UNIVERSITY OF LIVERPOOL

Characterising the Role of Anterior Gradient Protein 2 in Cell Adhesion and Metastasis

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Christopher John Clarke

September 2014

Acknowledgements

I would like to thank my supervisors Dr Roger Barraclough, Prof Philip Rudland and Dr Dong Barraclough for all their help and support throughout this project. I would also like to acknowledge the help of several people, all of whom are from the Institute of Integrative Biology, University of Liverpool, unless otherwise stated. I would like to acknowledge Dr Pryank Patel for his work on the AGR2 structure and Dr Marie Phelan for her help with structural analysis, Marjorie Howard and Dr Thomas Jowitt (both of the Univeristy of Manchester) for performing the SEC-MALLS experiments presented herein, Dr Deborah Simpsons for running the mass spectrometry analyses for this project, Dr Olga Vasieva for her help with Ingenuity Pathway Analysis, Prof Rob Beynon for his help and advice with the mass spectrometry experiments, and Dr Ed Yates, Prof Brian Campbell and Dr Mark Wilkinson for their help and advice in determining the post-translational modification of AGR2.

Abstract

Anterior gradient protein 2 (AGR2) is overexpressed in many cancers, in particular those of the breast, ovary, prostate and pancreas. This overexpression tends to correlate with a poor disease prognosis for patients, probably as a result of increased incidence of metastases, and indeed, AGR2 has been shown to play a role in cell adhesion in different cell lines. Whilst AGR2 is a member of the protein disulphide isomerase (PDI) family of endoplasmic reticulum (ER) chaperones, there are also multiple reports of the secretion of AGR2 by cultured cells and the protein is also found in the serum and urine of cancer patients. AGR2 is a dimer, and the dimerisation interface is characterised here. A short, flexible region of the protein is also described, that seems to be involved in the regulation of the adhesion activity of AGR2. Through the use of stable cell lines expressing wild type and several mutant AGR2 proteins, it is shown here that wild-type AGR2 expression is associated with changes in cell adhesion and anchorage-independent cell growth, that are lost upon inhibition of protein dimerisation, mutation of the putative active site or a change in the sub-cellular localisation of the AGR2. Furthermore, the secretion of AGR2 has been confirmed through the discovery of a glycosylated, extracellular form of the protein. The expression of AGR2 is shown to be associated with a number of changes to the secreted proteome, with effects on both pro-adhesive and anti-adhesive proteins that may contribute to the role of AGR2 in cell adhesion. Overall, through combined biochemical, cell biology and mass spectrometry approaches, potential mechanisms whereby AGR2 might contribute to cell adhesion and metastasis are explored and discussed.

Contents

List of figures

List of tables

Abbreviations

Å	angstrom
AGR2	anterior gradient protein 2
APS	ammonium persulfate
AREG	amphiregulin
BCA	bicinchoninic acid
Bcl-2	B-cell lymphoma-2
bp	base pair
BSA	bovine serum albumin
CD	cluster of differentiation
CHO	chinese hamster ovary
CM	conditioned medium
CMV	Cytomegalovirus
CRABP	cellular retinoic acid-binding protein
CV	column volume
DAPI	4',6-diamidino-2-phenylindole
dH_2O	distilled water
DMEM	Dulbecco's modified Eagle's medium
DSB	disulphide bond
DTT	dithiothreitol
E-cadherin	epithelial cadherin
ECL	enhanced chemiluminescence
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
EF1α	elongation factor 1α
EGFR	epithelial growth factor receptor
ELISA	enzyme-linked immunosorbant assay
EMT	epithelial-mesenchymal transition
ER	endoplasmic reticulum
ERK	extracellular signal-regulated kinase
ERp*X*	endoplasmic reticulum protein *X* kDa
EV	empty vector
FA	focal adhesion

FAK	focal adhesion kinase
FBS	foetal bovine serum
FERM	4.1 protein (F), ezrin (E), radixin (R) moesin (M)
FGF	fibroblast growth factor
FPLC	fast protein liquid chromatography
HED	2-hydroxyethyl disulphide
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HER2	human epidermal growth factor receptor 2
HIF	hypoxia-inducible factor
HMG	high mobility box
HSP	heat shock protein
IF	immunofluorescence
Ig	immunoglobulin
IL	interleukin
IMPACT-TWIN	Intein Mediated Purification with an Affinity Chitin-binding Tag-Two Intein
IP	immunoprecipitation
IPA	Ingenuity Pathway Analysis
IPTG	isopropyl β-D-1-thiogalactopyranoside
kDa	kilo Dalton
LB	Luria-Bertani
LC-MS	liquid chromatography-mass spectrometry
LDHA	lactate dehydrogenase A
MAPK	mitogen-activated protein kinase
MCF7A	Michigan Cancer Foundation 7 adherent
MEGF10	multiple epidermal growth factor-like domains protein 10
MEK1	MAPK/ERK kinase 1
MEM	Minimum essential medium
MET	mesenchymal-epithelial transition
MHC	major histocompatibility complex
MUC	mucin
N-cadherin	neural cadherin
NCAM	neural cell adhesion molecule
NEAA	non-essential amino acids

NLS	nuclear localisation sequence
N-LS	N-lauroylsarcosine
NoLS	nucleolar localisation sequence
OD	optical density
OR	oestrogen receptor
PAI-1	plasminogen activator inhibitor-1
PAS	periodic acid-Schiff
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDGF	platelet-derived growth factor
PDI	protein disulphide isomerase
PFA	paraformaldehyde
PI3K	phosphoinositide 3-kinase
PIP_2	phosphatidylinositol 4,5-bisphosphate
PTM	post-translational modification
PVDF	polyvinylidene difluoride
R37E	Rama 37-elongated
Rama 37	rat mammary 37
Rap1	repressor/activator protein 1
RIPA	radio immunoprecipitation assay
RLU	relative light units
RO	reverse osmosis
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC-MALLS	size-exclusion chromatography-multi angle laser-light scatterring
SMAD	mothers against decapentaplegic
SOC	super optimal broth with catabolite repression
Sox	SRY-related HMG box
SR	shape ratio
SRY	sex-determining region Y
TAE	Tris-acetate-EDTA
TBS	Tris-buffered saline
TBS-T	TBS-Tween

TCA	trichloroacetic acid
TEMED	tetramethylethylenediamine
TFA	trifluoroacetic acid
TGFβ	transforming growth factor β
TMX	thioredoxin-related transmembrane protein
tPA	tissue plasminogen activator
uPAR	urokinase-type plasminogen activator receptor
v/v	volume/volume
VEGF	vascular endothelial growth factor
w/v	weight/volume
WISP2	Wnt-inducible secreted protein 2
WT	wild type
YAP1	yes-associated protein 1
ZEB	Zinc-finger E-box-binding

Chapter 1

Introduction

1.1 Cancer metastasis

Metastatic spread of tumour cells is one of the final hallmarks of cancer described by Hanahan and Weinberg [1]. It represents a defining step in the oncogenic process that is responsible for the majority of patient deaths, sometimes accounting for 90 % of fatalities [2-4]. The metastatic cascade is a highly complex process, and it is this complexity that presents such a major challenge for the treatment of metastatic disease, and thus understanding the different mechanisms that underpin metastasis is of great importance to the advancement of disease treatment.

1.1.1 Events in the metastatic cascade

Normal cells must undergo a number of changes to become tumourigenic [1], and in the same way tumour cells must generally undergo several changes to become metastatic. Changes in cell adhesion play a key role in metastasis, as cells must release cell-cell and cell-matrix adhesions in order to break free from the main tumour mass, transmigrate into blood or lymph vessels and then adhere to the vessel endothelium at secondary sites, in order to extravasate and establish new tumours [5]. However, to fully understand this process, the factors contributing to each step of this cascade must be considered.

1.1.1.1 Loss of cell-cell adhesion

Cell-cell contacts are maintained by three classes of intercellular adhesion complexes: adherens junctions, desmosomes and tight junctions [6]. Adherens (Fig. 1.1) are made up of two families of proteins, cadherins and catenins [7]. The transmembrane cadherins (including E-, N-, P- and R-cadherins [8]) form the anchoring interactions of adherens between adjacent cells, and they interact in a homotypic and calcium-dependent fashion [9, 10]. The cytoplasmic catenin proteins (α-, β- and p120-catenin) provide a link between the cadherin molecules and the cytoskeleton and also perform a number of signalling roles. p120-catenin in particular, which binds to the membrane adjacent region of the cadherin cytoplasmic domain [11], is thought to stabilise the adherens junction, possibly through prevention of cadherin internalisation [11-13]. p120-catenin is also required to promote cadherin-based adhesion [14, 15], and is an important regulator of cell migration through regulation of Rac (Ras-related C3 botulinum toxin substrate) and Cdc42 (cell division control homolog 42), members of the Rho GTPase family [16]. Binding of β-catenin to both α-catenin and to cadherins is mediated by phosphorylation [17, 18], and this positional regulation is important given the function of non-adherens-bound β-catenin as a co-activator of the T cell factor/lymphoid enhancer factor (TCF/LEF) transcription factor in the Wnt pathway [19].

α-catenin is classically thought of as the connection between β-catenin and the cadherin molecule, due to its ability to bind both β-catenin [20] and actin bundles [21]. However, its role is probably more complex, and it may be that binding-induced clustering of cadherins increases the local concentration of α-catenin, which is then able to dimerise

and inhibit lamellipodia formation, a crucial membrane structure involved in migration [22] (see Section 1.1.1.3). In this way α-catenin promotes a non-motile phenotype, required for the maintenance of epithelial linings.

Desmosomes (Fig. 1.1) are largely responsible for coupling intercellular adhesion molecules to the intracellular intermediate filament system, thus providing resistance to mechanical stresses [23, 24]. Like cadherins in adherens junctions, the desmosomal cadherins (desmocollins 1-3 and desmogleins 1-4) are transmembrane proteins, but they exist as desmocollin-desmoglein heterodimers. These heterodimers interact in largely homotypic fashion, in that the desmocollin molecule from one hereodimer interacts with the desmocolin molecule in the opposing heterodimer, with the same being true for the desmoglein molecules [25-27] (Fig. 1.1). Desmoplakin binds intermediate filaments and links them to desmosomal cadherins [23], and is essential for desmosome function, as demonstrated by the complete loss of skin integrity by its deletion [28]. Plakoglobin (sometimes called γ-catenin) and plakophilins 1-3 bind the desmosomal cadherins themselves and also desmoplakin, thus completing the connection between intermediate filaments and intercellular junctions [23]. Plakoglobin is also thought to play a role in crosstalk between desmosomes and adherens [29, 30], and can also influence the Wnt pathway in a similar manner to β-catenin [31].

Tight junctions (Fig. 1.1) are important structures for maintaining cell polarity by forming a physical barrier that separates apical and basolateral membrane proteins and prevents them from mixing [32]. They are composed of the transmembrane occludin, claudin, tricellulin and junction adhesion molecule (JAM) proteins. The claudin molecules play an important role in the second function of tight junctions, that of controlling the passage of ions between adjacent cells [33]. In the cytoplasm, the zonula occludens (ZO) proteins 1-3 interact with the cytoplasmic tails of occludin [34-36], and act as scaffold and signalling proteins, as well as providing a link to the actin cytoskeleton [32, 34, 37]. These ZO proteins are also known to play roles in cell proliferation [38] and probably also migration [39].

Downregulation of cell-cell adhesion molecules is a hallmark of cells undergoing epithelial-mesenchymal transition (EMT), the process whereby, in cancer, epithelial cells gain a more migratory and invasive phenotype that allows them to spread and metastasise (see Section 1.2). Loss of cadherins not only results in loss of anchoring adhesion between epithelial layers, but liberates β-catenin from the junction complex, which can then promote transcription of pro-EMT genes under the influence of Wnt signalling [40, 41]. A similar picture exists for desmosomes, where loss of intracellular desmosome components

such as desmoplakin and plakophilins leads to increase tumour cell invasion and metastasis [24, 42] by similarly promoting Wnt/β-catenin signalling [43]. Loss of occludin and claudin from tight junctions is also evident in metastatic cells [44-47], where loss of tight junction integrity leads to loss of cell polarity and may also promote proliferation, through loss of ZO-mediated sequestration of pro-proliferative transcription factors (reviewed in [48]). Thus, maintaining the integrity of cell-cell junctions plays an important role in preventing tumourigenesis.

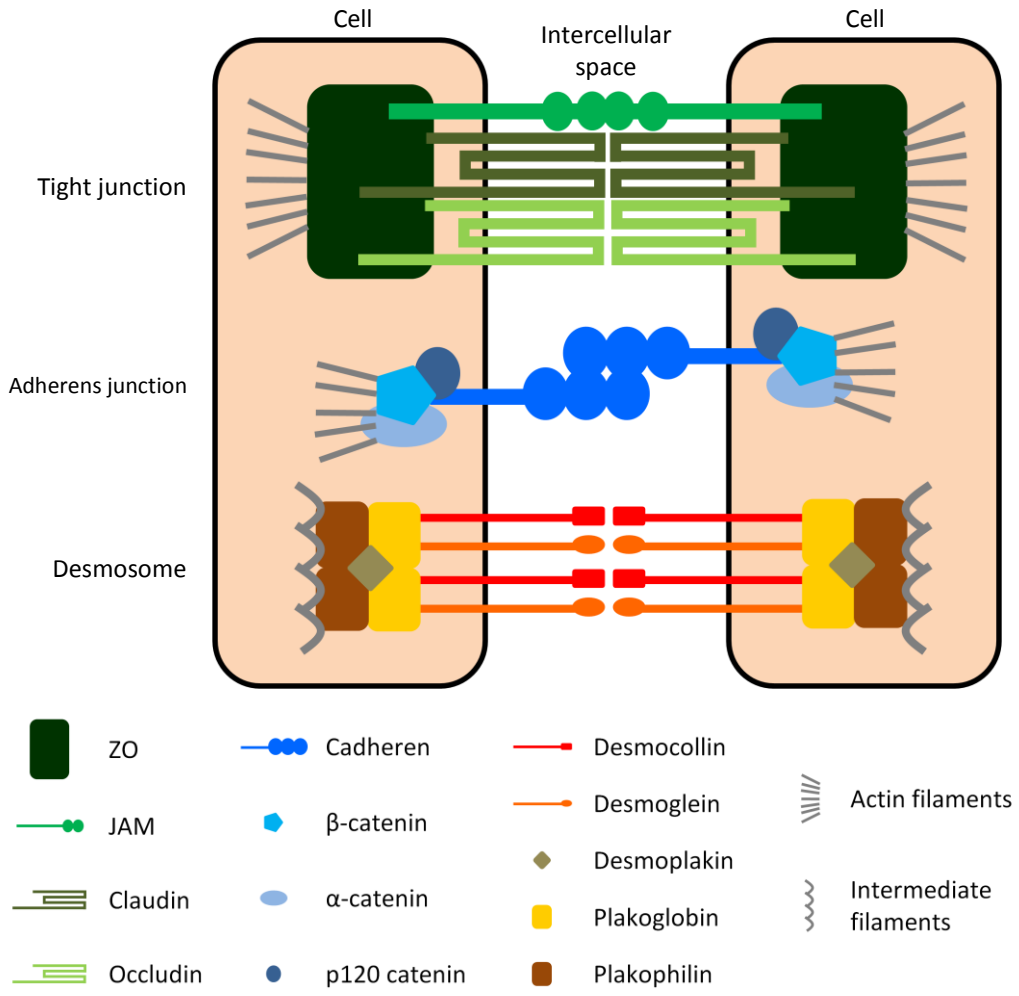


Figure 1.1. ***Epithelial intercellular junctions.*** The architecture of representative tight, adherens and desmosome junctions are shown. ZO: zonula occludens. JAM: junction adhesion molecule. Figure adapted from Knights *et al.,* 2012 [6].

1.1.1.2 Altered cell-matrix adhesion

Interactions between cells and their surrounding matrix are responsible for not only the spatial and structural arrangement of cells within tissues, but are also required for a plethora of cell processes, including proliferation, survival and differentiation [49-51]. The integrin family of cell surface receptors play an integral part in this process, mediating the attachment of cells to various extracellular matrix (ECM) ligands. Integrin molecules are heterodimeric complexes of an α and β subunit [52], of which there exists 18 α subunits and 8 β subunits that can combine to form 24 different integrin receptors (Fig. 1.2), that show different tissue distribution and different ligand specificities [53, 54]. This broad ligand binding has been attributed to the relatively short ligand recognition sites of integrins [55], the most widely studied of which is the RGD motif (consisting of Arg-Gly-Asp residues), which is found in a number of ECM components, namely fibronectin, vitronectin, fibrinogen and von Willebrand factor [56].

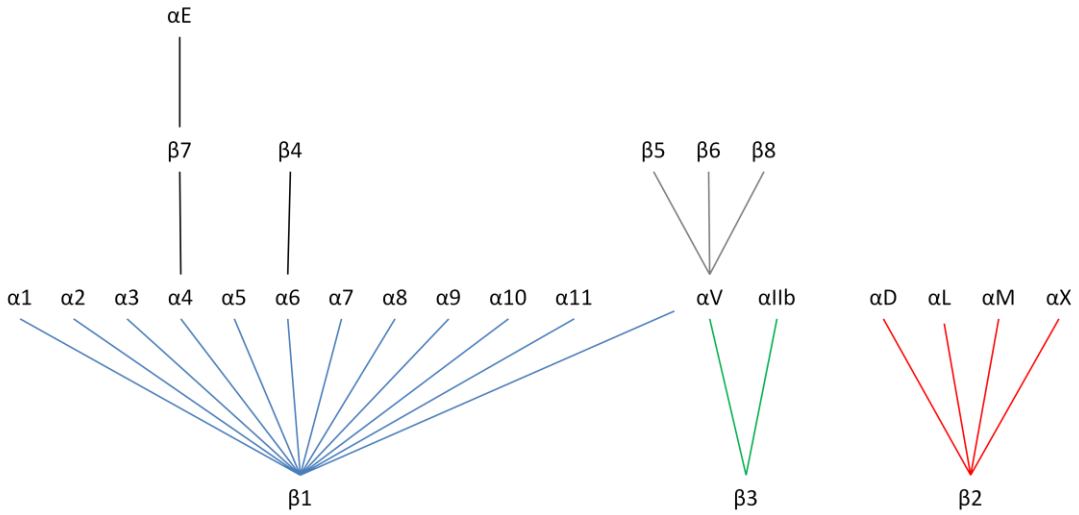


Figure 1.2. ***Members of the vertebrate integrin family*****.** The 24 possible combinations of α and β subunits are displayed, with different colours representing groups containing an identical subunit. Figure adapted from Bouvard *et al.,* 2001 [57].

Structurally, integrins contain large globular extracellular domains, coupled to a single-pass transmembrane region and a short cytoplasmic tail, usually of approximately 60 residues [52, 55, 58]. Integrins can exist in two different states: a bent, inactive or low-

affinity state, and in an upright, active or high affinity state [50, 55]. Switching from an inactive to active state is initiated by the binding of extracellular ligands (termed 'outside-in' signalling) [55] or through binding of cytoplasmic adapter proteins such as talin (known as 'inside-out' signalling) [59-61]. The talin proteins (talin-1 and -2) play a major role in 'inside-out' signalling, along with the related kindlin proteins (kindlin-1 to -3). Talin is a large cytoskeletal protein containing an N-terminal FERM (4.1 protein (F), ezrin (E), radixin (R) moesin (M)) domain linked to a C-terminal rod-like domain [62]. Within the FERM domain sits the F3 domain that mediates talin binding to the integrin β subunit cytoplasmic tail which is responsible for talin-mediated activation of integrins [63-65]. The mechanism of activation relates to an interaction between the talin F3 domain and the juxtamembrane region of the β subunit cytoplasmic tail upon talin binding, which causes a conformational change in the β subunit, causing it to extend away from the α subunit [66]. As association of the α and β subunits maintains the integrin in an inactive conformation [67], their talin-induced disassociation leads to receptor activation. Binding of the F3 domain also disrupts a salt bridge between the α and β subunits, further promoting their disassociation [68].

Several pathways can promote the talin-mediated activation of integrins. Non-integrin-bound talin largely exists in the cytoplasm in an auto-inhibited form, whereby the crucial F3 domain is bound to a region in the C-terminal rod domain [69-71]. Phosphatidylinositol 4,5-bisphosphate (PIP_2), most likely generated through PIP-kinase type 1γ [72], can interrupt this inhibitory binding and promote recruitment of talin to the membrane and subsequent integrin activation through the F2 region in the FERM domain [68, 69]. The generation of PIP_2 is also implicated in the Rap1A (repressor/activator protein 1A) GTPase 'inside-out' pathway of integrin activation. Recruitment of Rap1A to the plasma membrane, probably through PIP_2, in turn recruits Rap1-GTP-interacting adaptor molecule (RIAM) to membrane-bound Rap1A [73, 74]. RIAM contains an N-terminal talin-binding domain, which promotes talin recruitment to the membrane, leading to integrin activation [74, 75]. A third pathway of intracellular integrin activation relates to the focal adhesion kinase (FAK)-mediated recruitment of talin to the membrane [76]. This strays from the classical role of post-talin-recruited FAK (see Section 1.1.1.3), but is only an emerging concept, and as such little is known about the mechanism involved, except that this pathway may be regulated through p190RhoGEF (guanine nucleotide exchange factor), which has been shown to bind to FAK [77].

The second class of proteins involved in 'inside-out' signalling, the kindlins, are not as well characterised as the talins, but are known to contain a similar FERM domain to talin [78], containing the F3 domain that mediates kindlin binding to the integrin β subunit tail

[79-81]. *In vitro* at least, a ternary complex is formed between kindlin, the talin head FERM domain and integrin β tail [82, 83], but kindlin binding to integrins does not seem to promote recruitment of talin, nor enhance talin-integrin binding [83, 84]. However, kindlins are required for the activation of β1 integrins at least, as their activation is impaired by kindlin knockdown [85-87]. Kindlins have therefore been suggested to act as talin co-activators, influencing events that occur post-talin binding to integrins [84]. The nature of this effect is not clear, but may relate to further dissociation of the α and β transmembrane domains, through simultaneous kindlin binding to the integrin β subunit tail and membrane lipids [88-90]. Another proposed function of kindlins is displacement of integrin β subunit-bound inhibitory proteins. Evidence for this stems from the overlapping binding sites on integrin β subunits for kindlins and the inhibitory proteins filamin and ICAP1 (integrin cytoplasmic domain–associated protein 1) [91, 92]. In this way, talins and kindlins act together in the activation of integrin receptors.

During the metastatic cascade, tumour cells tend to undergo a switching of integrin receptors, where receptors mediating adhesion of epithelial cells to the basement membrane are downregulated (e.g. α6β4 [93]) and β1 integrins tend to be upregulated, promoting the disruption of E-cadherin junctions and thus downregulating cell-cell contacts [94, 95]. Expression of fibronectin-binding integrins such as α5β1 in transformed cells promotes cell migration, as fibronectin also tends to be upregulated by these cells [96, 97]. Similarly, non-β1 integrins are also displayed by tumour cells, which can also promote fibronectin-dependent migration (e.g. αvβ6 [98]), but this switching of integrins can also promote changes in adhesion related to EMT. Using several cancer cell lines, including lung, liver and colon, both the aforementioned αvβ6 and also αvβ8 were shown to promote the release of active TGFβ (a major inducer of EMT (see Section 1.2)) from its latent extracellular form [99, 100]. TGFβ is secreted in an inactive form bound to both the latency-associated peptide (LAP) and the latent-TGFβ-binding protein (LTBP) (reviewed in [101]). The LAP- TGFβ interaction is non-covalent, and TGFβ can be released through proteolytic degradation of LAP, or through integrin binding to the LAP, which is thought to cause a conformational change in the LAP, resulting in TGFβ release (reviewed in [101]). αvβ6 was shown to release TGFβ in a protease-independent manner [99], while the action of αvβ8 was protease dependent [100], but both led to a local increase of active TGFβ. Thus integrin switching in tumour cells can perpetuate changes to cell adhesion.

Additionally, the role of integrins in oncogenesis can also involve their intracellular signalling functions. This is illustrated by integrin α6β4, which is often overexpressed in a number of cancers [102] and activation of which promotes cell proliferation [103]. Its

cytoplasmic domain also acts as a scaffold for hepatocyte growth factor (HGF)-mediated signalling, amplifying pro-invasive signals [104], and α6β4 integrin also similarly signals through Rac to promote invasion [105]. Integrin α6β4-mediated activation of NF-κB also promotes cell survival [106, 107].

Thus, changes to integrin receptors during tumourigenesis do not simply result in loss or gain of adhesion to a specific substrate, but additionally affect many other pathways that contribute to the overall metastatic process. Like changes in cell-cell adhesion, alterations to cell-matrix adhesions also promote EMT, thus coupling loss of cell adhesion at the primary tumour with increased migratory potential.

1.1.1.3 Cell migration

The formation of focal adhesions (FAs), the structures through which cells generate traction forces to allow cell migration, progresses through defined stages [52] and is dependent on both integrin-mediated adhesions to the ECM and also the flow of actin into and out of cell membrane protrusions, namely lamellipodia [108]. These membrane projections are driven by polymerisation of branched actin filaments at the cytosolic face of the membrane [109], and this polymerisation of actin at the leading edge of lamellipodia is also coupled to a rearward, retrograde flow of actin filaments mediated by both the force exerted by actin on the cell membrane and myosin II [110-112]. Initial or nascent focal adhesions are formed within the lamellipodia [108, 113], where talin is then recruited to stabilise the interaction between the clustered integrins and the ECM [114]. Recruitment of talin also starts the FA maturation process, initially preventing the breakdown of the nascent FA by slowing the rate of retrograde actin flow, which otherwise causes the disassembly of a proportion of the nascent FAs [108]. This stabilisation is important given the dynamic nature of FA assembly and disassembly, where substrate stiffness and mechanical load on the individual FA contribute to their turnover [115-117].

Clustered integrins can exert a small amount of traction force on the matrix and at the same time, the retrograde actin flow pulls on the integrin-bound talin, in turn pulling on β1 integrin subunit, which leads to unfolding of the extracellularly-bound fibronectin [118]. This results not only in the strengthening of the interaction between talin and actin [119], but the stretching of talin reveals binding sites for other focal adhesion proteins, including vinculin, FAK and paxilllin, which promote the maturation of the nascent FA [120-123]. In particular, vinculin enhances the binding of actin to FAs and promotes clustering of ECM-

bound integrins [124, 125], and is also involved in the regulation of contractility [126]. FAK couples talin binding to cell cycle progression [127], but also promotes the eventual disassembly of the FA, which allows the cell to migrate [76, 127].

During FA maturation, the contractile force provided by myosin II contributes to the maturation of the FA through stretching of the bound talin and influences on FAK signalling [121, 128, 129]. Myosin contraction also pulls clustered integrins inwards [130], which, along with other mechanical stresses, also serves to promote the maturation of the filamentous actin network into higher order structures known a stress fibres [131-133]. These stress fibres are composed of 10 to 30 actin filaments which are bundled together through crosslinking by α-actinin [134]. This generates small bundles of actins that are then joined end-to-end through myosin II [133]. Two broad types of stress fibre are formed: transverse fibres that tend to form to the rear of the leading edge [135-137] and are at first not tethered to FAs [137], and dorsal and ventral stress fibres that are formed at FA sites through the actions of RhoA [138], the downstream effector ROCK (Rho-associated protein kinase) and the formin mDia1 (reviewed in [139]). The transverse and dorsal/ventral fibres eventually fuse and provide contractile forces [133, 140], but the mechanism involved in their assembly is largely unknown.

Further maturation of FAs is characterised by the gradual recruitment of increasing levels of zyxin [141] and also by its recruitment to stress fibres [142]. The size of the FA, which grows over time under the influence of contractile and mechanical stresses, at this point reaches its peak and they no longer provide great traction forces [141, 143]. The role of these mature FAs evolves into one of cell anchoring [144, 145], maturing further into fibrillar adhesion which are characterised by the presence of tensin, lack of vinculin and replacement of stress fibre interaction back to interactions with thinner actin filaments [146-149]. They also play a role in the organisation and modulation of extracellular fibronectin [149].

Directional migration mediated by FAs necessitates the disassembly of FAs (particularly at the trailing edge of the cell), but the processes involved in disassembly are less well understood than those involved in assembly. The assembly and disassembly of FAs at the leading edge is largely regulated by the forces generated from polymerising and de-polymerising actin [108, 150]. In terms of disassembly of mature FAs, one major regulatory mechanism involves their removal by endocytosis [151, 152], but microtubules are also implicated in their disassembly, by causing relaxation of the actin filaments around FAs, destabilising them and aiding in their endocytosis [152-154]. Src and FAK are important mediators of FA disassembly [146, 155]. They probably increase myosin II contraction

through activation of mitogen-activated protein kinase (MAPK), which in turn activates myosin light chain kinase (MLCK) [156], which is then able to phosphorylate myosin light chain and promote actomyosin contraction [157]. This contractility mediates the recruitment of endocytic machinery and the calpain-2 protease [158], which breaks down talin and FAK, severing the link between integrins and actin and promoting their turnover [159, 160]. Endocytosed integrins can be degraded but are more often recycled back to the leading edge (reviewed in [161]), where they can again bind extracellular ligands [162].

1.1.1.4 Angiogenesis, intra- and extravasation

Cell-cell and cell-matrix interactions regulate the tethering to, but also migration and invasion through, the ECM (as outlined above). In order to disseminate and metastasise however, tumours tend to engage the formation of new blood vessels (angiogenesis), through which they can then gain access to a host of secondary colonising sites. The initiation of angiogenesis is related to the dysregulated growth of the primary tumour, where tumours reaching a certain size become hypoxic [163]. This generates hypoxia inducible factors (HIFs), which in turn promote the production of members of the vascular endothelial growth factor (VEGF) family [164], the best characterised inducers of angiogenesis [165]. The release of VEGFs (VEGF-A, -B, -C, D, and –E) is also brought about by their proteolytic release from the ECM by MMP9 [166]. Their effects are mainly mediated on the vessel endothelium [167], promoting proliferation, migration and invasion of endothelial cells ([168, 169]. Interestingly, autocrine production of VEGF also promotes EMT, linking the vascularisation of the tumour with the induction of a more migratory and metastatic cell type [170].

A number of other secreted and membrane-bound ligands contribute to the induction of angiogenesis, including the fibroblast growth factor (FGF) family, the platelet-derived growth factor (PDGF) family the angiopoietin family [171]. FGFs tend to promote endothelial cell proliferation and migration [172], whilst PDGF and the angiopoietins are important for the recruitment of pericytes and vascular smooth muscle cells, which promote vessel maturation [173-175]. The angiopoeitins are also involved in the recruitment of macrophages to the site of angiogenesis [176, 177], further promoting vessel development (see below).

Release of TGFβ is also involved in angiogenesis, although it appears that low levels promote angiogenesis, whereas high levels slow down the process and promote vessel

maturation [178]. Similarly, the transmembrane Delta/Jagged and Notch pathways dampen down angiogenesis through modulation of VEGF activity and limit the levels of vessel branching [179]. The CXC chemokines can also be pro-angiogenic (e.g. CXCL1 and CXCL2) or anti-angiogenic (e.g. CXCL4), through the induction or repression of endothelial cell proliferation and chemotaxis, respectively (reviewed in [180]).

The release of VEGF and other growth factors and cytokines from tumour cells also allows the recruitment of hematopoietic cells, which support the formation of new vessels [181, 182]. These tumour-associated macrophages and neutrophils provide a ready-supply of pro-angiogenic factors such as VEGF and FGF and promote matrix degradation, thus favouring invasion of endothelial cells towards the tumour (reviewed in [183] and [184]). They also manipulate and downregulate the anti-tumour response. Cancer-associated fibroblasts are also recruited to sites of angiogenesis, where they provide not only pro-angiogenic factors such as VEGF [185], but crucially remodel the ECM, without-which the vessel lumens fail to form [186].

1.1.1.4.1 Intravasation

The vessels formed by tumour angiogenesis are very leaky, due to the defects in or absence of endothelial basement membrane, abnormal growth and shape of endothelial cells and increased space between cell junctions [187, 188]. These poorly-formed vessels favour the intravasation of migrating tumour cells into these vessels [189]. This is potentiated by the local release of factors such as TGFβ and VEGF, which further promote vessel permeability through breakdown of endothelial cell junctions, thus creating openings between endothelial cells [190-192].

Intravasation is regulated by similar factors governing migration and ECM degradation, such as N-WASP [193], MMT4-MMP [194] and ADAM12 [195]. Both MMT4-MMP and ADAM12 act to disrupt vessel integrity, with ADAM12 appearing to cleave vascular endothelial cadherin, thus weakening the lateral attachments between endothelial cells [194, 195]. Notch signalling also appears to be involved in the intravasation process, as suppression of Notch signalling prevents intravasation, at least in colon cancer [196]. Induction of endothelial cell apoptosis by tumour cells also contributes to enhanced vascular permeability [190].

1.1.1.4.2 Extravasation

Once within the vessel, tumour cells tend to associate with platelets and coagulation factors such as fibrin and thrombin, which serve to protect the cell from the shear stress of blood flow but also affect its arrest at secondary sites [197, 198]. These secondary sites tend to be in small capillaries and are thought to result from the size-restriction imposed on the tumour cell by the size of the capillary [199-201]. This arrest allows adhesion interactions to take place between the tumour cell and capillary endothelium [190, 202]. These initial interactions are mediated by endothelial (E)-selectin binding of ligands expressed on the tumour cell, notably haematopoietic cell E-selectin/L-selectin ligand (HCELL), P-selectin glycoprotein ligand 1 (PSGL1), mucin 1, CD24 and galectin-3-binding protein (LGALS3BP) [190, 202, 203]. N-cadherin may also be important in these initial interactions [202, 204].

More stable interactions are subsequently formed between the tumour cell and endothelium, mainly mediated by CD44 [205, 206] and a number of integrins (reviewed in [207]). Endothelium-bound tumour cells are then able to extravasate from the vessel via the same mechanisms used to intravasate, but must overcome one final hurdle, that of the endothelial cell basement membrane [207]. Presumably this is achieved through the release of matrix-degrading proteins, although this has not been shown experimentally.

Opposing changes in cell adhesion thus seem to contribute to intra- and extravasation. On the one hand, cell must become less adherent in order to leave the primary tumour site, but must then be more adhesive in order to bind to the vessel endothelium or risk never exiting the bloodstream. However, these processes appear to be regulated by different classes of adhesion molecules; cadherins and integrins mediate the former, while the latter is regulated by glycoproteins and lectins. Thus, in the context of tumour dissemination, the favourable loss of cell-cell and cell-ECM adhesions need not impact on the requirement to form adhesion-based interactions with the vessel endothelium that allows colonisation of secondary sites.

1.2 Epithelial-mesenchymal transition

As hinted at above, many of the processes that favour metastasis, especially those relating to adhesion and migration, are linked with epithelial-mesenchymal transition, a characteristic feature of cancers arising from epithelial tissues where there is a loss of epithelial-like characteristics and the acquisition of mesenchymal-like properties [208-210]. These changes are mediated by a number of pathways and culminate, at least in the context of carcinogenesis, in the production of highly motile and invasive cells [208, 209]. EMT is not necessarily a pathological process however, and occurs in embryo and tissue development as well as in wound healing [211]. It is also becoming clear that EMT is a plastic process which can be reversible, with cells also able to undergo only partial EMT [210].

1.2.1 Induction of cellular changes in EMT

As mentioned in section 1.1.1.1, intercellular adhesion complexes are involved in maintaining epithelial cell polarity. Polarisation of the apical compartment is achieved through interactions of cell junctions with two complexes: the Crumbs complex (consisting of Crumbs protein, PALS1 (protein associated with Lin-7 1) and PATJ (PALS1-associated tight-junction protein) and partitioning-defective (PAR) complex (comprising PAR3, PAR6 and atypical protein kinase C (aPKC)) [212]. Similarly, the basolateral compartment contains the Scribble complex (Scribble, Discs large (DLG) and lethal giant larvae (LGL)) [212]. When EMT is initiated, there is a downregulation of cell junction proteins which disrupts the interactions with the cell polarity complexes, ultimately leading to a loss of cell polarity [213, 214]. Thus the loss of cell-cell adhesions promotes loss of cell polarity, and the loss of E-cadherin in particular (which is cleaved and degraded [214]) results in release of β-catenin and p120 catenin, which are further involved in driving the EMT process through their effects on gene transcription [215, 216] (see below). The loss of cell polarity has knock-on effects on actin cytoskeleton dynamics, allowing increased migration through the generation of increased numbers of lamellipodia [209, 214, 217].

The migratory behaviour of cells undergoing EMT is typically further modulated by the ‘cadherin-switch’, where the repression of E-cadherin is coupled with the expression of neural (N)-cadherin [214, 218]. As described in Section 1.1.1.1, cadherins regulate cell-cell

adhesion, and as N-cadherin is also expressed on mesenchymal cells, cell-cell interactions of cells undergoing EMT with the mesenchyme are promoted, favouring migration away from the epithelial lining [210]. Furthermore, these homotypic N-cadherin interactions are weaker than homotypic E-cadherin interactions, further favouring migration and invasion of the surrounding stroma [219]. Importantly, N-cadherin expression also induces expression of neural cell adhesion molecule (NCAM) [214, 220], through which focal adhesion assembly is promoted, thus further promoting cell migration [221]. The loss of E-cadherin is further potentiated by changes to the make-up of intermediate filaments, with the upregulation of vimentin at the expense of cytokeratins, which results in a loss of cytokeratin-mediated transport of E-cadherin to the membrane [213, 222]. Vimentin-mediated interactions with motor proteins may also promote an increase in cell migration [223]. Also, as outlined in section 1.1.1.2, EMT also drives alterations to cell-ECM adhesion, promoting a switch from basement membrane-binding integrins such as α6β4 [93] to fibronectin-binding integrins such as α5β1 [96, 97], thus favouring detachment from the basement membrane and driving migration.

1.2.2 Inducers of EMT

1.2.2.1 TGFβ

Both TGFβ and bone morphogenic protein (BMP) ligands bind to and exert their actions through two TGF receptors, TGFRI and TGFRII [224, 225]. Binding of TGF ligands to TGFRII promotes its association with TGFRI, leading to the modulation of the activity of the SMAD (mothers against decapentaplegic) family of intracellular signalling molecules, which can be activating or inhibitory (reviewed in [226]). The major effects of TGFβ signalling are mediated by the SMAD-dependent upregulation of a number of transcription factors that regulate the expression of EMT genes, including Snail1 and 2 [227, 228], ZEB1 (Zinc-finger E-box-binding 1) [229] and Twist1 [230] (see below). Some genes are upregulated by SMADs without the need for EMT transcription factors, including fibronectin and vimentin [231-233].

TGFβ also has SMAD-independent functions, signalling through phosphoinositide 3-kinase (PI3K)/Akt, RhoA GTPase and MAPK pathways to promote EMT [234]. EMT is prevented by pharmacological inhibition of PI3K, suggesting that it plays an important role in relaying pro-EMT signals [235]. Similarly, inhibition of the extracellular signal-regulated

kinase (ERK) and p38 MAPK pathways inhibits the TGFβ-mediated EMT, and the c-Jun N-terminal kinases (JNK) pathway is also thought to be involved [234, 236, 237]. The effects of these pathways on EMT seem to revolve around increasing Snail activity and promoting the expression of N-cadherin and MMPs, as well as the repression of E-cadherin, although their exact roles are not yet clear [210, 238-241].

Another SMAD-independent mechanism of TGFβ-induced EMT involves RhoA, where effects are mediated in two opposing ways, related to the involvement of RhoA in both cell-cell junction and lamellipodia formation [242, 243]. In the initial stages of EMT, TGFβ promotes the SMURF1 (SMAD ubiquitination regulatory factor 1)-dependent ubiquitination and degradation of RhoA, promoting the breakdown of cell-cell junctions [244] whereas in the later stages of EMT, TGFβ signalling events promote RhoA activation and lamellipodia formation, thus promoting migration [245]. The mechanism involved in switching between RhoA degradation and RhoA activation is not clear, however.

It is interesting that TGFβ -SMAD signalling mediates non-pathological EMT during organ development [234, 246], and in normal cells it generally has growth inhibitory and pro-apoptotic effects, and can thus be considered a tumour-suppressive pathway [234, 247]. Thus in tumourigenesis, it appears that crosstalk between the TGFβ pathway with other pathways such as MAPK, PI3K and RhoA (as outlined above), effectively suppress the growth-inhibitory functions of TGFβ [245, 247, 248], presumably by overriding the cytostatic TGFβ effects through their own overwhelming pro-proliferative and anti-apoptotic signals, but also through mediating nuclear translocation of SMADs through phosphorylation [249]. In this way, EMT regulatory function of TGFβ signalling remains in the absence of its cytostatic effects, promoting tumour cell EMT [247, 248].

1.2.2.2 Growth factor receptors

In addition to the TGFβ-mediated activation of PI3K/Akt and MAPK pathways described above, the activation of these pathways by growth factor receptors can also induce EMT. Here, oncogenic activation of Raf and Ras seem to be important, promoting Snail1 and 2 expression and Rho activation [241, 250]. FGF1 can induce EMT in tumour cells through expression of Snail2, dissolution of desmosomes and expression of pro-migratory integrins and MMPs [251-253]. Similarly, PDGF promotes cell junction breakdown and E-cadherin downregulation in tumour cells [254], whilst VEGF can also promote EMT through Snail and possibly also Twist [254, 255]. In turn, Snail also upregulates VEGF expression

[256]. Overall, involvement of FGFs, PDGF and VEGF in both EMT and angiogenesis provides a useful mechanism in the context of tumour development, where the induction a cell type with altered adhesion properties and a more migratory phenotype is coupled to a means of disseminating to secondary sites.

1.2.2.3 Induction by developmental pathways

Developmental pathways can promote EMT in tumour cells, probably as a result of their involvement in physiological EMT mechanisms. The Wnt pathway, which is involved in embryo formation [257], induces EMT in tumour cells through β-catenin-mediated gene expression and also by increasing the stability of Snail1 [215, 258]. Likewise, the Delta-like/Jagged-Notch pathway enhances Snail2 expression, which has been shown to contribute to lung cancer invasion [259, 260].

1.2.3 Transcriptional regulation of EMT

As alluded to above, inducers of EMT converge on the modulation of several key transcription factors that drive the cellular changes associated with EMT. These transcription factor families, namely of the Snail family (Snail1 and 2), ZEB family (ZEB1 and 2) and the basic helix-loop-helix (bHLH) family (including Twist1 and 2), not only control each other's expression but also co-operate with each other to promote the expression of EMT genes, whilst at the same time repressing epithelial gene transcription [261, 262]. It is these transcription factors that can directly change the adhesion properties of cells that are characteristic of EMT, repressing many intercellular adhesion molecules such as E-cadherin, occludin, claudin, desmoplakin and plakophilin, whilst at the same time promoting expression of N-cadherin and fibronectin, amongst others (reviewed in [210]).

1.2.4 MicroRNA regulation of EMT

EMT is not only controlled at the protein level and the role of microRNAs in the control of EMT is becoming ever more apparent. Many of the major transcription factors described above are also regulated at the mRNA level through the actions of microRNAs,

including Snail 1 and 2, ZEB 1 and 2 and Twist, as is the expression of epithelial junction proteins such as E- and N-cadherin (reviewed in [263]). It is therefore through the balance of these microRNAs and the protein factors outlined above that progression or inhibition of EMT is controlled.

1.2.5 Mesenchymal-epithelial transition (MET)

There is increasing evidence to suggest that metastatic cells that have undergone EMT to migrate and colonise secondary site subsequently undergo the reverse process of MET, in order to establish tumours at these secondary sites. This stems from several observations (reviewed in [264]), including that metastases tend to contain cells with epithelial-like phenotypes [265-267] and that in animal models, epithelial-type cells rather than EMT-type cells were better able to form metastases [268, 269]. The mechanism of induction of MET is not yet clear, but it has been suggested that it might occur due to the loss of pro-EMT signals at the metastatic site, thus reverting these cells back to their 'default' epithelial phenotype [270, 271], or potentially through more pro-MET-type signals originating from the presumably non-transformed stroma at the secondary site, as evidenced by the apparent differentiation of prostate cancer cells cultured in the presence of normal hepatocytes [272]. An element of the 'seed and soil' hypothesis might also be involved here, where the secondary target site must provide a suitable microenvironment (i.e. soil) to support the induction of MET in the metastatic cell (i.e. the seed), to allow tumour growth and establishment of the secondary tumour [273, 274]. Regardless of the mechanism of induction, it appears that, in the opposite manner to EMT, the MET process is characterised by the re-acquisition of E-cadherin expression [275], and thus presumably more epithelial-like adhesion properties. The regulation of MET does appear to be similar to EMT however, in that it is associated with downregulation of Snail, ZEB and Twist proteins (reviewed in [264]), suggesting that the transcriptional activity, or loss thereof, of these transcription factors are important mediators of MET.

1.2.6 EMT and anoikis

Anoikis is a form of apoptosis that occurs when cells detach and lose contacts with the ECM [276]. As EMT promotes migration of cells away from their tissue of origin and

hence causes disruption to the ECM-cell adhesion contacts, resistance to anoikis is necessarily linked to the EMT process, and is a required step in the metastatic cascade [277]. Like classical apoptosis, anoikis is activated through both intrinsic and extrinsic pathways that converge on the activation of caspases [278], but a caspase-independent pathway also exists (see below).

1.2.6.1 Intrinsic pathway of anoikis

The intrinsic pathway is regulated by members of the Bcl-2 (B-cell lymphoma-2) family, which lie downstream of the PI3K/Akt and Ras pathways regulated by FAK and Src, including Mcl-1 (induced myeloid leukemia cell differentiation protein 1) and Bcl-$X_{L,}$ as well as caveolin-1 (Cav-1) and 14-3-3ζ (reviewed in [279]). Mcl-1 and Bcl-X_L both block the release of cytochrome c from the mitochondrial membrane, promoting cell survival [280] and overexpression of both is associated with increased resistance to apoptosis and decreased patient survival [281, 282]. The overexpression of Cav-1 has been reported in lung cancer, where it probably promotes the activity of Mcl-1 by preventing its degradation [283]. On the other hand, 14-3-3ζ is thought to contribute to anoikis-resistance through blocking pro-apoptotic signal from Bad [284] and p53 [285]. The involvement of 14-3-3ζ in anoikis-resistance is particularly interesting, as it is associated with increased potential for metastasis [286], whilst also increasing the risk of tumour recurrence [287].

1.2.6.2 Extrinsic pathway of anoikis

The extrinsic apoptotic pathway of anoikis is activated by ligation of Fas ligand with death receptors on the cell surface [288]. Active Fas ligand can be bound to the ECM [289] and thus cells migrating through the matrix presumably risk activation of the extrinsic apoptotic pathway. Resistance to the extrinsic pathway is thought to be mediated by cFLIP (cellular FLICE (FADD (Fas-associated death domain)-like IL-1β-converting enzyme)-inhibitory protein), whose overexpression is associated with increased metastasis and can also confer anoikis-resistance [290, 291]. It functions through competitive inhibition of death receptor-induced procaspase-8 activation, thus preventing the initiation of the caspase cascade and cell death [292].

1.2.6.3 Caspase-independent pathway of anoikis

A third pathway of apoptosis is regulated by the actions of the mitochondrial protein Bit-1 (Bcl-2 inhibitor of transcription 1). Upon disengagement of integrin-ECM interactions, Bit-1 translocates to the cytoplasm and causes the relocalisation of TLE1 (transducin-like enhancer of Split 1) from the nucleus to the cytoplasm and its subsequent proteasomal degradation [293, 294]. TLE1 is a repressor of endonuclease expression, and its export from the nucleus promotes endonuclease expression and subsequent cell death [295]. Reduction in cytoplasmic Bit-1 promotes tumour progression in breast cancer and appears to promote anoikis resistance in breast cancer cell lines [293, 296].

It is unclear whether one of these three pathways of anoikis in particular is more important for preventing apoptosis of metastatic cells over any other, and further investigation will be required to determine whether they occur together or separately. Regardless of the mechanism involved, it is clear that, like many other regulatory processes, anoikis is an obstacle that must be overcome by tumour cells in order to metastasise. In addition to the pathways described above, oncogenic activation of intracellular signalling pathways such as PI3K, ERK and RhoA, also promote resistance to anoikis, by providing pro-survival signals that can 'override' the pro-apoptotic signals from loss of cell adhesion [269]. In this way, the pro-EMT signals provided through these same signalling pathways (see Section 1.2.2.2) also promote the survival of these metastatic cells, as only anoikis-resistant cells will be able to colonise secondary sites [297]. In this way, the transformation and metastatic spread of tumour cells is linked by the crosstalk between core intracellular signalling pathways that have multiple and complementary outcomes on tumour cell fate, which eventually lead to the promotion of tumour cell dissemination.

1.3 Protein folding and secretion

The processes identified so far that contribute to metastasis often involve dynamic changes in protein abundance, and thus must necessarily involve regulation of the protein folding machinery, as well as the regulation of the secretory pathway in the context of extracellular and cell surface proteins such as TGFβ and integrins. As regulators of this protein homeostasis, it is evident that protein chaperones can be exploited in the metastatic process to increase the output of correctly folded proteins that promote various aspects of tumour progression. Indeed, protein chaperones have been directly linked to oncogenesis, where inhibition of heat shock protein 70 (HSP70) stimulates cancer cell apoptosis [298] and HSP90 expression is associated with neoplasia and poor prognosis of breast and gastric cancers [299]. It is also interesting that several HSP90 substrates include known oncogenes such as Abl, Src and Akt [300], and the chaperone AGR2 (anterior gradient protein 2) is also associated with metastasis [301-309]. Thus the control of protein folding is an important aspect of tumour biology.

1.3.1 Protein disulphide isomerases and protein folding within the endoplasmic reticulum

Many proteins require the assistance of protein folding chaperones, in particular proteins destined for secretion or the cell surface, which also contain several modifications that are mediated by enzymes and chaperones within the ER and Golgi [310]. The assumption of correct disulphide bonds within these proteins is often the rate-limiting step in the folding process within the ER [311] and it is estimated that approximately 30 % of proteins travelling through the ER contain at least one disulphide bond [312]. Disulphide bonds are important structural determinants, as they limit the possible structural conformations of a given protein due to their covalent nature [313]. Formation of disulphide bonds in ER proteins is regulated by members of the protein disulphide isomerase (PDI) family of chaperones, which also catalyse their reduction and isomerisation. PDI itself is the archetypal member of the enzyme family, and was the first enzyme involved in protein folding to be discovered [314]. Structurally, the 20 PDI family members identified to date have a similar feature, in that they contain a thioredoxin-like fold, based on a βαβαβαββα secondary structure arrangement, but otherwise their size and domain structure varies quite considerably [315, 316]. In general, PDI proteins tend to

contain at least one catalytic domain or 'a' domain, and can contain one or more catalytically-inactive or 'b' domains [315] (Fig.1.3). These b domains tend to be involved in substrate binding or recruitment of co-factors and as such have only limited sequence and structural homology [312, 315].

