INTERROGATION OF STRUCTURAL VARIATIONS OF **PRO-**ONCOGENIC AGR2 PROTEIN IN SUPPORTING ESOPHAGEAL CANCER PROGRESSION

NURUL HANANI BINTI BADEROL HISHAM

FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR

2022

INTERROGATION OF STRUCTURAL VARIATIONS OF **PRO-**ONCOGENIC AGR2 PROTEIN IN SUPPORTING ESOPHAGEAL CANCER PROGRESSION

NURUL HANANI BINTI BADEROL HISHAM

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE (BIOTECHNOLOGY)

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR

2022

UNIVERSITI MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **NURUL HANANI BINTI BADEROL HISHAM**.

Matric No: **17069806/2**.

Name of Degree: **MASTER OF SCIENCE (BIOTECHNOLOGY)**.

Title of Dissertation ("this Work"):

INTERROGATION OF STRUCTURAL VARIATIONS OF PRO-ONCOGENIC AGR2 PROTEIN IN SUPPORTING ESOPHAGEAL CANCER PROGRESSION

Field of Study:

BIOTECHNOLOGY.

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every right in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that anyreproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyrightwhether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date 7th February 2022

Subscribed and solemnly declared before,

Witness's Signature Date 7th February 2022

Name:
Designation:

INTERROGATION OF STRUCTURAL VARIATIONS OF PRO-ONCOGENIC AGR2 PROTEIN IN SUPPORTING ESOPHAGEAL CANCER PROGRESSION

ABSTRACT

A non-hormone-dependent cancer, esophageal adenocarcinoma was classified as the sixth deadliest cancer with an elevated mortality rate resulting from failure detection at an onset of the disease hence causing the tumor to be at either an advanced or metastatic stage upon diagnosis. Such cancer progression initially reported to be due to many factors including the emergence of pro-oncogenic protein such as the AGR2 protein, known to play roles mainly in the folding of proteins as well as maintaining proteostasis in the ER. The poor understanding on the structural functions of the AGR2 protein in supporting esophageal cancer progression has led to the design of this research which involved constructing four AGR2 protein variants comprised of the AGR2 – WT variant as to compare the basal expression to the other mutated variants, the AGR – KDEL variant to enhance localization of the protein in the ER, the mAGR2 – ΔKTEL variant to enhance secretion of the protein into the extracellular environment and the ΔNterm AGR2 – KDEL variant to enhance the dimerization of the protein. All constructs of AGR2 protein which were AGR – WT, AGR2 – KDEL, mAGR2 – ΔKTEL and ΔNterm AGR2 – KDEL were cloned into the pSF-CMV vector before basal expression of the AGR2 protein variants were determined by transiently transfecting them into the FLO-1 esophageal adenocarcinoma cell line and conduction of immunoblotting and immunofluorescence. The intracellular expressions of the FLO-1 showed the highest expression observed by the AGR2 – WT variant followed by AGR – KDEL. Such expressions were further validated by immunofluorescence. Reduced extracellular expressions of the AGR2 variants in FLO-1 were observed only in two variants, the AGR2 – WT and the AGR – KDEL, while no extracellular expression was observed for the mAGR2 – ΔKTEL variant.

The expressions of the ΔNterm AGR2 – KDEL variant either intracellularly or extracellularly could not be validated either by Western blot or immunofluorescence due to the lack of transfection efficiency. The accession on the influence of the AGR2 protein variants towards proliferation and migration of the FLO-1 observed from wound healing assay revealed significant elevations for FLO-1 proliferation and migration in AGR2 – WT, AGR2 – KDEL and mAGR2 – ΔKTEL (p<0.05). The viability of the FLO-1 transfected to AGR2 – WT, AGR2 – KDEL and mAGR2 – ΔKTEL through conduction of MTT assay revealed a significant result (p<0.05) supported by an increase in the absorbances of the cells with respect to time in contrary to Alamar Blue assay conducted which revealed an insignificant result (p>0.05) suggesting that the transfected AGR2 protein variants were independent towards the viability of the FLO-1.

Keywords: AGR2 protein, Esophageal adenocarcinoma, AGR2 – WT, AGR2 – KDEL, mAGR2 – ΔKTEL, ΔNterm AGR2 – KDEL, Wound healing assay, MTT assay, Alamar Blue assay.

KAJIAN KESAN VARIASI STRUKTUR PRO-ONKOGENIK PROTEIN AGR2 DALAM MENYOKONG KEMAJUAN BARAH ESOFAGUS

ABSTRAK

Sejenis barah tanpa kebergantungan hormon atau barah esofagus telah diklasifikasikan sebagai menduduki tangga keenam penyumbang kepada kematian ekoran peningkatan kes kematian yang berlaku disebabkan oleh kegagalan mengenal pasti kehadiran ketumbuhan tersebut semasa di awal peringkat penyakit seterusnya menyebabkan ketumbuhan tersebut berada pada tahap kronik atau metastatik semasa diagnosis. Kemajuan barah tersebut dilaporkan berlaku ekoran beberapa faktor dan antaranya termasuklah kehadiran protein yang bersifat pro-onkogenik seperti protein AGR2 yang menjalankan tugas lipatan protein serta menseimbangkan proteostasis ER. Kekurangan pemahaman terhadap fungsi struktur yang terdapat di dalam protein AGR2 serta sumbangannya terhadap kemajuan barah esofagus telah membawa kepada perangkaan kajian ini yang melibatkan penghasilan empat varian protein AGR2 yang merangkumi varian AGR2 – WT yang diwujudkan sebagai penanda aras kepada varian mutasi yang lain, varian AGR2 – KDEL yang dirangka bagi mengukuhkan penempatan protein AGR2 di dalam ER, varian mAGR2 – ΔKTEL yang dirangka bagi meningkatkan perembesan protein ke persekitaran ekstraselular dan varian ΔNterm AGR2 – KDEL yang dirangka bagi mengukuhkan struktur dimer protein tersebut. Kesemua viariasi AGR2 protein iaitu AGR – WT, AGR2 – KDEL, mAGR2 – ΔKTEL dan ΔNterm AGR2 – KDEL diklon ke dalam vektor pSF-CMV sebelum aras ekspresi varian-varian tersebut dinilai dengan mentransfeksikan varian-varian tersebut secara sementara ke dalam talian sel esofagus adenokarsinoma atau FLO-1 yang kemudiannya akan melalui proses pemblotan imun berserta pemblotan imun pendarfluor. Ekspresi intrasel bagi FLO-1 yang ditransfeksi dengan varian AGR2 – WT merekodkan tahap ekspresi tertinggi yang kemudiannya

diikuti oleh AGR2 – KDEL. Ekspresi varian – varian tersebut telah disahkan menerusi pemblotan imun pendarfluor. Ekspresi ekstraselular varian-varian AGR2 di dalam FLO-1 sementara tiada ekspresi ekstraselular direkodkan bagi varian mAGR2 – ΔKTEL. Ekspresi bagi varian ΔNterm AGR2 – KDEL sama ada di dalam intasel atau ekstraselular kedua-duanya tidak dapat dipastikan melalui pemblotan imun dan pemblotan imun pendarfluor lantaran kekurangan transfeksi yang efisien. Pengenalpastian terhadap kebergantungan varian-varian protein AGR2 terhadap percambahan dan migrasi FLO-1 telah dijalankan melalui ujian penyembuhan luka yang memperlihatkan peningkatan signifikan terhadap percambahan dan migrasi FLO-1 di dalam varian AGR2 – WT, AGR2 – KDEL dan mAGR2 – ΔKTEL (p<0.05). Daya maju sel-sel FLO-1 yang ditransfeksi dengan varian AGR2 – WT, AGR2 – KDEL dan mAGR2 – ΔKTEL dan diuji dengan ujian MTT menunjukkan keputusan signifikan (p<0.05) yang disokong oleh peningkatan bacaan penyerapan seiring dengan masa manakala sebaliknya ujian Alamar Biru yang dijalankan menunjukkan keputusan yang tidak signifikan (p>0.05) sekaligus mencadangkan bahawa varian protein AGR2 yang ditransfeksi tidak mempengaruhi daya maju sel-sel FLO-1.

Kata kunci: Protein AGR2, Esofagus adenokarsinoma, AGR2 – WT, AGR2 – KDEL, mAGR2 – ΔKTEL, ΔNterm AGR2 – KDEL, Ujian Penyembuhan Luka, Ujian MTT, Ujian Alamar Biru.

ACKNOWLEDGEMENTS

First and foremost, all praises and gratitude go to Allah, the Almighty, most merciful, and passionate for the wisdom, strength and good health He bestowed upon me throughout conducting the research until I am successfully capable of completing this research despite all the hardships and obstacles faced.

My deepest sense of gratitude goes to my supervisor and co-supervisor, Dr. Nikman Adli Bin Nor Hashim and Dr. Mohamad Aimanuddin Mohtar for their thoughtful guidance and supervision since the very beginning until the end of this research.

For my father, Baderol Hisham bin Hamzah, my mother, Noor Akma binti Jaffar as well as my siblings, thank you for being there, providing me with your utmost moral support, love and care in all aspects of my life. May Allah bless all of you at all times.

Huge thanks to UKM Medical Molecular Biology Institute (UMBI) acquaintances especially Asmaa, Nisa, Nurul Nadia and Syaza for the constant teaching and assisting in laboratory works. May Allah ease our next journey.

Special thanks to Universiti Malaya (UM) and UKM Medical Molecular Biology Institute (UMBI) for this remarkable and memorable collaboration.

Finally, thank you to all who have contributed directly or indirectly to this research. The research would not be completed, and this thesis would not be as it is without all these utmost assistances.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

±	:	Addition and subtraction operations
%	:	Percentage
^{o}C	:	Degree centigrade
18A4	:	Murine anti-AGR2 monoclonal antibody
AE	:	Buffer
AGR – KDEL	:	AGR2 protein with KDEL motif
AGR1	:	Anterior gradient 1 protein
AGR2	:	Anterior gradient 2 protein
AGR2 – WT	:	Wild type of AGR2 protein
AGR3	:	Anterior gradient 3 protein
AKT	:	Protein kinase B
AL	:	Lysis buffer
ANOVA	:	Analysis of variance
APPL1	:	Akt/protein kinase B-binding protein (an adaptor protein)
ATF6	:	Activating transcription factor 6
ATL	:	Lysis buffer
BCA	:	Bioinchoninic assay
BMI	:	Body mass index
BSA	:	Bovine serum albumin
C14B	:	Aptamer C14B
C4.4A	:	Glycosyl-phosphatidyl-inositol-linked receptor
C81	:	Single cysteine
CA125	:	Cancer antigen 125
CAFs	:	Cancer-associated fibroblasts

CCRT	:	Concurrent chemoradiotherapy
CD59	:	A glycoprotein
CDKN2A	:	Cyclin-dependent kinase inhibitor 2A gene
CDX2	:	Caudal-type homeobox 2
CHI3LI	:	Chitinase-3-like protein 1
CHN	:	Chimerin 1
CHOK1	:	Subclone of Chinese hamster ovary
CI	:	Confidence interval
cm	:	Centimetre
CO_2	:	Carbon dioxide
COOH	:	Carboxylic group
CPHS	:	Protein hydrolysate
CSSC	:	Recombinant chaperone protein
CTC	:	Circulating tumor cells
CXXC	:	C-cysteine, X-any protein
CXXS	:	Fold-independent redox motif
DAPI	:	4′,6-diamidino-2-phenylindole
DC	:	Direct current
DH5α	:	Competent *E. coli*
DMEM	:	Dulbecco's Modified Eagle Medium
DMSO	:	Dimethyl sulfoxide
DNA	:	Deoxyribonucleic acid
DTT	:	Dithiothreitol
EAC	:	Esophageal adenocarcinoma
eAGR2	:	Extracellular AGR2 protein
EALYK	:	Dimer interface

EB	:	Elution buffer
ECM	:	Extracellular matrix
EDTA	:	Ethylenediaminetetraacetic acid
EGFR	:	Epithelial growth factor receptor
ELISA	:	Enzyme-linked immunosorbent assay
EpCAM	:	Epithelial cellular adhesion molecule
ER	:	Endoplasmic reticulum
ERAD	:	Endoplasmic reticulum (ER)-associated protein degradation
ERdj3	:	Soluble ER Luminal Protein
ERK	:	Extracellular-signal regulated kinases
ERp18	:	ER-resident oxidoreductase
ERP44	:	Endoplasmic reticulum protein 44
ESCA	:	Esophageal carcinoma
ESCC	:	Esophageal squamous cell carcinoma
FBS	:	Fetal bovine serum
FGF	:	Fibroblast growth factor
FLO – 1	:	Human esophageal adenocarcinoma cell line
FOXA1	:	Forkhead Box A1
FOXA2	:	Forkhead Box A2
g	:	Gram
GA	:	Golgi apparatus
GEPIA2	:	Gene expression profiling interactive analysis
GERD	:	Gastroesophageal reflux disease
GLOBOCON	:	Global Cancer Observatory
hAgr2	:	Human *AGR2* gene
HE4	:	Human epididymis protein 4

HEPES	:	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, Zwitterionic sulfonic acid buffering agent
HIF – 1	:	Hypoxia induced factor-1
iAGR2	:	Intracellular AGR2 protein
IRE1	:	Inositol-requiring enzyme 1
IRE1α	:	Inositol-requiring transmembrane kinase endoribonuclease-1α
KDEL	:	K – lysine; D – aspartic acid; E – glutamic acid and L – leucine
KMT2D	:	Lysine methyltransferase 2D
KOH	:	Potassium hydroxide
$Kras^{G12D}$	:	A biomarker
KTEL	:	K – lysine; T – tyrosine; E – glutamic acid and L – leucine
KVEL	:	K – lysine; V – valine; E – glutamic acid and L – leucine
LB	:	Luria-Bertani
LOX5	:	Arachidonate 5-lipoxygenase
mAGR2 – ΔKTEL	:	Mature form of AGR2 protein without KTEL motif
MCF7	:	Human breast cancer cell line
mg	:	Milligram
mg/l	:	Milligram per litre
mg/m^2	:	Milligram per square metre
mg/ml	:	Milligram per millilitre
MG132	:	Carbobenzoxy-l-leucyl-l-leucyl-l-leucinal
ml	:	Millilitre
mm	:	Millimetre
mRNA	:	Messenger RNA
MTT	:	3-(4, 5-Dimethyl-2-thiazolyl)-2, 5-diphenyl-2H-tetrazolium bromide

MUC1	:	Mucin 1
MUC2	:	Mucin 2
MUC5AC	:	Mucin 5AC
Myc	:	Family of regulator gene
N3	:	Neutralization buffer
NaCl	:	Sodium chloride
Nco1	:	Restriction site
NEB	:	Restriction enzyme buffer
NFE2L2	:	Nuclear factor erythroid 2
ng/µl	:	Nanogram/microlitre
NH_2	:	Amine
nm	:	Nanometre
nm	:	Nanometre
NMR	:	Nuclear magnetic resonance
OD	:	Optical density
Opti – MEM	:	Optimized minimal essential medium
OR	:	Odds ratio
P1	:	Plasmid resuspension buffer
P2	:	Lysis buffer
p53	:	Tumor protein *p53* gene
PanIN – 3	:	Pancreatic intraepithelial neoplasia 3
PB	:	Phosphate buffer
PBS	:	Phosphate buffer solution
PDI	:	Protein disulfide isomerase
PDIA17	:	17^{th} member of the PDI family
PE	:	Washing buffer

PEBP4	:	Phosphatidylethanolamine binding protein 4
PET	:	Pesitron emission tomography
PFA	:	Paraformaldehyde
pH	:	Potential of hydrogen
pMA – T	:	Vector backbone
Prod1	:	Protein regulating limb regeneration in salamanders
PSA	:	Prostate-specific antigen
pSF – CMV	:	CMV promoter plasmid
RB	:	Retinoblastoma-associated protein
rcf	:	Relative centrifugal force
rpm	:	Revolutions per minute
RT – PCR	:	Reverse transcription polymerase chain reaction
SAG – 2	:	Secreted cement gland protein
SDS – PAGE	:	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SELEX	:	Systematic evolution of ligands by exponential enrichment
SMAD4	:	Mothers against decapentaplegic homolog 4
Smad4	:	Tumor suppressor gene
SP	:	Secretory pathway
SRP	:	Signal recognition particle
TBS	:	Tris-buffered saline
TBST	:	Tween 20-tris buffered saline
TCGA	:	The Cancer Genome Atlas
TFP	:	Three-finger protein
TGF – β	:	Transforming growth factor
Tip60	:	Histone acetyltransferase
TMED2	:	Transmembrane emp24 domain-containing protein 2

TNFAIP3	:	Tumor necrosis factor induced protein 3
TNM	:	Tumor niche microenvironment
TNS	:	Tensin protein, 2-p-toluidinylnaphthalene-6-sulfonate
TP53	:	Tumor protein p53
TxIYY	:	Penta–peptide docking motif
TXNDC12	:	Thioredoxin domain containing 12 proteins
UNG1	:	Uracil-DNA glycosylase
uPAR	:	Urokinase-type plasminogen activator receptor
UPR	:	Unfolded protein response
UPS	:	Unconventional protein secretion
USP19	:	Ubiquitin carboxyl-terminal hydrolase 19
VEGF	:	Vascular endothelial growth factor
WHO	:	World Health Organization
XAG – 2	:	Xenopus cement gland-specific gene
XbaI	:	Restriction site
XBP1s	:	X-box binding protein 1
ΔNterm – AGR2 KDEL	:	AGR2 protein without N-terminal with KDEL motif
µg/ml	:	Microgram per millilitre
µg/µl	:	Microgram per microlitre
µl	:	Microlitre
µm	:	Micrometre
µM	:	Micromolar

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Research background

Among the most common types of cancer worldwide, the esophageal cancer recently reported to rank ninth in the world in respect to its total number of incidences of 604,100 cases in 2020 alone. This cancer was classified as the sixth most deadly cancer worldwide with the occurrence of 544,076 deaths as of 2020, according to a report released by the Global Cancer Observatory (GLOBOCON) by Sung *et al*. (2021) The incidence of esophageal cancer which contributed to an up to 5.5% total number of deaths worldwide is worrying due to the high mortality rate in which one of the reasons was due to the fact that they failed to be detected at an early stage resulting in advanced or metastasized tumours upon diagnosis (Uhlenhopp *et al*., 2020). For instance, Uhlenhopp *et al*. (2020) reported that in the United States, only 18% of the reported cases of esophageal cancer was found to remain residing at the primary site . The remaining percentage of esophageal cases failed to be detected at either advanced or metastasized state. This cancer type in addition was known to have a significant distribution across all countries with male dominating the number of incidences as compared to the female. Esophageal cancer was also found to be common for those ages 55 and above constituting to an up to 88% of the total cases.

The esophageal cancer generally categorized into two different types which were the esophageal adenocarcinoma (EAC) and esophageal squamous cell carcinoma (ESCC). The differences in these two types of esophageal cancer lie in many aspects and one of it happened to be the position of the carcinoma. EAC was known to occur at the distal part of the esophagus and gastroesophageal junction while the ESCC was usually occurred at the proximal two thirds of the esophagus (Uhlenhopp *et al*., 2020). Other than that, EAC was found to be distinctive from the ESCC through its prevalence. The ESCC was

reported to be the most prevalence type worldwide as compared to the EAC. Even so, in developed countries such as the US and few other Western countries, the EAC recently was becoming the most prevalence type such that in the US alone, EAC contributed to an up to 64% cases followed by the ESCC with 31% cases and the rest 5% was either the basal, transitional or unspecified carcinomas. The common risk factors reported to be associated to the ESCC included alcohol consumption, tobacco and nitrosamines while the EAC was reported to occur due to the Barrett's esophagus, a condition in which the esophagus lining the mouth to the stomach was badly damaged by the acid reflux resulting in the thickening of the lining and swollen. In addition, gastroesophageal reflux disease, obesity and tobacco consumption were also reported to contribute to the emergence of EAC. Patients of both types of esophageal cancer required radiotherapy and stenting as palliative treatment as well as definitive chemoradiotherapy for ESCC and neoadjuvant or perioperative chemotherapy for EAC in which both were then followed by surgery as a curative treatment.

There were many factors that could promote to the emergence of cancer. One of the factors was the occurrence of protein with pro-oncogenic properties. AGR2 or anterior gradient 2 protein was a type of protein encoded by the human *AGR2* gene. This protein was found to be residing in the endoplasmic reticulum and played an important role in the protein folding mechanism. They were also involved in the formation of disulfide bonds within protein which included catalyzing, breaking, and isomerization of bonds. Even so, such protein belonging to the protein disulfide isomerase (PDI) family was reported to have abnormal functions such as they were capable of regenerating amphibian's limbs and could promote cancer metastasis in humans.

AGR expression could be found in both hormone-dependence cancers which included breast cancer cell line, ovarian cancer (Armes *et al*., 2013; Park *et al*., 2011) and prostate cancer (Kani *et al*., 2013; Neeb *et al*. 2014; Vitello *et al.*, 2016) as well as in non-hormone

dependent cancers such as in esophageal cancer (O'Neill *et al*., 2017), gastric cancer (Zhang *et al*., 2016), lung cancer (Milewski *et al.*, 2017) as well as head and neck cancer (Ma *et al*., 2015).

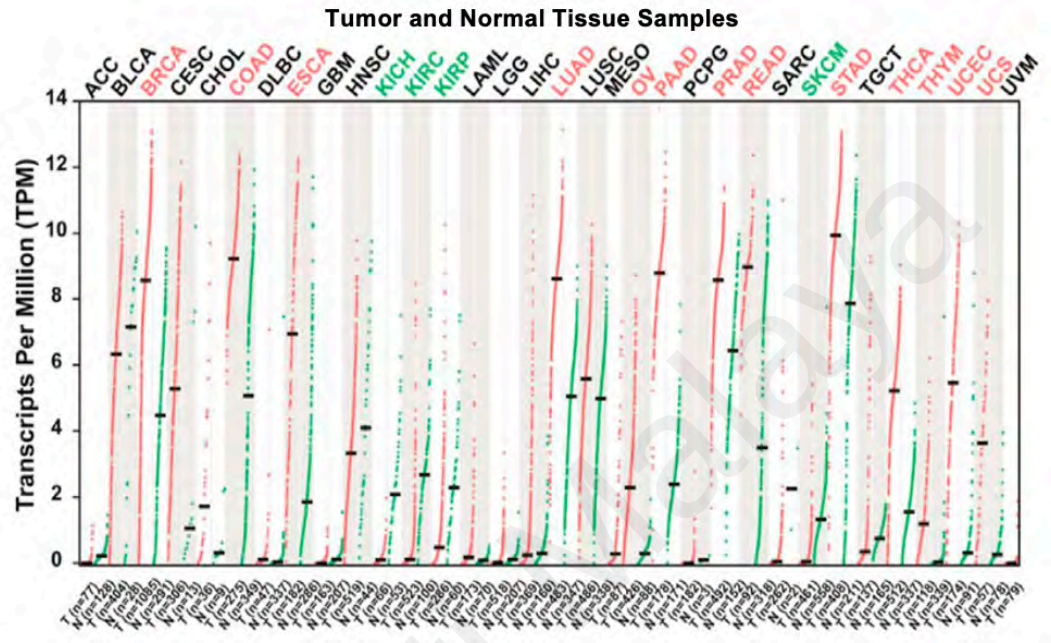


Figure 1.1: Dot plots summarizing different *AGR2* mRNA expression in varieties of normal and cancer tissue samples extracted from the Gene Expression Profiling Interactive Analysis 2 (GEPIA2) web tool. The types of cancer labelled in red represented the upregulation of AGR2 while the reduced expression of AGR2 was signified by the green label (Taken from Mohtar *et al*., 2020).

Dumartin *et al*. (2011) reported the occurrence of AGR2 both in the ER or intracellular (iAGR2) and on the surface or extracellular (eAGR2) of cancer cells via immunofluorescence and flow cytometry analyses. Both the intracellular and extracellular roles of AGR2 was reported to be highly attributed to its unique structure. This protein which was initially found in breast cancer cell line, MCF7 had a primary sequence which comprised of a cleavable signal sequence that lie within residues 1 and 20 of the N-terminal and responsible for the protein import into the endoplasmic reticulum (ER). The next unique feature of this protein was the C – terminal known as KTEL, an abbreviation for K – Lysine; T – Tyrosine; E – Glutamic acid and L – Leucine which served as an ER retention sequence. AGR2 protein also had single CXXS motif, a

catalytically active pseudo – thioredoxin domain which was suggested to have role in the formation of mixed disulfides with intestinal mucins, MUC2 as well as MUC1 and MUC5AC (Schroeder *et al*., 2012). Located at the N-terminal was the amino acid 21 – 40 which was unfolded and primarily responsible for the cell adhesion properties of the AGR2. Another functional domain in AGR2 protein was the folded part which formed a dimer via specific intermolecular salt bridges through EALYK motif at E60 – K64 position. The monomeric AGR2 (E60A) or the native homodimeric AGR2 was also found to be responsible for the cell adhesion properties of AGR2 protein.

To date, the effort to search for the ways to repress the growth and metastasized of cancer cells was still ongoing despite the amount of findings accumulated as cancer progression was known to differ among patients suffering from the disease depending on the types and severity of the cancer. Therefore, since AGR2 protein was known to have role in promoting the growth of cancer cells, we aimed to establish four AGR2 structural variants based on the existing mutational studies in order to study the effect of altering specified AGR2 structures towards the intracellular and extracellular role of AGR2 protein in supporting esophageal cancer cell progression. We established four AGR2 protein variants based on the previous report on the functional structure of the AGR2 protein. The four variants included the wild type of AGR2 (AGR – WT) which served as a comparison to the other mutated variants, AGR2 protein with the substitution of canonical KTEL motif into the KDEL motif (AGR – KDEL) which was designed to enhance ER localization, the mature form of AGR2 underwent deletion of signal peptide as well as the KTEL motif (mAGR2 – ΔKTEL) to enhance ER secretion and AGR2 protein which underwent deletion of N-terminal and bearing the KDEL motif (ΔNterm – AGR2 KDEL) to enhance the protein dimer. Such four variants were introduced into the esophageal adenocarcinoma cell line, FLO-1 and from these AGR2-transfected cell line, the basal expressions of both iAGR2 and eAGR2 were determined in esophageal

adenocarcinoma and the effects of the transfected variants towards the proliferation and migration of the adenocarcinoma esophageal were determined.

1.2 Problem statement

The function of AGR2 protein in esophageal cancer cell was not fully understood. Recent evidences showed that AGR2 had dual functions. First, AGR2 was expressed predominantly in the endoplasmic reticulum and this had been associated with its intracellular (iAGR2) role as a catalyst during the ER proteostasis. Second, AGR2 could be secreted extracellularly (eAGR2) and remodelled extracellular matrix and epithelial integrity. Hence, in this proposed study, we aimed to understand the dual roles of AGR2 protein by studying the reported structural variants in esophageal cancer cell line. Specifically, four structural AGR2 variants from previous studies would be selected in this study. The intracellular and extracellular expression of these AGR variants would be assessed in esophageal cancer cell models. This was followed by functional assays to determine whether the expression of the variants had any phenotypic effect on the esophageal cancer cells. From here, we were able to understand the mechanism of how AGR2 protein was retained and secreted in cancer.

1.3 Research questions

1. Which AGR2 constructs could be selected to study the AGR2 functions based on the current AGR2 structural variations and other mutational studies?
2. What were the basal level of expressions of both intracellular and extracellular of the said AGR2 protein variants in esophageal cancer cell model?
3. What were the phenotypic effects of the AGR2 variants expression in esophageal cancer cell model?

1.4 Aim and objectives

1.4.1 Aim

To understand the effects of structural variation of AGR2 protein in the progression of esophageal cancer.

1.4.2 Objectives

1. To construct vectors for the expression of recombinant AGR2 protein variants in mammalian cells.
2. To determine the level of intracellular and extracellular expression of AGR2 protein variants expressed from the esophageal cancer cells.
3. To determine the viability and migratory potential of esophageal cancer cells expressing AGR2 protein variants.

