DESIGN AND CHARACTERIZATION OF THERAPEUTIC PEPTIDES TARGETING THE RECEPTOR BINDING DOMAIN OF THE SPIKE GLYCOPROTEIN FROM SARS-COV-2

HO CHIAN

FACULTY OF SCIENCE AND NATURAL RESOURCES
UNIVERSITI MALAYSIA SABAH
2023

DESIGN AND CHARACTERIZATION OF THERAPEUTIC PEPTIDES TARGETING THE RECEPTOR BINDING DOMAIN OF THE SPIKE GLYCOPROTEIN FROM SARS-COV-2

HO CHIAN

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

FACULTY OF SCIENCE AND NATURAL RESOURCES
UNIVERSITI MALAYSIA SABAH
2023

UNIVERSITI MALAYSIA SABAH

BORANG PENGESAHAN STATUS TESIS

JUDUL : **DESIGN AND CHARACTERIZATION OF THERAPEUTIC PEPTIDES TARGETING THE RECEPTOR BINDING DOMAIN OF THE SPIKE GLYCOPROTEIN FROM SARS-COV-2**

IJAZAH : **SARJANA SAINS**

BIDANG : **BIOTEKNOLOGI**

Saya **HO CHIAN**, Sesi **2020-2023**, mengaku membenarkan tesis Sarjana ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:-

1. Tesis ini adalah hak milik Universiti Malaysia Sabah
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (/):

☐	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA 1972)
☐	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
/	TIDAK TERHAD	

Disahkan Oleh,

Ho Chian

ANITA BINTI ARSAD
PUSTAKAWAN KANAN
UNIVERSITI MALAYSIA SABAH

HO CHIAN
MS2011004T

(Tandatangan Pustakawan)

Tarikh : 16 August 2023

(Prof. Dr. Lee Ping Chin)
Penyelia

DECLARATION

I hereby declare that the material in this thesis is my own except for quotations, equations, summaries, and references, which have been duly acknowledged.

16 May 2023

Ho Chian

Ho Chian
MS2011004T

CERTIFICATION

NAME	:	**HO CHIAN**
MATRIC NO.	:	**MS2011004T**
TITTLE	:	**DESIGN AND CHARACTERIZATION OF THERAPEUTIC PEPTIDES TARGETING THE RECEPTOR BINDING DOMAIN OF THE SPIKE GLYCOPROTEIN FROM SARS-COV-2**
DEGREE	:	**MASTER OF SCIENCE**
FIELD	:	**BIOTECHNOLOGY**
VIVA DATE	:	**16 MAY 2023**

CERTIFIED BY:

Signature

SUPERVISOR
Prof. Dr. Lee Ping Chin

ACKNOWLEDGEMENT

I would like to take this opportunity to express my deepest gratitude to my supervisor Prof. Dr. Lee Ping Chin for her invaluable guidance, endless patience, and unwavering encouragement throughout my academic journey.

I would like to thank Dr. Eric Chong Tzyy Jiann, Mr. Ng Jun Wei and the laboratory staff at the Faculty of Science and Natural Resources for their precious time spent on providing advice, assistance and guidance while carrying out the laboratory experiments for this research. I would also like to thank Dr. Yew Chee Wei and the laboratory staff at Biotechnology Research Institute in providing computer access during this research.

Finally, I would like to thank my family and friends for the love, understanding, support, and encouragement shown me from the start of my academic journey until its completion.

