOM70 and ORF9b-TOM70 complex was pipetted into a four-faced 1 cm optical path quartz cuvette and placed in the fluorimeter, with temperature maintained at 25 ^{o}C. Fluorescence intensity was measured at 300-400 nm at an excitation at 295 nm. Maximum wavelength was determined from data analysis.

3.2.7 Differential scanning calorimetry (DSC)

DSC is used for measuring directly the thermodynamic parameters that characterize biomolecules. This thermo-analytical technique determines the thermal transitions of protein alone or in complex with other protein or ligands. Tm, the temperature at the middle of the transition, is used as an indicator of thermo-stability and generally, the higher the Tm, the more stable the protein is. Therefore, proteins with higher Tm are less susceptible to unfold at relatively lower temperatures (Bruylants et al., 2005).

In the DSC system, there are two cells; the sample cell and the reference cell, both being supplied with the same temperature. The reference cell contains the same buffer as that of the protein sample, therefore, during scanning, the instruments monitor a small temperature difference between the reference cell and the sample cell that contains the protein of interest in an identical buffer.

As the temperature of both cells is increased, thermally induced processes occurring in the sample cell result in heat being generated or consumed and this

produces a change in the temperature difference relative to the reference cell. Also, there are presence of heaters on the sample cell surface that work in a feedback circuit to input additional electrical power in order to return the temperature difference to its initial value (Johnson, 2013).

This method was used to determine the thermodynamic parameters and stability of TOM70 and ORF9b-TOM70 complex. Both proteins were kept in the same buffer (25 mM Tris HCl pH 8.0; 200 mM NaCl), and at the same concertation (20 µM). Blank scans from 20-90 ^{o}C had the reference cell and the sample cell containing the buffer. After this, blank from the sample cell was removed and replace with protein sample (Bittar et al., 2003) andscaned from 20-90 ^{o}C. Experiments were performed using a *MicroCal VP-DSC (Malvern),* and data were analyzed using *Origin*® *VP-DSC* software. Data were normalized by subtracting the buffer from the sample, and fitted to generate values of Tm and enthalpy change.

3.2.8 Size exclusion chromatography coupled with multi-angle light scattering (SEC-MALS)

In this technique, size exclusion chromatography (SEC), used to determine of the molecular mass (MM) of proteins and protein-protein complexes in solution is coupled with multiangle light scattering (MALS), an absolute technique that determines the MM of an analyte in solution from basic physical equations (Some et al., 2019). This device was set up with a 25 mL Superdex 200 10/300 GL (*GE Healthcare*) connected to a ÄKTA pure protein purification system (*Amersham Pharmacia Biotech*), that controls the flow and reads the absorbance of protein at 280 nm. This is then coupled with two other devices: *miniDAWN TREOS (Wyatt Technologies)*, which uses three light detectors to detect light scattering at 3 different angles; and *Optilab T-rEX (Wyatt Technologies),* which reads the diffractive index of samples to determine MM (Wyatt, 1993).

In this study, this method was used to determine the molecular mass of individual proteins, and to determine protein-protein interaction. For determination of proteins' MM, 500 µL of samples filtered with 0.22 µM filters was injected into system through the SEC chromatographic column. Flow was maintained at 0.5 mL/min and maximum pressure of 1.5 Mpa with their respective GF buffers. Data was analyzed with ASTRA *(Wyatt Technologies)* software.

In the case of protein-protein interaction studies, interaction between (i) Hsp90 and TOM70, (ii) Hsp90 and ORF9b-TOM70, (iii) C-Hsp90 and ORF9b, and (iv) C-Hsp90 and ORF9b-TOM70 was determined. Protein mixtures were incubated for 4h at 4 ^{o}C at same concentration. Post incubation, the samples were applied and elution fractions were collected and analyzed with 12% SDS-PAGE. Characteristic profiles or samples such as MM, and SDS-PAGE band of each peak was used to determine the formation of complex and protein-protein interactions.

3.2.9 Pull-down assay

The pull-down assay is an *in vitro* technique used to determine physical interactions between two or more proteins. It is efficient in confirming protein–protein interaction or identifying new interacting biomolecules. The method involve the localization of a tagged protein (bait) on an affinity ligand specific to the tag, such as Nickel, creating a solid phase (prey) that interact with the bait. In this case, his-tagged proteins are used for bait to capture TEV protease-cleaved protein.

In this study, 3 mL 20 µM His-tagged TOM70 and ORF9b-TOM70 complex were immobilized in separate experiments into the 5 mL HisTrap™ HP (GE Healthcare) nickel column. This was followed by 3 CV wash with buffer A to make sure unbound protein is washed away from the column. Then, 3 mL of 20 µM untagged Hsp90 was injected into the column and incubated for 30 min at 4 ^{o}C. After incubation, the column was washed again with 3 CV wash with buffer A to removed protein that failed to interact. Finally, interacting proteins were eluted with 100% buffer B containing 500 mM imidazole. Samples were collected at each stage and analyzed by SDS-PAGE. See table 3 for buffer description.

4.0 RESULTS

4.1 Bioinformatics

The 97 amino acid (aa) sequence of ORF9b, analyzed using *Protparam Expassy*™, calculated a MM of 10.8 kDa and the theoretical PI 6.59. The protein has equal positive and negatively charged residues (Asp + Glu = 11; Arg + Lys = 11). The protein does not contain any tryptophan in its residue; therefore, the extinction coefficient is determined to be 1490 (mg/ml) $^{-}$1 cm $^{-}$1.

A sequence alignment tool was used for a pairwise alignment of SARS-CoV-2 ORF9b and its homologue from SARS-CoV. The similarity index was approximately 72.4% identity in amino acid composition and 84.7% similarity in amino acid groups (Figure 9).

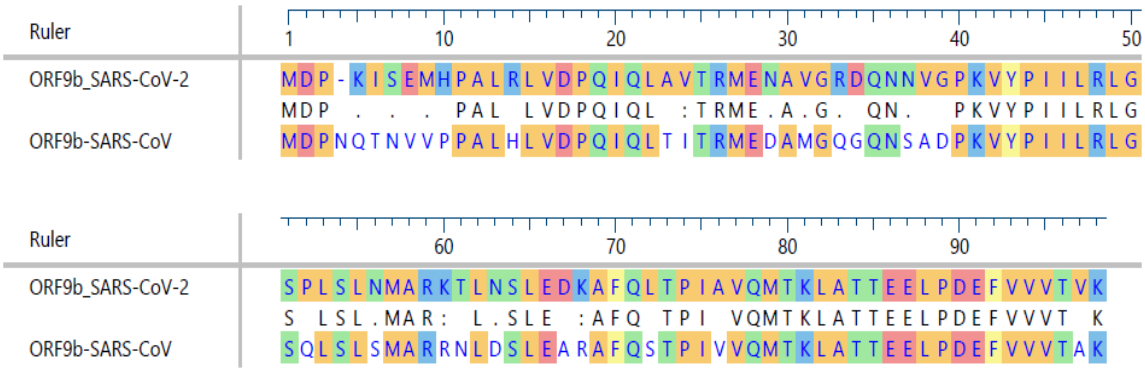


Figure 9: Pairwise alignment of ORF9b from SARS-CoV-2 and SARS-CoV. The middle residue sequence shows identical residues while single and double dots show similarity.

A sequence chart of all proteins used in this study is presented below (Figure 10).

Figure 10: Sequence chat of expressed protein. All proteins are expressed to contain the sequence shown. Purple letters show the TEV protease cleavage site, while blue letters show the polyhistidine tag. TOM70 is a cytosolic fragment of the whole protein from residue 111-608. C-Hsp90 is the C-terminal domain of Hsp90 from reside 566-732.

4.2 Expression and purification

Expression tests (pilots) were carried out on ORF9b to determine the optimal conditions for expression. In this step, the protein was expressed in the *E. coli* system at different temperatures (20 and 37 ^{o}C) and at different durations of induction (1 h, 2 h, 3 h, 4 h, and overnight). Fractions of induced medium were collected, centrifuged, and lysed to separate soluble and insoluble fractions. These fractions were analyzed by SDS-PAGE (Figure 11). The proteins expressed were all present in the soluble fractions, and the best conditions for expression was determined to be at 37 ^{o}C and 4 hrs induction time. These conditions were then repeated at large-scale protein expression for ORF9b.

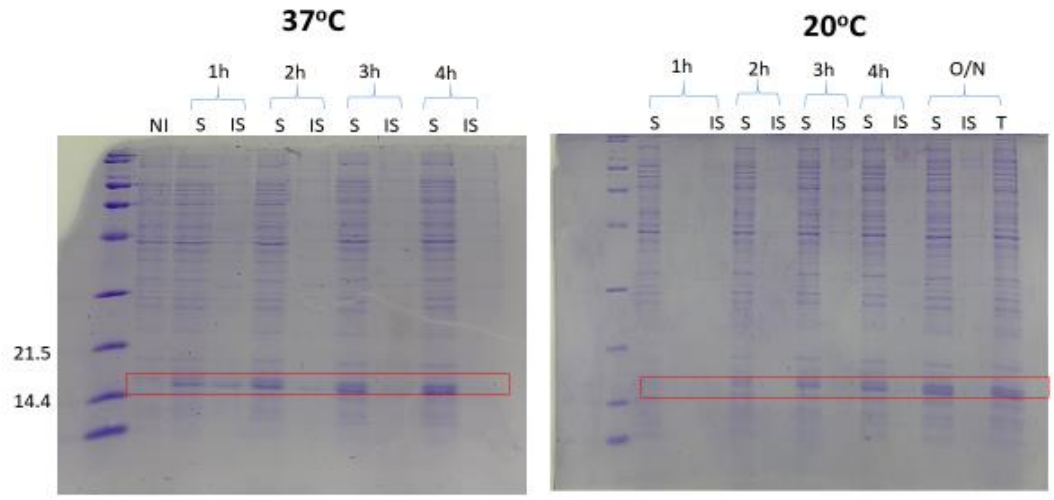


Figure 11: Expression test of ORF9b at different temperatures and duration. The plasmid pET28a-ORF9b was transformed into *E. coli* BL21 cell, expressed in LB medium containing 30 µg/mL kanamycin. The lysed samples obtained at these conditions were analyzed by 12% SDS-PAGE. NI: samples not induced, S: soluble fraction after centrifugation, IS: insoluble fraction, T: total protein.

Other proteins were expressed according to the conditions stated above (Table 2) and purified using a two-stepped purification method of affinity chromatography and gel filtration chromatography. Table 4 shows the percentage purity of protein after GF chromatography as analyzed by Image J, and the amount of protein yield per liter of induction.