CHAPTER 2: LITERATURE REVIEW

In 2020, cancer generally had been classified as the first and second leading cause of death for patients with below than 70 years old in 112 countries and ranked third and fourth in the remaining 23 countries by the World Health Organization (WHO). With respect to the Global Cancer Statistics (GLOBOCAN) 2020, Bray *et al.* (2018) provided an estimation on the incidence number of cancer cases as well as number of mortalities reported due to cancers worldwide. The data which was produced by the International Agency for Research on Cancer had shown that in 2020 alone, approximately 19.3 million of new cancer cases with an up to 10 million cancer deaths had been reported (Bray *et al.*, 2018). Out of ten highest incidences of cancer, esophageal cancer was ranked as ninth signifying that their incidences was also worrying such that they contributed to one in 18 cancer deaths as reported in 2020 (Bray *et al.*, 2018).

2.1 Esophageal cancer

2.1.1 Incidence, mortality and survival rates

Comparing the incidence rate between different types of esophageal cancer, the incidence of EAC contributed to only 11% while ESCC was reported to be an up to 87% from the total number of esophageal cancer cases in the world as reported by the Arnold *et al*. (2015). According to a study of a standardized-age done by Uhlenhopp *et al*. (2020), the incidence rate for both types of esophageal cancer however, was found to be the highest in Eastern Asia with 12.2 per 100,000 people reported to suffer from the disease. Such high incidence was followed by the Eastern and Southern Africa with 8.3 and 7.4 out of 100,000 reported with the disease respectively. In contrast, lowest rate on the incidence of esophageal cancer was reported to be in Central America with only 0.98 per 100,000

people while 78% of all reported cases of esophageal cancer actually came from Asia with 49% cases were observed in China alone (Bray *et al.*, 2018). Bray *et al.* (2018) in the same study also investigated the incidence of esophageal cancer in the US. From the study, US recorded low incidence with 5.6 per 100,00 cases. Within the US, the Caucasians had recorded the highest confirmed cases within 2012 – 2016 time points and the EAC cases were recorded to be higher than the ESCC. Comparing the incidence of esophageal cancer in between genders, Uhlenhopp *et al*. (2020) reported that higher incidence of esophageal cancer was recorded in men as compared to the women with 70% of the cases worldwide was seen to occur in men. Males' higher prevalence of esophageal cancer was partially explained by the occurrence of abdominal obesity and smoking habit in addition to the gastrophageal reflux disease (GERD) which contributed to the onset of this disease. Men were reported to develop ESCC at four times higher risk than women and EAC at seven to ten times higher risk as compared to the women. Both sexes however, showed an increment in the incidence rate as proportional to the increase of the age.

In terms of mortality, the estimated age-standardized mortality rates was 5.5 per 100,000 cases. In 2020 alone, 544,076 mortality cases due to esophageal cancer were recorded worldwide and out of them, 374,313 cases comprised of male patients while the rest 169,763 cases were among the females (Sung *et al*., 2021). Globally, the mortality rate for age-standardized of both sexes was 10.7 per 100,000 people and turned out to be the highest in Eastern Asia.

Esophageal cancer was known for its poor prognosis which resulted from the late diagnosis. Just like any other cancer incidences, the survival rate of esophageal cancer also decreased with the increase of the cancer stage. The reported survival rate for both regional and distant metastasis of esophageal cancer was five years with 25.1% and 4.8% survival rate respectively. Despite the differences in incidence rates between males and

females however, the average 5 years rate of survival was somehow the same and males recorded 19.4% survival cases while female recorded 21.5% survival rate (Uhlenhopp *et al*., 2020).

2.1.2 Pathogenesis and risk factors

The onset of esophageal cancer in patients was often reported due to the chronic inflammation in the esophagus which resulted in disruption of both normal cell signalling as well as growth of the cells. In terms of pathogenesis, EAC basically rose from the Barrett esophageal which occurred when preneoplastic tissue such as squamous esophageal epithelium was being substituted with columnar intestinal-type mucosa (Smyth *et al.*, 2017). The occurrence of gastro-esophageal reflux of acid also contributed to the onset of EAC. Long-term exposure of esophageal mucosa towards acid or bile might result in the emergence of reactive oxygen species and nitric oxide which resulted in detrimental effects to the DNA or transversion from A base to C base and classified as early pathogenesis of esophageal cancer (Dvorak *et al*., 2011). Barrett esophageal was identified by having preneoplastic lesion in which such genetic alterations may resulted in carcinogenesis. Smyth *et al.* (2017) suggested two mechanisms on the emergence of EAC from Barrett esophageal. The first mechanism proposed was the loss of tumor suppressor genes such as *CDKN2A* and *TP53* just like ESCC with addition to mutation occurring in *SMAD4* and interference of chromatin modifying enzyme. Weaver *et al.* (2014) for instance also supported the mutations of *CDKN2A* and *TP53* as early tumor development for EAC. The second mechanism of development of EAC from Barrett esophageal was the chromosomal instability occurring extensively which contributeed to the reduction of *p53* regulation due to the loss of heterozygosity of 17p and resulted in aneuploidy thus increasing malignant progression (Stachler *et al*., 2015). Other mechanisms apart from the main two were chromothripsis and kataegis which also

contributed to the transformation of Barrett esophageal to invasive EAC (Smyth *et al*., 2017). Martinez *et al*. (2016) in another research had proved a greater clonal diversity in Barrett esophageal to be the mechanism involved in advancement of Barrett esophageal to invasive EAC.

In contrast to EAC, development of ESCC was often associated with basal cell hyperplasia and dysplasia before reaching the carcinoma in situ stage. ESCC tended to occur when there was an occurrence of dysregulation of *TP53*, genes that encoded for cell cycle regulators such as *cyclin-dependent kinase inhibitor 2A* (*CDKN2A*) as well as retinoblastoma-associated protein (RB) in which high level of *CDKN2A* and RB contributed to the transformation of esophageal lesions from inflammation to cancer (Müller *et al*., 2014). During this transition, *TNFAIP3* which encoded for tumor necrosis factor induced protein 3 as well as *CHN* which encoded for chimerin 1 were found to be expressed at higher level. Smyth *et al*., (2017) even reported that a The Cancer Genome Atlas (TCGA) data showing that points mutations and indels in *TP53*, *KMT2D* which encoded for lysine methyltransferase 2D and *NFE2L2* which encoded for nuclear factor erythroid 2 had been detected to give rise to esophageal cancer.

Several risk factors had been reported to be associated to the esophageal cancer incidences and among them included smoking. Smoking was determined as the major cause of both EAC and ESCC. Huang & Yu (2018) reported that the incidence of EAC among men with smoking history (OR = 2.10; 95% CI, 1.71 – 2.59) was significantly higher than women (OR = 1.74; 95% CI, 1.21 – 2.51). Men who quitted smoking in addition also reported to have higher risk of suffering from EAC as compared to those who never smoked and this was due to the chemicals presented in cigarette smoke which passed through the esophageal cancer thus resulting in Barrett's esophagus condition.

Second risk factor associated to the emergence of esophageal cancer was alcohol consumption. Severe alcohol consumption had higher risk of rising into esophageal

cancer due to the acetaldehyde which occurred when ethanol was metabolized by alcohol dehydrogenase and formed interaction with DNA thus resulting in DNA adducts that could induce gene mutation. The excessive alcohol intake to an up to 170 g or more per week might cause upregulation of ESCC but not EAC among esophageal cancer patients (Huang & Yu, 2018). Other important risk factor for esophageal cancer was the occurrence of gastroesophageal reflux disease (GERD). Patients diagnosed with GERD had an increased risk of progressing into EAC. Obesity was also reported to contribute to the increase of EAC. Individuals of high body mass index (BMI) and increased abdominal obesity tended to be at risk of developing EAC but decreased risk of developing ESCC. Huang & Yu (2018) reported increased risk of EAC among obese males (OR = 2.4; 95% CI, 1.9 – 3.2) and females (OR = 1.9; 95% CI, 1.5 – 2.5).

2.1.3 Symptoms, screenings and treatments

Patients diagnosed with esophageal cancer usually only showed symptoms when the obstructive lesion or stricture had grown into advanced local or metastatic stage. Symptoms such as dysphagia or difficulty in swallowing, odynophagia or pain in swallowing, hoarseness or cough and progressive weight loss could be seen in most patients. Certain patients also suffer from serious symptoms such as blood vomiting, and pass melaena in which dark stool occurs as a result of internal bleeding of upper gastrointestinal tract (Smyth *et al*., 2017).

The gold standard used for the screening of esophageal cancer was endoscopy. The T (size and extent of the main tumor or primary tumor) and N (number of nearby lymph nodes that had cancer) stages of esophageal cancer could be identified using endoscopic ultrasonography while the M (whether the cancer had metastasized) stage could only be determined by pesitron emission tomography (PET) or PET – CT. Early ESCC could be detected using endoscopy and chromoendoscopy in which Lugol's iodine dye was applied

for better resolution of face mucosa. Transnasal endoscopy could be applied for detection of Barrett's esophageal which usually developed into EAC. Histological confirmation was also recommended for detection of esophageal cancer. Cytokeratin 5, cytokeratin 6 and p63 were markers that could be used in histochemical or immunohistochemical for detecting ESCC while periodic acid-Schiff staining, cytokeratin 7, cytokeratin 20 could be used for detection of EAC (Wong & Chu, 2012).

Moving on to the treatment of esophageal cancer, Merkow *et al*. (2012) reported on the use of multimodality neoadjuvant concurrent chemoradiotherapy (CCRT) in treating esophageal cancer. The survival rate among the patients suffering from the advanced esophageal cancer had showed a convincing improvements with the use of CCRT together with surgery, according to a study done by Hagen *et al.* (2012) on a randomized trials. Apart from that, postoperative adjuvant therapy also showed significant improvement in the patients' survival rates at the same time reducing the period of admission to the hospital among the EESC patients. The efficiency of chemotherapy on patients with ESCC could be measured through administration of nimotuzumab (200 mg) weekly, paclitaxel (175 mg/m^2) on the first day and cisplatin (30 mg/m^2) on first and second days with the cycle repeating every 3 weeks over six cycles. The overall response rate was recorded at 51.8% with patients' survival rate to an up to 20.2 months (95% CI, 11.5 – 28.9 months) (Hsu *et al*., 2016). Meanwhile, as for the EAC patients, neoadjuvant chemoradiation therapy had been implemented widely in an attempt to treat locally advanced or lymph node positive tumours. A study by Alnaji *et al*. (2016) revealed that out of 205 patients bearing the T2 – T4 or lymph nodes positive EAC which underwent neoadjuvant chemoradiation therapy had resulted in 86% of three years overall survival with 80% recurrence–free survival rates. This further supported the capability of neoadjuvant chemoradiation therapy in treating EAC.

2.2 AGR2 protein

AGR2 or anterior gradient – 2 protein was one of the proteins belonging to the protein disulfide isomerase (PDI) family and found to be residing in the endoplasmic reticulum (ER). Being the 17th member of the PDI family, AGR2 was also known as PDIA17. Generally, AGR2 was among the one-third of the proteome which passed through the secretory pathway (SP) before entering the ER in which several chaperone families of proteins including PDI family were accumulated. Within the PDI family itself was the AGR2 protein which was generally known to be involved in the formation of disulfide bonds as well as responsible for maintaining the quality control of proteins found in the ER.

Strolling down the lane into which it first discovered, the protein which also termed as secreted cement gland protein SAG-2 homolog was originally found in *XAG-2* gene of *Xenopus laevis* through several attempts of dissecting embryos of different ages. The expression of such gene was mainly focused on the anterior region of dorsal ectoderm corresponding to cement gland anlage. AGR2 protein had been encoded in humans by the *AGR2* gene from the AGR2 subfamily which comprised of AGR2, AGR3 and AGR1 or TXNDC12 (Delom *et al*., 2020). Such protein was also found to be more prone towards alterations due to its localization on the 7p21.3 chromosome. Basically, there were 8 exons that made up the *hAgr2* gene with capability of rising into seven transcript variants in which only four could encode for protein products including AGR2 – 201, AGR2 – 202, AGR2 – 203 and AGR2 – 204 (Moidu *et al*., 2020). The presence of AGR2 protein in amphibians was not solely in *X. laevis* however, as reviewed by Delom *et al*. (2020) in salamander, AGR2 protein aided limbs regeneration from dedifferentiated cells as well as stem cells. Meanwhile, in humans, the first occurrence of AGR2 protein as described by Thompson & Weigel (1998) was in estrogen-receptor-positive breast cancer cells, MCF-7. As advancement in research progressed, more and more expressions of AGR2

protein were found in cancers which further proved AGR2's role as an epithelial barrier. Despite being an ER-resident protein however, several studies somehow managed to prove their localization in other subcellular locations which promoted its interactions with other cellular proteins for instance, proteins involved in signal transduction as well as proteostasis control. Besides, the localization of AGR2 protein even found to be in nucleus (Li *et al*., 2015), cytoplasm (Yu *et al*., 2012), mitochondria (Liu *et al*., 2016), extracellular matrix (Fessart *et al*., 2016), surface of the cell as well as in blood and urine (Fessart *et al*., 2016).

In the review by Delom *et al*. (2020), AGR2's role as redox chaperone was suggested in folding secreted proteins of Alpha-1 antitrypsin and mucins as well as their collaboration with protein substrates for the formation of mixed disulfides. Another important role of AGR2 in normal tissues could be seen in the regulation of total protein loads occurring in cells. This could be demonstrated for instance by the normal mammary gland in which AGR2 was highly expressed during two situations which were during late pregnancy and lactation period (Verma *et al*., 2012). During the occurrence of these two periods, more protein folding, and secretory elements were needed hence, AGR2 proteins were upregulated to aid in the increased of protein loads.

After it first discovery in MCF-7 breast cancer cell line, AGR2's existence began to be traced in both hormone-dependence as well as non-hormone dependence cancers. Examples of the hormone-dependence cancers included ovarian cancer and prostate cancer while the non-hormone dependence cancers included in esophageal cancer, lung cancer as well as the head and neck cancer. Due to their pro-oncogenic properties, any overexpression of AGR2 protein was said to promote clonogenic growth of cancer cells thus lengthening the survival of such cancer cells. In contrast, depletion of AGR2 could result in decreasing of cells proliferation and migration such as the one described by Xue

et al. (2018) whereby depletion of AGR2 resulted in reduction of phosphor-AKT level in non-small cell lung cancer.

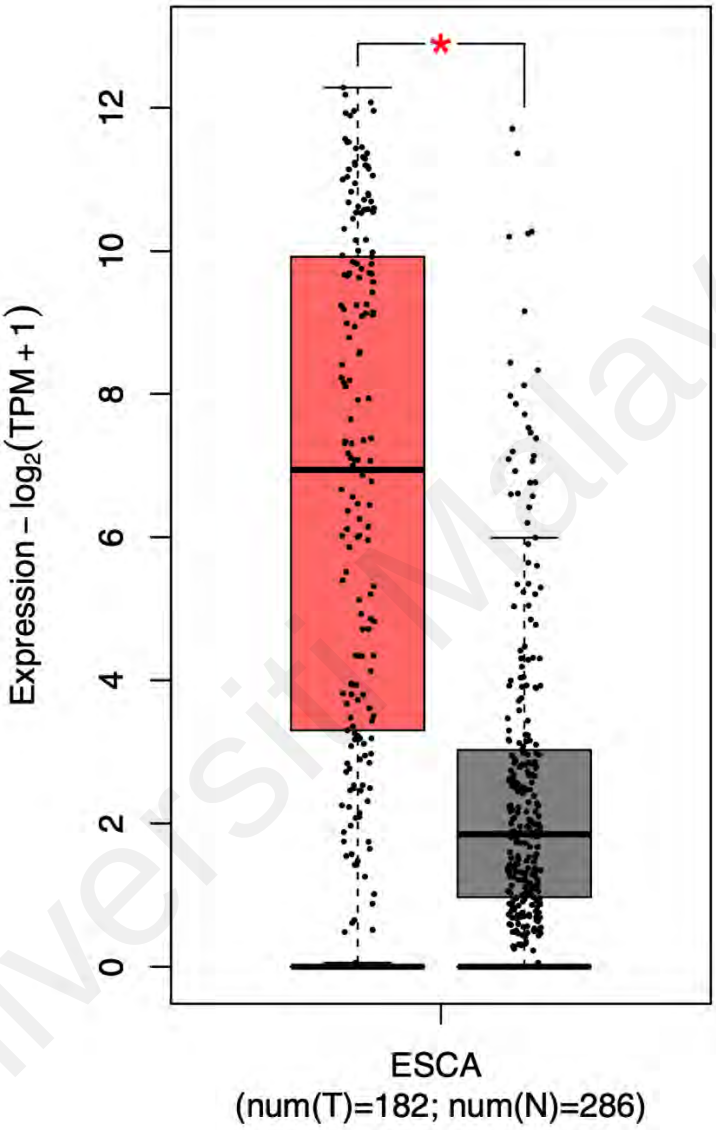


Figure 2.1: Box plot showing signature score calculated by mean value of log_2(TPM + 1) for distribution of AGR2 protein in esophageal tumors (ESCA). The red box signified the tumors (T) while normal tissues (N) were represented by grey (Taken from GEPIA, 2018).

Moidu *et al*. (2020) in their paper described on an analysis done by The Cancer Genome Atlas (TCGA) PanCancer Atlas studies in which they found out that AGR2 protein was mostly upregulated in the cancer landscape and that any mutations of AGR2

had implication towards development of cancer cells such that mutating AGR2 protein could decrease the progression of cancer cells.

2.2.1 Structural roles of AGR2 protein

A total of 175 amino acids became the fundamental in construction of AGR2 protein resulting in a protein with molecular mass of 17 kDa (Delom *et al*., 2020).

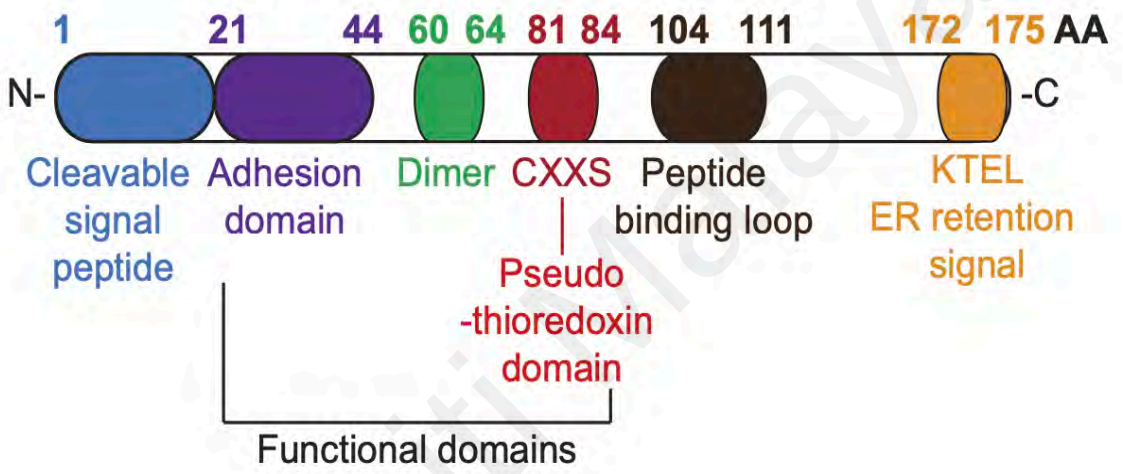


Figure 2.2: Primary structure of AGR2 protein. The functional domains and important amino acids which aided in AGR2 function were represented by the colored figure. AA, amino acid; ER, endoplasmic reticulum; K – lysine; T – tyrosine; E – glutamic acid; L – leucine (Taken from Delom *et al*., 2018).

Figure 2.2 depicted the full-length structure of AGR2 protein. From amino acid 1 to 20 lied the cleavable signal sequence NH_2 – terminal followed by the adhesion domain from amino acid 21 to 44 which found to be responsible for cell adhesion property of the protein. The first 20 amino acids bearing the hydrophobic signal peptide sequence at N-terminus was crucial in importing AGR2 protein into the ER. Another functional domain which lied from amino acid 81 to 84 was the pseudo-thioredoxin domain or CXXS active domain motif which functioned for the AGR2 protein's oxidation and reduction reactions (Galligan & Peterson, 2012). Such motif also responsible for AGR2 classification in the PDI family. Basically, thioredoxin fold was made up of CXXC motif and two dual

cysteines and said to be prominence in shuffling of disulphide bonds during the maturation of protein. Even so, AGR2 protein actually comprised of only single cysteine (CXXS) resulting in lowered oxidoreductase activity but found to be efficient in isomerization of disulphide bonds at the same time having specialized functions in the ER (Moidu *et al*., 2020). Such thioredoxin fold was responsible for the motif's placement as part of oxidoreductase family. ERP44 was the example of oxidoreductase bearing the CHHS motifs.

At amino acid 172 to 175 on the other hand was the KTEL COOH-terminal motif which was identical to both KDEL and KVEL ER retention sequences (Gupta *et al*., 2012). The C-terminal KTEL was the one responsible in targeting AGR2 into the ER. The KTEL was basically a classic linear peptide motif known as ER retention motif consisting of the tetrapeptide sequence of lysine (K) – threonine (T) – glutamic acid (E) – leucine (L). KTEL was said to be the variant of ER localizing motif, KDEL. KDEL motif contributed to the transport of proteins in ER during the protein secretion process into the extracellular space. Proteins bearing KDEL motif often formed interaction with KDEL receptors in the cis-Golgi complex specifically at the intermediate compartments. Such binding that occurred in return allowed for conformational changes of KDEL receptor which guided the ligand-containing receptor into the vesicles which consequently would be returned to the ER. Any increasing of the pH may cause the KDEL peptide motif to dissociate from KDEL receptor.

Then, AGR2 protein basically existed as homodimer structure through motif EALYK at position E60 – K64 as shown by Patel *et al*. (2013) via the nuclear magnetic resonance (NMR) crystallography. Such dimer was stabilized by Glu60 – Lys64 interface and the AGR2 protein subunits were found to be interacted with each other in a reverse parallel form and this occurred due to the salt bridge formed from E60 carboxylate as well as K64 ammonium groups. Even so, any mutagenesis occurring at the E60A dimerization motif

could result in the loss of its homodimer structure. AGR2 protein dimerization interface in addition was positioned far from the catalytic CXXS motif such that it could stimulate any disulfide exchanges between the client proteins (Delom *et al*., 2020). Formation of homodimer was also possible through oxidation-dependent homodimerization occurring through C81-mediated disulfide bond formation resulting in distinct re-orientation of homodimer (Ryu *et al*., 2013). Such homodimerization of AGR2 was also made possible upon chemical peroxide exposure.

Any mutations occurring which resulted in replacement of single C81 cysteine to alanine could prevent hydrogen peroxide from catalyzing the homodimerization of AGR2 protein. The establishment of labile sulfenic acid intermediate was due to the result of intermolecular disulfide bond formed from low level of chemical oxidants. High level of oxidants in contrast stimulated the formation of sulfinic and sulfonic acid that could obstruct homodimerization of AGR2 protein. This further strengthened the fact that single cysteine (C81) of AGR2 protein was responsible for oxidant-responsive moiety for the oxidation role of AGR2 protein as well as its monomeric-homodimer state (Delom *et al*., 2020). Such homodimerization was even claimed to be linked with the NH_2 terminus such that deletion of disorder NH_2-terminal located at amino acid 21 – 45 could result in an improved stability of AGR2 protein homodimer (Gray *et al*., 2013). Even so, homodimerization did not actually contribute to the cell adhesion properties of AGR2 protein as AGR2 protein often formed interaction with cell receptors via thioredoxin motif to aid in cell adhesion (Delom *et al*., 2020). The equilibrium formed between the monomer and homodimer on the other hand was crucial for AGR2 signaling. Most of the abundant AGR2 homodimer were positioned in the ER and found to be crucial for the folding of cysteine-rich client protein. The nature of AGR2 protein which reacted to oxidizing conditions also contributed to the equilibrium of monomer and homodimer. Such equilibrium was often disturbed by redox imbalance environment formed from

redox signaling mediators which existed in the ER of cancer cells, stimulating reactive oxygen species thus contributed to AGR2's oncogenic functions.

Besides, Maurel *et al.* (2019) proved involvement of AGR2 homodimer in dictating ER proteostasis such that they played a role as a sensor. Such role was further improvised by TMED2 which served as an enhancer. Both *in vivo* and *in vitro* interaction conducted had successfully resulted in segregation of AGR2 protein and TMED2 upon ER stress. The interaction between AGR2 and TMED2 even included K66 and Y111 amino acid residues.

Being a protein under PDI family, AGR2 through phylogenetic analysis was found to have a CPHS active site sequence. However, the activity of such sequence remained undetermined. Certain other proteins in PDI family also showed chaperone activity due to the hydrophobic b and b domains apart from redox motif. PDI proteins tended to reside in ER and sometimes escaped to function at the cell surface or in extracellular matrix.

Ten phosphorylation sites for AGR2 protein were also discovered using UniProt and PhosphoSite including S36p, Y111p, T113p, T114p, S119p, Y124p, S146p, Y150p, Y152p and T157p in which mostly were acquired from both cancerous tissues and cell lines (Consortium, 2018; Hornbeck *et al*., 2015).

2.2.2 AGR2 interactomics

Several research had proved AGR2's capability of binding to many proteins including nuclear, cytosolic and plasma membrane proteins. Out of all proteins, AGR2 protein basically interacted the most with plasma membrane proteins. This could be due to involvement of AGR2 protein in secretory pathways as well as proteostasis. In the review by Delom *et al*. (2020), an attempt had been done to study AGR2 protein by investigating the peptide motif bounded to AGR2 protein. Such research had been conducted since AGR2 was the only PDI which was equipped with the specific peptide-binding activity.

This could support the hypothesis that AGR2 protein could bind to proteins of the same motif if they were capable of binding to sequence-specific penta-peptide motif. Hence, a consensus AGR2 peptide motif was constructed and found to be abundance in membrane-associated protein.

AGR2 protein also found to be responsible in bringing EpCAM through secretory pathway at the same time making sure the proper maturation of such receptor protein, a role known as membrane trafficking. AGR2's role towards its client protein was substantiated using in-vitro and cell-based assays as well as using hydrogen deuterium mass spectrometry (Delom *et al*., 2020). O'Neill *et al*. (2017) in a different paper showed that both AGR2 and EpCAM were classified as cancer-associated-proteins such that an immunohistochemistry conducted on esophageal adenocarcinoma tissues resulted in high expression of both proteins. Previous study proved interaction occurring between AGR2 protein and EpCAM to be in the ER such that EpCAM actually established a model client protein for AGR2 protein hence making them involved in regulation of specific protein trafficking just like EGFR. This again highlighted AGR2's role in secretory pathway, a mechanism known as AGR2 proteostasis interactome which could increase the successful chances of using AGR2 protein as target for anti-cancer therapy.

Delom *et al*. (2020) in their paper even highlighted the presence of AGR2 both in intracellular (iAGR2) and extracellular (eAGR2) of ER as found on pancreatic cancer cells via immunofluorescence and flow cytometry. Such position of eAGR2 on the surface of the cell could further support the hypothesis that AGR2 protein could be attached to client receptor protein. Fessart *et al*. (2016) even showed that eAGR2 could form interaction with extracellular matrix such that they were secreted in microenvironment and this included in tumor microenvironment as reviewed by Delom *et al*. (2018). This could also be seen for instance in a research by Hong *et al*. (2013) where regulating hypoxia induced factor-1 (HIF-1) by incorporating exogenous AGR2

protein could stimulate angiogenesis. Fessart *et al*. (2016) in addition discovered that eAGR2 protein was responsible for microenvironmental pro-oncogenic regulator for both epithelial morphogenesis and tumorigenesis. eAGR2 protein at the same time capable of triggering insulin-like growth factor-1 receptor-induced cell proliferation, migration as well as transitioning of epithelial to mesenchymal via interaction of membrane ER – α in breast cancer cells (Li *et al*., 2016). Tian *et al*. (2018) on the other hand proved eAGR2's capability of stimulating migration and metastasis via Wnt11 signalling in colorectal cancer.

Being a protein in PDI family, AGR2 protein was also found to perform interaction with a protein responsible for base-excision repair in mitochondria known as UNG1 (Liu *et al*., 2016). UNG1 was prone to oxidative damage such that their highly expression could result in increasing cell resistance in order to protect mRNA from oxidative stress. From here, it was found that any interactions occurring between UNG1 as well as AGR2 protein could improve the stability of UNG1 at the same time enhancing the enzymatic activity occurring during DNA repair mechanism. Chevet *et al.* (2013) in their paper stated that AGR2 protein served as an intermediate in between ER and tumor development such that their interactome in cancer could increase proteostasis hence causing the tumor cells to withstand the production of abnormal protein thus allowing tumor development process to take place.