The PDI a domains house the active site, with the consensus sequence CxxC [312, 315]. In general, the N-terminal Cys in this sequence forms mixed disulphides with the substrate protein, as its thiol group tends to become deprotonated at physiological pH, whilst the C-terminal Cys is required for subsequent release of the substrate [312, 317]. PDIs containing only a single Cys in the active site can also form mixed disulphides with target proteins, as exemplified by PDILT (protein disulphide isomerise-like protein of the testis) and ERp44 [318-320]. However, it appears that at least two CxxC-containing catalytic domains are required for efficient isomerisation reactions [321, 322]. Interestingly, PDIs need not be catalytically active towards thiol groups to be denoted as PDI family members, as demonstrated by two family members containing thioredoxin-like folds but no active site Cys residues, ERp27 and ERp29 [323, 324]. These proteins are obviously not redox-active and are therefore probably involved in a different type of chaperone activity than the CxxC and CxxS PDI proteins. Classification as a PDI therefore relates more to the presence of a structurally similar thioredoxin fold than a necessarily redox-active domain.

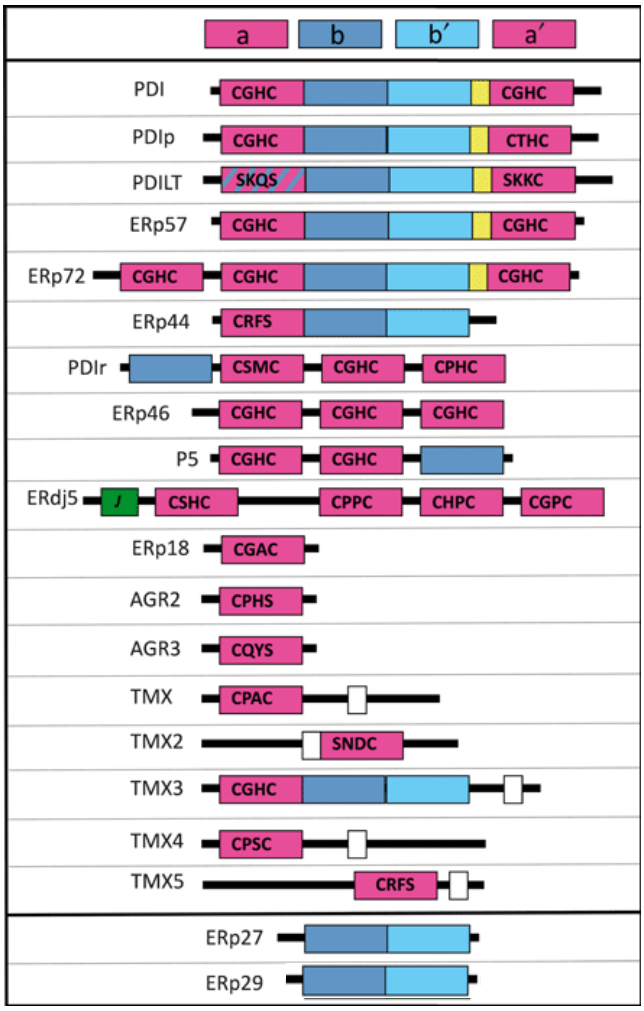


Figure 1.3. ***Overview of PDI family domain structures.*** a domains are shaded in purple and b domains in blue. a' and b' denotes their positions in the sequence. Yellow boxes represent linker regions, which probably increase the flexibility between a and b domains, and the green box denotes the ERdj5 DnaJ domain, a protein interaction domain. White boxes denote transmembrane regions. The hashed PDILT domain denotes its similarity to other a domains. Figure adapted from Kozlov *et al.,* 2010 [315].

1.3.2 Intrinsically disordered regions in protein chaperones

Intrinsically disordered regions (IDRs) are flexible sections of proteins that have no defined 3-dimensional conformation, at least in the absence of interactions with other proteins [325, 326]. These IDRs tend to contain a high proportion of polar amino acids, whereas hydrophobic residues are underrepresented, thus favouring the solvent exposure

of these regions over their internalisation towards the protein hydrophobic core [327]. Importantly, IDRs of 30 residues in length or more are reported to be present in over a third of all eukaryotic proteins, pointing to the possible global importance of these regions [328].

The functions of these IDRs are wide-ranging, including roles in signal transduction, regulation of subunit assembly, and transcription and translation [325, 329-331]. This functional flexibility stems from their structural flexibility, which provides a freedom of movement lacking in more structured regions of proteins, allowing increased sampling of the environment, which leads to increased numbers of interactions. These linear, unfolded regions are also more accessible to other molecules than tightly folded structured domains, providing larger interaction surfaces [332]. This trade off between increased chances of interactions at the expense of rigid interaction domains creates IDR-mediated interactions that are weak but specific, ideal for regulatory-based functions [332]. These interactions should allow rapid but specific activation and deactivation of interaction events, and are thought to be mediated by binding-coupled folding, where the unstructured region assumes a defined fold upon binding to a target protein [326, 333].

The characteristics of IDRs are also particularly suited to the roles of protein chaperones. One study into the prevalence of IDRs in RNA and protein chaperones found that, in the dataset used, around 80 % of protein chaperones possessed an IDR of at least 30 residues in length, and 50 % had an IDR of at least 40 residue in length [334]. In contrast to RNA chaperones however, longer IDRs of more than 40 residues in length seem to appear less frequently in protein chaperones, possibly reflecting the need for an overall more ordered structure that still remains pliable, that presumably must accommodate a bulkier protein substrate than the less bulky RNA [334, 335]. The prevalence of IDRs in chaperone molecules probably stems from the requirement to constantly probe the surrounding environment for the presence of misfolded proteins through the process of ‘fly-casting’ [333] and also the need for chaperones to bind multiple substrate molecules, most likely in different folded states, where a highly structured region would not allow this range of conformational substrate interactions [334]. A further advantage of substrate binding through IDRs is that binding tends to prevent aggregation of substrate proteins due to the general polar nature of these regions that aid in protein solubilisation [333, 336].

The importance of IDRs in chaperone activity is highlighted by the RNA chaperone α-crystallin, where mutations reducing the flexibility of its IDR reduce its chaperone activity, showing a functional link between IDRs and chaperone function [337]. Similarly, the importance of IDRs in the contribution to disease states was highlighted in a study of missense mutations in the UniProt database, where over 20 % of all mutations were found

within IDRs [338]. Interestingly, 20 % of these IDR mutations resulted in a shift towards a more ordered structure, providing evidence that it is the inherent flexibility of these regions that contribute significantly to IDR-containing protein function.

1.3.3 ER-Golgi transport, protein glycosylation and secretion

1.3.3.1 ER-Golgi transport

The transport of proteins destined for secretion or the cell membrane is mediated by the COPII (coat protein II) complex (reviewed in [339]). COPII-coated vesicles capture and deliver protein cargo to the *cis*-Golgi network, but a retrograde transport system from the Golgi back to the ER also exists for the retrieval of ER proteins that have 'escaped' to the Golgi [339, 340]. This retrograde transport is mediated by COPI-coated vesicles [339, 340], but the cargo itself is recognised by the ERD2 (ER retention-defective complementation group 2) family of receptors in the *cis*-Golgi [341]. In humans this family is made up by ERD21, 22 and 23 [342], and these receptors recognise cargo for retrograde transport through binding to C-terminal HDEL motifs (and derivatives thereof) in relevant ER-luminal proteins and K(X)KXX and RXR motifs in transmembrane ER proteins [339, 342]. The ERD2 receptors appear to bind directly to the COPI coat to promote their transport back to the ER [341, 343], where release of the cargo protein from the ERD2 receptors appears to be mediated by the rise in pH in the ER [344].

1.3.3.2 N-linked glycosylation

N-linked glycosylation occurs co-translationally in the ER, where a 14-moiety core glycan is added to target Asn residues by the OST (oligosaccharyl-transferase) complex, after which trimming occurs by glucosidases and mannosidases (reviewed in [345]). Although these 'high mannose' N-glycans can continue unchanged along the secretory pathway, they are not often found in differentiated cell types, but are prevalent in stem cells [346, 347]. Instead, these glycan chains are usually edited as the protein travels through the Golgi network. The glycans are first trimmed by the actions of Golgi-mannosidases (mannosidase IA, IB, IC and II), which generate structures containing either 5 or 3 mannose residues, which in turn affect the later glycan processing steps [347, 348].

The trimmed glycans then become substrates for glycosyltransferases (GTs), which catalyse the addition of N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), galactose, fucose and sialic acid residues, leading to the formation of complex, branched glycan chains [347, 349].

GTs are located throughout the Golgi, from the *cis*- to *trans*-face [347, 350] and in general, the *cis*-compartment contains GTs involved in initiating O-glycosylation (see below) (as well as the N-glycan mannosidases), the medial-Golgi contains GTs for branching both N- and O-glycans, and in the *trans*-compartment and *trans*-Golgi network, GTs cap the glycan chains through the addition of terminal galactose, sialic acid or fucose residues [347, 349, 351]. However, the story in reality is much more complex, as there are an estimated 200 to 250 GTs residing in the Golgi and their localisations appear to overlap or change between cell types [352-354]. Thus the regulation of glycan maturation is still not totally clear, but probably involves recognition of specific structures for individual GTs, the sequestration of different GTs into different Golgi cisternae and sub-compartments, as well as complex COP-mediated intra-Golgi transport of GTs [349, 352, 355]. Another interesting regulatory mechanisms probably stems from gradual acidification of the Golgi from *cis* to *trans* [356], as Golgi pH has been shown to affect both the activity and distribution of Golgi GTs [357, 358].

1.3.3.3 O-linked glycosylation

Several types of O-glycosylation of Ser and Thr residues exist, including O-fucosylation, O-galactosylation, O-glucosylation and O-mannosylation, but the majority of O-glycosylations is made up by mucin-type O-glycans [347, 359]. As the name suggest, this type of glycosylation is predominant in mucin proteins, where extensive glycosylation allows them to form mucus gels which protect epithelial linings (reviewed in [360]). Mucin-type O-glycosylation is initiated by the transfer of a GalNAc residue onto Ser or Thr residues within the Golgi [359, 361], although studies have shown that the addition of GalNAc can be initiated in the ER during oncogenesis [362, 363]. An important feature of mucin-type O-glycosylation is that there exists 15 or more GalNAc GTs that catalyse the initial transfer of the GalNAc residue, whereas other types of O-glycosylation use just one or two enzymes, and N-glycosylation uses the OST complex [345, 347, 359]. The purpose of so many GTs is not clear, though it might provide functional redundancy within the system, or could be the cause of, or result of, the lack of observed consensus sequence for O-glycosylation, such

that different GTs are required to recognise different substrates [347, 359, 364]. The addition of the GalNAc residues then allows the extension and branching of the glycan chain through the actions of other Golgi GTs as described above for N-glycosylation.

1.3.3.4 Protein secretion

Secretory and plasma membrane proteins that have travelled through the Golgi eventually make their way to the *trans*-Golgi network (TGN), a tubular membrane network that acts as the major sorting point for proteins exiting the Golgi (reviewed in [365]). The TGN provides the vesicles required for transport to the plasma membrane, and loaded TGN vesicles are transported to the plasma membrane along microtubules, where they encounter the exocyst complex. This octameric complex is involved in several vesicle trafficking processes, but in the context of secretion it promotes the tethering of the vesicle to the plasma membrane and the subsequent release of the cargo (reviewed in [366]). Interestingly, the exocyst complex also appears to be involved in the regulation of membrane protrusions and cell migration, possibly through promoting the generation of branched actin networks [367-369] (see Section 1.1.1.3). Furthermore, the exocyst is found at E-cadherin-containing cell junctions in normal epithelia [370, 371] and is required for E-cadherin delivery to the membrane in *Drosophila* [372]. However, the exocyst is redistributed to sites of membrane protrusions upon dissolution of cell-cell junctions in rat prostate cancer cells, where it is associated with adhesion complexes [373]. It was also shown to regulate the delivery of β1 integrins to the leading edge of migrating MDA-MB-231 and HeLa cells [374]. The exocyst complex may play an important role in cell-cell adhesion, cell-ECM adhesion and the metastatic spread of tumours.

1.4 Anterior Gradient protein 2

First identified in a screen of oestrogen responsive genes in the MCF7A mammary adenocarcinoma cell line [375] and then shortly afterwards as a protein expressed in murine goblet cells in the intestinal tract [376], anterior gradient protein 2 (AGR2) is the human homologue of the *Xenopus laevis* XAG-2 (*Xenopus* anterior gradient 2) protein. XAG-2 functions along with XAG-1 to regulate the formation of the mucus-secreting cement gland during *Xenopus* embryogenesis [377]. The human homologue of XAG-1 is ERp18/AGR1 [378], but apart from displaying PDI activity, its specific function is unknown [379, 380]. The AGR family is completed by AGR3 [381-383], which, in ovarian cancer, is associated with differentiation and also appears to mediate cisplatin resistance, but its function in normal tissue is largely unknown [384-386].

1.4.1 Regulation of AGR2 expression

The AGR2 gene is situated on chromosome 7p21.3 [387]. Two mRNA transcripts have been reported for AGR2, but the functional consequences of the transcript containing an additional 5' exon are not known [384, 388]. Oestrogen was one of the first regulators of AGR2 gene expression identified, in the original discovery of AGR2 by Thompson and colleagues [375]. Here they found that AGR2 was expressed in the oestrogen receptor (OR)-positive MCF7A mammary cell line but absent in the OR-negative mammary cell line MDA-MB-231. Several studies have since shown an association of AGR2 expression with OR in breast cancer [309, 389, 390], and its expression has similarly been shown to be regulated by androgens in prostate cancer [309, 388, 390, 391].

Other important regulators of AGR2 expression have also been identified in a variety of experimental systems, including the forkhead box transcription factors Foxa1, Foxa2, Foxp1 and Foxp4 [305, 384, 392-394], as well as hepatic nuclear factor 1 which belongs to the same transcription factor family [384]. The Foxa family is a key regulator of the development of organs arising from the endoderm which include the lung, prostate, pancreas and liver (reviewed in [395]), where knockout of Foxa2 is embryonically lethal [396] and Foxa1 knockout results in early neonatal death [397]. Foxp proteins appear to be involved in the regulation of heart, lung, gut and oesophageal development, where again, loss of Foxp1 or Foxp4 are both embryonically lethal [398-402]. The fact that AGR2

expression is found in the human foetal liver [403] and also promotes the development of the lung epithelium [392] also suggests that AGR2 plays a role in organ morphogenesis. It is also interesting that Foxp1 enhances oestrogen-driven gene transcription in MCF7A breast cancer cells [404], whilst, in a panel of breast cancer cell lines, Foxa1 was found to be required for OR binding to chromatin [405], thus further cementing AGR2 as a hormone responsive gene.

Direct binding of the Sox10 (SRY (sex-determining region Y)-related HMG (high mobility box) box 10) transcription factor to the AGR2 promoter has been demonstrated during otic vessel formation in the zebrafish embryo, where Sox10 knockdown also reduced AGR2 expression [406]. Similarly, a binding site for Sox9 is predicted in the AGR2 promoter [407], but no experimental evidence exists of its regulation of AGR2 expression. Interestingly, AGR2 knockdown results in increased levels of Sox9 expression in the murine gut, suggesting that AGR2 is a repressor of Sox9 [408, 409]. Sox9 may be a stem cell marker [410], and thus, its repression by AGR2 suggests that AGR2 promotes differentiation of Sox9-positive cells in the murine gut [409].

Whilst the transcriptional regulators mentioned above suggest a role for AGR2 in tissue development, the demonstrated binding of the oncogenesis-associated zinc finger transcription factor ZNF217 to the AGR2 promoter in the metastatic lung cancer cell line Ntera2 [411] points to a role for AGR2 in cancer and metastasis (see below). In line with this, AGR2 expression is also regulated by TGFβ (itself implicated in oncogenesis and EMT, see section 1.2.2.1) through the SMAD4 transcription factor in pancreatic neoplasia [412].

Aside from the specific transcriptional regulators noted above, expression of AGR2 is also controlled by physiological stress. Serum-starvation in combination with hypoxia was found to induce AGR2 expression *in vitro*, and was mediated at least in part through the ERK1/2 pathway in the breast cancer cell line, MDA-MB-231 [413]. A further study indicated that this AGR2 induction was HIF1α-dependent, and that together these two proteins contributed to angiogenesis in glioblastoma [414]. AGR2 is also expressed in response to ER stress, where it protects cells from stress-induced cell death [408, 415-419]. This induction appears to be dependent on the ATF6 (activating transcription factor 6) and IRE (inositol-requiring protein 1) elements of the unfolded protein response [417], the process activated to restore ER homeostasis (reviewed in [420]). This implies that AGR2 might function as an ER chaperone.

1.4.2 Normal functions of AGR2

1.4.2.1 AGR2 and mucin expression

AGR2 (and indeed AGR3 and ERp18/AGR1) are members of the PDI family of ER chaperones, based on their structural homology to PDIs [378] (Fig. 1.3), and the recently solved structure of AGR2 demonstrated that the protein does indeed contain a thioredoxin-like fold [421] (see Appendix 1), a defining feature of PDI family members [315]. However, AGR2 possesses the divergent CxxS active site motif [315].

The major function of AGR2 was illustrated through a mouse knockout model, where loss of AGR2 led to loss of intestinal mucus production through the loss of mucin 2 (MUC2) expression [408, 422]. This suggests that AGR2 contributes to the process of protecting the intestinal epithelial layer from insult, further demonstrated by the decrease in TFF3 (trefoil factor 3) levels in the intestinal tract of *Agr2*$^{-/-}$ mice [422], a secretory protein that plays a role in mucosal repair [423]. Furthermore, AGR2 was shown to form thiol-dependent interactions with MUC2 *in vitro* [422], but whether this reflects an active role for AGR2 in disulphide bond formation or just an association with free-thiol groups is not clear. It is clear however that AGR2 plays a role in the differentiation of intestinal goblet cells, and also in maintaining both the distribution and integrity of mucin secretory goblet and antimicrobial Paneth cells within the linings of the intestine and stomach [384, 408, 409, 424-426].

AGR2 was similarly shown to be involved in the production of two other gel forming mucins, MUC5B and MUC5AC [418, 427]. In an *Agr2*$^{-/-}$ mouse model of asthma, mice lacking AGR2 showed defects in the production of airway mucus in response to allergen treatment, with a concurrent increase in ER-trapped MUC5B and MUC5AC and the appearance of ER stress markers [418]. The authors also reported a physical association between immature mucins and AGR2, further supporting the notion that AGR2 acts as a chaperone for mucin proteins.

There is evidence that AGR2 is involved in the regulation of the expression of a further mucin, MUC1 [412]. MUC1 is a cell surface mucin that is not involved in gel formation, but is involved in maintaining cell polarity and cell-cell adhesions [428-431], and as such is often aberrantly expressed in cancer (reviewed in [432]). Likewise, there is also a positive correlation between AGR2 and the expression of another cell surface mucin, MUC4, in pancreatic adenocarcinoma, where the elevated levels of both proteins together were associated with lower grade tumours [433]. This suggests that MUC4 may be a

protective factor against cancer, but overexpression of MUC4 also induces metastasis [434, 435], thus the exact role of MUC4 is probably context dependent [436].

1.4.2.2 Extracellular role of AGR2

AGR2 generally localises to the ER [417, 422, 437-439], although several studies have reported the secretion of AGR2 into cell culture medium [388, 426, 440], but whether this occurs merely as a consequence of cell lysis and subsequent release of abundant intracellular AGR2 has not been addressed. However, the abundance of AGR2 in mucus all along the murine gastrointestinal tract [426, 441, 442], as well as in human colonic mucus [426] probably suggests that AGR2 can be secreted *in vivo*. Others have also reported the association of AGR2 with the plasma membrane [302, 437, 443-445]. The consequences of these non-ER localisations of AGR2 are not clear, but several ER chaperone proteins, such as calreticulin, BiP and PDI, are also known to be secreted or membrane-bound [446-449], so this phenomenon is not AGR2-specific.

The newt homologue of AGR2 (nAG) is also secreted and nAG secretion is involved in limb regeneration via intercellular signalling through its surface receptor Prod1 [450, 451]. Although there was initial promise of a similar mechanism occurring in mammalian cells when a yeast two-hybrid screen reported an interaction of AGR2 with C4.4a (a structural homologue of the most closely related human counterpart of Prod1, urokinase-type plasminogen activator receptor (uPAR) [452, 453]) and the cell surface protein dystroglycan [389], there has been no further evidence that there exists a *bona fide* interaction between these molecules *in vitro* or *in vivo* in mammals. Nonetheless, there is some evidence for a role of secreted AGR2 in humans, as a mutant, highly secreted form of AGR2 expressed in A375 malignant melanoma cells upregulated the production of several proteins, notably dysferlin and WNK1 (with no lysine 1) [438]. These proteins promote membrane repair [454] and cell proliferation [455], respectively, and thus may support the notion of AGR2 contributing to intestinal epithelium homeostasis [384, 406, 408, 409, 425, 426]. Furthermore, conditioned medium from AGR2-secreting pancreatic cancer cells was shown to promote their proliferation and invasion, which was lost upon AGR2 knockdown [440]. It is not clear however if these effects are directly caused by extracellular AGR2 or by the loss of secretion of as yet unidentified factors under the control of AGR2.

1.4.3 AGR2 in cancer

Several regulators and potential substrates for AGR2 highlighted so far are known to play a role in tumourigenesis, and AGR2 itself has been associated with the progression of a number of different cancer types, although hormone-dependent cancers such as breast and prostate cancers have been the most studied, given the regulation of AGR2 by oestrogens and androgens (see above).

1.4.3.1 Hormone-dependent cancers

AGR2 seems to be particularly associated with the progression of breast cancer, possibly due to its role in breast tissue development, where it mediates proliferation of the mammary epithelium [456]. Multiple reports have identified its expression as a marker of disease stage and as a prognostic factor for patient survival in breast cancer, where increased expression of AGR2 is associated with reduced patient survival [309, 443, 457-462]. Importantly, AGR2 is also implicated in the resistance of breast cancers to drug treatment. AGR2 was shown to mediate resistance to cisplatin [463], where tumour xenografts stably-expressing AGR2 were insensitive to cisplatin-mediated growth arrest, in contrast to isogenic xenografts not expressing AGR2. Tamoxifen treatment leads to upregulation of AGR2 expression, directly through an OR and Akt-dependent manner [459, 460], but also in an OR-independent fashion through Foxa1 [393]. Thus, AGR2 expression is also predictive of Tamoxifen resistance [460]. On the other hand, treatment of patients with aromatase inhibitors such as letrozole and anastrozole downregulates the expression of AGR2 [464] (presumably through loss of oestrogen-mediated gene transcription) and in one patient cohort, those who responded the best to letrozole treatment corresponded to those patients that saw the largest suppression of AGR2 expression [459]. Therefore combination therapies including drugs affecting AGR2 expression might have a favourable outcome for breast cancer patients.

A similar role for AGR2 in the progression of prostate cancer is also apparent, with a number of studies again showing a correlation between AGR2 and the occurrence of prostate cancer, as well as decreased patient survival [388, 391, 444, 465-469]. Interestingly, studies in prostate cancer cells have revealed that knockdown of AGR2 can promote cellular senescence [470], but also that ErbB3 binding protein 1 (EBP1) acts as a negative regulator of AGR2 expression [305]. EBP1 is often downregulated in prostate

cancers [471, 472], and seems to inhibit AGR2 expression through inhibition of Foxa1 and 2. As downregulation of EBP1 is associated with resistance to hormone treatment [471] and is also presumably associated with upregulation of AGR2 expression [305], AGR2 could be a valid target for reducing hormone treatment resistance in prostate cancer, in a similar fashion to that seen in breast cancer.

The presence of AGR2 acts as a similar prognostic factor for other hormone-dependent cancers. It is a marker of higher grade ovarian cancers, and can also be used to distinguish between serous and mucinous-type ovarian cancers [301, 473-476]. Pancreatic cancers are also sensitive to oestrogens [477], and the expression of AGR2 also appears to correlate with appearance of neoplastic or pre-neoplastic pancreatic cells [302, 412, 433, 437, 440, 478-480].

The prevalence of elevated AGR2 levels in these cancers (but also in some non-hormone-dependent cancers, see below) may relate to its pro-proliferative and anti-apoptotic effects. AGR2 mediates growth-regulatory effects in several cancer cell lines [301, 393, 440, 459], where AGR2 knockdown results in decreased or inhibited growth. More strikingly, AGR2 knockdown also results in apoptosis of several cancer cell lines [415, 445, 459, 470, 481] but in the context of breast cancers at least, this only appears to be true for OR-positive cell lines [482]. This suggests that some cancers may be dependent on AGR2 for survival. AGR2 appears to mediate these effects, at least in part, through upregulation of the cell cycle protein cyclin D_1, as well as the pro-survival protein survivin [482], although how it influences the cellular abundance of these proteins is not clear. Additionally, AGR2 is also an inhibitor of p53 activity, preventing activation of pro-apoptotic protein transcription through inhibiting its phosphorylation [438, 445]. Again, whether these effects are mediated through direct interactions or are the result of influences on upstream regulators is unknown.

1.4.3.2 Non-hormone dependent cancers

The occurrence of AGR2 expression in non-hormone dependent cancers is not as widely reported, presumably due to the strong effects of oestrogens and androgens on AGR2 expression. Nonetheless, AGR2 also appears to play a role in carcinogenesis in other tissues, generally those that contain a secretory epithelium.

AGR2 has been associated with the progression of oesophageal carcinomas, where it tends to be upregulated relative to normal tissue [394, 419, 445, 481, 483-485]. As

oesophageal carcinomas have been reported to display a characteristic inhibition of p53 [486], AGR2 may promote disease progression through inhibition of p53 activation [445]. Similarly, there is a correlation between AGR2 expression and disease progression in lung cancers [303, 487-491].

In cancers originating from the liver, the role of AGR2 appears to depend on the specific sub-type of carcinoma. As mentioned previously, AGR2 is expressed during liver and biliary tree morphogenesis [403] and seems to be maintained in normal liver [492], as well as in extrahepatic and some intrahepatic cholangiocarcinomas, where AGR2 expression correlates with mucus production [403]. AGR2 was similarly overexpressed in fibrolamellar carcinomas, but not in primary hepatocellular carcinomas (HCC), or at least only very rarely [403, 492]. However, high levels of AGR2 were found in metastatic HCC cells [304], suggesting that different selection pressures exist on AGR2 expression, depending on the cell type of origin. Furthermore, Yu and colleagues demonstrated that AGR2 directly interacted with elements of the MAPK and caspase pathways using co-immunoprecipitation experiments, and that this may contribute to the induction of metastasis in these cells [304]. However, the cellular expression levels of these potential interacting proteins were not investigated, so it is unclear if AGR2 affects the abundance of these proteins (in a similar way to mucins) or whether interaction with AGR2 results in some activation or inhibitory action. Any activity not based on protein abundance suggests an additional, chaperone-independent role for AGR2. Likewise, in colorectal carcinomas, where AGR2 expression also correlates with mucus production and can distinguish between mucinous and non-mucinous carcinomas, elevated expression of AGR2 can predict both better and poorer patient outcome, based on the specific cancer sub-type [493-495]. However, it is possible that the increase in serum AGR2 mRNA in colorectal cancer patients identified by Valladares-Ayerbes and colleagues [496] could also reflect the expression of AGR2 in circulating tumour cells as reported previously [306], thus again implicating AGR2 in the metastatic process.

1.4.3.3 AGR2 downregulation in cancer

Whilst the majority of reports demonstrate AGR2 expression to be an indicator of poor patient outcome, a number of studies have also demonstrated a correlation between high levels of AGR2 expression and better disease prognosis. Two individual studies of breast cancer and prostate cancer showed that AGR2 expression correlated with improved

patient survival [497, 498]. However, this is not entirely unexpected, given the role of AGR2 in differentiation of breast, intestinal and stomach tissues [409, 422, 424, 439, 456], where in the latter, loss of AGR2 leads to dedifferentiation and metaplasia [409, 439]. It is also important to note that the aforementioned studies on breast and prostate cancers did not stratify groups based on hormone receptor status. The fact that hormone receptor positive cancers are generally more treatable than receptor-negative cancers [499], and that AGR2 expression correlates with hormone receptor expression, may explain the apparent protective effect of AGR2. There are similar reports of AGR2-protective effects in ovarian cancer and colorectal cancer [493, 500], but this also positively correlates with cell differentiation in ovarian cancer [500] and probably also in colorectal cancer, given the role of AGR2 in the development and maintenance of the intestinal epithelium (see Section 1.4.2.1). Also in biliary tract tumours, expression of AGR2 correlates with cell differentiation, such that AGR2 can be downregulated during progression of these tumours [501].It appears then that, in tissues where AGR2 is normally expressed and functions normally, AGR2 tends to be lost during tumourigenesis due to the loss of a differentiated cell phenotype, whereas in tissues where AGR2 is not normally, or only weakly expressed, AGR2 overexpression leads to pro tumourigenic effects, including increased proliferation and survival.

1.4.4 AGR2 and metastasis

In addition to the association of AGR2 with different cancer types outlined above, a number of studies have demonstrated differential expression of AGR2 in primary and metastatic tumour cells [301-309]. This association of elevated AGR2 levels in metastatic cells, coupled with the reported secretion of AGR2 and its correlation with disease progression and patient survival, has led to the exploration of AGR2 as a potential prognostic and diagnostic biomarker. Several studies have explored urinary or serum AGR2 levels in this role, with encouraging results for not only tumour-secreted AGR2 [465, 469, 475, 476, 479, 480, 488, 502-504], but also for the detection of circulating tumour cells within patient serum, based on the presence of AGR2 mRNA [306].

The mechanism whereby AGR2 might induce metastasis however, is not yet clear, although the EGFR ligand amphiregulin (AREG), which plays a crucial role in promoting mammary gland development [505], may play a role here. AGR2 induces expression of the AREG, through the Hippo pathway transcriptional co-activator YAP1 (yes-associated protein

1) [506], by inhibiting YAP1 phosphorylation, which otherwise sequesters it in the cytosol [507]. Whether a direct interaction occurs between AGR2 and YAP1 has not been determined. AREG is often overexpressed in a number of cancers (including breast, prostate, lung and colorectal cancers (reviewed in [508]), where it acts as a growth and survival factor by signalling through the EGF receptor and also by upregulating anti-apoptotic Bcl-X_L protein whilst downregulating the pro-apoptotic Bac and Bax proteins [508-511]. It also mediates resistance to anti-EGFR therapy, probably through drug-induced activation of AREG transcription and thus increased anti-apoptotic signalling [512-514]. Importantly, AREG has pro-EMT effects, decreasing cell-cell adhesion and increasing migration and invasion, through downregulation of E-cadherin and the upregulation of matrix-degrading enzymes including MMP2 and MMP9 (reviewed in [508]). Therefore, an AGR2-AREG axis may be important in promoting cancer cell dissemination. Additionally, the demonstration that the ECM-degrading proteases cathepsin B and D are upregulated by AGR2 in pancreatic cancer cells [437] could feasibly also contribute to the pro-metastatic ability of AGR2, by promoting breakdown of the surrounding matrix and thus facilitating invasion.

Metastases promoted by AGR2 therefore probably relate to increased cell migration and invasion of AGR2-expressing cells [301, 305, 307, 308, 415, 437, 481]. Furthermore, a direct link between AGR2 and metastasis has also been demonstrated in a rat model of breast cancer, where normally benign rat mammary cells were induced to form lung metastases through the expression of AGR2 [309]. In this model, the forced expression of AGR2 enhanced the rate of adhesion of these cells to the substratum, which as outlined in section 1.1.1.4.2, might be important for circulating tumour cell attachment to the endothelium at secondary metastatic sites. A further study in prostate cancer cells demonstrated a similar effect, where knockdown of AGR2 in PC3 cells resulted in a decreased rate of adhesion to the substratum [515]. Thus AGR2 is associated with changes in cell adhesion, a process which is not only important for tumourigenesis in the context of adhesion at secondary sites, but also due to the fact that changes in adhesion are implicated in multiple steps in the metastatic cascade, including cell-cell adhesion, cell-matrix adhesion, cell migration and anchorage-independent cell growth. AGR2-regulated adhesion could thus be a critical central component of the metastatic cascade.

1.5 Project aims

Given its role in cancer progression, the overall aim of the present project is to investigate the mechanism involved in AGR2-promoted metastasis. More specifically, the aims of the project are: (1) To explore the biophysical properties of AGR2 protein along with its effects on cell adhesion, in order to uncover and identify regions of AGR2 protein that might contribute to its adhesion-related activities, which are associated with its induction of metastasis. (2) To study the extent and nature of secreted AGR2, and (3) To identify significant changes to the intracellular and secreted proteomes arising from the overexpression of AGR2, to shed light on possible downstream targets of AGR2 overexpression, with the longer term goal of identifying targets for anti-AGR2 drug development.

Chapter 2

Materials and Methods

2.1 Reagents and Materials

All chemicals were obtained from Sigma or Fisher Scientific unless otherwise stated. All cell culture plastic ware was obtained from Corning unless otherwise stated.

2.2 Cell Culture

All cell culture media and supplements were obtained from Gibco, unless otherwise stated. All cell lines were grown in Dulbecco's Modified Eagle's Medium (DMEM) containing 4500 mg/mL glucose, non-essential amino acids (NEAA), 3.7 g/L sodium bicarbonate and phenol-red. All growth media were supplemented with 4 mM L-glutamine, 100 units/mL

penicillin and 100 µg/mL streptomycin. Where added, foetal bovine serum (FBS) was from Lonza. Cells were maintained at 37° C in a humidified atmosphere containing 10 % CO_2.

2.2.1 Cell lines

	Description	Culture Medium Supplements
Rama 37	Benign, DMBA (dimethylbenzanthracene)-derived rat mammary adenoma [516]	5 % (v/v) FBS, 10 ng/mL insulin, 10 ng/mL hydrocortisone
MCF7A	Metastatic human mammary adenocarcinoma	10 % (v/v) FBS, 10 µg/mL insulin

Table 2.1. ***Parental cell lines and growth media.***

The Rama 37 cell line was employed for this study with a view to utilising the cell lines created below in a syngeneic rat model to investigate the metastatic effects of AGR2 in an *in vivo* model, as performed previously using WT AGR2-expressing Rama 37 cells [309].

2.2.2 Stably-transfected cell lines

All stably-transfected cell lines were grown in Rama 37 medium as described Table 2.1, with the addition of 1 mg/mL G418 disulphate (Melford) as a selection agent.

	AGR2 Sequence Covered (amino acid number)	Mutations
Rama 37 Empty Vector	None	None
Rama 37 WT AGR2	Full length (1-175)	None
Rama 37 E60A AGR2	Full length (1-175)	Glutamate to alanine at position 60
Rama 37 C81S AGR2	Full length (1-175)	Cysteine to serine at position 81
Rama 37 Δ1-20 AGR2	(21-175)	Deletion of N-terminal signal peptide
Rama 37 Δ21-40 AGR2	(41-175)	Deletion of N-terminal residues 21-40
Rama 37 ΔKTEL AGR2	(1-171)	Deletion of four C-terminal residues (KTEL)

Table 2.2. ***Rama 37 stably-transfected cell lines.***

2.2.3 Thawing and subculturing cells

Cells were rapidly thawed in a 37°C waterbath and plated in 10 mL Rama 37 medium in 10 cm culture dishes. After 24 h, the medium was changed to fresh growth medium.

Cells were subcultured every 2-3 days upon reaching 70-80 % confluence. For cuboidal Rama 37 cell lines, the cell monolayer was washed twice with PBS and incubated in 1 mL Versene (0.02 % EDTA (Ethylenediaminetetraacetic acid) in PBS)for 10 min in a 37°C incubator. For elongated Rama 37 cell lines, cells were incubated for only 3 min. After removing the Versene, cells were then incubated in 1 mL Trypsin solution (0.05 % in Versene) for 3 min in a 37°C incubator or until all cells had detached. Cells were collected in Rama 37 medium and subcultured at a ratio of 1:3 to 1:8, as required.

For the subculturing of MCF7A cells, the cell monolayer was washed twice in PBS and incubated in 1 mL Trypsin solution for 2 min in a 37°C incubator until all cells had detached. Cells were then collected in MCF7A medium and subcultured at a ratio of 1:2 to 1:3, as required.

2.2.4 Freezing cells

Frozen cell stocks were made using cells grown to 90-100 % confluence. Cells were collected as described in Section 2.2.3, and then centrifuged at 1000 x *g* at room temperature. The cell pellet was then resuspended in 2 mL freezing medium (complete growth medium supplemented with 20 % (v/v) FBS and 7.5 % (v/v) DMSO) per 10 cm culture dish, and 1 mL cell suspension added to each cryovial (Star Labs). Cryovials were placed in a Cryo freezing container (Nalgene) overnight at -80° C to ensure a cooling rate of approximately -1°C/min and then transferred to long-term storage at -135° C.

2.2.5 Creation of cell lines

2.2.5.1 Isolation of elongated Rama 37 cells (R37E)

Rama 37 cells can undergo spontaneous transformation from a cuboidal to an elongated morphology [516]. Elongated cells attach much more weakly to the cell culture

vessel (see Chapter 4) and so to isolate this elongated subpopulation of Rama 37 cells, confluent cultures of cells in 10 cm cell culture dishes were treated with Versene for 10 min at 37°C and then gently agitated, causing the most weakly attached cells to detach from the plate surface. These cells were gently aspirated and placed in Rama 37 medium in 6-well cell culture plates. Upon reaching confluency, the cells were transferred to 10 cm dishes and allowed to reach confluency again. These dishes were then treated with Versene again for 5 min at 37°C and again gently agitated. Detached cells were collected into 6-well cell culture plates and expanded, as before. This process was repeated a third time, such that cells had been selected from three rounds of Versene treatment, and the cells detaching from this final Versene treatment were expanded and named Rama 37-elongated (R37E).

2.2.5.2 Transfection and generation of stably-expressing cell lines

Cells were transfected using FuGENE6 transfection reagent (Promega). Twenty-four hours prior to transfection, Rama 37 cells were seeded at 0.3 x 10^6 cells/well in 6-well cell culture plates in order to reach 50-60 % confluence. On the day of transfection, cells were washed twice with PBS and the medium replaced with Opti-MEM reduced serum medium (Gibco). DNA (either 2 µg PiggyBac plasmid DNA, 0.8 µg PiggyBac transposase plasmid DNA or both) was complexed in a 1:3 ratio with FuGENE6, and complexes were incubated with cells at 37°C for 6 h. At this point the medium containing the complexes was aspirated and replaced with fresh Rama 37 medium.

For the generation of stably-expressing cell lines, cells reaching 80-100 % confluency 24 h post-transfection were transferred from 6-well cell culture plates to two 10 cm cell culture plates. Forty-eight hours post-transfection, medium was replaced with selection medium containing 1 mg/mL G418 (Melford), and fresh medium was added to cells every 2-3 days for 10-14 days until cells had grown sufficiently to allow freezing of cell stocks. For cells transfected with shRNA constructs, the same procedure was followed except that selection was performed in Rama 37 medium supplemented with 250 µg/mL Hygromycin B.

2.2.6 Analysis of cell morphology using ImageJ

For measuring the morphological characteristics of cells, images were taken of cultured cells using a DC5000 CMEX microscope camera (Euromex). Images were opened in ImageJ [517], converted to 8-bit images and then set in black-and-white by entering the 'Threshold' menu and selecting the default settings. Cell morphology was measured using the 'Analyze particles' function with the inclusion of 'shape descriptors'. Cells were considered if they fell within the 100-5000 pixel^2 size range and cells residing on the edges of the image were not considered in the analysis. Morphology was measured by the aspect ratio (AR) value, which measures the longest axis divided by the shortest axis. AR values were measured for all cells fitting the size criteria and the mean of these AR values used as the shape ratio (SR) for each cell line. A representative analysis is shown in Fig. 2.1.

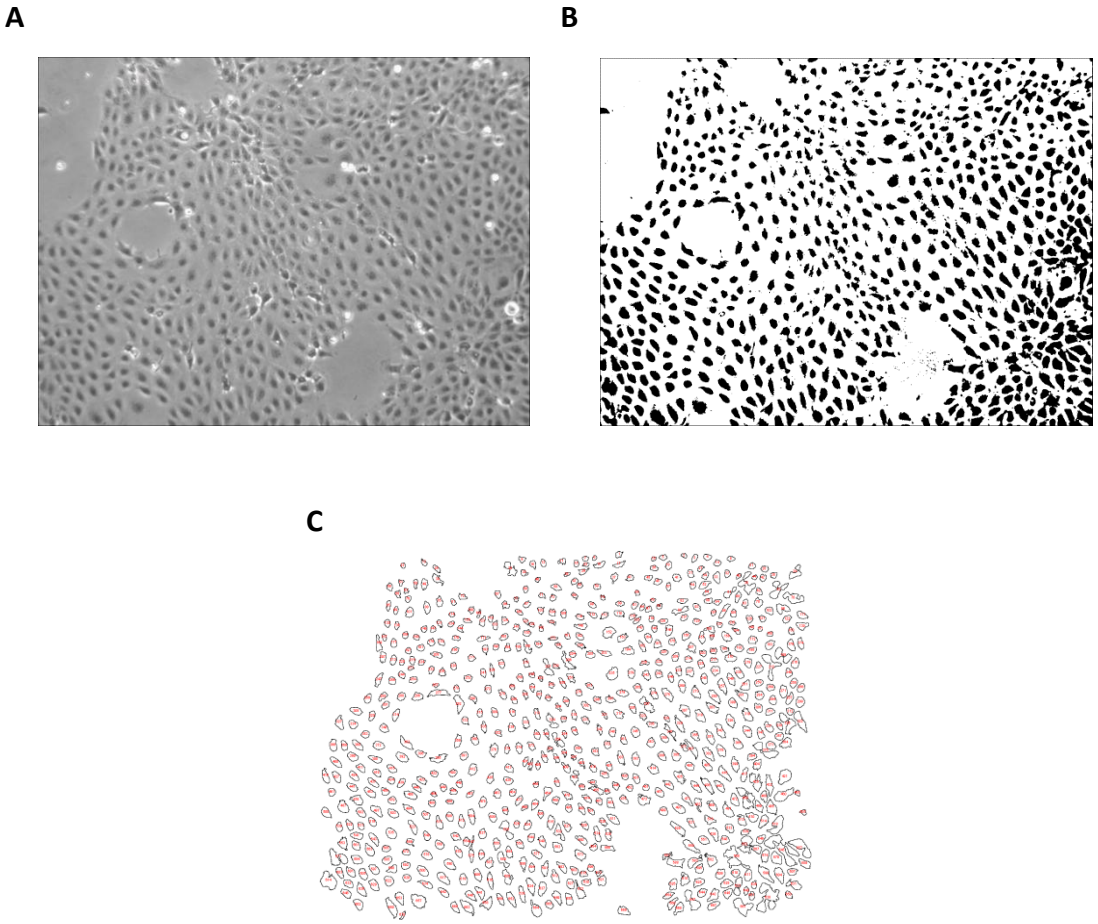


Figure 2.1. ***Analysis of cell morphology using ImageJ.*** Images of cultured cells **(A)** were converted to Threshold images **(B)** to delineate individual cells. Cells fitting the criteria outlined in the main text were then measured across the shortest and longest axes **(C)** to give an aspect ratio value.

2.3 Cell-based assays

2.3.1 Adhesion rate assay

Cells were grown to 70-80 % confluence and collected by Versene and trypsin treatment in Rama 37 medium. Cells were counted with a Z1 Coulter® particle counter (Beckman Coulter) and resuspended to 2 x 10^5 cells/mL. 1 mL of this cell suspension was then plated into each well of a 24-well plate and left in a 37°C incubator for the time indicated in the figure legends. After the allocated time, medium was aspirated from wells and adhering cells were washed twice with 1 mL PBS. Adhering cells were detached by adding 250 µl Versene for 10 min in a 37°C incubator, followed by the addition of 250 µl 0.05 % (v/v) trypsin/Versene for 5 min in a 37°C incubator, without aspirating off the Versene solution. Cells were then collected by the addition of 500 µL medium, leading to a cell suspension of 1 mL that was subsequently counted using a Z1 Coulter® particle counter (Beckman Coulter). The number of adhering cells was calculated as a percentage of the 2 x 10^5 cells that were added to each well.

2.3.2 Adhesion rate assay on protein-coated plates

For protein coating experiments, 500 µL of 2 µM purified protein in PBS buffer (pH 7.4) was coated onto each well of a 24-well plate the day before the adhesion experiment. Plates were placed in a 37°C incubator uncovered overnight, until all liquid had evaporated. Prior to use, wells were washed with 1 mL PBS and allowed to dry again, before the addition of cells. The assay was then performed in the manner described above.

For peptide coating experiments, 100 µl of 2 µM peptide and 7.5 nM AGR2 protein were coated onto 96-well non-cell culture treated plates in 50 µl carbonate buffer (50 mM carbonate, 150 mM NaCl, pH 10.6) in the same manner described above. Wells were then washed once with 100 µL PBS and allowed to dry at room temperature. Cells were grown to 70-80 % confluence and collected by Versene and trypsin treatment in Rama 37 medium. Cell were counted with a Z1 Coulter® particle counter (Beckman Coulter) and resuspended to 3 x 10^5 cells/mL. 100 µl of this cell suspension was then plated into each well of the peptide- or protein-coated 96-well plate and left in a 37°C incubator for 1 h. Medium was

aspirated from wells and adhering cells were washed twice with 100 µL PBS and quantified using crystal violet staining.

For staining, cell were fixed in 100 % ethanol for 10 min at room temperature and then incubated with 0.05 % (w/v) crystal violet in 20 % (v/v) methanol/distilled water (dH_2O) for 10 min at room temperature. Following 5 washes of dH_2O, cell-bound crystal violet was released through incubation with 100 µL/well of 10 % (v/v) acetic acid for 5 min at room temperature, with moderate agitation. Staining intensity was measured at 570 nm using a Spectramax plus384 (Molecular Devices) plate reader, and the staining of treated cells was expressed as a percentage of staining in control (non-protein coated) wells. Each condition was plated in duplicate.

2.3.3 Cell detachment assays

2.3.3.1 Short term trypsin resistance assay

The assay was performed in the same manner as the rate of adhesion assay (Section 2.3.1), up to and including the two washes of PBS after the allocated incubation time, but two sets of wells in separate plates were used: one control plate to measure the number of adhering cells in the allocated time, and one treatment plate to measure the adhesion strength of these cells. At this point, the control plate was quantified in exactly the same manner as the rate assay (Section 2.3.1). For the treatment plate, cells were incubated with 250 µL 0.0125 % (v/v) trypsin/Versene in a 37°C incubator for 5 min. The solution was aspirated and cells were gently washed with 1 mL PBS, and the remaining adhering cells were counted using a Z1 Coulter® particle counter (Beckman Coulter) as described in Section 2.3.1. The number of adhering cells was expressed as the number of cells still adhering to the treatment plate as a percentage of those adhering to the control plate. The process is illustrated in Fig. 2.2.

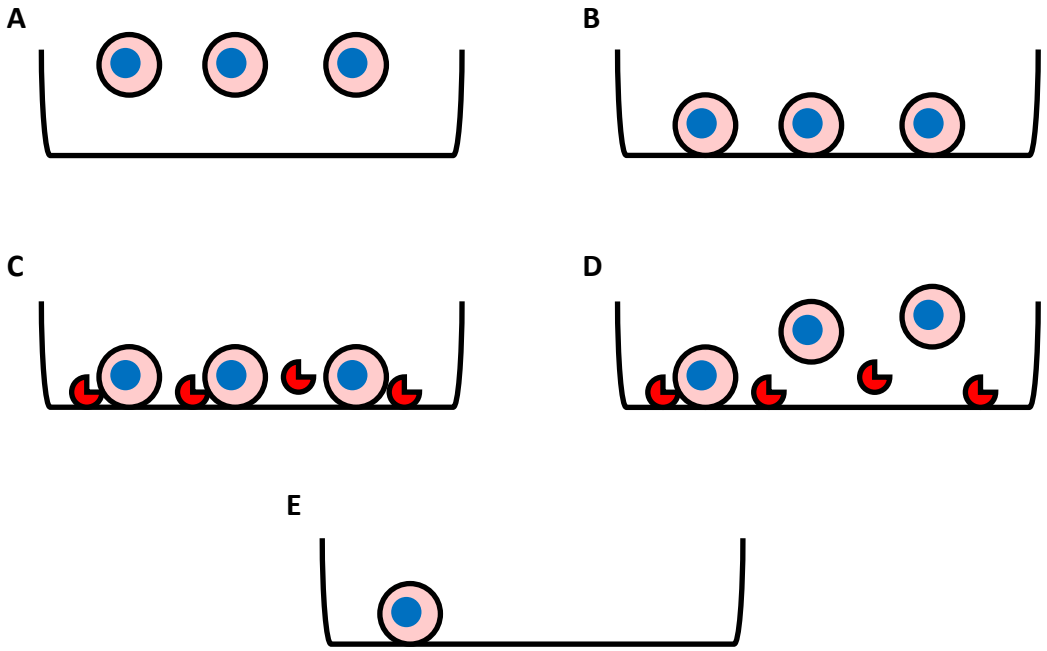


Figure 2.2. ***Short term trypsin resistance assay.*** Cells are plated **(A)** and allowed to adhere to the plate surface for the indicated time **(B)**. Bound cells are then treated with low strength trypsin for 5 min **(C)**, and detached cells are washed away with PBS **(D)**. Remaining, trypsin-resistant cells **(E)** are then counted using an automated cell counter. Red shapes represent trypsin molecules.

2.3.3.2 Monolayer trypsin resistance assay

Cells were plated into two sets of 24-well plates and left overnight to reach 70-80 % confluence by the next day, with one plate acting as the control and the other the treatment plate. The next day, cells in the control plate were collected and counted as described above. For the treatment plate, cells were incubated with 250 µL 0.0125 % (v/v) trypsin/Versene in a 37°C incubator for 5 min. The solution was aspirated and cells were gently washed with 1 mL PBS, and the remaining adhering cells were released and counted using a Z1 Coulter® particle counter (Beckman Coulter) as described in Section 2.3.1. The number of adhering cells was expressed as the number of cells still adhering to the treatment plate after treatment with 0.0125 % (v/v) trypsin, as a percentage of those adhering to the control plate. The process is illustrated in Fig. 2.3.