Table 4: Protein purity and yield

Protein	Purity (%)	Yield (mg/L)
ORF9b	97	46
TOM70	96	72
Hsp90	91	26
C-Hsp90	95	36
ORF9b-TOM70	96	64

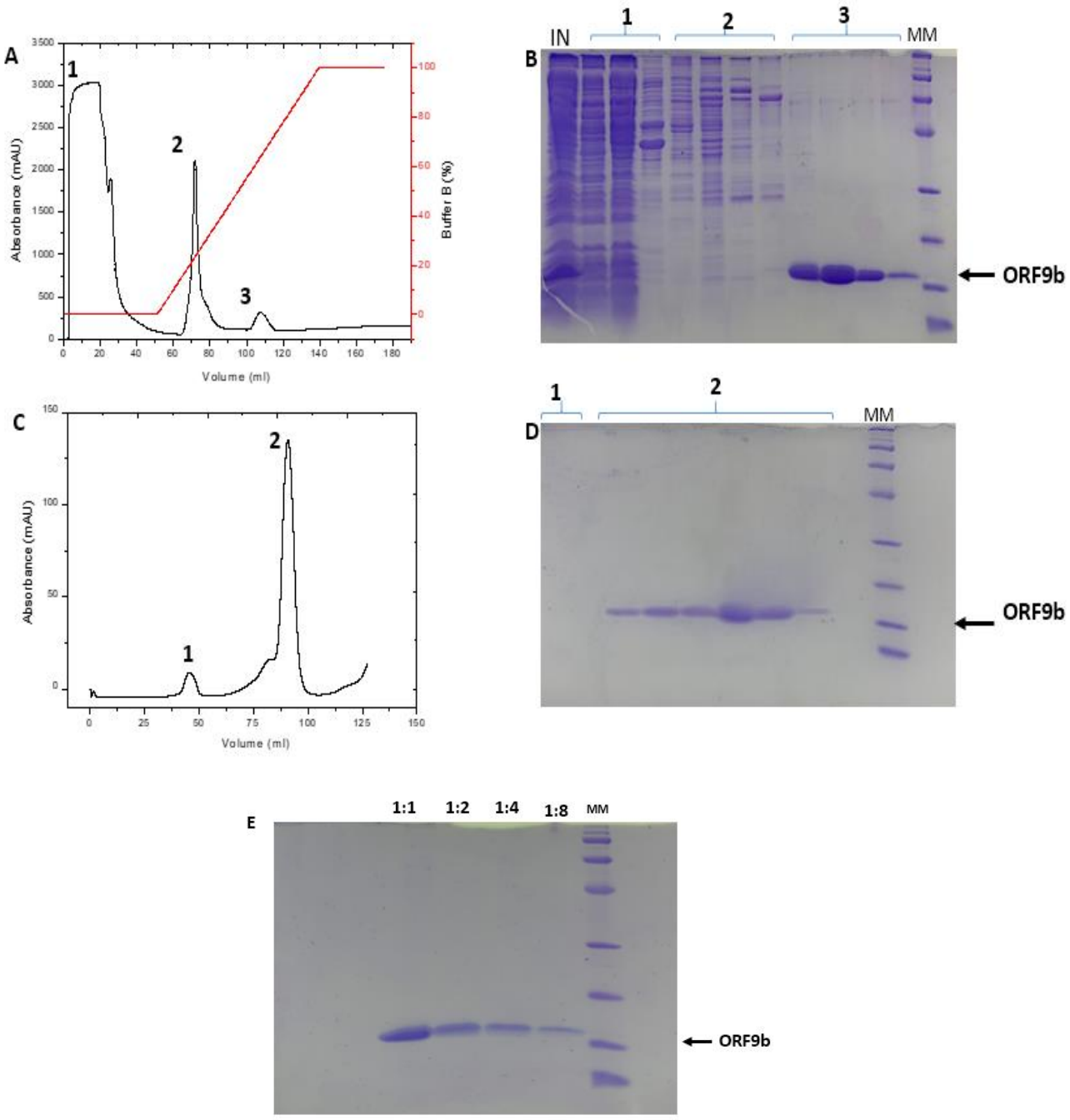


Figure 12: Purification of ORF9b by affinity and gel filtration chromatographies. **(A)** ORF9b affinity chromatogram, showing three pics. The red line indicates the graduation percentage of 500 mM buffer B. **(B)** SDS-PAGE analysis of the fractions from affinity chromatography. IN, input, and peaks 1,2,3 are shown. Peak 3 corresponds with the ORF9b of interest. **(C)** Gel filtration chromatogram of ORF9b shows two peaks **(D)** SDS-PAGE analysis of peaks from the collected sample from GF chromatography. Peak 2 corresponds to a pure fraction of ORF9b. 97 % protein purity was achieved **(E)** SDS-PAGE analysis of ORF9b after dialysis. Protein was fractionated into dilutions in the above ratios (ORF9b:Loading buffer), analysis from image J showed a reduction in density of the bands as a result of dilution (75, 12, 8 , 4 % respectively). MM indicates molecular weight marker.

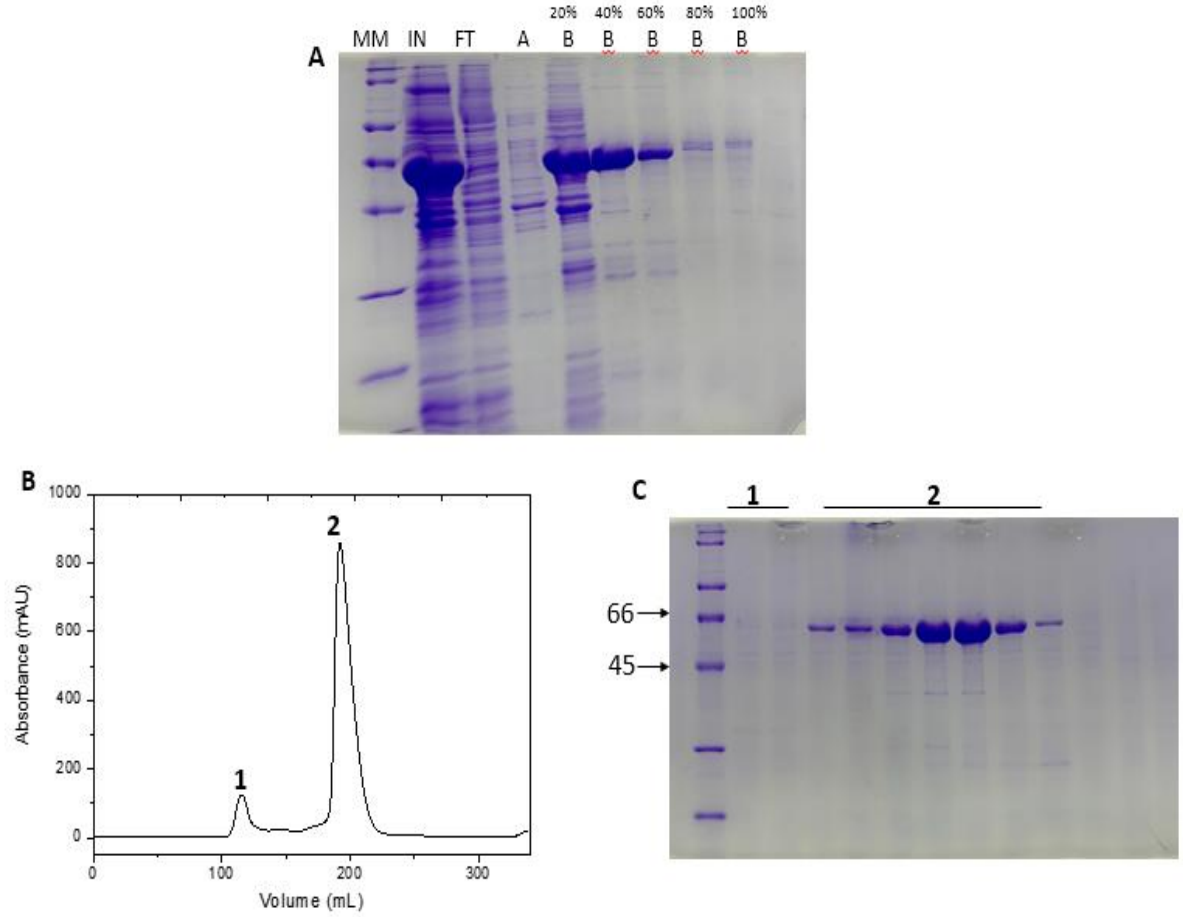


Figure 13: Purification of TOM70 by affinity and gel filtration chromatographies. (A) SDS-PAGE analysis of the fractions from affinity chromatography. MM: molecular mass marker; IN: sample input; FT:flow-through; A: Buffer A; B: Buffer B containing 500 mM imidazole. Protein was eluted by 20, 40, and 60% 500 mM imidazole **(B)** Gel filtration chromatogram of TOM70 showing two peaks **(C)** SDS-PAGE analysis of peaks collected from GF chromatography. Peak 2 corresponds to a pure fraction of TOM70. 96 % protein purity was achieved as determined by Image J.

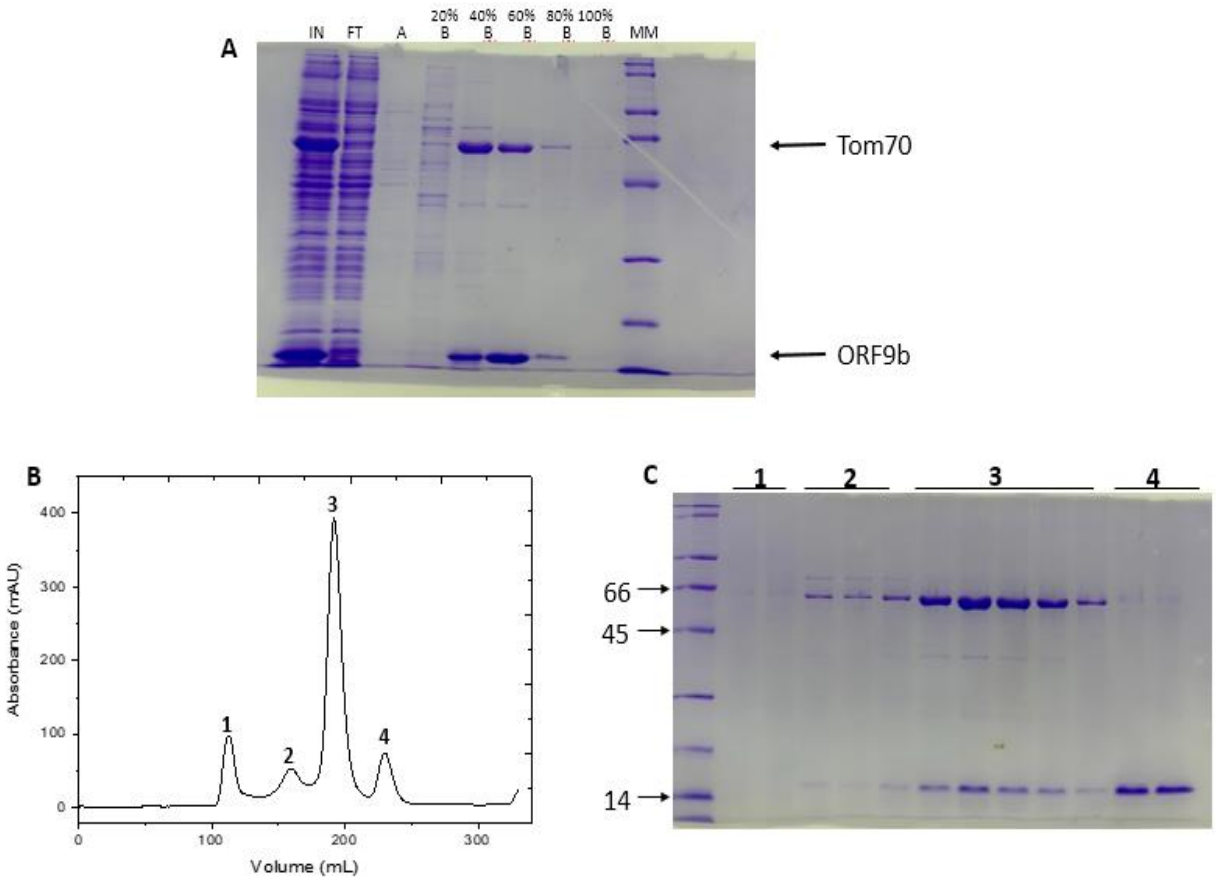


Figure 14: Purification of co-expressed ORF9b-TOM70 by affinity and gel filtration chromatograpies. (A) SDS-PAGE analysis of the fractions from manual affinity chromatography. MM: molecular mass marker; IN: sample input; FT:flow-through A: Buffer A; B: Buffer B containing 500 mM imidazole. Protein was eluted by 40 and 60% 500 mM imidazole **(B)** Gel filtration chromatogram of ORF9b-TOM70 showing 4 peaks **(C)** SDS-PAGE analysis of peaks collected from gel filtration chromatography. Peak 3 corresponds to a pure fraction of the ORF9b-TOM70 complex. 96 % protein purity was achieved as determined by Image J.