Li *et al*. (2015) in their research showed that AGR2 could bind and stabilize hypoxia induced factor 1, HIF-1α such that in breast cancer, there were differences in chemoresistance levels which were due to chemical hypoxia-induced doxorubicin resistance and could increase chemoresistance in such cancer cells.

As reviewed by Delom *et al*. (2020), a mitochondrial protein known as Reptin was also classified as one of the proteins which form interaction with AGR2 protein. Such protein which had been discovered from yeast two-hybrid screen was a member of AAA+

family and found to be involved in multiprotein complexes linked to transcription, nonsense-mediated RNA decay as well as DNA damage response. The protein belonging to the RuvBl1/2 superfamily had ATP binding motifs as among one of the features which allowed it to form interactions with proteins involved in cancer for instance, Tip60, APPL1, Pontin, Myc and telomerase haloenzyme complexes. Both binding motifs as well as AGR2's substrate-binding loop were equally important as determinants that formed AGR2-Reptin complex. In cancer growth on the other hand, Reptin was found to be highly produced in primary breast cancer biopsy specimens and this could be because they could also function as prometastatic transcription protein (Delom *et al*., 2020).

Arumugam *et al.* (2015) in their research reported a receptor known as glycosyl-phosphatidyl-inositol-linked receptor or C4.4A which had formed synergy with eAGR2 protein resulting in chemoresistance, proliferation and migration of pancreatic ductal adenocarcinoma. This could further support eAGR2's capability in stimulating metastasis of tumor via interaction occurring between eAGR2, secreted metastasis-associated GPI – anchored C4.4A protein together with α-dystroglycan extracellular domain which stimulated adhesion of receptors and extracellular matrix interaction. C4.4A in addition was a structural homologue of urokinase-type plasminogen activator receptor or uPAR which was said to be closest to Prod1 human homologue just like CD59. Such Prod1 which was a part of three-finger protein (TFP) superfamily could also form interaction with eAGR2, and this could be seen in limb regeneration of salamander. Anyhow, in the absence of CD59, cells could actually be resistant to eAGR2's function. This further emphasized that the interaction in between eAGR2 and cell surface could regulate adhesion at the same time stimulating tumor cell dissemination.

Dong *et al*. (2015) clarified the importance of interaction occurring between AGR2 and epithelial growth factor receptor (EGFR) before the arriving of receptor at plasma membrane and Golgi apparatus. Such interaction occurred when AGR2 provided an

expression for EGFR signaling which in turn became possible target in treating both chronic as well as neoplastic pancreatic disease. Meanwhile, AGR2's prominence role in pancreatic ductal adenocarcinoma was seen when they were activated downstream in mutant $Kras^{G12D}$ such that deletion of tumor suppressor gene known as *Smad4* occurring during the PanIN-3 or late stage of pancreatic cancer.

Meanwhile, in VEGF and FGF signaling, as stated by Delom *et al*. (2018), eAGR2 which was AGR2 protein residing outside of cancer cells were responsible for cells' aggressiveness. Such extracellular role of AGR2 protein also involved in stimulating angiogenesis and invasion of fibroblast-coordinated tumor cell. AGR2 could bind directly to both VEGF and FGF at the same time increasing activity of tumor angiogenesis. Various signaling routes of eAGR2 protein hence potentially allowed them to be a therapeutic target.

AGR2's role as chaperone protein allowed them to have protein interaction and specific sequence peptide binding activity. Linear motif or functional module made up from small stretches of amino acid basically stored most of the polypeptide information especially in higher eukaryotes. Linear motif generally promoted weak but specific protein-protein docking as mentioned by Tompa *et al*. (2014). Peptide-phage combinatorial library on the other hand was actually the tool used to elaborate linear motif docking repository of target protein. Brychtova *et al*. (2015) reported an appearance of penta-peptide docking motif or TxIYY via such peptide-phage combinatorial library when AGR2 protein was exposed to the combinatorial peptide screen. Such peptide motif was useful for purifying AGR2 protein especially from tissue biopsies as well as useful for activation of p53 tumor suppressor which occurred when AGR2 protein became deactivated through activation of factors involved in suppressing p53. The penta-peptide docking was not affected by the mutation done on AGR2 protein in stabilizing monomer

or dimer isoforms (Gray *et al.*, 2013). This further showed the prominence role of AGR2's sequence-specific peptide binding activity in oncogenic setting.

2.2.3 Secretion of AGR2 protein into the extracellular environment

Despite being in the ER, AGR2 protein however, turned out to be capable of escaping the ER and localized outside for instance on the surface of plasma membrane. Even so, their function as cell surface AGR2 protein remained undetermined. Secretome was basically the term used to express the secreted proteins released into the extracellular microenvironment in which most of them were responsible for many processes such as extracellular matrix remodelling, proteostasis, immune responses, cellular homeostasis as well as developmental processes. Such secretome together with components of extracellular microenvironment were usually capable of developing into cancers (Lu *et al*., 2012).

Research by Fessart *et al*. (2016) proved the occurrence of eAGR2 protein in the extracellular medium of non-tumoral organoids which had the capability of converting non-tumoral organoids to tumor organoids in the human lung epithelial cells. Such lung cancer growth was also promoted through high expression of both intracellular and recombinant AGR2 protein introduced into the cells. The ability of changing cell polarization as well as promoting metastatic activities were also determined in eAGR2 protein.

Conventional secretory pathways which included ER and Golgi apparatus (GA) were the regular route for secretion of protein. Another route known as unconventional protein secretion (UPS) which provided as an alternative for proteins to pass ER – GA path despite not acquiring signal sequence. As for AGR2 protein, the fact that such protein was equipped with highly conserved N-terminal signal peptide (Met1 – Ala20) especially in their variants allowing them to go through classical secretory pathway (SP) and not the

unconventional pathway or non-classical SP. Even so, the duration for AGR2 either to be in ER or outside of ER was unidentified. In addition, it was undetermined whether AGR2 protein could bypass ER – GA and travelled via UPS route (Moidu *et al*., 2020).

UPR which was activated by ER stress had the potential to increase secretion of AGR2, a hypothesis that was yet to be proved. A protein known as USP19 protein bore TxIYY, together with AGR2 binding peptide motif was said to be associated to UPS such that they could activate misfolding-associated protein secretion which confined misfolded cytosolic proteins, a mechanism known for controlling the quality of protein (Lee *et al*., 2016). Hence, AGR2 protein might be capable of functioning via USP19.

The signal peptide sequence in N-terminus of AGR2 which was hydrophobic somehow responsible for directing AGR2 protein into the ER. Such sequence targeted polypeptides from the ribosome to the ER with the aid of signal recognition particle (SRP) before they were being cleaved by the signal peptidase in ER therefore, forming a protein. Basically, full-length AGR2 proteins were positioned in the ER while the mature AGR2 protein without the N-terminal leader sequence could be found in the nucleus. For proper protein folding and maturation, co-translational cleavage should be conducted and usually by signal peptidase from the signal peptide of protein. Any proteins which did not have signal peptide would not enter the ER resulting in subcellular mislocalization. Hence, deleting KTEL could lead to secretion of AGR2 protein into the extracellular media.

Glycosylation process was basically undergone by post-translation secretory proteins that entered the ER. The types of sugar being added into proteins led into their grouping of either N-glycosylated or O-glycosylated. Clarke *et al.* (2015) in their research showed that eAGR2 secreted in human was the one among the O'-glycosylated protein.

2.2.4 AGR2 systematic mutations

The C-terminal ER-retention KTEL motif in AGR2 protein which was also known as H/KDEL was a prominence motif specifically for ER localization. Any other proteins that possessed such motif could block ER export at the same time formed interaction with KDEL receptors located at the intermediate compartment or cis-Golgi for ER retrieval. Meanwhile, variants existed from the KDEL motif were intended for proteins to remain in the ER. This also applied to KTEL motif having reduced binding affinity for KDEL receptor. Such capability of being in other subcellular compartments as said to occur due to the ER-retention motif KTEL (Brychtova *et al.,* 2015; Gupta *et al.,* 2012). Mutagenesis in any motifs of AGR2 protein was usually done to study any effects that it could bring to the cell ecosystem as a whole. AGR2 protein secretion occurred due to deletion of KTEL. Magnitude of AGR2 protein localization in the ER could be enhanced by mutating the KTEL to KDEL (Moidu *et al.,* 2020). For instance, wild type KTEL was capable of expressing amphiregulin and CDX2 instead of KDEL (Gupta *et al.*, 2012). Deleting KTEL somehow could prevent AGR2 protein from stimulating cell growth in clonogenic assays at the same time reducing the p53 protein transcriptional response to DNA damage. Such deletion also resulted in more secretion of AGR2 protein as compared to the wild-type hence crucial for AGR2 protein secretion.

KTEL motif also played prominence role in maintaining AGR2 protein structure as a research by Jia *et al.* (2018) proved that deletion of KTEL mutant caused recombinant AGR2 to be unable to produce. Grassme *et al.* (2016) in another research revealed the released of protein secretion into cell culture media could be increased by adding myc-tag at the nAG C-terminal. Such addition could cause KTEL motif to be hidden resulting in low binding affinity towards KDEL receptor such that protein could not be retained in the cells thus secreted into the media. Hence, epitope-tag was usually fused upstream of the KTEL sequence in most of the plasmid vectors used for AGR2 protein so that it could improve their localization and functional assays (Moidu *et al.,* 2020).

eAGR2 protein which was secreted from the ER was basically a functional protein and was not regarded as products of cell death or lysis. AGR2 protein could still be secreted at the same tendency as the wild-type even when KTEL motif was deleted or mutated wholly (Fessart *et al.,* 2016). Both AGR2 – KTEL as well as AGR2 – KDEL mutants showed similarity in terms of eAGR2 protein secretion levels. This meant that such motif might not be crucial to be used in AGR2 protein secretion and that the AGR2 protein released was a type of soluble protein. AGR2 protein secretion also occurred due to their KTEL – ER retention signal and when Cys was being replaced by Ser at amino acid 81. Such substitution occurring at Cys81Ser hence found to be prominence in managing AGR2 retention in the ER (Bergström *et al.,* 2014). In general, AGR2 protein relied on KTEL retention signal and single Cys81 for ER retention and mutation occurring at Cys to Ser resulted in higher levels of AGR2 secreted.

In terms of structure, AGR2 protein could form dimeric structure such that oxidation-dependent homo-dimerization occurred via Cys81-mediated disulphide bond formation could result in re-orientation of the dimer into different conformation (Ryu *et al.,* 2013). This was further supported by Clarke *et al*. (2016) in which they had proved that mutation of Cys81 to Ala could prevent peroxide from catalyzing AGR2 protein dimerization at the same time proposing role of cysteine for covalent dimer formation. AGR2 cysteine residue (Cys81) could form mixed disulfide bonds with its client protein, mucin. AGR2 – MUC2 complex was the product of the bond formed between Cys81 and N- and C-terminal domains of MUC2 which was rich in cysteine. Such interaction however, diminished with the instigation of Cys81Ser AGR2. AGR2 protein even formed interaction with Cys81 resulting in their homodimerization via intermolecular disulphide bond (Clarke *et al*., 2016; Ryu *et al*., 2013). Bergström *et al*. (2014) in their research proved that protein secretion could be controlled through the occurrence of single cysteine in AGR2 thioredoxin-like domain. Any mutation observed in Cys81Ser or ΔKTEL

contribute to AGR2's secretion in an ovarian cell line, the CHO – K1 cell line hence, emphasized on association of KTEL motif and single Cys in secretion.

Apart from that, dimer could be stabilized through deletion of the first 45 amino acids such that the N-terminal disordered region which was located at amino acid 20 – 40 which was crucial for AGR2 protein dimeric structure. In an attempt to study the association in between dimerization of AGR2 protein mutants and AGR2 protein secretion, the generated monomeric AGR2 E60A mutant resulted in higher rate of secretion as compared to the AGR2 wild-type. The dimeric AGR2 45 mutant was found to be retained in the cells. Both results revealed by the Maurel *et al*. (2019) showed that changing of AGR2 protein monomeric or dimeric form could influence AGR2 protein export into the extracellular environment. As mentioned, AGR2 protein bearing dimerization motif contributed to its dimeric structure. Deletion of N terminal 45 amino acids could enhance the affinity of the dimer as reported by Patel *et al*. (2013). This was due to the fact that N-terminus had negative regulatory role which could reduce dimer affinity. Manipulating the AGR2 protein dimer stability hence could be done because peptides from such N-terminal disordered region controlled the *in trans* stability of the dimer such that altering the stability could actually bring affect towards the metastatic activities of the AGR2 protein. AGR2 protein which in normal condition appeared to be in homodimer turned out to form a complex with ERAD machinery under ER stress which caused the dimer to dissociate in the attempt of isolating misfolded proteins outside of the ER (Maurel *et al*., 2019). Somehow, this event caused activation of pro-inflammatory signals which caused AGR2 to be released into the extracellular environment.

AGR2 protein was not redundant to ER-resident oxidoreductase (ERp18) as a study by Chevet *et al*. (2013) via OMICS platform highlighted only AGR2 was the oncogenic signaling molecule, having clear phenotype which unable to be replaced by AGR3 or ERp18. ERp18 having mutation from CSSC motif to CXXS had the ability of trapping

dithiothreitol-sensitive client proteins. The same mutation however was not suitable to be used on disulphide sensitive AGR2 intermediates. Even so, performing single cysteine mutation in AGR2 protein somehow could weakened mucin interactions as shown in murine suffering from inflammatory diseases. As single cysteine in AGR2 greatly contributed to its disulphide exchange activities therefore, they could be used as a target in gene editing in order to study the thioredoxin activity and the effects that it could bring on the ER homeostasis and ER-associated protein degradation (ERAD) control during both normal and diseased states.

As highlighted in the paper by Brychtova, Mohtar, Vojtesek & Hupp (2015) AGR2 expression tend to elevate when stress occurred in the ER. The wild-type (WT) AGR2 protein bearing the full sequence was able to co-localize together with the ER marker in immunofluorescence with the KDEL mutant having more significant co-localization. Not only deleting the KTEL sequence of AGR2 could promote its secretion somehow, the ER retention motif also responsible for AGR2's escaped into the ER. In ER, AGR2 formed association with the ribosomes via polypeptides.

2.2.5 iAGR2 and eAGR2 in cancer

Overexpression of iAGR2 was often said to be associated with tumor development. Proteomic analysis conducted on ER-bound ribosome revealed that AGR2 protein was bound to the newly synthesized cargo proteins. AGR2 protein was also found to have prominence role in the regulation of ER as well as quality control such that when being chemically-induced by ER stress, the expression of AGR2 in liver cancer cells increased (Moidu *et al*., 2020). This further support that AGR2 protein's expression significantly increased during the occurrence of cancer in order to bear with the high production of proteins to meet secretory demands. High protein demands, inflammatory cytokines and environmental toxins were categorized as the intracellular or extracellular signals that

could disturb ER proteostasis apart from the accumulation of misfolded and unfolded proteins in the lumen of ER. Unfolded protein response or UPR was a mechanism that was activated when the cells in the ER underwent stress due to abundance of protein folding in the ER. Such mechanism could contribute to manifestation of cancer. AGR2 protein expression somehow was also controlled by the UPR via IRE1α and ATF6 arms. In a research by Wang *et al*. (2019), increasing of UPR markers along with IRE1α and ATF6 expression would reduce the AGR2 protein expression during the proteasome inhibition or also known as MG132 treatment. In another research by Dumartin *et al*. (2017), AGR2 protein expression was shown to be increased through application of tunicamycin while high expression of XBP1s which was a form of ER stress marker resulted in silence of AGR2 protein. Genereux *et al*. (2015) in another research showed that secretion of ER chaperone protein known as ERdj3 could activate the UPR. Secreted ERdj3 was usually responsible for the binding of misfolded proteins occurring at the extracellular space, reduce the accumulation of proteins at the same time decreased the proteotoxicity of toxic prion protein. When the ER chaperone machine became saturated, the ERdj3 protein tended to be secreted together with the reduced proteins and this resulted in extracellular chaperoning of the proteotoxic unfolded proteins.

Comparing various types of cancer, breast cancer happened to have poor survival rate as high AGR2 protein expression was detected. Rodriguez *et al.* (2018) successfully showed that AGR2 protein expression together with LOX5 could serve as prognostic markers in the analysis of prostate cancer tissues.

eAGR2 was also found to be useful in cancer diagnoses as well as prognoses through detection of eAGR2 in the patients' body fluids. Wayner *et al*. (2012) for instance showed the use of eAGR2 in prostate cancer as a marker using the ELISA assay that could detect AGR2 protein in the voided urine belonging to the patients. Bu *et al*. (2011) in another research on prostate cancer patients revealed that AGR2 protein transcripts in urine

sediments of patients were detected via RT – PCR. Both researches somehow proved that AGR2 protein was useful enough to be used as cancer marker as their expression pattern almost imitated the prostate-specific antigen (PSA). Hence, AGR2/PSA expression ratio was worth to be considered to differentiate prostate cancer patients effectively. Even so, the mere use of AGR2 protein or PSA was not sufficient enough for identification of cancer samples since AGR2 could also be found in plasma and circulating tumor cells (CTC). AGR2 protein at the same time did not correspond to PSA level. Ho *et al*. (2016) even reported occurrence of AGR2 protein in voided urine of bladder cancer patients. Another research conducted by Edgell *et al*. (2010) also reported detection of AGR2 in serous and non-serous tumors of ovarian cancer such that their expression were higher in stages II and III when detected using ELISA. Early detection of ovarian cancer with an up to a year before diagnosis somehow also possible by using AGR2 protein detection together HE4, CHI3LI, PEBP4 and CA125 cancer markers with an up to 95.4% specificity and 85.7% sensitivity besides using CA125 alone (Whitwell *et al*., 2020).

These days, the development of antibodies against AGR2 protein begun to be conducted. The systematic Evolution of Ligands by Exponential enrichment (SELEX) for instance had constructed an aptamer C14B with the length of 87 nucleotide and capable of binding against AGR2 protein (Wu *et al.*, 2012). Another research by Arumugam *et al.* (2015) showed that the growth of pancreatic cancers together with their metastasis rate could be reduced by murine blocking antibody which targeted AGR2 consequently suppressing AGR2 client protein, C4.4a. Another research by Guo *et al*. (2016) successfully constructed humanized antibody 18A4 which also served AGR2 as a target. Such antibody was found to inhibit xenograft tumor growth. Following this, another study by Guo *et al*. (2017) proved 18A4's capability of detecting eAGR2 such that mixing 18A4 antibody with bevacizumab could inhibit the growth of tumor in ovarian cancer xenograft model. One recent study by Negi *et al*. (2019) proved that the 18A4 antibody could

suppress the growth of lung cancer due to their ability to activate *p53* pathway thus decreasing the cell proliferation rate.

AGR2 protein which could be induced by estrogens played a role as oncogene in hormone-dependent breast cancer and found to be induced by tamoxifen which was a drug that had agonist effects on AGR2 protein Brychtova *et al*. (2015) in their paper even highlighted on AGR2 protein as estrogen agonist in which their expression could result in poor prognosis especially in breast cancer patients. AGR2 protein's expression thus could be used to determine patients' drug resistance.

Gray *et al*. (2014) showed that alterations done on cell line that produced AGR2 protein allowed it to reduce p53 response towards cisplatin as well as enhancing cell proliferation. They were also reported to be capable of providing resistance towards growth inhibiting molecules.

The occurrence of AGR2 protein in prostate cancer was stimulated by androgens and its expression was associated to the growth of tumor during adaptation to the need of protein quality control (Bu *et al*., 2013; Chanda *et al*., 2014; Kani *et al*., 2013). Neeb *et al*. (2014) reported detection of alternative splicing in *agr2* gene of prostate cancer. Such finding contributed to the proposed mechanism on ways of shuffling regulatory domains as well as the affect that it could bring on protein's cell signaling function. Deleting the N-terminal leader sequence for example could decrease ER localization at the same time enhanced stability of dimer. Deleting C-terminal ER- retention site on the other hand contributed to an increase in the protein released from transfection or from isogenic cells (Gray *et al.*, 2014). Such detection of AGR2 in androgen-stimulating cells in serum or urine of patients thus could be used to diagnose prostate cancer.

Although *agr2* gene expression was detected in estrogen or a drug mimicking estrogen called tamoxifen however, certain expressions showed lack of responses especially for EGFR-dependent and AKT signaling pathways as reported by Hrstka *et al*. (2013).

Hormone independent pathways somehow contributed to the expression of *agr2* and *agr3* genes in cancers. A study done by Gray *et al*. (2012) on ovarian cancer showed that *agr3* gene was capable of dominating *agr2* gene. The use of monoclonal antibody on AGR3 protein allowed ovarian cancers to produce both AGR3 and AGR2 proteins. FOXA1 and FOXA2 were the hormone-independent transcription factors responsible for stimulating AGR2 in ovarian systems.

In esophageal adenocarcinoma, the TGF-β /AGR2 axis contributed a role in the onset of cancer. For instance, AGR2 protein as viewed on proteomics screen managed to identify factors that suppressed the p53 tumor suppressor activity in Barrett's esophagus. Any alterations involving TGF-β could suppress AGR2 protein and affect both secretory and protein quality control David *et al*. (2014). Such expression data suggested the importance of quality control protein folding mechanisms in synthesis, folding as well as trafficking of proteins in varieties of cancers.

Other significant role of AGR2 lied in pathophysiological stress in cancer cells. Increasing of AGR2 mRNA via ERK-dependent signaling pathway was said to occur when treating cancer cell lines with hypoxic stress. Other than that, AGR2 mRNA was also induced by treating cancer cells with thapsigargin, tunicamycin and other reducing agents such as ATF6 and IRE1-dependent signaling which stimulated ER stress.

A tumor basically not only composed of cancer cells but also varieties of host cells for instance cancer stem cells, immune cells, extracellular matrix (ECM), endothelial cells as well as cancer-associated fibroblasts (CAFs) in which together they would form a microenvironment known as tumor niche microenvironment (TNM). The surrounding cells contributed to the advancement of tumor by supplying necessary elements hence making cancer as part of systemic disease instead of solely genetic or cellular disease. Such TNM was crucial for both cellular communications as well as the progression of tumor itself. Any changes occurring in the TNM could bring effects to the TNS. Firstly,

any alterations occurring in the ECM components or expression of matrix metalloproteases could result in destabilization of the TNM thus causing an elevated cancer invasion (Leung & Brugge, 2012). Secondly, migration, proliferation as well as tumor angiogenesis could be stimulated by TNS through modulation of messengers for instance like chemokines and cytokines as well as adhesion protein such as membrane receptors and integrins (Paltridge *et al*., 2013). Thirdly, any changes detected in the TNM could influence the immune response of the TNM (Paltridge *et al*., 2013) at the same time modulating tumor immunogenicity (Meeusen *et al*., 2017). AGR2 tumor niche secretome on the other hand was the disruption of the TNM by AGR2 which resulted in modulating proteins such as extracellular matrix proteins, matrix metalloproteinases, inflammatory factors, growth factors, integrins, immunogenic factors as well as pro-angiogenic factors (Delom *et al*., 2018).

CHAPTER 3: METHODOLOGY

Objective 1: To construct vectors for the expression of recombinant AGR2 protein variants in mammalian cells.

3.1 Plasmid cloning

The open frame region of AGR2 protein basically consisted of 175 amino acids, including a signal peptide at 1 – 20 amino acids and a mature peptide at the subsequent 21 – 175. The nucleotide sequence was obtained from the Ensembl.org (Gene accession: NP_006399.1) and viewed using SnapGene viewer. The complementary DNA (cDNA) of four variants of AGR2 protein, AGR2 – WT, AGR2 – KDEL, mAGR2 – ΔKTEL and ΔNterm AGR2 – KDEL insequence was as shown in Figure 3.1 and chemically synthesized using Gene Art Gene synthesis (Thermo Fisher Scientific). The 5' and 3' ends of each variant contain *NcoI* and *XbaI* restriction sites respectively and cloned into pMA-T holding vectors. The resulting plasmids were then sub-cloned into the multiple cloning sites containing the same restriction sites in pSF-CMV mammalian expression plasmids.

A

AGR2 – WT

5′ -ATG GAG AAA ATT CCA GTG TCA GCA TTC TTG CTC CTT GTG GCC CTC TCC TAC ACT
CTG GCC AGA GAT ACC ACA GTC AAA CCT GGA GCC AAA AAG GAC ACA AAG GAC TCT CGA
CCC AAA CTG CCC CAG ACC CTC TCC AGA GGT TGG GGT GAC CAA CTC ATC TGG ACT CAG
ACA TAT GAA GAA GCT CTA TAT AAA TCC AAG ACA AGC AAC AAA CCC TTG ATG ATT ATT
CAT CAC TTG GAT GAG TGC CCA CAC AGT CAA GCT TTA AAG AAA GTG TTT GCT GAA AAT
AAA GAA ATC CAG AAA TTG GCA GAG CAG TTT GTC CTC CTC AAT CTG GTT TAT GAA ACA
ACT GAC AAA CAC CTT TCT CCT GAT GGC CAG TAT GTC CCC AGG ATT ATG TTT GTT GAC
CCA TCT CTG ACA GTT AGA GCC GAT ATC ACT GGA AGA TAT TCA AAT CGT CTC TAT GCT
TAC GAA CCT GCA GAT ACA GCT CTG TTG CTT GAC AAC ATG AAG AAA GCT CTC AAG TTG
CTG AAG ACT GAA TTG TAA- 3′

Figure 3.1: cDNA sequence of AGR2 protein variants. The sequence in blue indicated substitution of nucleotide. (A) AGR2 – WT sequence. (B) AGR2 – KDEL sequence. (C) mAGR2 – ΔKTEL sequence. (D) ΔNterm AGR2 – KDEL sequence (Taken from Mohtar, 2017).

B

AGR2 – KDEL

5′ -ATG GAG AAA ATT CCA GTG TCA GCA TTC TTG CTC CTT GTG GCC CTC TCC TAC ACT
CTG GCC AGA GAT ACC ACA GTC AAA CCT GGA GCC AAA AAG GAC ACA AAG GAC TCT CGA
CCC AAA CTG CCC CAG ACC CTC TCC AGA GGT TGG GGT GAC CAA CTC ATC TGG ACT CAG
ACA TAT GAA GAA GCT CTA TAT AAA TCC AAG ACA AGC AAC AAA CCC TTG ATG ATT ATT
CAT CAC TTG GAT GAG TGC CCA CAC AGT CAA GCT TTA AAG AAA GTG TTT GCT GAA AAT
AAA GAA ATC CAG AAA TTG GCA GAG CAG TTT GTC CTC CTC AAT CTG GTT TAT GAA ACA
ACT GAC AAA CAC CTT TCT CCT GAT GGC CAG TAT GTC CCC AGG ATT ATG TTT GTT GAC
CCA TCT CTG ACA GTT AGA GCC GAT ATC ACT GGA AGA TAT TCA AAT CGT CTC TAT GCT
TAC GAA CCT GCA GAT ACA GCT CTG TTG CTT GAC AAC ATG AAG AAA GCT CTC AAG TTG
CTG AAG GAT GAA TTG TAA- 3′

C

mAGR2 – ΔKTEL

5′ -GAT ACC ACA GTC AAA CCT GGA GCC AAA AAG GAC ACA AAG GAC TCT CGA CCC AAA
CTG CCC CAG ACC CTC TCC AGA GGT TGG GGT GAC CAA CTC ATC TGG ACT CAG ACA TAT
GAA GAA GCT CTA TAT AAA TCC AAG ACA AGC AAC AAA CCC TTG ATG ATT ATT CAT CAC
TTG GAT GAG TGC CCA CAC AGT CAA GCT TTA AAG AAA GTG TTT GCT GAA AAT AAA GAA
ATC CAG AAA TTG GCA GAG CAG TTT GTC CTC CTC AAT CTG GTT TAT GAA ACA ACT GAC
AAA CAC CTT TCT CCT GAT GGC CAG TAT GTC CCC AGG ATT ATG TTT GTT GAC CCA TCT
CTG ACA GTT AGA GCC GAT ATC ACT GGA AGA TAT TCA AAT CGT CTC TAT GCT TAC GAA
CCT GCA GAT ACA GCT CTG TTG CTT GAC AAC ATG AAG AAA GCT CTC AAG TTG CTG TAA
- 3′

D

ΔNterm AGR2 – KDEL

5′ -ATG GAG AAA ATT CCA GTG TCA GCA TTC TTG CTC CTT GTG GCC CTC TCC TAC ACT
CTG GCC CCC CAG ACC CTC TCC AGA GGT TGG GGT GAC CAA CTC ATC TGG ACT CAG ACA
TAT GAA GAA GCT CTA TAT AAA TCC AAG ACA AGC AAC AAA CCC TTG ATG ATT ATT CAT
CAC TTG GAT GAG TGC CCA CAC AGT CAA GCT TTA AAG AAA GTG TTT GCT GAA AAT AAA
GAA ATC CAG AAA TTG GCA GAG CAG TTT GTC CTC CTC AAT CTG GTT TAT GAA ACA ACT
GAC AAA CAC CTT TCT CCT GAT GGC CAG TAT GTC CCC AGG ATT ATG TTT GTT GAC CCA
TCT CTG ACA GTT AGA GCC GAT ATC ACT GGA AGA TAT TCA AAT CGT CTC TAT GCT TAC
GAA CCT GCA GAT ACA GCT CTG TTG CTT GAC AAC ATG AAG AAA GCT CTC AAG TTG CTG
AAG GAT GAA TTG TAA- 3′

Figure 3.1, continued.