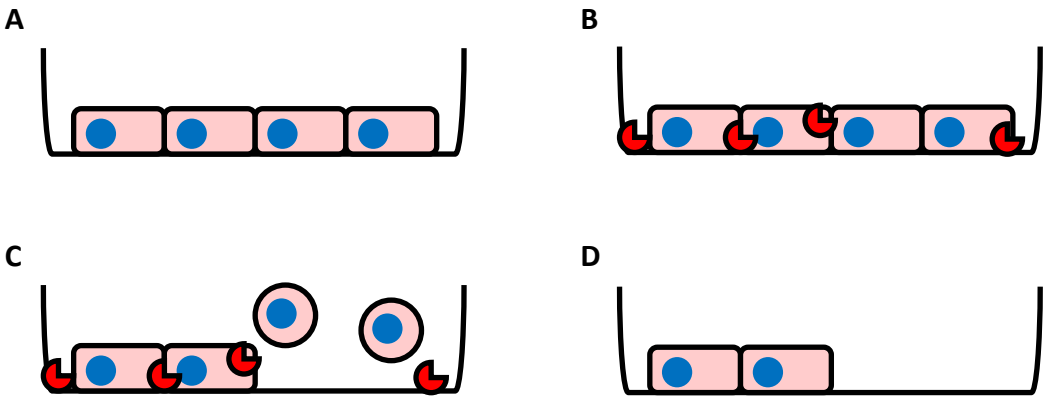


Figure 2.3. ***Monolayer trypsin resistance assay.*** Cells are plated and allowed to reach 70-80 % confluence overnight **(A)** and then treated with low strength trypsin for 5 min **(B)**. Detached cells are washed away with PBS **(C)** and the remaining, trypsin-resistant cells **(D)** are then released and counted using an automated cell counter. Red shapes represent trypsin molecules.

2.3.3.3 Centrifugation assay

Assays were performed as described in [518, 519]. Cells were grown to 70-80 % confluence and then collected using Versene and trypsin as described above. Cells were resuspended to 3 x 10^5 cells/mL in normal medium, and 100 µL/well of this cell suspension were plated into 2 plates. Both sets of plates were then centrifuged at 10 x *g* for 5 min in a swing out plate rotor to sediment all cells and ensure they were in contact with the substratum. Plates were then left at room temperature for 5 min, before gently adding 240 µL medium to all wells. Plates were sealed with sticky-back 96-well plate covers, avoiding the formation of air bubbles within wells. The treatment plate was then placed upside down and centrifuged at 800 x *g* for 10 min at room temperature. During this time, the control plate was kept the correct way up at room temperature without centrifugation. Following centrifugation, medium was aspirated from both sets of plates, and attached cells were quantified using crystal violet staining, as described in Section 2.3.2. The assay is illustrated in Fig 2.4.

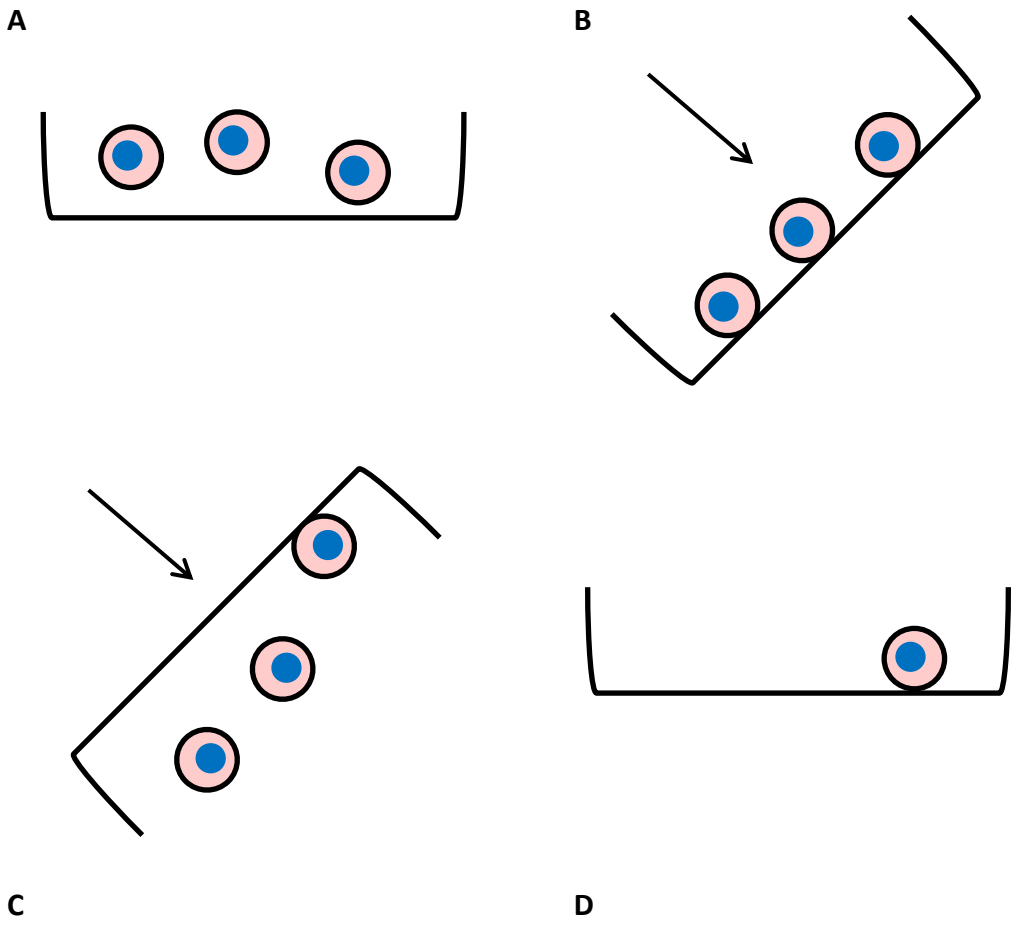


Figure 2.4. ***Centrifugation cell detachment assay.*** Cells are plated into 96-well plates **(A)** and centrifuged at 10 x *g* for 5 min to ensure all cells are in contact with the substratum **(B)**. Plates are sealed and centrifuged upside down for 10 min at 800 x *g* **(C)** and cells that remained attached **(D)** are quantified using crystal violet staining. Arrows indicate the direction of centrifugal force

2.3.3.4 Migration assay

Cells were grown to 70-80 % confluence, washed twice with PBS and incubated with serum-free medium for 2 h in a 37°C incubator. Cells were then collected in fully supplemented medium containing only 1 % FBS, resuspended to 1 x 10^5 cells/mL and then 100 µL (1 x10^4 cells) were added to the upper chamber of a Transwell insert, with 8.0 µm pore size. Inserts were placed into 24-well plates containing 600 µL fully supplemented medium containing 5 % (v/v) FBS, and cells were allowed to migrate for 16 h in a 37°C incubator. Cells were subsequently stained using a Reastain Quick Diff kit (Reagena) and

counted using a microscope. All cells were counted from each membrane and each condition was plated in duplicate.

2.3.3.5 Anchorage-independent growth assay

50 µl of semi-solid agar base layers consisting 0.5 % agar (w/v) in normal growth medium were plated into 96-well plates by diluting 2.5 % (w/v) autoclaved Ultrapure CM agarose (Affymetrix)/dH_2O 1:4 into 1.2 x concentrated normal growth medium. 1.2 x medium was made by adding all supplements for Rama 37 medium or G418 medium (as described in Section 2.2.1) required to make 50 mL medium to only 40 mL DMEM. Plates were covered and placed at 4°C for 5 min to ensure agar solidification. During this time, 1.5 % (w/v) MacroSieve Clone CM low melting temperature agarose (Flowgen Biosciences) was prepared by microwaving in dH_2O and subsequently kept in a 40°C waterbath until needed.

For the top, cell-containing layer, cells were grown to 70-80 % confluence, collected and resuspended to 2.5 x 10^4 cells/mL in 1.2 x normal medium. This cell suspension was then mixed in a 1:4 ratio with warm 1.5 % (w/v) agarose and 50 µL (1 x 10^3 cells) were gently pipetted onto the agar base layer. Plates were then incubated at 4°C for 10 minutes to ensure agarose solidification, and this cell layer was then overlaid with 50 µL normal medium. Cells were kept in a 37°C, 10 % (v/v) CO_2 incubator for four weeks, and medium was changed once a week by gently pipetting off the top medium layer and gently adding new medium using a multi-channel pipette. After four weeks, colonies were imaged using a DC5000 CMEX microscope camera (Euromex) and subsequently quantified using CellTiter Glo 2.0 cell viability reagent (Promega), as described by the manufacturer. Using a TriStar LB 941 multimode microplate reader (Berthold Technologies), 50 µl of CellTiter Glo 2.0 reagent was dispensed into each well, still containing the 50 µl medium overlay. Samples were gently agitated for 2 min and then left to incubate at room temperature for 10 min. The signal in relative light units (RLU) was collected for 2 seconds. RLU is proportional to the number of viable cells, and the value from test samples was normalised to the RLU signal obtained from control WT AGR2 clone 2 cells to allow comparison across experiments. Cells were plated in quadruplicate, onto clear (Corning) and opaque (Berthold Technologies) 96-well cell culture plates for imaging and quantification by luminometry, respectively.

2.4 Immunocytochemistry

2.4.1 Preparation of paraformaldehyde

4 % (w/v) paraformaldehyde (PFA) was made by adding 4 grams PFA powder to approximately 90 mL of sterile PBS. The solution was heated and several drops of concentrated NaOH were added until the PFA had completely dissolved. The pH was then adjusted back to pH 7.4 and the solution was sterile filtered through a 0.22 µm filter before being aliquotted and frozen at -20°C. Frozen aliquots were kept for a maximum of 4 weeks.

2.4.2 Cell fixing and staining

One day before imaging, CultureSlides (Beckman Coulter) were coated with approximately 2.5 µg bovine fibronectin in PBS per cm^2 by evaporation at room temperature for 3-4 h. Slides were then washed once in PBS and left to air dry. Cells were collected and 3 x 10^4 cells were plated per well and left to grow overnight.

The next day, cells were washed once with PBS and then fixed in 4 % (w/v) PFA for 10 min at room temperature. Cells were then washed twice in PBS before incubation for 20 min with 50 mM NH_4Cl at room temperature to block any remaining reactive aldehyde groups. For permeabilisation, cells were treated with 0.5 % (v/v) Triton X-100 in PBS for 10 min at room temperature and were then blocked with 2 % (w/v) BSA, 3 % (w/v) skimmed milk powder in 0.5 % (v/v) Triton X-100/PBS for 1 h at room temperature. The blocking solution was aspirated off and cells were stained with a 100 µL solution of the antibodies shown in Table 2.3 in blocking buffer, overnight at 4°C. For co-staining experiments, both antibodies were incubated with cells at the same time.

The next day, antibody solutions were aspirated off and cells were washed 4 times with 0.5 % (v/v) Triton X-100/PBS before the addition of 100 µL secondary antibody solution (see Table 2.3) in blocking buffer for 30 min at room temperature. Cells were washed 4 times in 0.5 % (v/v) Triton X-100/PBS before incubation with 0.1 µg/ml DAPI for 10 min at room temperature, and subsequently washed twice more in PBS. CultureSlides were then soaked in 100 % methanol for 10 min to separate the glass slide from the plastic wells, and coverslips were mounted using Hydromount mounting medium (National Diagnostics) and allowed to air dry for at least 15 min at room temperature. Slides were

subsequently imaged using an EVOS® FL microscope imaging system (Life Technologies), equipped with DAPI, GFP and RFP light cubes.

For experiments involving non-permeabilised cells, all steps described above using 0.5 % (v/v) Triton X-100/PBS were replaced with PBS only, except for the permeabilisation step which was omitted.

Primary Antibodies				
Antibody	Host Species	Company	Product Code	Dilution Used
AGR2	Mouse	Abcam	ab56703	1:200
AGR2	Rabbit	Proteintech Group	12275-1-AP	1:200
CD44	Mouse	Thermo Scientific	MS-668-P0	1:20
Giantin	Rabbit	Abcam	ab24586	1:500
MEK1	Rabbit	Abcam	ab32091	1:200
Paxillin	Rabbit	Abcam	ab32048	1:500
PDI	Rabbit	Abcam	ab3672	1:100

Secondary Antibodies			
Antibody	Company	Product Code	Dilution Used
AlexaFluor488 Goat anti-mouse IgG	Life Technologies	A-11029	1:500
AlexaFluor488 Goat anti-rabbit IgG	Life Technologies	A-11034	1:500
AlexaFluor568 Goat anti-rabbit IgG	Life Technologies	A-11036	1:500

Table 2.3*. Antibodies used for immunocytochemistry.*

2.4.3 Cell surface staining

Staining of cell surface proteins was based on the method of Zhuang *et al.,* 2008 [520]. Slides were prepared and cells plated and allowed to grow overnight as described in Section 2.4.2. All subsequent steps were performed on ice with ice-cold reagents. Cells were incubated with primary antibodies (Table 2.3) for 1 h in staining medium consisting of DMEM supplemented with 10 % (v/v) FBS, 100 units/mL penicillin and 100 µg/mL streptomycin, 4 mM L-glutamine, 10 mM HEPES and 15 mM sodium azide. Inclusion of azide helps prevent internalisation of surface antigens [521]. Following this incubation, the staining solution was aspirated and cells were gently washed 3 times with a wash solution consisting of PBS supplemented with 1 mg/mL glucose, 10 mM HEPES and 15 mM sodium azide. Cells were then incubated with the relevant secondary antibodies (Table 2.3) for 30 min in staining medium, before gently washing 3 times in wash solution. Cells were then fixed on ice with 4 % (w/v) PFA at room temperature, before being washed twice with PBS.

Slides were then stained with DAPI at room temperature and mounted and imaged as described above.

2.5 Bacterial transformation and growth

2.5.1 Preparation of agar plates

LB agar
1 % (w/v) tryptone
0.5 % (w/v) yeast extract
1 % (w/v) NaCl
1.5 % (w/v) agar

Stock solutions of 100 mg/mL ampicillin (Sigma) and 30 mg/mL kanamycin (Calbiochem) in dH_2O were filter sterilised through a 0.22 µm filter and stored at -20°C. Autoclaved LB-agar (Merck) was melted in a microwave oven and allowed to cool to approximately 50°C and ampicillin or kanamycin was added to a final concentration of 100 µg/mL or 30 µg/mL, respectively. Agar was then quickly dispensed into 10 cm Petri dishes (Sterilin) and allowed to set at room temperature for at least 15 min. Unused plates were sealed and kept at 4°C no longer than 4 weeks.

2.5.2 Transformation

SOC medium
2 % vegetable peptone
0.5 % yeast extract
10 mM NaCl
2.5 mM KCl
10 mM $MgCl_2$
10 mM $MgSO_4$
20 mM glucose

LB broth
1 % (w/v) tryptone
0.5 % (w/v) yeast extract
1 % (w/v) NaCl

50 µL of One-shot ® Top10 chemically competent *E .coli* cells (Invitrogen) or 50 µL of One-shot ® BL21 star ™ chemically competent *E. coli* cells (Invitrogen) were used for the expression of plasmid DNA and recombinant proteins, respectively. Cells were thawed on

ice and incubated with DNA on ice for 30 min before being heat shocked at 42°C for 30 seconds. Cells were then placed back on ice for 5 min before the addition of 250 µL SOC medium (New England Biolabs) and agitation at 220 rpm for 1 h at 37°C. 100 µL of transformation mixture was then plated onto LB-agar plates and left overnight in a humidified incubator at 37°C.

Starter cultures were prepared by inoculating 5 mL or 100 mL LB broth (Merck) with a single colony picked from transformed agar plates and agitating at 220 rpm overnight in a 37°C incubator. The volume of LB broth in the culture vessels did not exceed 25 % of the total vessel volume, in order to allow adequate aeration of the culture.

2.5.3 Glycerol stocks

800 µL of overnight bacterial culture was mixed with 200 µL sterile glycerol by gentle pipetting, and tubes were immediately frozen at – 80°C.

2.6 Cloning, PCR and DNA amplification

2.6.1 Expression vectors

A previously-constructed pBK-CMV vector containing AGR2 (pBK-CMV-AGR2, kindly provided by Dr Dong Barraclough, Institute of Ageing and Chronic Disease, University of Liverpool) was used as a template for subcloning WT and mutant AGR2 into the PiggyBac expression vector (System Biosciences, see Appendix 2), as described in Section 2.6.3. For recombinant protein expression in *E. coli*, two previously constructed expression vectors were used: an IMPACT-TWIN1 vector (Intein Mediated Purification with an Affinity Chitin-binding Tag-Two Intein, New England Biolabs) expressing residues 21-175 of AGR2 ($AGR2_{21-175}$) with a cleavable N-terminal chitin-binding tag (kindly provided by Dr Dong Barraclough, Institute of Ageing and Chronic Disease, University of Liverpool), and a Champion™ pET151 Directional TOPO® vector (Life Technologies) expressing residues 41-175 of AGR2 ($AGR2_{41-175}$) with a cleavable N-terminal polyhistidine tag (Patel, 2013, see Appendix 1).

2.6.2 Primer design

Primers were designed using the Invitrogen OligoPerfect ™ Designer [736]. Melting temperature was set at 60°C (range: 57-63°C) and GC content at 50 % (range: 20-80 %). After generation, primers were manually edited by addition or deletion of bases within the sequence so that the final base at the 3' end was either guanine or cytosine, to aid anchoring of the primer to the template strand. A 5' XbaI restriction site was added to forward primers where required and a 5' EcoRI site added to the reverse primer, where required (Table 2.4). Where restriction sites were added, further guanine and/or cytosine nucleosides were added to the 5' end of the restriction site to aid better digestion of the sequence, as described by New England Biolabs [737]. Primers were then synthesised by Integrated DNA Technologies (IDT).

Primer	Sequence	Tm (°C)
WT AGR2 Forward	GCTCTAGAATGGAGAAAATTCCAGTGTCAGC	60.4
WT AGR2 Reverse	CCGGAATTCTTACAATTCAGTCTTCAGCAACTG	60.7
E60A AGR2 Forward	CTGGACTCAGACATATGAAGAAGCTCTATATAAATCCAAGAC	60.8
E60A AGR2 Reverse	GTCTTGGATTTATATAGAGCTTCTTCATATGTCTGAGTCCAG	60.8
Y63A AGR2 Forward	ACTCAGACATATGAAGAAGCTCTAGCTAAATCCAAGACAAGCAACAAACC	65.4
Y63A AGR2 Forward	GGTTTGTTGCTTGTCTTGGATTTAGCTAGAGCTTCTTCATATGTCTGAGT	65.4
K64A AGR2 Forward	GACATATGAAGAAGCTCTATATAAATCCAAGACAAGCAACAAACCC	62.6
K64A AGR2 Reverse	GGGTTTGTTGCTTGTCTTGGATTTATATAGAGCTTCTTCATATGTC	62.6
Δ1-20 AGR2 Forward	GCTCTAGAATGAGAGATACCACAGTCAAACCTGG	61.9
Δ21-40 AGR2 Forward (incl. signal sequence)	GCTCTAGAATGGAGAAAATTCCAGTGTCAGCATCTTGCTCCTTGTGGCCCTCTC CTACACTCTGGCCCCCCAGACCCTCTCCAGAGG	73.0
ΔKTEL AGR2 Reverse	CCGGAATTCTTAGAGAGCTTTCTTCATGTTGTCAAGC	62.6

Table 2.4. ***Primers used for AGR2 subcloning and mutagenesis.*** Restriction sites (if present) are underlined.

2.6.3 Sub-cloning of AGR2

The open reading frame (ORF) containing only the AGR2 coding sequence was amplified from the pBK-CMV-AGR2 into the PiggyBac transposon vector (catalogue no. PB533A-2, see Appendix 2) using the primers detailed in Table 2.4, and using REDTaq® ReadyMix™ PCR reaction mix (Sigma) with the method described by the manufacturer, but using the following cycling parameters:

PCR reaction mix	
Template	10 ng
Forward primer	125 ng
Reverse primer	125 ng
REDTaq ReadyMix	12.5 µL
PCR water	to 25 µL

PCR cycling parameters			
Step	Temp. (°C)	Time (seconds)	
Initial denaturation	95	120	
Denaturation	95	30	} 30 cycles
Annealing	60	30	
Elongation	72	30	
Final elongation	72	420	

After amplification, the reaction mix was run on a 0.7 % (w/v) agarose gel (see Section 2.7) to separate the amplified insert from parental plasmid DNA. The insert was excised and purified using a QuiaQuick Gel Extraction kit (Qiagen), as described by the manufacturer. 1 µg each of the PiggyBac vector and insert were then digested with XbaI and EcoRI HF enzymes (New England Biolabs) for 1 h at 37°C, as described by the manufacturer. Enzymes were then heat inactivated for 20 min at 65°C and the digested DNA was purified using a QiaQuick PCR Cleanup kit (Qiagen). 0.025 pmol of digested vector was mixed with 0.075 pmol digested insert and ligated overnight at room temperature using T4 DNA ligase (New England Biolabs), as described by the manufacturer. The ligase was inactivated by incubation at 65°C for 10 min and One-shot ® Top10 chemically competent *E.coli* cells were then transformed with 5 µL of ligation mixture as described in Section 2.5.2.

2.6.4 Mutagenesis

Due to the size of the PiggyBac expression vector (approximately 7,400 bp with AGR2 insert), point mutations were introduced into the pBK-CMV-AGR2 expression vector and pET151 expression vector by site-directed mutagenesis using the QuickChange Lightning kit (Agilent), and mutated inserts were then subcloned into the PiggyBac vector as described above. The following reaction conditions were used for site-directed mutagenesis:

PCR reaction mix	
Template	10 ng
Forward primer	125 ng
Reverse primer	125 ng
REDTaq ReadyMix	12.5 µL
10 x reaction buffer	5 µL
dNTP mix	1 µL
QuickSolution	1.5 µL
QuickChange Lightning enzyme	1 µL
PCR water	to 50 µL

Site-directed mutagenesis cycling parameters			
Step	Temp. (°C)	Time (min.)	
Initial denaturation	95	2	
Denaturation	95	1	18 cycles
Annealing	55	1	
Elongation	68	14	
Final elongation	68	14	

Methylated, parental DNA was removed with treatment with *Dpn1* enzyme for 1 h at 37°C, and 2 µL *Dpn1*-digested plasmid DNA was transformed into XL10-Gold ultracompetent *E.coli* cells in the same way as described for One-shot ® Top10 chemically competent cells in Section 2.5.2.

For the Δ1-20 AGR2 and ΔKTEL AGR2 truncation mutations, AGR2 was amplified from the pBK-CMV-AGR2 vector using the appropriate primers detailed in Table 2.4. For the Δ21-40 AGR2 truncation, a forward primer containing the AGR2 signal sequence (Table 2.4) was used to amplify the 41-175 region of AGR2 from the pET151 expression vector. In this way, the signal sequence was joined directly to residue 41, effectively deleting the 21-40

region. All truncation mutations were amplified using REDTaq ® ReadyMix ™ PCR reaction mix and were subsequently purified, digested and ligated as described in Section 2.6.3.

The presence of DNA inserts was checked by submitting purified DNA (see Section 2.6.5) to restriction digest with XbaI and EcoRI as described in Section 2.6.3. Appropriate base changes were confirmed by DNA sequencing of purified DNA (Beckman Coulter Genomics).

2.6.5 DNA purification

For use in cloning, mutagenesis and sequencing reactions, transformed cells were grown as 5 mL cultures as described in Section 2.5.2. Cultures were then centrifuged at 3,500 rpm in a Sigma 4K15 refrigerated centrifuge for 10 min at 4°C to pellet the bacteria, and plasmid DNA was isolated using a QIAprep® spin mini kit (Qiagen), using the protocol described by the manufacturer. Purified DNA was stored at -20°C.

For use in transfection of mammalian cells, transformed cells were picked from a single colony on an agar plate and grown in 100 mL LB broth for 20 h in a 37°C incubator with agitation at 220 rpm. Cells were then pelleted by centrifugation at 3,500 rpm as above and plasmid DNA purified using a Plasmid Midi Prep kit (Qiagen), using the protocol described by the manufacturer. Purified DNA was stored at -20°C.

DNA concentration was measured by absorbance at 260 nm using a NanoDrop 1000 spectrophotometer (Thermo Scientific) and DNA purity was monitored by the ratio of absorbance at 260 nm to 280 nm.

2.6.6 Agarose gel electrophoresis

TAE buffer (50 x)
2 M Tris-HCl, pH 8.0 1 M glacial acetic acid 50 mM EDTA

Agarose gels were prepared by adding the required amount of agarose (Bioline) into 1 x TAE buffer, diluted from a concentrated stock with dH_2O. Agarose was dissolved by heating in a microwave oven, and when cooled to approximately 50°C, Midori Green

nucleic acid dye (Nippon Genetics) was added at a final dilution of 1:20,000 and the solution was poured into a horizontal casting tray. After solidification, gels were submerged in 1 x TAE buffer in a horizontal electrophoresis tank (Geneflow). Samples diluted 1:5 with 6 x DNA loading buffer (New England Biolabs) were loaded, along with 5 µL of a 100 bp DNA ladder (New England Biolabs) or 1 kbp ladder (Fermentas) and gels were run until the desired separation was achieved at 100 V at room temperature. Bands were visualised using a UV transilluminator and gels were imaged using a Bio-Rad Gel Doc EQ system.

2.7 Recombinant protein purification

2.7.1 Recombinant protein expression and cell lysis

One-shot ® BL21 star ™ chemically competent E. coli cells were transformed and 100 mL starter cultures were prepared as described in Section 2.5.2. These cultures were diluted into 500 mL LB broth so as to reach a starting optical density (O.D.) at 600 nm of 0.1. Cultures were then placed in a 37°C incubator with agitation at 220 rpm for approximately 2 h, until reaching an O.D. 600 nm of between 0.6 and 0.8. At this point, 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG), previously made up in dH_2O, sterile filtered through a 0.22 µm filter and frozen at -20°C, was added at a final concentration of 1 mM and cultures were grown for 16-20 h in an 18°C incubator with agitation at 180 rpm.

Cultures were subsequently centrifuged at 4,500 rpm in a Fiberlite™ F9-4 x 1000y rotor for 20 min at 4°C, and pelleted bacteria were resuspended in a minimal volume of the appropriate lysis buffer on ice (see below). Resuspended cell pellets were frozen at -80°C for future use or lysed using a Soniprep 150 plus sonicator (MSE), set at the highest amplitude. Lysis was performed on ice and was achieved by 10 successive cycles of 30 s. sonication followed by 60 s. rest. Lysates were subsequently centrifuged at 15,000 rpm in a Sorvall SS-34 Rotor (Thermo Scientific) for 30 min at 4°C to pellet cell debris. Supernatants were then filtered through a 0.22 µm filter on ice and protein was purified using the appropriate method described below.

2.7.2 Purification of AGR2 $_{21-175}$ from IMPACT-TWIN$_1$ vector

IMPACT-TWIN$_1$ vector lysis buffer
20 mM HEPES, pH 8.0 500 mM NaCl 1 mM EDTA

Chitin column elution/cleavage buffer
20 mM HEPES, pH 6.5 500 mM NaCl 1 mM EDTA

Chitin column stripping buffer
300 mM NaOH 1 % (w/v) SDS

Chitin column regeneration buffer
300 mM NaOH

All steps were performed under gravity flow and all buffers were made with ultrapure water. All buffers (including ultrapure water) were filtered and degassed through a 0.22 µm bottle top filter using a vacuum pump prior to use. Chitin resin slurry (New England Biolabs) was poured into an empty PD-10 column (GE Life Sciences) to form a 6 mL bead bed. Beads were washed with 10 column volumes (CV) ultrapure water and equilibrated with 10 CV lysis buffer, as described by the manufacturer. Filtered protein expression supernatant (see above) was passed through the column 3 times at room temperature, each time collecting the flowthrough on ice before re-applying it to the column. The column was then extensively washed with 20 CV of lysis buffer before being flushed with 3 CV of elution buffer. The column was left in elution buffer overnight at room temperature to allow the cleavage reaction to take place (Fig. 2.5).

After overnight incubation, cleaved protein was eluted with 2 CV elution buffer and 1 ml elution fractions were collected. Protein concentration was measured using a NanoDrop 1000, reading absorbance at 280 nm. This value was corrected using the theoretical extinction co-efficient at 1 g/L (1.203) predicted from the Expasy ProtParam server [738]. The five most concentrated elution fractions were combined, filtered through a 0.22 µm filter and subjected to gel filtration (see below).

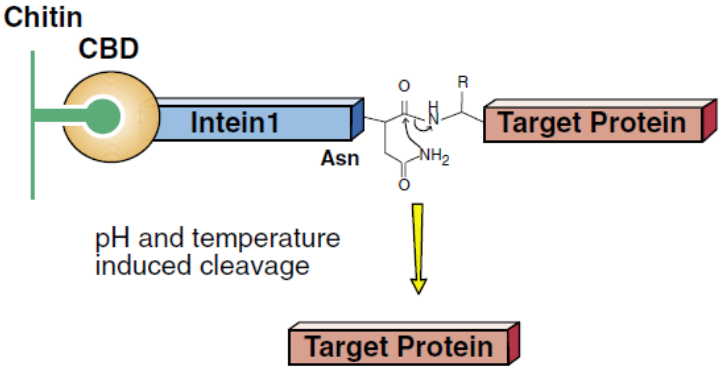


Figure 2.5. ***Mechanism of pH-induced cleavage of Intein-tag.*** By acidifying the pH on the chitin column, the C-terminal Asn residue in the Intein-tag undergoes cyclisation of its side-chain, resulting in cleavage of the peptide bond between the tag and target protein. Reproduced from NEB.com [737].

2.7.2.1 Regeneration of chitin beads

Chitin-binding tags and uncleaved protein were stripped from the column by passing 3 CV stripping buffer through the column at room temperature. The column was then washed with 3 CV regeneration buffer, soaked for 30 min in this buffer and then washed with another 7 CV regeneration buffer. The column was washed with 20 CV ultrapure water and then 5 CV 20 % (v/v) ethanol, before storing the column at 4°C. Columns were regenerated after each separate purification and any one column was regenerated no more than five times.

2.7.3 Purification of AGR2 $_{41\text{-}175}$ from pET151 vector

pET151 vector lysis buffer
20 mM Tris-HCl, pH 7.5 500 mM NaCl 20 mM imidazole

HisTrap FF elution buffer
20 mM Tris-HCl, pH 7.5 500 mM NaCl 500 mM imidazole

HisTrap FF stripping buffer
20 mM Tris-HCl, pH 7.5 500 mM NaCl 50 mM EDTA

All steps were performed using an ÄKTA purifier FPLC system unless otherwise stated and all buffers were made with ultrapure water. All buffers (including ultrapure water) were filtered and degassed through a 0.22 µm bottle top filter using a vacuum pump prior to use.

5 mL HisTrap FF columns (GE Life Sciences) were charged with nickel by passing 2 CV 100 mM $NiSO_4$ through the column and excess $NiSO_4$ was removed with 4 CV washes ultrapure water. Filtered protein expression supernatant (see Section 2.7.1) was manually loaded onto the column at room temperature using a syringe. Flowthrough was collected on ice and passed through the column a second time. The column was then connected to the ÄKTA purifier and washed with 5 CV pET151 vector lysis buffer. Elution buffer was added on a gradient up to 100 % (v/v) over 10 CV and His-tagged eluted protein was collected using an automated fraction collector. Removal of all column-bound protein was achieved by passing through 5 CV elution buffer.

Protein concentration was measured using a NanoDrop 1000, reading absorbance at 280 nm. This value was corrected using the theoretical extinction co-efficient at 1 g/L (1.325) predicted from the Expasy ProtParam server [738]. The five most concentrated elution fractions were combined and the histidine tag was cleaved by incubation with recombinant TEV (tobacco etch virus) protease (kindly provided by Prof. Lu-Yun Lian, Institute of Integrative Biology, University of Liverpool) and 1 mM dithiothreitol (DTT). Imidazole is inhibitory to TEV activity [522], therefore the cleavage reaction was performed with simultaneous dialysis. 5 mL purified protein with TEV protease was first dialysed against 1 L gel filtration buffer (see below) for 4 h at room temperature and then dialysed again against fresh gel filtration buffer overnight at 4°C. Protein was then dialysed against 1 L fresh gel filtration buffer for a third time for 4 h at room temperature the next day. The extent of histidine tag cleavage was then assessed by running 5 µL of the cleavage reaction on an SDS-PAGE gel and staining with Coomassie brilliant blue. If cleavage was insufficient, the reaction was allowed to proceed for a further 1-4 h at room temperature.

The histidine tag and residual uncleaved protein was removed from the sample by manually passing it once through a 5 mL HisTrap FF column at room temperature using a syringe. Flowthrough containing cleaved protein was collected on ice and protein concentration measured as described above. Suitable fractions were concentrated using Amicon® Ultra-15 centrifugal spin cartridges (Millipore), by centrifuging at 4,500 rpm at 4°C in a Sigma 4K15 refrigerated centrifuge until the volume was reduced to 5 mL or less. This concentrated protein was then subjected to gel filtration (see below).

2.7.3.1 Column regeneration

Following each purification, columns were washed with 4 CV ultrapure water and the $NiSO_4$ was stripped from the column by manually passing through 4 CV stripping buffer. Columns were washed with 4 CV ultrapure water and 4 CV 20 % (v/v) ethanol was passed through, before storage at 4°C. Before the next use, columns were first washed with 4 CV ultrapure water before charging with $NiSO_4$.

2.7.4 Gel filtration

Gel filtration buffer
20 mM sodium phosphate, pH 6.5 150 mM NaCl

All buffers were made with ultrapure water and all buffers (including ultrapure water) were filtered and degassed through a 0.22 µm bottle top filter using a vacuum pump prior to use.

A Superdex 75 26/60 column (GE Life Sciences) was equilibrated overnight at room temperature with gel filtration buffer, using an ÄKTA purifier. Up to 5 mL of purified protein was then manually loaded onto the system using a syringe and the system was run at room temperature. Protein was collected using an automated fraction collector and the concentration of protein in these fractions was measured as described in Section 2.7.2. Purity of the collected fractions was assessed by running 5 µL of each fraction on an SDS-PAGE gels and staining with Coomassie brilliant blue (see Section 2.8). Samples were concentrated using Amicon® Ultra-15 centrifugal spin cartridges (Millipore), by centrifuging at 4,500 rpm at 4°C in a Sigma 4K15 refrigerated centrifuge if required. The gel filtration column was washed with 1.5 CV ultrapure water and then 1.5 CV 20 % ethanol and stored at room temperature.

2.7.4 Size-exclusion chromatography – multi-angle laser-light scattering (SEC-MALLS)

Protein samples were prepared as described above, without a gel filtration step. Except where noted, all samples were buffer-exchanged into 20 mM $NaPO_4$, 150 mM NaCl using a PD-10 desalting column (GE Life Sciences). Columns were washed with 25 mL ultrapure water under gravity flow and equilibrated with 25 mL buffer. 2.5 mL purified protein was added and eluted with 3.5 mL equilibration buffer, all of which was collected and the protein concentration measured using a NanoDrop 1000. Columns were washed with 25 mL ultrapure water and then 25 mL 20 % (v/v) ethanol, and stored at room temperature. Except where noted, samples were provided at a concentration of 1 mg/mL.

Samples for SEC-MALLS were sent to the University of Manchester for analysis, where they were loaded onto a Superdex 75 10/300 GL column and analysed using EOS 18-angle laser photometer (Wyatt Technology), an rEX refractive index detector (Optilab) and a UV-2077 Plus UV/Vis spectrophotometer (Jasco). Simultaneous measurement of hydrodynamic radius was achieved through a coupled quasi elastic light-scattering detector. Molar mass measurements were performed using both Astra 5.3.2.16 software (Wyatt Technology) and the "three detector method".

2.8 SDS-PAGE

2.8.1 Cell lysis

RIPA lysis buffer
50 mM Tris-HCl, pH 6.8
150 mM NaCl
2 mM EDTA
1 % (v/v) Tergitol® (NP-40, nonyl phenoxypolyethoxylethanol-40)
0.5 % (w/v) sodium deoxycholate
0.1 % (w/v) SDS
Protease inhibitor cocktail

RIPA buffer was prepared in dH_2O and kept at 4°C for up to 4 weeks. Protease inhibitor cocktail (cOmplete EDTA-free, Roche) was prepared in dH_2O, aliquotted and frozen at -20°C. Protease inhibitor cocktail was added to lysis buffer immediately prior to use.

Cells were grown to 70-80 % confluence, placed on ice and washed twice with ice-cold PBS. Cells were scraped into RIPA buffer and placed on an end-over-end mixer for 30 min at 4°C. Lysates were then centrifuged at 12,000 x g for 20 min at 4°C in an Eppendorf 5415R refrigerated centrifuge. Supernatants were collected, transferred to fresh tubes on ice and frozen at -80°C.

2.8.2 Protein concentration quantification

The protein concentrations of cell lysates prepared using RIPA buffer were measured using a BCA protein assay kit (Pierce Biotechnology). A standard curve was produced using a BSA standard (Pierce) diluted in lysis buffer, as described by the manufacturer. Protein samples were dilute 1 in 5 also using lysis buffer. 10 µL BSA protein standard or protein sample was pipetted in triplicate into a 96-well plate. 200 µL BCA reagent was then added using a multi-channel pipette, briefly mixed using a plate shaker and then covered and incubated at 37°C for 30 min. After this time, the plate was left for 5 min at room temperature before reading the absorbance at 562 nm using a Spectramax plus384 (Molecular Devices) plate reader. The concentration of the protein samples was determined from their mean absorbance values, using a linear standard curve produced from the mean absorbance values of the BSA protein standards.

For samples from conditioned medium (CM) collected in DMEM medium, protein concentration was also measured using the BCA method. For CM collected in Opti-MEM medium (see Section 2.12.1), protein concentration was measured by Bradford assay due to an interfering substance within the medium. As with the BCA assay, a standard curve was produced using a BSA standard (Pierce) diluted in Opti-MEM medium. 5 µL BSA protein standard or undiluted CM samples were pipetted in triplicate into a 96-well plate. 200 µL Bradford reagent was then added using a multi-channel pipette, briefly mixed using a plate shaker and incubated at room temperature for 10 min. Absorbance was read at 595 nm and the concentration of the protein samples was determined from their mean absorbance values, using the linear region of the standard curve produced from the mean absorbance values of the BSA protein standards.

For recombinant protein, protein concentration was determined using a NanoDrop 1000 as described previously.

2.8.3 Preparation and running of polyacrylamide gels

Stacking gel
250 mM Tris-HCl, pH 6.8
0.1 % (v/v) SDS
4 % or 5 % acrylamide* (37.5:1 acrylamide:bis-acrylamide)
0.06 % (w/v) ammonium persulfate (APS)#
0.015 % (v/v) TEMED

Resolving gel
750 mM Tris-HCl, pH 8.9
0.1 % (v/v) SDS
8 %, 10 % or 15 % acrylamide* (37.5:1 acrylamide:bis-acrylamide)
0.06 % (w/v) ammonium persulfate (APS)#
0.015 % (v/v) TEMED

* final percentage of acrylamide used as required.

10 % (w/v) solution made up in dH_2O and kept at 4°C for up to 4 weeks.

SDS-PAGE running buffer
50 mM Tris
192 mM glycine
0.1 % (w/v) SDS
pH 8.3

3 x SDS-PAGE loading buffer$
187.5 mM Tris-HCl, pH 6.8
6 % (v/v) SDS
30 % (v/v) glycerol
300 mM DTT
0.625 % (v/v) bromophenol blue

3 x SDS-PAGE loading buffer for conditioned medium samples$
187.5 mM Tris-HCl, pH 6.8
1 % (v/v) SDS
30 % (v/v) glycerol
300 mM DTT
0.625 % (v/v) bromophenol blue

$Aliquotted and frozen at -20°C.

Acrylamide gels were poured into 1 mm-thick glass plates in a vertical casting apparatus (Bio Rad). Resolving gels were overlaid with 200 µL butanol to flatten the gel surface and prevent drying out of the gel, and left to polymerise at room temperature for 45 min. Butanol was washed away using dH_2O and stacking gel was poured over the set resolving gel. 10- or 15-well combs were immediately inserted into the stacking gel and gels were left to polymerise for 15 min at room temperature. Residual unpolymerised acrylamide was removed from the wells using dH_2O, and gels were then placed in a vertical gel electrophoresis tank (Bio Rad) and submerged in 1 x SDS running buffer.

All protein samples were made up to the same volume with dH_2O before the addition of 3 x SDS loading buffer in a 1:2 ratio. Precipitated protein from CM samples

already contain 2 % (w/v) SDS (see Section 2.12.2), so a modified loading buffer containing a lower percentage SDS was added to these samples so that the final concentration of SDS remained at 2 % (w/v). All samples were mixed by vortex and were then heated in a 98°C heating block for 5 min before being briefly centrifuged at 5,000 rpm in an Eppendorf 5415D centrifuge prior to loading on the gel. Gels were run at 220 V for 45-60 min at room temperature, until the bromophenol blue dye front reached the bottom of the gel.

2.8.4 Coomassie brilliant blue staining

Coomassie brilliant blue stain
0.1 % (w/v) Coomassie brilliant blue
50 % (v/v) methanol
7 % (v/v) glacial acetic acid

Destain
20 % (v/v) methanol
10 % (v/v) glacial acetic acid

Electrophoresed gels were removed from their glass plates and rinsed twice in reverse osmosis (RO) water. Gels were then submerged in Coomassie brilliant blue stain for at least 30 min at room temperature on an oscillating platform. Gels were then rinsed several times in RO water to remove all excess stain and then submerged in destain buffer until protein bands became visible and background staining was suitably reduced. Destain buffer removed and fresh buffer was added as required. Gels were scanned using an ImageScanner III gel scanner (GE Life Sciences).

2.8.5 Western blotting

Western blot transfer buffer*
120 mM Tris
192 mM glycine
20 % (v/v) methanol

TBST-T
20 mM Tris-HCl, pH 7.5
150 mM NaCl
0.1 % (v/v) Tween20

* Made fresh.

For the analysis of mammalian cell lysates, 10 µg of protein were loaded per well. For CM samples, either 20 µL of sample were loaded per well for qualitative experiments or samples were equally loaded based on the maximum amount of protein that could be loaded for the least concentrated sample in the experiment, e.g. if the least concentrated

sample had a concentration of 0.1 µg/mL, 20 µL of this sample was loaded on the gel, along with 2 µg/well of all other samples.

After SDS-PAGE, proteins were transferred onto Immobilon-P PVDF membrane (Millipore) via wet transfer. Gels were removed from their casting plates and rinsed twice in RO water before being submerged in freshly prepared transfer buffer for 15 min at room temperature. During this time, PVDF membrane was activated by submersion in 100 % methanol for 30 s. at room temperature, before being submerged in transfer buffer, along with sponges and filter paper. The transfer apparatus was assembled, whilst maintaining the whole apparatus submerged in transfer buffer to prevent the generation of air bubbles. The apparatus was then placed in a vertical electrophoresis tank and filled with transfer buffer and a -80°C freezer pack. The tank was placed in an ice tray and electroblotting was performed at 100 V for 90 min.

After transfer, membranes were rinsed twice with RO water and stained with ATX Ponceau S stain (Fluka) to check the efficiency of transfer. Membranes were then briefly washed in TBS-Tween 20 (TBS-T) to remove the stain. Membranes were then blocked for 1 h at room temperature in blocking buffer (5 % (w/v) skimmed-milk powder (Marvel) in TBS-T), on an oscillating platform. Membranes were then placed in 50 mL tubes and incubated with primary antibody (Table 2.5) in blocking buffer overnight at 4°C on a rolling mixer.

The following day, membranes were washed 3 x 10 min in TBS-T at room temperature. They were then placed in 50 mL tubes and incubated with secondary antibody (Table 2.5) in blocking buffer for 1 h at room temperature on a rolling mixer. Membranes were then washed a further 3 x 10 min in TBS-T. Excess liquid was removed and membranes were incubated with 1 mL Amersham™ ECL™ substrate (GE Life Sciences) per membrane, as described by the manufacturer. Excess reagent was removed and membranes were exposed to Super RX X-ray film (Kodak) for 1-60 min and then manually developed, in a dark room. If signal was weak or non-existent after 60 min exposure, membranes were briefly rinsed in TBS-T and incubated with 1 mL Amersham™ ECL Select™ substrate (GE Life Sciences) for 5 min at room temperature, before being re-exposed to X-ray film.

Primary Antibodies				
Antibody	Host Species	Company	Product Code	Dilution Used
AGR2	Mouse	Millipore	MABC48	1:5,000
LDHA	Rabbit	Cell Signaling Technology	2012	1:2,000
β-actin	Mouse	Sigma	A5316	1:20,000

Secondary Antibodies			
Antibody	Company	Product Code	Dilution Used
HRP-conjugated Rabbit anti-mouse IgG	Dako	PO260	1:5,000
HRP-conjugated Swine anti-rabbit IgG	Dako	PO217	1:5,000

Table 2.5. ***Antibodies used for Western blotting.***

2.8.6 Densitometry

Developed X-ray film was scanned using a flatbed scanner (Ricoh). Densitometry was performed on developed blots or scanned SDS-PAGE gels using ImageStudio Lite software (LI-COR Biosciences).

2.9 Immunoprecipitation (IP)

IP wash buffer
10 mM Tris-HCl, pH 7.5
150 mM NaCl
2 mM EDTA
1 % (v/v) Triton X-100

2 x SDS-PAGE loading buffer
125 mM Tris-HCl, pH 6.8
4 % (w/v) SDS
30 % (v/v) glycerol
200 mM DTT
0.415 % (w/v) bromophenol blue

Cell were lysed in RIPA buffer as described in Section 2.8.1. 100 µL Protein A/G bead slurry (Pierce) was centrifuged for 15 s. at 10,000 rpm in an Eppendorf 5415D centrifuge, the buffer removed and the beads washed 3 times with 1 mL PBS, with centrifugation for 15 s. at 10,000 rpm in between washes. 200 µL cleared lysate was added to the washed beads and incubated for 1 h on an end-over-end mixer at 4°C to remove proteins that bind non-specifically to the Protein A/G beads. The sample was then centrifuged for 5 min at 10,000 rpm at 4°C in a Eppendorf 5415R refrigerated centrifuge.

The supernatant was placed into a fresh tube on ice, 5 µL AGR2 antibody (Millipore) was added and the tube placed on an end-over-end mixer overnight at 4°C.

The next day, 30 µL Protein A/G bead slurry was washed 3 times in PBS as above and the lysate/antibody mix was incubated with bead slurry for 4 h at 4°C on an end over end mixer. Beads were centrifuged for 5 min at 10,000 rpm in an Eppendorf 5415R refrigerated centrifuge at 4°C and the supernatant removed. Beads were then washed 3 times with 1 mL IP wash buffer, with mixing by brief vortexing and centrifugation for 5 min at 10,000 rpm at 4°C in between washes. After the final wash and centrifugation, 30 µL 2 x SDS loading buffer was added to the beads and heated for 5 min in a 98°C heat block. Samples were centrifuged for 5 min at 10,000 rpm in an Eppendorf 5415D centrifuge and the supernatant was loaded onto an SDS-PAGE gel and analysed by Western blotting as described in Section 2.8.

2.10 Direct ELISA assay

ELISA wash buffer
50 mM $NaPO_4$, pH 7.2
150 mM NaCl
0.05 % (v/v) Tween20

AGR2 protein and peptide were coated onto 96-well non-tissue culture-treated plates in the same way as described in Section 2.3.2. After coating, plates were washed once with 100 µL PBS and then incubated with 200 µL Superblock blocking buffer (Thermo Scientific) for 2 h at room temperature. Blocking buffer was removed and wells were incubated with 100 µL AGR2 antibody (Millipore, Table above) diluted 1 in 5000 in Superblock T20 blocking buffer (Thermo Scientific) at 4°C overnight.

The next day, wells were washed 6 times with 100 µL wash buffer and then incubated with 100 µL HRP-conjugated Rabbit anti-mouse IgG (Dako) diluted 1 in 10,000 in Superblock T20 blocking buffer for 2 h at room temperature. Wells were then washed a further 6 times in 100 µL wash buffer. Bound antibody was detected by the addition of 80 µL 1-Step Ultra TMB-ELISA substrate (Thermo Scientific) and colour was allowed to develop for up to 30 min at room temperature. The reaction was quenched by the addition of 80 µL 2 M sulphuric acid and plates were read at 450 nm in a Spectramax plus384 plate reader (Molecular Devices).

2.11 Biochemical assays for PDI activity

2.11.1 Insulin turbidity assay

The assay was based on the method described by Holmgren, 1979 [523]. In a 96-well clear plate, reagents were added to a final concentration of 50 mM $NaPO_4$ (pH 7.0), 2mM EDTA, 1 mg/mL insulin from bovine pancreas (Sigma) and either 0.5 µM recombinant thioredoxin from *E. coli* (Sigma) or 5 µM recombinant AGR2 (prepared as described in Section 2.7) in a final reaction volume of 200 µL. The reaction was started by the addition of DTT to a final concentration of 1 mM and the absorbance at 650 nm was measured every 5 min for 90 min in a Spectramax plus384 plate reader (Molecular Devices) set at 25°C. A control reaction without thioredoxin or AGR2 was also included. Reactions were carried out in duplicate.

2.11.2 Glutaredoxin assay

The assay was based on the method described by Holmgren, 1979 [524]. In a 96-well clear plate, reagents were added to a final concentration of 100 mM Tris-HCl (pH 8.0), 2 mM EDTA, 1 mM GSH, 0.4 mM NADPH, 100 µg/mL BSA and 6 µg/mL recombinant human glutathione reductase from *E.coli* (Sigma) in a final reaction volume of 200 µL. The substrate 2-hydroxyethyl disulphide (HED) was then added to a final concentration of 0.7 mM and after a 2-min incubation at room temperature, recombinant human glutaredoxin-1 from *E.coli* (Sigma) or recombinant AGR2 (prepared as described in Section 2.7) was added to a final concentration of 5 µM. The absorbance at 340 nm was then measured every 2 min for 120 min in a Spectramax plus384 plate reader (Molecular Devices) set to 25°C. A control reaction without glutaredoxin or AGR2 was also included. Reactions were carried out in duplicate.