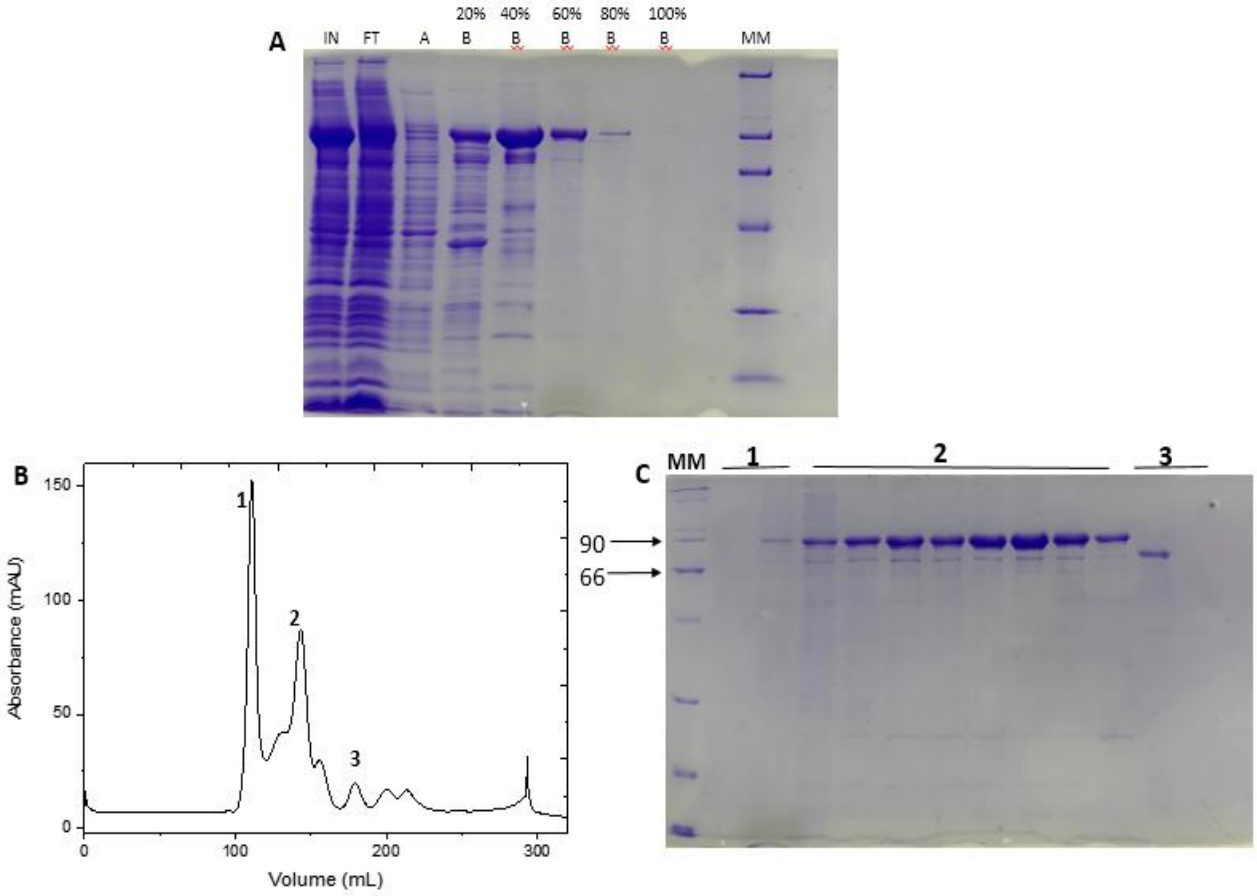


Figure 15: Purification of Hsp90 by affinity and gel filtration chromatographies. (A) SDS-PAGE analysis of the fractions from affinity chromatography. MM: molecular mass marker; IN: sample input; FT:flow-through; A: Buffer A; B: Buffer B containing 500 mM imidazole. Protein was eluted by 20, 40, and 60% 500 mM imidazole **(B)** Gel filtration chromatogram of Hsp90 showing 3 major peaks **(C)** SDS-PAGE analysis of peaks collected from gel filtration chromatography. Peak 2 corresponds to a pure fraction of Hsp90. 91 % protein purity was achieved, as determined by Image J.

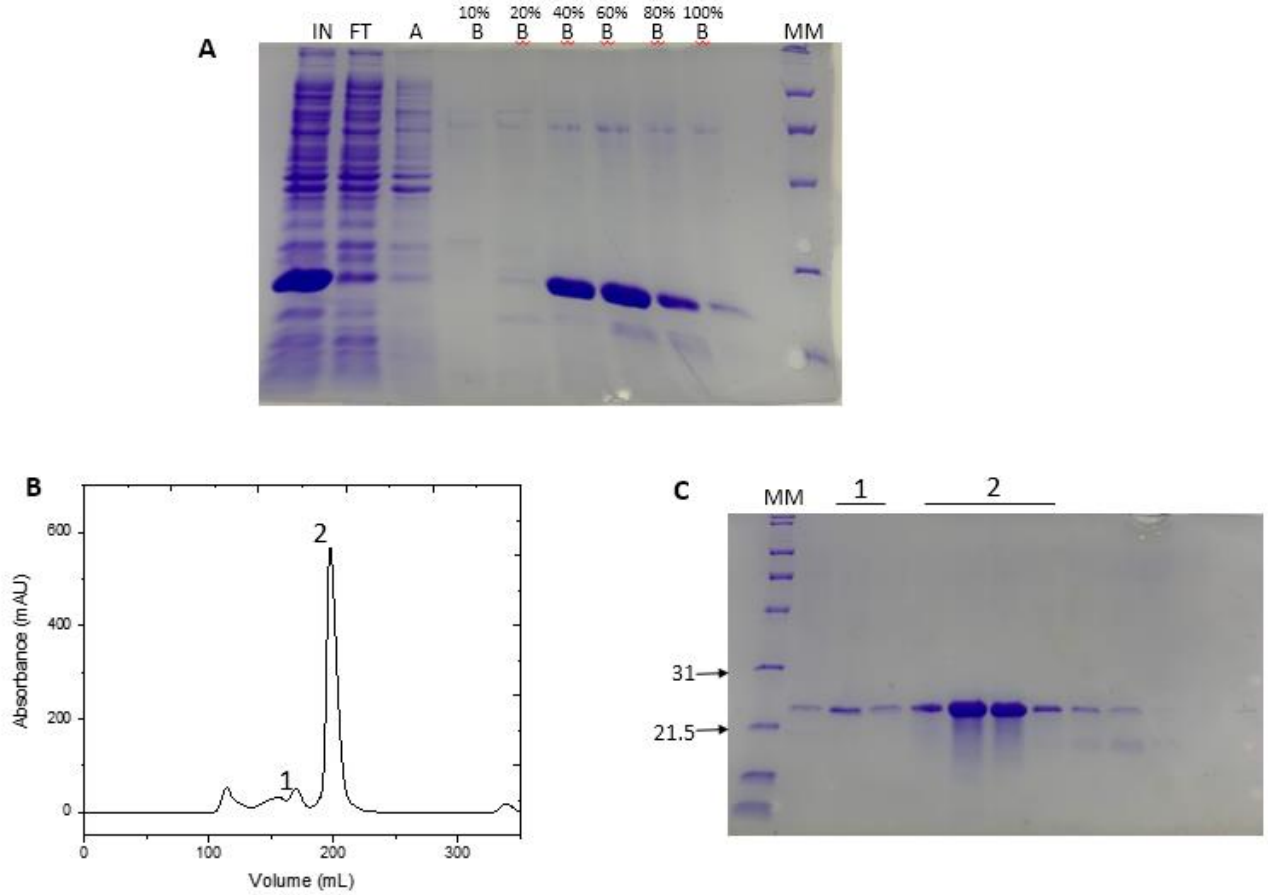


Figure 16: Purification of C-Hsp90 by affinity and gel filtration chromatographies. (A) SDS-PAGE analysis of the fractions from manual affinity chromatography. MM: molecular mass marker; IN: sample input; FT:flow-through; A: Buffer A; B: Buffer B containing 500 mM imidazole. Protein was eluted by 40, 60, and 80% 500 mM imidazole **(B)** Gel filtration chromatogram of C-Hsp90 showing 2 major peaks **(C)** SDS-PAGE analysis of peaks collected from gel filtration chromatography. Peak 2 corresponds to a pure fraction of C-Hsp90. 95 % protein purity was achieved as determined by Image J.

The poly histidine tag of ORF9b and Hsp90 was successfully cleaved with TEV-protease, with a second affinity purification showing that each protein was free of the His tag. This was further confirmed by western blot against anti-His antibody. The cleaved protein did not react with the anti-His antibody, and no protein band was visible (Figure 17).

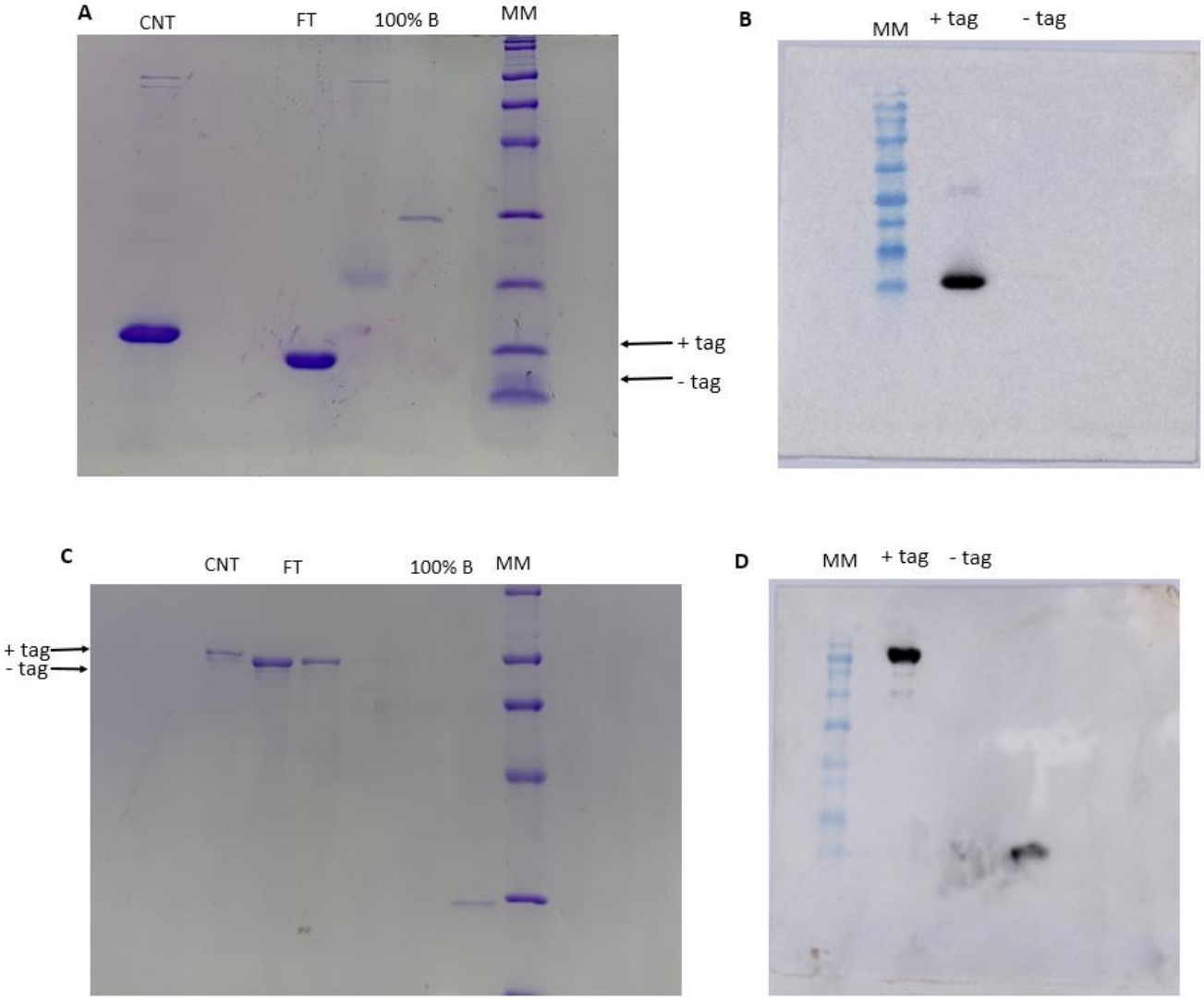


Figure 17: TEV protease cleavage of his-tag. (A) SDS-PAGE analysis of the second affinity chromatography of ORF9b after incubation with TEV-protease. CNT: control (His-ORF9b), MM: molecular mass marker; FT: flow-through (ORF9b).; 100% B: Buffer B containing 500 mM imidazole. **(B)** Western blot analysis against anti-his antibody. +tag: purified His-ORF9b; -tag: purified ORF9b after treatment with TEV. **(C)** SDS-PAGE analysis of the second affinity chromatography of Hsp90 after incubation with TEV-protease. CNT: control (His-Hsp90), MM: molecular mass marker; FT: flow through (Hsp90).; 100% B: Buffer B containing 500 mM imidazole. **(D)** Western blot analysis against anti-his antibody. +tag: purified His-Hsp90; -tag: purified Hsp90 after treatment with TEV.