3.1.1 Digestion

To clone the AGR2 protein variants into pSF-CMV vector, both the pSF-CMV vector as well as the AGR2 protein variants cloned into holding vector, pMA-T vector were subjected to double digestion using restriction enzymes and compatible buffers as suggested in the manufacturer's guideline. The holding vector, pMA-T was digested at the *Nco1* and *XbaI* restriction sites which hold the four protein variants, AGR2 – WT, AGR2 – KDEL, mAGR2 – ΔKTEL and ΔNterm AGR2 – KDEL. Digestion set up was prepared as in Table 3.1 and incubated for 2 – 3 hours at 37^{O}C (Memmert). To isolate the inserts from its remaining and restriction enzymes as well as purifying the digested vector

from residual nicked and supercoiled vector DNA, gel purification was conducted by mixing the product of digestion with DNA loading buffer and electrophosed them on 1% agarose gel for 40 – 60 minutes. The single bands from the double digested vector as well as inserts were excised under the UV light. DNA extraction was conducted, and the concentration of recovered DNA was determined.

Table 3.1: DNA digestion reaction.

Components	Final concentration
Inserts (variants) / vector	1.5 – 2.0 µg
Restriction buffer	10x
Restriction enzyme 1	10 units (1µl)
Restriction enzyme 2	10 units (1µl)
	Total volume = 20 µl

3.1.2 DNA extraction

To examine the purity and concentration of inserts and vectors, DNA extraction was first conducted using QIAamp DNA mini kit (Qiagen). The excised vector as well as inserts were placed in 1.5 ml microcentrifuge tube each. A volume of 180 µl buffer ATL and 20 µl proteinase K were added and mixed well by vortexing occasionally during incubation at 56^{O}C for 1 – 3 hours until complete lysis had been achieved. Then, 200 µl of buffer AL was added and mixed thoroughly by vortexing at 15 seconds. Mixture was incubated at 70^{O}C for 10 minutes before centrifuged (Eppendorf) briefly as to remove any drops coming from the lid. This was followed by addition of 200 µl of absolute ethanol which would then vortex for another 15 seconds. Mixture was centrifuged briefly again to remove any drops from the lid. The mixture was then pipetted into QIAamp DNA mini spin column into 2 ml collection tube and added with 500 µl buffer AW1. Mixture

was centrifuged at 14,000 rpm for 3 minutes and the collection tube was discarded. The QIAamp DNA mini spin column was placed in a 1.5 ml microcentrifuge tube and 200 µl of buffer AE was added before incubated at room temperature for a minute. Mixture was centrifuged at 8,000 rpm for a minute to elute the DNA. About 1 µl or a drop of the DNA was checked for its purity and concentration using the NanoDrop 2000 spectrophotometer (Thermo Scientific).

3.1.3 Ligation

Ligation of the digested inserts and vector extracted previously was done using T4 DNA ligase buffer (NEB). The molar ratio of vector to insert was in between 1:1 to 1:4. Amount of insert added was calculated using the formula below:

$$\textit{Amount of insert (ng)} = \frac{\textbf{Amount of vector (ng) X Size of insert (kb)}}{\textbf{Size of vector (kb)}} x \quad \textit{Molar ratio of } \frac{\textbf{Insert}}{\textbf{Vector}} \tag{3.1}$$

Ligation reaction was prepared as in Table 3.2 and conducted on ice before incubated at 37^{O}C water bath for 3 hours or left at room temperature for overnight.

Table 3.2: DNA ligation reaction.

Components	Final concentration
Vector	100 ng
Insert	As calculated
Buffer (NEB)	1x
T4 DNA ligase (NEB)	1µl

3.2 Transformation

3.2.1 Bacterial growth culture

For miniprep or small DNA preparation, a single colony of DH5α bacteria cells grown on a Luria-Bertani (LB) agar (Sigma-Aldrich) plate was picked and grown in 5 ml LB medium (Sigma-Aldrich) containing ampicillin (Sigma) and incubated overnight at 37^{O}C on a shaker set at 200 rpm (New Brunswick Scientific). From the bacterial culture prepared, dilution from 1/500 to 1/1000 was conducted and transferred into 200 ml medium containing ampicillin (100 μg/ml) and incubated at 37^{O}C on a shaker set at 200 rpm for 12 – 16 hours.

Table 3.3: Compositions of LB agar and broth for bacterial culture.

LB agar		**LB medium**	
10 g	Tryptone	10 g	Tryptone
10 g	NaCl	10 g	NaCl
5 g	Yeast extract	5 g	Yeast extract
	Up to 1 liter of distilled water	15 g	Agar per liter of distilled water

Table 3.4: Antibiotic used for growing of bacteria.

Antibiotic	**Stock concentration**	**Storage**	**Working dilutions**
Ampicillin (Sigma)	100 mg/ml	-20^{O}C	100 μg/ml

3.2.2 Preparation of DH5α competent cells

A single colony of DH5α competent cells grown on an agar containing ampicillin was picked and grown in 5 ml LB medium containing appropriate antibiotic and incubated

overnight at 37^{O}C on a shaker set at 200 rpm. A volume of 250 µl of bacterial suspension was inoculated into 100 ml LB broth and incubated at 37^{O}C on a shaker set at 200 rpm until the suspension reached 0.4 at OD_{600nm}. The subsequence steps were conducted on ice. Cells were then centrifuged at 4^{O}C with 4,000 rpm for 15 minutes. Pellet was resuspended with 32 ml cold competent buffer I before incubated on ice for 10 minutes. Cells were again centrifuged at 4^{O}C with 4,000 rpm for 15 minutes and pellet was resuspended with 4 ml cold competent buffer II. A volume of 50 µl was then transferred into a microcentrifuge tube that had been stored at -80^{O}C beforehand. They were then kept at -80^{O}C before transferring them into liquid nitrogen for long-term storage.

3.2.3 Transformation of DH5α competent cells

DH5α competent cells were thawed on ice for 10 minutes. About 1 – 5 µl of 100 ng plasmid DNA was added into the cells. Tube was flicked for a couple of times to ensure a thorough mix before incubated on ice for 30 minutes. Mixture was then heat-shocked at 42^{O}C for 10 seconds and then placed on ice for 5 minutes. A volume of 950 µl of LB medium was pipetted into the tube and incubated at 37^{O}C on a shaker set at 200 rpm for an hour. Cells were mixed well by flicking and inverting and ten-serial fold dilution was done using the same medium. Then, 50 – 100 µl from each dilution was pipetted into a pre-warmed LB agar plate and left to grow by incubated invertedly for overnight at 37^{O}C.

3.2.4 Purification and quantification of plasmid DNA

For small-scale or miniprep plasmid DNA purification, DH5α bacteria was grown as described in 3.2.1 and harvested by centrifugation at 25^{O}C for 3 minutes at 8,000 rpm. Supernatant was then discarded. QIAprep Spin Miniprep Kit (Qiagen) was used to purify the plasmid DNA. The bacterial cells pellet was resuspended in 250 µl of P1 buffer before transferred to a microcentrifuge tube. A volume of 250 µl of P2 buffer was added and

mixed well by inverting the tube for 4 – 6 times. This was followed by addition of 350 µl of N3 buffer and mixtures were mixed well by inverting the tube again for 4 – 6 times. Mixtures were then centrifuged at room temperature (25 ^{O}C) for 10 minutes at 13,000 rpm until pellet was formed. Then, 800 µl of supernatant was added into the QIAprep 2.0 spin column and centrifuged briefly for 30 – 60 seconds before discarding the flow. The spin column was then washed with 0.5 ml of PB buffer and centrifuged for 30 – 60 seconds. Flow was discarded. The spin column was washed with 0.75 ml of PE buffer and centrifuged for 30 – 60 seconds. Flow was discarded and column was once again centrifuged at full speed for 60 seconds. Column was then transferred to a 1.5 ml microcentrifuge tube and elution of DNA was conducted by adding 50 µl of EB buffer at the center of the column and left to stand for 60 seconds before centrifuged for another 60 seconds. The purified plasmid was then validated by sending them for sequencing.

Objective 2: To determine the level of intracellular and extracellular expression of AGR2 protein variants expressed from the esophageal cancer cells.

3.3 Cell culture

All cell culture works were performed under Biosafety Cabinet Class II using standard aseptic tissue culture procedures. Only molecular biology grade reagents and sterile consumables were used. Filtering and autoclaving of selected reagents and consumables were done for sterilization purposes.

3.3.1 Recovery of cells

Esophageal adenocarcinoma cancer cell line (FLO-1) used in this research was obtained from the UKM Medical Molecular Biology Institute (UMBI). The FLO-1 cell line stored in cryotube in liquid nitrogen was recovered by thawing them at 37^{O}C water

bath. The tube was gently swirled in loop motion in the water bath until only small amount of ice was left and then transferred into 100 mm cell culture plate containing 9 ml media (Dulbecco's Modified Eagle Medium (DMEM)) (Nacalai Tesque) before incubated at 37^{O}C with 5% carbon dioxide (CO_2) gas supplied (Eppendorf). The 500 ml – DMEM was supplemented beforehand with 10% fetal bovine serum (FBS) (Sigma) together with 1% penicillin and streptomycin (Penstrep) (Nacalai Tesque). Changing of media was conducted 24 hours post incubation to remove any excess of DMSO. The cell line was then grown and maintained as described in 3.3.2.

3.3.2 Cell line and maintenance

The FLO-1 cell line grown in the humidified incubator at 37^{O}C with continuous supplied of 5% CO_2 gas was maintained in 6-cm cell culture dish followed by subculturing for 1 – 3 times per week depending on the degree of confluency and assay purposes. Splitting of cells was conducted after the cells had reached 80 – 90% confluency rate. Screening of cells under the inverted microscope (Nikon) were done before splitting as to ensure that the cells were not subjected to any contaminations. This was followed by washing the cells with sterile 1x PBS (Nacalai Tesque) before sub-culturing as to remove any residues from the previous media. They were then trypsinized through addition of prewarmed 0.5 – 1 ml of 2.5 g/l Trypsin 1 mmol/l – EDTA (Nacalai Tesque). During trypsinization process, FLO-1 cell was incubated at 37^{O}C with 5% CO_2 supplied for 3 – 5 minutes to allow them to fully detach from the cell culture dish. Cell counting was done in which trypsinized cells were mixed with 0.4% Tryphan Blue Stain (Labtech) at the ratio of 1:1 with the purpose of staining any dead cells and consequently loading them into Luna cell counting slide (Labtech) and counted using Luna Automated BF Cell Counter (Labtech). After counting of cells, they were then passage into new cell culture

dish at the ratio of 1:5 to 1:20 trypsin to media, depending on the size of cell culture dish used or desired rate of confluency achieved by the cells.

3.3.3 Storage of cells

Storage of cells was usually conducted after the cells grown in cell culture dish had reached 90% confluency. Cells were first trypsinized with 1 ml trypsin-EDTA before diluting to a final volume of 10 ml. Cells were then transferred to a 15 ml falcon tube and centrifuged at 10,000 rpm for 5 minutes to form cell pellet. The cell pellet was resuspended in a freezing medium composed of 50% FBS, 20% DMSO and 30% serum-free media. This was followed by aliquoting the cells to 1 ml before transferring them into cryotubes which later will be placed into a gradient freezing container (Nalgene) containing isopropanol for 24 hours at -80^{O}C before transferring them into liquid nitrogen for long-term storage.

A

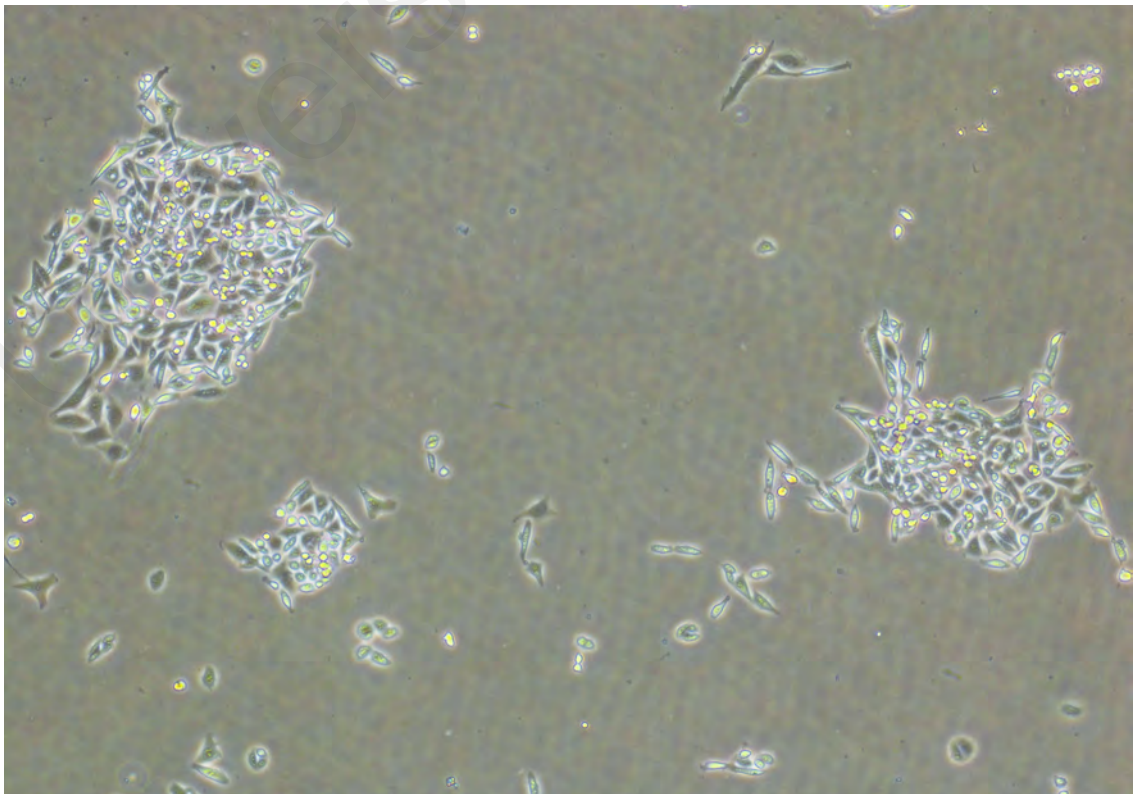
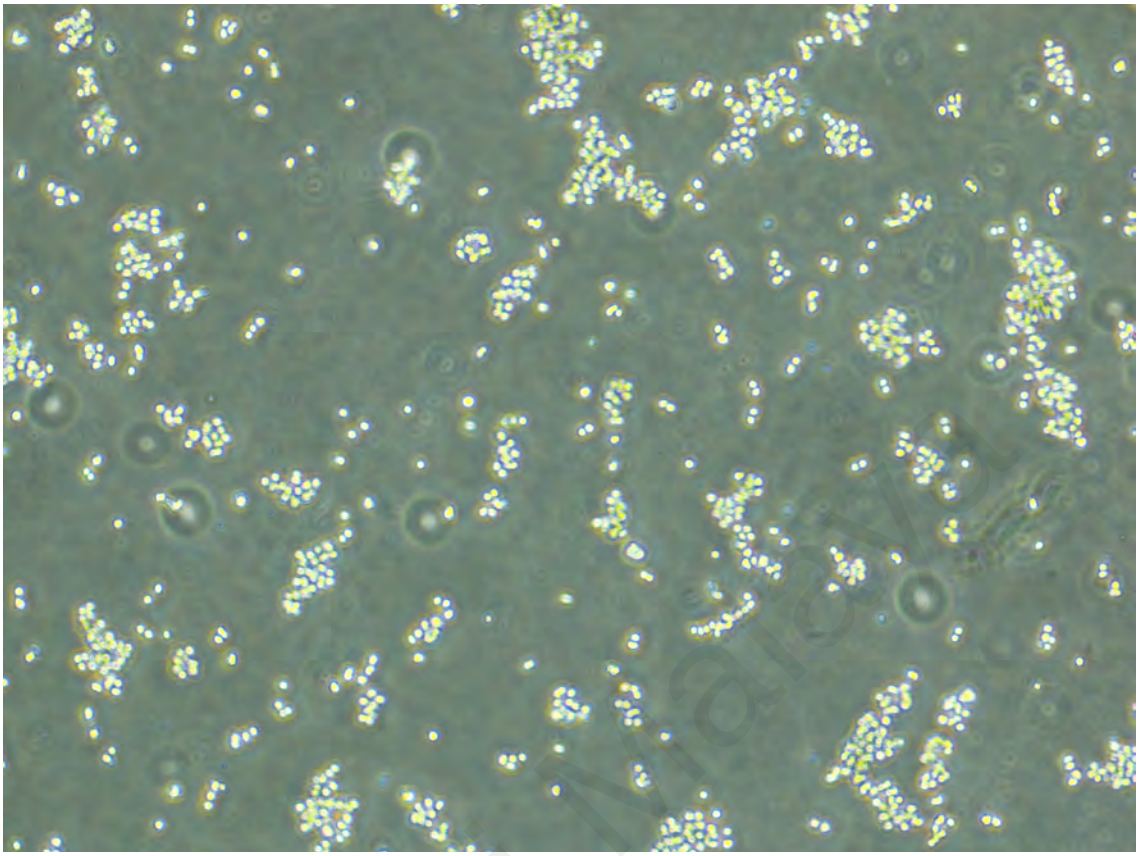


Figure 3.2: FLO-1 cell line. (A) Post thawing. (B) Trypsinized cells. (C) Confluent cells at 4x magnification. (D) Confluent cells at 10x magnification.

B



C

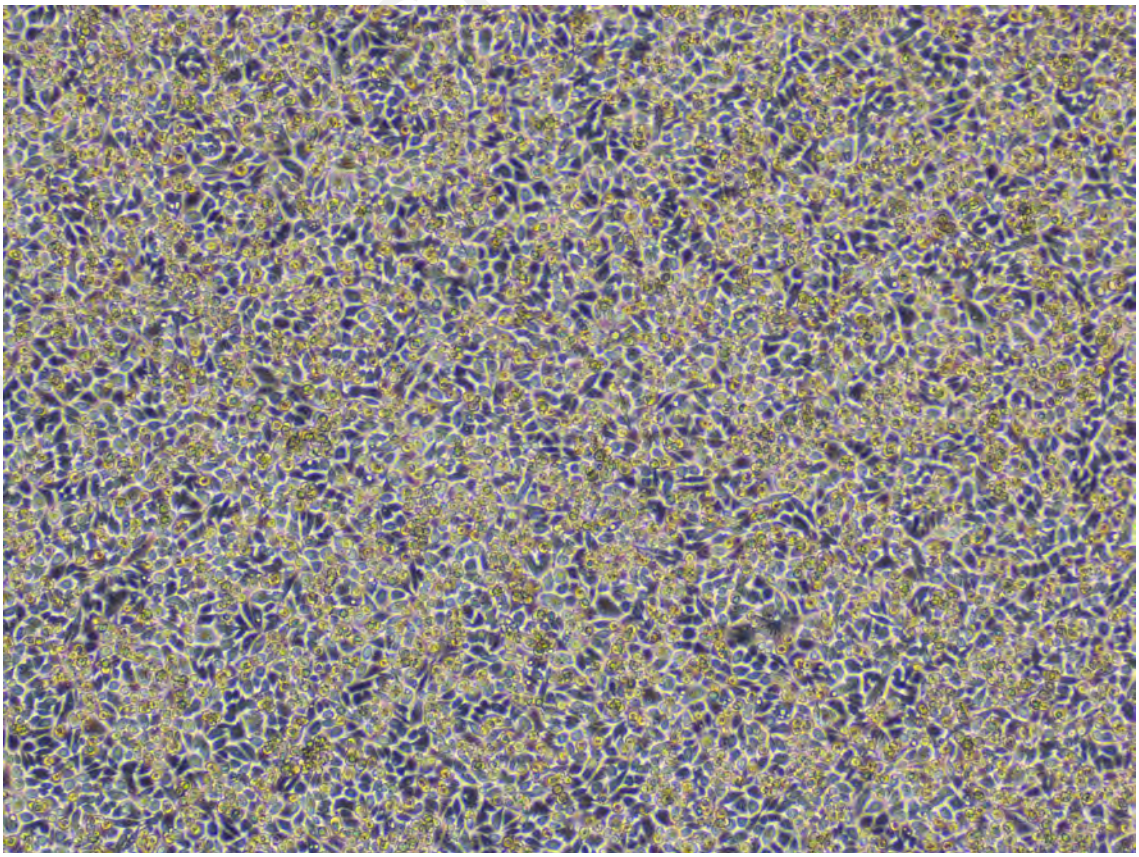


Figure 3.2, continued.

D

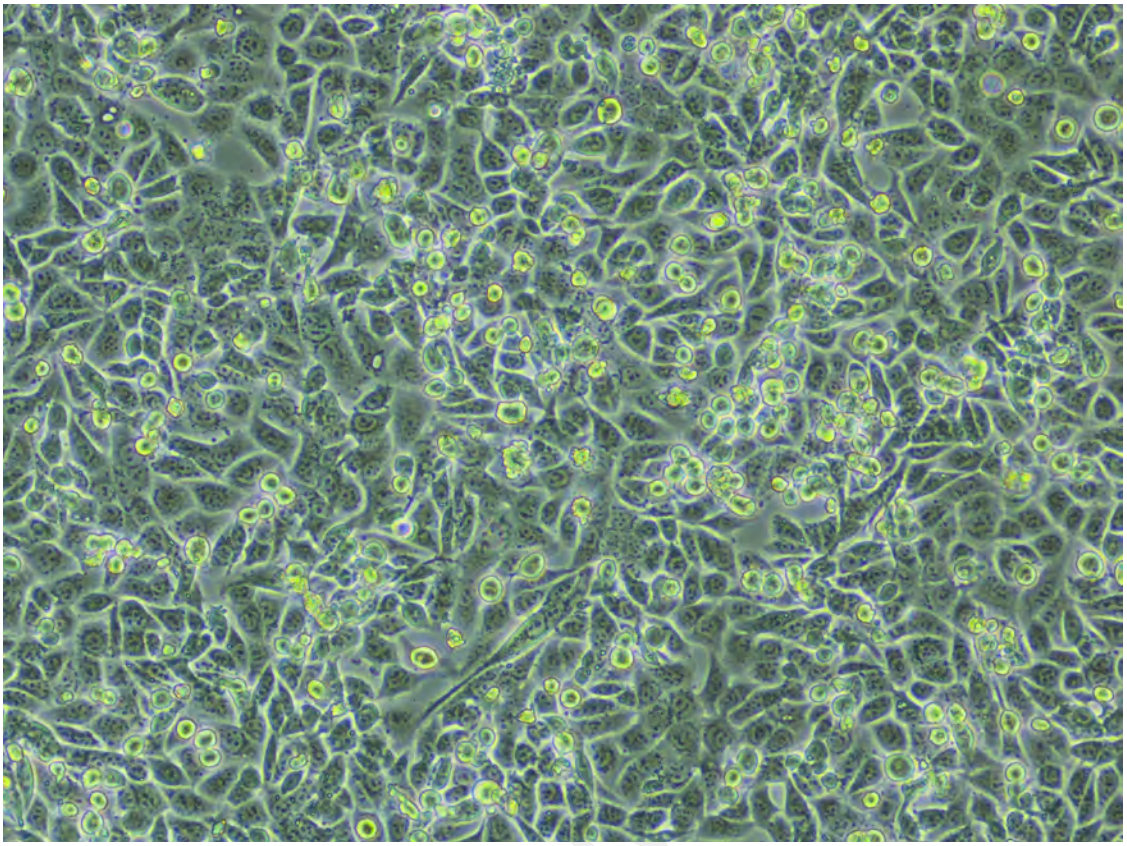


Figure 3.2, continued.

3.3.4 Transient transfection

To examine intracellular and extracellular expression of AGR2 protein variants in FLO-1 cell line, the cells were first grown to a confluency of 80 – 90% before seeding into 1 x 6 well – plate and incubated at 37^{O}C with 5% CO_2 supplied until 70% confluency was achieved on the following day. Changing of media was done before transfecting FLO-1 cell with AGR2 protein variants using Attractene transfection reagent (Qiagen). Mixture of 1 µg DNA, 100 µl of 1x Opti-MEM (Gibco), a type of reduced serum medium or alternatively serum-free DMEM and 5 µl of Attractene were initially prepared in a sterile microcentrifuge tube and mixed well before incubated at 25^{O}C for 15 minutes. Volume of AGR2 protein variants pipetted was as in Table 3.5. Such mixtures were then pipetted dropwise into the FLO-1 cancer cell line and incubated at 37^{O}C and 5% CO_2 for 24 – 48 hours. Successful transfection would be determined via western blot.

Table 3.5: Volume of AGR2 protein variants per 1 µg.

AGR2 protein variants	Concentration of variants (ng/µl)	Volume per 1µg (µl)
AGR2 – WT	696	1.4
AGR2 – KDEL	660	1.4
mAGR2 – ΔKTEL	730	1.5
ΔNterm AGR2 – KDEL	695	1.4

3.3.5 Harvesting

The FLO-1 cell line which had undergone transfection with AGR2 protein variants and achieved 80 – 90% confluency was considered suitable for harvesting. Before harvesting, the existing media was aspirated and each well containing the transfected cell line was rinsed with cold 1x sterile PBS of similar volume as the previous media. The 6-well plate containing transfected cells were positioned on ice before 1 ml of cold 1x PBS was pipetted into the first well and scraping of cells was conducted using a cell scraper. The plate was observed under the inverted microscope to confirm for fully detachment. Cell lysates were then transferred to a labelled 1.5 ml microcentrifuge tube. The process was repeated for the rest of the wells containing AGR2 protein variants. The cell lysates were then centrifuged at 3,000 rpm for 5 minutes at 4^{O}C. Supernatant was then discarded, and cell pellet was lysed using suitable lysis buffer as further described in 3.3.6 or stored at -80^{O}C until further use.

3.3.6 Cellular lysis and protein extraction

Cell pellet was removed from -80^{O}C and thawed on ice. Urea lysis buffer was pipetted into the cell pellet at a volume of approximately 40 – 50 µl or 4 times the size of the pellet and mixed well. Samples were incubated on ice for 20 – 30 minutes before centrifuged at

10,000 rpm for 15 minutes at 4^{O}C. Supernatant was then collected into a new microcentrifuge tube on ice. Cell lysates were then proceeded for protein concentration determination via Bioinchoninic (BCA) assay as described in 3.4.1 or stored at -80^{O}C until further use.

Table 3.6: Cell lysis buffer.

Urea lysis buffer	
7 M	Urea
0.1 M	DTT
0.1%	Triton X-100
25 mM	NaCl
20 mM pH 7.6	HEPES – KOH
1 mM	Benzamidine
7 M	Urea
1x	Protease inhibitor mix

Table 3.7: Protease inhibitor mix.