2.12 Collection and analysis of cell secretomes

2.12.1 Collection of condition medium (CM)

Cells were grown to 30-40 % confluence in normal medium and then washed 5 times in PBS supplemented with 0.9 mM $CaCl_2$ and 0.5 mM $MgCl_2$ to remove serum proteins whist limiting the amount of cell detachment. Cells were then incubated for 24 h in a 37°C, 10 % (v/v) CO_2 incubator in 7 mL Opti-MEM medium without phenol-red, supplemented with final concentrations of 25 mM glucose, 3.7 g/mL sodium bicarbonate, 1.8 mM $CaCl_2$, 4 % (w/v) NEAA, 100 units/mL penicillin, 100 µg/mL streptomycin and 10 ng/mL hydrocortisone, so that this serum-free medium was as close to the formulation of normal Rama 37 medium as possible.

After 24 h, CM was collected on ice, cells washed with 1 mL ice-cold supplemented Opti-MEM medium and the combined CM and wash were centrifuged at 1000 x *g* for 5 min at 4°C in a Sigma 4K15 refrigerated centrifuge to pellet any large cellular debris. Supernatants were then centrifuged at 100,000 x *g* for 1 h at 4°C in a in a Sorvall Combi Plus ultracentrifuge (DuPont Instruments). Supernatants were removed and stored at -80°C for future use or directly subjected to trichloroacetic acid (TCA) precipitation.

2.12.2 TCA precipitation

Solubilisation buffer
50 mM Tris-HCl, pH 8.0 2 % (w/v) SDS

100 % (w/v) TCA solution was prepared by adding 227 mL ultrapure water to 500 g TCA powder. The powder was dissolved by gentle swirling at room temperature and the solution was subsequently stored at 4°C.

On ice, N-lauroylsarcosine (N-LS) solution was added to CM samples to a final concentration of 0.1 %* (w/v), vortexed and then TCA was added to a final concentration of 7.5 %* (w/v). Samples were briefly mixed by vortexing and incubated at -20° for 2 h. After this time, samples were thawed at room temperature and 1.5 mL fractions were sequentially centrifuged at 10,000 x *g* for 10 min at 4°C in an Eppendorf 5415R refrigerated centrifuge in the same tube, until the whole sample had been processed. The pellet containing precipitated proteins was then washed twice with 1 mL ice-cold acetone, with

mixing by vortexing and centrifugation at 10,000 x *g* for 10 min at 4°C. After the second wash, the acetone was removed and excess acetone was allowed to evaporate for 5-10 min at room temperature in a fume hood. Pellets were incubated with 10 µL (or 1 volume) 200 mM NaOH on ice for 5 min to aid protein solubilisation [525]. 90 µL (or 9 volumes) 1.1 x solubilisation buffer was added and the samples were subjected to two cycles of mixing by vortexing and a 10 min incubation in a sonicating waterbath at room temperature. The samples were then heated for 5 min in a 98°C heat block before a final 10 min incubation in a sonicating waterbath at room temperature. Samples were then centrifuged for 1 min at 13,000 rpm in an Eppendorf 5415D centrifuge to pellet any remaining insoluble material and the supernatant frozen at -80°C. Supernatants were analysed by Western blot as described in Section 2.8 .

* Concentrations are indicated for the final concentration of both reagents, after the addition of both of these reagents.

2.12.3 Periodic acid-Schiff (PAS) staining

CM from eight 10 cm cell culture plates collected from cells in supplemented Opti-MEM medium as described above and combined, but instead of TCA precipitation, CM was concentrated down to approximately 1 mL using Amicon® Ultra-15 centrifugal spin cartridges (Millipore) by centrifuging at 4,500 rpm at 4°C in a Sigma 4K15 refrigerated centrifuge. Samples were concentrated in this manner rather than by TCA precipitation as precipitated protein could not be efficiently immunoprecipitated. The concentrated sample was subjected to immunoprecipitation (IP) as described in Section 2.9, except that 10 µL AGR2 antibody was added and the next day the sample was incubated with 100 µL Protein A/G bead slurry. After IP, the beads were incubated with 30 µL 2x SDS sample buffer, heated for 5 min in a 98°C heat block and run on an SDS-PAGE gel as described in Section 2.8.

All PAS staining steps were performed at room temperature. Following SDS-PAGE, gels were fixed in 50 % (v/v) methanol for 30 min and then washed for 3 x 5 min in 3 % (v/v) acetic acid. Gels were then incubated with 1 % (w/v) freshly prepared sodium periodate (Sigma) in 3 % (v/v) acetic acid for 30 min in the dark and then washed 3 more times with 3 % (v/v) acetic acid for 5 min each, prior to the addition of Schiff's Fuchsin-sulphite reagent (Sigma) for 15 min. Gels were then washed a final time in 3 % (v/v) acetic acid for 5 min and

then washed several times in RO water. Gels were scanned using an ImageScanner III gel scanner (GE Life Sciences) and were then counter-stained with Coomassie brilliant blue as described in Section 2.8.4, before being scanned again.

2.12.4 Enzymatic deglycosylation of proteins

For deglycosylation of secreted proteins, CM from eight 10 cm cell culture plates collected from cells in supplemented Opti-MEM medium were combined and concentrated down to approximately 0.5 mL using Amicon® Ultra-15 centrifugal spin cartridges (Millipore) as described above. For deglycosylation of intracellular proteins, cells were grown to 70-80 % confluence and lysed using RIPA buffer as described in Section 2.8.1. Two 10 cm plates were lysed into 200 µL buffer.

Deglycosylation was performed using a Protein Deglycosylation Mix (New England Biolabs) using the method described by the manufacturer. For CM samples, 9 µL concentrated CM was added to 1 µL denaturing buffer, whereas 5 µl cell lysate was added to 1 µL denaturing buffer and made up to 10 µL with dH_2O. Both types of sample were subsequently treated in the same manner. Samples were heated in a heat block at 98°C for 10 min, chilled on ice and then briefly centrifuged. Reaction buffer (50 mM NaCl, 20 mM Tris-HCl (pH 7.5) and 0.1 mM EDTA) and NP-40 were added to a final concentration of 1 x and 1 % (v/v), respectively, before the addition of 2.5 µL deglycosylation enzyme cocktail in a final reaction volume of 25 µl. Samples were then incubated in a 37°C waterbath for 4 h. After incubation, samples were briefly centrifuged before the addition of 3 x SDS loading buffer and samples were then subjected to analysis by Western blot as described in Section 2.8.

2.12.4.1 PNGase F treatment of proteins

Samples were resuspended in reaction buffer consisting of a final concentration of 50 mM $NaPO_4$ (pH 7.5), 40 mM DTT and 0.1 % (w/v) SDS, and heated at 98°C for 5 min. After brief centrifugation, NP-40 was added to a final concentration of 0.75 % (v/v), before the addition of PNGase F (Peptide-*N*-Glycosidase F, Europa Bioproducts) to a final concentration of[312, 377, 383, 457, 458, 526-535] 50 mU/mL. Samples were incubated overnight at room temperature. Deglycosylation of α-acid glycoprotein (α-AG) was used as

a positive control for these experiments. Note that concentrations denote final concentrations, after the addition of all reagents. PNGase F enzyme and recombinant α-AG protein were kindly provided by Dr Mark Wilkinson (Institute of Integrative Biology, University of Liverpool).

2.12.5 λphosphatase treatment of proteins

Samples were resuspended in NEBuffer for protein metalloproteinases (PMP, New England Biolabs) consisting of a final concentration of 50 mM HEPES (pH 7.5), 10 mM NaCl, 2 mM DTT and 0.01 % (v/v) Brij 35 detergent. $MnCl_2$ was added at a final concentration of 1 mM. Recombinant λphosphatase was then added to a final concentration of 10 µg/mL and samples were incubated at 30°C for 30 min. Dephosphorylation of recombinant Mps1 (monopolar spindle 1) was used as a positive control for these experiments. Recombinant λphosphatase and recombinant Mps1 were kindly provided by Dr Patrick Eyers (Institute of Integrative Biology, University of Liverpool).

2.13 LC-MS analysis

2.13.1 Analysis of conditioned medium

CM was collected and processed as described in Section 2.12.1 from quadruplicate plates and samples were sent for analysis by the Centre of Proteome Research, Institute of Integrative Biology at the University of Liverpool. Protein concentration was determined by Bradford assay as described in Section 2.8.2. Equal amounts of protein were prepared in a 1 mL reaction volume and captured onto 10 µL Strataclean beads (Agilent Technologies) by vortexing for 1 min. The process was repeated to capture proteins from 2 mL sample onto 20 µl beads. Beads were pelleted by brief centrifugation and then washed in 1 mL 25 mM NH_4HCO_3 before being resuspended in fresh 25 mM NH_4HCO_3. Samples were subsequently incubated with a final concentration of 0.05 % (v/v) RapiGest SF™ (Waters) in 25 mM NH_4HCO_3 for 10 min at 80°C, before the addition of 60 mM DTT for 10 min at 60°C and then 178 mM idoacetamide for 30 min at room temperature in the dark. Gold mass spectrometry grade trypsin (Promega, 1 µg) was then added and samples were incubated

at 37°C for 45 min and then centrifuged for 30 min at 17,000 x *g* at room temperature. 10 µL of supernatant was then transferred to a Total Recovery Vial (Waters) and placed in an instrument auto sampler. 1 µL of sample was injected onto a trap column (Acclaim® PepMap 100, Dionex/Thermo Scientific) in buffer containing 0.1 % (v/v) trifluoroacetic acid (TFA) and 2 % (v/v) acetonitrile at a flow rate of 5 mL/min. After 3 min, the trap column was set in-line with an Easy-Spray PepMap® RSLC analytical column (Dionex/Thermo Scientific), and peptides were separated over a linear gradient up to 40 % (v/v) acetonitrile in 0.1 % (v/v) formic acid, at a flow rate of 300 nL/min, using an UltiMate® 3000 Nano HPLC system (Dionex/Thermo Scientific). MS/MS analysis of eluted peptides was then performed on a Q Exactive™ Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Scientific).

Progenesis LC-MS label-free quantification software (Nonlinear Dynamics) was used for protein quantification. Data were normalised using an in-built algorithm that considers all peptide features and compares them to a chosen reference sample. The software assumes that most of the peptide abundances will not change between samples and calculates a global scaling factor, based on the mean scaling factor for each feature and iterating the process to remove any influencing outlying values. The top 5 MS/MS spectra for each peptide were then searched using the Mascot search engine and database (Matrix Science), and using a 1 % false discovery rate. The peptide score threshold for this false discovery rate was then used as a cut-off score for peptides within Progenesis.

For CM experiments, samples could not be normalised to protein concentration due to the presence of non-cell derived culture medium proteins at a concentration of 15 µg/mL, which masked the concentration of cell-derived proteins. Therefore, for normalisation of these samples, protein was captured onto Strataclean beads and digested as described above. The digests were then analysed using a 1 h elution gradient LC-MS method to obtain an approximate normalisation value from Progenesis. Using this value, normalised amounts of sample were then loaded and analysed over a 2 h elution gradient method, and differences in protein abundance identified using Progenesis. Subcellular distribution of proteins, pathway analysis and construction of regulatory/interaction networks was performed using Ingenuity Pathway Analysis software (Qiagen).

2.13.2 Analysis of whole cell lysates

Cells were grown for two passages in DMEM without phenol-red, supplemented with 5 % (v/v) FBS, 4 % (v/v) NEAA, 4 mM L-glutamine, 10 ng/mL insulin, 10 ng/mL

hydrocortisone, 100 units/mL penicillin and 100 µg/mL streptomycin. Upon reaching 70-80 % confluence, cells from five 10 cm cell culture plates were washed 3 times with PBS and detached by treatment with an enzyme-free dissociation buffer (Enzyme-free cell dissociation solution (PBS-based), Millipore) for 30-45 min in a 37°C, 10 % (v/v) CO_2 incubator. Cells were collected in PBS, combined and centrifuged for 5 min at 1000 x *g* at room temperature to remove the dissociation reagent. Cells were washed once more in PBS, centrifuged for 5 min as before and resuspended in 10 mL PBS. 100 µL of this cell suspension was counted using a Z1 Coulter® particle counter (Beckman Coulter) and 1 x 10^7 cells were transferred to a fresh tube. These cells were pelleted by centrifugation for 5 min as before, the supernatant removed and the cell pellets were snap-frozen in liquid nitrogen and stored at -80°C until needed. Cells were collected in this manner in quadruplicate.

Samples were sent for analysis by the Protein Function Group, Institute of Integrative Biology, at the University of Liverpool. Cell pellets were thawed, resuspended in 25 mM NH_4HCO_3 and lysed by sonication. DNA was removed by digestion with 25 units benzonase nuclease (Novagen). Protein concentration was then determined by Bradford assay and 100 µg of protein was digested and analysed as described for CM samples, with 500 ng being injected onto the trap column, as above. Whole cell samples were directly separated over a 2 h gradient as they had already been normalised to protein concentration.

2.14 Statistics

Statistical analyses were performed using the SPSS software package, version 20 (IBM).

Chapter 3

Characterisation of Recombinant AGR2 and AGR2-expressing Stable Cell Lines

3.1 Introduction

AGR2 is a member of the protein disulphide isomerase (PDI) family of endoplasmic reticulum (ER) chaperones, based on its sequence homology to the PDI ERp18 and its housing of a divergent PDI active site motif with the sequence CxxS, which normally sits within a thioredoxin fold [378]. Several other structural features within the AGR2 sequence have been identified and are outlined in Fig. 3.1A. The protein contains an N-terminal signal peptide that promotes its translocation into the ER [536], where it is mostly resident [417, 422, 437-439]. At the C-terminus, there is an ER-retention sequence (KTEL) that mediates retrograde transport from the Golgi to the ER, further maintaining the protein within the ER

[342]. More recently, the solving of the structure of AGR2 by nuclear magnetic resonance (NMR) revealed that the protein does indeed contain a thioredoxin fold [421], where the active site motif lies at the N-terminus of α-helix 2 (Fig. 3.1A and B), as in other family members [315]. Importantly, the structure revealed that AGR2 exists in monomer-dimer equilibrium [421], with AGR2 dimerisation subsequently also shown in other reports [416, 438]. The determination of the AGR2 dimer structure also revealed the nature of the dimerisation interface, with the complex being assembled with anti-parallel facing subunits that interact along α-helix-1 (Fig. 3.1B and C), and also suggested the involvement of charge interactions between a glutamate and lysine residue in dimerisation [421]. Homodimerisation plays a role in the regulation and function of the PDI ERp29, an unusual PDI that contains a thioredoxin fold but no active site cysteines [315]. Like AGR2, ERp29 is also thought to dimerise through charge-based interactions, with hydrogen bonding probably occurring between the aspartate side-chain on one subunit with a glycine residue on the opposing subunit, such that mutation of the aspartate residue alone was enough to prevent dimerisation [537]. As this dimerisation was of crucial importance to ERp29 chaperone function [537], a similar dimerisation-dependent mechanism may regulate AGR2 function, and thus it will be important to decipher the forces governing its dimerisation.

Through the afore-mentioned structural studies of AGR2, an unstructured, highly flexible region of the protein was also identified, initially within the residues immediately following the cleavable signal peptide (and thus termed the ’21-40’ region [421]), but later predicted to extend up to and including Thr57 (Fig. 3.1A). Whilst 3-dimensional structures can be attributed to these intrinsically disordered regions [326], it is noteworthy that the published structure for AGR2 does not contain the flexible 21-40 region. There are difficulties in assigning a structure to such disordered regions in the context of a protein in dynamic equilibrium, which relate to the necessity of labelling core hydrophobic residues for NMR. Thus signals from the charged N-terminal flexible region of AGR2 were not obtained [421, 538].

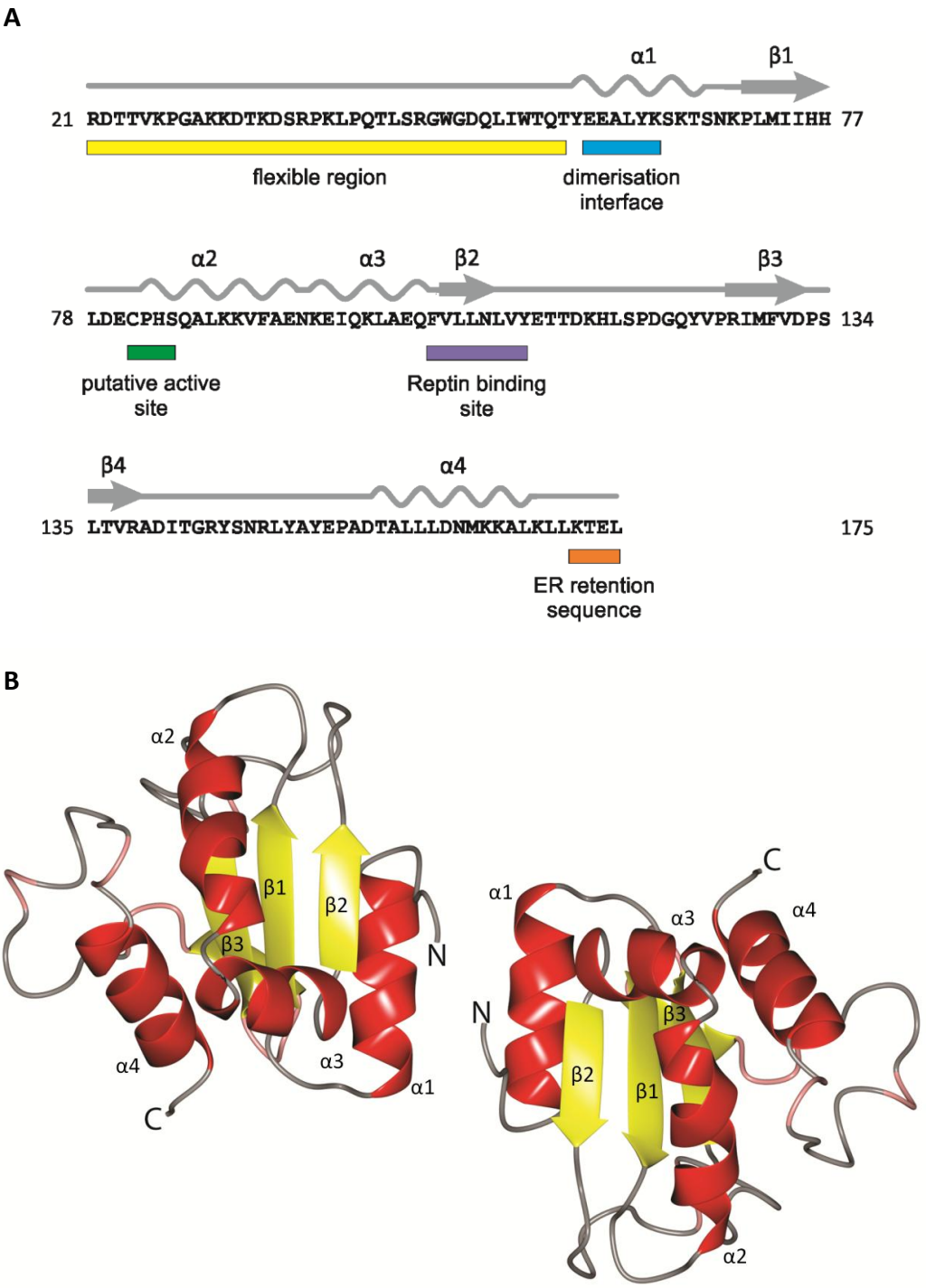


Figure 3.1. ***AGR2 can form a homodimer***. **(A)** Positions of notable regions within the AGR2 primary structure, relative to secondary structure elements (adapted from Patel *et al.,* 2013 [421]). Note that the signal peptide comprising residues 1-20 is not shown and the binding site for Reptin refers to the minimum sequence requirement for the reported binding of AGR2 to the histone acetyltransferase component Reptin [539]. **(B)** Solution structure of AGR2 derived by NMR (reproduced from Patel *et al.,* 2013 [421]). Note that the small β-sheet 4 is obscured by α-helix 2.

To date, biochemical assays have failed to detect PDI activity for recombinant AGR2 (i.e. in assays of disulphide bond reduction or isomerisation, or assays of thiol oxidation), although recombinant AGR2 does promote cell attachment when coated on a plastic substratum [309]. Similarly, AGR2-expressing cells showed an enhanced rate of cell adhesion compared to non-expressing cells [309, 515], and further studies have shown that AGR2 can promote migration and invasion *in vitro* [301, 414, 415, 440, 481], indicating that AGR2 expression has functional consequences related to adhesion and migration.

3.1.1 Chapter objectives

The aims of this chapter are to explore the factors affecting the extent of AGR2 dimerisation and thus the mechanism whereby AGR2 is able to form a dimer. In light of the difficulties of assaying recombinant AGR2 activity, the aim of this chapter is also to generate and characterise a panel of cell lines stably expressing structural mutant AGR2 proteins, with a view to determining their contributions to AGR2 biological activity.

3.2 Results

3.2.1 Characterisation of recombinant AGR2

3.2.1.1 Identification of residues involved in AGR2 dimerisation

AGR2 can form a homodimer in solution, as revealed by both NMR and size-exclusion chromatography-multiple angle laser light scattering (SEC-MALLS, [421]). Furthermore, paramagnetic relaxation experiments and residual dipolar coupling (RDC)-assisted modelling of symmetry [538] suggested that AGR2 subunits most likely dimerise along α-helix 1 of both subunits [421] (see Appendix 1). An electrostatic interaction between Glu60 and Lys64 in this α-helix 1 on opposing subunits was also predicted to be responsible for maintaining protein dimerisation. The presence of a tyrosine (Tyr63) on each chain within the proposed dimerisation interface also raised the possibility that hydrophobic interactions between these two tyrosine rings may further contribute to the extent of AGR2 dimerisation (Fig. 3.2).

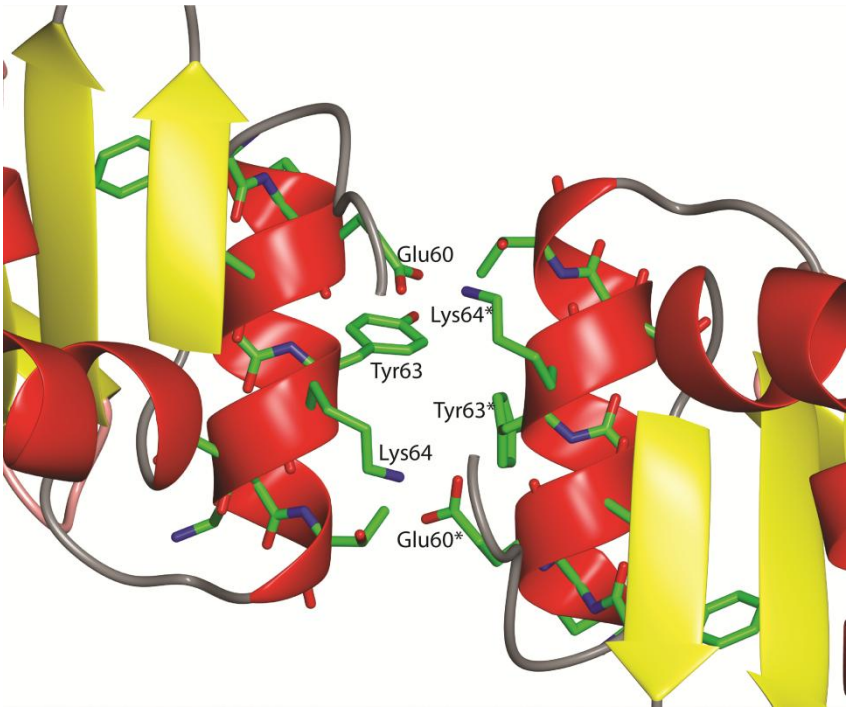


Figure 3.2. ***Proposed interface of the AGR2 dimer***. The residues thought to be involved in protein dimerisation along α-helix 1 are labelled, with asterisks (*) marking residues from a second subunit. Reproduced from Patel *et al*., 2013 [421].

The relative contributions of the above-mentioned residues to the dimerisation of AGR2 were tested on recombinant protein using mutagenesis and SEC-MALLS (Fig. 3.3). Furthermore, as one study has reported that AGR2 dimerises through Cys81 in the putative active site [416], the oligomerisation status of $AGR2_{41-175}$ with a mutation of this cysteine to a serine (C81S $AGR2_{41-175}$) was also analysed (Fig. 3.3E). As noted previously, residues 21-40 of AGR2 form part of an intrinsically disordered region of the protein. Regions with such a flexibility are a challenge for NMR structural determination [540] and, given the already challenging nature of determining the structure of a homodimer [538], a truncated form of AGR2 was used for structural and subunit studies. This truncated protein, $AGR2_{41-175}$, lacks both the signal peptide and 21-40 flexible region.

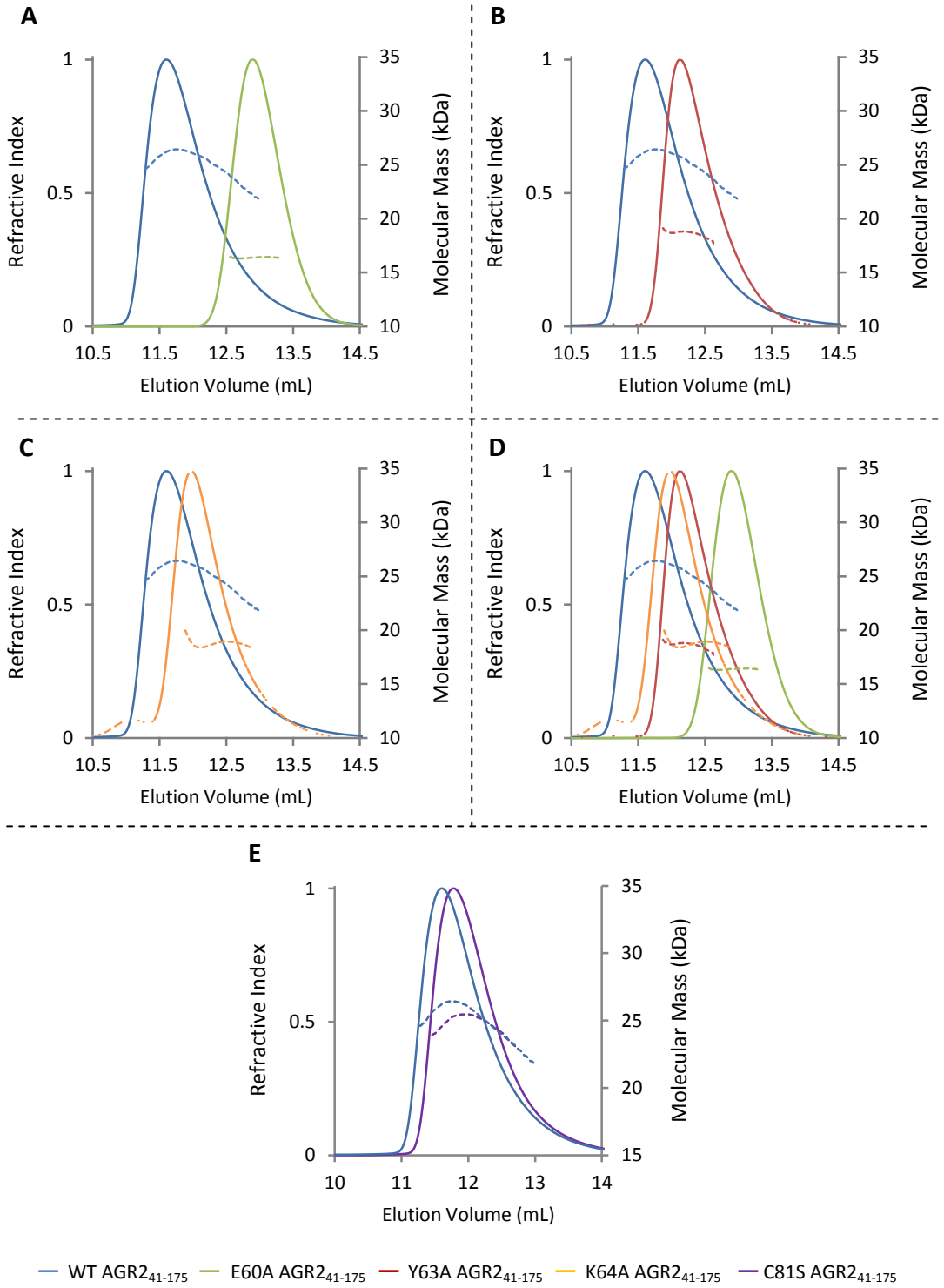


Figure 3.3. ***Determination of residues involved in AGR2 dimerisation***. Recombinant AGR2 proteins were purified and subjected to SEC-MALLS analysis. The oligomerisation of wild-type (WT) AGR2$_{41-175}$ was compared to **(A)** E60A AGR2$_{41-175}$, **(B)** Y63A AGR2$_{41-175}$, **(C)** K64A AGR2$_{41-175}$, **(D)** all three interface mutations and **(E)** C81S AGR2$_{41-175}$. The refractive indices of the measured proteins are represented by solid lines, whilst the corresponding molecular masses of the protein species across these peaks are shown by dashed lines. Experiments were performed at 150 mM NaCl.

The elution profile of WT $AGR2_{41-175}$ yielded an asymmetrical peak, consistent with a sample of heterogeneous molecular mass. Indeed, molecular mass across the major isolated peak (dashed blue line, Fig. 3.3A) was not uniform, with a mean mass of 26.0 kDa. Together, these data imply that recombinant AGR2 exists in a monomer-dimer equilibrium, rather than as a pure dimer. The mean molecular mass of 26.0 kDa does suggest that the dimer is favoured, given that it is closer to the expected mass of 32.2 kDa for a dimer than 16.1 kDa for a monomer of $AGR2_{41-175}$ [738].

Strikingly, mutation of the Glu60 residue to Ala (E60A) resulted in complete ablation of dimerisation potential (Fig. 3.3A). The protein here gave a symmetrical elution profile (indicative of uniform molecular mass) with a corresponding mean molecular mass of 16.4 kDa across the peak, the expected mass for a monomeric $AGR2_{41-175}$ protein. On the other hand, mutation of the Tyr63 residue to Ala (Y63A) produced an asymmetrical peak similar to that of WT $AGR2_{41-175}$ (Fig. 3.3B), but with an average molecular mass of 18.9 kDa. This suggests that this mutant protein is still in monomer-dimer equilibrium, but that the monomer is more favoured than for WT protein, given that the average molecular mass of the sample was very close to that of monomeric $AGR2_{41-175}$. These changes in equilibrium and average molecular mass for Y63A $AGR2_{41-175}$ would indicate that Tyr63 also contributes to AGR2 dimerisation.

Surprisingly, mutation of the Lys64 residue (the predicted residue involved in salt bridge formation with Glu60) to Ala did not completely disrupt dimerisation in the same way as the E60A mutation (Fig. 3.3A and C). The elution profile of K64A $AGR2_{41-175}$ was more reminiscent of Y63A $AGR2_{41-175}$, with a very similar average molecular mass of 18.7 kDa. This points to the protein being largely monomeric, with a limited ability to dimerise. AGR2 possessing the C81S mutation behaved in almost exactly the same manner as WT $AGR2_{41-175}$ (Fig. 3.3E), with a very similar elution profile and mean molecular mass (26.2 kDa and 25.4 kDa, respectively), indicating that Cys81 does not play a role in $AGR2_{41-175}$ dimerisation. Taken together, these data support the proposed model of AGR2 dimerisation, by demonstrating the importance of the three residues E60, Y63 and K64 in α-helix 1 to dimer formation, and that dimerisation is not mediated through Cys81, as reported by others [416].

3.2.1.2 Full length AGR2$_{21-175}$ dimerises in the same way as truncated AGR2$_{41-175}$

Within the cell, AGR2 is synthesised with an N-terminal signal peptide that is cleaved upon translocation into the ER [309, 541]. This mature form of the protein, devoid of signal peptide, is denoted as AGR2$_{21-175}$. In order to determine whether the natural form of AGR2 (AGR2$_{21-175}$) exhibits the same dimerisation properties as those observed with AGR2$_{41-175}$, recombinant WT AGR2$_{21-175}$ and E60A AGR2$_{21-175}$ were purified and subjected to SEC-MALLS analysis.

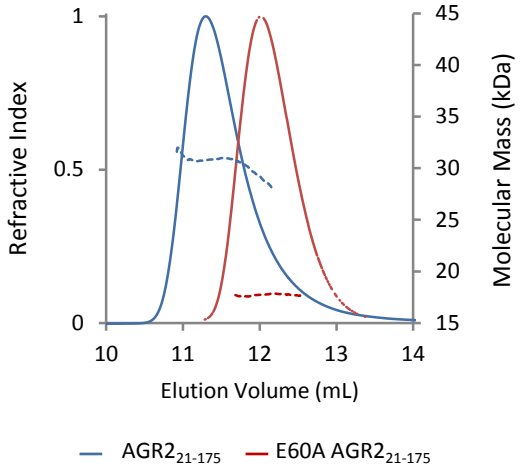


Figure 3.4. ***AGR2$_{21-175}$ exists in monomer dimer equilibrium, mediated by Glu60.*** Wild-type (WT) AGR2 (AGR2$_{21-175}$) and **(A)** AGR2 bearing a Glu to Ala mutation at position 60 (E60A AGR2$_{21-175}$) were purified and subjected to subunit analysis using SEC-MALLS. The refractive indices of the measured proteins are represented by solid lines, whilst the corresponding molecular masses of the protein species across these peaks are shown by dashed lines. Experiments were performed at 150 mM NaCl.

WT AGR2$_{21-175}$ displayed a very similar elution profile to that of WT AGR2$_{41-175}$, also showing an asymmetrical peak with a non-uniform molecular mass, again suggestive of monomer-dimer equilibrium (Fig. 3.4). The mean molecular mass across the peak was 30.9 kDa, again closer to the 35.6 kDa theoretical mass of a dimer than 17.8 kDa, the theoretical mass of a monomer. It is interesting that the mean molecular masses for AGR2$_{21-175}$ and AGR2$_{41-175}$ were 80.7 % and 86.8 %, respectively, of the theoretical masses for dimers of their respective species. This might suggest that both proteins dimerise to approximately the same extent, with AGR2$_{21-175}$ possibly producing a slightly more favoured dimer.

Mutating the Glu60 residue of AGR2$_{21-175}$ (E60A AGR2$_{21-175}$, Fig. 3.4) had the same effect as with the truncated AGR2$_{41-175}$ protein, in that there was a uniform molecular mass across the peak (17.2 kDa) which matched the expected mass for an AGR2$_{21-175}$ monomer

(17.8 kDa). This is consistent with the observation made with $AGR2_{41-175}$ protein [421], and suggests that $AGR2_{21-175}$ and $AGR2_{41-175}$ behave in essentially the same manner, with dimerisation occurring through the α-helix 1 interface (amino acids Glu60 to Lys64).

3.2.1.3 Protein concentration contributes to the extent of AGR2 dimerisation

Protein concentration can be a factor in determining both the extent of dimerisation in proteins exhibiting monomer-dimer equilibria and also in their biological activity [542, 543]. This might be particularly relevant in the context of tumourigenesis, where abnormal overexpression of proteins such as AGR2 would increase the local concentration of protein and potentially affect its oligomerisation state. To this end, $AGR2_{21-175}$ was purified and analysed by SEC-MALLS at three different protein concentrations (Fig. 3.5).

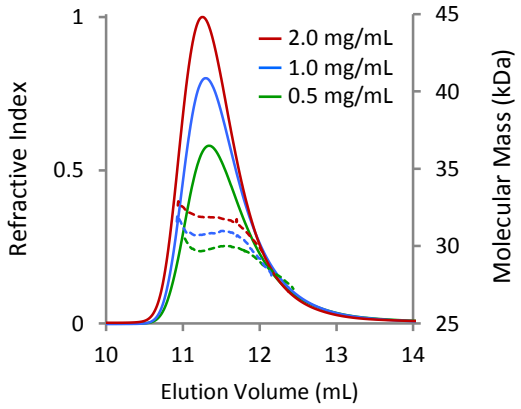


Figure 3.5. ***Protein concentration affects monomer-dimer equilibrium of AGR2.*** Recombinant wild-type $AGR2_{21-175}$ was purified and analysed at three different concentrations by SEC-MALLS. The refractive indices of the measured proteins are represented by solid lines, whilst the corresponding molecular masses of the protein species across these peaks are shown by dashed lines. Experiments were performed at 150 mM NaCl.

There was a small effect of protein concentration on AGR2 oligomerisation, with protein at 0.5, 1.0 and 2.0 mg/mL all displaying a very similar elution profile. There was a small upward shift in mean molecular mass across the peaks with increasing protein concentration, rising from 29.9 kDa at 0.5 mg/mL, to 31.0 kDa at 1.0 mg/mL and then further to 31.8 kDa at 2.0 mg/mL. These data provide some evidence that an increase in local concentration of AGR2 could favour dimer formation to some extent in the cell, but

without information on physiological concentrations of AGR2, it is difficult to say whether this has any functional significance *in vivo*.

3.2.1.4 The flexible 21-40 region stabilises the AGR2 dimer

To further explore the type of forces driving dimerisation of AGR2, the potential of higher salt concentration to disrupt the dimer was tested, given that the complex is held together through predominantly ionic interactions.

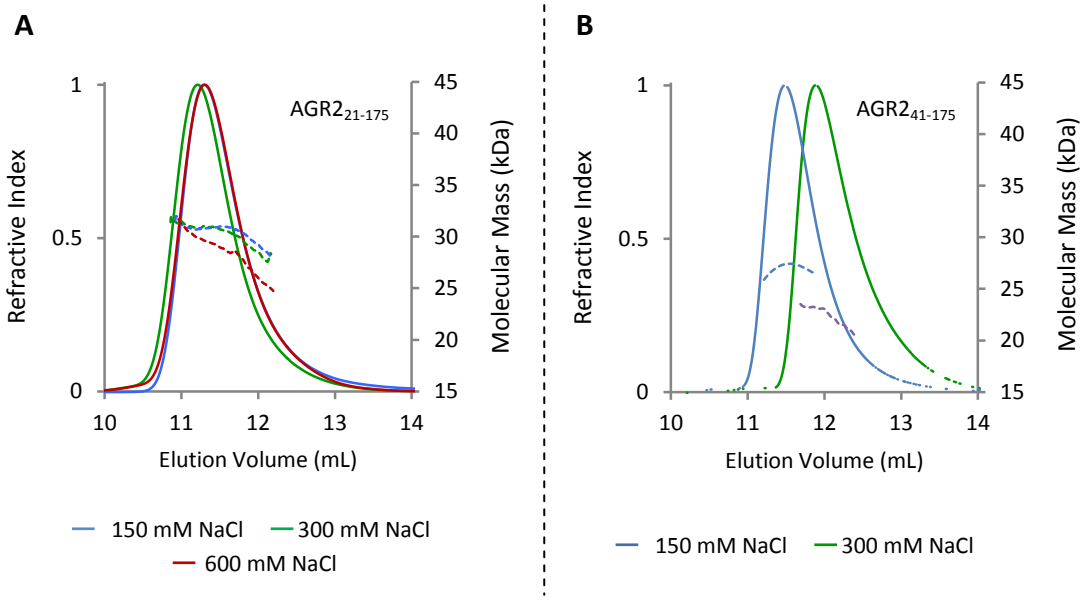


Figure 3.6. ***The AGR2$_{41-175}$ dimer is more susceptible to dissociation from high salt concentration than AGR2$_{21-175}$ dimer.*** **(A)** Wild-type (WT) full length AGR2 (AGR2$_{21-175}$) and **(B)** WT truncated AGR2 (AGR2$_{41-175}$) were purified and the concentration of NaCl adjusted as shown. The proteins were then subjected to SEC-MALLS analysis. The refractive indices of the measured proteins are represented by solid lines, whilst the corresponding molecular masses of the protein species across these peaks are shown by dashed lines.

Interestingly, there was no difference in mean molecular mass of AGR2$_{21-175}$ in 150 mM or 300 mM NaCl (Fig. 3.6A), with values of 30.9 kDa for 150 mM NaCl and 31.0 kDa for 300 mM NaCl. However, the presence of 600 mM NaCl did reduce the mean molecular mass to 29.5 kDa, indicating a small change in monomer-dimer equilibrium at this high salt concentration.

In contrast, $AGR2_{41-175}$ in only 300 mM NaCl (Fig. 3.6B), exhibited a much greater shift in mean molecular mass than was observed even at 600 mM NaCl for $AGR2_{21-175}$, reducing the mean molecular mass from 27.2 kDa to 22.6 kDa. Such a large shift in molecular mass supports the notion of the complex forming through ionic interactions, but the dramatic difference in resistance to salt concentration observed between $AGR2_{21-175}$ and $AGR2_{41-175}$ suggests that the presence of the flexible region stabilises the AGR2 dimer, possibly through protecting the dimer interface from solvent attack. However, while these experiments were performed in buffer containing non-physiological concentrations of NaCl, intracellular KCl is present at a concentration of around 140 mM [544], and so the relevance of these results to *in vivo* AGR2 protein may depend on whether K^+ ions have a different effect than Na^+ ions on AGR2 dimerisation. With this in mind however, even at 150 mM NaCl, there was little difference in the monomer-dimer equilibria of $AGR2_{21-175}$ and $AGR2_{41-175}$ (see Section 3.2.1.2), and so *in vivo*, the 21-40 region probably does not play a significant role in AGR2 dimerisation. Whether these experiments are evidence of a protective effect of the 21-40 region on the dimer interface could be an interesting avenue of investigation.

3.2.1.5 AGR2 does not function in biochemical assays of PDI function

Having established the consequences of mutations in the dimer interface on the dimerisation of AGR2, the functional consequences of these mutations were next investigated through the use of biochemical assays. PDI proteins catalyse the reduction, formation and isomerisation of disulphide bonds (DSBs) in substrate proteins within the ER [311]. Measuring this activity in *in vitro* biochemical assays facilitates the assessment of changes to enzyme activity as a result of any protein modifications. Therefore, the PDI activity of AGR2 was explored, with a view to pinpointing the functional consequences of the loss of protein dimerisation and loss of the flexible 21-40 region.

The insulin turbidity assay is a measure of dithiol reductase activity, i.e. the ability to catalyse the reduction of DSBs. When the interchain DSB between insulin α- and β-chains is reduced, the β-chain precipitates and this can be measured at 650 nm (reviewed in [311]). AGR2 was unable to catalyse this DSB reduction at a measurable rate, as it showed no difference in reaction rate compared to the control reaction without enzyme (Fig. 3.7A). Thioredoxin, a cellular reductase [545], exhibited similar activity to that reported before [524] in control reactions carried out at the same time (Fig 3.7A).

AGR2 also showed no isomerase activity as measured by scrambled RNAse assays (Pryank Patel, personal communication). Here, an active PDI should promote the correct re-folding of RNAse with mispaired disulphides, producing an active form of RNAse [546]. This lack of isomerase activity is perhaps not unexpected given that isomerase activity generally requires two CxxC-containing catalytic domains [321, 322].

Some proteins containing the single-cysteine CxxS active site motif are able to function as glutaredoxins [547, 548], contributing to cellular redox reactions through catalysing the reduction of DSBs between glutathione and protein thiol groups (reviewed in [548]). However, AGR2 did not display any glutaredoxin activity, showing no difference in reaction rate compared to the negative control reaction (Fig. 3.7B).

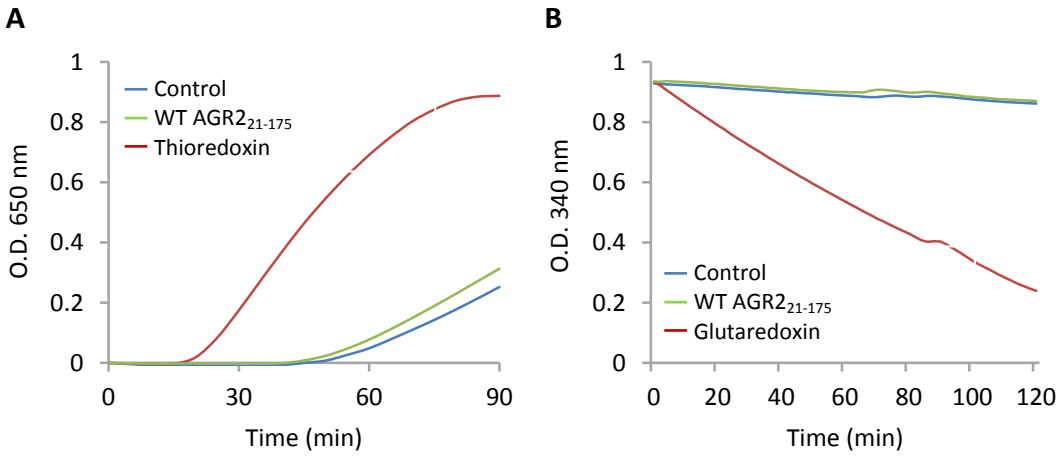


Figure 3.7. ***AGR2 is inactive in biochemical assays of PDI activity.*** **(A)** Insulin turbidity assay. The reduction of the interchain disulphide bond between insulin α- and β-subunits by DTT was measured in the absence (blue) and presence of thioredoxin (red) or WT AGR2 $_{21-175}$ (green). Precipitation of the β-chain as a result of DSB reduction was measured at 650 nm. **(B)** Glutaredoxin assay. The reduction of DSBs within the HED substrate was monitored by the consumption of NADPH, measured at 340 nm. Both panels show representative outputs from experiments performed using duplicate wells.

This lack of PDI-like enzyme activity may indicate that AGR2 does not in fact function as a PDI. This creates a problem for studying the functional consequences of changes to AGR2 oligomerisation and other potentially important domains, given the lack of suitable assays to assess their activity. AGR2 does, however, promote cell adhesion [309, 421, 515], and so in the absence of a biochemical assay for AGR2 activity, cell-based biological assays are the major tool for determining the effects of sequence- and structure-

based alterations to AGR2 on its activity. In this context, the generation and characterisation of AGR2-expressing cell lines are an important means of relating AGR2 structure and activity. Therefore, cell lines stably-expressing WT AGR2 and AGR2 mutant proteins were created, with a view to testing their adhesion properties as a measure of AGR2 activity.

3.2.2 Characterising AGR2-expressing stable cell lines

3.2.2.1 Creating cell lines to study AGR2 function **in vitro**

AGR2 contains several protein domains that might be relevant to its function, including its putative cysteine-containing thioredoxin domain (C81) and its C-terminal ER-retention sequence (KTEL), that controls not only its localisation but also plays a role in anchorage-independent cell growth [445, 481, 482]. In addition to these domains, the structural work described in Section 3.2.1 demonstrates that protein dimerisation (crucially mediated through Glu60) and also the intrinsically disordered, flexible 21-40 region may also play important roles in AGR2 function. Therefore, rat mammary cells (Rama 37) stably expressing wild-type (WT) and mutant forms of AGR2 were created to investigate AGR2 structure/function relationships (Table 3.1). In addition to the mutants described above, a further mutant protein devoid of signal peptide was also created, in order to block entry of AGR2 into the ER and investigate whether any or all of its functions are dependent on residency within the ER.

Mutant	*Domain mutated*	*Expected effect on protein*
E60A AGR2	Dimerisation interface	Monomeric
C81S AGR2	Putative thioredoxin domain	Potentially catalytically inactive
Δ1-20 AGR2	Signal sequence	Prevention of entry into the ER
Δ21-40 AGR2	Flexible region	Deletion of flexible region
ΔKTEL AGR2	ER-retention sequence	Enhanced secretion

Table 3.1. ***AGR2 mutant proteins stably expressed by Rama 37 cell lines.***

3.2.2.2 Transposase-based plasmid integration

Rama 37 cell lines stably expressing both WT and mutant AGR2 protein under the control of a cytomegalovirus (CMV) promoter had previously been constructed [309], but these cells appeared to lose expression of these proteins upon subculturing. The reasons for this are not clear, but may be related to the propensity of CMV promoters to become silenced during long term culture [549-552]. Therefore, new stably-expressing cell lines were created, using a more robust and efficient transposase-based system, under the control of the robust EF1α promoter, which appears to be much less frequently silenced [551].

The PiggyBac transposon vector system has several key features that promote efficient mammalian cell transfection. In addition to the EF1α promoter mentioned above, the main attraction of the system is that, through the co-transfection of a transposase-expressing plasmid with the vector expressing the gene of interest (GOI), the entire GOI expression cassette is inserted into the host genome at specific AT-rich sites, through a "cut and paste" mechanism [553]. These AT-rich regions are not well characterised in higher eukaryotes, but in prokaryotes and yeast, they tend to be found at origins of DNA replication, where the weaker A-T interactions relative to G-C interactions facilitate DNA strand separation (reviewed in [554]). For the most part, these sites are therefore presumably located outside of coding regions, and thus insertion of a GOI into these regions of DNA should limit the likelihood of inducing unwanted mutagenesis into the host genome through disruption of existing gene architecture.

The transposase-catalysed insertion of the GOI into the host genome permits not only a higher transfection efficiency (Fig. 3.8) but, due to core insulator sites flanking the expression cassette (Appendix 2), the integrity of the cassette is conserved and should not be broken during genome integration. A final important feature of the system is the use of an internal ribosome entry site (IRES) that allows the expression of an antibiotic resistance gene (here, neomycin) on the same mRNA as the GOI, at its 3' end. Thus, only cells expressing the GOI will be resistant to the selection agent, ensuring that the population of cells demonstrating antibiotic resistance do also express the GOI. Using this system, stable cell lines were produced rapidly and reliably (Fig. 3.8A).

The benefits of working with the PiggyBac transposase system are outlined in Fig. 3.8. There was a greater transfection efficiency when co-transfecting the GOI vector with the transposase plasmid, rather than using the GOI vector alone, represented by the relative increase in AGR2 expression 24 h after transfection (Fig. 3.8A). Expression was

monitored over the active period of the transposase (72 h), but appears to be lost during this period, most likely due to the effects of the growth of non-transfected cells 'diluting' the apparent amount of AGR2 protein.

It is interesting that after selection in G418, the stable cell line created using transposase co-transfection seemed to produce more AGR2 protein than the cell line created without transposase (Fig. 3.8A), probably reflecting the more efficient process of using the transposase, but also the fact that without the enzyme, chromosomal integration will be more random, and so chromosomal insertion points may affect cellular components involved in gene expression and protein synthesis, such that protein production is less efficient. It is also important to note that transposase co-transfection allowed the much more rapid production of a stably-expressing cell line, with the vector-only cell line requiring an extra 7-10 days growth to reach the same confluence as transposase co-transfected cells. This again most likely stems from the increased transfection efficiency of the transposase system, producing a higher percentage of transfected cells so that more survive under selection and confluency is reached much more quickly.