4.3 Structural and biophysical characterization of ORF9b

Studies of ORF9b using SEC-MALS showed that the MM is 22.0 ± 1.0 kDa (Figure 18A), confirming the homodimeric assembly of the protein (ORF9b is predicted to be 10.8 kDa). However, looking into the structural characteristic and secondary structure assembly, CD revealed a spectrum with more signal below 205 nm, indicating that the protein is predominantly composed of random coils (Figure 18B). Incubation with TFE, an inductor of secondary structure in proteins, caused no relevant modification in the CD spectrum (Figure 18B).

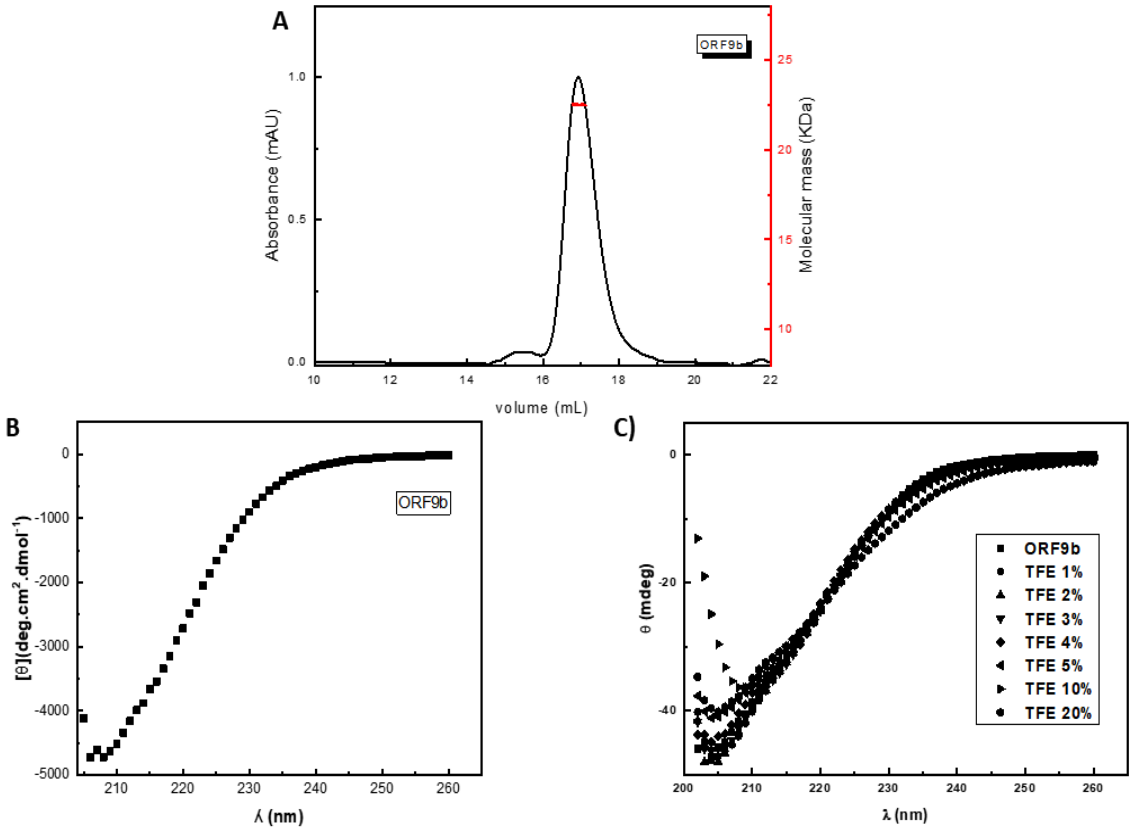


Figure 18: Conformational characterization of ORF9b. (A) SEC-MALS analysis of ORF9b using *miniDAWN TREOS* and *Optilab T-rEX*, and data was analyzed with ASTRA *(Wyatt Technologies)* software. Calculated MM is 22.0 ± 1.0 kDa. **(B)** Circular dichroism spectrum with mean residue elipticity [θ] measured from 205-260, spectrum indicate a random coil predomination. **(C)** CD spectra of ORF9b after treatment with TFE.

4.4 Characterization of structural, biophysical and thermal stability of ORF9b-TOM70 complex

The MM of ORF9b-TOM70 complex was determined by SEC-MALS as 67.0 ± 2.0. This indicates that a complex was formed with 1:1 (ORF9b:TOM70) stoichiometry (Figure 19A). The absorbance of the ORF9b-TOM70, TOM70, and ORF9b elution fractions alone by SEC-MALS showed a shift in the peak of elution of these proteins (Figure 19B).

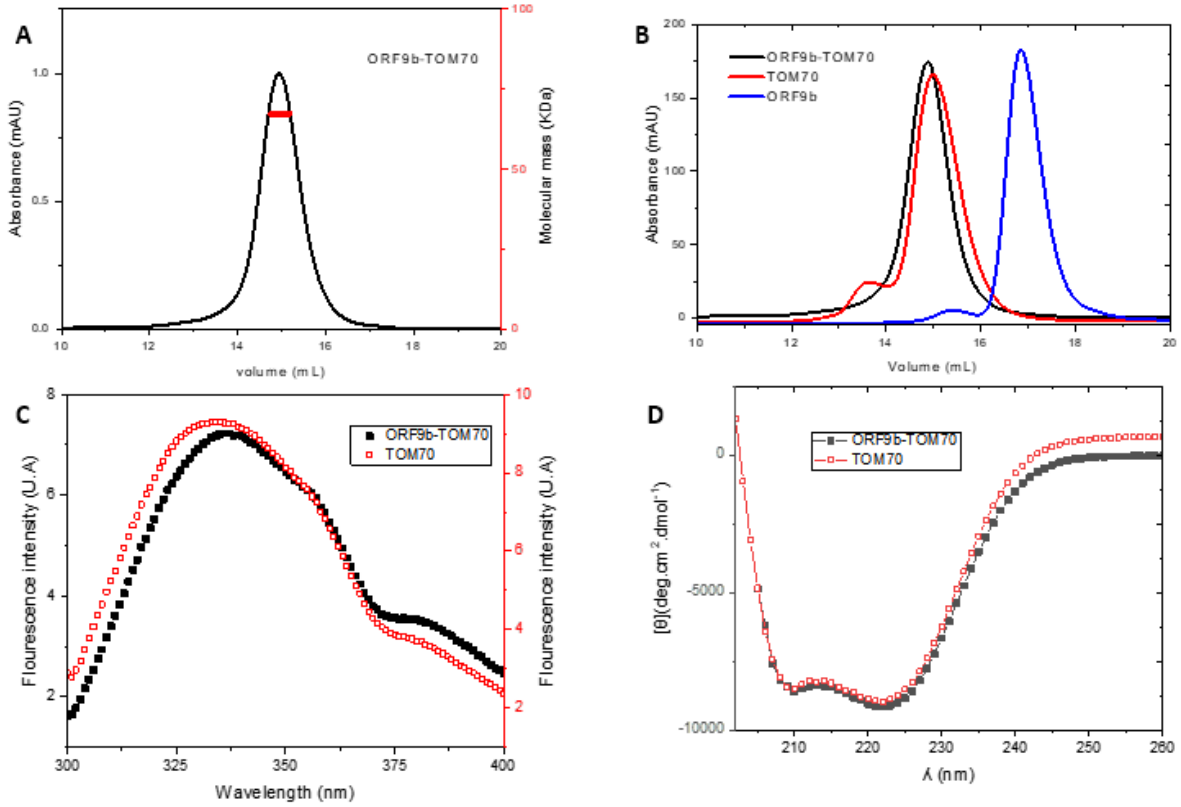


Figure 19: Characterization of ORF9b-TOM70 complex. (A) SEC-MALS analysis of ORF9b using *miniDAWN TREOS* and *Optilab T-rEX* and data were analyzed with ASTRA *(Wyatt Technologies)* software. The calculated molecular mass was 67.0 ± 2.0 kDa. **(B)** SEC-MALS experiment showing the elution volume of each individual protein and the complex. Elution is dependent on the molecular mass; ORF9b-TOM70 was eluted before TOM70 alone, indicating the presence of ORF9b in the complex. **(C)** Fluorescence emission spectra of ORF9b-TOM70 and TOM70 were obtained with proteins kept in GF buffer (Table 3). On excitation at 295 nm, the maximum wavelength determined for ORF9b-TOM70 and TOM70 alone is 338 ± 1 nm and 339 ± 1 nm, respectively. **(D)** CD spectra of complex and TOM70 alone measured at 200-260 nm showing α-helical structure with minimum signals at 222 and 208 nm.

Furthermore, tryptophan fluorescence was measured at 300-400 nm with an excitation at 295 nm. The maximum wavelength determined for ORF9b-TOM70 and TOM70 alone is 338 ± 1 and 339 ± 1 nm, respectively, indicating Trp was well buried (Figure 19C). CD was used to determine the folding and secondary structure characteristics of the ORF9b-TOM70 complex, which was compared to TOM70 alone; CD revealed two minimum signals of -9000 and -8500 deg.cm^2.dmol^{-1} at 222 and 208 nm, respectively, indicating that the protein is predominantly composed of α-helices. This is similar to the TOM70 signals (Figure 19D).

Another method used with CD investigates the thermal stability. The temperature range scan at 222 nm from 20 to 90 ^{o}C was used to determine the Tm, temperature at the middle of the transition, value. Tm values is 58 ± 1 ^{O}C for ORF9b-TOM70 complex., The results indicate that ORF9b formed a stable complex with TOM70 (Figure 20). However, a return scan from 90 to 20 ^{o}C was unable to refold both proteins back to their secondary conformation (Not shown).

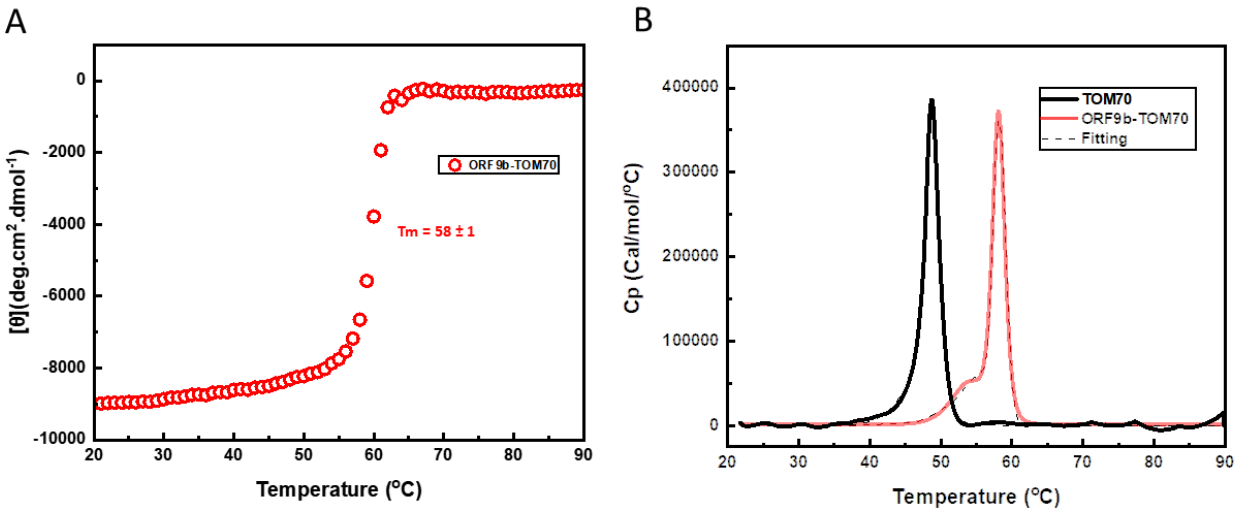


Figure 20: Thermal-induced unfolding. (A) Thermal-induced unfolding of ORF9b-TOM70 complex. The unfolding curve was followed at 222 nm in 2 mm cuvette at a scan rate of 1 °C/min from 20 to 90 °C. The Tm, temperature at the middle of the transition, calculated is 58 ± 1 °C. **(B)** Typical DSC curves for TOM70 and ORF9b-TOM70 after subtraction of the buffer baseline. The solid curves represent the data, dashed curves indicate fitting according to two independent non-two-state transitions. See Table 4 for the thermodynamic parameters

DSC data were normalized and fitted according to DSC non-Two-State model fit having two peaks, while the first peaks were insignificant, the Tm, temperature at the middle of the transition, values calculated for the significant peaks are 48.8 and 58.2 ºC for TOM70 and ORF9b-TOM70, respectively (Figure 20B). The calorimetric enthalpy ΔH_{cal} and van't Hoff enthalpy – ΔH_{vH} were calculated (Table 5).