10x Protease inhibitor mix	
200 µg/ml	Leupeptin
10 µg/ml	Aprotinin
20 µg/ml	Pepstatin
10 mM	Benzamidine
100 µg/ml	Soybean trypsin inhibitor
20 mM	Pefabloc
10mM	EDTA

3.3.7 Media concentration

To examine AGR2 protein extracellular expression, FLO-1 cell were first grown into 60 – 70% confluency, transfected with the protein variants as described in 3.3.4 and incubated for 24 – 48 hours until 80 – 90% confluency was achieved. Media from the transfected cells were then collected into 15 ml falcon tube and mixed with acetone at the ratio of 1:4 media collected to acetone. The acetone was cold beforehand in -20^{O}C for 2 hours. Media and acetone were then kept at -20^{O}C for 24 hours or overnight and consequently centrifuged at 13,000 rcf for 10 minutes at 4^{O}C. Supernatant was discarded and 200 – 1000 µl of pre-warmed media was added depending on the pellet size and resuspended well until it dissolved in the media. The concentrated media was then transferred to microcentrifuge tube and stored at -80^{O}C until further use.

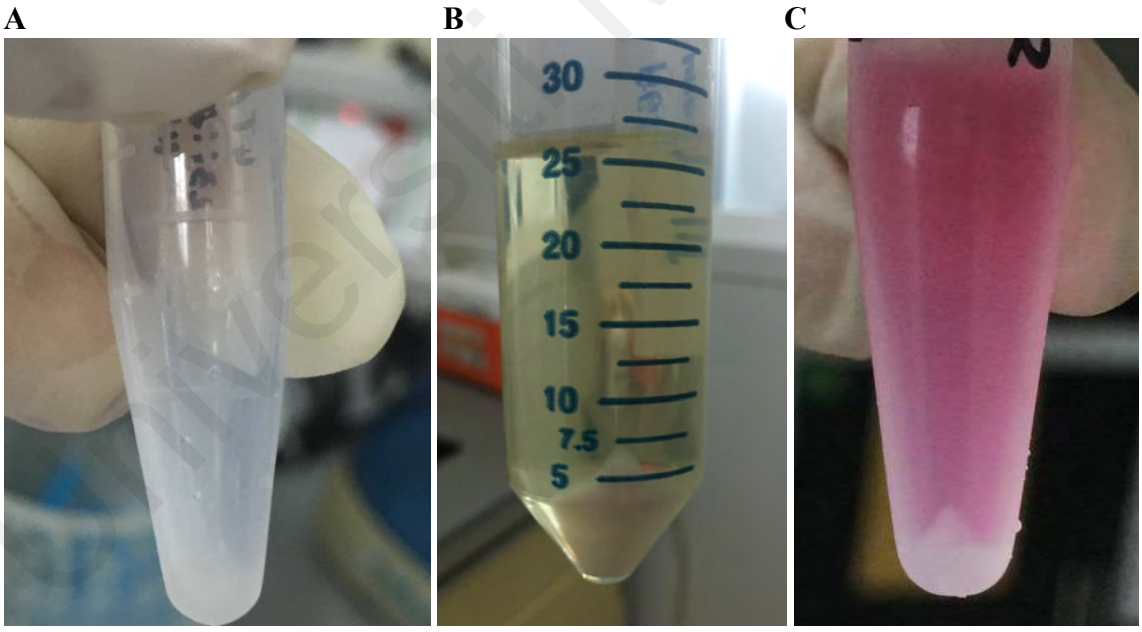


Figure 3.3: Transfecting and harvesting. (A) Transfected cell pellet obtained from harvesting. (B) Pellet from transfected media and acetone. (C) Concentrated media re-suspended in DMEM.

3.4 Quantification and expression validation

3.4.1 Bioinchoninic (BCA) protein quantification assay

BCA protein quantification assay was conducted using a kit (Thermo Scientific). BCA standards, A, B, C, D, E, F, G, H and I were prepared through serial dilution of 2 mg/ml Bovine Serum Albumin (BSA) as in Table 3.8. BCA reagent was prepared at a ratio of 1:50 reagent A to reagent B and mixed well using vortex. The volume of BCA reagent to be prepared depended on the total number of standards (A – I) as well as samples with volume of 200 µl BCA reagent needed per well. Standards and samples were then mixed with BCA reagent in a 96-well plate as in Table 3.9 and incubated at 37^{O}C for 30 minutes. Protein concentrations were then determined by placing the plate in microplate reader (Varioskan Flash Thermo Scientific) and measured at OD_{562nm}. From the absorbance obtained, graph of standard linear curve was plotted and R^2 value as well as equation were obtained. Concentration of transfected AGR2 protein variants were determined by subtracting the absorbance values with blank absorbance and substituting them into the equation obtained from the graph. To prepare the samples for western blot, the concentration of transfected AGR2 protein variants obtained would be multiplied by the dilution factor based on the BCA assay conducted and suitable volume of samples buffer needed was determined. Mass of each transfected AGR2 protein variants was then determined through division with the dilution factor. Volume of water needed to dilute the samples was finalized from the suitable total volume of samples and sample buffer.

Table 3.8: Serial dilution of BSA for BCA standards.

Vial	Volume of diluent (µl)	Volume and source of BSA (µl)	Final BSA concentration (µg/ml)
A	0	300 of stock	2000

Table 3.8, continued.

Vial	Volume of diluent (µl)	Volume and source of BSA (µl)	Final BSA concentration (µg/ml)
B	125	375 of stock	1500
C	325	325 of stock	1000
D	175	175 of vial B dilution	750
E	325	325 of vial C dilution	500
F	325	325 of vial E dilution	250
G	325	325 of vial F dilution	125
H	400	100 of vial G dilution	25
I	400	0	0 (Blank)

Table 3.9: Volume of samples and standards prepared for BCA assay.

Standards (A – I)	Intracellular AGR2 protein variants	Extracellular AGR2 protein variants
200 µl BCA reagent	200 µl BCA reagent	200 µl BCA reagent
25 µl standard	23 µl sample	25 µl sample
	2 µl miliQ water	

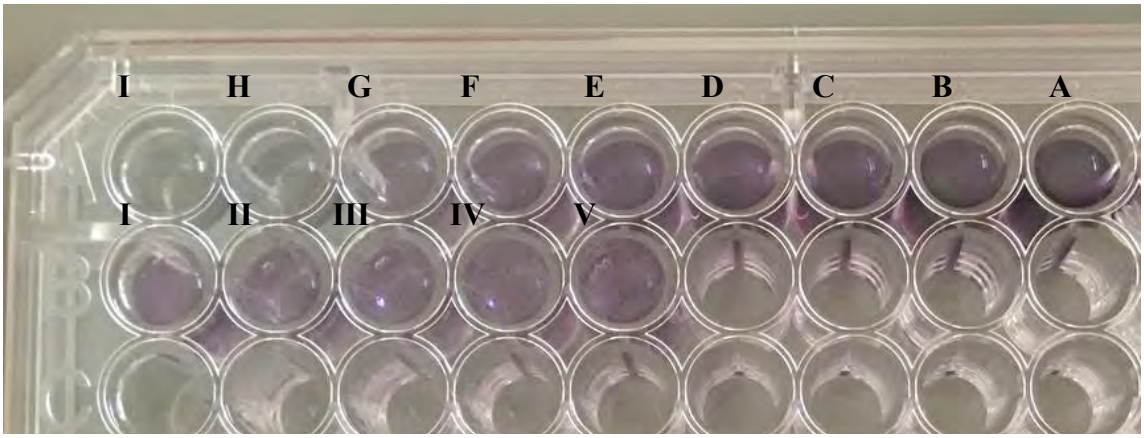


Figure 3.4: BCA assay.

3.4.2 SDS – PAGE

Two sets of polyacrylamide gels consisting of stacking and resolving gels were prepared in the laboratory as described in Table 3.10. Gels were cast using Mini-PROTEAN Tetra Cell apparatus (BioRad) using 1.5 mm gap in between glass plates with respect to the volume of loading gel in order to allow for polymerization at room temperature. Resolving gel was prepared and left to polymerize before addition of stacking gel onto it. A 10-well or 15-well comb was chosen depending on the total number of samples to be run and inserted into the stacking gel and left to polymerize to form the wells for loading the gels.

Before conducting SDS – PAGE, protein lysate obtained from 3.3.6 was adjusted until 20 – 50 µg and mixed with 4x Laemmli SDS sample buffer (GeneTex) at the ratio of 1:1 sample to sample buffer. The 4x SDS sample buffer was prepared as in Table 3.11. Note that the maximum volume of gel (mixture of samples and sample buffer) loaded into each well depended on the size of the well and for 1.5 mm glass gap with 1 x 10-well, the recommended loading gel prepared should not exceeded 40 µl with maximum protein load of 50 µg per band. For extracellular expression, post-transfection concentrated media was mixed with 4x sample buffer at ratio of 1:1. Mixture of samples and samples' buffer were then centrifuged and placed on thermal block at 95^{O}C for 5 minutes to remove any untargeted proteins. They were then left to cool and centrifuged before loaded into the wells. Then, 5 µl of protein ladder (SMOBIO) were loaded as marker.

Table 3.10: Volume of reagents needed for the preparation of two gels comprising of 20 ml of 12% resolving gel and 5 ml of 5% stacking gel.

Reagents	12% Resolving gel (ml)	5% Stacking gel (ml)
Distilled water	6.6	3.4
30% acrylamide mix	8.0	0.83

Table 3.10, continued.

Reagents	12% Resolving gel (ml)	5% Stacking gel (ml)
1.0 M Tris (pH 8.8)	5.0	-
1.0 M Tris (pH 6.8)	-	0.63
10% SDS	0.2	0.05
10% Ammonium persulfate	0.2	0.05
TEMED	0.008	0.005

Table 3.11: SDS sample, running and transfer buffers.

4x SDS sample buffer	10x SDS running buffer		1x SDS transfer buffer	
250 mM Tris – HCl (pH 6.8)	30 g 25 mM	Tris base	30 g 25 mM	Tris base
8% SDS	144 g 192 mM	Glycine	144 g 192 mM	Glycine
40% glycerol	10 g 0.1%	SDS	20%	Methanol
Bromophenol blue		-		-
8% β – mercapthoethanol (added prior to use)		-		-

The set up for western blot was prepared before loading the gels and ladder into the wells. The 1x SDS running buffer was prepared and diluted from 10x SDS running buffer as in Table 3.11 and poured into the buffer tank. The running apparatus was closed with an electrode lid connected to an external DC voltage power supply via anode and cathode probes before switched on. Proteins were then separated by electrophoresis at 80 – 120V or until all bands were separated at stacking gel (top gel) and increased to 130 – 150V once they had entered resolving gel (bottom gel). The SDS – PAGE was run for a total of 60 – 90 minutes or until all bands had reached the bottom of the gel and that ladder had separated.

3.4.3 Western blot

Traditional wet transferring process or wet electroblotting was done by preparing blotting sandwich comprising of foam pad, filter paper, membrane, gel, filter paper and foam pad respectively. The set up was sealed to a gel holder cassette after separation of proteins was done. Proteins would be transferred to a 0.2 µm nitrocellulose blotting membrane (Amersham Protran, GE Healthcare, Life Science) using Mini Trans-Blot Electrophoretic Transfer Cell (BioRad) with cold 1x SDS transfer buffer prepared as in Table 3.11 filling up the buffer tank. Extra care should be taken while handling the nitrocellulose membrane as the membrane should not be touched with bare hands. An ice pack was added, and the electrode lid was connected to the anode and cathode of the electrode module. A magnetic stirrer bar was also placed in the buffer tank and the electrode module was positioned on the magnetic stirrer. Transferring process was run at 300 mA for 90 minutes.

After transferring, the membrane was carefully removed from the gel holder cassette and transferred to a plate. Membrane was washed briefly for three times with 1x Tween – TBS (TBST) prepared as in Table 3.12. A 5% skimmed milk in 1x TBST was also

prepared. Membrane was then blocked with milk for an hour at room temperature on a shaker or for overnight in a cold room. Primary antibody was diluted in 5% skimmed milk as shown in Table 3.13. Membrane was incubated with the prepared primary antibody and incubated overnight at $4^{O}C$. The primary antibody would then discard, and membrane was washed three times with 1x TBST with 5 minutes soaked on shaker for each wash. This was followed by incubation with diluted secondary antibody for an hour at room temperature. Membrane was once again washed three times with 1x TBST with 5 minutes soaked on a shaker for each wash. SuperSignal West Pico PLUS chemiluminescent substrate (Thermo Fisher Scientific) was prepared with the 1:1 ratio of luminol to hydrogen peroxide as in Table 3.14, pipetted onto membrane and left for a minute at room temperature before visualized using ChemiDoc MP Imaging System (BioRad).

Table 3.12: PBS – Tween 20 and blocking buffers.

PBS – Tween20 buffer	Blocking buffer
1x PBS	1x PBS
0.1% Tween20	0.1% Tween20
	5% skimmed milk

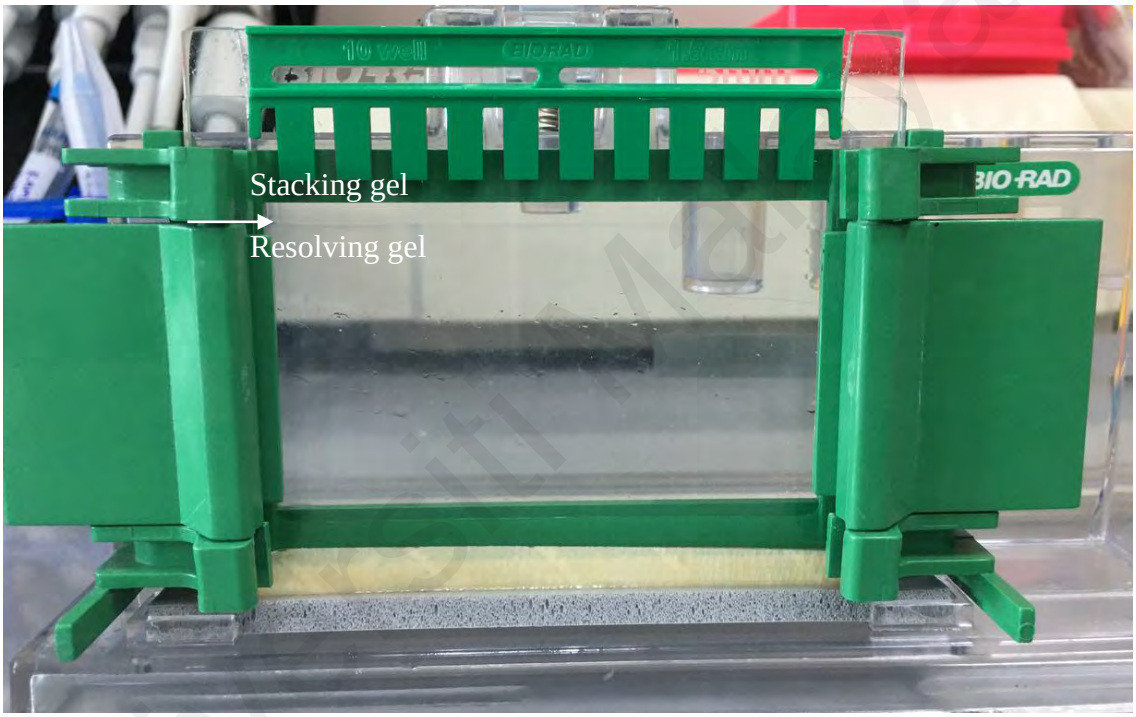
Table 3.13: Antibodies and β – actin.

Antibody	Type	Dilution
Primary anti-AGR2 (Santa Cruz Biotechnology)	Mouse monoclonal	1:1000
Secondary anti-AGR2 (Dako)	Rabbit anti-mouse	1:1000
β – actin (Santa Cruz Biotechnology)	Mouse monoclonal	1:1000

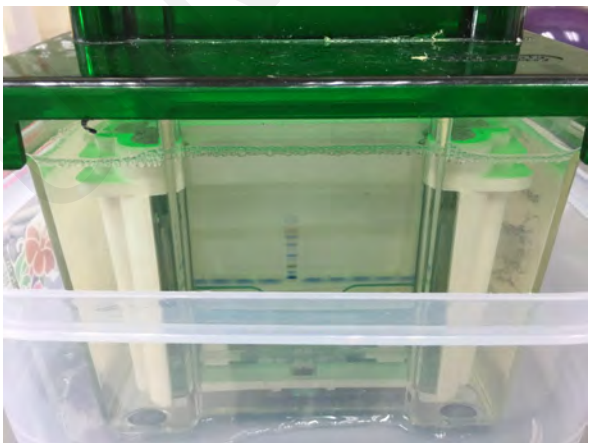
Table 3.14: Chemiluminescent prepared for viewing a membrane.

Reagent	Volume (µl)
Luminol	500
Hydrogen peroxide	500

A



B



C

Figure 3.5: SDS – PAGE and Western blot. (A) Gel preparation. (B) SDS – PAGE. (C) Traditional wet transfer process.

After visualizing, the secondard antibody was discarded, and membrane was washed three times with 1x TBST for 5 minutes on each wash. For each wash, the membrane was soaked in 1x TBST on a shaker to ensure a thorough wash. This was followed by incubation with diluted β – actin for an hour at room temperature. Membrane was washed three times with 1x TBST with 5 minutes soaked on a shaker for each wash before blotting with diluted secondary antibody for an hour at room temperature. Secondary antibody was discarded, and membrane was once again washed three times with 1x TBST with 5 minutes soaked on a shaker for each wash. SuperSignal West Pico PLUS chemiluminescent substrate was prepared with 1:1 ratio of luminol to hydrogen peroxide as in Table 3.14, pipetted onto membrane and left for a minute at room temperature before visualized using ChemiDoc MP Imaging System.

3.4.4 Immunofluorescence

To conduct immunofluorescence, cells were sub-cultured into a 1 x 8-cell chamber slide. Cells were left to grow until 50% confluency was achieved before transfected as described in 3.3.4. Media was removed and cells were washed three times with 1x PBS. Loading chamber was removed from the chamber slide and the slide was placed onto parafilm in a plate wrapped with aluminium foil. Cells were fixed with 4% paraformaldehyde (PFA) for 15 minutes at room temperature. Cells were washed three times with 1x PBS and permeabilized with 0.1% triton in PBS for 10 minutes at room temperature. Cells were washed three times with 1xPBS and blocked with 3% BSA in 1x PBS for 30 minutes at room temperature. Primary antibody was diluted in 3% BSA in 1x PBS at the ratio of 1:200 antibody to BSA. Coverslip containing cells were incubated with the diluted antibody for overnight in a cold room. The coverslips were washed with 1x PBS for three times with 5 minutes incubation per wash. Secondary antibody which had been conjugated to fluorophore (Alexa Fluor 488) was diluted in 3% BSA in 1x PBS

at the ratio of 1:1000 antibody to BSA and incubated at room temperature for an hour. The cells were washed with 1x PBS for three times with 5 minutes incubation per wash. The nuclei of the cells were then stained with the diluted DAPI nuclear stain (Invitrogen) and incubated at room temperature for a minute before mounted with coverslip using 10 µl or a drop of mounting media (Dako) and viewed using fluorescent microscope (Nikon).

Table 3.15: Antibodies and staining for immunofluorescent imaging.

Antibody/staining	Type	Dilution
Primary antibody	Mouse monoclonal	1:200
Secondary antibody (Alexa Fluor 488)	Goat anti-mouse	1:1000
DAPI	Nuclear stain	-

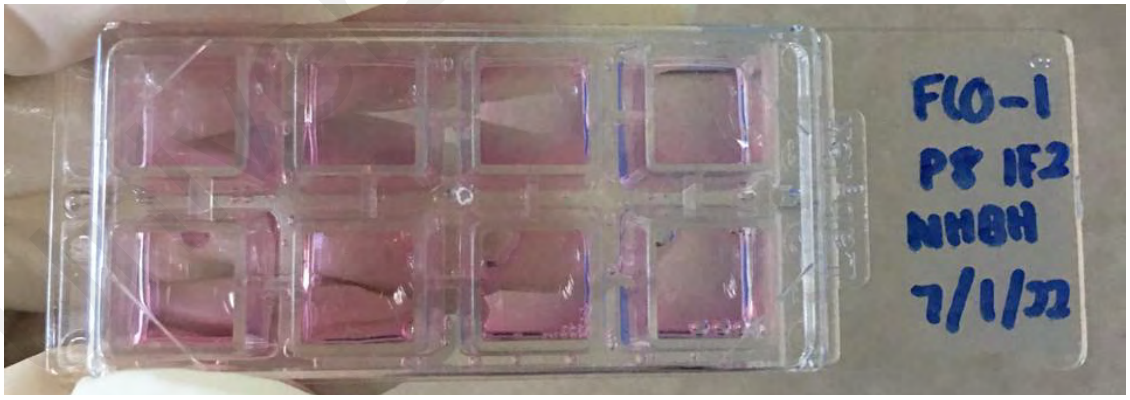


Figure 3.6: Chamber slide containing transfected FLO-1 cell line for immunofluorescence.

Objective 3: To determine the viability and migratory potential of esophageal cancer cells expressing AGR2 protein variants.

3.5 Functional assays

3.5.1 Wound healing assay

To conduct wound healing assay, FLO-1 cell line was first grown in a 10 cm – cell culture dish and left to grow until it achieved 80 – 90% confluency. Cells were then trypsinized with 1 ml of trypsin-EDTA for 5 minutes at $37^{O}C$ with 5% CO_2 incubator. 11 ml of media was pipetted into a new 10-cm cell culture dish while another 10 ml media was pipetted into the trypsinized cells. The trypsinized cells were resuspended well in the media before transferred into the new dish-containing media and resuspended again. From the 10-cm dish, 12 ml media containing trypsinized cells were then seeded into 1 x 6-well plate with 2 ml media and trypsinized cells occupying each well while the rest of the cells were incubated for maintaining purpose. Cells in 1 x 6-well plate were left to grow at $37^{O}C$ with 5% CO_2 incubator until it achieved 60 – 70% confluency. Transfection was conducted as described in 3.3.4 and incubated for 48 hours until cells achieved 90% or full confluency.

Changing of media was done before conducting the assay. Scratching of cells was conducted using a 100 µl pipette tip with a change of different tip for each AGR2 protein variants. The best scratching area was chosen and marked before captured under the inverted microscope projected to a computer display. Such step was repeated after 24, 48 and 72 hours and the time taken for the transfected FLO-1-AGR2 protein variants to migrate until the scratching gap was fully closed was recorded. The area of cell-free region was measured using Image J software and the percentage of wound closure was

calculated with relative to 0 hours before analyzed using two-way ANOVA via GraphPad Prism software version 8.0.2 (263).

3.5.2 3-(4, 5-Dimethyl-2-thiazolyl)-2, 5-diphenyl-2H-tetrazolium bromide (MTT) assay

To conduct MTT assay, cells were sub-cultured into a 1 x 6-well plate and transfected as described in 3.3.4. Changing of media was done before conducting the MTT assay. Cells were trpsinized with the ratio of 1:4 trypsin to media. A volume of 30 µl of trypsinized cells were transferred into a microcentrifuge tube and mixed with 30 µl 0.4% tryphan blue and loaded onto the hemocytometer for cell counting purpose. Dilution of cells were done from 1 x 10^4 – 1 x 10^6 cells/ml. In five replicates, cells were seeded at 10^3 – 10^5 cells/well with 100 µl/well for each dilution. Cells were incubated at 37^OC with 5% CO_2 incubator for 2 – 24 hours. The subsequent procedures were conducted in dark as the reagent used was light-sensitive. A volume of 10 µl of MTT reagent was added and cells were incubated for 4 – 24 hours at 37^OC with 5% CO_2 incubator. The plate was wrapped in aluminium foil beforehand. Once purple precipitate was visible under the microscope, 100 µl detergent was then added into all wells. Microplate should not be shaken after adding the detergent. Cells were incubated overnight at 37^OC with 5% CO_2 incubator in dark. Plate cover was then removed, and absorbance of each well was read at 570 nm. All procedures were repeated for day 2 and day 4. The absorbances obtained was then analyzed using two-way ANOVA via GraphPad Prism software version 8.0.2 (263).

3.5.3 Alamar Blue assay

To conduct Alamar Blue assay, cells were sub-cultured into a 1 x 6-well plate and transfected as described in 3.3.4. Changing of media was done before conducting the

assay. Cells were trpsinized with the ratio of 1:4 trypsin to media. A volume of 30 µl of trypsinized cells were transferred into a microcentrifuge tube and mixed with 30 µl 0.4% tryphan blue and loaded onto the hemocytometer for cell counting purpose. Dilution of cells were done from 1 x 10^4 – 1 x 10^6 cells/ml. In five replicates, cells were seeded at 10^1 – 10^3 cells/well with 200 µl/well for each dilution. Cells were incubated at 37^OC with 5% CO_2 incubator for 24 hours. The subsequent procedures were conducted in dark as the reagent used was light-sensitive. Alamar blue reagent was diluted with the same media used to grow cell line at ratio of 1:10. A volume of 120 µl of diluted reagent was pipetted into the cells seeded in 96-well plate. The plate was incubated at 37^OC with 5% CO_2 incubator for 90 minutes. Plate cover was then removed, and absorbance of each well was read at 570 nm. All procedures were repeated for day 2 and day 4. The absorbances obtained was then analyzed using two-way ANOVA via GraphPad Prism software version 8.0.2 (263).

CHAPTER 4: RESULTS

4.1 Objective 1: To construct vectors for the expression of recombinant AGR2 protein variants in mammalian cells.

AGR2 protein which was initially found to be existing in MCF-7 breast cancer cell line also had expression in other cancer tumors including pancreas, stomach, and colon cancer. Despite their prominence role in protein folding machinery, AGR2's role in cancer was undeniably remarkable. Hence, in the first phase of our research, we aimed to introduce the recombinant AGR2 protein variants comprised of AGR2 – WT, AGR2 – KDEL, mAGR2 – ΔKTEL as well as ΔNterm AGR2 – KDEL into the FLO-1 cancer cell line. The esophageal adenocarcinoma, FLO-1 that we used in this study did not express the AGR2 protein (Wang *et al*., 2014) therefore, it allowed for clear and further exploration of the significant functional role of each of the AGR2 protein variant constructed. The four variants of AGR2 protein were successfully synthesized chemically and the altered sequence of each variant was as presented in appendix B. The first variant, AGR – WT was characterized by having the full length of AGR2 protein wild type. They served as the reference to the other three mutated variants. As for the second variant, the AGR2 – KDEL was established by introducing point mutations to the canonical KTEL thus resulting in the change from KTEL motif to KDEL motif. For this construct, we aimed to enhance the localization of the recombinant AGR2 protein in the intracellular environment of the FLO-1 cancer cells. The third variant comprised of mature AGR2 protein which was characterized by lacking the first 20 amino acids signal peptide as well as the KTEL motif and aimed to enhance the secretion of AGR2 protein into the extracellular environment. The fourth variant on the other hand was characterized through deletion of unfolded N-terminal of the AGR2 protein at 21 – 40 amino acid domains as

well as point mutations also resulting in the change of KTEL motif to KDEL motif. Both alterations were done to significantly enhance the function of AGR2 protein dimer as AGR2 protein was known to be existed as homodimer. Each variant established in our research was based on their roles as described previously by Brychtova *et al*. (2015); Delom *et al*. (2020); Delom *et al*. (2018).

The concentration of AGR2 protein constructs successfully ligated into the pSF-CMV-Amp vectors was as presented in Table 4.1. All four designated constructs of AGR2 protein were verified with Sanger sequencing and the chromatograms were as presented in Figure 4.1. All sequencing data observed from the sequencing chromatogram of the four AGR2 protein variants, the AGR2 – WT, AGR2 – KDEL, mAGR2 – ΔKTEL as well as ΔNterm AGR2 – KDEL were found to have evenly-spaced nucleotide peaks with minimum baseline noises indicating good sequencing results.

Table 4.1: Concentration of plasmids harboring AGR2 protein constructs.