Another interesting feature of the transposase system is highlighted in Fig. 3.8B. The amount of vector DNA transfected into Rama 37 cells was varied from the 2 µg recommended by the manufacturer to 12 µg. Regardless of the amount of DNA transfected, all stable cell lines showed the same steady-state levels of protein, possibly reflecting the saturation of AT-rich insertion sites available within the host genome. This meant that low amount of DNA could be used, requiring less transfection reagent, leading to lower amounts of cell death, with the ultimate outcome of further speeding up the generation of stable cell lines.

A potential pitfall in the generation of stable cell lines is the loss of protein expression over time, as mentioned above. Therefore, the expression of AGR2 from a PiggyBac-created cell line was monitored over a ten week period (Fig. 3.8C). These cells grow rapidly and were subcultured every two days, and whole cell lysates were collected after the indicated number of passages. Expression of AGR2 protein was maintained at a similar level throughout the time-course, with possibly a slight reduction in expression after 35 passages. This continued expression of AGR2 over a relatively long culture period alleviated any concerns that expression of the protein might be repressed in the long term, and highlights the robustness of the EF1α promoter. Together, these data demonstrate the PiggyBac system to be an effective tool for both transient and stable transfection of cells.

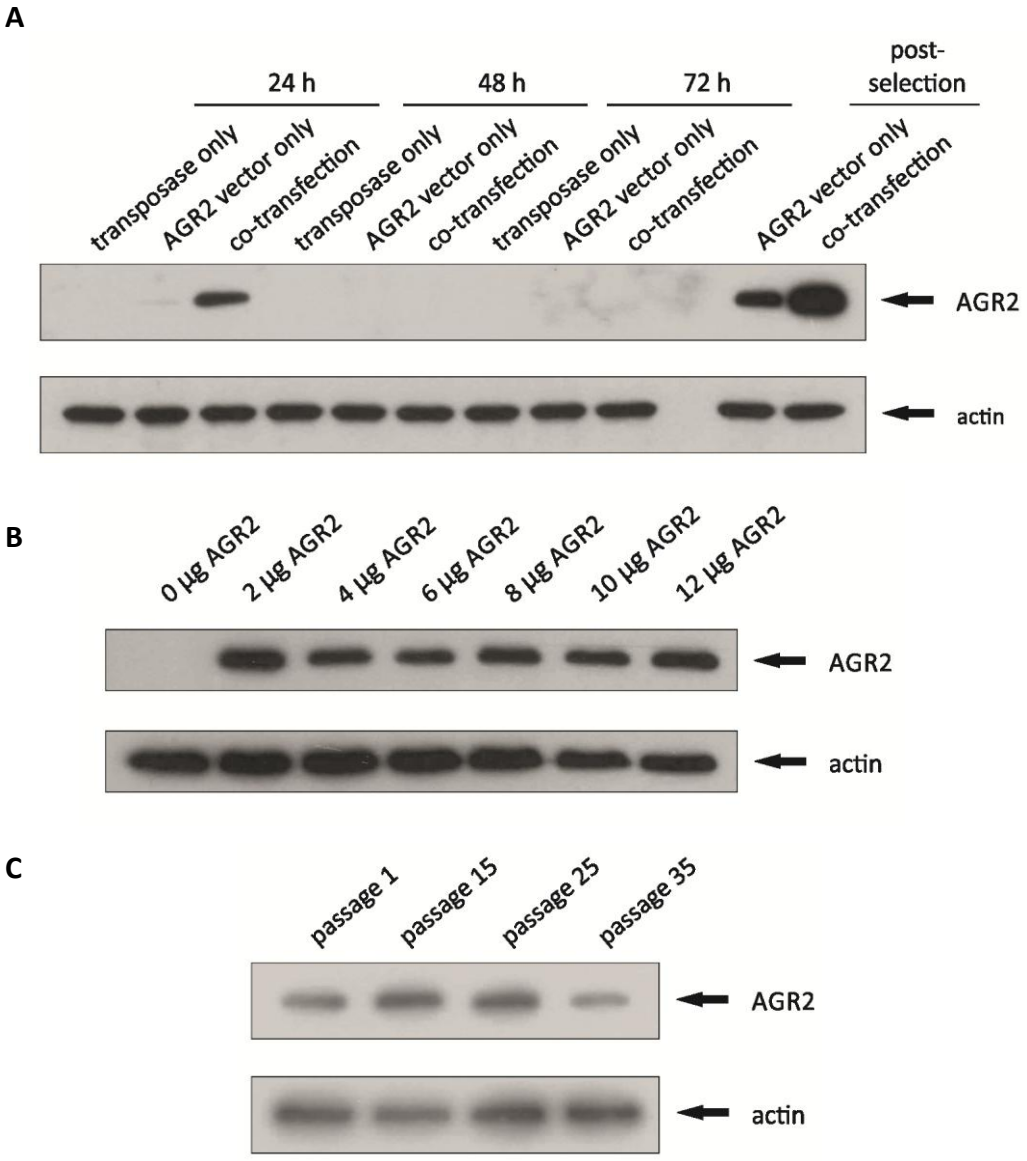


Figure 3.8. ***The PiggyBac transposon system is a robust system for the generation of stable cell lines.*** **(A)** Rama 37 cells were transiently transfected with either transposase plasmid only, AGR2-containing PiggyBac vector only or co-transfected with both. Cells were harvested and probed for AGR2 expression at the indicated timepoints. Cells were then subjected to G418 selection to establish stable cell lines, and probed for AGR2 expression. Transposase-only transfected cells do not survive selection, as expected. **(B)** Rama 37 cells were co-transfected with the indicated amount of AGR2-containing PiggyBac plasmid and transposase plasmid. After G418 selection, cells were harvested and probed for AGR2 expression. **(C)** Rama 37 cells stably expressing AGR2 were cultured for 10 weeks (35 passages) and harvested and probed for AGR2 at the indicated passages.

3.2.2.3 Generation of mutant AGR2-expressing cell lines

Having established the PiggyBac system as a suitable tool for the generation of stable cell lines, other cell lines expressing the mutant AGR2 proteins outlined in Table 3.1 were created (Fig. 3.9). Surprisingly, given the consistent efficiency of transfection (Fig. 3.8B), mutant proteins displayed a wide range of expression levels, with only the Δ21-40 AGR2 mutation being expressed at a comparable level to WT protein (Fig. 3.9A). These differences in steady-state protein levels therefore likely reflect differences in protein stability or turnover within the cell, rather than differences in transfection efficiency.

In order to be effective tools for examining the effects of mutations on protein activity, these cells lines must express similar levels of proteins to discount any effects mediated by protein concentration. Therefore, multiple clonal cell lines were created from each pooled mutant cell line, to isolate cell lines expressing similar protein levels across all six cell lines. The three clonal cell lines expressing the highest amount of protein for each mutation were identified (Fig. 3.9B and C).

As seen with the pooled cell lines, overall steady state levels of mutant proteins were largely lower in clonal mutant cell lines than for clonal WT cell lines, possibly as a result of decreased protein or mRNA stability. Nevertheless, clonal cell lines expressing similar amounts of protein were found, providing tools for subsequent experiments. The two major exceptions to this were Δ1-20 AGR2 clones and ΔKTEL AGR2 clones, where steady state levels were substantially lower (Fig. 3.9C). Given that ΔKTEL AGR2 lacks an ER-retention sequence, the protein would be expected to be secreted to a greater degree than WT protein and this could explain the apparent low intracellular steady state levels in these cells. Indeed, when the levels of secreted AGR2 were measured by Western blot for WT clone 2 and ΔKTEL clone 8, there was no detectable secreted WT AGR2 but ΔKTEL AGR2 was highly secreted (Fig. 3.10).

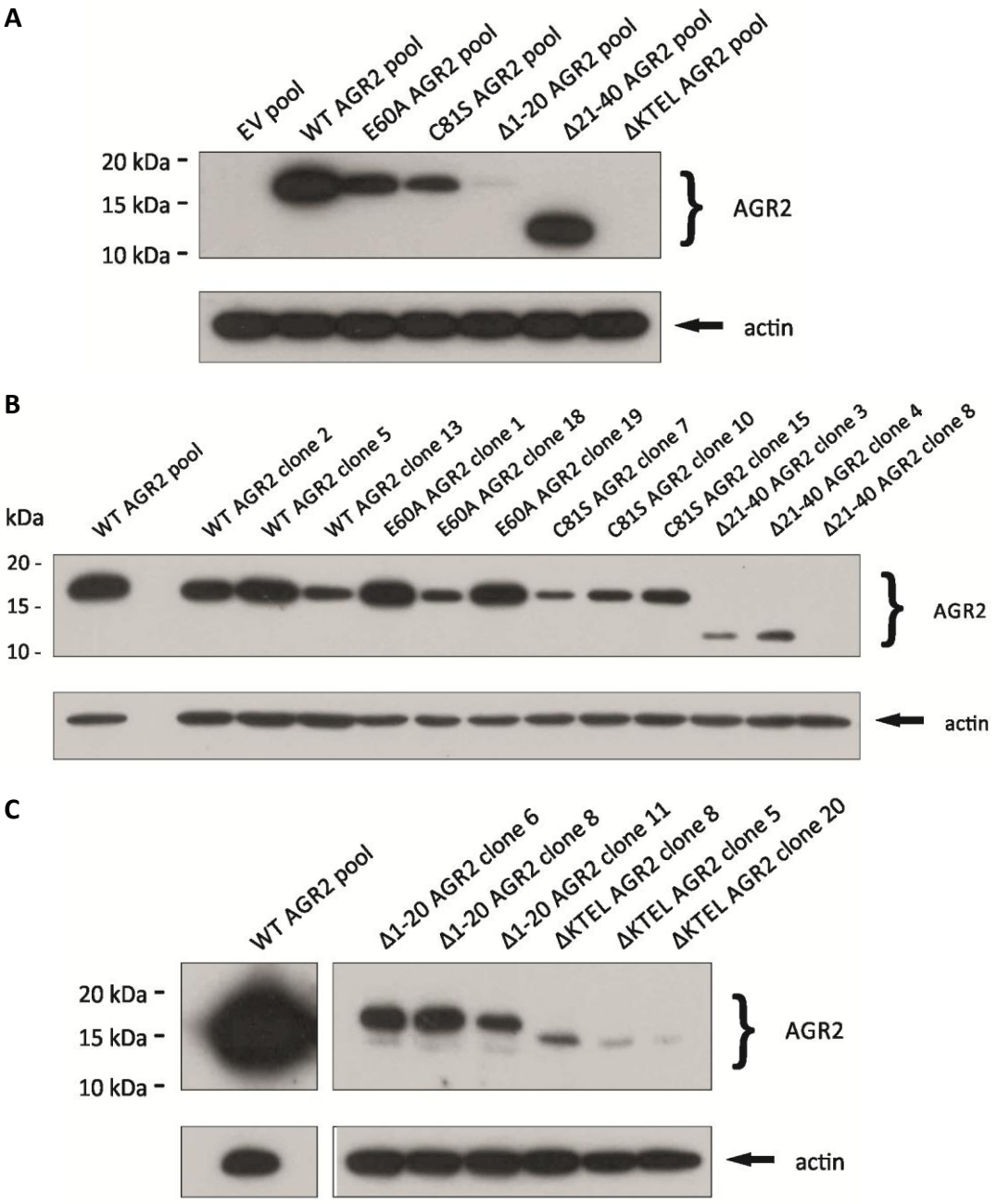


Figure 3.9. ***Generation of AGR2-expressing stable cell lines*****. (A)** Equal numbers of Rama 37 cells were transfected with equal amount of PiggyBac vector containing the indicated AGR2 mutation and subjected to selection for 10-14 days. Cells were harvested and probed for AGR2 by Western blot. **(B and C)** From the cells in (A), clonal cell lines were created using a limiting dilution technique, where the pooled cell lines are subjected to serial dilutions, until one cell per well is obtained. After expansion, the clonal cells were harvested and probed for AGR2 by Western blot. The three highest expressing clonal cell lines for each mutant and WT AGR2 are shown, with expression in WT pool cells for comparison. **(C)** AGR2 expression levels in Δ1-20 AGR2 and ΔKTEL AGR2 clones. WT pool expression is shown from a different part of the same blot.

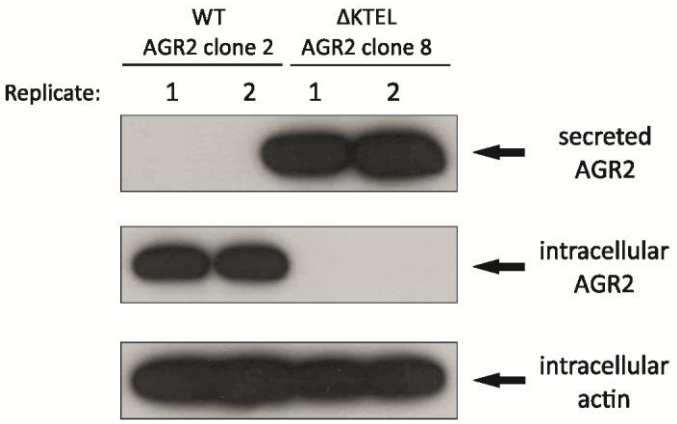


Figure 3.10. ***ΔKTEL AGR2 is highly secreted compared to WT AGR2.*** Whole cell lysates from WT and ΔKTEL AGR2 cells grown in normal medium were collected and probed for AGR2. For measuring secreted AGR2, conditioned medium (CM) was collected from cells of approximately equal confluency, in Opti-MEM-based serum-free medium. As CM loading could not be normalised to protein concentration due to the high concentration of transferrin in the Opti-MEM medium (15 μg/mL) which masks the concentration of cell-derived proteins, only equal volumes of CM were loaded. However, samples were collected in duplicate to limit the effects of any differences in protein concentration on any observed differences. Equal amounts of cell lysate, measured by BCA assay, were loaded and intracellular actin levels were used as a loading control for these samples

The low expression of Δ1-20 AGR2 is not as easy to explain. The protein is essentially the same as WT protein (it is in fact fully mature WT AGR2) except that it does not gain access into the ER. Assuming that there is no difference in the extent of gene expression between these two proteins, it may be that AGR2 itself requires one or more ER chaperones to help it fold correctly and so lack of entry into the ER limits its correct folding, possibly resulting in lower levels of AGR2 protein as incorrectly folded protein will presumably be degraded. Alternatively, its apparently low expression may be related to its sub-cellular localisation, as outlined below.

3.2.2.4 Sub-cellular localisation of WT and mutant AGR2

In order to further characterise the stable cell lines expressing AGR2 and attempt to explain the differences in steady state levels between the various mutants, the sub-cellular localisations of the proteins were investigated using fluorescence microscopy. AGR2 is an ER-resident protein [417, 422, 437-439], but has also been reported to be secreted [388, 426, 440]. Therefore, to examine whether mutations to AGR2 might affect its sub-cellular localisation, but also to confirm that Δ1-20 AGR2 and ΔKTEL AGR2 do not enter the ER or are secreted, respectively (Table 3.1), cells were co-stained with AGR2 and one of three sub-cellular markers: PDI for the ER [555], giantin for the Golgi [556] and MEK1 for the cytoplasm [557]. Secreted proteins must generally travel through the Golgi towards the cell exterior (see Section 1.3.3.1), and thus co-localisation of AGR2 with a Golgi marker was used as an indicator of proteins potentially destined for secretion. However, it should be noted that non-secreted, ER proteins can be briefly resident in the Golgi before retrograde transport back into the ER [342].

In cells expressing WT AGR2 and the E60A, C81S and Δ21-40 AGR2 mutations, the protein localised largely to the ER, as demonstrated by co-localisation with PDI (Fig. 3.11A-C and E). There was no evidence of Golgi-located protein nor cytoplasmic protein, consistent with AGR2 being an ER-resident protein. Interestingly, some cells did appear to show some nuclear staining, but in the absence of a suitable nuclear protein marker this cannot be confirmed. However, it is interesting that in Δ1-20 AGR2 cells, AGR2 localised not only to the cytoplasm as expected (as evidenced by co-localisation with MEK1) but also showed strong nuclear staining (Fig. 3.11D). To ensure that this nuclear staining was not artefactual, empty vector (EV)-transfected cells were also stained for AGR2. These cells express no detectable amounts of AGR2 by Western blot (Fig. 3.9A) and whilst they showed no nuclear AGR2 staining, they did display a very weak staining for AGR2 in the nucleoli (Fig. 3.11H). However, given the lack of predicted nucleolar localisation sequence (NoLS) for AGR2 (Appendix 3), this nucleolar staining is most likely a staining artefact. Therefore, it appears that Δ1-20 AGR2 does indeed localise to the nucleus, and, along with the predicted weak NLS for AGR2 (Appendix 4), these data may support the observed weak nuclear localisation of WT AGR2, indicating that it can enter the nucleus as previously reported [467].

In ΔKTEL cells, the low abundance of intracellular protein meant that it was difficult to determine whether there was co-localisation between AGR2 and PDI (Fig. 3.11F). Despite this, ΔKTEL cells were the only cell type where co-localisation of AGR2 with the Golgi

marker, giantin was observed, consistent with the loss of the ER-retention sequence leading to onward transport to the Golgi and increased protein secretion.

Overall, these imaging data show these cell lines to be suitable tools for exploring the functional consequences of AGR2 mutation, as only the mutation designed to alter protein localisation (Δ1-20 AGR2 and ΔKTEL AGR2) actually did so, with other mutations localising to the same sub-cellular compartment as WT protein.

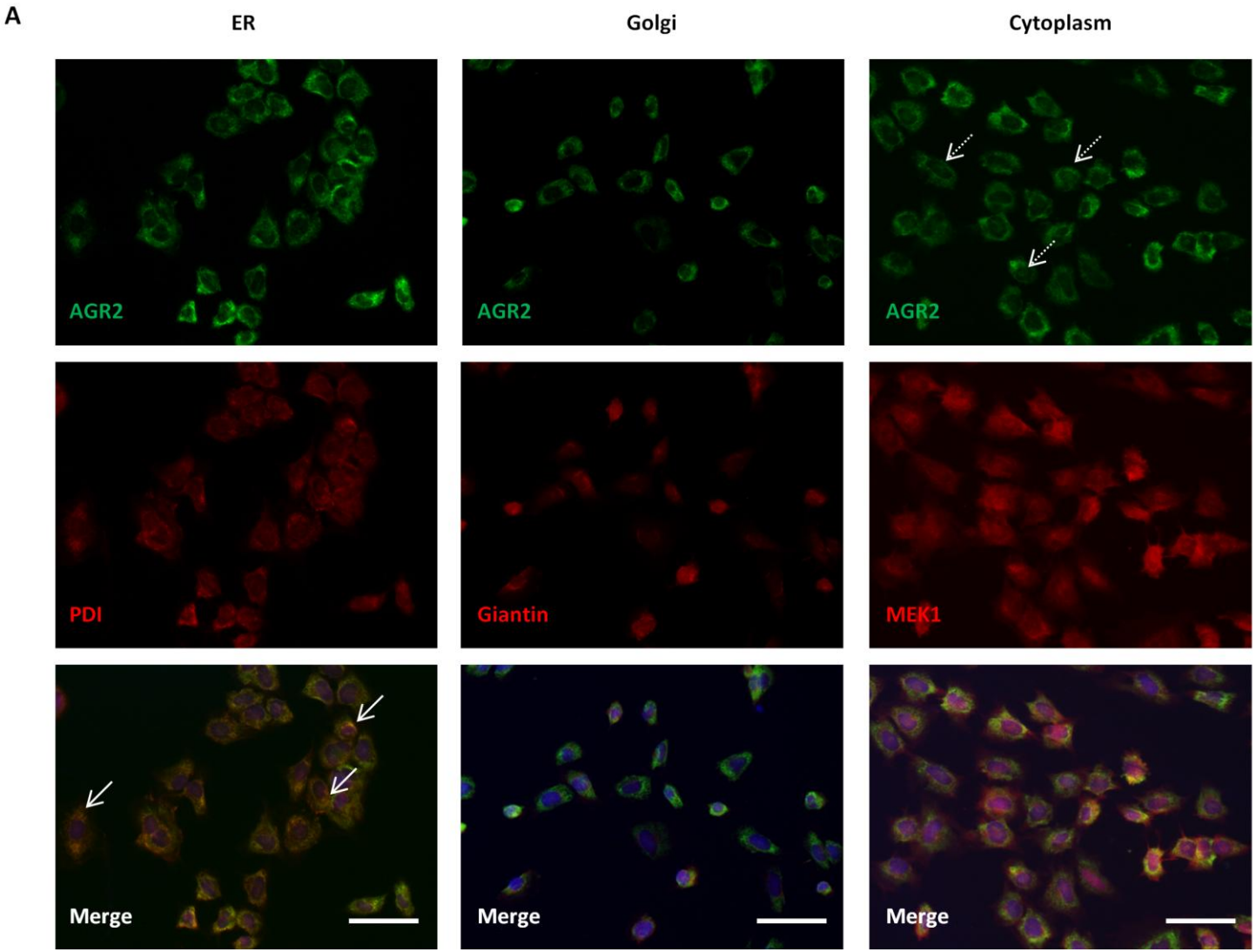


Figure 3.11. Cellular localisation of WT and mutant AGR2 protein. Cells were fixed with 4% PFA and permeabilised with 1% triton in PBS. Cells were co-stained with AGR2 (green) and the organelle markers PDI (for ER staining), Giantin (for Golgi staining) or MEK1 (for cytoplasmic staining), all in red. Representative images are shown for clonal cell lines expressing **(A)** WT AGR2, **(B)** E60A AGR2, **(C)** C81S AGR2, **(D)** Δ1-20 AGR2, **(E)** Δ21-40 AGR2, **(F)** ΔKTEL AGR2 and **(G)** empty vector. Nuclei (blue) were stained with DAPI. Solid white arrows indicate areas of co-localisation. Dashed arrows indicate possible nuclear AGR2 staining. Scale bar = 50 µm.

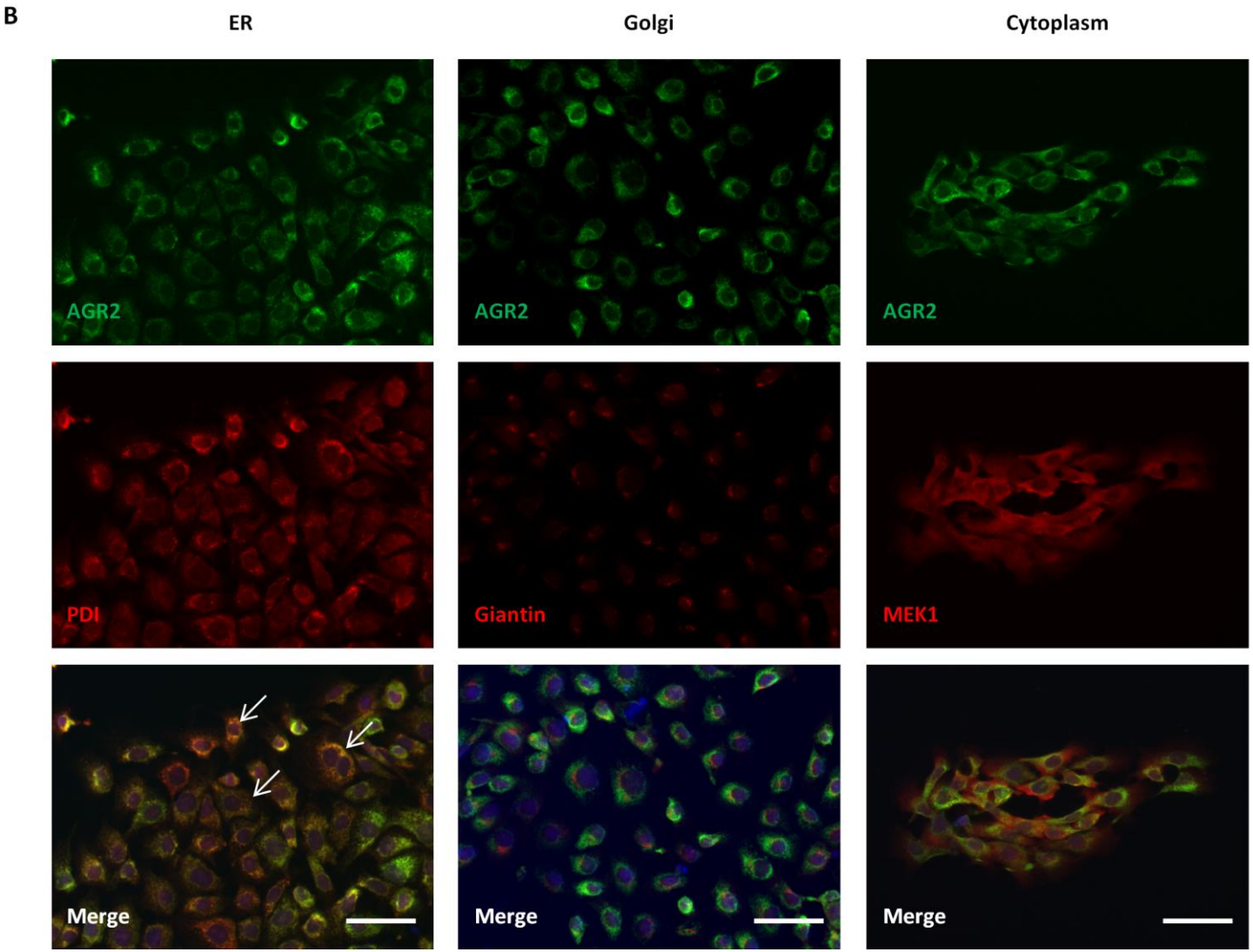


Figure 3.11. Cellular localisation of WT and mutant AGR2 protein (cont.). Cells were fixed with 4% PFA and permeabilised with 1% triton in PBS. Cells were co-stained with AGR2 (green) and the organelle markers PDI (for ER staining), Giantin (for Golgi staining) or MEK1 (for cytoplasmic staining), all in red. Representative images are shown for clonal cell lines expressing **(A)** WT AGR2, **(B)** E60A AGR2, **(C)** C81S AGR2, **(D)** Δ1-20 AGR2, **(E)** Δ21-40 AGR2, **(F)** ΔKTEL AGR2 and **(G)** empty vector. Nuclei (blue) were stained with DAPI. Solid white arrows indicate areas of co-localisation. Dashed arrows indicate possible nuclear AGR2 staining. Scale bar = 50 µm.

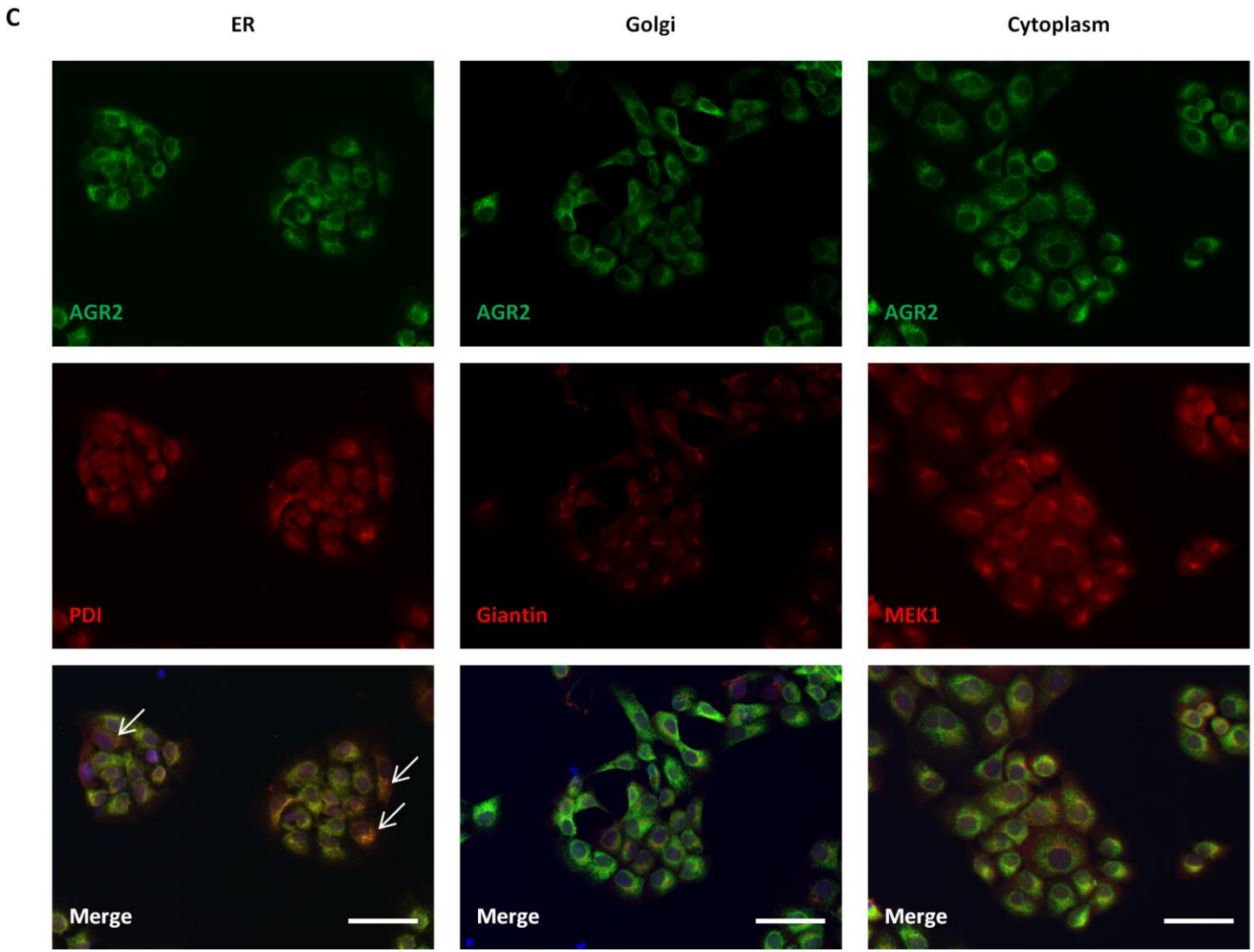


Figure 3.11. Cellular localisation of WT and mutant AGR2 protein (cont.). Cells were fixed with 4% PFA and permeabilised with 1% triton in PBS. Cells were co-stained with AGR2 (green) and the organelle markers PDI (for ER staining), Giantin (for Golgi staining) or MEK1 (for cytoplasmic staining), all in red. Representative images are shown for clonal cell lines expressing **(A)** WT AGR2, **(B)** E60A AGR2, **(C)** C81S AGR2, **(D)** Δ1-20 AGR2, **(E)** Δ21-40 AGR2, **(F)** ΔKTEL AGR2 and **(G)** empty vector. Nuclei (blue) were stained with DAPI. Solid white arrows indicate areas of co-localisation. Dashed arrows indicate possible nuclear AGR2 staining. Scale bar = 50 µm.

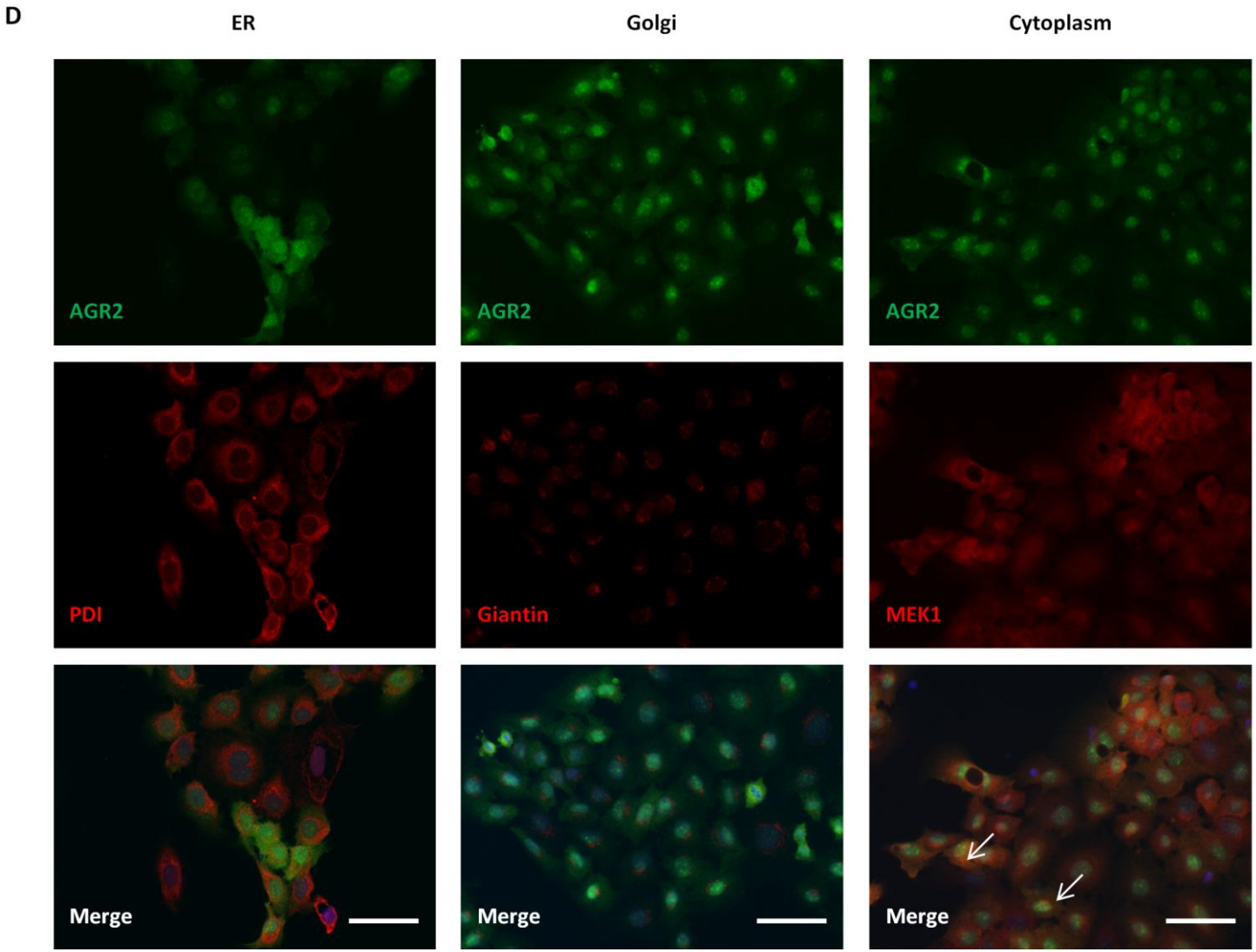


Figure 3.11. Cellular localisation of WT and mutant AGR2 protein (cont.). Cells were fixed with 4% PFA and permeabilised with 1% triton in PBS. Cells were co-stained with AGR2 (green) and the organelle markers PDI (for ER staining), Giantin (for Golgi staining) or MEK1 (for cytoplasmic staining), all in red. Representative images are shown for clonal cell lines expressing **(A)** WT AGR2, **(B)** E60A AGR2, **(C)** C81S AGR2, **(D)** Δ1-20 AGR2, **(E)** Δ21-40 AGR2, **(F)** ΔKTEL AGR2 and **(G)** empty vector. Nuclei (blue) were stained with DAPI. Solid white arrows indicate areas of co-localisation. Dashed arrows indicate possible nuclear AGR2 staining. Scale bar = 50 µm.

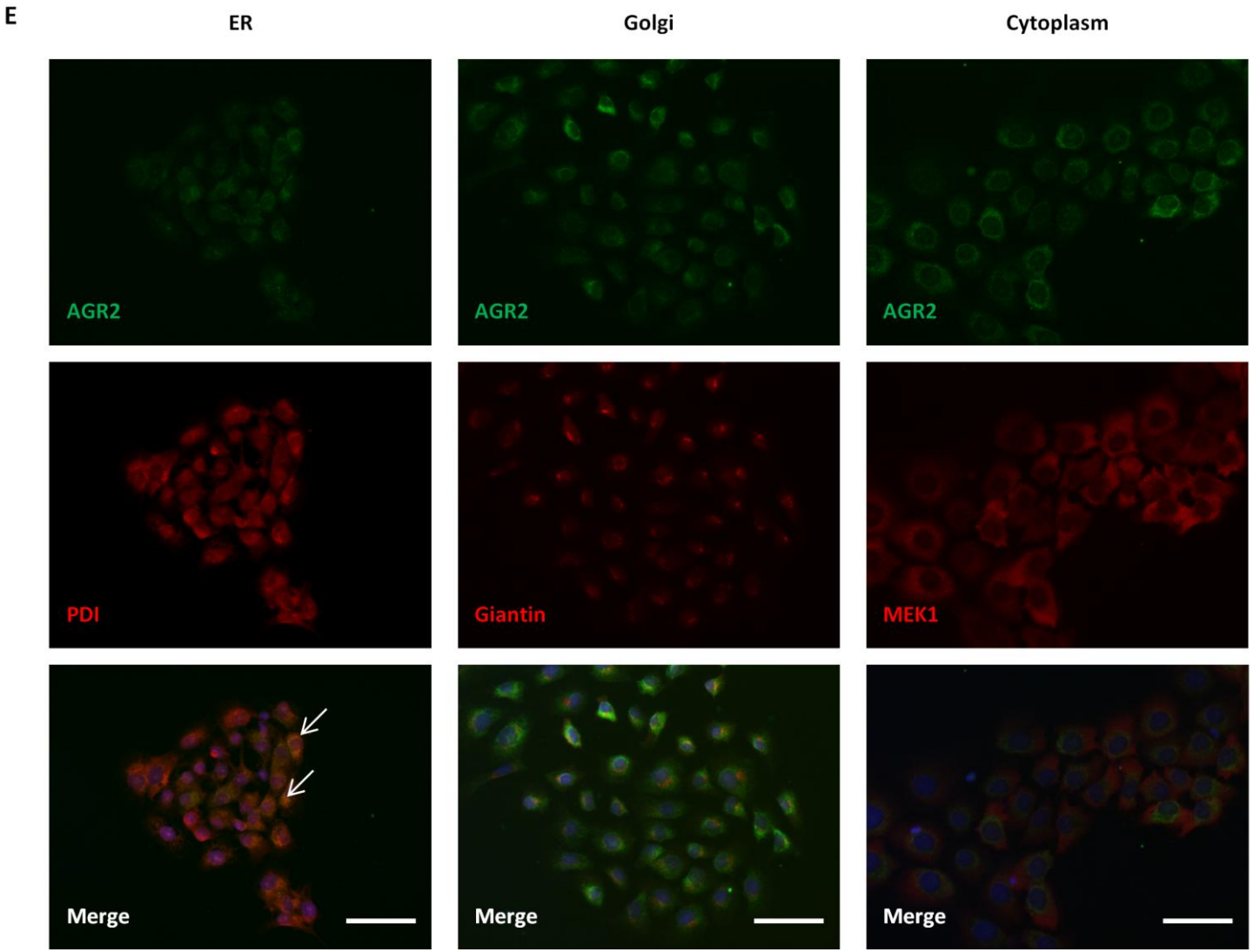


Figure 3.11. Cellular localisation of WT and mutant AGR2 protein (cont.). Cells were fixed with 4% PFA and permeabilised with 1% triton in PBS. Cells were co-stained with AGR2 (green) and the organelle markers PDI (for ER staining), Giantin (for Golgi staining) or MEK1 (for cytoplasmic staining), all in red. Representative images are shown for clonal cell lines expressing **(A)** WT AGR2, **(B)** E60A AGR2, **(C)** C81S AGR2, **(D)** Δ1-20 AGR2, **(E)** Δ21-40 AGR2, **(F)** ΔKTEL AGR2 and **(G)** empty vector. Nuclei (blue) were stained with DAPI. Solid white arrows indicate areas of co-localisation. Dashed arrows indicate possible nuclear AGR2 staining. Scale bar = 50 µm.

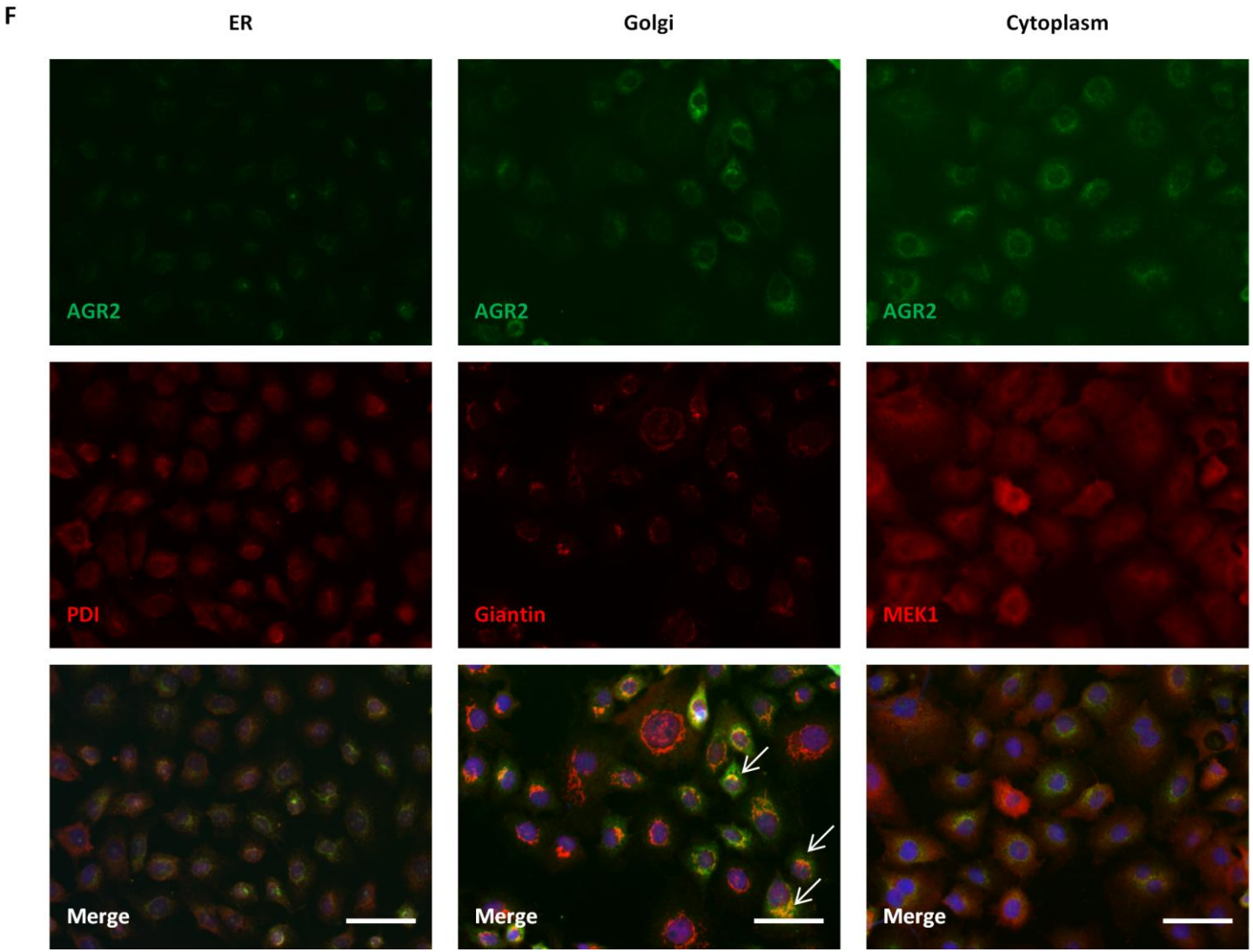


Figure 3.11. Cellular localisation of WT and mutant AGR2 protein (cont.). Cells were fixed with 4% PFA and permeabilised with 1% triton in PBS. Cells were co-stained with AGR2 (green) and the organelle markers PDI (for ER staining), Giantin (for Golgi staining) or MEK1 (for cytoplasmic staining), all in red. Representative images are shown for clonal cell lines expressing **(A)** WT AGR2, **(B)** E60A AGR2, **(C)** C81S AGR2, **(D)** Δ1-20 AGR2, **(E)** Δ21-40 AGR2, **(F)** ΔKTEL AGR2 and **(G)** empty vector. Nuclei (blue) were stained with DAPI. Solid white arrows indicate areas of co-localisation. Dashed arrows indicate possible nuclear AGR2 staining. Scale bar = 50 µm.

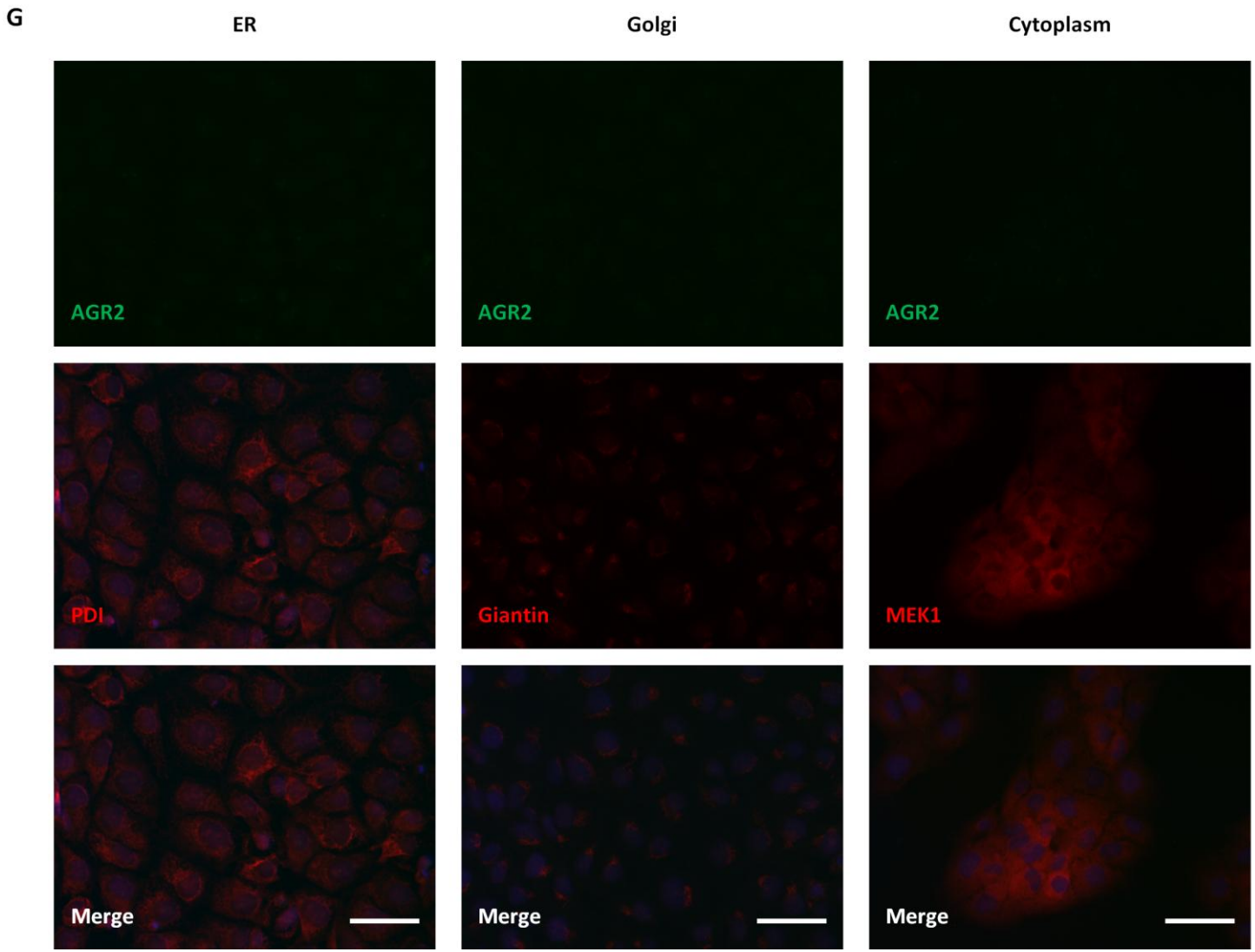


Figure 3.11. Cellular localisation of WT and mutant AGR2 protein (cont.). Cells were fixed with 4% PFA and permeabilised with 1% triton in PBS. Cells were co-stained with AGR2 (green) and the organelle markers PDI (for ER staining), Giantin (for Golgi staining) or MEK1 (for cytoplasmic staining), all in red. Representative images are shown for clonal cell lines expressing **(A)** WT AGR2, **(B)** E60A AGR2, **(C)** C81S AGR2, **(D)** Δ1-20 AGR2, **(E)** Δ21-40 AGR2, **(F)** ΔKTEL AGR2 and **(G)** empty vector. Nuclei (blue) were stained with DAPI. Solid white arrows indicate areas of co-localisation. Dashed arrows indicate possible nuclear AGR2 staining. Scale bar = 50 µm.

3.3 Discussion

3.3.1 Molecular characterisation of AGR2 dimer formation

It has been demonstrated here that AGR2 exists in a monomer-dimer equilibrium (also see Appendix 1), a fact that has been subsequently supported by the work of others [416, 558]. Dimerisation is also observed in other PDIs and oxidoreductases, notably PDI itself [559, 560], ER oxidoreductin 1α (Ero1α) [561] and Ero1β [562]. Interestingly, ERp29, along with the *Legionella pneumophila* disulphide bond A2 (DsBA2) protein, require dimerisation for activity [537, 563]. Dimerisation of AGR2 could therefore also be important for its activity, given that AGR2 is also a member of the PDI family [378].

The AGR2 dimer complex was predicted to be held together by salt bridges between Glu60 and Lys64 residues on opposing subunits, but only mutation of the Glu60 totally inhibited dimer formation. However, the shift towards a monomer species seen with Tyr63 mutation points to an important role for Tyr63 in AGR2 dimerisation. The aromatic ring interactions measured between the phenol side chains of the Tyr63 residues (Appendix 5) could provide some stabilisation to the structure [564], but another implication of the presence of these aromatic rings is the possibility of electrostatic interactions between these electron-dense π-orbitals, and the positively-charged side chain of Lys64. These cation-π interactions are found in a number of protein structures [565] and can often be found at homodimer interfaces [566]. Importantly, these interactions have been calculated to provide greater stabilisation than salt bridge interactions [567]. However, using the CaPTURE server for measurement of cation-π interactions [565], these lysine-tyrosine interactions do occur, but are not predicted to be energetically significant in AGR2 (Appendix 6) nor were these interactions detected for $AGR2_{41\text{-}175}$ using the Protein Interaction Calculator (PIC) server (Appendix 5).