Table 5: DSC thermodynamic parameters for TOM70 and ORF9b-TOM70 complex

Protein	Tm_1 (°C)	ΔH_1 (kcal/mole/°C)	Tm_2 (°C)	ΔH_2 (kcal/mole/°C)
TOM70	46±1	295 ± 20	49±1	930 ± 60
ORF9b-TOM70	54±1	340 ± 30	58±1	820 ± 50

4.5 Protein-protein interaction: Pull-down assay

In order to specify whether ORF9b, by binding to TOM70 affect its binding to Hsp90, we employed a pull-down protein-protein interaction approach. In this method, the interaction of Hsp90 with ORF9b-TOM70 or TOM70 was studied by using both proteins consisting of a poly histidine-tag which can specifically bind to the Nickel affinity column representing a bait, and thereby incubating with Hsp90 free of histidine-tag (previously cleaved by TEV-protease). Our results showed that Hsp90 failed to bind with the ORF9b-TOM70 complex, as it was observed from SDS-PAGE analysis of each fraction that Hsp90 was eluted with wash buffer after incubation (Figure 21A), but Hsp90 firmly binds with TOM70, and both was only eluted as a complex after washing with buffer containing imidazole (Figure 21B).

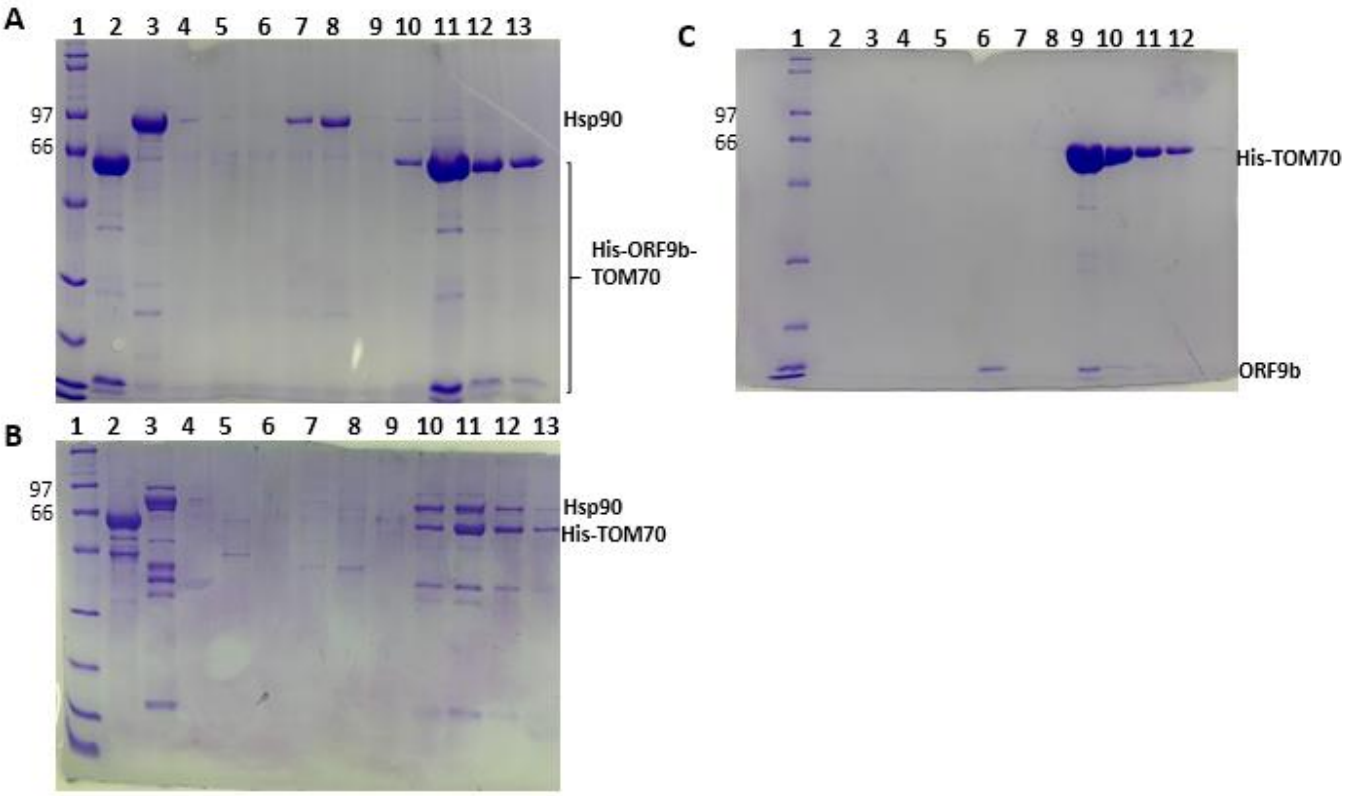


Figure 21: Pull down assay. (A) SDS-PAGE analysis of pull-down experiment of His-ORF9b-TOM70 complex with Hsp90. 1. Molecular mass marker; 2. His-ORF9b-TOM70 (bait); 3. Hsp90; 4. Flow through His-ORF9b-TOM70 (bait); 5-6. Wash with buffer A; 7. Flow through Hsp90; 8-9. Wash with buffer A; 10-13. Elution with buffer B (500 mM imidazole). Flow through of His-ORF9b-TOM70 shows that the protein was bound to the nickel column, which remained bound upon two steps of wash. Flow through of Hsp90 shows that Hsp90 failed to bind to the complex, and was eluted by buffer A wash. Elution by buffer B containing imidazole remove bound protein from the column **(B)** SDS-PAGE analysis of pull-down experiment of His-TOM70 with Hsp90. 1. Molecular mass marker; 2. His-TOM70 (bait); 3. Hsp90; 4. Flow through His-TOM70 (bait); 5-6. Wash with buffer A; 7. Flow through Hsp90; 8-9. Wash with buffer A; 10-13. Elution with buffer B (500 mM imidazole). Flow through of His-TOM70 shows that the protein is bound to the nickel column, which remain bound upon two steps of wash. Flow through of Hsp90 shows that Hsp90 was bound to TOM70, and was not dissociated upon two steps buffer A wash. Elution by buffer B containing imidazole confirms the presence of bound His-TOM70 with Hsp90. **(C)** SDS-PAGE analysis of pull-down experiment of His-TOM70 with ORF9b. 1. Molecular mass marker; 2. Flow through His-TOM70 (bait); 3-4. Wash with buffer A; 5. Flow through ORF9b; 6-8. Wash with buffer A; 9-12. Elution with buffer B (500 mM imidazole). Flow through of His-TOM70 shows that the protein is bound to the nickel column, which remain bound upon two steps of wash. Flow through of ORF9b shows that ORF9b was bound to TOM70, and partially dissociated upon two steps buffer A wash.

4.6 Protein-protein interaction: Analytical gel filtration

Analytical gel filtration performed to study the protein-protein interaction shows between Hsp90 and ORF9b-TOM70 complex showed independent peaks corresponding to the molecular mass of Hsp90 and the ORF9b-TOM70 complex not undergoing any interaction when analyzed by SDS-PAGE (Figure 22A). However, there was an interaction between Hsp90 and TOM70, which peak analyzed correspond with the molecular mass of the complex (Figure 22B). Furthermore, Hsp90 was incubated with 50 mM ATP-γs and ADP before SEC-MALS, similar independent peaks was observed (Figure 23). In further steps, we used the C-terminal Hsp90 (C-Hsp90) which is previously characterized to bind TOM70 to confirm the hypothesis that Hsp90 failed to bind ORF9b-TOM70 complex. SEC-MALS chromatogram and MM determination analyzed by SDS-PAGE shows that C-Hsp90 failed to bind ORF9b-TOM70 as opposed a stable complex formed with TOM70 (Figure 24).

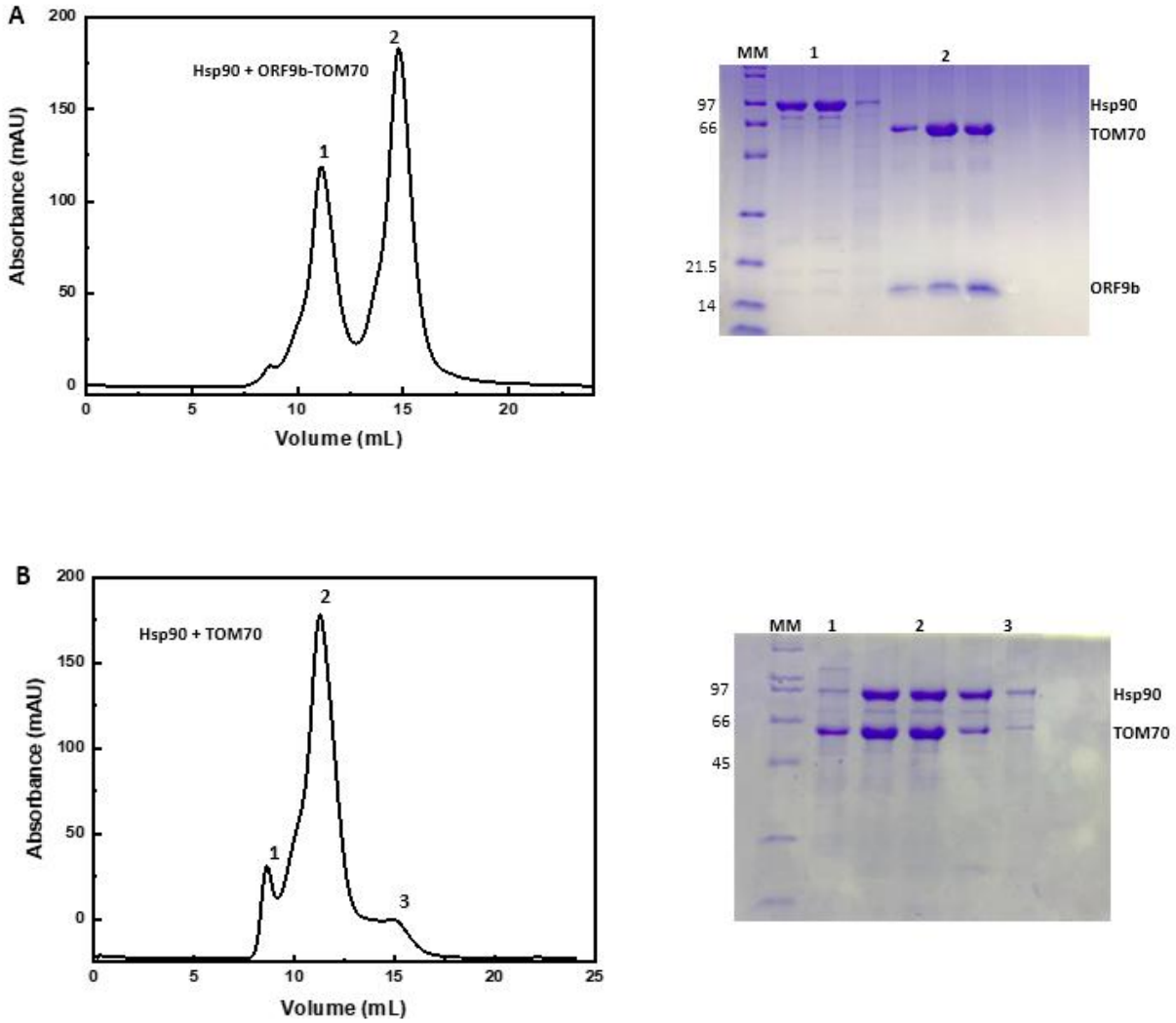


Figure 22: SEC-MALS Protein-Protein (A) SEC-MALS analysis of Hsp90 and ORF9b-TOM70 show two peaks, and SDS-PAGE analysis of peaks corresponding to the molecular mass of Hsp90 and ORF9b-TOM70 complex. **(B)** SEC-MALS analysis of Hsp90 and TOM70 show one major peaks, and SDS-PAGE analysis of peaks corresponding to the molecular mass of Hsp90 and TOM70 complex.