AGR2 protein constructs	Concentration of variants (ng/µl)
AGR2 – WT	696
AGR2 – KDEL	660
mAGR2 – ΔKTEL	730
ΔNterm AGR2 – KDEL	695

A

B

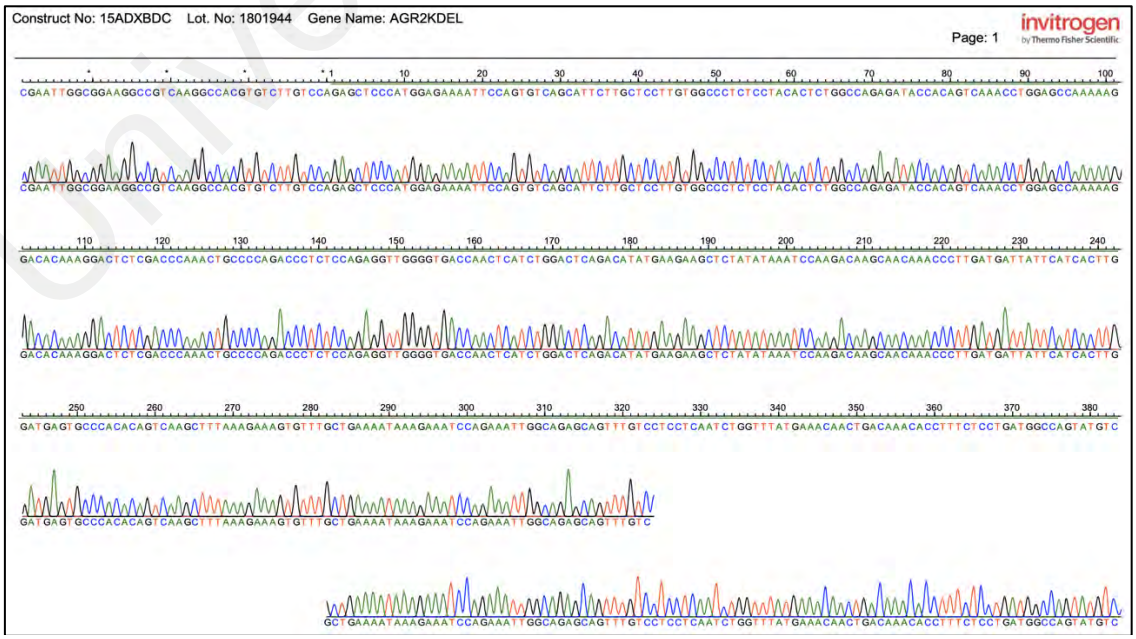
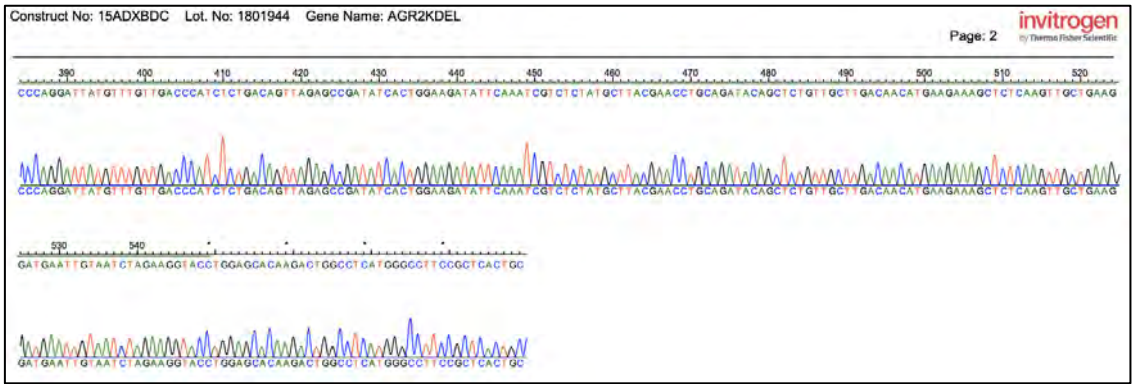
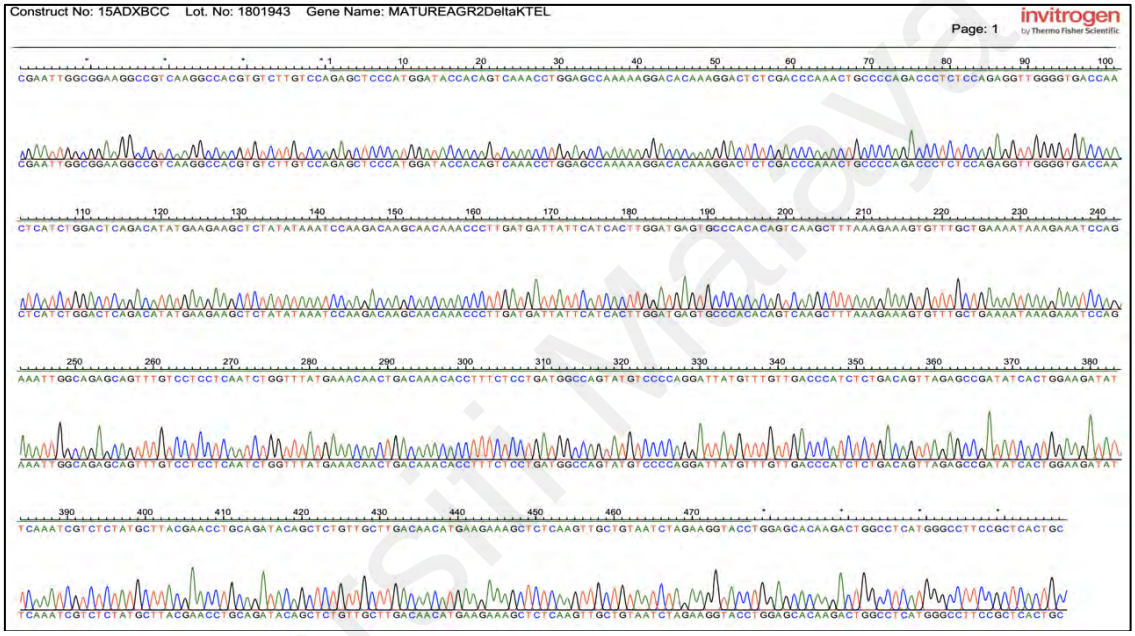


Figure 4.1: Sequencing chromatogram of AGR2 protein variants. (A) AGR2 – WT. (B) AGR2 – KDEL. (C) mAGR2 – ΔKTEL. (D) ΔNterm AGR2 – KDEL (Taken from Mohtar, 2017).



C



D

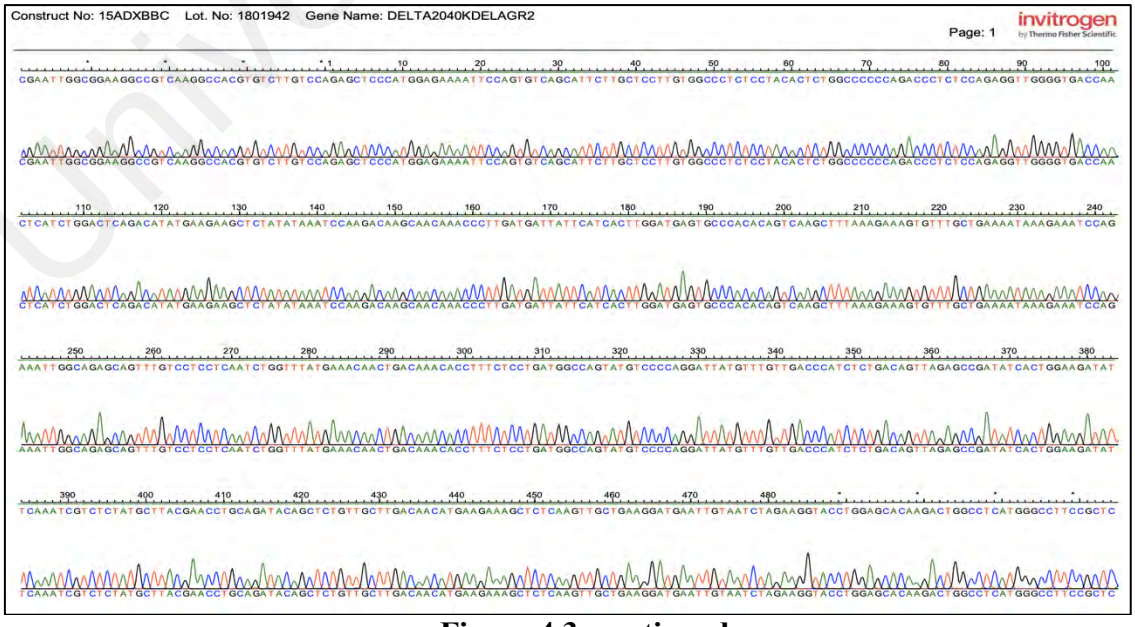


Figure 4.3, continued.

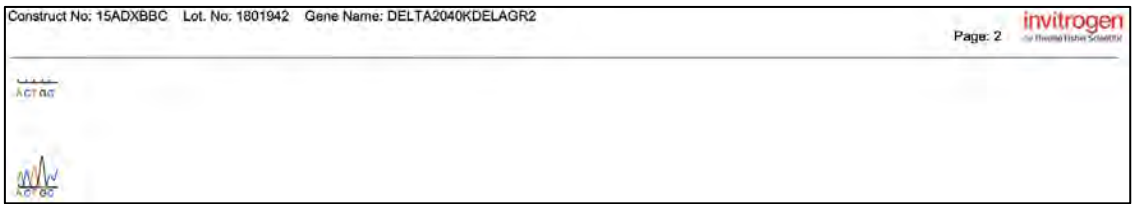


Figure 4.3, continued.

4.2 Objective 2: To determine the level of intracellular and extracellular expression of AGR2 protein variants expressed from the esophageal cancer cells.

4.2.1 Intracellular expression of AGR2 protein variants in esophageal adenocarcinoma.

In the second part of our research, we aimed to determine the protein expression levels of the constructed AGR2 protein variants in the FLO-1 esophageal adenocarcinoma. The western blot analysis was conducted 24 hours post transfection using the mouse monoclonal AGR2 antibody revealed the level of intracellular AGR2 protein variants expression in FLO-1 esophageal adenocarcinoma as depicted in Figure 4.2 (A).

From the results that we obtained, expressions of the recombinant AGR2 proteins transfected into the FLO-1 cancer cell line were seen to be present only for the three of the variants which were AGR2 – WT, AGR2 – KDEL and mAGR2 – ΔKTEL with no expression observed for the ΔNterm AGR2 – KDEL. The blotting showed that AGR2 – WT had the strongest expression in the intracellular of FLO-1 cell line as compared to the other recombinant variants, AGR2 – KDEL and mAGR2 – ΔKTEL and this was as reflected in the graph in Figure 4.2 (B). The graph depicted that the expression of AGR2 – WT was high in relative to β – actin expression with 0.72 ratio of AGR2 – WT to β – actin as compared to the other ratios of other mutated AGR2 variants with β – actin. The presence of such expression verified our hypothesis such that the AGR2 – WT bearing the canonical KTEL motif was capable of binding to the KDEL motif of the ER resulting in the expression of AGR2 – WT in the intracellular of FLO-1 cell.

Next, our immunoblotting revealed that FLO-1 transfected with AGR2 – KDEL also showed expression in the intracellular of the cell, but such expression was observed to be lowered as compared to the AGR2 – WT. The lower expression was depicted in the graph in Figure 4.2 (B) such that in relative to β – actin, the AGR2 – KDEL recorded 0.69 ratio of the size and intensity of AGR2 – KDEL to β – actin as compared to the AGR2 – WT. This, however, did not correspond to our hypothesis in which we hypothesized that mutating the KTEL motif to KDEL motif would enhance the AGR2 protein localization in the ER of the FLO-1 cell thus resulting in stronger expression intracellularly for the construct bearing the KDEL motif as compared to the wild type bearing the canonical KTEL motif.

The same blotting also revealed that mAGR2 – ΔKTEL transfected into the FLO-1 cancer cell line had even lower expression signified by the lower intensity of band as compared to both first and second construct of the recombinant AGR2 protein. The lower expression measured was signified by the ratio value of 0.13 mAGR2 – ΔKTEL to β – actin as plotted on the graph in Figure 4.2 (B).

In addition, on the same blot, we determined that the ΔNterm AGR2 – KDEL was not expressed in the FLO-1 cell line even though deletion of the N-terminal should at least result in low expression intracellularly instead of no expression. This did not support our hypothesis on mutating the N-terminal of the AGR2 protein as well as introducing point mutation to substitute the KTEL motif with the KDEL motif in order to enhance dimerization of the AGR2 protein could result in expression in both, the intracellular and extracellular environment of the FLO-1.

A

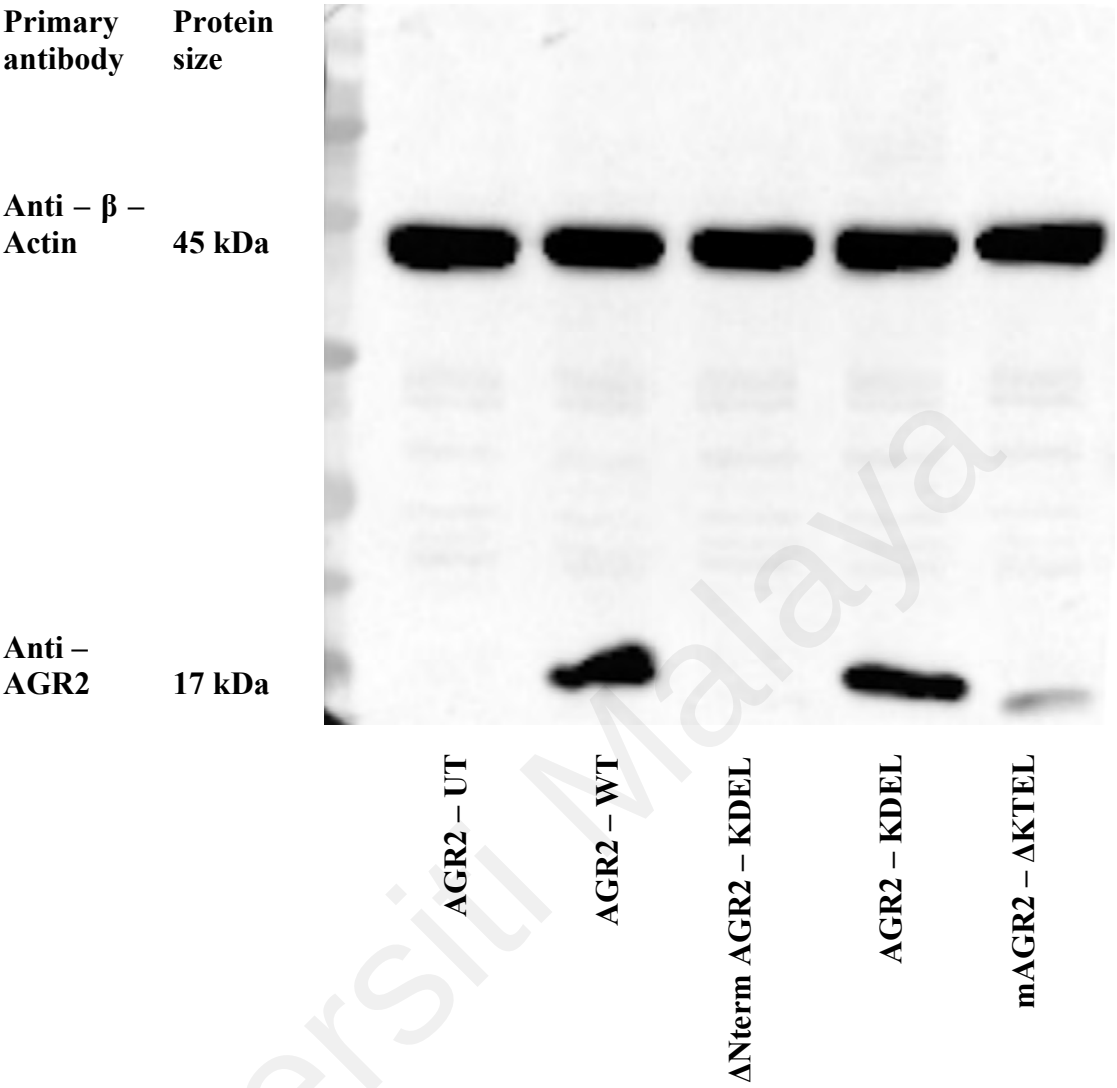


Figure 4.2: Intracellular expression of the AGR2 protein variants in the FLO-1 cancer cell line. Immunoblotting was conducted 24 hours post transfection. AGR2-UT denoted for untransfected AGR2 protein. (A) Western blot analysis showing presence of bands on FLO-1 transfected with AGR2 – WT, AGR2 – KDEL, mAGR2 – ΔKTEL with no expression observed for ΔNterm AGR2 – KDEL. (B) Bar graph showing ratios of transfected intracellular AGR2 protein variants to β – actin. Statistical test conducted using multiple t-test on the ratio of AGR2 protein variants expression in relative to β – Actin in comparison with the untransfected cell revealed *p>0.05.

B

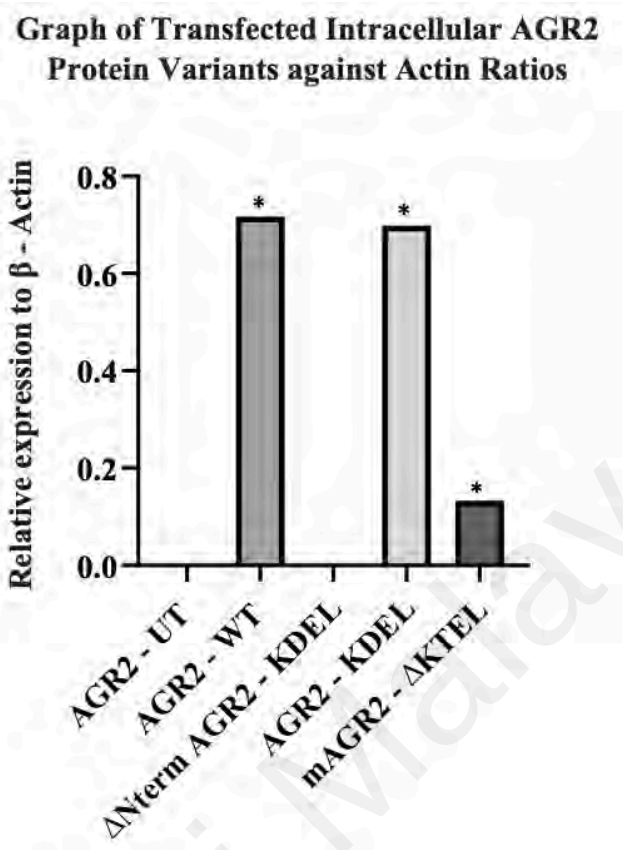


Figure 4.2, continued.

4.2.2 Extracellular expression of AGR2 protein variants in esophageal adenocarcinoma.

Other than that, previous studies had also proved the detection of AGR2 protein in the conditioned media from cancer cells which signified AGR's secretion into the extracellular environment suggesting role that was yet to be identified by previous research. The finding of AGR2 protein in the tumor niche microenvironment was common among the protein in the PDI family as they were characterized by having the thioredoxin folds which were known to be crucial for them to act extracellularly in which majorly at the cell membrane. Therefore, in the second part of our research, we also aimed to determine the expression levels of the four constructed AGR2 protein variants, eAGR2 – WT, eAGR2 – KDEL, emAGR2 – ΔKTEL and ΔNterm eAGR2 – KDEL at the

extracellular environment of the FLO-1 esophageal adenocarcinoma. The level of expression of the AGR2 protein variants in the extracellular of the FLO-1 was as depicted in Figure 4.3 (A).

From the results that we obtained, expressions of the transfected recombinant AGR2 proteins in the extracellular environment of the FLO-1 cancer cell line were seen to be present only for the two out of four recombinant variants. Such variants were eAGR2 – WT and eAGR2 – KDEL. Our immunoblotting clearly depicted that no expression was observed at the secreted media from the FLO-1 cell transfected to another two variants, emAGR2 – ΔKTEL and ΔNterm eAGR2 – KDEL. The presence of bands which signified the extracellular expression of the recombinant AGR2 protein variants in FLO-1 secretome was as reflected in graph in Figure 4.3 (B).

In relative to β – actin, the eAGR2 – WT recorded 0.72 ratio of the size and intensity of band. The same relative expression of AGR2 – WT with β – actin in both intracellular and extracellular of the FLO-1 cell was in concordance with our hypothesis such that the first variant that we were trying to investigate which in our case was AGR2 – WT, would be half-retained in the intracellular of the FLO-1 cell line and at the same time secreted into the extracellular environment of the FLO-1.

In the same blotting, we also verified the presence of band in the extracellular of FLO-1 cell transfected with AGR2 – KDEL. The ratio of the size and intensity of band measured was 0.38 and this was reported to be lowered than the intracellular expression of such variant. The lower expression in the extracellular of FLO-1 as compared to the intracellular supported our hypothesis on mutating the KTEL motif to the KDEL motif resulted in more binding of AGR2 protein to the ER reflecting high retaining of such protein in the intracellular as compared to in the secretome of the FLO-1.

As for the emAGR2 – ΔKTEL, absence of band could be observed from the same membrane. We initially expected that the variant would have high expression at the

extracellular environment. Instead, the result that we obtained was in contrast with our hypothesis in which deleting the signal peptide and KTEL motif from the full length of the AGR2 protein could enhance AGR2 protein secretion into the extracellular environment of the FLO-1.

A

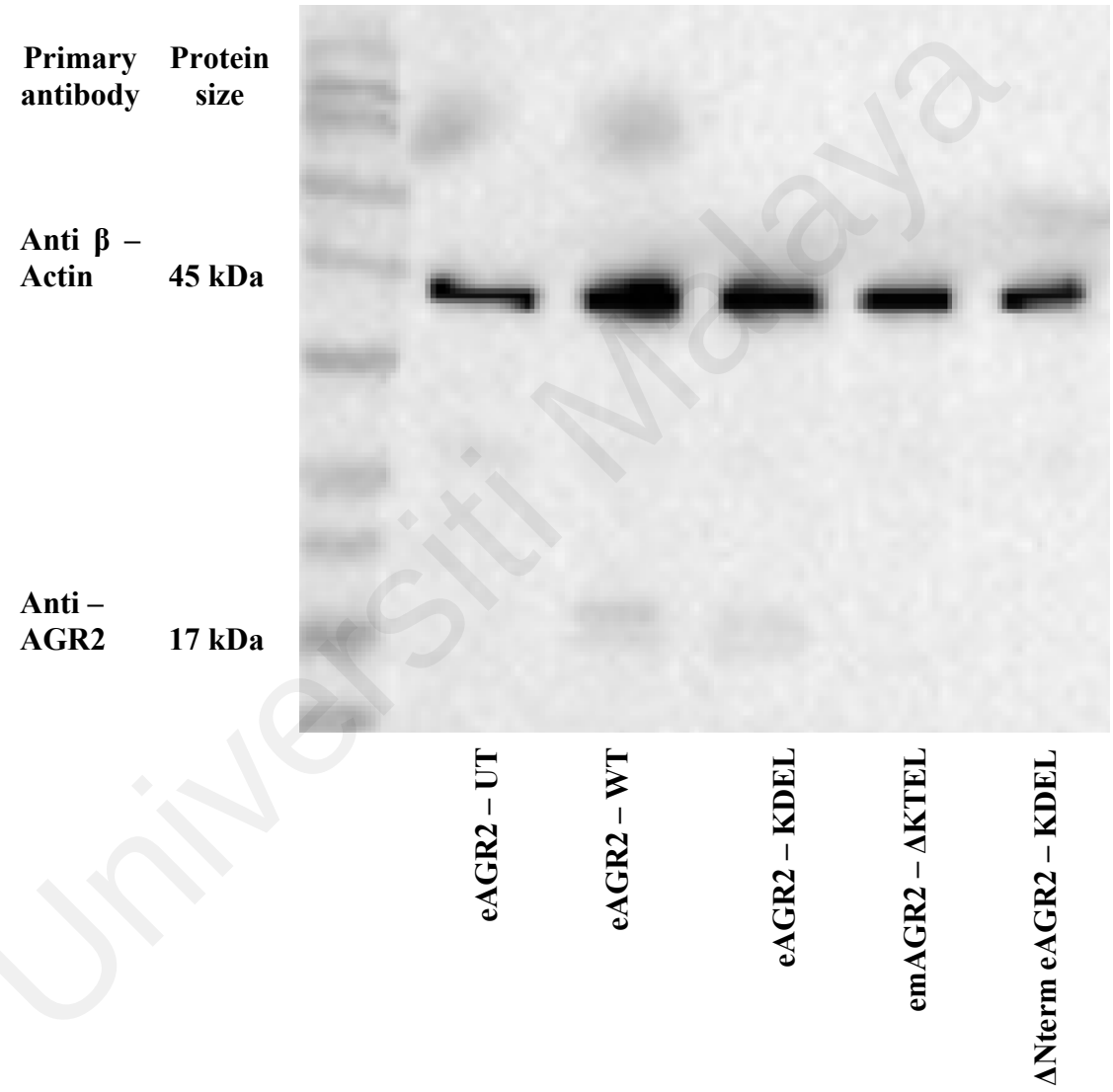


Figure 4.3: Extracellular expression of the FLO-1 cancer cell line transfected with AGR2 protein variants. Immunoblotting was conducted 24 hours post transfection. (A) Western blot analysis showing presence of bands at low intensity for eAGR2 – WT and eAGR2 – KDEL with no expressions observed in extracellular of FLO-1 transfected with emAGR2 – ΔKTEL and ΔNterm eAGR2 – KDEL. (B) Bar graph showing ratios of transfected extracellular AGR2 protein variants to β – actin. Statistical test conducted using multiple t-test on the ratio of AGR2 protein variants expression in relative to β – Actin in comparison with the untransfected cell revealed *p>0.05.

B

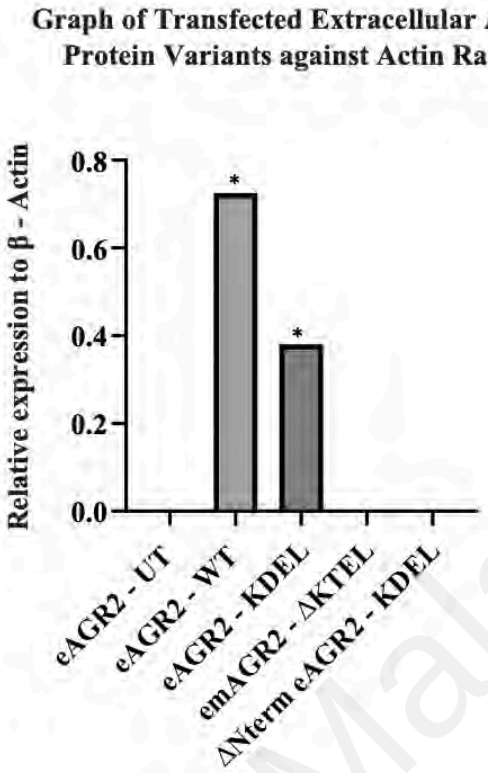


Figure 4.3, continued.

Meanwhile, as the function of the dimer motif of the AGR2 protein was more significant in the extracellular environment hence, we hypothesized that the FLO-1 transfected with ΔNterm eAGR2 – KDEL construct should resulted in higher expression in the extracellular environment. Instead, the immunoblotting conducted revealed an in-contrast result such that no expression was detected in the extracellular environment of the FLO-1 transfected with ΔNterm eAGR2 – KDEL.

4.2.3 Immunofluorescence for intracellular localization of AGR2 protein variants in esophageal adenocarcinoma.

To further assess the intracellular localization of our four constructed AGR2 protein variants, AGR2 – WT, AGR2 – KDEL, mAGR2 – ΔKTEL as well as ΔNterm AGR2 – KDEL in the FLO-1 cell line, such transfected FLO-1 were then processed for

immunofluorescence microscopy. From the immunofluorescence conducted, the intracellular localization of AGR2 protein variants in the FLO-1 cell line was as depicted in Figure 4.4.

Basically, the immunofluorescence that we conducted successfully validated the intracellular expression of the AGR2 protein variants that we obtained from the immunoblotting. From Figure 4.4, it was proved that the FLO-1 cell line that we used in this study did not express AGR2 protein and this was determined from the image of the untransfected AGR2 (AGR2-UT) in which only nuclei of the esophageal adenocarcinoma cells were stained while the other parts of the cells were not stained with the anti-AGR2 antibody which had been conjugated to a fluorophore, a type of fluorescent dye used to stain the AGR2 protein. The same procedure conducted also revealed that three of our constructed protein variants, AGR2 – WT, AGR2 – KDEL and mAGR2 – ΔKTEL transfected into FLO-1 cell line had expression in the intracellular environment of the esophageal adenocarcinoma cancer cells, and this was in parallel with the results that we obtained from the immunoblotting. Despite the low transfection efficiency however, certain FLO-1 cell transfected with the AGR2 – WT, AGR2 – KDEL and mAGR2 – ΔKTEL were observed to fluoresce in green. As for the ΔNterm AGR2 – KDEL variant, our immunofluorescence results also revealed that no expression was observed in the intracellular of the FLO-1 cell with only nuclei being stained instead of the cytoplasm indicating that no AGR2 protein variant was detected since AGR2 was and ER-residence protein. This was as reflected in our immunoblotting such that no band was observed from the membrane blotted with anti-AGR2 antibody.