Tyrosine residues have also been shown to contribute to the stabilisation of protein molecules through hydrogen bonding from their hydroxide groups, for example, intrachain hydrogen bonding between tyrosine and glutamate residues stabilise RNase Sa from *Streptomyces aureofaciens* [568]. The close proximity of two tyrosine and two glutamate residues in the dimer interface raised the possibility that inter-chain hydrogen bonding could occur between these residues. However, while such intra-chain Tyr-Glu hydrogen bonds exist in the AGR2 structure (Appendix 7), there were no measured inter-chain hydrogen bonds between tyrosine and glutamate residues (Appendix 5 and 7). However,

the AGR2 dimer structure was determined using a labelling scheme involving heavy-isotope labelling of leucine, isoleucine and valine residues. As a result, the exact positions of amino acid side-chains on residues other than these three could not be determined [421] (Appendix 1). Therefore, as tyrosine side-chains exhibit high degrees of flexibility and can exist in a number of conformations [569], the uncertainty about the orientation of the side-chain within the dimer interface means that the presence of these inter-chain hydrogen bonds cannot be totally discounted, as the measurements by the PIC and HBPLUS servers (Appendix 5 and 7) were based on the orientations shown in the NMR structure in Fig. 3.2. It would be interesting therefore to test the oligomerisation state of AGR2 with a tyrosine to phenylalanine mutation at Tyr63, both in the presence and absence of Lys64. This should shed some light on whether hydrogen bonding and/or π-stacking interactions contribute to dimer formation. Overall, considering these factors and the oligomerisation data presented above, the complete loss of dimerisation after Glu60 mutation can probably be attributed to the loss of both salt bridge formation with Lys64 and also loss of any potential hydrogen bonding with Tyr63. The presence of aromatic ring stacking interactions between Tyr63 residues probably also contributes to dimerisation, although these forces alone are not sufficient to hold the complex together. The incomplete loss of dimerisation with the loss of Lys64 may provide some support for the existence of hydrogen bonding between Tyr63 and Glu60, as these relatively weaker interactions could allow limited dimerisation, but this will require further investigation as outlined above.

3.3.2 Contributions to AGR2 dimer formation

In addition to the actual interactions involved in promoting AGR2 dimerisation at the homodimer interface, other factors influencing the extent of dimerisation were also explored. Higher concentrations of AGR2 protein promote increased dimerisation of the protein *in vitro*, as demonstrated here and by others [558]. Gray and colleagues [558] in fact demonstrated a greater effect of AGR2 concentration on dimer formation using 5-2500 µg/mL AGR2 concentrations than seen here in the 500-2000 µg/mL concentration range. The functional significance of these observations however is not yet clear, as they were observed at non-intracellular concentrations of NaCl and the local intracellular concentration of AGR2 is unknown. The contribution of the 21-40 flexible region to AGR2 dimer formation must also be considered. Whilst the dimerisation potential of $AGR2_{21\text{-}175}$ and $AGR2_{41\text{-}175}$ appeared to be relatively similar, the much higher susceptibility of $AGR2_{41\text{-}175}$

to dimer dissociation at high salt concentration points to a possible contribution of the flexible region in dimer stabilisation, but as mentioned in Section 3.2.1.4, the relevance of the contribution of the flexible region at physiological salt concentration is unclear. The 21-40 region has been implicated in dimer stability [558], but the authors found that deletion of residues 1-45 of AGR2 led to a more stable dimer, as evidenced by higher rates of spontaneous dimer formation seen on SDS-PAGE gels.

It is difficult to correlate these opposing findings, and whilst it seems unusual that a dimer held together through electrostatic interactions would remain virtually intact during denaturing SDS-PAGE [558], a low amount of these SDS-resistant dimers have been observed during recombinant protein production for the present study, and there does appear to be a higher percentage of these dimers in AGR2 $_{41\text{-}175}$ (Appendix 8). Conversely, recombinant E60A AGR2$_{21\text{-}175}$ did not seem to produce these dimers, suggesting that they do occur through electrostatic interactions. SDS-resistant dimers have been previously observed in other proteins, with the authors suggesting that this might result from unfavourable SDS interactions with negatively charged regions of the protein [570]. The region surrounding Glu60 contains a relatively high proportion of negatively charged and hydrophobic residues (Fig. 3.1A), which would not interact favourably with negatively-charged SDS and thus a small percentage of dimers might remain bound through electrostatic interactions of the dimerisation interface. However, the relatively high amount of positive charge within the 1-40 and 1-45 regions should provide favourable interactions with the negatively charged SDS, thus helping to 'rip open' the dimer in the majority of species. Deletion of this region would result in a loss of this dimer-disrupting interaction, increasing the percentage of dimers able to complex through the Glu60 region. Therefore, analysis by SDS-PAGE may not be the optimal method for investigating such a dynamic process and could possibly help dimer formation through electrostatic exclusion. Further experimentation will be required to account for these differences.

It is important to highlight that any role in dimerisation played by the 21-40 region appears to be in the context of high, non-physiological Na^+ concentration, although this concentration was similar to physiological KCl concentration [544]. While the data presented here support the notion that AGR2 dimerises through electrostatic interactions, the effects of high salt concentration demonstrate not only that these interactions appear to be fairly robust, but also that the presence of the 21-40 region stabilises the AGR2$_{21\text{-}175}$ dimer relative to the AGR2$_{41\text{-}175}$ dimer. Whether this has any physiological significance is not clear. A second point to address is the report from Ryu and colleagues [416] of AGR2 dimerising through its single cysteine. This observation was based on crosslinking in

cultured cells using a membrane-permeable crosslinker, where mutation of Cys81 led to a loss of crosslinked AGR2 homodimers. It has been demonstrated here however that in the presence of only recombinant AGR2, dimerisation is not affected by mutation of Cys81 (see also Appendix 1). The authors also reported the requirement of AGR2 Cys81 for interaction with glucose-related protein 78 kDa (GRP78), but it is not clear whether these interactions occur via free thiol-groups in GRP78 as previously reported [571], or whether it is the reported thiol-dependent dimerisation of AGR2 that is required for this interaction. With this in mind, it seems unlikely that AGR2 would dimerise through the same residue required to interact with one of its substrates MUC2 [422], unless this acts as another form of AGR2 regulation. It may be that, rather than allowing dimerisation, Cys81 allows similar thiol-dependent AGR2-GRP78 interactions, thereby bringing two AGR2 molecules together long enough to be crosslinked. This might also suggest that *in vivo*, AGR2 may exist primarily as a monomer. A possible alternative is that the apparent dimer observed in the experiments described by Ryu and colleagues [416] is a crosslinked heterodimer of AGR2 with a similar sized protein.

Overall, there is conflicting evidence on the dimerisation of AGR2, but it appears that in isolation, AGR2 dimerises through a disulphide-independent mechanism [421, 558]. The dimerisation of AGR2 within the cell may require further investigation, perhaps through the use of fluorescence resonance energy transfer (FRET) in living cells as carried out previously [572] to demonstrate association (or lack of association) of mixed WT, E60A and C81S proteins.

3.3.3 Biochemical activity of AGR2

Having explored the factors influencing AGR2 dimerisation, the effects of this dimerisation on the function of AGR2 were tested through the use of biochemical assays. However, AGR2 did not appear to function in any biochemical assays for PDI-related activity, showing no activity in reductase, isomerase and glutaredoxin assays. This brings into question the enzymatic activity of the protein. It is possible that AGR2 might act as just an oxidase, but this seems somewhat unlikely given that PDIs with a single Cys to Ser mutation in the active site become trapped in disulphide bonded intermediates [317]. While the cysteine from the second AGR2 subunit might provide the necessary reducing power to break the AGR2-substrate interaction [317], its position on the opposite side of

the molecule as indicated in the three-dimensional structure of AGR2 (Fig. 3.1) makes this geometrically hard to achieve.

A clue to the activity of AGR2 might come from ERp44, another PDI family member containing a CxxS active site [320]. Like AGR2, ERp44 is upregulated in response to ER stress, indicating an involvement in protein folding [320]. Interestingly, rather than functioning as a PDI, ERp44 appears to contribute to protein folding through interactions with free thiol groups on substrate proteins and the oxidoreductase Ero1α [320, 573]. This stems from the initial observation that an exposed thiol in the tail of immunoglobulin (Ig) μ chain mediates retention of the chain in the ER [574], and it was later determined that this retention was dependent on reversible DSB formation with ERp44 [573]. The ERp44-mediated retention of Ero1 in the ER then points to a mechanism whereby ERp44 links unpaired thiols with an oxidoreductase, thereby aiding the correct folding of the substrate proteins.

It is possible therefore that AGR2 may function in a similar manner. MUC2 is a thiol-dependent substrate of AGR2 [422], and its large size and high number of DSBs presumably requires extended interactions with PDIs and chaperones to fold correctly. Therefore binding of free thiol groups by AGR2 may both prevent the secretion of an incorrectly folded MUC2 and other substrates by retaining them within the ER, and could also act as a 'signpost' to other PDIs, directing them to sites of unpaired cysteines. The fact that the Cys81 residues of AGR2 are located on opposite sides of the dimer suggest that it could hold adjacent unpaired thiols on substrate proteins in close proximity for oxidation by PDIs, or alternatively use one of these Cys81 thiols to bind a PDI and target its activity to a particular free thiol. This idea is supported to some extent by the reported Cys81-dependent interaction of AGR2 with GRP78 [416] as discussed above, although it is not clear if these two proteins directly interact through thiol-dependent interactions.

3.3.4 AGR2-expressing cell lines as a tool for exploring protein domains involved in protein function

The lack of AGR2 activity in biochemical assays, as discussed above, requires the use of a different assay to measure the effects of dimerisation and other protein domains on AGR2 function. AGR2 has been shown to be involved in mediating cell adhesion [309, 515] and so stable cell lines expressing various mutant forms of AGR2 were created and characterised here with a view to later use in adhesion-based assays. However, it is crucial

that mutations did not affect sub-cellular distribution of the protein (unless intentionally done so) to avoid any potential effects caused by protein localisation rather than the desired mutation itself.

WT and mutant AGR2 protein localised largely to their expected cellular compartments, with the majority of protein localising to the ER and ΔKTEL protein localising to the Golgi, consistent with its high levels of secretion. Unexpectedly, Δ1-20 AGR2, which is expected to localise to the cytoplasm as a result of losing its ER entry sequence, was abundantly found within the nucleus. This was also observed, but to a much lesser extent, with WT AGR2. The reasons for this nuclear accumulation are not clear, but it is interesting that two putative but overlapping nuclear localisation sequences (NLS) at the C-terminus of AGR2 are predicted (Fig. 3.12), that are also predicted to localise the protein to both the nucleus and cytoplasm (Appendix 4). These NLS's are dependent on importin for nuclear localisation [575], which acts as a receptor for NLS-containing proteins and promotes their interaction with the nuclear pore complex, thus facilitating their transport across the nuclear membrane (reviewed in [576]).

MEKIPVSAFLLLVALSYTLARDTTVKPGAKKDTKDSRPKLPQTLSRGWGDQLIWTQTYEE
ALYKSKTSNKPLMIIHHLDECPHSQALKKVFAENKEIQKLAEQFVLLNLVYETTDKHLSP
DGQYVPRIMFVDPSLTVRADITGRYSNRLYAYEPADTALLLDNMKKALKLLKTEL

Figure 3.12. ***Prediction of AGR2 nuclear localisation sequences (NLS) by cNLS Mapper server [575].*** The two predicted sequences overlap, thus one is highlighted in red and the other underlined.

Whether this sequence is active in promoting nuclear import of AGR2 is not clear, but it is interesting that Pohler and colleagues [445] noted that C-terminal truncation of AGR2, which included most of the putative NLS, led to a switch from promoting oesophageal cancer cell survival to reducing it, relative to non-AGR2 expressing cells. Whether this equates with a loss of AGR2 nuclear accumulation would be interesting to determine. Additionally, AGR2 is reported to interact with Reptin, a member of the NuA4 histone acetyltransferase complex [539]. Reptin is found both in the nucleus and cytoplasm, although its role in the cytoplasm is not clear (reviewed in [577]). This interaction with Reptin may provide evidence of nuclear AGR2, or at least non-ER AGR2, although the subcellular localisation of this AGR2-Reptin interaction was not determined. Further

investigation of the AGR2-Reptin interaction could provide crucial evidence as to whether AGR2 can indeed enter the nucleus.

This nuclear accumulation of AGR2 offers some new insights into the possible roles of AGR2 within the cell, but overall the cell lines created in this chapter offer suitable tools for the study of the effects AGR2 mutations on its function. Apart from Δ1-20 AGR2, all mutations localised to their expected compartments and as Δ1-20 AGR2 was utilised to examine the dependence of AGR2 residency within the ER on its function, it still remains a viable tool to achieve this aim. Overall, these cell lines should provide an adequate tool for investigating AGR2 function in a cellular context.

3.4 Conclusions

In this chapter, the dimerisation of AGR2 has been explored in detail and the residues involved in this process have been determined, although the exact way that AGR2 dimerises requires a little more investigation. The demonstration that AGR2 does not appear to display any PDI activity raises the possibility of an enzyme-independent function of the protein that might involve sequestration of free thiol groups. Determining the exact mechanisms of AGR2 dimerisation may be crucially important in the design of any potential inhibitors of AGR2, should the inhibition of dimerisation prove an effective manner to inhibit AGR2 activity. In a similar vein, investigating any enzyme activity of AGR2 has similar prospects for the design of anti-AGR2 therapeutics. In line with this, the creation of a panel of mutant-AGR2-expressing cell lines in this chapter provides tools for the exploration of the contributions of these structural features to AGR2 activity, in the absence of other suitable, biochemical-based assays.

Chapter 4

The Role of AGR2 in Cell Adhesion

4.1 Introduction

Expression of AGR2 has been linked to reduced patient survival in a several cancers [309, 433, 443, 457, 458, 467, 474]. A number of studies have also linked AGR2 to cell migration and invasion, both through actions of endogenous protein [301, 415, 440, 481] and also through the stimulatory actions of extracellular recombinant protein [414], demonstrating that AGR2 can increase the migratory and invasive potential of cells *in vitro*. The elevated levels of AGR2 observed in metastatic cells of differing origin further implicate AGR2 in the metastatic process [302-304, 306, 469], with the strongest evidence for an association of AGR2 with metastasis coming from the ability of AGR2 to confer a metastatic phenotype on an otherwise benign cell line in a rat model [309].

Cell metastasis and cell adhesion are inexorably linked, due to metastatic cells requiring reduced cell-cell adhesion and cell-matrix interaction in the primary tumour, as well as enhanced adhesion of circulating tumour cells to the vasculature at sites of secondary colonisation (reviewed in [578]). Accordingly, AGR2-expressing rat mammary tumour cells show a high rate of cell adhesion to a substratum [309], which was later confirmed to be an AGR2-dependent effect in cultured prostate cancer cells, through targeted knockdown of AGR2 [515]. Similarly, when coated onto a plastic substratum, extracellular AGR2 is able to promote adhesion of rat mammary cells, suggesting that AGR2 could also have an extracellular role in cell adhesion [309], as there are also several reports of its secretion by cells in culture [388, 426, 440].

In addition to these effects on cell adhesion, AGR2 has also been demonstrated to play a role in anchorage-independent cell growth, an important characteristic that metastasising tumour cells must acquire in order to avoid cell death induced by loss of contact with the extracellular matrix (anoikis), that occurs upon migration away from the primary tumour site [279]. AGR2 was shown to confer resistance to anoikis in oesophageal, lung and breast cancer cell lines [445, 481, 482], as well as mouse NIH3T3 cells [481], as knockdown of AGR2 resulted in loss of anchorage-independent growth [445, 481, 482], whereas forced expression of AGR2 resulted in gain of anchorage-independent cell growth [445, 481]. Furthermore, the anti-apoptotic protein survivin has been implicated in mediating AGR2-driven anoikis resistance in breast cancer cells, as it is upregulated by AGR2 [482] and inhibits caspase-3 [579], a major effector of anoikis [580]. Additionally, resistance to anoikis in AGR2-expressing oesophageal cancer cells is driven by inhibition of p53 activity through blocking an activating phosphorylation event at Ser15 [445]. Thus AGR2 is implicated in two adhesion-based roles that can favour tumour progression.

4.1.4 Chapter objectives

Using the AGR2-expressing cell lines described in the previous chapter, the aim of the current chapter is to determine the contribution of several sequence and structural features of AGR2 to its effect on cell adhesion and anchorage-independent cell growth. As extracellular recombinant AGR2 is also known to affect the adhesion of Rama 37 cells, the aim here is also to investigate whether extracellular and intracellular AGR2 affect cell adhesion in different ways.

4.2 Results

4.2.1 Effects of extracellular AGR2 on rate of cell attachment

Previous studies have shown that AGR2 promotes cell adhesion when coated onto cell culture plates [309]. In the present experiments, recombinant WT AGR2 ($AGR2_{21\text{-}175}$) coated onto cell culture plates increased the percentage of applied Rama 37 cells attaching to these plates in 60 min from 18.6 % in uncoated wells to 45.0 % in coated wells (Fig. 4.1), values that were similar to those obtained previously [309], but achieved over a 60 min rather than a 30 minute time period. In contrast, coating with an equimolar amount of BSA resulted in a significant reduction in the percentage of cells adhering compared to uncoated controls (2.9 % vs. 18.6% of cells adhering, $p < 0.001$, Student's t-test). A similar effect of BSA has been reported previously using Chinese hamster ovary (CHO) cells [581].

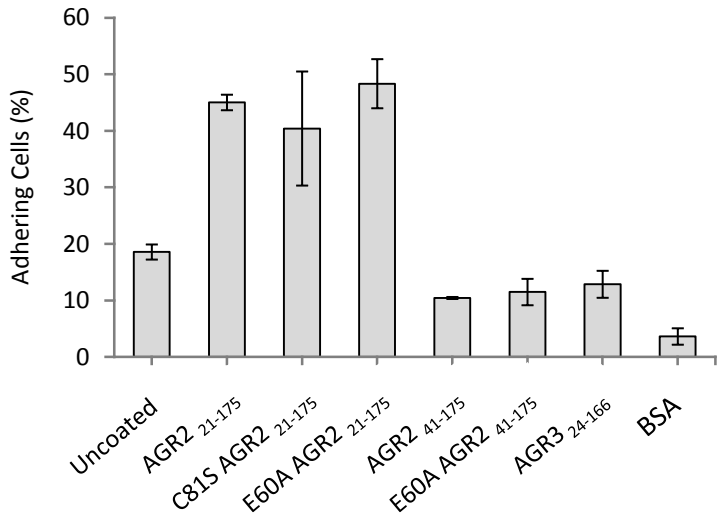


Figure 4.1. ***Extracellular AGR2 promotes cell adhesion*.** A 2 μM solution of each of the indicated proteins was coated onto plastic cell culture multiwell dishes overnight. Rama 37 cells were then allowed to adhere to the coated wells for 60 min and the number of bound cells was determined using an automated cell counter. The number of adhering cells is expressed as a percentage of the total number of cells added to each well at the start of the experiment. The means from duplicate wells from 3 independent experiments are shown, ±SE. The numbers in subscript denote the amino acids present in the relevant proteins, while prefixes denote point mutations present in these proteins (see main text for description). BSA: bovine serum albumin.

Given that AGR2 can be a dimer and is a member of the PDI family, the importance of dimerisation or the putative CxxS active site to AGR2-promoted cell adhesion was investigated, using monomeric AGR2 protein (E60A $AGR2_{21-175}$) and active site-null AGR2 protein (C81S $AGR2_{21-175}$, see Chapter 3), respectively. There was no significant difference in adhesion rate compared to WT protein with either C81S $AGR2_{21-175}$ (p = 0.593, Student's t-test) or with E60A $AGR2_{21-175}$ (p = 0.772). By contrast, coating the wells with AGR2 lacking the flexible 21-40 amino acid region (see Chapter 3), either in dimeric form ($AGR2_{41-175}$) or monomeric form (E60A $AGR2_{41-175}$), led to a significant decrease in adhesion rate compared to $AGR2_{21-175}$-coated wells (p < 0.001 in both cases). While this decrease alone could result from $AGR2_{41-175}$ and E60A $AGR2_{41-175}$ proteins not adhering to the wells, the fact that coating with these two proteins also led to a significant decrease in adhesion relative to uncoated wells (p = 0.001 and p = 0.021, respectively) suggests that they must be indeed be present to actively reduce adhesion relative to uncoated wells. Therefore, these results suggest that amino acids 21-40, present in $AGR2_{21-175}$, but not in $AGR2_{41-175}$, may be important in AGR2-enhanced cell adhesion on a plastic substratum. As further evidence of the importance of these residues, coating wells with WT AGR3 protein ($AGR3_{24-166}$), which is closely related to AGR2 in sequence, but naturally lacks the region corresponding to AGR2 residues 21-40 [378], not only failed to enhance adhesion but also decreased adhesion relative to uncoated wells (p = 0.05) as observed with $AGR2_{41-175}$.

The flexible AGR2 21-40 region contains seven positively charged amino acids (two arginine and five lysine residues) and only three negatively charged residues (three aspartate residues). This raises the possibility that the increase in cell adhesion observed with AGR2 could be the result of a non-specific interaction simply caused by the presence of increased positive charge, in a manner akin to coating with poly-lysine [582]. However, experiments where wells were coated with solutions of equal concentration of 21-40 peptide or an unrelated peptide showed no difference in adhesion between uncoated and peptide-coated wells, suggesting that the overall positive charge of the flexible region is not responsible for the increased adhesion seen with $AGR2_{21-175}$- coated wells (Appendix 9).

To ensure that the observed effects on adhesion from peptide coating were not due to inefficient peptide coating of the wells, indirect ELISA assays were performed on AGR2- and peptide-coated wells to determine the amount of peptide attached to the wells after coating, using an antibody raised against an epitope in this 21-40 region, which thus recognises both $AGR2_{21-175}$ protein and 21-40 peptide [475]. These assays indicated that it was not possible to coat the 21-40 peptide at as high a level as that achieved with 2 µM $AGR2_{21-175}$ protein solution, although this conclusion assumes that the antibody binds

equally well to AGR2 protein and to the 21-40 peptide. Further calibration experiments indicated that, on average, a 7.5 nM solution of $AGR2_{21-175}$ protein needed to be coated onto wells in order to achieve the same ELISA signal intensity as a 2 µM peptide solution coating (Appendix 10). Thus, wells coated with a 7.5 nM $AGR2_{21-175}$ protein solution or 2 µM 21-40 peptide solution were then subjected to an adhesion assay as above, with slight modifications due to the low numbers of adhering cells observed, as a consequence of coating very low amounts of AGR2 protein (see Materials and Methods for full description).

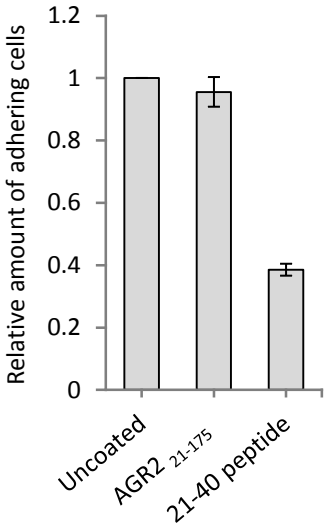


Figure 4.2. ***Amino acids 21-40 of AGR2 alone are not sufficient to promote cell adhesion*****.** In order to reach equal levels of coating (as determined by indirect ELISA assay), a 2 µM solution of 21-40 peptide and 7.5 nM solution of $AGR2_{21-40}$ protein in coating buffer were coated onto non-cell culture plates overnight. Rama 37 cells were then allowed to adhere to the plates for 60 min and the number of adhering cells was quantified using crystal violet staining, measured by optical density at 570 nm. The intensity of cell staining for coated plates is expressed relative to intensity of cell staining in uncoated wells. The means from triplicate wells from 3 independent experiments are shown, ±SE.

Coating with just 7.5 nM $AGR2_{21-175}$ solution had no effect on cell adhesion compared to uncoated wells (p = 0.448), but a corresponding equal amount of 21-40 peptide attached to the plate resulted in a significant decrease in cell adhesion (61 % reduction in relative number of adhering cells, as indicate by crystal violet staining; p = 0.001, Fig. 4.2). Together, these data provide evidence that the 21-40 peptide alone is not able to support cell adhesion, but rather that the adhesion-promoting activity of AGR2 protein requires the presence of both the globular part of the protein (residues 41-175) and the flexible region composed of residues 21-40.

4.2.2 Effects of intracellular AGR2 on cell attachment

The effects of AGR2 expression on the adhesion of rat mammary (Rama 37) cells were investigated, to further explore the AGR2-mediated changes to cell behaviour that might contribute to the metastatic potential of AGR2-transfected Rama 37 cells [309].

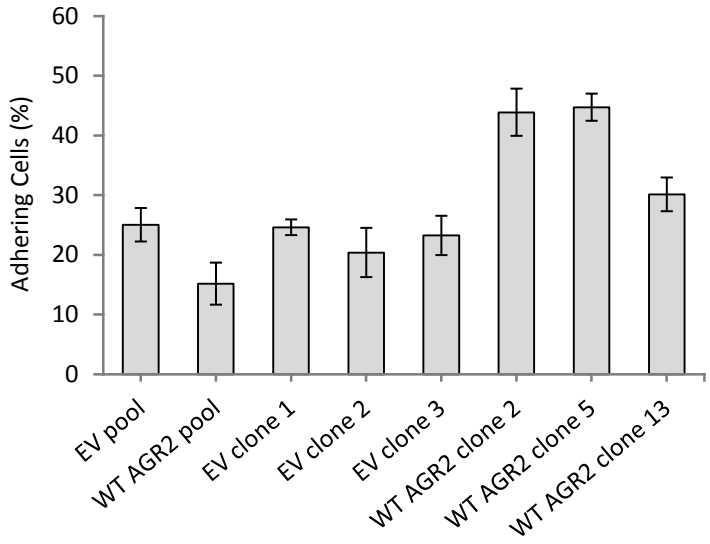


Figure 4.3. ***Clonal AGR2-expressing cells exhibit a greater rate of cell attachment than control cells*.** Wild type (WT) AGR2-expressing and empty vector (EV) control cells were allowed to adhere for 30 min to cell culture plates and the number of bound cells was determined using an automated cell counter. The number of adhering cells is expressed as a percentage of the number of cells added to each well. The means from duplicate wells from 3 independent experiments are shown, ±SE.

Cell attachment assays performed on uncoated cell culture plates using both pooled and clonal WT-AGR2-expressing and empty vector (EV) control cells are shown in Fig. 4.3. There was a reduction in the number of WT AGR2 pool cells attaching compared to EV pool cells, but this was not statistically significant (p = 0.131, Student's t-test). Three separate EV clone cells showed levels of attachment that were not significantly different from each other (smallest p value = 0.445) or EV pool cells (smallest p value = 0.465), but there was a significant difference between EV clones and WT AGR2 clones 2 and 5 (highest p value = 0.016), but not between EV clones and WT AGR2 clone 13 (smallest p value = 0.167). Clonal WT AGR2 cells not only attached more rapidly than EV clonal cells, but also exhibited resistance to trypsin once attached to the substratum, as demonstrated

qualitatively by increased trypsin incubation times required to detach these cells for counting. AGR2-dependent trypsin resistance was also observed during experiments with coated protein in Section 4.2.1, when being collected for counting. In order to better understand this process and to determine whether resistance to trypsin is a feature of AGR2-expressing cells and separate from the rate of cell attachment, cell lines were subjected to quantitative cell detachment assays: cells adhering more strongly to the plate will not be dislodged by trypsinisation to the same extent as loosely adhering cells (see Materials and Methods for full description). Preliminary experiments to determine the optimal concentration of trypsin in this assay are shown in Appendix 11.

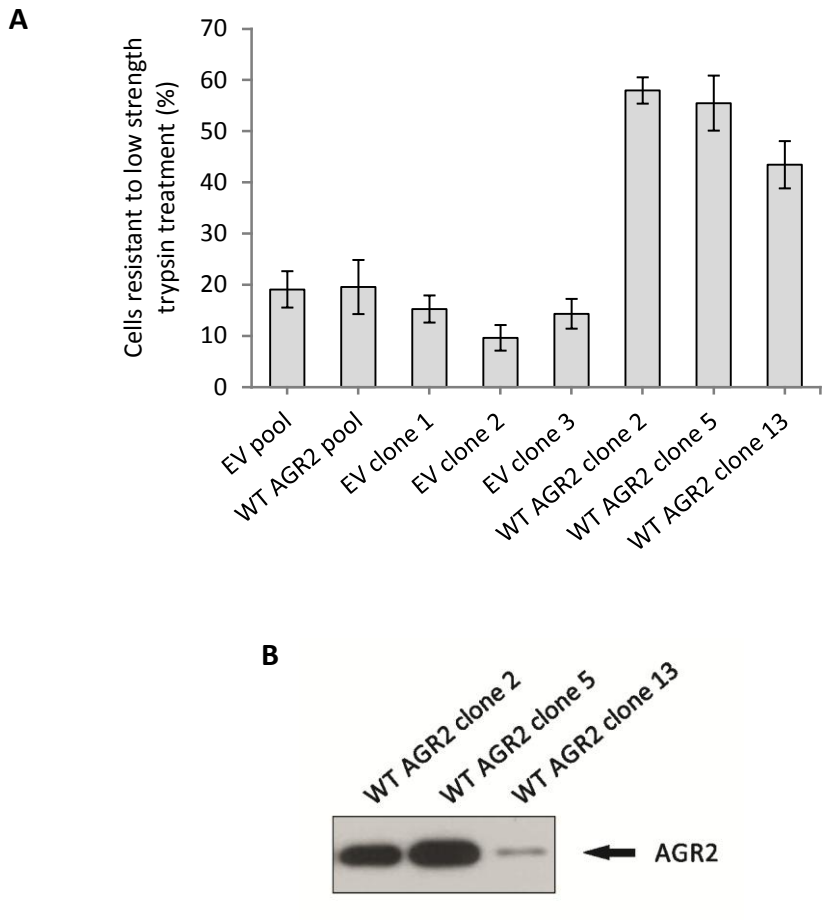


Figure 4.4*. Clonal AGR2-expressing cells are more resistant to trypsin than control cells.* **(A)** Wild type (WT) AGR2-expressing and empty vector (EV) control cells were allowed to adhere for 30 min to cell culture plates. Attached cells were then treated with 0.0125 % (v/v) trypsin for 5 min and the remaining attached cells were counted with an automated cell counter after removal with 0.05 % (v/v) trypsin. The percentage of cells resistant to 0.0125 % (v/v) trypsin treatment represents the proportion of applied cells attaching in 30 min that remained adherent after 0.0125 % (v/v) trypsin treatment. The means from duplicate wells from 3 independent experiments are shown, ±SE. **(B)** Western blot of AGR2 expression levels from the three clonal AGR2 cell lines used.

Clonal cells expressing WT AGR2 were significantly more resistant to removal by trypsin than EV clonal or pooled cells (largest p value = 0.014, Student's t-test), with the greatest difference observed between EV clone 2 and WT AGR2 clone 2 ($p < 0.001$). There was no significant difference in the proportion of trypsin-resistant cells between WT AGR2 and EV pool cells ($p = 0.946$), with EV clonal cells demonstrating an average strength of adhesion that was not significantly different from that of pooled cells (smallest p value = 0.095, Fig. 4.4A). WT AGR2 clone 13, which expresses the lowest amount of AGR2 protein

of the three clones tested in the trypsin-based assay (Fig. 4.4B), showed a lower resistance to trypsin than either clones 2 or 5, although this was just outside statistical significance (p = 0.052 and p = 0.164, respectively). WT AGR2 clone 13 cells were still significantly more resistant to trypsin than EV clone cells, however (largest p value = 0.006). These results suggest that increased expression of AGR2 correlates broadly with increased resistance to trypsin (Fig. 4.5), and a statistically significant correlation was indeed observed between these two cell properties (Spearman's σ = 0.880, p = 0.021).

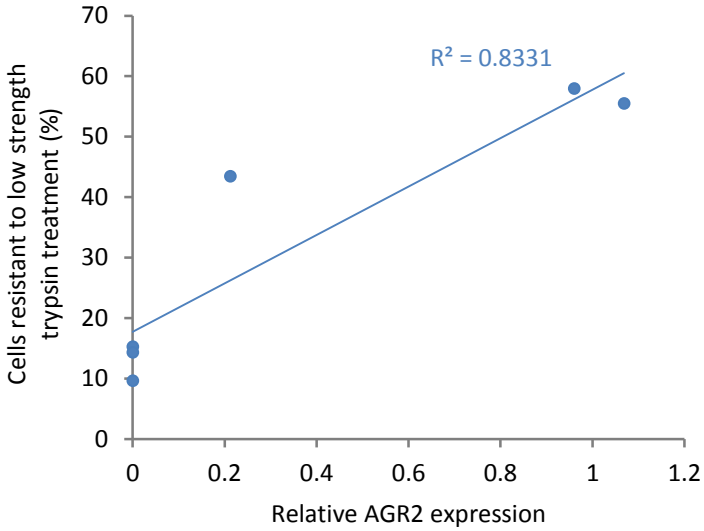


Figure 4.5. ***Resistance to trypsin detachment correlates with AGR2 expression.*** The percentage of trypsin-resistant clonal WT AGR2 and clonal EV cells shown in Fig. 4.4A were plotted against the steady-state levels of AGR2 in these cells, as previously determined in Chapter 3.

In order to check that the above results were not due to the nature of the assay used, the two cell lines showing the greatest difference in trypsin resistance (EV clone 2 and WT AGR2 clone 2) were subjected to an alternative, centrifugation-based assay which measures the strength of cell adhesion based on the resistance of cells to detachment from a substratum by a direct perpendicular force [518, 519]. Here, a similar pattern was observed to that seen in the trypsin-resistance assay, with an approximate 2-fold increase in the proportion of WT AGR2 clone 2 cells remaining attached relative to EV clone 2 cells (Appendix 12). Although the difference was not as marked as with the trypsin-resistance assay, the results of this second assay confirm the enhanced strength of attachment of AGR2-expressing cells.

4.2.3 The effects of AGR2 on cell attachment change over time

The enhanced strength of adhesion of WT AGR2 clonal cells shown in section 4.2.2 did not match an observation that during routine subculture of cells, WT AGR2 clonal cells detached much more readily than EV clonal cells. To investigate this further, the sensitivities of EV clone 2 and WT AGR2 clone 2 cells to trypsin were compared over a 6 h time course (Fig. 4.6).

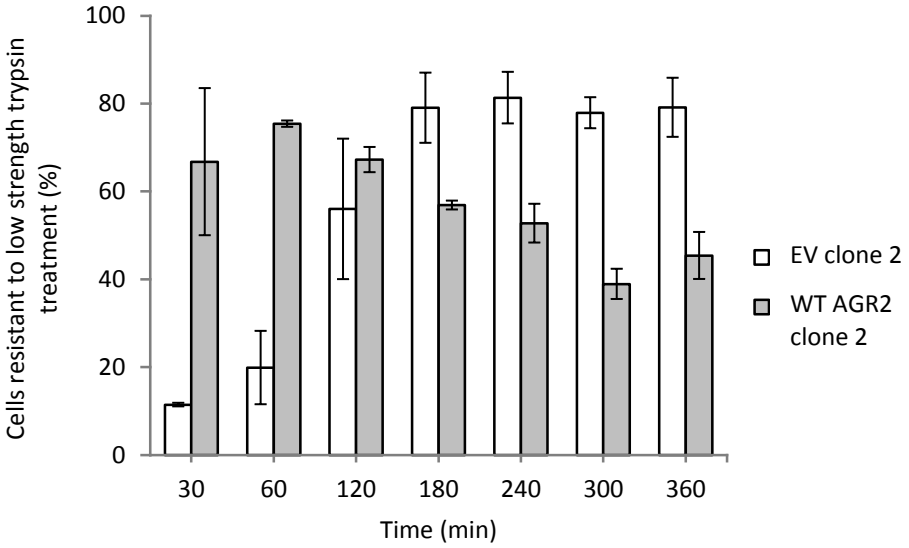


Figure 4.6. ***Time course of strength of adhesion*****.** Wild type (WT) AGR2-expressing and empty vector (EV) control cells were allowed to adhere to cell culture plates for the indicated times. Cells were then treated with a 0.0125 % (v/v) trypsin for 5 min and the remaining adhering cells were counted with an automated cell counter. The percentage of cells resistant to 0.0125 % (v/v) trypsin treatment represents the proportion of applied cells, which had attached by the indicated time points, that remained adherent after 0.0125 % (v/v) trypsin treatment. A control plate, seeded simultaneously with the treatment plate but not treated with low strength trypsin, was used to determine the number of cells present in wells at each time point, to control for any effects of cell division over the timeframe of the experiment. Note that means from duplicate wells from a single experiment only are shown and so error bars show SD only.

As expected, WT AGR2 clone 2 cells were very resistant to trypsin within the first 30 to 60 min, with 75 % of cells remaining attached at 60 min. Conversely, EV clone 2 cells were not very resistant to trypsin during these early time points, with only 20 % of cells remaining attached at 60 min. EV clone 2 cells gradually became more resistant to trypsin

over the next 300 min, with significantly more cells remaining attached by 180 min (p = 0.019, Student's t-test) and continuing at this level until the end of the assay. Conversely, there was a gradual fall in the resistance to trypsin of WT AGR2 clone 2 cells over the 60 min to 360 min time points, falling from 75 % at 60 min to 45 % at 360 min, although this was just outside statistical significance (p = 0.075), but this may relate to the small sample size.

Over this 6 h period, cells begin to divide and spread over the surface of the cell culture well, suggesting that adhesion strength might change more dramatically if cells are allowed to form a monolayer. Therefore, cells were grown overnight to 70-80 % confluence and subjected to the same cell detachment assay (Fig. 4.7).

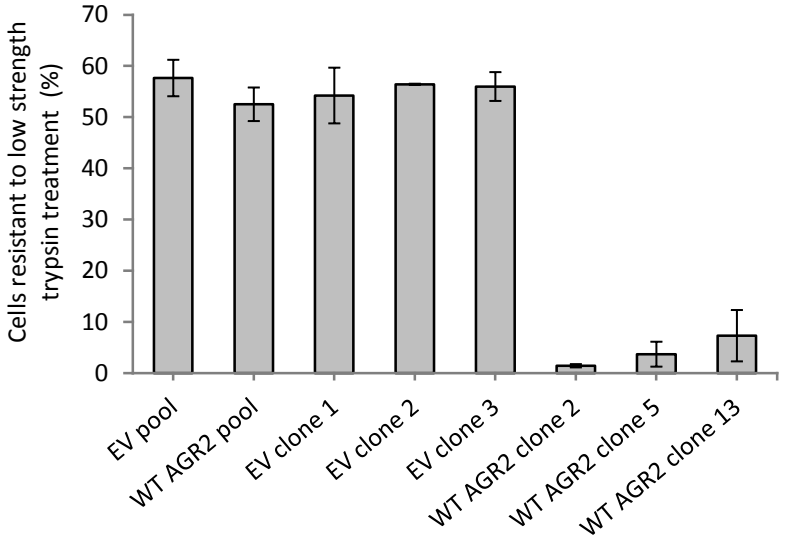


Figure 4.7. ***Clonal WT AGR2 cells are less resistant to trypsin than control cells when allowed to form a monolayer.*** Wild type (WT) AGR2-expressing and empty vector (EV) control cells were grown overnight to 70-80 % confluence and then trypsinised in 0.0125 % (v/v) trypsin for 5 min. The remaining adhering cells were counted with an automated cell counter. The number of trypsin-resistant cells is expressed as a percentage of the total number of cells in each well. The means from duplicate wells from 3 independent experiments are shown, ±SE.

As predicted, WT AGR2 clonal cells were significantly less resistant to trypsin than EV clonal cells once these cells had formed monolayers (largest p value = 0.003, Student's t-test), exemplified by the fact that 56 % of EV clone 2 cells remained adherent compared to just 1 % of WT AGR2 clone 2 cells. Notably, WT AGR2 clone 13 cells, which express lower levels of AGR2 than clone 2 or 5 cells, were also significantly less resistant to trypsin than

the three EV cloned cell lines (largest *p* value = 0.01) and were not statistically significantly different from clones 2 and 5 ($p = 0.362$ and $p = 0.552$, respectively). This might suggest that the effects of AGR2 on adhesion of monolayer cells can occur at lower AGR2 levels than the effects observed on single cells (Fig. 4.4).

A possible explanantion for the strong difference in adhesion between EV and WT AGR2-expressing cells is that, at least in the context of WT AGR2 clone 2 and EV clone 2 cells, clonal WT AGR2-expressing cells produce much lower levels of the ECM protein fibronectin than EV clones (Fig. 4.8). As such, EV clone cells will have more matrix to adhere to, thus making them more resistant to trypsinisation. EV clones and WT AGR2 clones have a markedly different morphology (see Section 4.2.4 below) whereas EV and WT AGR2 pooled cells have a very similar morphology, and as the latter two cell lines showed no difference in the level of fibronectin production (Fig. 4.8), it may be that changes in fibronectin production (and hence resistance to trypsin) are related to cell morphology.

Overall, these experiments indicate that clonal WT AGR2 cells initially attach more rapidly and in a more trypsin-resistant manner than their AGR2-negative counterparts, but this trypsin resistance reduces over time until they become more easily detached once a monolayer is formed.

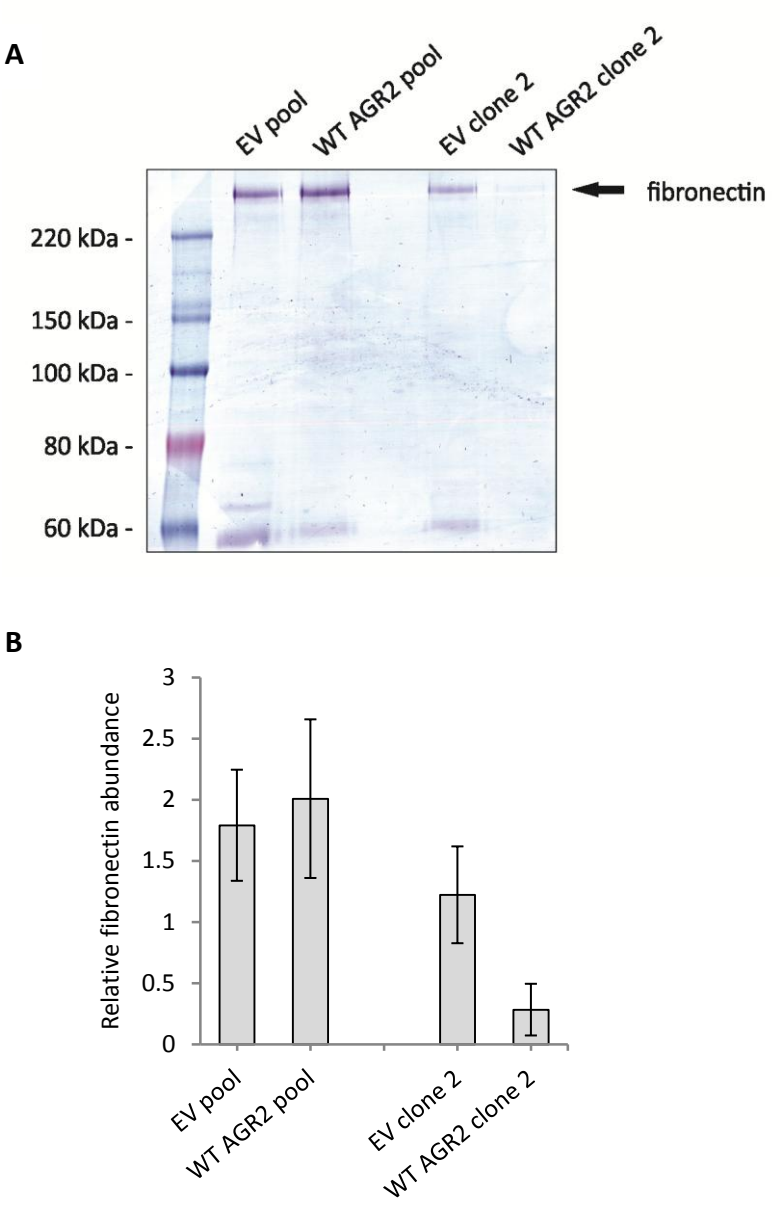


Figure 4.8. ***Comparison of secreted fibronectin from empty vector and WT AGR2-transfected cell lines.*** **(A)** Cells were grown to 70-80 % confluence, washed in PBS supplemented with $CaCl_2$ and $MgCl_2$ (see Materials and Methods) and incubated in serum- and phenol red-free Rama 37 medium for 4 h (cells incubated for longer begin to die). Protein was precipitated by TCA and resuspended in an equal volume solubilisation buffer. Protein concentration was measured by BCA assay and equal amounts of protein were loaded in each well. Samples were subjected to SDS-PAGE and stained with Coomassie brilliant blue. The indicated high molecular weight band was confirmed as fibronectin through peptide mass fingerprinting. **(B)** Fibronectin abundance was measured by densitometry and the mean relative abundance of fibronectin from 2 replicate experiments is shown, ±SD. There was no significant difference in fibronectin between pool cells (p = 0.739, Student's t test) or clonal cells (p = 0.133), although there was a noticeable difference in fibronectin secretion between clonal cell lines.

4.2.4 AGR2 expression correlates with an elongated cell morphology

As a cell line of epithelial origin, Rama 37 cells have a cuboidal morphology that is maintained after transfection with empty vector and clonal selection (Fig. 4.9A). However, some WT AGR2 clonal cells exhibit a more elongated morphology that is seemingly related to their level of AGR2 expression (Fig. 4.9B and C). This difference in morphology might explain both the striking differences in adhesion observed between EV and AGR2 clonal cell lines, and also the relationship between resistance to trypsin and the level of AGR2 expression observed in Fig. 4.4.

To quantify this morphological change, eight clonal cell lines from a randomly-picked panel were selected, all with differing levels of AGR2 expression, and analysed using ImageJ (see Materials and Methods). The degree of cell elongation (shape ratio, SR) was expressed as the ratio between the length of the longest and shortest cross sections of each cell in the x/y plane: the greater this value, the more elongated the morphology. This value was plotted against the level of AGR2 protein, measured by Western blot.

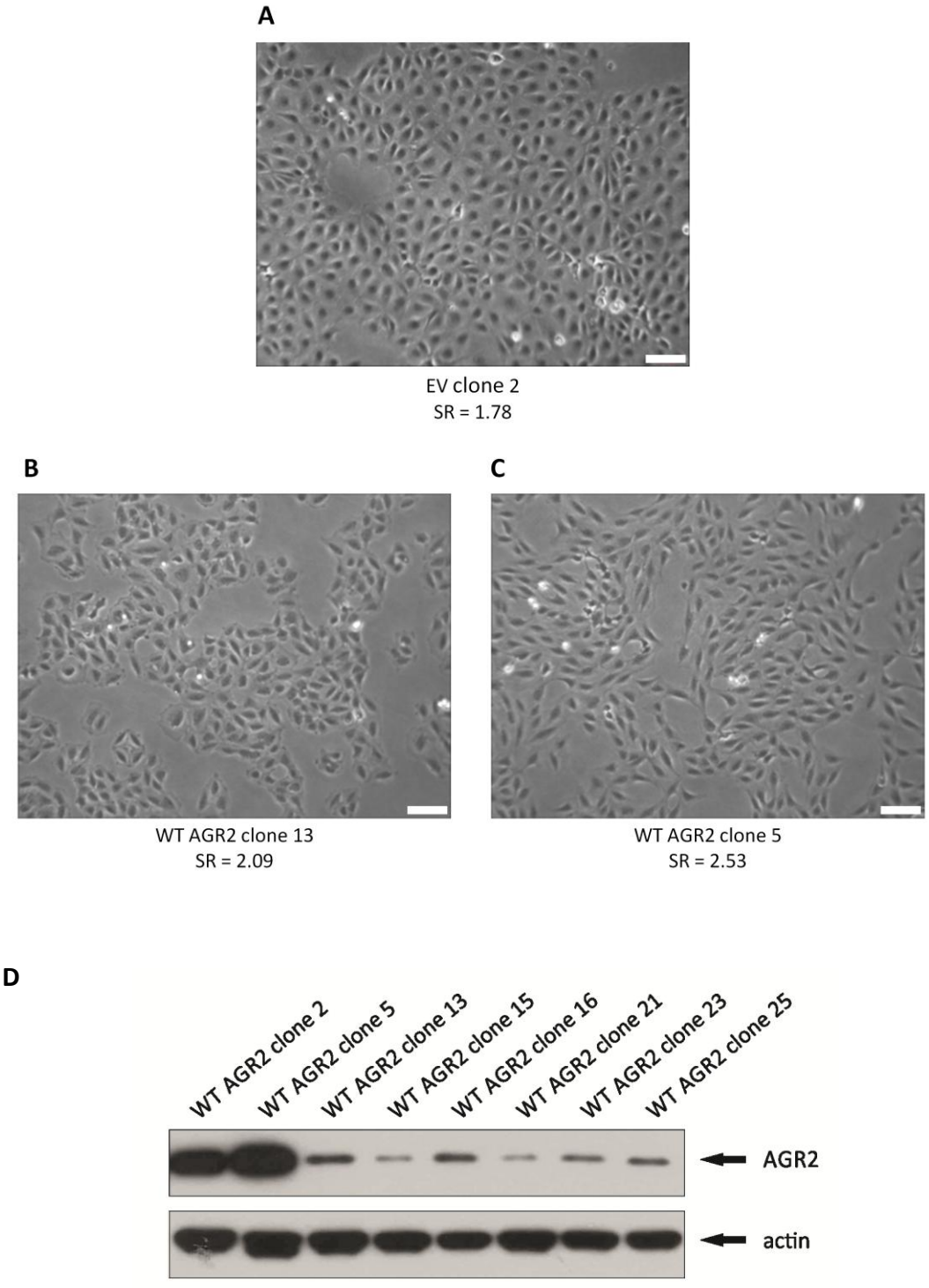


Figure 4.9 (continued overleaf). ***AGR2 expression correlates with cell morphology.*** Images representing different cell morphologies: **(A)** empty vector clone 2, cuboidal; **(B)** WT AGR2 clone 13, slightly elongated; and **(C)** WT AGR2 clone 5, very elongated. **(D)** Western blot of AGR2 expression levels across eight clonal cell lines. Actin was used as a loading control.