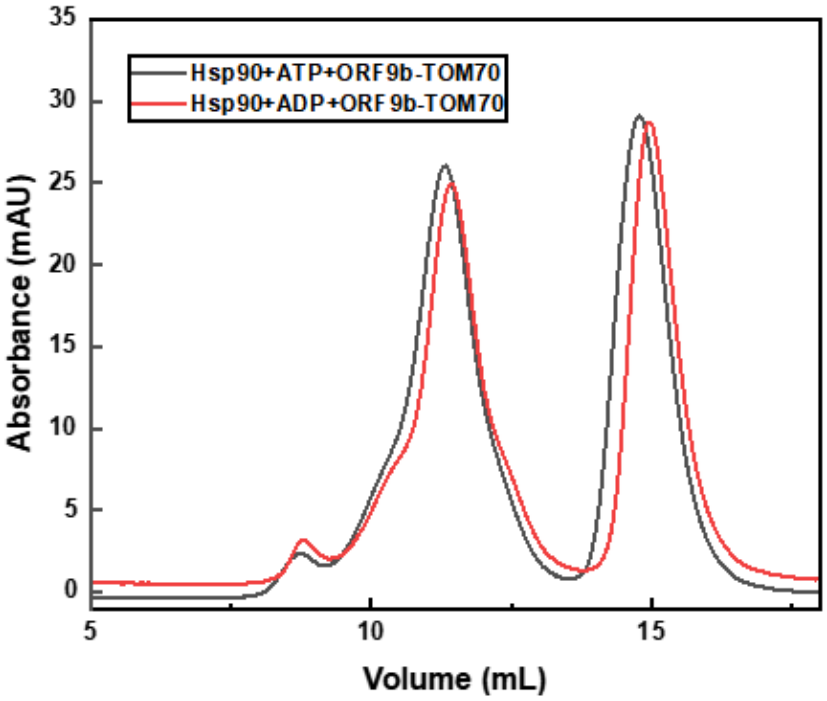


Figure 23: SEC-MALS protein-protein interaction. Hsp90 and ORF9b-TOM70 complex in the presence of non-hydrolysable ATP and ADP**.** Shows two independent peaks indicating the absence of interaction between the two proteins.

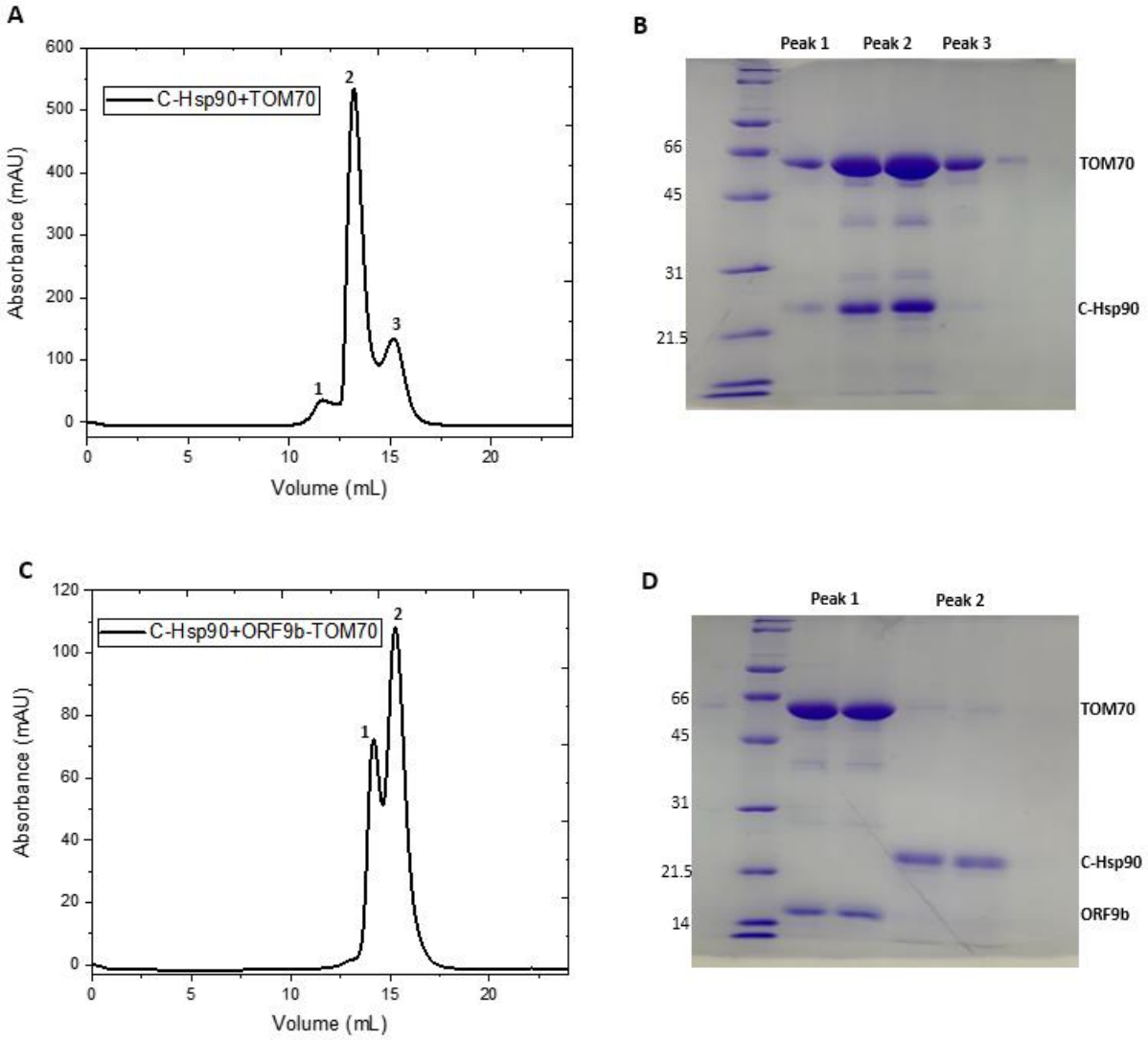


Figure 24: SEC-MALS Protein-Protein interaction. (A) SEC-MALS analysis of C-Hsp90 and TOM70 showing one major peak. **(B)** SDS-PAGE analysis of peaks corresponding to the molecular weight of C-Hsp90 in complex with TOM70 complex. **(B)** SEC-MALS of ATP and ADP bound Hsp90 with complex. **(C)** SEC-MALS analysis of C-Hsp90 and ORF9b-TOM70 showing two peaks. **(D)** SDS-PAGE analysis of peaks corresponding to the molecular weight of C-Hsp90 and ORF9b-TOM70 complex.

5.0 DISCUSSION

5.1 Structure and conformation of ORF9b

Accessory proteins of SARS-CoV-2 are encoded by individual ORFs. One of them is ORF9b, which, like other ORFs, is involved in modulating the response to infection (Wang et al., 2021). This response triggers a cascade that leads to the production of the antiviral protein interferon and includes the mitochondrial antiviral signaling (MAVS) protein. ORF9b likely disrupts the response by binding to the mitochondrial outer membrane protein TOM70 (Kreimendahl and Rassow, 2020, Wang et al., 2021). Thus, a detailed understanding of the mechanism of interaction has great potential to be relevant in the development of therapeutic strategies against COVID-19 (Gao et al., 2021). To contribute to such understanding, this investigation aimed to produce, characterize and investigate the interactions of ORF9b and the cytosolic fraction of human TOM70 by a combination of experimental tools (Batista et al., 2015).

Results from SEC-MALS confirm the homodimeric assembly of ORF9b which is consistent with crystal structure of ORF9b from SARS-CoV and SARS-CoV-2 (PDB id: 2CME, 6Z4U). The ORF9b conformation described by CD analysis shows that protein fold in a shape mainly consisting of random coils or disordered conformation.

The CD spectrum of ORF9b was characteristic of a disordered polypeptide. However, several structural characteristics of the ORF9b structure using other techniques, such as electron microscopy and cryo-EM, indicate that ORF9b is a 2-fold symmetric dimer constructed from two adjacent twisted β sheets forming strands contributed by both monomers with highly interlocked architecture reminiscent of a handshake (Gao et al., 2021). This is not a disagreement because the CD result revealed that a major part, but not all, of the protein was disordered, which is supported by the analysis of the protein sequence by PSIPRED that confirmed that a large part of the sequence may conform into random coils. More interestingly, ORF9b is believed to belong to the family of proteins referred to as fold switchers (Porter, 2021). Fold-switching proteins remodel their secondary structures and change their functions in response to environmental stimuli. We approached to see if by treating ORF9b with trifluoroethanol (TFE), a compound that alters the propensity of certain

polypeptides/proteins to form secondary structure, especially to helical conformation (Correa and Ramos, 2009), it will change the global structure of ORF9b. ORF9b retained its original conformation, and this shows that if at all ORF9b is a fold switcher, it only attains this with a small portion of the sequence participating in the switch. It is likely that the β-sheets that are best used to describe the homodimeric structure of the protein which are connected by a large random coil region (PDB; 6Z4U).

5.2 Thermodynamic analysis of ORF9b-TOM70 complex.

Co-expression of ORF9b and TOM70 produced a complex of two proteins corresponding with a monomer of both forming the complex. This is consistent with all two solved structures of the ORF9b-TOM70 complex by Cryo-EM (Gordon et al., 2020) and X-ray crystallography (Gao et al., 2021). Furthermore, CD spectra comparison of the ORF9b-TOM70 complex with TOM70 shows that the latter maintains its proper conformation even when bound with ORF9b.

Thermodynamic studies of the complex and comparison with TOM70 alone showed a strikingly more stable ORF9b-TOM70 complex. The broad transitions were due to the unfolding of the main species (as shown by SEC-MALS and CD experiments) and curve fitting of the full profile, using two endotherms, generated their T_ms . Although thermal-induced unfolding profiles were poorly reversible (data not shown), the T_{m2} values strongly supported the conclusion that TOM70 bound to ORF9-b was more stable than TOM70 alone, indicating that the interaction generates well-packed protein interfaces (Ramos and Ferreira, 2005). The increased stability of the ORF9b-TOM70 complex suggests that the binding of ORF9b further stabilized the structure of TOM70.

5.3 Protein interaction analysis

To specify whether ORF9b, by binding to TOM70, affects its binding to Hsp90, we employed a pull-down protein–protein interaction approach. Our results showed that Hsp90 binds TOM70 but fail to bind the ORF9b-TOM70 complex The oligomeric states of the full-length Hsp90 and ORF9b-TOM70 complexes showed that they eluted as independent peaks. To gather more information, this experiment was repeated with Hsp90 treated with 50 mM nonhydrolyzable ATP and ADP, and similar independent peaks were observed, showing that Hsp90 failed to bind with the ORF9b-TOM70 complex either in the presence or absence of nucleotides.

The MEEVD motif of Hsp90 specifically binds the cytosolic TPR (tetratricopeptide) clamp of TOM70 (Fan et al., 2006), and in a previous work, we studied the interaction of the C-terminal domain of Hsp90 (C-Hsp90) with TOM70 and found a second site of interaction (Zanphorlin et al., 2016). Thus, the interaction of this domain with the ORF9b-TOM70 complex was investigated to add information on the mechanism of inhibition. We showed an analytical gel filtration chromatogram of a mixture of ORF9b-TOM70 and C-Hsp90, consisting two distinct peaks analysed by SDS–PAGE, indicating no complex formation, a result similar to that of full-length Hsp90. However, as expected, the mixture of C-Hsp90 with TOM70 alone eluted as a single peak with a molecular mass of ~110 kDa, and peaks analysed by SDS–PAGE showed the two protein bands present in the peak, indicating the same stoichiometry arrangement of 2:1 (C-Hsp90:TOM70).