Ho Chian
16 May 2023

ABSTRACT

Coronavirus Disease 2019 (COVID-19), a global pandemic which first emerged in Wuhan City, Hubei Province, China, in December 2019, was caused by a novel *Betacoronavirus*, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). To date, there are only two Food and Drug Administration (FDA)-approved therapeutic drugs to counter COVID-19. Hence, development for COVID-19 treatment needs to be broadened. This study aims to explore an alternative treatment for COVID-19 using therapeutic peptides. This study hypothesised that therapeutic peptides that can bind with the receptor binding domain (RBD) of spike glycoprotein (S protein) of SARS-CoV-2 might inhibit the entry of the virus into host cells and prevent infection. Bioinformatics tools include PyMOL, UCSF chimera, PEP-FOLD3, HADDOCK (High Ambiguity Driven protein-protein DOCKing), and molecular dynamics simulation (MDS) are first utilised for peptide design. PyMOL and UCSF chimera were used to determine the amino acid residues involved in the binding interface. A total of 291 peptides were designed, and PEP-FOLD3 was used to generate the peptide models. Docking analysis on the peptides and RBD of S protein was performed via HADDOCK. The peptide-RBD complexes of peptides 65, 66, and 189 showed the highest HADDOCK score and hence were chosen for 200 ns of MDS analysis. MDS results showed that all three peptide-RBD complexes were compact and stable, and their interaction energy being -480.38 ± 118.61 kJ/mol, -410.92 ± 126.31 kJ/mol, and -338.49 ± 97.88 kJ/mol, respectively. The peptides were then subjected to a binding affinity test and competitive assay via ELISA (enzyme-linked immunosorbent assay). For the binding affinity test, ELISA was performed using the hemagglutinin (HA)-tagged peptides in a series of concentrations bound to RBD of S protein on a 96-well plate. A competitivee assay was performed using His-tagged ACE2 protein incubated with the peptides. Peptide 65 showed the best result with bioinformatics analysis and binding affinity test with a K_D of 43.26 μM. However, the results of the competitive assay showed the that binding of peptide 65 to S protein is weak compared to the ACE2 protein. Bioinformatics tools can be utilised for preliminary analysis of the peptides designed and enables choosing of peptides with the best results for further laboratory testing. Peptide 65 can be further modified to improve its inhibiting ability and further develop into therapeutic drugs to counter COVID-19.

ABSTRAK

REKAAN DAN PENCIRIAN PEPTIDA TERAPEUTIK YANG MENSASARKAN DOMAIN PENGIKAT RESEPTOR BAGI GLIKOPROTEIN SPIKE DARIPADA SARS-CoV-2

Penyakit koronavirus 2019 (COVID-19), pandemik global yang pertama kali muncul di bandar Wuhan, Provinsi Hubei, China pada Disember 2019, disebabkan oleh sejenis Betacoronavirus baru, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Sehingga kini, hanya terdapat dua ubat terapeutik yang diluluskan oleh Food and Drug Administration (FDA) untuk melawan COVID-19. Oleh itu, pembangunan untuk rawatan COVID-19 perlu diperluaskan. Kajian ini bertujuan untuk meneroka rawatan alternatif untuk COVID-19 dengan penggunaan peptida terapeutik. Adalah dihipotesiskan bahawa peptida terapeutik yang boleh mengikat dengan domain pengikat reseptor (RBD) spike glikoprotein (protein S) SARS-CoV-2 mungkin dapat menghalang kemasukan virus ke dalam sel perumah dan mencegah jangkitan. Alat bioinformatik seperti PyMOL, UCSF chimera, PEP-FOLD3, HADDOCK (High Ambiguity Driven protein-protein DOCKing), dan simulasi dinamik molekul digunakan terlebih dahulu untuk merekakan peptide-peptida. PyMOL dan UCSF chimera digunakan untuk menentukan residu asid amino yang terlibat antara muka pengikat. Sebanyak 291 peptida telah direka dan PEP-FOLD3 digunakan untuk menjana model peptida. Analisis dok pada peptida dan RBD protein S dilakukan melalui HADDOCK. Kompleks peptida-RBD bagi peptida 65, 66, dan 189 menunjukkan skor HADDOCK tertinggi dan telah dipilih untuk 200 ns analisis MDS. Keputusan MDS menunjukkan bahawa ketiga-tiga kompleks peptida-RBD adalah padat dan stabil, dan tenaga interaksi masing-masing ialah -480.38 ± 118.61 kJ/mol, -410.92 ± 126.31 kJ/mol, dan -338.49 ± 97.88 kJ/mol. Peptida-peptida ini kemudiannya ditujukan kepada ujian afiniti mengikat dan ujian kompetitif melalui ELISA (ujian imunosorben berpaut enzim). Untuk ujian afiniti mengikat, peptida tag dengan hemagglutinin (HA) dalam satu siri kepekatan diikat pada RBD protein S pada plat 96-telaga. Untuk ujian kompetitif, kehadiran protein ACE2 dengan tag His dieram dengan peptide dalam satu siri kepekatan. Peptida 65 menunjukkan keputusan terbaik dengan analisis bioinformatik serta ujian afiniti mengikat dengan K_D 43.26 µM. Walau bagaimanapun, keputusan ujian kompetitif menunjukkan bahawa pengikatan peptida 65 dengan protein S adalah lemah berbanding dengan protein