E

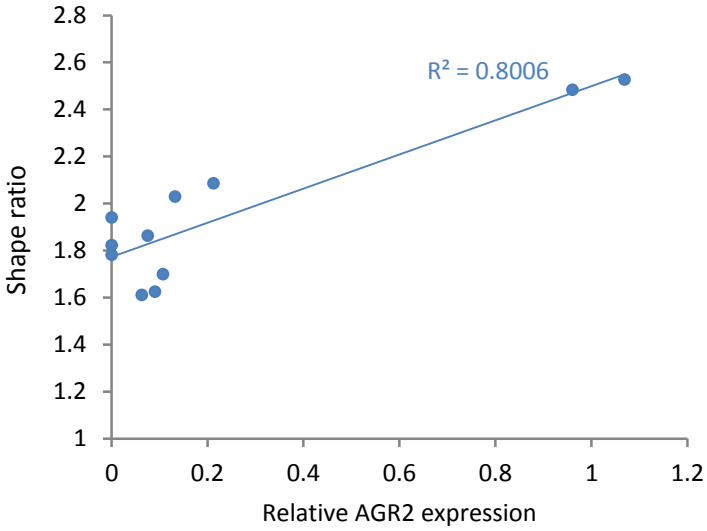


Figure 4.9 (cont.). ***AGR2 expression correlates with cell morphology.*** **(E)** Scatterplot of relative AGR2 expression and their relevant morphologies, expressed as a shape ratio (see main text), in eight WT AGR2-expressing clonal cell lines and three non-AGR2 expressing (EV) clonal cell lines. AGR2 levels were normalised to actin levels for each clone, and then this normalised AGR2 expression was used to plot expression relative to WT AGR2 clone 2. At least 200 cells were analysed for each clone and the average SR value from these is plotted.

It is clear from the cell images that there is a change in morphology associated with AGR2-expressing cells (Fig. 4.9A-C). This difference in shape ratio was statistically significant between EV clone 2 and both WT clone 5 and 13 cells ($p < 0.001$ in both cases, Student's t-test), as well as between EV clone 2 and WT clone 2 ($p < 0.001$). There was no significant difference between WT clone 2 and 5 ($p = 0.652$), although these were both significantly different to WT clone 13 ($p < 0.001$ in both cases). This association of AGR2 expression with SR is further illustrated by the positive linear correlation observed between these two factors in Fig. 4.9E (Spearman's $\sigma = 0.633$, $p = 0.037$).

4.2.5 Cell morphology contributes to cell adhesion properties

In order to find out whether cell shape ratio was related to the observed changes in cell adhesion, a subpopulation of elongated cells was presently isolated from cuboidal Rama 37 cells and expanded (see Materials and Methods). These AGR2-negative cells

(Appendix 13), termed Rama 37-elongated (R37E), were tested in adhesion assays as above, offering a comparison between elongated, AGR2-positive cells (WT AGR2 clone 2) and elongated, AGR2-negative cells (R37E).

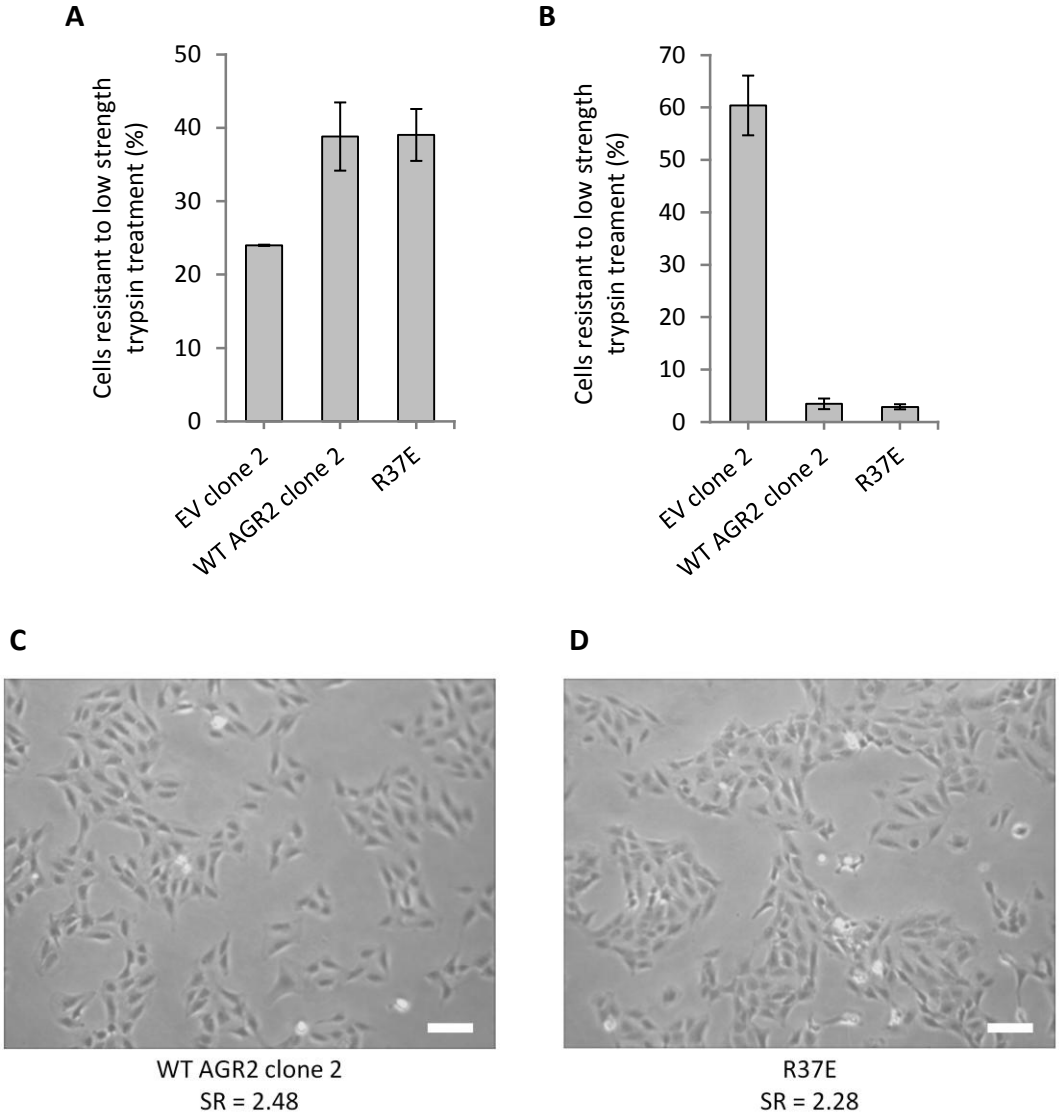


Figure 4.10. ***Comparison of adhesion properties of elongated AGR2-positive and –negative cells.*** **(A and B)** Wild type (WT) AGR2-expressing cells and R37E cells were allowed to attach for 30 min **(A)** or allowed to form a monolayer overnight **(B)** and then trypsinised in a 0.0125 % (v/v) trypsin for 5 min. The remaining adherent cells were counted with an automated cell counter. The percentage of cells resistant to low strength trypsin treatment represents the proportion of cells attaching at the indicated time points that remained adherent after dilute trypsin treatment. The means from duplicate wells from 3 independent experiments are shown, ±SE. Images and SR values of **(C)** WT AGR2 clone 2 cells and **(D)** R37E cells. SR values were calculated for at least 200 cells for each cell type, and the average SR value is shown. SR: shape ratio. Scale bar: 100 µm

The sensitivities of R37E and WT AGR2 clone 2 cells to 0.0125 % (v/v) trypsin were virtually identical. In the short term (Fig. 4.10A), WT AGR2 clone 2 and R37E cells were both more resistant to 0.0125 % (v/v) trypsin than EV clone 2 cells on average, although this difference was just outside statistical significance (p = 0.086 and 0.051, respectively, Student's t-test). There was no significant difference to resistance of 0.0125 % (v/v) trypsin between WT AGR2 clone 2 and R37E cells, however (p = 0.972). Similarly, when grown to 70-80 % confluence and subjected to the detachment assay (Fig. 4.10B), WT AGR2 clone 2 and R37E cells were both highly significantly less resistant to 0.0125 % (v/v) trypsin than EV clone 2 cells (p = 0.008 and 0.009, respectively), but not significantly different from each other (p = 0.643). These data confirm the relationship between adhesive behaviour and cell morphology, given the similarity in morphologies and SR values of the two cell lines (Fig. 4.10C and D). Together, these data point to an AGR2-associated change in cell morphology as being a major contributing factor determining cell adhesion properties in these cells.

4.2.6 Effects of AGR2 mutations on cell morphology and adhesion

The establishment of several stable cell lines expressing mutant forms of AGR2 (see Chapter 3) were used as a tool to investigate the contribution of key regions of AGR2 protein to AGR2-induced in cell adhesion. The effects of mutating the thioredoxin domain (C81S AGR2), preventing dimerisation (E60A AGR2), inhibiting ER entry (Δ1-20 AGR2), deleting the flexible region (Δ21-40 AGR2) and promoting AGR2 secretion (ΔKTEL AGR2) were explored. As protein expression levels in mutant AGR2 pooled cell lines were largely lower than those observed in WT AGR2-expressing cells (see Chapter 3), only mutant clonal cell lines were assayed as they displayed a wider range of expression levels.

The resistance to trypsin of three separate clonal cell lines was assayed for each mutation, 30 min after plating (Fig. 4.11). Due to the large number of cells involved in these assays, cells were divided into three groups and three replicate experiments were performed for one group before passing on to the next. EV clone 2 and WT clone 2 cells were included in each set of assays as controls for the attachment behaviour of cuboidal and elongated cells, respectively. However, whilst the data between triplicate experiments were reproducible (as judged by the relatively tight standard error bars), there was marked variation in the behaviour of EV clone and WT clone 2 cells between sets of experiments (i.e. experiments A, B and C in Fig. 4.11). This was particularly true of EV clone 2 cells, where the percentage of cells resistant to trypsin varied from below 10 % in Fig. 4.11A to around

25 % in Fig. 4.11B and C. WT clone 2 cells behaved more consistently, with the proportion of trypsin-resistant cells varying from only 55 % to 45 %. Thus, the data from these individual experiments are shown separately.

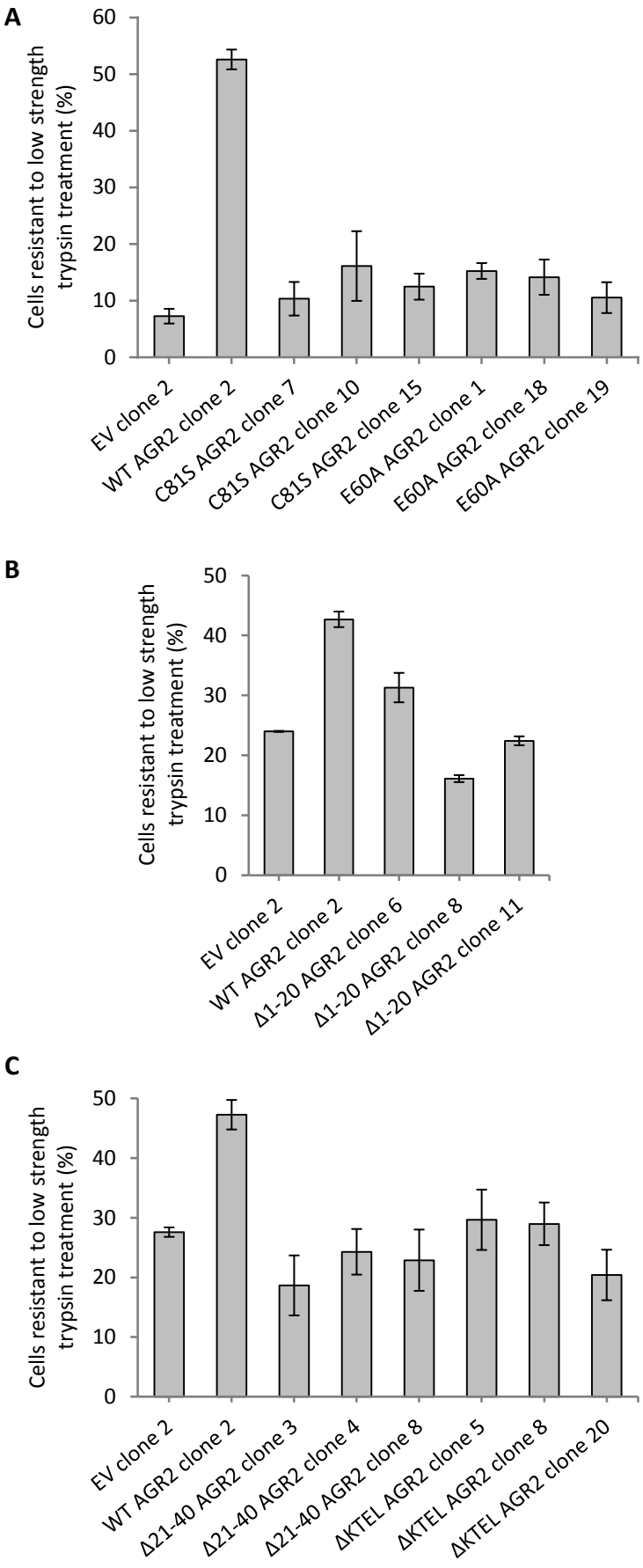


Figure 4.11. ***Initial attachment properties of mutant-AGR2-expressing clonal cell lines.*** The indicated cells were allowed to adhere for 30 min to cell culture plates. Cells were then trypsinised in 0.0125 % (v/v) trypsin for 5 min and the remaining adhering cells were counted with an automated cell counter. The percentage of adhering cells represents the proportion of cells adhering after 30 min that remained adhering after weak trypsin treatment. **(A)** E60A AGR2 and C81S AGR2 clones. **(B)** Δ1-20 AGR2. **(C)** Δ21-40 AGR2 and ΔKTEL AGR2 clones. The means from duplicate wells from 3 independent experiments are shown, ±SE.

Despite this variability in control cells, the data showed that mutant cell lines behaved on the whole in a more similar way to EV clone 2 cells than WT clone 2 cells. However, these cell lines express a range of steady-state AGR2 protein levels (Chapter 3), and so to separate out any potential effects caused by differences in AGR2 levels within these cells, cell lines with similar steady-state AGR2 levels were considered for further analysis (Fig. 4.12A and B). In order to compare cell lines across different experiments, the number of trypsin-resistant cells were normalised to the levels observed with WT clone 2 cells in each set of experiments. This was reasoned to provide the most appropriate normalisation factor, given the high variability of EV clone 2 cells, and allowed WT AGR2 clone 13 cells to be directly compared to cells expressing mutant AGR2 proteins, due to similar levels of AGR2 protein expression.

Considering the trypsin sensitivity of these cell lines in the context of similar AGR2 levels (Fig. 4.12C), both E60A AGR2 clone 18 and C81S AGR2 clone 10 cells showed significantly lower attachment compared to WT AGR2 clone 13 (p = 0.006 and p = 0.037, respectively, Student's t-test). Conversely, attachment levels of Δ1-20, and ΔKTEL AGR2 cells were not significantly different from WT AGR2 clone 13 cells (p = 0.827, and 0.140, respectively), whilst Δ21-40 exhibited a borderline significant reduction, p = 0.061). Comparison with EV clone 2 attachment levels is difficult, given the differences in attachment levels between experiments. Due to this fact, these data must be interpreted with some caution, but point to dimerisation and the thiol group of Cys81 as being important in mediating this resistance to trypsin, at least during these initial stages of cell attachment.

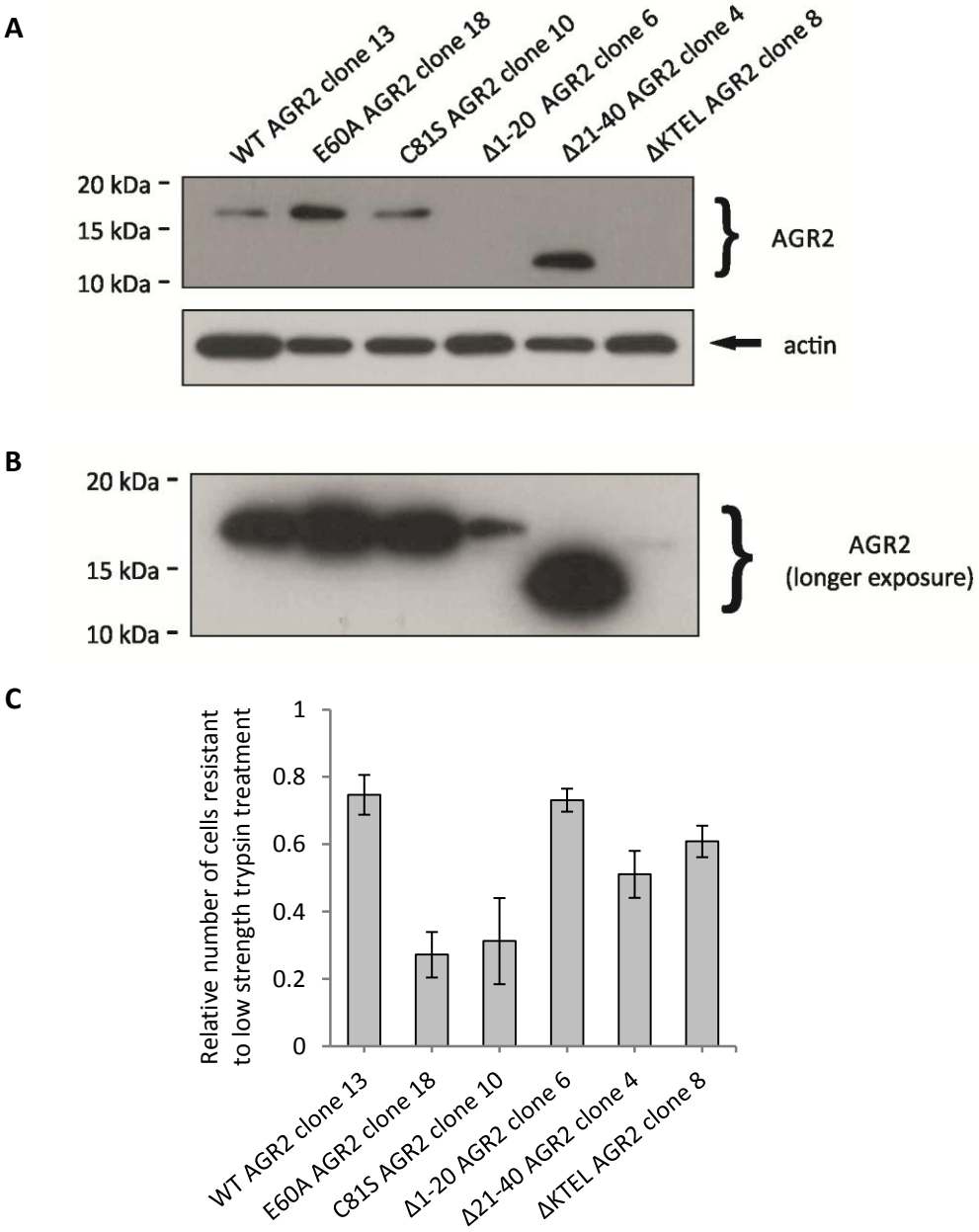


Figure 4.12. ***Initial attachment behaviour of mutant-AGR2-expressing clonal cell lines with similar steady-state protein levels.*** **(A)** Western blot showing AGR2 protein levels in the most comparable clonal cell line of each listed mutation. Actin was used as a loading control **(B)** Longer exposure showing expression levels of the two weakest expressing cell lines. Note the difference in size of Δ21-40 AGR2 and ΔKTEL AGR2. For quantitation of protein abundance, see Appendix 14. **(C)** The indicated cells were allowed to adhere for 30 min to cell culture plates. Cells were then trypsinised in 0.0125 % (v/v) trypsin for 5 min and the remaining adhering cells were counted with an automated cell counter. The proportion of trypsin-resistant cells is expressed relative to the number of trypsin-resistant WT clone 2 cells in each experiment shown in Fig. 4.10. The means from duplicate wells from 3 independent experiments are shown, ±SE.

To further explore the effects of AGR2 mutations on trypsin sensitivity, mutant cell lines were also assayed for their resistance to trypsin once allowed to form a monolayer (Fig. 4.13A-C). The attachment levels of EV clone 2 and WT clone 2 cells in this assay were much more consistent than seen in the short term assay above. To filter out any effects of differences in steady-state AGR2 levels, the trypsin sensitivity of cells with similar steady-state AGR2 levels were compared (Fig. 4.13D). The data here were normalised to the attachment levels of EV clone 2 cells as they were slightly more consistent across experiments than WT clone 2 attachment levels, although normalising to WT clone 2 gave a near identical relative attachment pattern (Appendix 15).

Comparing these cell lines, only Δ21-40 AGR2 clone was not significantly different to WT AGR2 clone 13 (p = 0.095, Student's t-test). Comparisons of mutants with EV clone 2 must be done in the context of each individual set of experiments, given that EV clone 2 was used as a normalisation factor to allow comparison to WT AGR2 clone 13. In this way, only Δ21-40 AGR2 clone 4 and ΔKTEL AGR2 clone 8 were significantly different to EV clone 2 (p = 0.006 and 0.004, respectively).

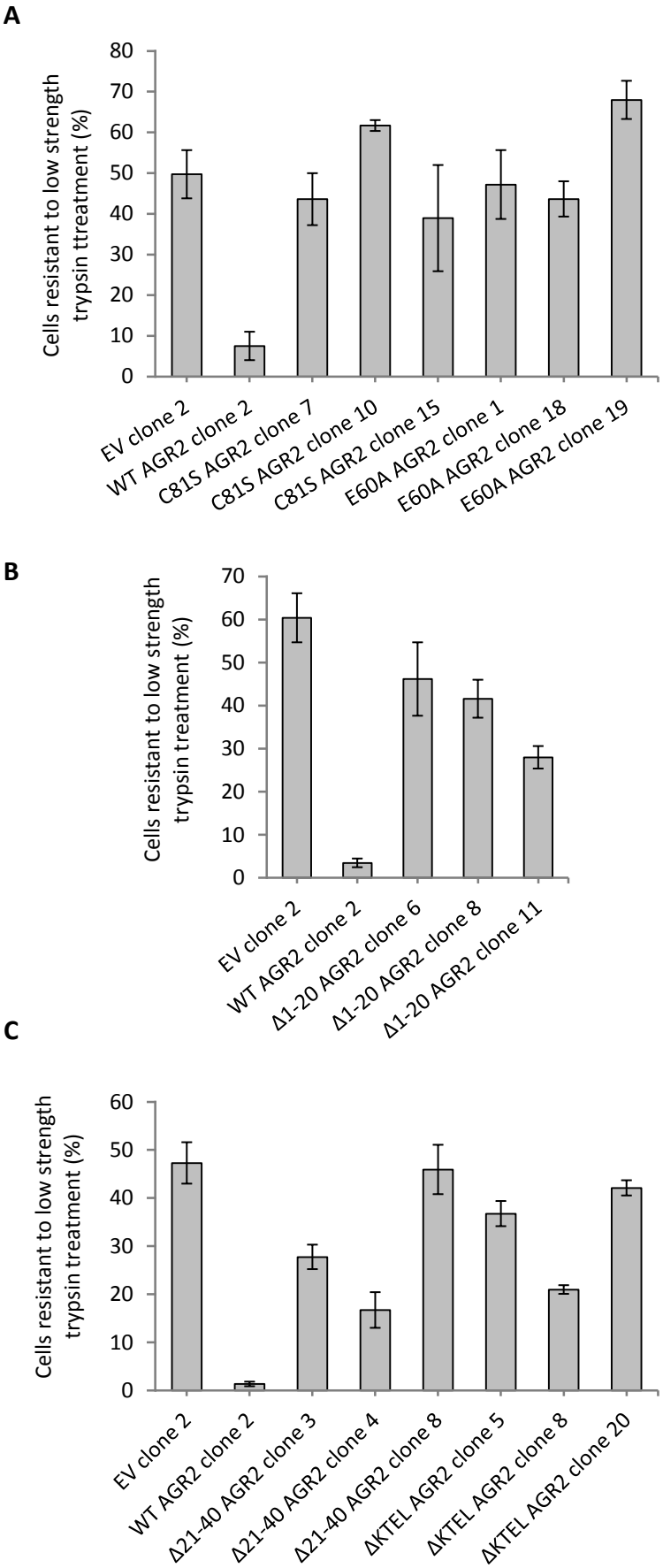


Figure 4.13 (continued overleaf). ***Monolayer attachment properties of mutant-AGR2-expressing clonal cell lines.*** The indicated cells were grown overnight to 70-80 % confluence and then trypsinised in 0.0125 % (v/v) trypsin for 5 min. The remaining adhering cells were counted with an automated cell counter. The number of trypsin-resistant cells is expressed as a percentage of the total number of cells in each well. **(A)** C81S AGR2 and E60A AGR2 clones. **(B)** Δ1-20 AGR2. **(C)** Δ21-40 AGR2 and ΔKTEL AGR2 clones. The means from duplicate wells from 3 independent experiments are shown, ±SE.

D

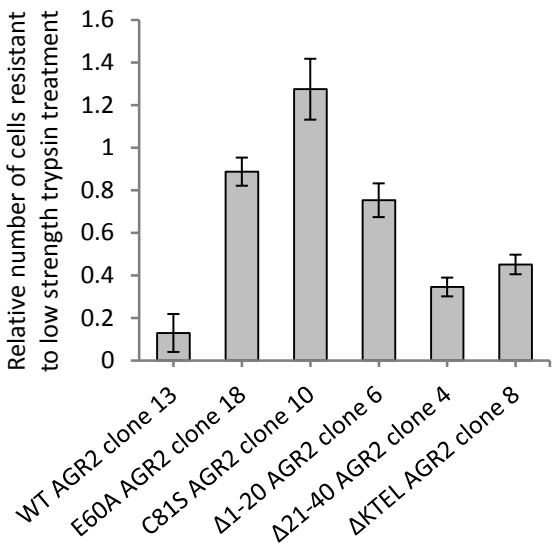


Figure 4.13 (cont.). ***Monolayer attachment properties of mutant-AGR2- expressing clonal cell lines.*** **(D)** The number of trypsin-resistant cells is expressed as a percentage of the total number of cells in each well. The proportion of trypsin-resistant cells is expressed relative to the number of trypsin-resistant EV clone 2 cells in each of (A), (B) and (C). The means from duplicate wells from 3 independent experiments are shown, ±SE.

Given the association of elongated morphology with trypsin sensitivity (Fig. 4.10), the correlation between morphology (measured by shape ratio, SR) and the relative trypsin sensitivity of all clonal cell lines was investigated (Fig. 4.14). There was a statistically significant correlation between SR and relative trypsin sensitivity in both the initial attachment (Spearman's $\sigma = 0.478$, $p = 0.028$) and cell monolayer assays (Spearman's $\sigma = -0.631$, $p = 0.002$). This correlation was not conserved when considering only the cell lines with similar state-steady protein levels in the initial attachment assay (Spearman's $\sigma = 0.486$, $p = 0.329$) but the correlation was even more pronounced for this sub-set of cells in the monolayer assay (Spearman's $\sigma = -0.943$, $p = 0.005$). These correlations suggest that the main factor affecting the resistance to trypsin for all of these cells in the monolayer adhesion assay is their morphology, but not in the cell suspension assay. Therefore, the use of changes in shape ratio were explored as a means of determining the effects of mutation on AGR2 function in the monolayer assay.

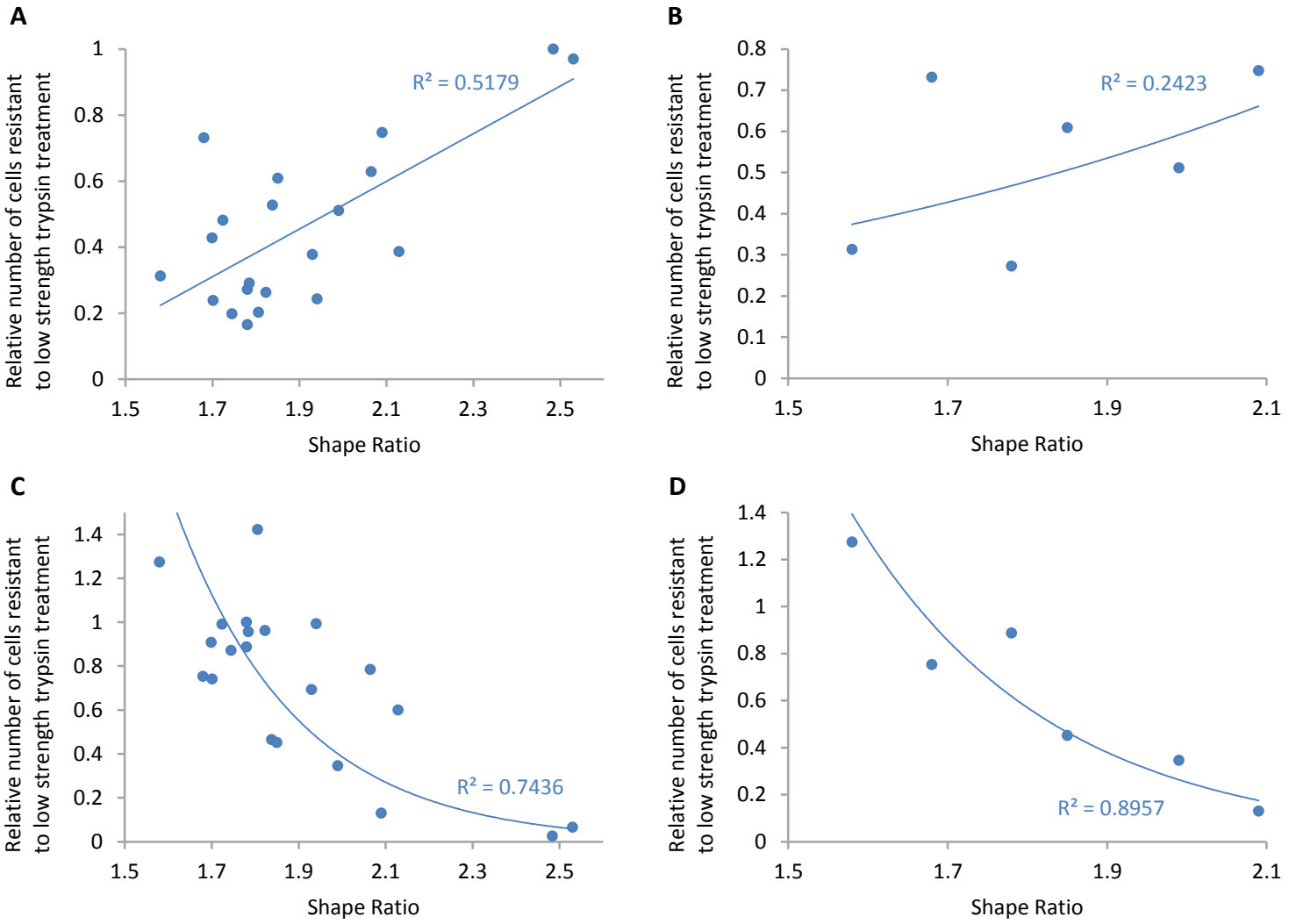


Figure 4.14. ***Correlation between cell morphology and attachment behaviour for AGR2-expressing clonal cells lines.*** Correlation was measured between **(A)** all AGR2-expressing clonal cell lines and initial attachment, **(B)** clonal cell lines of similar steady-state protein levels (shown in Fig. 4.12) and initial attachment, **(C)** all AGR2-expressing clonal cell lines and monolayer attachment, **(D)** clonal cell lines of similar steady-state protein levels (shown in Fig. 4.13D) and monolayer attachment. Cell morphology was measured by shape ratio (SR, see Materials and Methods) for at least 150 cells for each cell line and the mean SR value was plotted against the mean relative resistance to trypsin as measured in Fig. 4.11 for A and B, and as in Fig. 4.13D for C and D.

The SR values from the three EV clones used in these experiments were combined, to avoid any bias given the slight variation in their SR values (Appendix 16). WT AGR2 clone 13 SR was significantly higher than the combined SR of the three EV clones, but only Δ21-40 AGR2 clone 4 had an SR value that was also significantly higher than EV clones (Table 4.1). Conversely, C81S AGR2 clone 10 and Δ1-20 AGR2 clone 6 had significantly lower SR values than EV clones, probably highlighting the importance of the thiol group and ER residence in particular in the function of AGR2, respectively. All mutant clones had an SR value that was significantly lower than WT AGR2 clone 13.

Cell line	Mean SR	Statistical difference compared to EV clones (*p* value)	Statistical difference compared to WT clone 13 (*p* value)
EV clones	1.85	-	< 0.001
WT AGR2 clone 13	2.09	< 0.001	-
E60A AGR2 clone 18	1.78	0.506	< 0.001
C81S AGR2 clone 10	1.58	< 0.001	< 0.001
Δ1-20 AGR2 clone 6	1.68	0.001	< 0.001
Δ21-40 AGR2 clone 4	1.99	0.029	0.021
ΔKTEL AGR2 clone 8	1.85	0.698	< 0.001

Table 4.1. ***Statistical analysis of differences in shape ratio (SR) values across mutant-AGR2-expressing clonal cell lines with similar steady-state protein levels.*** The SR values of at least 200 cells for each cell line were measured and the mean SR values are shown. The mean SR value across all three EV clones was used (see text). See Appendix 17 for representative cell morphology images and corresponding SR values.

Overall, these results suggest that dimerisation, the Cys81 thiol group, and protein localisation are important for AGR2-associated changes to cell morphology, as perturbation of any of these features resulted in a complete loss of the elongated cell type associated with WT AGR2 expression. The 21-40 region on the other hand is less important, given that its SR ratio was significantly different to both EV clones and WT13, possibly pointing to a less efficient AGR2-driven process. It is important to note however that it has not been expressly demonstrated that AGR2 is the cause of cell elongation. Several attempts were made to knockdown AGR2 expression in WT AGR2 clonal cells to address this notion, all of which were ultimately unsuccessful (Appendix 18).

4.2.7 Mutation of AGR2 completely ablates anchorage-independent cell growth

Aside from changes to the adhesion properties of tumour cells, anchorage-independent growth is a necessary requirement for metastasis, and *in vitro* assays of anchorage-independent growth give clues to the metastatic potential of cells *in vivo* (reviewed in [583]). AGR2 is known to play a role in this process, as knockdown of AGR2 severely reduced anchorage-independent growth of oesophageal cancer SEG-1 cells [481], as well as several breast cancer cell lines [482]. Expression of AGR2 also induces anchorage-independent growth in NIH3T3 and lung cancer H1299 cells [445, 481].

To this end, the ability of Rama 37 cells expressing WT and mutant AGR2 were assayed for anchorage-independent growth by the use of a soft agar assay. Only the three WT AGR2-expressing cell lines exhibited anchorage-independent cell growth, with all EV and AGR2 mutant clonal cell lines unable to grow in the soft agar (Fig. 4.15). The growth of R37E cells was also monitored as a control for elongated cell morphology. Representative images and quantitation of colony formation for cell lines expressing similar steady-state levels of AGR2 are shown in Fig. 4.16, along with EV clone 2 and R37E. Importantly, R37E cells were also able to form colonies, and there were on average approximately 2.5 times more viable, colony-forming R37E cells than WT AGR2 clone 13 cells (Fig. 4.16I), suggesting that these cells grow more efficiently in soft agar than AGR2-expressing cells with a similar morphology. Overall, these data further highlight the importance of cell morphology in bestowing pro-metastatic characteristics on Rama 37 cells, with only the most elongated cell types used here able to grow in soft agar, including those expressing AGR2. Whether continued expression of AGR2 creates some sort of survival disadvantage (as hinted at by the apparently more efficient growth of AGR2-negative, but elongated, R37E cells) will require further investigation.

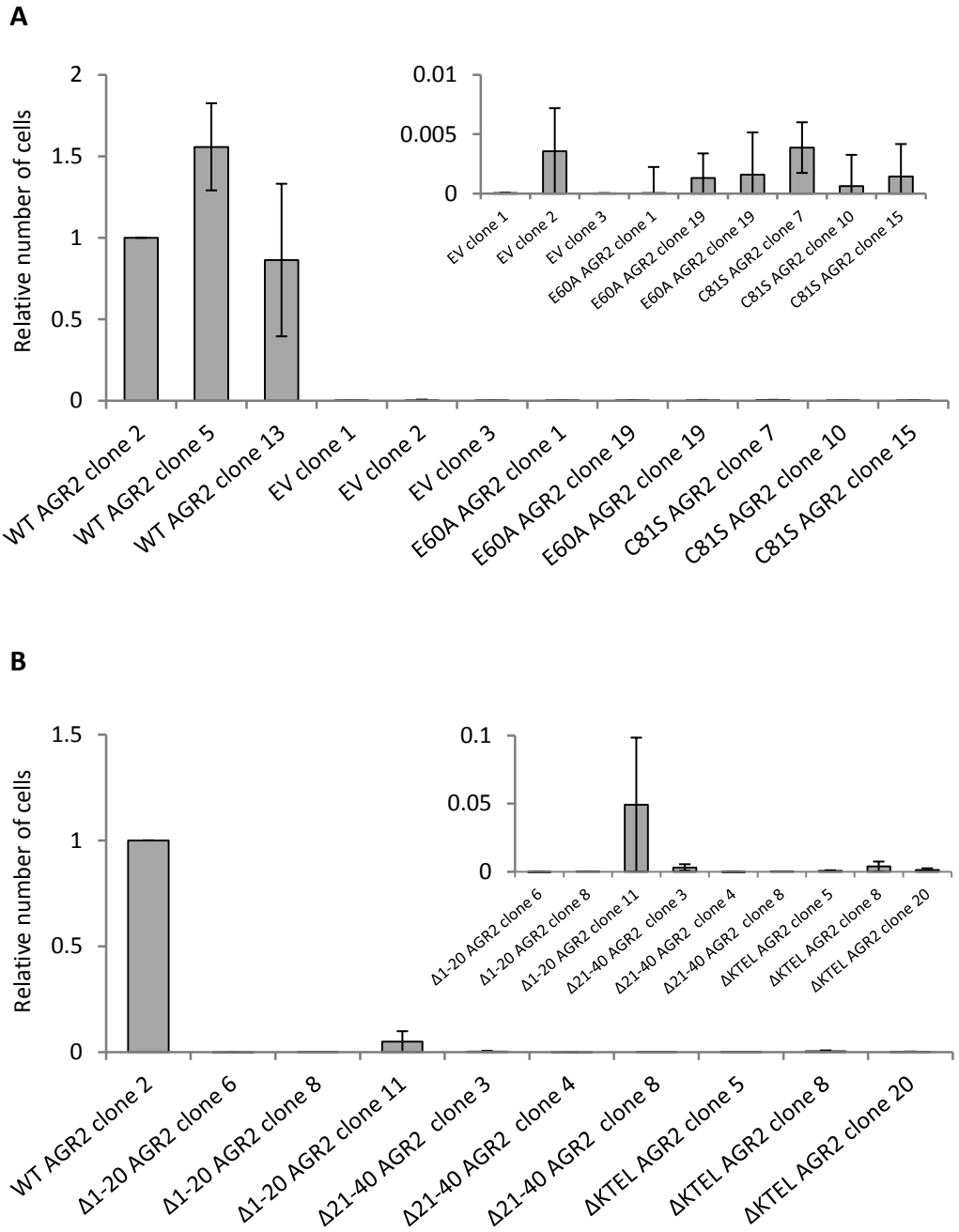


Figure 4.15. ***Anchorage-independent growth assay of highest expressing clonal cell lines.*** The indicated cells were plated into 0.3 % agarose and left to form colonies for 4 weeks. After that time, the number of viable cells was measured using bioluminescence (see Materials and Methods). Cells were plated in triplicate and the means of the luminescence relative to WT AGR2 clone 2 from three independent experiments are shown, ±SE. **(A)** Empty vector, WT AGR2, E60A AGR2 and C81S AGR2 clones. **(B)** Δ1-20 AGR2, Δ21-40 AGR2 and ΔKTEL AGR2 clones. Insets: relative cell numbers for non-WT AGR2 cells using smaller scale.

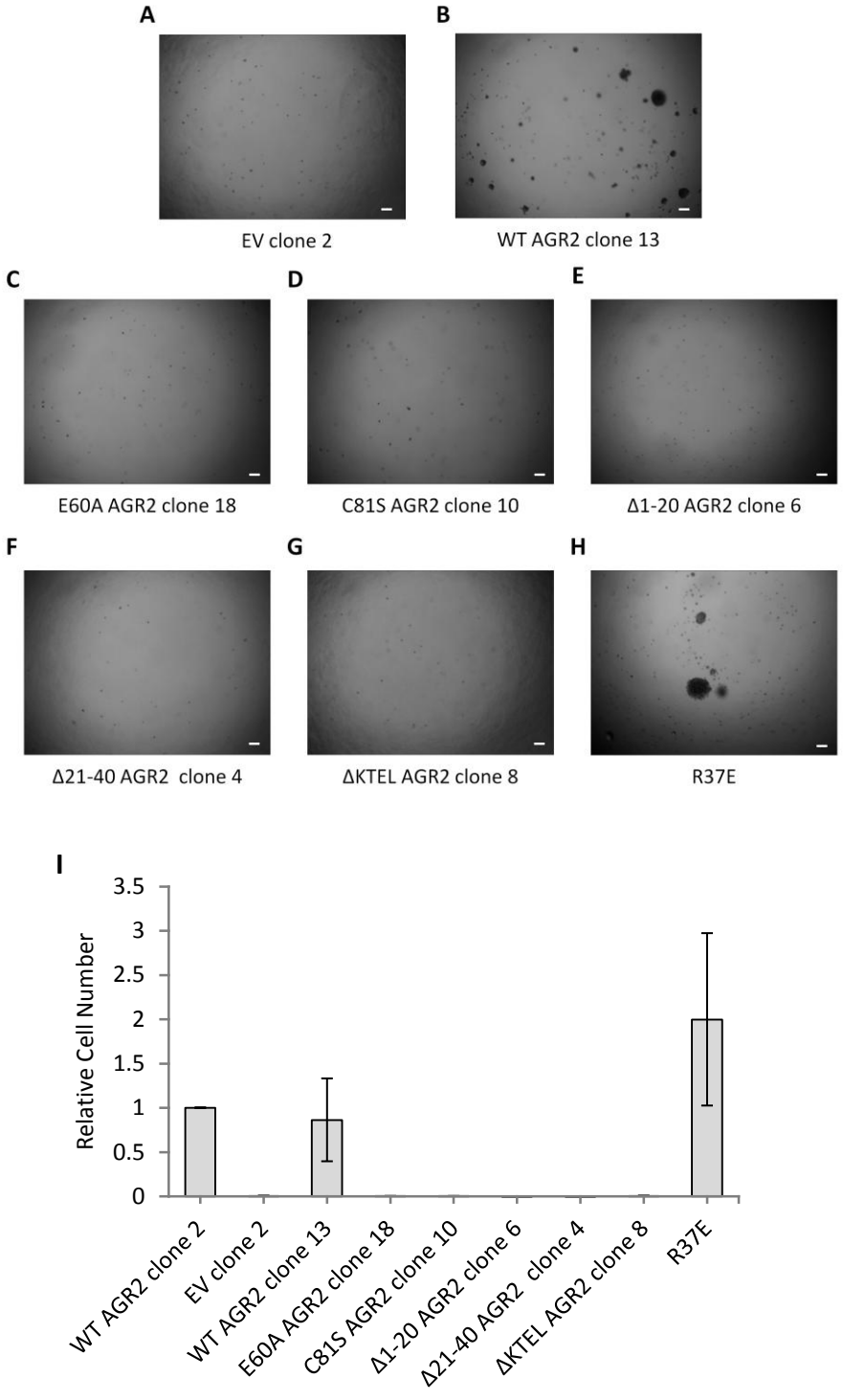


Figure 4.16. ***Anchorage-independent growth of Rama 37 cells.*** The indicated cells were plated into 0.3 % agarose and left to form colonies for 4 weeks. Representative images of colony formation in wells containing **(A)** EV clone 2, **(B)** WT AGR2 clone 13, **(C)** E60A AGR2 clone 18, **(D)** C81S AGR2 clone 10, **(E)** Δ1-20 AGR2 clone 6, **(F)** Δ21-40 AGR2 clone 4, **(G)** ΔKTEL AGR2 clone 8 and (**H**) R37E. The number of viable cells was quantified (**I**) using bioluminescence (see Materials and Methods). Cells were plated in triplicate and the means of the luminescence relative to WT AGR2 clone 2 from three independent experiments are shown, ±SE. Scale bar: 100 µm.

4.2.8 Clonal WT AGR2-expressing cells are highly migratory

Having established that clonal AGR2-expressing cells display differences in adhesion and anchorage-independent growth relative to EV clonal cells, the migratory capacity of these cells were compared. Over a 16 h period, WT clone 2 cells migrated to a much greater extent than EV clone 2 cells, with an average of 140 times more WT AGR2 clone 2 cells migrating across the Transwell insert than EV clone 2 cells (Fig. 4.17). These data highlight that clonal WT AGR2-expressing cells possess several characteristics that are advantageous in the context of cancer metastasis, notably changes to cell adhesion, increased migration potential and anchorage-independent cell growth.

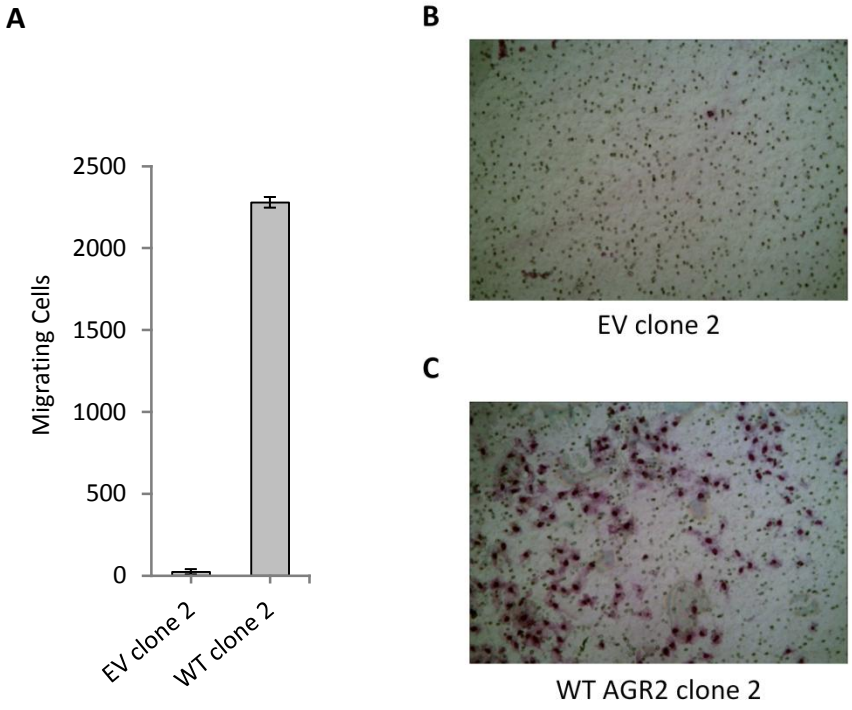


Figure 4.17. ***Clonal WT AGR2 cells are highly migratory in the Boyden Chamber assay.*** **(A)** Cells were plated into the top of Transwell inserts in 1 % FBS medium and allowed to migrate for 16 h towards the lower compartment containing 5 % FBS. Cells were then fixed and stained, and all cells from each membrane were counted under a microscope. Data shown is from duplicate wells from a single experiment, ±SD. Representative fields of view from Transwell membranes for **(B)** EV clone 2 cells and **(C)** WT AGR2 clone 2 cells following 16 h migration assay. Data collected in collaboration with Kathryn McCall, Institute of Integrative Biology, University of Liverpool.

4.3 Discussion

4.3.1 AGR2-cell morphology correlation

It has been demonstrated here that there exists a positive correlation between WT AGR2 expression and elongated cell morphology, and AGR2 has indeed been previously linked to elongated morphology in prostate cancer PC3 cells [515]. If AGR2 were the cause of this morphology change, and that this would presumably be through induction of an EMT-type transformation, this could provide an explanation for the association of AGR2 with metastasis [301, 305, 307, 309, 440, 481]. The fact that the elongated phenotype is only observed in clonal but not pooled WT AGR2 cells may be the result of some as yet unidentified positive feedback mechanism. The EGFR ligand amphiregulin (AREG) might play a role here, given its upregulation by AGR2 and its promotion of metastasis and invasion through an autocrine loop [506, 584, 585]. It may be that, in the context of a heterogeneous population of cells, signals from other, more cuboidal cells (with lower AGR2 expression) act to balance out pro-EMT signals such as amphiregulin, maintaining the majority of the population in a more cuboidal morphology. When a clonal population is created from a highly expressing cell, there may be a loss of this 'off' switch, pushing the cell towards a permanent elongated morphology.

An alternate explanation for the differences between pooled and clonal cells relates to the MET process. There is evidence to suggest that metastatic cells must revert back to epithelial–type cells in order to allow outgrowth of a secondary tumour, stemming from observations that secondary metastases from a number of adenocarcinomas (including breast, ovarian and prostate) seem to be made up of cells that are more differentiated than their tumour of origin (reviewed in [264]). The reasons for this MET change are not clear, but suggest that there is some disadvantage to the mesenchymal phenotype in terms of maintaining a population of cells, rather than individual metastatic cells. It is interesting that the clonal WT AGR2 cells in the present experiments grow at a slower rate than pooled cells in culture (personal observation) and clonal and pooled AGR2-expressing cells yield tumours that are slower growing *in vivo* than EV-transfected controls [309], hinting at a growth disadvantage for these cells in mixed populations. It may be then that, due to this growth disadvantage, elongated WT AGR2 cells, although present at a low level within the pooled population (personal observation), are unable to establish themselves as the dominant cell type within this pooled WT AGR2 population. Therefore,

since much of the adhesion behaviour of these AGR2-expressing cells can be explained by morphology, the cuboidal morphology of the majority of the WT AGR2 pooled cells may mask any change induced by AGR2. Indeed, WT AGR2 pool 2 cells and EV pool cells show no difference in overall morphology, with SR values of 1.88 and 1.89, respectively. It is probably for these reasons that no differences in adhesion are observed between WT AGR2 and EV pool cells.

In line with this, and in order to better understand the role of AGR2 with regard to cell morphology, several attempts were made to knockdown AGR2 expression in both pooled and clonal cells to see whether their morphology would revert someway back towards cuboidal, as would be expected through an AGR2-dependent MET event. Unfortunately, despite the use of two different transposase vectors designed to remove the gene cassette containing AGR2 that was originally transfected into the cells, and the use of four shRNA plasmids individually, no knockdown could be achieved (Appendix 18).