In a previous study, one of the main experiments utilized to identify the second site of interaction between Hsp90 and TOM70 used specific peptides from these proteins in a competition inhibition assay involving protein translocation by the Hsp90-TOM70 system (Zanphorlin et al., 2016). One of the peptides, Hsp90-GYSRMEEVD, contains the canonical MEEVD Hsp90 motif that binds to TPR motifs, corresponding to approximately 40% translocation inhibition (Zanphorlin et al., 2016). Other peptides, Hsp90-TLRQKAEADKNDKSVKDLVILLY and Tom70-LLADKVLKLLGKEKAKEKY, which correspond to a second site of interaction between Hsp90 and TOM70, inhibit translocation by approximately 40 and 35%, respectively (Zanphorlin et al., 2016). Additionally, the combination of the two Hsp90 peptides results in a translocation inhibitory level similar to that of a competition

inhibition assay by adding the Hsp90 protein (Zanphorlin et al., 2016). These results clearly indicate that, together, the two sites account for the strength of the whole interaction and that each contributes equally to the Hsp90-TOM70 interaction. Gao and colleagues (Gao et al., 2021) argued that occupation of ORF9b at the C-terminal of TOM70 may severely disrupt binding of the MEEVD motif at the N-terminus. We suggest that our findings add important additional information to their model to generate a more detailed one.

We suggest that our findings add important additional information to their model to generate a more detailed one. Therefore, we need to consider the crystal structure of the ORF9b-TOM70 complex (PDB: 7DHG) to highlight the interaction sites (Figure 6A). In this illustration, we mapped the following binding regions on TOM70: 1) in blue, the ORF9b interacting residues E549, H583, E477, Q379, Q381, D545, H515, Q594, and V556 and the residue E580 seen by cryo-EM structure (7KDT). 2), in yellow, the residues that interact with the MEEVD motif of Hsp90, adapted from the crystal structure of yeast TOM71 in complex with the Hsp82 C-terminal fragment (PDB: 3FP2) (Li et al., 2009). 3), in cyan, Trp residues. 4), in magenta, the identified second functional interaction site with Hsp90, which includes the sequence LLADKVLKLLGKEKAKEKY located on helix α7 ((Zanphorlin et al., 2016); Figure 25).

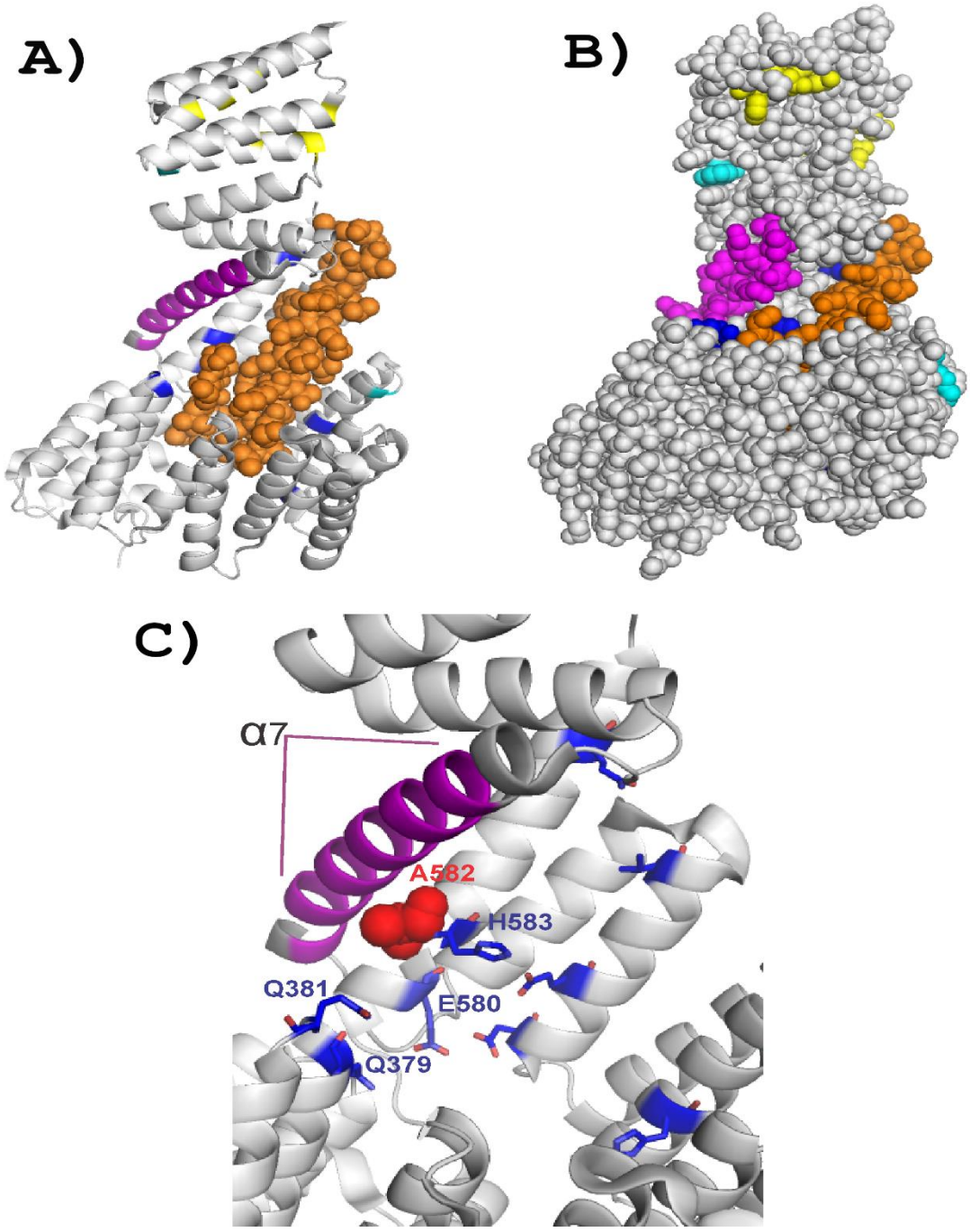


Figure 25. TOM70 structural arrangement and regions of interaction. (**A**) Crystal structure of the ORF9b-TOM70 complex (PDB: 7DHG). Helices of TOM70 and all residues of ORF9b. Residues marked in yellow show the TPR domain that interacts with the MEEVD motif of Hsp90, adapted from the crystal structure of yeast TOM71 in complex with the Hsp82 C-terminal fragment (PDB: 3FP2). Magenta shows the region (located on helix α7) from the newly identified second functional interaction site with Hsp90. Blue shows the ORF9b interacting residues on TOM70, which include residues E549, H583, E477, Q379, Q381, D545, H515, Q594, V556, and E580. Cyan, Trp residues. **B**) Surface model of the crystal structure of the ORF9b-TOM70 complex (see A) for details. **C**) Closer view of helix α7 (second site of interaction with Hsp90) and residues H583, E580, Q379, and Q381 at the binding region of ORF9b. Highlight of

residue Ala582 (red) that is in close contact with helix α7 (magenta), in which lies the second site of interaction with Hsp90), and is found mutated in a disease caused by poor translocation of key proteins to mitochondria

This is a direct indication that the second binding site in TOM70 (the first site being the one binding the MEEVD motif) is involved in the inhibition process, emerging from the observation that this region is in close proximity to the residues H583, E580, Q379 and Q381 (Fig. 25, in blue), i.e., the binding region of ORF9b. Therefore, binding of ORF9b to TOM70 is likely to perturb the conformational position of helix α7, which contains the second site of interaction with Hsp90. This hypothesis is strongly supported by the superimposition of the crystal structures of TOM70 bound to ORF9b with that of free TOM71, which indicates structural rearrangements on helices α7 and α8 upon binding (Gao et al., 2021). In fact, the superimposition shows that the region comprising helices α7 and α8 is stabilized by the binding of ORF9b (Gao et al., 2021). Therefore, a model materializes in which the binding of ORF9b to TOM70 perturbs not only the region that binds the Hsp90 MEEVD motif but also the region of helix α7, which lies at the second site of interaction with Hsp90

In conclusion, the combination of the results shown in this work with those previously published (Zanphorlin et al., 2016, Gao et al., 2021) give support to this model and suggest that ORF9b binding is very efficient in inhibiting Hsp90 and thus evading the immune response because it targets not one but two sites of interaction between TOM70 and Hsp90. The fact that the two sites are equally efficient in the binding process between Hsp90 and TOM70 (Zanphorlin et al., 2016) highlights the effectiveness of the inhibitory effect caused by ORF9b. This finding is relevant because the ORF9b-TOM70 interaction is an attractive target for the development of therapeutic strategies in COVID-19 infection since it has been previously shown that TOM70 is involved in a signaling cascade that ultimately leads to the induction of type I interferons (IFN-I). SARS-CoV-2 can evade host immunity through the downregulation of IFN production in the host cell (Blanco-Melo et al., 2020, Jiang et al., 2020, Han et al., 2021).

This cascade depends on the recruitment of Hsp90-bound proteins to the N-terminal domain of TOM70, thus recruiting TBK1/IRF3 to mitochondria, and disruption of this interaction impairs the activation of TBK1 and IRF3, which are critical steps to the mitochondrial antiviral system and induction of type I interferons (IFN-I) (Liu et al., 2010, Wei et al., 2015). Therefore, our results provide several lines of evidence that support the findings that inhibition of Hsp90 binding to TOM70 by ORF9b is the reason for the reduced IFN-I and overall host immune evasion of SAR-CoV-2 (Brandherm et al., 2021).

6.0 CONCLUSIONS

- SARS-CoV-2 ORF9b was a soluble protein which can be expressed in *E. coli* system
- SARS-CoV-2 ORF9b was purified as a homodimeric protein with molecular mass of 22.3 KDa. This protein has a high portion of random coils in its secondary structure
- Coexpression of ORF9b and TOM70 produced a stable ORF9b-TOM70 complex
- ORF9b-TOM70 complex is more heat stable than TOM70 alone
- ORF9b-TOM70 prevents the binding of Hsp90
- Inability of Hsp90 to bind ORF9b-TOM70 complex provided a major explanation for the SARS-CoV-2 evasion of host innate immunity and antiviral mechanism via inhibition of interferon activation cascade.

REFERENCES

Alanagreh, L., F. Alzoughool and M. Atoum (2020). "The Human Coronavirus Disease COVID-19: Its Origin, Characteristics, and Insights into Potential Drugs and Its Mechanisms." Pathogens **9**(5).

Batista, F. A., L. M. Gava, G. M. Pinheiro, et al. (2015). "From Conformation to Interaction: Techniques to Explore the Hsp70/Hsp90 Network." Curr Protein Pept Sci **16**(8): 735-753.

Bittar, E. R., F. R. Caldeira, A. M. Santos, et al. (2003). "Characterization of -trypsin at acid pH by differential scanning calorimetry." Braz J Med Biol Res **36**(12): 1621-1627.

Blanco-Melo, D., B. E. Nilsson-Payant, W. C. Liu, et al. (2020). "Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19." Cell **181**(5): 1036-1045.e1039.

Borges, J. C. and C. H. Ramos (2005). "Protein folding assisted by chaperones." Protein Pept Lett **12**(3): 257-261.

Brandherm, L., A. M. Kobaš, M. Klöhn, et al. (2021). "Phosphorylation of SARS-CoV-2 Orf9b Regulates Its Targeting to Two Binding Sites in TOM70 and Recruitment of Hsp90." Int J Mol Sci **22**(17).

Bruylants, G., J. Wouters and C. Michaux (2005). "Differential scanning calorimetry in life science: thermodynamics, stability, molecular recognition and application in drug design." Curr Med Chem **12**(17): 2011-2020.

Cascella, M., M. Rajnik, A. Aleem, et al. (2021). Features, Evaluation, and Treatment of Coronavirus (COVID-19). StatPearls. Treasure Island (FL), StatPearls Publishing

Cerveny, L., A. Straskova, V. Dankova, et al. (2013). "Tetratricopeptide repeat motifs in the world of bacterial pathogens: role in virulence mechanisms." Infect Immun **81**(3): 629-635.

Chang, L., Y. Yan and L. Wang (2020). "Coronavirus Disease 2019: Coronaviruses and Blood Safety." Transfus Med Rev **34**(2): 75-80.

Chen, Y., Q. Liu and D. Guo (2020). "Emerging coronaviruses: Genome structure, replication, and pathogenesis." J Med Virol **92**(4): 418-423.

Correa, D. H. and C. H. Ramos (2009). "The use of circular dichroism spectroscopy to study protein folding, form and function." Afr J Biochem Res **3**: 164-173.

da Silva, V. C. and C. H. Ramos (2012). "The network interaction of the human cytosolic 90 kDa heat shock protein Hsp90: A target for cancer therapeutics." J Proteomics **75**(10): 2790-2802.