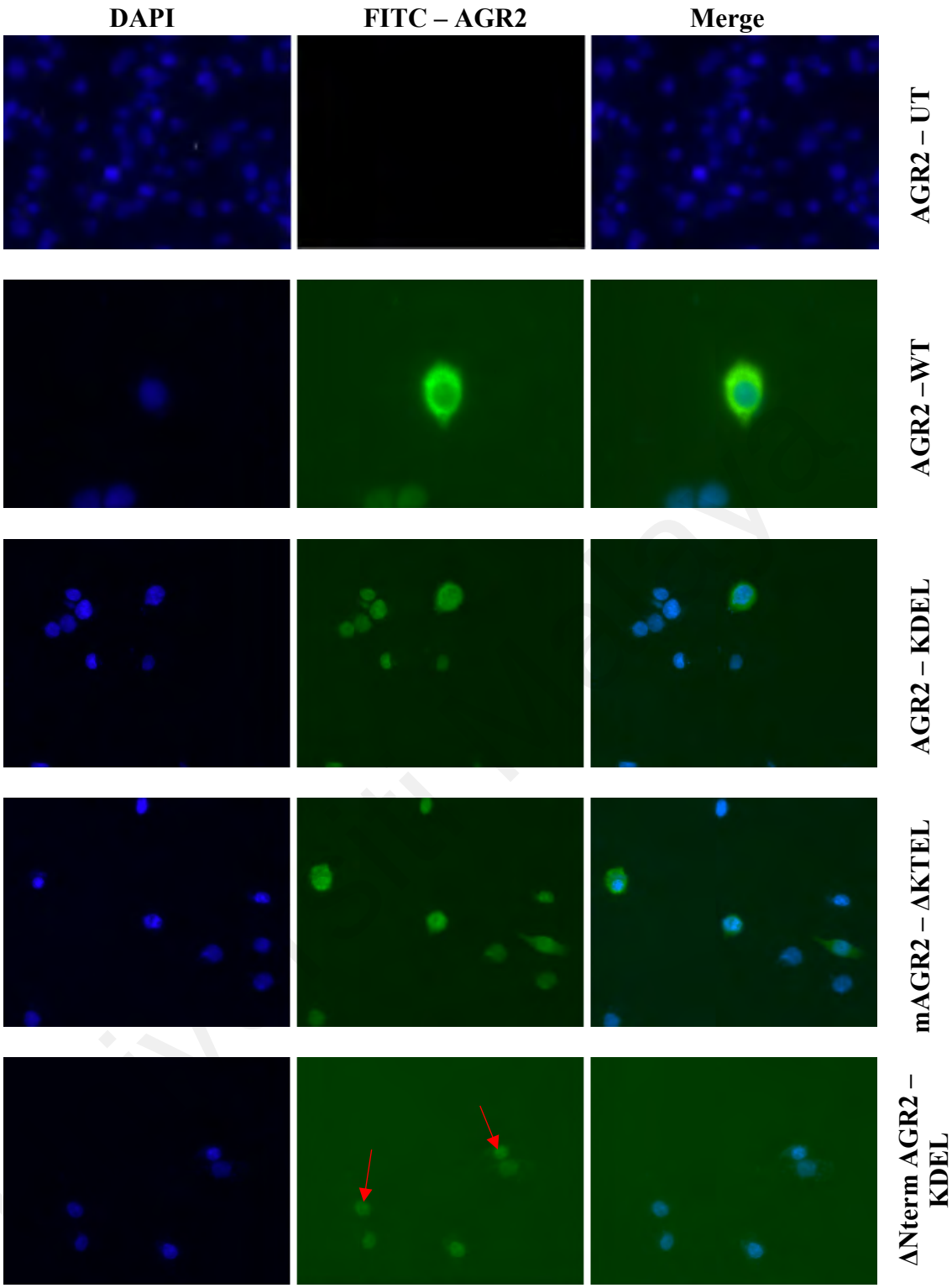


Figure 4.4: Intracellular localization of transiently transfected AGR2 protein variants in the FLO-1 esophageal adenocarcinoma subjected to immunofluorescence observed at 40x magnification. The AGR2 protein was unexpressed in the untransfected FLO-1. Mutating the AGR2 proteins resulted in expression for three constructed variants, AGR2 – WT, AGR2 – KDEL and mAGR2 – ΔKTEL with no expression seen for ΔNterm AGR2 – KDEL construct. Red arrow depicted only nuclei of the cell showed signal with no occurrence of AGR2 protein which should be localized in the outside of nucleus in the cytoplasm of the cell.

4.3 Objective 3: To determine the viability and migratory potential of esophageal cancer cells expressing AGR2 protein variants.

4.3.1 Wound healing assay.

Previous study by Zhu *et al*. (2017) reported that AGR2 enhanced proliferation and migration of fibroblasts and keratinocytes. Therefore, in the third phase of our research, we aimed to determine the phenotypic effects of esophageal adenocarcinoma cells expressing either full length of AGR2 protein, AGR2 – WT or modified AGR2 protein variants, AGR2 – KDEL, mAGR2 – ΔKTEL or ΔNterm AGR2 – KDEL. The first assay that we conducted to study such phenotypic effects of AGR2 protein variants was the wound healing assay. We conducted such assay to observe any significance differences between the four transfected variants of AGR2 protein in promoting the proliferation and migration of the FLO-1 cell line used in this study. The results that we obtained was as depicted in Figure 4.5.

Generally, FLO-1 cancer cell line both either transfected or untransfected with AGR2 protein variants took a total of 72 hours to migrate into the cell-free region. In the presence of AGR2 protein, however, we observed that the FLO-1 cell did not show much significant difference ($p>0.05$) in proliferation and migration rate as compared to the untransfected cell line. However, comparing among the four transfected variants, three of the variants, AGR2 – KDEL, mAGR2 – ΔKTEL and ΔNterm AGR2 – KDEL showed an increase in proliferation and migration rate in different time points as compared to the wild type of AGR2 protein ($p<0.05$). This was shown after 24 hours and consequently 48 hours of being introduced to a scratch such that the FLO-1 transfected with AGR2 – KDEL and mAGR2 – ΔKTEL proliferated and migrated faster into the cell-free region.

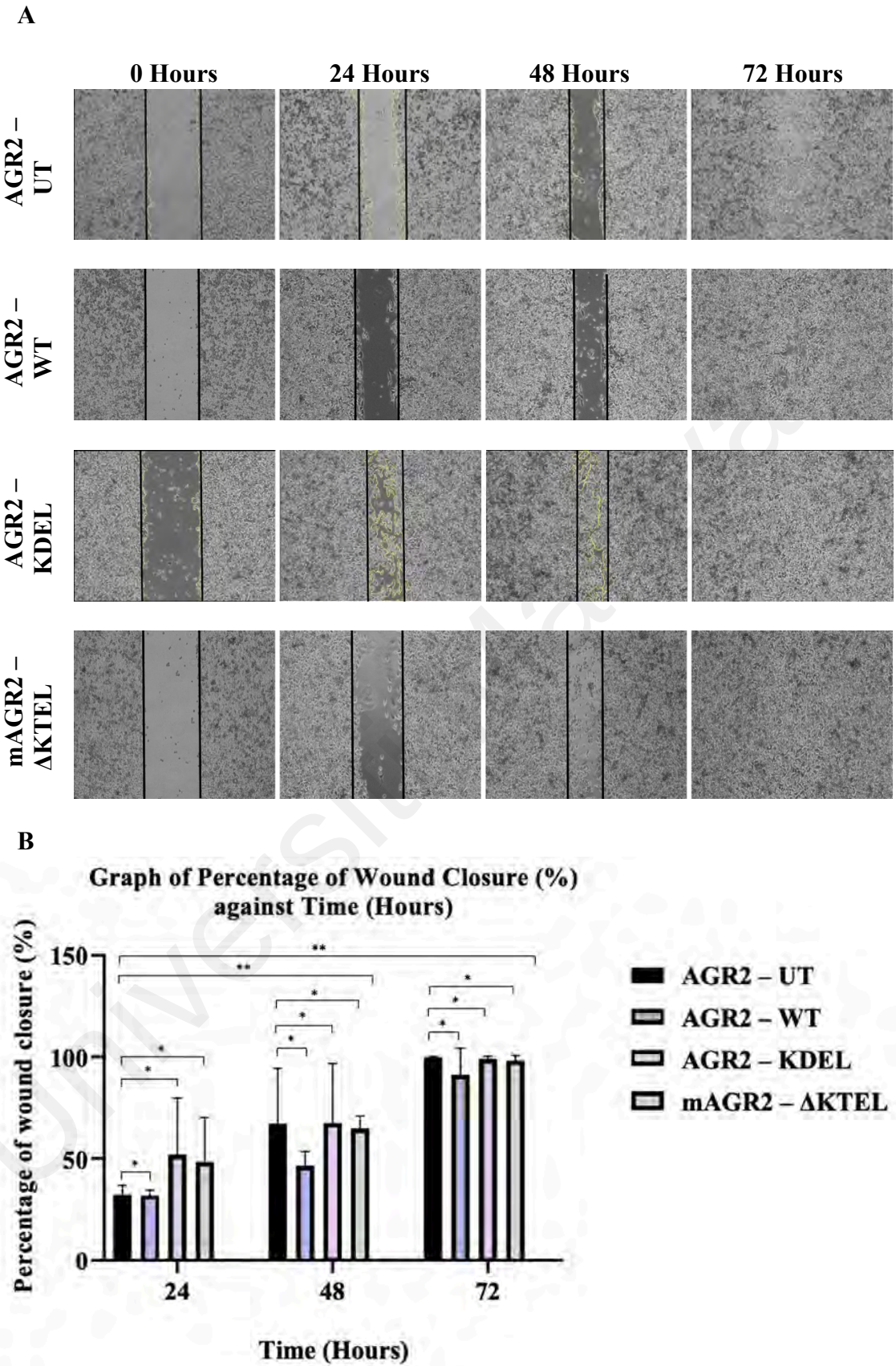


Figure 4.5: Assessing proliferating and migrating effects of transfecting wild type and mutated AGR2 protein variants in FLO-1 cell line. (A) Representative figures from wound healing assay conducted on FLO-1 cell line transfected with AGR2 protein variants demonstrating cell migration into the cell-free region (outlined) was accelerated with the presence of mutated variants of AGR2 – KDEL and mAGR2 –

ΔKTEL but not for the wild type (AGR2 – WT) as compared to the non-targeting control (AGR2 – UT) after 24 and 48 hours introduced to a wound. Figures were observed at 10x magnification. (B) Bar graph illustrating the percentage of wound closure at indicated time points for the wound healing assay conducted. Data was expressed as mean ± standard deviation. Statistical test conducted using two-way analysis of variance (two-way ANOVA) revealed*p>0.05 for comparison in between untransfected and transfected AGR2 protein variants to achieve full wound closure and **p<0.0001 for comparison of wound closure on different time points.

We also measured the area of the cell-free region and calculated the percentage of closure and from here, we determined that the three mutated variants migrated and proliferated until the cell-free region almost fully covered with the cells resulting in percentage of closure of AGR2 – KDEL (99.1%), and mAGR2 – ΔKTEL (98.5%) after the last time point (72 hours) while the FLO-1 transfected to AGR2 – WT (91.4%) as compared to the untransfected AGR2 (99.6%).

4.3.2 MTT (3[4,5-dimethylthiazol-2-yl]-2,5 diphenyl tetrazolium bromide) assay.

To study the viability of the esophageal cancer cell line after transfected with the AGR2 protein variants, we carried out MTT assay which could detect any mitochondrial activity in the FLO-1 cell line via the conversion of MTT (3[4,5-dimethylthiazol-2-yl]-2,5 diphenyl tetrazolium bromide) into formazan crystals. The constant mitochondrial activity of cells was proportional to the number of viable cells hence, allowed us to determine the percentage of cell viability without the cell counting measure. The results of MTT assay performed was as depicted in Figure 4.6.

From the results obtained, transfecting the FLO-1 cell line with either AGR2 wild type, AGR2 – WT or modified variants, AGR2 – KDEL and mAGR2 – ΔKTEL however, did not result in much significant difference as compared to the untransfected variants (p<0.05). Despite that, the viability of the FLO-1 cell either untransfected or transiently transfected to AGR2 protein variant was observed to be significantly increased (p<0.05) at 48- and 96-hours' time interval as seen from the increased of OD_{570nm} values.

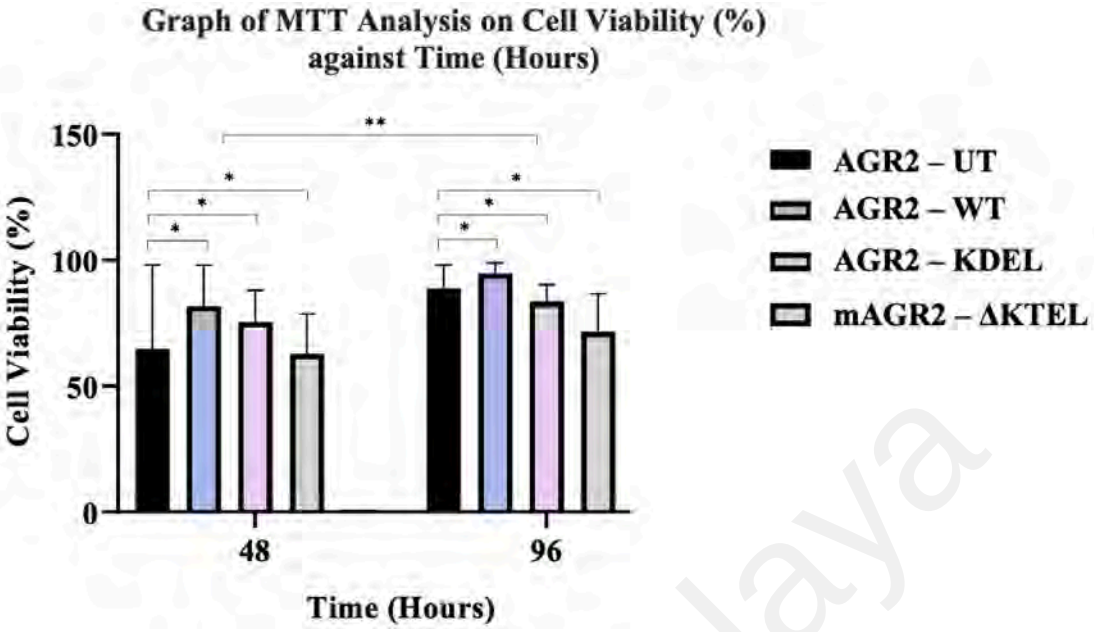


Figure 4.6: Bar graph illustrating the percentage of cell viability at indicated time points from the MTT assay conducted. Data was expressed as mean ± standard deviation. Statistical test conducted using two-way ANOVA revealed *p<0.05 for comparison of cell viability in between untransfected as well as transfected AGR2 protein variants and ** p<0.05 for comparison of cell viability in between time points.

4.3.3 Alamar Blue assay.

Apart from MTT assay, we proceeded to determine the viability of the esophageal adenocarcinoma transfected with our AGR2 protein constructs through conduction of Alamar Blue assay. For this assay, the resazurin compound would serve as an oxidation-reduction indicator which would undergo changes resulting from reduction of cellular metabolic process.

The FLO-1 tested was introduced to the Alamar Blue reagent which could detect oxidation level from the cell's respiration hence, signifying viability of the cells. The results of Alamar Blue assay performed was as depicted in Figure 4.7. From the results obtained, we determined that there was no significant different (p>0.05) observed on the viability of FLO-1 cell either untransfected or transfected with AGR2 protein variants. However, the viability of the untransfected FLO-1 cell as well as three of the transfected

variants, AGR2 – WT, AGR2 – KDEL, and mAGR2 – ΔKTEL was observed to be decreased at 48- and 96-hours' time interval.

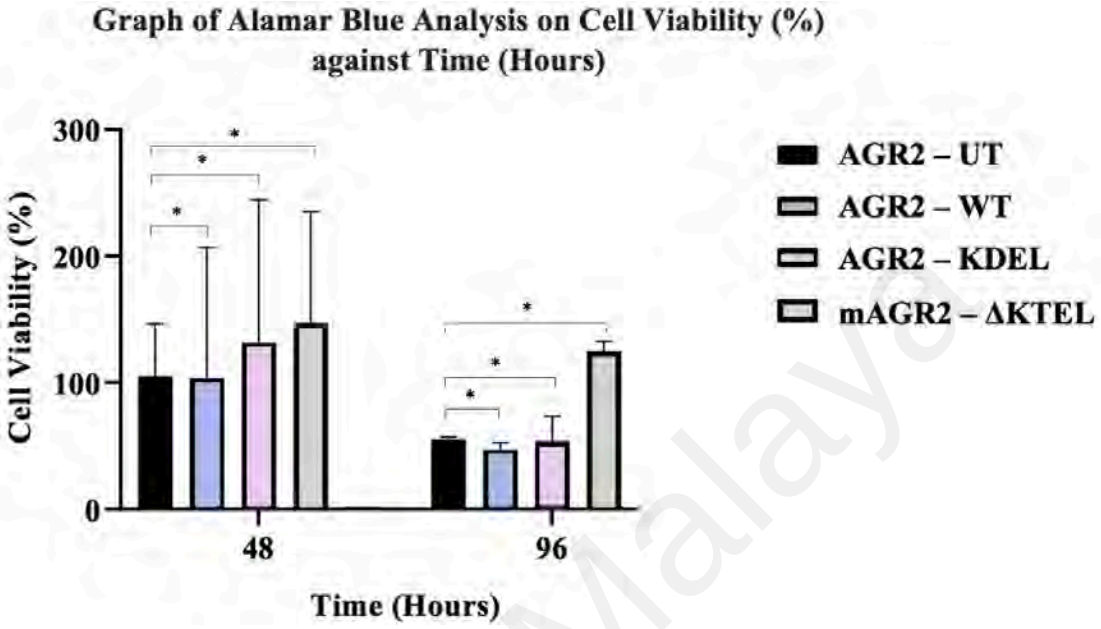


Figure 4.7: Bar graph illustrating the percentage of cell viability at indicated time points from the Alamar Blue assay conducted. Data was expressed as mean ± standard deviation. Statistical test conducted using one-way ANOVA revealed p>0.05.

CHAPTER 5: DISCUSSION

Prior research had documented on the emerging roles of the pro-oncogenic AGR2 protein in advancing various types of cancer ranging from hormone dependent to non-hormone dependent cancers. Despite the abundance of study associated to AGR2 protein however, the ambiguous expression of such protein in the esophageal adenocarcinoma was still not fully established hence, leading us to design the research on the interrogation of structural variations of pro-oncogenic AGR2 protein in supporting esophageal cancer progression. Based on the previous evidences on the role of AGR2 protein in supporting cancer cell progression either intracellularly or extracellularly, we constructed three specific objectives that would lead us to understand the mechanism of how AGR2 protein was retained and secreted in esophageal cancer through generation of recombinant AGR2 protein variants. Such recombinant AGR2 protein variants constructed allowed us to determine the selected structural function that made up the AGR2 protein, suggesting their prominence role towards retention or secretion of the protein in our selected cancer cell line. From here, the expression levels of the constructed AGR2 protein variants in the intracellular and extracellular environment of the FLO-1 esophageal adenocarcinoma was assessed. Consequently, the phenotypic effects in terms of proliferation and migration of the cells expressing AGR2 protein variants on esophageal cell line was observed in addition to the viability of the cells post-transfected to our four designated variants.

To fulfil the first and second designated objectives in our research, we constructed the first variant which was the wild type of AGR2 protein (AGR2 – WT) as to compare the basal expression of the other three mutated AGR2 protein sequences as they were introduced to the esophageal adenocarcinoma, FLO-1 cancer cell line. Through the immunoblotting that we had conducted on the unexpressed AGR2 FLO-1 cell line, the AGR2 – WT transfected was proved to be localized in the intracellular of the cells,

signified by the presence of bands on the membrane. The expression of the wild type of AGR2 protein (AGR2 – WT) in the extracellular environment was also confirmed through the same procedure such that reduced basal expression signified by low intensity of band could be observed from the membrane. These two findings supported that AGR2 was an ER-residing protein and had dual functions both in the ER (intracellular) and the extracellular matrix (ECM) hence, would be expressed intracellularly and extracellularly just as reported by Delom *et al*. (2018). Such localization of the AGR2 – WT in the intracellular environment of the FLO-1 was possible because the full length of the wild-type AGR2 protein bore the non-canonical KTEL COOH-terminal motif from 172 to 175 aa and this was similar to the KDEL and KVEL motif of the ER therefore, allowing them to bind to the KDEL receptor of the ER just as reported by Gupta *et al*. (2012). This binding, hence, allowed AGR2 protein to play their prominence role in maintaining the homeostasis of the ER or also known as ER proteostasis. Our immunofluorescence microscopy also revealed the intracellular localization of AGR2 – WT signified by the green fluoresce of the FLO-1 cell which was in parallel with our finding from immunoblotting. The localization of the AGR2 – WT in the extracellular environment of the FLO-1 cell on the other hand was possible because retention of AGR2 protein in the cell was reported to be non-associated to the KTEL motif through a study done by Fessart *et al*. (2016). This was in proportional to the AGR2's role in the ER of cancer cells as well as their recently suggested role in the secretome of tumor cells. Therefore, regardless on the presence or absence of KTEL motif, the full length of AGR2 wild type could still be secreted in the extracellular of the FLO-1.

The second variant, AGR2 – KDEL was established by introducing point mutation to the complementary DNA sequence of the AGR2 protein resulting in mutation of the KTEL motif to the KDEL motif in order to enhance AGR2 localization. This variant was constructed based on the study by Gupta *et al*. (2012) which revealed that ER residence

was achieved by proteins with a carboxyl terminal KDEL or KSEL instead of the KTEL motif. The immunoblotting that we conducted revealed that mutating non-canonical KTEL to KDEL resulted in lower expression of the AGR2 – KDEL on both intracellular and extracellular of FLO-1 as compared to the intracellular and extracellular of the FLO-1 transfected to the wild type AGR2 protein. Ideally, mutating the KTEL to KDEL shall resulted in higher expression in the intracellular of FLO-1 cells due to the increase in the retention of AGR2 protein in the ER of the cells. The lower expression that we obtained could be due to the fact that the FLO-1 cell line that we used in this study did not express the AGR2 protein hence, transiently transfecting the AGR2 protein variant with KDEL motif instead of the non-canonical KTEL motif into the FLO-1 cell line might result in different affinities for the KDEL receptors of the ER thus impacting the subcellular localization of the AGR2 protein. The immunofluorescence that we conducted also verified the presence of AGR2 – KDEL in the intracellular of the FLO-1 through the fluoresce cells consequently reflecting our findings from immunoblotting. On the other hand, we determined that the presence of expression for AGR2 – KDEL in the extracellular of FLO-1 just like in the FLO-1 transfected with AGR – WT fully supported the study by Fessart *et al*. (2016) which proved that extracellular function of AGR2 protein was independent of the KTEL motif. Therefore, our study on mutating the KTEL to KDEL showed an in concordance result such that there were expression observed in the extracellular of FLO-1 transfected to AGR2 – KDEL.

Constructing the third variant had been done based on the study by Bergström *et al*. (2014) in which mature AGR2 protein mutated through deletion of signal peptide and KTEL motif could enhance AGR2 secretion into the extracellular environment of the cells. The western blot that we conducted revealed that the mature form of AGR2 protein without the non-canonical KTEL motif was expressed in the intracellular of the FLO-1 but not expressed in the extracellular. This finding was in concordance with our

hypothesis such that the mature form of AGR2 without signal peptide and KTEL motif resulted in half retention of AGR2 protein in the intracellular environment of the FLO-1 hence, the lower intensity of the band. The presence of expression in the intracellular environment was in parallel with the result that we obtained from immunofluorescence such that the FLO-1 transfected to the mAGR2 – ΔKTEL was seen to fluoresce in green. Such expression in the intracellular supported a recent study by Moidu *et al*. (2020) which demonstrated that the mature form of AGR2 protein even with the KTEL motif completely deleted or mutated could cause AGR2 still be secreted at relatively equal or lower levels as compared to the wild-type AGR2 although the former has a significantly higher secretion. Higher expression of the mAGR2 – ΔKTEL on the other hand should be seen in the extracellular of the FLO-1 as compared to the wild-type instead of no expression just as we observed from our immunoblotting. The absence of expression of the mAGR2 – ΔKTEL in the extracellular of the FLO-1 which revealed that the mature form of AGR2 protein without the KTEL retention site was incapable of secreting into the extracellular environment of the esophageal cancer as the esophageal cancer initially did not express any AGR2 protein hence introducing the mature form of AGR2 even without the KTEL motif was insufficient for the cells to build machineries to allow the mutated form of AGR2 protein to be secreted into the extracellular environment of the esophageal adenocarcinoma.

The ΔNterm AGR2 – KDEL variant designed by deleting the N-terminal motif of the AGR2 protein and substitution of KDEL motif to enhance the dimer was conducted based on the study by Bergström *et al*. (2014) on the dimerization of AGR2 protein. Through the western blot that we had conducted, the ΔNterm AGR2 – KDEL variant showed no expression both in the intracellular and extracellular of the FLO-1 cell line. This, however, was not correspond to a previous study which reported that deleting the N-terminal motif of the AGR2 protein and substitution of KDEL motif had expression in

the intracellular of an AGR2 binding protein, EpCAM. The structural function of the dimer motif belonging to the AGR2 protein was found to be more significant in the extracellular environment as reported by Maurel *et al*. (2019), however, in our research we could not prove the expression of such mutated variant of AGR2 (ΔNterm AGR2 – KDEL) as there was no expression found on both membranes such that the mutated variant was not expressed in the intracellular and extracellular of the FLO-1. One possible reason that could explain the absence of expressions of ΔNterm AGR2 – KDEL in both the intracellular and extracellular of the FLO-1 was probably due to the low to no transfection efficiency achieved since our study basically implemented transient transfection mean in studying the basal expression of AGR2 protein in FLO-1. This was as reflected in the graph of ratio of intracellular ΔNterm AGR2 – KDEL with actin as in Figure 4.2 as well as graph of ratio of extracellular ΔNterm AGR2 – KDEL with actin as in Figure 4.3. Our immunofluorescence results also clearly depicted that localization of the ΔNterm AGR2 – KDEL in the FLO-1 cell line could not be proved as none of the FLO-1 cell were found to successfully bind to the AGR2-conjugated fluorophore antibody resulting in no fluoresce cells seen in the cytoplasm of the cells.

Due to unavoidable circumstances as well as limited time allocated for the research, we were not capable of optimizing our transfection procedure and merely followed protocol on the concentration of DNA as well as transfection reagent as suggested by the manufacturers. Alternatively, to enhance the transfection efficiency, we recommended on the cloning of plasmid equipped with GFP tag so that selection of FLO-1 carrying our targeted inserts which was in this case was our AGR2 protein variant could be done before carrying out immunoblotting and immunofluorescence. Single sorting of the cell would also be recommended using single cell flow cytometry (FACS) to further enhance the efficiency of the transfection hence, resulting in better results of immunoblotting and immunofluorescence.

To validate the difference in the expression of our transfected AGR2 protein variants, we conducted statistical test as to compare our transfected cell line with the untransfected cell line. Our statistical test however, showed that the ratios of relative expression for both intracellular and extracellular AGR2 proteins to the protein loading control, β – actin was found to be statistically insignificant (p>0.05). Despite that, we convinced that we had proved the difference in the expression of our transfected AGR2 protein variants from the presence of bands from the immunoblotting as well as immunofluorescence conducted.