ACE2. Alat bioinformatik boleh digunakan untuk analisis awal peptida yang direka dan membolehkan pemilihan peptida dengan hasil terbaik untuk ujian makmal selanjutnya. Peptide 65 boleh diubah suai lagi untuk meningkatkan keupayaan perencatannya supaya ia dapat dihasilkan dengan lebih lanjut lagi untuk menjadi ubat terapeutik untuk melawan COVID-19.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3CL^{pro}	-	3C-like Protease
AA	-	Amino Acid
ACE2	-	Angiotensin-converting Enzyme 2
ACTIV	-	Accelerating COVID-19 Therapeutic Interventions and Vaccines
ADMET	-	Absorption, Distribution, Metabolism, Excretion, and Toxicity
AIR	-	Ambiguous Interaction Restraint
AKI	-	Acute Kidney Injury
ALP	-	Alkaline Phosphatase
ARDS	-	Acute Respiratory Distress Syndrome
ATP	-	Adenosine Triphosphate
AVPdb	-	Antiviral Peptide Database
BBB	-	Blood Brain Barrier
BC	-	Binet-Cauchy
BLI	-	Bio-layer Interferometry
CAPRI	-	Critical Assessment of Predicted Interactions
CDC	-	Centers for Disease Control and Prevention
CH	-	Central Helix
CNS	-	Central Nervous System
COVID-19	-	Coronavirus Disease 2019
CoVs	-	Coronaviruses
CT	-	Cytoplasmic Terminal
E	-	Envelope
ECMO	-	Extracorporeal Membrane Oxygenation
ELISA	-	Enzyme-Linked Immunosorbent Assay
EMA	-	European Medicines Agency
EUA	-	Emergency Use Authorization
Fc	-	Fragment Crystallizable
FCC	-	Fraction of Common Contact
FDA	-	Food and Drug Administration
FNIH	-	Foundation for The NIH
FP	-	Fusion Peptide
FS	-	Femtosecond

Fu	-	Fraction Unbound
GDT	-	Global Distance Test
GIT	-	Gastrointestinal Tract
GLP-1	-	Glucagon-like Peptide-1
GM-CSF	-	Granulocyte-macrophage Colony-stimulating Factor
GnRH	-	Gonadotropin-releasing Hormone
GRAVY	-	Grand Average of Hydropathicity Index
GROMACS	-	Groningen MAchine for Chemical Simulation
HA	-	Hemagglutinin
HADDOCK	-	High Ambiguity Driven Protein-protein DOCKing
hERG	-	Human Ether-à-go-go-related Gene
HIA	-	Human Intestinal Absorption
HR	-	Heptad Repeat
HRP	-	Horseradish Peroxidase
IC_{50}	-	Half Maximal Inhibitory Concentration
IFN	-	Interferon
IL	-	Interleukin
IL-1RA	-	Interleukin-1 Receptor Antagonist
i-RMSD	-	Interface RMSD
JAK	-	Janus Kinase
K_D	-	Dissociation Constant
LC_{50}	-	Lethal Concentration Causing 50% Death
LD_{50}	-	Lethal Dose Causing 50% Death
LJ	-	Lennard-Jones
LOAEL	-	Lowest Observed Adverse Effect
LRA	-	Ligand-receptor Interaction Assay
LSPR	-	Localised Surface Plasmon Resonance
M	-	Membrane
M^{pro}	-	Main Protease
MCP-1	-	Macrophage Inflammatory Protein 1
MDS	-	Molecular Dynamics Simulation
MERS	-	Middle East Respiratory Syndrome
MM-GBSA	-	Molecular Mechanics Generalised Born Surface Area
MRTD	-	Maximum Recommended Tolerated Dose