Duong-Ly, K. C. and S. B. Gabelli (2014). "Gel filtration chromatography (size exclusion chromatography) of proteins." Methods Enzymol **541**: 105-114.

Edelhoch, H. (1967). "Spectroscopic determination of tryptophan and tyrosine in proteins." Biochemistry **6**(7): 1948-1954.

Fan, A. C., M. K. Bhangoo and J. C. Young (2006). "Hsp90 functions in the targeting and outer membrane translocation steps of Tom70-mediated mitochondrial import." J Biol Chem **281**(44): 33313-33324.

Gao, X., K. Zhu, B. Qin, et al. (2021). "Crystal structure of SARS-CoV-2 Orf9b in complex with human TOM70 suggests unusual virus-host interactions." Nat Commun **12**(1): 2843.

Gava, L. M., D. C. Gonçalves, J. C. Borges, et al. (2011). "Stoichiometry and thermodynamics of the interaction between the C-terminus of human 90kDa heat shock protein Hsp90 and the mitochondrial translocase of outer membrane Tom70." Arch Biochem Biophys **513**(2): 119-125.

Gordon, D. E., J. Hiatt, M. Bouhaddou, et al. (2020). "Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms." Science **370**(6521).

Gordon, D. E., G. M. Jang, M. Bouhaddou, et al. (2020). "A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug-Repurposing." bioRxiv.

Greenfield, N. J. (2006). "Using circular dichroism spectra to estimate protein secondary structure." Nat Protoc **1**(6): 2876-2890.

Han, L., M.-W. Zhuang, J. Deng, et al. (2021). "SARS-CoV-2 ORF9b antagonizes type I and III interferons by targeting multiple components of the RIG-I/MDA-5-MAVS, TLR3-TRIF, and cGAS-STING signaling pathways." Journal of medical virology **93**(9): 5376-5389.

Hartl, F. U. (1996). "Molecular chaperones in cellular protein folding." Nature **381**(6583): 571-579.

Hartl, F. U., A. Bracher and M. Hayer-Hartl (2011). "Molecular chaperones in protein folding and proteostasis." Nature **475**(7356): 324-332.

Hartl, F. U. and M. Hayer-Hartl (2009). "Converging concepts of protein folding in vitro and in vivo." Nat Struct Mol Biol **16**(6): 574-581.

Haslbeck, M. and E. Vierling (2015). "A first line of stress defense: small heat shock proteins and their function in protein homeostasis." J Mol Biol **427**(7): 1537-1548.

Helmy, Y. A., M. Fawzy, A. Elaswad, et al. (2020). "The COVID-19 Pandemic: A Comprehensive Review of Taxonomy, Genetics, Epidemiology, Diagnosis, Treatment, and Control." J Clin Med **9**(4).

Hoppe, T. and E. Cohen (2020). "Organismal Protein Homeostasis Mechanisms." Genetics **215**(4): 889-901.

Ji, C. H. and Y. T. Kwon (2017). "Crosstalk and Interplay between the Ubiquitin-Proteasome System and Autophagy." Mol Cells **40**(7): 441-449.

Jiang, H. W., H. N. Zhang, Q. F. Meng, et al. (2020). "SARS-CoV-2 Orf9b suppresses type I interferon responses by targeting TOM70." Cell Mol Immunol **17**(9): 998-1000.

Johnson, C. M. (2013). "Differential scanning calorimetry as a tool for protein folding and stability." Arch Biochem Biophys **531**(1-2): 100-109.

Kreimendahl, S. and J. Rassow (2020). "The Mitochondrial Outer Membrane Protein Tom70-Mediator in Protein Traffic, Membrane Contact Sites and Innate Immunity." Int J Mol Sci **21**(19).

Laskowska, E., D. Kuczyńska-Wiśnik and B. Lipińska (2019). "Proteomic analysis of protein homeostasis and aggregation." J Proteomics **198**: 98-112.

Li, J., X. Qian, J. Hu, et al. (2009). "Molecular Chaperone Hsp70/Hsp90 Prepares the Mitochondrial Outer Membrane Translocon Receptor Tom71 for Preprotein Loading." Journal of Biological Chemistry **284**(35): 23852-23859.

Liu, X. Y., B. Wei, H. X. Shi, et al. (2010). "Tom70 mediates activation of interferon regulatory factor 3 on mitochondria." Cell Res **20**(9): 994-1011.

Obermann, W. M., H. Sondermann, A. A. Russo, et al. (1998). "In vivo function of Hsp90 is dependent on ATP binding and ATP hydrolysis." J Cell Biol **143**(4): 901-910.

Pace, C. N., F. Vajdos, L. Fee, et al. (1995). "How to measure and predict the molar absorption coefficient of a protein." Protein Sci **4**(11): 2411-2423.

Park, A. and A. Iwasaki (2020). "Type I and Type III Interferons - Induction, Signaling, Evasion, and Application to Combat COVID-19." Cell Host Microbe **27**(6): 870-878.

Porath, J. (1992). "Immobilized metal ion affinity chromatography." Protein Expr Purif **3**(4): 263-281.

Porter, L. L. (2021). "Predictable fold switching by the SARS-CoV-2 protein ORF9b." Protein Sci **30**(8): 1723-1729.

Radli, M. and S. G. D. Rüdiger (2018). "Dancing with the Diva: Hsp90-Client Interactions." J Mol Biol **430**(18 Pt B): 3029-3040.

Ramos, C. (2011). "Molecular chaperones and protein quality control." Protein Pept Lett **18**(2): 100.

Ramos, C. H. and S. T. Ferreira (2005). "Protein folding, misfolding and aggregation: evolving concepts and conformational diseases." Protein Pept Lett **12**(3): 213-222.

Ramos, C. H. I. and K. S. Ayinde (2020). "Are Hsp90 inhibitors good candidates against Covid-19?" Curr Protein Pept Sci.

Raran-Kurussi, S., S. Cherry, D. Zhang, et al. (2017). "Removal of Affinity Tags with TEV Protease." Methods Mol Biol **1586**: 221-230.

Redondo, N., S. Zaldívar-López, J. J. Garrido, et al. (2021). "SARS-CoV-2 Accessory Proteins in Viral Pathogenesis: Knowns and Unknowns." Front Immunol **12**: 708264.

Scheufler, C., A. Brinker, G. Bourenkov, et al. (2000). "Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine." Cell **101**(2): 199-210.

Schneider, C. A., W. S. Rasband and K. W. Eliceiri (2012). "NIH Image to ImageJ: 25 years of image analysis." Nat Methods **9**(7): 671-675.

Shiber, A. and T. Ravid (2014). "Chaperoning proteins for destruction: diverse roles of Hsp70 chaperones and their co-chaperones in targeting misfolded proteins to the proteasome." Biomolecules **4**(3): 704-724.

Some, D., H. Amartely, A. Tsadok, et al. (2019). "Characterization of Proteins by Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS)." J Vis Exp(148).

Spriestersbach, A., J. Kubicek, F. Schäfer, et al. (2015). "Purification of His-Tagged Proteins." Methods Enzymol **559**: 1-15.

Tiroli-Cepeda, A. O. and C. H. Ramos (2011). "An overview of the role of molecular chaperones in protein homeostasis." Protein Pept Lett **18**(2): 101-109.

Totura, A. L. and R. S. Baric (2012). "SARS coronavirus pathogenesis: host innate immune responses and viral antagonism of interferon." Curr Opin Virol **2**(3): 264-275.

Wang, S., T. Dai, Z. Qin, et al. (2021). "Targeting liquid-liquid phase separation of SARS-CoV-2 nucleocapsid protein promotes innate antiviral immunity by elevating MAVS activity." Nat Cell Biol **23**(7): 718-732.

Wankhede, N. L., M. B. Kale, A. B. Upaganlawar, et al. (2022). "Involvement of molecular chaperone in protein-misfolding brain diseases." Biomed Pharmacother **147**: 112647.

Wei, B., Y. Cui, Y. Huang, et al. (2015). "Tom70 mediates Sendai virus-induced apoptosis on mitochondria." J Virol **89**(7): 3804-3818.

Wyatt, P. J. (1993). "Light scattering and the absolute characterization of macromolecules." Analytica Chimica Acta **272**(1): 1-40.

Yin, W., C. Mao, X. Luan, et al. (2020). "Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir." Science **368**(6498): 1499-1504.

Young, J. C., N. J. Hoogenraad and F. U. Hartl (2003). "Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70." Cell **112**(1): 41-50.

Young, J. C., I. Moarefi and F. U. Hartl (2001). "Hsp90: a specialized but essential protein-folding tool." J Cell Biol **154**(2): 267-273.

Zanphorlin, L. M., T. B. Lima, M. J. Wong, et al. (2016). "Heat Shock Protein 90 kDa (Hsp90) Has a Second Functional Interaction Site with the Mitochondrial Import Receptor Tom70." J Biol Chem **291**(36): 18620-18631.

Zeytuni, N. and R. Zarivach (2012). "Structural and functional discussion of the tetra-trico-peptide repeat, a protein interaction module." Structure **20**(3): 397-405.

APPENDIX

COMISSÃO INTERNA DE BIOSSEGURANÇA
UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE QUÍMICA
Cx. Postal 6154 - Campinas - 13083-970 SP - Brasil
http://www.iqm.unicamp.br

Solicitação/Protocolo: CHIR2021-5

Nome do(a) aluno(a): Kehinde Sulaimon Ayinde

Nível: Mestrado

Período: 11/01/2021 a 11/04/2022

Nome do(a) orientador(a): Prof. Dr. Carlos Henrique Inácio Ramos

Título: Produção de proteínas chaperonas recombinantes humanas e estudo de sua interação com outras proteínas.

Resumo: Este projeto visa o estudo da interação entre chaperonas recombinantes humanas, Hsp90 e TOM70, e a proteína recombinante orf9b de SARS-COV-2. Todas as proteínas (genes) são de nível de biossegurança 1, não patogênicas e serão clonadas em vetor pET23 e produzidas em *Escherichia coli* BL21(DE3) em volumes de no máximo 500 mL.
Os equipamentos a serem utilizados para a produção das proteínas são incubadora, shaker, centrífuga e sonicador.
Todo meio líquido ou sólido passa primeiro por uma descontaminação com hipoclorito de sódio por pelo menos 30 minutos. Em seguida, esse e qualquer outro material passa por um processo de descontaminação por autoclave. Uma vez que a descontaminação é feita o material é limpo ou descartado seguindo regras adequadas a um laboratório de química.

O projeto de pesquisa acima descrito, a ser conduzido pelo(a) bolsista(a) de mestrado Kehinde Sulaimon Ayinde, recebeu autorização desta Comissão Interna de Biossegurança.

Presidente da CIBio-IQ: Profa. Dra. Denize Cristina Favaro
20/10/2021

cibioiq@unicamp.br
Telefone: 19 3521 3002

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE QUÍMICA

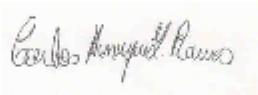
Campinas 01 de outubro de 2019

CARTA-COMPROMISSO DE ORIENTAÇÃO

Declaro que, no caso de aprovação do candidato Ayinde Kehinde Sulaimon, passaporte número A10590872, atualmente residente em Ipaja Lagos, Nigéria, no processo seletivo para ingresso no Curso de Pós-graduação em Biologia Funcional e Molecular, nível Doutorado, assumirei a orientação do aluno, garantindo-lhe todas as condições científicas, técnicas e de infra-estrutura para o desenvolvimento de seu projeto de pesquisa, dentro do prazo estabelecido pelo Regulamento dos Cursos de Pós-Graduação do Instituto de Biologia da UNICAMP (Deliberação CEPE-A-2/02, de 05/02/2002).
Estou ciente que o descumprimento dos termos desta declaração poderá comprometer a possibilidade de futuras orientações junto ao Curso de Pós-Graduação em Biologia Funcional e Molecular do IB/UNICAMP.

Colocando-me à disposição para fornecer outras informações, subscrevo-me.

Cordialmente,



Carlos Henrique Inácio Ramos

Professor Titular – Departamento de Química Orgânica
cramos@iqm.unicamp.br
Fone: 55-19-3521-3096