In the final part of our research, we aimed to determine the phenotypic effects in terms of proliferation and migration of the cells expressing our AGR2 protein variants on esophageal cancer cells. One of the mean that had been applied in carrying out this objective was via wound healing assay conducted. Such assay basically involved introducing a wound to the cells by scratching the cells and examining the percentage of wound closure at indicated time points. We wanted to study as to whether there was any significant differences in terms of proliferation and migration of the FLO-1 transfected to our four AGR2 protein variants. This was because as suggested by Bergström *et al*. (2014), AGR2 protein was proposed to act as a pro-oncogene and involved in the regulation of cell proliferation. An in vitro study by Verma *et al*. (2012) also revealed that the estrogen-responsive AGR2 as an oncogene that upregulated cell proliferation resulting in the advancement of lobuloalveolar in the mammary gland thus further suggesting pro-proliferative effects of AGR2 in supporting tumorigenesis. The results that we obtained basically revealed that our transfected AGR2 variants, AGR – KDEL and mAGR2 – ΔKTEL enhanced the proliferation and migration of the cells after subjected to scratching as compared to the untransfected cell line as well as AGR2 – WT. This was clearly depicted from the representative images captured at 48 and 96 hours time interval post scratching phase such that the percentage of wound closure was found

to be significantly increased (p<0.05). Despite the significant results that we obtained however, the AGR2 – WT did not show much significant difference in migrating and proliferating as compared to the untransfected FLO-1. This was seen from the 24 and consequently 48 hours' time interval such that the percentage of closure was almost the same as the untransfected cell line. Therefore, from our findings, we suggested that AGR2 – WT did not upregulate the proliferation and migration rate of the FLO-1 cells but in contrast the modified forms of AGR2 protein which were AGR – KDEL and mAGR2 – ΔKTEL significantly upregulated the proliferation and migration of FLO-1 cell line. The results that we obtained also suggested that basal expression of our transfected AGR2 protein in the intracellular as obtained from immunoblotting did not influence the proliferation and migration of the FLO-1 such that high expression did not result in an increase of the proliferation and migration of the FLO-1 in contrary to the finding by Gong *et al*. (2020) which reported that high expression of AGR2 could result in an increase in the cell proliferation, migration and invasion abilities in the endometrial carcinoma.

Then, we interested to find out as to whether transfecting four of our AGR2 constructs, AGR2 – WT, AGR – KDEL, mAGR2 – ΔKTEL and ΔNterm AGR2 – KDEL into the FLO-1 esophageal adenocarcinoma could influence the viability of the cells therefore, directing us to conduct another two functional assays in order to determine the percentage of cell viability via MTT and Alamar Blue assays. Our MTT assay showed that transfecting the FLO-1 cell line with AGR2 – WT, AGR2 – KDEL and mAGR2 – ΔKTEL did not result in much significant difference as compared to the untransfected FLO-1 (p<0.05) suggesting that the four constructed AGR2 protein variants might not influence the viability of the esophageal adenocarcinoma tested. Even so, the increase in cell viability in respect to time signified by higher absorbances recorded supported a study by Liu *et al*. (2018) which revealed that the presence of AGR2 – siRNA group elevated the

sensitivity of cells towards cisplatin. Our Alamar Blue assay showed that there was no significant different (p>0.05) observed on the viability of FLO-1 cell either untransfected or transfected with AGR2 protein variants. As the differences in the viability of FLO-1 either transfected or untransfected as observed from the conduction of MTT and Alamar Blue assays were not distinctive hence, we suggested that our AGR2 constructs did not influence the viability of our esophageal adenocarcinoma cell line.

Suggestions on the future researches would be stressed on optimizing the transfection efficiency through the use of GFP-tag cell line, single cell sorting of the post-transfected cells or creating a stable cell line equipped with existing designated variants of AGR2 protein, the four AGR2 – WT, AGR2 – KDEL, mAGR2 – ΔKTEL and ΔNterm AGR2 – KDEL.

CHAPTER 6: CONCLUSION

The existence of the AGR2 protein in extensive cancer types was predominantly reported to be overexpressed in most of the hormone-dependent types validating the fact that AGR2 protein was a pro-oncogenic which could highly influenced the progression of cancer cells. The structural function of AGR2 protein was scrutinized through designation of four variants in which each variant was established with significant functions as to serve as the comparative tool to the other mutated variant, the role hold by AGR2 – WT, enhancing AGR2 localization in the ER as depicted by AGR2 – KDEL, stimulating AGR2 secretion into the secretome, the role designed for mAGR2 – ΔKTEL as well as enhancing the dimerization motif of the AGR2 protein through designation of ΔNterm AGR2 – KDEL. Intracellular basal expression of the AGR2 – WT, AGR2 – KDEL and mAGR2 – ΔKTEL in the FLO-1 esophageal adenocarcinoma cell line was confirmed with the decreasing of the basal level of expression respectively while the extracellular expression was observed only in the FLO-1 transfected to the AGR2 – WT as well as AGR2 – KDEL, a finding that further validated the dual roles of AGR2 protein in the hallmark of cancer. Half of the mutated variants of the protein studied, the AGR2 – KDEL and mAGR2 – ΔKTEL also uncovered a finding on the upregulation on the proliferation and migration activities of the FLO-1 cell line. Overall, completion of this research enabled us to understand the structural role of AGR2 protein such that difference in AGR2 protein variants constructed resulted in different expression of AGR2 protein in the intracellular and extracellular environment of esophageal adenocarcinoma suggesting that unique motifs presented in the AGR2 protein structure came with prominence role such that altering the motifs could modulate cancer progression. Understanding the role of AGR2 protein allowed for implementation of AGR2 protein to be developed as a diagnostic marker as well as therapeutic target for cancer such that integrating AGR2

protein in biomarkers might contribute to increasing sensitivity and specificity of cancer detection thus useful for early cancer detection and synergistically improve existing treatment of cancers.

REFERENCES

Alnaji, R. M., Du, W., Gabriel, E., Singla, S., Attwood, K., Nava, H., . . . Kukar, M. (2016). Pathologic complete response is an independent predictor of improved survival following neoadjuvant chemoradiation for esophageal adenocarcinoma. *Journal of Gastrointestinal Surgery, 20*(9), 1541-1546.

Armes, J. E., Davies, C. M., Wallace, S., Taheri, T., Perrin, L. C., & Autelitano, D. J. (2013). AGR2 expression in ovarian tumours: a potential biomarker for endometrioid and mucinous differentiation. *Pathology, 45*(1), 49-54.

Arnold, M., Soerjomataram, I., Ferlay, J., & Forman, D. (2015). Global incidence of oesophageal cancer by histological subtype in 2012. *Gut, 64*(3), 381-387.

Arumugam, T., Deng, D., Bover, L., Wang, H., Logsdon, C. D., & Ramachandran, V. (2015). New blocking antibodies against novel AGR2–C4.4A pathway reduce growth and metastasis of pancreatic tumors and increase survival in mice. *Molecular Cancer Therapeutics, 14*(4), 941-951.

Bergström, J. H., Berg, K. A., Rodríguez-Piñeiro, A. M., Stecher, B., Johansson, M. E., & Hansson, G. C. (2014). AGR2, an endoplasmic reticulum protein, is secreted into the gastrointestinal mucus. *PLOS ONE, 9*(8), Article#e104186.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 68*(6), 394-424.

Brychtova, V., Mohtar, A., Vojtesek, B., & Hupp, T. R. (2015). *Mechanisms of anterior gradient-2 regulation and function in cancer.* Presented in Seminars in Cancer Biology, Kuala Lumpur, Malaysia.

Bu, H., Bormann, S., Schäfer, G., Horninger, W., Massoner, P., Neeb, A., . . . Sültmann, H. (2011). The anterior gradient 2 (AGR2) gene is overexpressed in prostate cancer and may be useful as a urine sediment marker for prostate cancer detection. *The Prostate, 71*(6), 575-587.

Bu, H., Schweiger, M. R., Manke, T., Wunderlich, A., Timmermann, B., Kerick, M., . . . Cato, A. C. (2013). Anterior gradient 2 and 3–two prototype androgen-responsive genes transcriptionally upregulated by androgens and by oestrogens in prostate cancer cells. *The FEBS Journal, 280*(5), 1249-1266.

Chanda, D., Lee, J. H., Sawant, A., Hensel, J. A., Isayeva, T., Reilly, S. D., . . . Singh, R. (2014). Anterior gradient protein-2 is a regulator of cellular adhesion in prostate cancer. *PLOS ONE, 9*(2), Article#e89940.

Chevet, E., Fessart, D., Delom, F., Mulot, A., Vojtesek, B., Hrstka, R., . . . Hupp, T. (2013). Emerging roles for the pro-oncogenic anterior gradient-2 in cancer development. *Oncogene, 32*(20), 2499-2509.

Clarke, C., Rudland, P., & Barraclough, R. (2015). The metastasis-inducing protein AGR2 is O-glycosylated upon secretion from mammary epithelial cells. *Molecular and Cellular Biochemistry, 408*(1), 245-252.

Clarke, D. J., Murray, E., Faktor, J., Mohtar, A., Vojtesek, B., MacKay, C. L., . . . Hupp, T. R. (2016). Mass spectrometry analysis of the oxidation states of the pro-oncogenic protein anterior gradient-2 reveals covalent dimerization via an intermolecular disulphide bond. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1864*(5), 551-561.

Consortium, U. (2018). UniProt: the universal protein knowledgebase. *Nucleic Acids Research, 46*(5), 2699.

Delom, F., Mohtar, M. A., Hupp, T., & Fessart, D. (2020). The anterior gradient-2 interactome. *American Journal of Physiology-Cell Physiology, 318*(1), C40-C47.

Delom, F., Nazaraliyev, A., & Fessart, D. (2018). The role of protein disulphide isomerase AGR2 in the tumour niche. *Biology of the Cell, 110*(12), 271-282.

Dong, A., Wodziak, D., & Lowe, A. W. (2015). Epidermal growth factor receptor (EGFR) signaling requires a specific endoplasmic reticulum thioredoxin for the post-translational control of receptor presentation to the cell surface. *Journal of Biological Chemistry, 290*(13), 8016-8027.

Dumartin, L., Alrawashdeh, W., Trabulo, S., Radon, T., Steiger, K., Feakins, R., . . . Lemoine, N. (2017). ER stress protein AGR2 precedes and is involved in the regulation of pancreatic cancer initiation. *Oncogene, 36*(22), 3094-3103.

Dumartin, L., Whiteman, H. J., Weeks, M. E., Hariharan, D., Dmitrovic, B., Iacobuzio-Donahue, C. A., . . . Timms, J. F. (2011). AGR2 is a novel surface antigen that promotes the dissemination of pancreatic cancer cells through regulation of cathepsins B and D. *Cancer Research, 71*(22), 7091-7102.

Dvorak, K., Goldman, A., Kong, J., Lynch, J. P., Hutchinson, L., Houghton, J. M., . . . Westra, W. M. (2011). Molecular mechanisms of Barrett's esophagus and adenocarcinoma. *Annals of the New York Academy of Sciences, 1232*(1), 381-391.

Edgell, T. A., Barraclough, D. L., Rajic, A., Dhulia, J., Lewis, K. J., Armes, J. E., . . . Autelitano, D. J. (2010). Increased plasma concentrations of anterior gradient 2 protein are positively associated with ovarian cancer. *Clinical Science, 118*(12), 717-725.

Fessart, D., Domblides, C., Avril, T., Eriksson, L. A., Begueret, H., Pineau, R., . . . Chevet, E. (2016). Secretion of protein disulphide isomerase AGR2 confers tumorigenic properties. *Elife, 5*, Article#e13887.

Galligan, J. J., & Petersen, D. R. (2012). The human protein disulfide isomerase gene family. *Human Genomics, 6*(1), 1-15.

Genereux, J. C., Qu, S., Zhou, M., Ryno, L. M., Wang, S., Shoulders, M. D., . . . Wiseman, R. L. (2015). Unfolded protein response-induced ER dj3 secretion links ER stress to extracellular proteostasis. *The EMBO Journal, 34*(1), 4-19.

Gong, W., Ekmu, B., Wang, X., Lu, Y., & Wan, L. (2020). AGR2-induced glucose metabolism facilitated the progression of endometrial carcinoma via enhancing the MUC1/HIF-1α pathway. *Human Cell, 33*(3), 790-800.

Grassme, K. S., Garza-Garcia, A., Delgado, J.-P., Godwin, J. W., Kumar, A., Gates, P. B., . . . Brockes, J. P. (2016). Mechanism of action of secreted newt anterior gradient protein. *PLOS ONE, 11*(4), Article#e0154176.

Gray, T. A., Alsamman, K., Murray, E., Sims, A. H., & Hupp, T. R. (2014). Engineering a synthetic cell panel to identify signalling components reprogrammed by the cell growth regulator anterior gradient-2. *Molecular BioSystems, 10*(6), 1409-1425.

Gray, T. A., MacLaine, N. J., Michie, C. O., Bouchalova, P., Murray, E., Howie, J., . . . Vojtesek, B. (2012). Anterior Gradient-3: a novel biomarker for ovarian cancer that mediates cisplatin resistance in xenograft models. *Journal of Immunological Methods, 378*(1-2), 20-32.

Gray, T. A., Murray, E., Nowicki, M. W., Remnant, L., Scherl, A., Muller, P., . . . Hupp, T. R. (2013). Development of a fluorescent monoclonal antibody-based assay to measure the allosteric effects of synthetic peptides on self-oligomerization of AGR2 protein. *Protein Science, 22*(9), 1266-1278.

Guo, H., Chen, H., Zhu, Q., Yu, X., Rong, R., Merugu, S. B., . . . Li, D. (2016). A humanized monoclonal antibody targeting secreted anterior gradient 2 effectively inhibits the xenograft tumor growth. *Biochemical and Biophysical Research Communications, 475*(1), 57-63.

Guo, H., Zhu, Q., Yu, X., Merugu, S., Mangukiya, H., Smith, N., . . . Rong, R. (2017). Tumor-secreted anterior gradient-2 binds to VEGF and FGF2 and enhances their activities by promoting their homodimerization. *Oncogene, 36*(36), 5098-5109.

Gupta, A., Dong, A., & Lowe, A. W. (2012). AGR2 gene function requires a unique endoplasmic reticulum localization motif. *Journal of Biological Chemistry, 287*(7), 4773-4782.

Ho, M. E., Quek, S.-I., True, L. D., Seiler, R., Fleischmann, A., Bagryanova, L., . . . Shimizu, Y. (2016). Bladder cancer cells secrete while normal bladder cells express but do not secrete AGR2. *Oncotarget, 7*(13), 15747.

Hong, X.-Y., Wang, J., & Li, Z. (2013). AGR2 expression is regulated by HIF-1 and contributes to growth and angiogenesis of glioblastoma. *Cell Biochemistry and Biophysics, 67*(3), 1487-1495.

Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., & Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Research, 43*(D1), D512-D520.

Hrstka, R., Brychtova, V., Fabian, P., Vojtesek, B., & Svoboda, M. (2013). AGR2 predicts tamoxifen resistance in postmenopausal breast cancer patients. *Disease Markers, 35*(4), 207-212.

Hsu, H.-Y., Chao, Y.-K., Hsieh, C.-H., Wen, Y.-W., Chang, H.-K., Tseng, C.-K., & Liu, Y.-H. (2016). Postoperative adjuvant therapy improves survival in pathologic nonresponders after neoadjuvant chemoradiation for esophageal squamous cell carcinoma: a propensity-matched analysis. *The Annals of Thoracic Surgery, 102*(5), 1687-1693.

Huang, F.-L., & Yu, S.-J. (2018). Esophageal cancer: risk factors, genetic association, and treatment. *Asian Journal of Surgery, 41*(3), 210-215.

Jia, M., Guo, Y., Zhu, D., Zhang, N., Li, L., Jiang, J., . . . Wang, M. (2018). Pro-metastatic activity of AGR2 interrupts angiogenesis target bevacizumab efficiency via direct interaction with VEGFA and activation of NF-κB pathway. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1864*(5), 1622-1633.

Kani, K., Malihi, P. D., Jiang, Y., Wang, H., Wang, Y., Ruderman, D. L., . . . Gross, M. E. (2013). Anterior gradient 2 (AGR2): blood-based biomarker elevated in metastatic prostate cancer associated with the neuroendocrine phenotype. *The Prostate, 73*(3), 306-315.

Lee, J.-G., Takahama, S., Zhang, G., Tomarev, S. I., & Ye, Y. (2016). Unconventional secretion of misfolded proteins promotes adaptation to proteasome dysfunction in mammalian cells. *Nature Cell Biology, 18*(7), 765-776.

Leung, C. T., & Brugge, J. S. (2012). Outgrowth of single oncogene-expressing cells from suppressive epithelial environments. *Nature, 482*(7385), 410-413.

Li, Z., Zhu, Q., Chen, H., Hu, L., Negi, H., Zheng, Y., . . . Li, D. (2016). Binding of anterior gradient 2 and estrogen receptor-α: Dual critical roles in enhancing fulvestrant resistance and IGF-1-induced tumorigenesis of breast cancer. *Cancer Letters, 377*(1), 32-43.

Li, Z., Zhu, Q., Hu, L., Chen, H., Wu, Z., & Li, D. (2015). Anterior gradient 2 is a binding stabilizer of hypoxia inducible factor-1α that enhances CoCl2-induced doxorubicin resistance in breast cancer cells. *Cancer Science, 106*(8), 1041-1049.

Liu, Q.-G., Li, Y.-J., & Yao, L. (2018). Knockdown of AGR2 induces cell apoptosis and reduces chemotherapy resistance of pancreatic cancer cells with the involvement of ERK/AKT axis. *Pancreatology, 18*(6), 678-688.

Liu, Z., Hu, Y., Gong, Y., Zhang, W., Liu, C., Wang, Q., & Deng, H. (2016). Hydrogen peroxide mediated mitochondrial UNG1-PRDX3 interaction and UNG1 degradation. *Free Radical Biology and Medicine, 99*, 54-62.

Lu, P., Weaver, V. M., & Werb, Z. (2012). The extracellular matrix: a dynamic niche in cancer progression. *Journal of Cell Biology, 196*(4), 395-406.

Ma, S.-R., Wang, W.-M., Huang, C.-F., Zhang, W.-F., & Sun, Z.-J. (2015). Anterior gradient protein 2 expression in high grade head and neck squamous cell carcinoma correlated with cancer stem cell and epithelial mesenchymal transition. *Oncotarget, 6*(11), Article#8807.

Martinez, P., Timmer, M. R., Lau, C. T., Calpe, S., del Carmen Sancho-Serra, M., Straub, D., . . . Mallant-Hent, R. C. (2016). Dynamic clonal equilibrium and predetermined cancer risk in Barrett's oesophagus. *Nature Communications, 7*(1), 1-10.

Maurel, M., Obacz, J., Avril, T., Ding, Y. P., Papadodima, O., Treton, X., . . . Hou, W. (2019). Control of anterior GR adient 2 (AGR 2) dimerization links endoplasmic reticulum proteostasis to inflammation. *EMBO Molecular Medicine, 11*(6), Article#e10120.

Meeusen, E., Lim, E., & Mathivanan, S. (2017). Secreted tumor antigens–immune biomarkers for diagnosis and therapy. *Proteomics, 17*(23-24), Article#1600442.

Merkow, R. P., Bilimoria, K. Y., McCarter, M. D., Chow, W. B., Ko, C. Y., & Bentrem, D. J. (2012). Use of multimodality neoadjuvant therapy for esophageal cancer in the United States: assessment of 987 hospitals. *Annals of Surgical Oncology, 19*(2), 357-364.

Milewski, D., Balli, D., Ustiyan, V., Le, T., Dienemann, H., Warth, A., . . . Kalin, T. V. (2017). FOXM1 activates AGR2 and causes progression of lung adenomas into invasive mucinous adenocarcinomas. *PLoS Genetics, 13*(12), Article#e1007097.

Moidu, N. A., Rahman, N. S. A., Syafruddin, S. E., Low, T. Y., & Mohtar, M. A. (2020). Secretion of pro-oncogenic AGR2 protein in cancer. *Heliyon, 6*(9), Article#e05000.

Müller, L. B., Meurer, L., Lopes, A. B., Antunes, L. C., Vanazzi, S., & Fagundes, R. B. (2014). Stepwise expression of CDKN2A and RB1 proteins in esophageal mucosa from patients at high risk for squamous cell carcinoma. *Applied Immunohistochemistry & Molecular Morphology, 22*(9), 669-673.

Neeb, A., Hefele, S., Bormann, S., Parson, W., Adams, F., Wolf, P., . . . Wilhelm, K. (2014). Splice variant transcripts of the anterior gradient 2 gene as a marker of prostate cancer. *Oncotarget, 5*(18), Article#8681.

Negi, H., Merugu, S. B., Mangukiya, H. B., Li, Z., Zhou, B., Sehar, Q., . . . Li, D. (2019). Anterior gradient-2 monoclonal antibody inhibits lung cancer growth and metastasis by upregulating p53 pathway and without exerting any toxicological effects: a preclinical study. *Cancer Letters, 449*, 125-134.

O'Neill, J. R., Pak, H.-S., Pairo-Castineira, E., Save, V., Paterson-Brown, S., Nenutil, R., . . . Hupp, T. R. (2017). Quantitative shotgun proteomics unveils candidate novel esophageal adenocarcinoma (EAC)-specific proteins. *Molecular & Cellular Proteomics, 16*(6), 1138-1150.

Paltridge, J. L., Belle, L., & Khew-Goodall, Y. (2013). The secretome in cancer progression. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1834*(11), 2233-2241.

Park, K., Chung, Y. J., So, H., Kim, K., Park, J., Oh, M., . . . Choi, Y.-L. (2011). AGR2, a mucinous ovarian cancer marker, promotes cell proliferation and migration. *Experimental & Molecular Medicine, 43*(2), 91-100.

Patel, P., Clarke, C., Barraclough, D. L., Jowitt, T. A., Rudland, P. S., Barraclough, R., & Lian, L.-Y. (2013). Metastasis-promoting anterior gradient 2 protein has a dimeric thioredoxin fold structure and a role in cell adhesion. *Journal of Molecular Biology, 425*(5), 929-943.

Rodríguez-Blanco, G., Zeneyedpour, L., Duijvesz, D., Hoogland, A. M., Verhoef, E. I., Kweldam, C. F., . . . Jenster, G. (2018). Tissue proteomics outlines AGR2 and LOX5 as markers for biochemical recurrence of prostate cancer. *Oncotarget, 9*(92), 36444.

Ryu, J., Park, S. G., Lee, P. Y., Cho, S., Kim, G. H., Kim, J.-H., & Park, B. C. (2013). Dimerization of pro-oncogenic protein anterior gradient 2 is required for the interaction with BiP/GRP78. *Biochemical and Biophysical Research Communications, 430*(2), 610-615.

Schroeder, B. W., Verhaeghe, C., Park, S.-W., Nguyenvu, L. T., Huang, X., Zhen, G., & Erle, D. J. (2012). AGR2 is induced in asthma and promotes allergen-induced mucin overproduction. *American Journal of Respiratory Cell and Molecular Biology, 47*(2), 178-185.

Smyth, E. C., Lagergren, J., Fitzgerald, R. C., Lordick, F., Shah, M. A., Lagergren, P., & Cunningham, D. (2017). Oesophageal cancer. *Nature Reviews Disease Primers, 3*(1), 1-21.

Stachler, M. D., Taylor-Weiner, A., Peng, S., McKenna, A., Agoston, A. T., Odze, R. D., . . . Leshchiner, I. (2015). Paired exome analysis of Barrett's esophagus and adenocarcinoma. *Nature Genetics, 47*(9), 1047-1055.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 71*(3), 209-249.

Thompson, D. A., & Weigel, R. J. (1998). hAG-2, the human homologue of the Xenopus laevis cement gland gene XAG-2, is coexpressed with estrogen receptor in breast cancer cell lines. *Biochemical and Biophysical Research Communications, 251*(1), 111-116.

Tian, S., Hu, J., Tao, K., Wang, J., Chu, Y., Li, J., . . . Li, Q. (2018). Secreted AGR2 promotes invasion of colorectal cancer cells via Wnt11-mediated non-canonical Wnt signaling. *Experimental Cell Research, 364*(2), 198-207.

Tompa, P., Davey, N. E., Gibson, T. J., & Babu, M. M. (2014). A million peptide motifs for the molecular biologist. *Molecular Cell, 55*(2), 161-169.

Uhlenhopp, D. J., Then, E. O., Sunkara, T., & Gaduputi, V. (2020). Epidemiology of esophageal cancer: update in global trends, etiology and risk factors. *Clinical Journal of Gastroenterology*, 1-12.

van Hagen, P., Hulshof, M., Van Lanschot, J., Steyerberg, E., Henegouwen, M. V. B., Wijnhoven, B., . . . Bonenkamp, J. (2012). Preoperative chemoradiotherapy for esophageal or junctional cancer. *New England Journal of Medicine, 366*(22), 2074-2084.

Verma, S., Salmans, M. L., Geyfman, M., Wang, H., Yu, Z., Lu, Z., . . . Andersen, B. (2012). The estrogen-responsive Agr2 gene regulates mammary epithelial proliferation and facilitates lobuloalveolar development. *Developmental Biology, 369*(2), 249-260.

Vitello, E. A., Quek, S.-I., Kincaid, H., Fuchs, T., Crichton, D. J., Troisch, P., & Liu, A. Y. (2016). Cancer-secreted AGR2 induces programmed cell death in normal cells. *Oncotarget, 7*(31), Article#49425.

Wang, D., Xu, Q., Yuan, Q., Jia, M., Niu, H., Liu, X., . . . Yuan, H. (2019). Proteasome inhibition boosts autophagic degradation of ubiquitinated-AGR2 and enhances the antitumor efficiency of bevacizumab. *Oncogene, 38*(18), 3458-3474.

Wang, D. H., Tiwari, A., Kim, M. E., Clemons, N. J., Regmi, N. L., Hodges, W. A., . . . Zhang, X. (2014). Hedgehog signaling regulates FOXA2 in esophageal embryogenesis and Barrett's metaplasia. *The Journal of Clinical Investigation, 124*(9), 3767-3780.

Wang, D. H., Tiwari, A., Kim, M. E., Clemons, N. J., Regmi, N. L., Hodges, W. A., . . . Souza, R. F. (2014). Hedgehog signaling regulates FOXA2 in esophageal embryogenesis and Barrett's metaplasia. *The Journal of Clinical Investigation, 124*(9), 3767-3780.

Wayner, E. A., Quek, S. I., Ahmad, R., Ho, M. E., Loprieno, M. A., Zhou, Y., . . . Liu, A. Y. (2012). Development of an ELISA to detect the secreted prostate cancer biomarker AGR2 in voided urine. *The Prostate, 72*(9), 1023-1034.

Weaver, J. M., Ross-Innes, C. S., Shannon, N., Lynch, A. G., Forshew, T., Barbera, M., . . . Dunning, M. J. (2014). Ordering of mutations in preinvasive disease stages of esophageal carcinogenesis. *Nature Genetics, 46*(8), 837-843.

Whitwell, H. J., Worthington, J., Blyuss, O., Gentry-Maharaj, A., Ryan, A., Gunu, R., . . . Zaikin, A. (2020). Improved early detection of ovarian cancer using longitudinal multimarker models. *British Journal of Cancer, 122*(6), 847-856.

Wong, H. H., & Chu, P. (2012). Immunohistochemical features of the gastrointestinal tract tumors. *Journal of Gastrointestinal Oncology, 3*(3), Article#262.

Wu, J., Wang, C., Li, X., Song, Y., Wang, W., Li, C., . . . Yang, C. J. (2012). Identification, characterization and application of a g-quadruplex structured DNA aptamer against cancer biomarker protein anterior gradient homolog 2. *PLOS ONE, 7*(9), Article#e46393.

Xue, X., Fei, X., Hou, W., Zhang, Y., Liu, L., & Hu, R. (2018). miR-342-3p suppresses cell proliferation and migration by targeting AGR2 in non-small cell lung cancer. *Cancer Letters, 412*, 170-178.

Yu, H., Zhao, J., Lin, L., Zhang, Y., Zhong, F., Liu, Y., . . . He, F. (2012). Proteomic study explores AGR2 as pro-metastatic protein in HCC. *Molecular BioSystems, 8*(10), 2710-2718.

Zhang, J., Jin, Y., Xu, S., Zheng, J., Zhang, Q., Wang, Y., . . . Zhao, Z. (2016). AGR2 is associated with gastric cancer progression and poor survival. *Oncology Letters, 11*(3), 2075-2083.

Zhu, Q., Mangukiya, H. B., Mashausi, D. S., Guo, H., Negi, H., Merugu, S. B., . . . Li, D. (2017). Anterior gradient 2 is induced in cutaneous wound and promotes wound healing through its adhesion domain. *The FEBS Journal, 284*(17), 2856-2869.