N	-	Nucleocapsid
NIH	-	National Institutes of Health
NS	-	Nanosecond
NSP	-	Non-structural Proteins
NTD	-	N-terminal Domain
OCT2	-	Organic Cation Transporter 2
PaO2	-	Arterial Partial Pressure of Oxygen
PCA	-	Principal Component Analysis
PEG	-	Polyethylene Glycol
pI	-	Isoelectric Point
PL^{pro}	-	Papain-like Protease
PNPP	-	P-nitrophenyl Phosphate Disodium Salt
PRODIGY	-	Protein Binding Energy Prediction
PTH1	-	Parathyroid Hormone 1
RBD	-	Receptor Binding Domain
RdRp	-	RNA-dependent RNA Polymerase
Rg	-	Radius of Gyration
RMSD	-	Root Mean Square Deviation
RMSF	-	Root Mean Square Fluctuation
RTC	-	Replication and Transcription Complex
S	-	Spike Glycoprotein
SA	-	Structural Alphabet
SARS	-	Severe Acute Respiratory Syndrome
SARS-CoV-2	-	Severe Acute Respiratory Syndrome Coronavirus 2
SD1/SD2	-	Subdomain 1/Subdomain 2
sOPEP	-	Optimised Potential for Efficient Structure Prediction
SpO2	-	Oxygen Saturation
SS	-	Signal Sequence
ssRNA	-	Single-stranded Positive-sense RNA
SuPAR	-	Soluble Urokinase Plasminogen Activator Receptor
SVM	-	Support Vector Machine
TM	-	Transmembrane Domain
TMB	-	3,3',5,5'-tetramethylbenzidine
TMPRSS2	-	Transmembrane Protease Serine 2

TNF-α	-	Tumour Necrosis Factor-α
UK	-	United Kingdom
VBM	-	Variant Being Monitored
vdW	-	van der Waals
VDss	-	Volume of Distribution
VOC	-	Variant of Concern
VOHC	-	Variant of High Consequence
VOI	-	Variant of Interest
WHO	-	World Health Organization

LIST OF APPENDIX

CHAPTER 1

INTRODUCTION

1.1 Background of The Research

The outbreak of coronavirus disease 2019 (COVID-19), first emerged in Wuhan city, Hubei Province, China in December 2019, was caused by a novel *Betacoronavirus*, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) (Wu *et al.*, 2020). On March 11, 2020, the disease was declared a global pandemic by the World Health Organization (WHO) (Cucinotta and Vanelli, 2020). As of 18 June 2023, over 768 million confirmed COVID-19 cases and 6.9 million deaths had been reported worldwide. This global pandemic has caused devastating social and economic disruption. To date, the number of positive cases as well as the number of deaths, has decreased globally, compared to previous years. However, WHO also noted that due to a reduction in testing and reporting, the number of recorded cases did not represent an accurate infection rate (COVID-19 Weekly Epidemiological Update on COVID-19-22 June 2023, 2023).

SARS-CoV-2 S protein and the host cell angiotensin-converting enzyme 2 (ACE2) receptor are the key components for infection of COVID-19. The presence of the ACE2 receptor in the vital organs, especially the lungs, enables viral infection in the host (Xu *et al.*, 2020; Jain, 2020). The treatment for COVID-19 is associated with the stages of the disease (Gandhi, 2021). For this purpose, the United States Food and Drug Administration (FDA) has approved two drugs for the treatment of COVID-19, the antiviral drug Veklury (remdesivir) in the early stages, and the immune modulator Olumiant (baricitinib), in the later stages (Coronavirus (COVID-19) | Drugs, 2022).