The reason for this failure is not clear, but the lack of activity from the transposase could be attributed to loss of enzyme recognition sites during plasmid integration, or some epigenetic change following transfection that now occludes the integration site. The shRNA plasmids used in knockdown experiments all target regions within the coding sequence of human AGR2 (Appendix 20), and so should not be affected by the fact that human AGR2 has been introduced into rat cells. This could be tested using a cell line expressing endogenous AGR2, e.g. MCF7A, to ensure these are indeed functional shRNA sequences. A more likely explanation for the lack of knockdown by shRNA plasmids relates to the different promoters driving expression of AGR2 and AGR2-shRNA. In the transfection experiments above, the AGR2 gene is transcribed from a highly active EF1α promoter and so the rate of AGR2 mRNA transcription and the abundance of AGR2 mRNA transcript may be too high to be efficiently knocked down by the relatively weaker U1 promoter-driven shRNA [551]. Therefore, one possible way to achieve substantial knockdown of AGR2 may be to subclone the shRNA sequences into a different vector, under the control of a stronger RNA polymerase III promoter such as CMV [586, 587], rather than the U1 RNA polymerase II promoter used here.

4.3.2 Adhesion properties of AGR2-expressing cells

To assess the effects of AGR2 on cell adhesion, a trypsin-resistance assay was used in the experiments presented here as a measure of the strength of cell attachment. Whilst

the assay does not directly measure forces required to overcome the strength of cell adhesion in the same way as the centrifugation assay, its use in other studies [588, 589] and the similar overall results seen between the trypsin and centrifugation assay here support its use as a proxy for strength of cell adhesion. Furthermore, in the context of Rama 37 cells at least, the centrifugation assay is only useful over short time periods (up to approximately 30 min after plating cells), at which point, cells could no longer be detached by centrifugation (personal observation). The trypsin resistance assay therefore permits the study of cell adhesion where stronger adhesive forces are involved, such as those of cells in monolayers. Using this assay, AGR2 expression was shown to initially enhance cell attachment immediately after plating of cells. In this short timeframe, the initial stages of cell adhesion are thought to be mediated through electrostatic interactions between the cell surface and substratum, mainly through the negatively-charged glycosaminglycan (GAG) hyaluronan (HA) [590, 591]. The presence of HA in the pericellular matrix is required for this early interaction, solely mediating the adhesion of chondrocytes up to 10 min after plating, but HA also promotes the formation of FAs in the longer term, leading to an overall more rapid HA-mediated anchoring of the cell [592]. After these initial adhesion events, HA is sequestered from the cell-substratum interface, and focal adhesions (FAs) then become the major determinant of adhesion strength [592]. Given the reduction in the number of FAs in clonal WT AGR2 cells (Appendix 21), these would be expected to adhere less strongly than EV cells after the initial HA-mediated adhesion phase, which was indeed the case. Thus while the time scale of progression from HA- to FA-mediated adhesion might be longer than observed for chondrocytes [592], the overall pattern observed here for Rama 37 cells is very similar. It is noteworthy that chondrocytes contain a thick pericellular coat with probably multiple layers of HA [591], and thus may adhere more rapidly than Rama 37 cells, providing a possible explanation for the difference in time scale for the switch between HA- to FA-mediated adhesion.

It will be interesting to determine whether HA does play a major role in the adhesion of clonal WT AGR2 cells, as metastatic cells have been shown to produce high levels of HA and that this HA mediates the adhesion of these cells to the endothelium, in a similar manner to immune cell recruitment [593-595]. Whilst the ultimate arrest of circulating tumour cells in a particular vessel may be due to the size restriction of the vessels that these cells arrive at [199-201], it is clear that increased cellular adhesion is a favourable characteristic for promoting this metastatic process, as chemical inhibition of HA synthesis or blocking HA activity through the use of a synthetic peptide reduced the metastatic potential of melanoma and osteosarcoma cells [596, 597]. Thus, AGR2-

associated increases in the adhesion of single cells or clumps of cells may partially explain its pro-metastatic activity, by promoting adhesion of these cells at secondary sites.

The fact that the strength of attachment of clonal WT AGR2 cells reverses over time is probably also related to the switch from HA- to FA-mediated adhesion. Additionally, whist EV clone 2 cells (and indeed other cuboidal Rama 37 cell lines) secrete a large amount of fibronectin, WT AGR2 clone 2 cells produce comparatively much lower amounts (Fig. 4.8). As such, this fibronectin-rich ECM deposited by EV cells will provide increased attachment for these cells over time [598, 599], compared to WT AGR2 clone 2 cells, where there will be little such ECM deposition. This could contribute to the switch from strong adhesion to weak adhesion over time for WT AGR2 cells, and vice versa for EV cells, as adhesion switches from HA- to FA- and ECM-mediated.

4.3.3 Roles of extracellular and intracellular AGR2 in cell adhesion

AGR2 is an ER-resident protein (Chapter 3) but can also be secreted by some cells [301, 388, 426, 440], and so the effects on adhesion of both intracellular and extracellular AGR2 were tested in this chapter. A previous study showed that extracellular AGR2 coated onto cell culture plates increased the rate of cell adhesion [309], as demonstrated here. However, it should be noted that AGR2 has an overall positive charge of + 4, which may contribute to cell attachment through electrostatic interactions as described above. Furthermore, the inhibition of adhesion caused by BSA could be explained by its high overall negative charge (- 10), causing unfavourable interactions between itself and the negatively-charged cell surface [591]. $AGR2_{41\text{-}175}$ and $AGR3_{24\text{-}166}$ have an overall neutral charge, which may account for the reduction in adhesion compared to uncoated wells as a result of masking favourable electrostatic interactions with the plate surface. The absence of a negative charge could account for the lack of total adhesion inhibition as seen with BSA. However, the salt concentration in culture medium can negate the effects of electrostatic interaction between cells and coated surfaces [600], and the fact that BSA inhibits binding, and fibronectin increases binding despite both bearing a negative charge [581] indicates that charge is not the only factor in determining cell attachment to protein-coated surfaces. Furthermore, the fact that coating with 21-40 peptide (which is positively charged) inhibited cell attachment further suggests that charge is not the only factor that influences cell adhesion in these experiments.

Further evidence for a role of extracellular AGR2 stems from studies of newt anterior gradient protein (nAG), a member of the anterior gradient family [601]. nAG is important for limb regeneration in the salamander, where it functions through the GPI-anchored surface receptor, Prod1 [450, 451]. Importantly, Prod1 is also implicated in mediating cell adhesion [451] and although Prod1 appears to be restricted to salamanders, the most closely related mammalian proteins are urokinase-type plasminogen activator receptor (uPAR) and CD59 [452, 602]. Both of these surface proteins are involved cell adhesion: uPAR mediates adhesion through interactions with integrins and vitronectin, and adhesion is enhanced when uPAR is also bound by uPA [603-605], whereas T-cell binding to CHO cells is mediated by CHO-expressed CD59 [606, 607]. Furthermore, AGR2 is reported to bind to C4.4a [389], a structural homologue of uPAR [453] as well as dystroglycan, a transmembrane glycoprotein linking the cytoskeleton to the ECM (reviewed in [608]). With this in mind, it is possible that AGR2 acts through one of these surface receptors to affect cell adhesion. In this context, it appears that the flexible region of the protein (comprising residues 21-40) is the key feature required to promote adhesion, though this must be in the presence of the rest of the protein as the 21-40 peptide alone did not promote adhesion. Such flexible regions are known to mediate protein-protein interactions (reviewed in [326]), and so it is possible that the 21-40 region might mediate receptor binding. However, the inhibition of adhesion observed with the 21-40 peptide suggests that this region alone is not sufficient to promote adhesion, pointing to the requirement of both the 21-40 region and the globular part of the protein (residues 41-175) in this function. It is interesting to note that deletion of part of this globular domain of AGR2 (the C-terminal α-helix and loop region) inhibits its ability to promote anchorage-independent growth [445], providing evidence that the AGR2 globular domain is at least involved in anchorage-independent growth. The extracellular role of AGR2 in adhesion, while probably requiring the combined actions of the flexible and globular regions of the protein, appears to be independent of both the Cys81 thiol group and AGR2 dimerisation, given the lack of changes to cell adhesion between these mutant and WT AGR2 protein-coated wells (Fig. 4.1). Further experimentation will be required to determine whether the entire globular domain is required for AGR2 to promote adhesion.

The effect of intracellular AGR2 on cell adhesion depends upon the nature of the cells. In cell monolayers, the effect of AGR2 can be largely defined by its effects on cell morphology, as there exists a clear correlation between cell morphology and cell attachment, as previously reported [609, 610]. These effects on attachment were generally reciprocal, such that cells that were strongly adherent upon initial plating were

subsequently weakly adherent upon forming a monolayer, and vice versa. In the context of WT AGR2 clonal cells, this weak attachment upon monolayer formation may be explained by a reduction in the number of focal adhesions anchoring these cells to the substratum (as described above). This weakened attachment is presumably related to their more migratory phenotype relative to cuboidal EV clone 2 cells (Fig. 4.17), where the deposition of FAs for gripping the substratum to allow cellular movement must be balanced by opposing events that allow the cell to leave its previous location [611]. Therefore, migrating cells require a higher rate of FA turnover, such that fewer FAs will be present at any given time compared to a non-migratory cell (reviewed in [612]). Furthermore, this high turnover rate presumably stops or severely reduces the formation of the more anchoring fibrillar adhesions [144, 145], further weakening the cell-substratum interaction.

In all WT and mutant AGR2 cell lines, attachment behaviour was correlated with cell morphology, regardless of the particular mutation involved (Fig. 4.13). This provides clues as to how intracellular AGR2 might affect cell morphology, and in turn cell adhesion. The variability in the attachment levels in short term assay makes it hard to draw any firm conclusions about the involvement of mutations in AGR2 function, but the far more consistent attachment behaviour of cells in monolayers offers a better readout of the effects of these mutations. Interestingly, while the 21-40 region appears to be important for the activity of extracellular AGR2, Δ21-40 AGR2 clone 4 cells were the only mutant AGR2-expressing cells to maintain a SR value that was significantly different from EV clone cells, although it was also significantly different from WT clone 13 cells. Coupled with a level of cell attachment in between the levels of WT AGR2 clone 13 and EV clone 2 cells, this suggests that the loss of the flexible region may reduce the activity of AGR2 and indeed, deletions of flexible regions can affect the efficiency of protein interactions, at least in the context of enzyme-substrate interactions [613]. Therefore, while not critical to AGR2 function, the loss of the flexible region may make AGR2 less efficient and thus less able to induce the strong morphological change observed with WT protein.

In both sets of attachment assays, E60A AGR2 clone 18 and C81S AGR2 clone 10 cells displayed attachment behaviour that was no different to EV clone 2 cells, and also displayed an SR value that was more cuboidal than EV clone cells. This points to dimerisation and the Cys81 thiol group playing a particularly important role in mediating the morphological changes induce by AGR2. It is important to note that at least one PDI family member (ERp29) similarly requires dimerisation to function in its role as a chaperone [537], and, whereas the evidence in Chapter 3 points to a lack of PDI activity for AGR2, the thiol-dependent chaperone activity of ERp44 provides a mechanism whereby AGR2 might

also promote the production of protein involved in morphology regulation (discussed in Chapter 3). Importantly, AGR2 is known to affect the production of several mucins, including MUC1 [412], MUC2 [408, 422], MUC5AC and MUC5B [418, 427], and forms thiol-mediated interactions with at least one of these proteins, MUC2 [301]. These mucins are large molecular weight glycoproteins which maintain cell polarisation, both enhance and decrease cell adhesion, and provide a protective epithelial surface lining (reviewed in [614]). Critically, MUC1, which is a marker of poor prognosis in several cancers [615-617] has been shown to induce EMT in pancreatic and breast cancer cells [618, 619], providing a direct mechanism whereby AGR2 might induce EMT. It will be interesting to examine the expression of MUC1 in WT AGR2-expressing clones and the role it may play in bringing about morphological changes in these cells.

The attachment behaviour of ΔKTEL AGR2 clone 8 cells provides an interesting link between intracellular and extracellular AGR2 activities. Whilst ΔKTEL AGR2 clone 8 cells have an SR value that is no different to EV clone cells, the differences in adhesion between these cells may point to the involvement of other factors that govern AGR2-induced changes to cell attachment, rather than morphology alone. The high levels of secreted AGR2 in these cells (see Chapter 3) raises the possibility that secreted AGR2 plays a role in the adhesion of these cells, especially given the increase in adhesion seen when AGR2 is coated onto the substratum (Section 4.2.1). Indeed, reports have shown that the secretion of other ER chaperones such as calreticulin and GRP78 can support tumourigenesis, affecting pro-survival pathways and ECM modulation [448, 449, 620, 621]. Similarly, secretion of the cytoplasmic chaperone Hsp90 promotes cell motility and metastasis through activation of EMT [622]. Therefore, AGR2 may affect adhesion in two ways: intracellular protein affects the production of morphology-related proteins, whilst secreted AGR2 could presumably affect adhesion via signalling through cell surface receptors.

4.4 Conclusions

In this chapter, intracellular production of AGR2 has been found to enhance the adhesion of newly attached cells but to reduce the adhesion of monolayers of cells. A correlation between AGR2 expression and an elongated cell morphology has been described which might form the basis of the effect of AGR2 on cell adhesion, as these changes in cell morphology appear to be a major contributor to the altered attachment and

anchorage-independent growth properties of these cells. This chapter opens up the possibility of understanding the proteins and pathways that are affected by AGR2 to bring about these changes.

Chapter 5

Proteome analysis of AGR2-expressing cells

5.1 Introduction

The expression of AGR2 has been reported in a number of different cancers and is associated with metastasis [309] and decreased patient survival (see Section 1.4.2.2.1), but yet its mechanism of action in cancer remains largely undefined. The protein belongs to the protein disulphide isomerase (PDI) family of chaperones based on its sequence homology to other PDI proteins, but contains a divergent single-cysteine active site [378]. This divergent motif does not appear to be catalytically active in terms of PDI activity, but does permit thiol-dependent interactions between the luminal PDI ERp44 and target substrates [320, 573]. Therefore, as PDIs and chaperones

generally promote the folding of client proteins [310], AGR2 may function by increasing the output of correctly folded proteins from the endoplasmic reticulum (ER), in either a general or substrate-specific manner.

Accordingly, AGR2 is known to be directly required for the secretion of the gel-forming mucin MUC2 in the gut [408, 422], with which it forms thiol-dependent interactions, supporting the notion that AGR2 can function as a protein chaperone. AGR2 is further implicated in the expression of two other gel forming mucins (MUC5AC and MUC5B), where it is reported to interact with immature mucin chains [418, 427]. The nature of these interactions has not been defined, but is likely to be also thiol-dependent, as AGR2 was shown to be associated only with immature mucin [418], where there are presumably a number of unpaired disulphides and thus free-thiols will be present. Additionally, immature MUC5AC co-immunoprecipitated with AGR2 [418], suggesting the presence of a reasonably stable interaction between the two. Similarly, the cell surface mucin, MUC1, has been shown also to co-immunoprecipitate with AGR2, and both proteins co-localise by immunofluorescence, seemingly in the ER in neoplastic pancreatic cells [412]. AGR2 is also associated with the expression of the surface mucin MUC4, but direct regulation by, or interaction with, AGR2 was not determined [433]. Both of these cell surface mucins are associated with cancer progression, where differences in expression levels or glycosylation patterns are observed, and thus these proteins may be useful biomarkers for monitoring disease progression [432, 623]. Additionally, AGR2 has been shown to upregulate the levels of the lysosomal proteases, cathepsins B and D in pancreatic cancer cells [437], which also led to an increase in their secretion and an increase in the invasive capability of these cells *in vitro*. Since a direct interaction with AGR2 could not be determined, the authors suggested that this may relate to the probable rapid and transient interaction between these proteins and AGR2. Such transient interactions may also be a reason why it has been possible to trap AGR2-mucin complexes, as these proteins contain numerous disulphide bonds and thus presumably either undergo extended interactions with AGR2 or are associated with multiple AGR2 molecules.

The abundances of intracellular proteins are also influenced by AGR2, with a recent proteomics screen identifying several differentially regulated proteins of the p53 pathway in AGR2-expressing cells compared to non-AGR2-expressing cells [438]. A previous report also demonstrated that AGR2 inhibits p53 phosphorylation and activation in response to UV damage [445]. Whether AGR2 interacts directly with elements of these intracellular pathways has not been determined, but it is interesting that AGR2 can also be secreted [166, 301, 388, 440], and thus could presumably influence these signalling pathways

through binding to cell surface receptors. AGR2 has indeed been reported to interact with the cell surface C4.4a and dystroglycan proteins in a yeast two-hybrid screen, but these interactions have yet to be validated in mammalian cells [389]. Thus, AGR2 appears to be able to both directly and indirectly regulate the abundances of proteins both inside and outside the cell. Whether the intracellular actions of AGR2 related solely to influences on the p53 pathway however, is not clear. Similarly, whilst AGR2 affects the secretion of several proteins, it is not known if these are AGR2-specific substrates or whether they reflect a more general upregulation of protein output. Therefore, given the largely unknown extent of AGR2 function, investigating AGR2-dependent global changes to the intracellular and extracellular proteomes may provide a better idea of processes and individual proteins regulated by AGR2, and in this way may help elucidate its role in cell adhesion and metastasis.

5.1.1 Chapter objectives

The aims of this chapter are to employ a mass spectrometry-based approach to identify proteins whose abundances might be influenced by the expression of AGR2 in Rama 37 cells, and which might themselves influence the adhesive/metastatic properties of the cells.

5.2 Results

5.2.1 AGR2-induced changes to the intracellular proteome

The Rama 37 cell line is a model system for highlighting AGR2-mediated effects on adhesion and metastasis [309] (Chapter 4), and thus a four-replicate mass spectrometric proteomic screen of differentially regulated proteins was performed on WT AGR2 positive and negative (EV) Rama 37 pool cells in order to investigate possible protein changes associated with the presence of AGR2. 2566 quantifiable proteins were detected in the comparative mass spectrometric screen between WT AGR2 and EV pool cells. Of these, only 5 proteins, identified by two or more unique peptides, showed a greater than two-fold difference in expression between the two cell pools (Table 5.1).

Upregulated in AGR2-positive cells				
Uniprot Accession	***Description***	***Fold change up***	***Anova (p)***	***Unique peptides***
D3ZIA7	Anterior gradient 2 (AGR2)	7.58	<0.01	3
Q4QRB8	Argininosuccinate lyase (ASL)	4.80	<0.01	2

Downregulated in AGR2-positive cells				
Uniprot Accession	***Description***	***Fold change down***	***Anova (p)***	***Unique peptides***
Q642B0	Glypican 4 (GPC4)	2.65	<0.01	3
P62966	Cellular retinoic acid-binding protein 1 (CRABP1)	2.07	<0.01	3
Q6P6Q2	Keratin, type II cytoskeletal 5 (KRT5)	2.02	<0.01	9

Table 5.1. ***Effects of AGR2 expression on the intracellular proteome of pooled transfected Rama 37 cells***.

AGR2 was identified as the protein with the highest fold-increase in AGR2-transfected cells, indicating that the present methodology can indeed detect the increased levels of AGR2 previously observed in these cells (Chapter 3). Other than AGR2 however, only argininosuccinate lyase (ASL), a component of the arginine metabolism pathway, was

upregulated in WT AGR2 cells. Cancers including melanomas, hepatocellular carcinomas and glioblastomas can display a growth dependence on arginine, and ASL is indeed upregulated in some of these cancers (reviewed in [624]), and therefore upregulation may provide a survival advantage for AGR2-expressing cells.

There were moderate decreases in only three proteins associated with AGR2 expression (Table 5.1): glypican-4, cellular retinoic acid-binding protein 1 (CRAPBP1) and cytokeratin 5. CRABP1 is involved in the binding and transport of retinoic acid (RA), a promoter of cell growth and differentiation [625]. Decreasing the levels of CRABP1 might decrease the availability of RA, in a similar manner to CRABP2 [626] and thus possibly promote de-differentiation or reduced growth of the cells, which may be important for the generation of cancer stem cells (reviewed in [627]). Glypican-4 is a GPI-anchored heparan sulphate proteoglycan (HSPG) that enhances insulin signalling through interaction with the insulin receptor [628], and is also a positive regulator of the Wnt/planar cell polarity (PCP) pathway, a pathway involved in promoting cell migration (reviewed in [629]). A reduction in glypican-4 may seem counter-intuitive given the increase in metastases observed in AGR2-overexpressing cells [309], but overexpression of glypican-4 has also been shown to prevent migration [630]. The authors suggest that, as with other Wnt-regulating HSPGs such as Knypek [631], glypican-4 may regulate the availability of Wnt ligands in the extracellular space, and thus by decreasing levels of glypican-4, the levels of available Wnt molecules are increased. This could therefore result in an increase in signalling through the Wnt/PCP pathway, which leads to re-organisation of actin filaments and promotes migration [629]. Overall, although these changes in levels of intracellular proteins were minimal, they hint at AGR2 affecting pathways involved in cell migration, differentiation and survival, all aspects which could favour the development of metastases. However, these changes will need to be validated and any causative link with changes in the levels of these proteins with an effect on metastasis will also have to be demonstrated.

As mentioned above, AGR2 has been shown to affect the secretion of several secreted proteins [408, 412, 418, 422, 427, 437] and so, given the limited effects on intracellular protein abundance observed in AGR2-expressing cells, the effects of AGR2 expression on secreted proteins was next explored.

5.2.2 AGR2-induced changes to the secreted proteome

5.2.2.1 Developing a methodology for the collection and analysis of cell secretomes

Studying the secreted proteome (or 'secretome') from cultured cells presents several challenges. These relate to the contamination of the pool of secreted proteins by intracellular proteins released from dead or dying cells, the low concentration of secreted protein within the culture medium and the masking of these proteins by serum proteins (e.g. BSA) used in most cell culture media [632]. Thus, experiments were carried out to establish methodologies that address these potential problems.

To avoid the effects of a large amount of contaminating serum proteins in both LC-MS and gel electrophoresis analyses of the secretomes, cells were incubated in serum-free Rama 37 medium, and the resulting conditioned medium (CM) was collected for analysis by LC-MS (see Materials and Methods). A total of 802 proteins were identified in the comparative CMs of WT AGR2 pool and EV pool cells (including differentially and non-differentially expressed proteins), and these proteins were subjected to further analysis using the Ingenuity Pathway Analysis (IPA) software package (Qiagen). Of the proteins identified, 627 (78.2 %) could be mapped to the IPA database, and information on the recorded subcellular localisation of these proteins could be collected for 527 (90.1 % mapped proteins, 71.1 % total proteins) of these (Fig. 5.1).

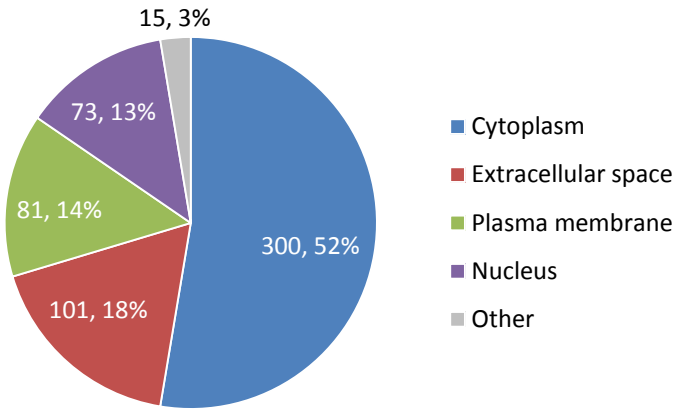


Figure 5.1. ***Distribution of proteins between intracellular and extracellular compartments from conditioned medium of WT AGR2 pool and EV pool cells incubated in serum-free DMEM medium.*** CM proteins were analysed by LC-MS and their cellular distribution determined using Ingenuity Pathway Analysis. Note that this encompasses all proteins identified in the CM of both cell types, not just differentially expressed proteins.

Only 101 (18 %) of the identified proteins were listed as belonging to the extracellular space, with 81 (14 %) proteins labelled as membrane proteins. Given that one proteomic study identified 18 of 22 transmembrane proteins from human mammary epithelial cells as having an extracellular shed form [633], at least some of the membrane proteins identified in the WT AGR2 pool and EV pool secretomes could also be shed-domains of membrane proteins. However, 373 (65 %) of the secretome proteins were of intracellular origin (cytoplasmic or nuclear), suggesting that there was substantial contamination of the secretome by intracellular proteins, presumably released from dead and dying cells.

Several studies investigating secreted proteins have used Opti-MEM serum-free medium (Gibco) for the collection of these proteins [634-638]. Opti-MEM is a chemically-defined, low protein concentration medium containing growth factors to support growth in the absence of FBS. To try and minimise the contamination of the secretome by intracellular proteins observed with serum-free DMEM medium, EV and WT AGR2 pool cells were incubated in Opti-MEM medium and the CM again analysed by LC-MS and IPA (Fig. 5.2).

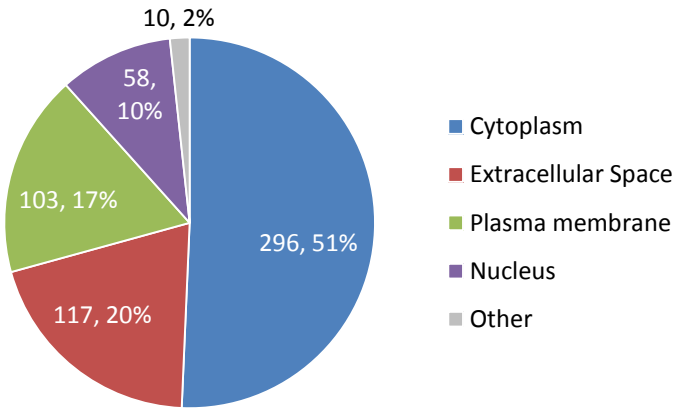


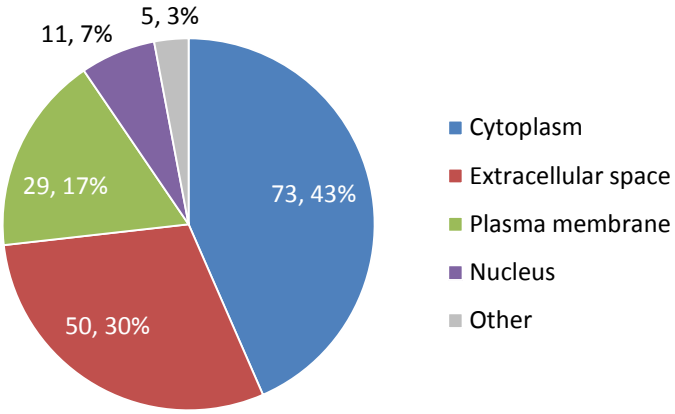
Figure 5.2. ***Distribution of proteins between intracellular and extracellular compartments from conditioned medium of WT AGR2 pool and EV pool cells incubated in serum-free Opti-MEM medium.*** CM proteins were analysed by LC-MS and their cellular distribution determined using Ingenuity Pathway Analysis. Note that this encompasses all proteins identified in the CM of both cell types, not just differentially expressed proteins.

A similar number of proteins were identified in the CM of Opti-MEM-incubated cells (783 proteins), and both the identity of these proteins and the overall distribution of proteins between cellular compartments was largely the same as from DMEM-incubated

cells. However, the overall concentrations of these proteins in the CMs, based on peptide abundances calculated by Progenesis software (see Materials and Methods), differed between DMEM- and Opti-MEM-incubated cells. This suggests that simply examining all the proteins identified in CMs may not be an appropriate method for determining the suitability of the CM collection method, and that the relative abundances of these proteins should be taken into account. This necessitates the use of a normalisation factor to compare protein abundances across different cell lines and CM collection methods. Therefore, one *bona fide* secreted protein, fibronectin [639, 640], was identified in both experiments shown in Fig. 5.1 and 5.2 that showed no significant difference in expression between EV or WT AGR2 cells, and could potentially be used as a normalisation factor for the levels of secreted proteins between WT AGR2 and EV pool cells. In addition, pilot data suggested that there was indeed no difference in the amount of fibronectin secreted from WT AGR2 and EV pool cells, as CM from both cell types analysed by SDS-PAGE showed comparable levels of fibronectin (Fig. 4.8). Similarly, there was no significant difference in intracellular levels of fibronectin between these cells (fold change: 1.12, $p = 0.078$, ANOVA). This suggests that the secreted levels of fibronectin could be used as a normalisation factor for the secretomes of WT AGR2 and EV pool cells.

With this in mind, the levels of proteins identified in CM collected from DMEM- and Opti-MEM-based serum-free media by LC-MS were normalised to the level of fibronectin in the CMs. The mean level of fibronectin across the four replicate samples for each cell type in each serum-free medium was used as a normalisation factor for the mean levels of each protein identified in the CM of each cell type. This yielded a fibronectin-normalised abundance value that was used to compare the abundances of CM proteins across cell types and incubation media. To determine the most appropriate collection method, the cellular localisation of the most abundant proteins in both collection media was assessed. The most abundant proteins were determined as those present at a level greater than 0.5 times that of fibronectin, retaining 168 proteins from DMEM-incubated cells and 15 from Opti-MEM incubated cells. In essence, the value itself of the cutoff point is not important, as long as it selects for the most abundant proteins; choosing a lower cutoff will simply increase the proportion of proteins included in the analysis that are released as a consequence of cell lysis, which should be present at a lower concentration than true secreted proteins.

A



B

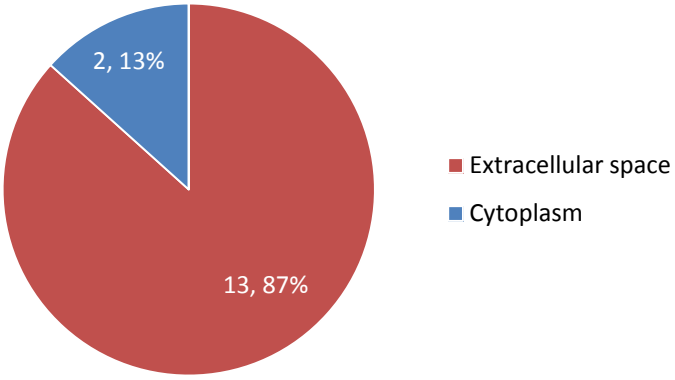


Figure 5.3. ***Comparison of distribution between intracellular and extracellular compartments of the most abundant proteins from conditioned medium of DMEM- and Opti-MEM-incubated WT AGR2 and EV pool cells.*** CM proteins were analysed by LC-MS and their cellular distribution determined using Ingenuity Pathway Analysis. Protein abundance was normalised to fibronectin, and the cellular localisation for most abundant proteins (present at a level ≥ 0.5 times the abundance of fibronectin) were determined using Ingenuity Pathway Analysis. **(A)** DMEM-incubated cells and **(B)** Opti-MEM-incubated cells.

Considering only the most abundant proteins, the proportion of proteins denoted as belonging to the extracellular space increased in DMEM-incubated cells rose from 18 % of all CM-derived proteins (Fig. 5.1) to 30 % of the most abundant proteins (Fig. 5.3A). However, 50 % of the most abundant proteins were still classified as cytoplasmic. In the CM from Opti-MEM-incubated WT AGR2 and EV pool cells, only 15 proteins out of all identified proteins from both cell types (i.e. including differentially and non-differentially expressed

proteins) were found at levels ≥ 0.5 times that of fibronectin (Fig. 5.3B). Of these, only 2 (13 %) were not classified as belonging to the extracellular space. Of these two cytoplasmic proteins, one (nucleobindin-1) is resident in the Golgi and the other (SRPRB) is an ER protein. Given the previously observed secretion of other ER proteins [448, 449, 620, 621], these could potentially also be secreted proteins, but overall, this shows that the majority of the most abundant proteins in the CM from Opti-MEM medium are indeed extracellular proteins. These observed differences in the proportion of extracellular and intracellular proteins with these two serum-free media indicate that Opti-MEM is the more suitable method for minimising contamination from intracellular proteins. While CM from these Opti-MEM-incubated cells contained many fewer proteins expressed at ≥ 0.5 times the level of fibronectin, this is not surprising given the generally low abundance of secreted proteins [632]. An independent validation of the method can also be seen in the levels of the cytoplasmic protein lactate dehydrogenase A (LDHA) relative to fibronectin in the extracellular space. Extracellular LDHA is routinely used as a measure of cell lysis [641, 642], and the average levels of LDHA relative to fibronectin in the CM from both EV and WT AGR2 cells were 0.12 for cells incubated in Opti-MEM and 8.44 for cells incubated in DMEM, a 70-fold difference in the relative amount of extracellular LDHA between these methods. This indicates a much higher level of cell lysis in DMEM-incubated cells and reinforces the suitability of Opti-MEM as the optimal medium for collecting CM.

5.2.2.2 Effects of AGR2 expression on secreted proteins

Using the LC-MS dataset collected from cells using the Opti-MEM protocol, the changes to secreted proteins associated with AGR2 expression were explored. A total of 758 quantifiable proteins were identified from the comparative secretomes of WT AGR2 positive and AGR2-negative (EV) pooled cells. The abundances of the identified proteins were normalised to the levels of fibronectin in each replicate, and the list reduced to proteins that were significantly differentially expressed ($p < 0.05$, Student's *t*-test) and identified by 2 or more unique peptides, yielding a list of 30 proteins differentially expressed between the CM of WT AGR2 and EV pool cells (Table 5.2). 63 % of identified proteins were classed as extracellular by IPA and 30 % as plasma membrane proteins. Only four of the differentially regulated proteins were upregulated in the AGR2-expressing cells (including AGR2 itself), whilst the remainder were downregulated in these cells, relative to the EV controls. The identification of AGR2 in CM suggests that the protein is indeed

secreted, although it was present at a relatively low concentration relative to fibronectin (0.002 times the concentration of fibronectin). However, as AGR2 was found to localise mostly to the ER (Chapter 3), this is not unexpected and suggests that only a relatively low amount of AGR2 is secreted. Pathway analysis revealed that the top functions that the differentially regulated proteins (both up- and down-regulated) are involved in were cellular movement and cellular development, in keeping with the roles of AGR2 in cell differentiation and metastasis. In particular, there was downregulation of several adhesion molecules (cadherin 1, 2 and 11, as well as LGALS3BP), although to what extent levels of membrane proteins in the CM reflect the actual membrane levels is not clear. There was down-regulation of basement membrane proteins (LAMA4, LAMB2, fibulin and LOXL3), coupled with an increase in the matrix-degrading protease ADAMTS1. It is also interesting that there were changes to several cytokines (CXCL1, CXCL2 and granulin). Overall, in the pool cells tested here, the presence of AGR2, rather surprisingly for a potential chaperone, does not appear to be associated with an increase in the level of many proteins, other than the secretion of two CXCL cytokines, which is potentially important for the recruitment of hematopoietic cells that are known to support tumour growth and invasion [184, 643]. However, AGR2 might be associated with the downregulation of adhesion molecules.

Table 5.2 (continued overleaf). Effects of AGR2 expression on the secretome of Rama 37 cells.

Upregulated in AGR2-expressing cells				
Uniprot Accession	***Description***	***Fold change up***	***t-test (p)***	***Unique peptides***
D3ZIA7	Anterior gradient 2 (AGR2)	22.48	<0.01	2
P30348	C-X-C motif chemokine 2 (CXCL2)	4.25	0.02	2
G3V6C6	Chemokine (C-X-C motif) ligand 1 (CXCL1)	3.64	0.03	2
Q9WUQ1	A disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS1)	3.49	<0.01	10

Downregulated in AGR2-expressing cells				
UniProt Accession	***Description***	***Fold change down***	***t-test (p)***	***Unique peptides***
D4ACD7	Follistatin-related protein 5 (FSTL5)	34.55	0.01	19
F1LR45	Sushi repeat-containing protein 2 (SRPX2)	5.51	0.02	4
F1MAP4	Multiple EGF-domains 10 (MEGF10)	4.52	0.048	7
D3ZP82	Lysyl oxidase-like 3 (LOXL3)	4.40	<0.01	10
F1MAH6	Cadherin-11 (CDH11)	4.25	<0.01	4
P63090	Pleiotrophin (PTN)	4.18	0.02	5
F1MAG8	Semaphorin 3D (SEMA3D)	3.70	<0.01	3
H9KVF9	Proprotein convertase subtilisin/kexin type 5 (PCSK5)	3.69	<0.01	2
Q642B0	Glypican 4 (GPC4)	3.45	<0.01	10
D3ZEU2	Laminin subunit alpha-4 (LAMA4)	3.21	<0.01	4
Q62894	Extracellular matrix protein 1 (ECM1)	3.09	<0.01	15

G3V803	Cadherin-2 (CDH2)	2.92	<0.01	10
O70513	Galectin-3-binding protein (LGALS3BP)	2.74	<0.01	5
Q9R0C5	N-acetylgalactosaminyltransferase 7 (GALNT7)	2.66	0.01	4
Q6P9Z6	Tumor-associated calcium signal transducer 2 (TACSTD2)	2.61	<0.01	2
Q4V8N0	Lipocalin 7 (TINAGL1)	2.59	<0.01	8
G3V8V1	Granulin (GRN)	2.49	<0.01	10
F1LRR7	Suppressor of tumorigenicity 14 (ST14)	2.33	0.01	13
P15800	Laminin subunit beta-2 (LAMB2)	2.14	<0.01	31
Q9R0T4	Cadherin-1 (CDH1)	2.13	<0.01	2
P29524	Plasminogen activator inhibitor 2 type A (SERPIN B2)	2.13	0.04	5
G3V6X1	Fibulin 2 (FBLN2)	2.12	0.01	22
P55053	Fatty acid-binding protein, epidermal (FABP5)	2.02	0.01	2
Q641Z6	EH domain-containing protein 1 (EHD1)	2.01	<0.01	2
G3V6B1	Transforming growth factor beta-2 (TGFB2)	2.01	<0.01	2
Q8R3Z7	EH-domain containing 4 (EHD4)	2.00	<0.01	4

Table 5.2 (cont.). Effects of AGR2 expression on the secretome of Rama 37 cells.

5.2.3 AGR2-induced changes to secreted proteins in clonal cell lines

As shown in Chapter 4, clonal cell lines expressing AGR2 exhibit altered cell adhesion properties compared to clonal AGR2-negative EV cell lines. The differences in secreted proteins between WT AGR2 clone 2 and EV clone 2 cells were therefore explored using LC-MS to investigate any changes that might explain these differences in cell behaviour (Table 5.3).

While the experiments involving the CM from WT AGR2 and EV pool cells in section 5.2.2.2 could be normalised to fibronectin due to its equal expression between cell lines, the same was not true for WT AGR2 clone 2 cells and EV clone 2 cells, where EV clone 2 cells secreted substantially more fibronectin that WT AGR2 clone 2 cells (Fig. 4.8). Thus, CM protein levels could not be normalised to fibronectin for these cell lines. Furthermore, due to the presence of significant amounts of transferrin in Opti-MEM medium, CM samples could not be normalised to protein concentration due to masking of actual secreted protein concentration by transferrin, although this is only a problem for clonal cell CM due to the lack of possible fibronectin normalisation. Similarly, normalising samples based on cell number, as reported in other studies [644, 645], were found to be poorly reproducible (Appendix 24). Therefore, normalisation for clonal CM samples in LC-MS experiments was achieved using the normalisation algorithm within the Progenesis software, which considers all peptides present within the sample to calculate a normalisation value (see Materials and Methods). However, WT AGR2 clone 2 cells have an average relative area 79 % that of EV clone 2 cells (as measured by ImageJ), and because cultures were placed into serum-free medium for CM collection upon reaching 30-40 % confluence (see Materials and Methods), WT AGR2 clone 2 cells at this confluence will contain more cells than 30-40 % confluent EV clone 2 cells. Therefore, there could be a disproportionate abundance of secreted proteins in CM from the WT AGR2 clone 2 cells. However, to overcome this potential problem, the threshold for a meaningful difference in a secreted protein between the two cell lines was increased from 2-fold to a more conservative 3-fold change in protein abundance. Using this cut-off, there are still many differences in upregulated proteins observed between EV clone 2 and WT AGR2 clone 2 cells (Table 5.3).

Many differentially expressed proteins were identified in the comparative CMs of WT AGR2 and EV clonal cells. However, it is important to note that the observed decreases in histones H1.2, H2B, H3 and H4 in the CM of WT AGR2 clone 2 cells (Table 5.3), coupled with the fact that a number of the AGR2-downregulated proteins also belong to the 60S ribosomal family of cytoplasmic proteins, may suggest that some of these differences only

arise from a difference in the extent of cell lysis between EV clone 2 and WT AGR2 clone 2 cells under serum-free conditions. The increased levels of extracellular LDHA found in the CM of EV clone 2 cells (3.82-fold increase relative to WT AGR2 clone 2, Table 5.3) also support this notion, and also suggests that WT AGR2 clone 2 cells are more resistant to cell lysis. However, without knowing the relative intracellular expression of LDHA between WT and EV cells (or indeed between any proteins identified in these secretomes), it is difficult to set a cutoff value for meaningful downregulation of proteins in the CM of these cell lines. Therefore, only actual extracellular proteins were considered in further analyses, as their presence in CM collected from Opti-MEM-incubated cells appears to represent actual secreted proteins, as described above.

It is interesting that the difference in the abundance of secreted AGR2 was lower between clonal cells than that observed with pooled cells (9.4 fold and 29.9 fold, respectively), despite very similar steady-state intracellular AGR2 levels between pooled and clonal WT AGR2 cells (see Chapter 3). This may indicate a general decrease in the secretory potential of elongated cells compared to cuboidal cells. Nevertheless, many more proteins were differentially secreted between clonal cells than were observed with pooled cells, probably reflecting the fundamental changes to cell morphology in WT AGR2 clonal cells. Of note, there was increased secretion of several proteases including ADAMTS2, ADAMTS8 and MMP19, but there was also a reduction in several other proteases, ADAMTS7, MMP3 and PRSS22 (Table 5.3). There was an upregulation of multiple collagen molecules, namely those of the collagen VI family but also members of the collagen I, II and VII families. At the same time, there was downregulation of several basement membrane proteins including agrin and laminins (LAMC2, LAMA3, LAMB3 and LAMA5), but also upregulation of LAMA4. There was downregulation of several members of the serpin family of protease inhibitors (serpin B2 (plasminogen activator inhibitor 2), B5 and B7), and also the SPINT1 protease inhibitor. However, there was also upregulation of serpin 1 (plasminogen activator inhibitor 1) and serpin F1.

The axon guidance protein draxin [646] showed the largest increase in expression, with an increase of 770-fold in WT AGR2 clone 2 cells, suggesting that AGR2 may be involved in draxin secretion. Axon guidance proteins can act as both oncogenes and tumour suppressors, with both positive and negative effects on apoptosis, proliferation and adhesion (reviewed in [647]). The strong effect of AGR2 expression on axon guidance proteins, namely draxin, SLIT2 and members of the semaphorin family (SEMA3C, SEMA3D and SEMA3E) might suggest not only that AGR2 might regulate the secretion of these

proteins, but that regulating this class of molecules plays a role in AGR2-associated tumourigenesis.

The RT1-AW2 Class I histocompatibility antigen showed the greatest decrease in expression in WT AGR2 clone 2 cells, displaying a 200-fold reduction (Table 5.3). RT1-AW2 is thought to be involved in presenting antigens through the MHC class 1 pathway, based on its similarity to other RT1 proteins [648] but as a transmembrane protein, it is not clear whether the decrease in RT1-AW2 in the CM of WT AGR2 clone 2 cells necessarily reflects a change in protein expression or a decrease in a potential soluble form of the protein, as has been observed with at least one other histocompatibility-related transmembrane protein, MICA (major histocompatibility complex (MHC) class I–related chain A) [649]. The implications of reduced RT1-AW2 are unknown but could feasibly help to reduce unfavourable recognition of metastatic cells by immune cells.

Overall, these changes to the secretome of WT AGR2 clone 2 cells point to a cell type with altered adhesion and migration, a reduction in basement membrane proteins and an increase in abundance and activity of several extracellular proteases. These cells also appear to deposit more collagen molecules and may also undergo changes to bypass immune surveillance mechanisms. Together, the influences on these proteins provide information about how AGR2 might induce changes to cell adhesion metastasis, although further experimentation will be required to determine to what extent these changes in the secretome are directly caused by AGR2 and which are the result of a change in cell morphology, and how much they contribute to AGR2-associated tumourigenesis. Like differences in intracellular proteins, these changes will also have to be validated and direct links between differentially expressed proteins and influences on cell adhesion and metastasis will have to be explored.

***Table 5.3*. Effects of AGR2 expression on the secretome of clonal Rama 37 cells**

Upregulated in AGR2-expressing cells				
UniProt Accession	***Description***	***Fold change up***	***Anova (p)***	***Unique peptides***
D3ZDG4	Draxin	770.52	<0.01	3
D3ZTE7	A disintegrin and metalloproteinase with thrombospondin motifs 2 (ADAMTS2)	151.30	<0.01	3
P24594	Insulin-like growth factor-binding protein 5 (IGFBP5)	120.36	<0.01	12
F1M7X3	Cadherin-13 (CDH13)	91.20	<0.01	3
O08628	Procollagen C-endopeptidase enhancer 1 (PCOLCE)	84.62	<0.01	16
D3ZUL3	Collagen alpha-1 (VI) chain (COL6A1)	79.86	<0.01	23
F1LNH3	Collagen alpha-2 (VI) chain (COL6A2)	56.29	<0.01	18
F1LS40	Collagen alpha-2 (I) chain (COL1A2)	47.73	<0.01	47
D4AAT1	A disintegrin and metalloproteinase with thrombospondin motifs 8 (ADAMTS8)	39.91	<0.01	11
P47853	Biglycan (BGN)	32.60	<0.01	12
Q9ERB4	Versican core protein (VCAN)	30.68	<0.01	15
Q01129	Decorin (DCN)	29.62	<0.01	4
D3ZQM5	Fibrillin-1 (FBN1)	28.12	<0.01	30
D3ZYI8	Secreted protein acidic and rich in cysteine (SPARC)	25.00	<0.01	10
P02454	Collagen alpha-1 (I) chain (COL1A1)	24.88	<0.01	17
F1LSW0	Serpin F1	24.09	<0.01	12
D3ZE24	EGF-containing fibulin-like extracellular matrix protein 2 (EFEMP2)	23.60	<0.01	14
P13941	Collagen alpha-1 (III) chain (COL3A1)	22.55	<0.01	15
P12843	Insulin-like growth factor-binding protein 2 (IGFBP2)	19.64	<0.01	6
Q8CHN8	Mannan-binding lectin serine protease 1 (MASP1)	16.84	<0.01	9
Q06880	Neuroblastoma suppressor of tumorigenicity 1 (NBL1)	16.32	<0.01	2

B5DFC9	Nidogen-2 (NID2)	14.79	<0.01	23
Q6IRK9	Carboxypeptidase Q (CPQ)	14.42	<0.01	3
P15087	Carboxypeptidase E (CPE)	14.17	<0.01	6
D3ZE04	Collagen alpha-1 (VII) chain (COL7A1)	13.63	<0.01	16
G3V7F3	WNT1 inducible signaling pathway protein 2 (WISP2)	12.66	<0.01	2
D4A115	Collagen alpha-3 (VI) chain (COL6A3)	12.58	<0.01	9
Q9R0K8	Stanniocalcin-2 (STC2)	11.67	<0.01	2
C0M4B0	Matrix metalloproteinase 19 (MMP19)	11.14	<0.01	6
D3ZAE6	Vasorin (VASN)	10.08	<0.01	7
Q63772	Growth arrest-specific protein 6 (GAS6)	10.03	<0.01	12
D3ZIA7	Anterior gradient 2 (AGR2)	9.40	<0.01	1
O70513	Galectin-3-binding protein (LGALS3BP)	9.30	<0.01	5
P20961	Plasminogen activator inhibitor 1 (SERPINE1)	9.19	<0.01	22
P03994	Hyaluronan and proteoglycan link protein 1 (HAPLN1)	9.11	<0.01	5
Q6P7A4	Prosaposin (PSAP)	8.85	<0.01	2
D4A9N1	HIPP-like protein-1 (HHIPL1)	7.66	<0.01	4
G3V6K1	Transcobalamin 2 (TCN2)	7.46	<0.01	3
F1LQC3	Collagen alpha-1 (XII) chain (COL12A1)	6.91	<0.01	41
P17491	Inhibin beta B (INHBB)	6.85	<0.01	3
P53813	Vitamin K-dependent protein S (PROS1)	6.63	<0.01	3
G3V928	Prolow-density lipoprotein receptor-related protein 1 (LRP1)	6.44	<0.01	5
G3V7Y9	Alpha-mannosidase 2 (MAN2A1)	6.07	<0.01	3
F1LTJ5	Basement membrane-specific heparan sulfate proteoglycan core protein (HSPG2)	5.95	<0.01	15
D3ZEU2	Laminin subunit alpha-4 (LAMA4)	5.83	<0.01	35
D3ZPA9	Lysosomal protective protein (CTSA)	5.70	<0.01	4
P43145	Adrenomedullin (ADM)	5.66	<0.01	4
F1M6X4	Uncharacterized protein	5.65	<0.01	2

Q5RJL6	Meteorin-like protein (METRNL)	5.19	<0.01	5
G3V9N9	Mannosidase 1, alpha (MAN1A1)	5.12	<0.01	2
F1LTE5	Neural cell adhesion molecule 1 (NCAM1)	4.88	<0.01	13
D3ZHA1	N-acetyllactosaminide beta-1,3-N-acetylglucosaminyltransferase (B3GNT1)	4.57	<0.01	6
P13852	Major prion protein (PRNP)	4.22	<0.01	3
Q01460	Di-N-acetylchitobiase(CTBS)	4.12	<0.01	3
Q63691	Monocyte differentiation antigen CD14 (CD14)	3.79	<0.01	3
F1LML8	N-acetylglucosamine-1-phosphotransferase subunit gamma (GNPTG)	3.58	<0.01	3
F1LWQ3	Follistatin-related protein 1 (FSTL1)	3.39	<0.01	16
P01346	Insulin-like growth factor II (IGF2)	3.39	<0.01	2
D4A2G6	Thrombospondin 2 (THBS2)	3.20	<0.01	10
Q6AYE5	Out at first protein homolog (OAF)	3.01	<0.01	4
F1M879	Platelet-derived growth factor subunit A (PDGFA)	2.82	<0.01	5
G3V6B1	Transforming growth factor beta-2 (TGFB2)	2.74	<0.01	4
P35053	Glypican-1 (GPC1)	2.69	<0.01	9
F1M7L7	Latent-transforming growth factor beta-binding protein 2 (LTBP2)	2.65	<0.01	25
P30121	Metalloproteinase inhibitor 2 (TIMP2)	2.64	<0.01	12
P07154	Cathepsin L1 (CTSL1)	2.58	<0.01	12
Q5RJR9	Serine (Or cysteine) proteinase inhibitor, clade H, member 1 (SERPINH1)	2.55	<0.01	5
G3V6X1	Fibulin 2 (FBLN2)	2.53	<0.01	27
F1LRG5	Galectin-1 (LGALS1)	2.52	<0.01	8
F1MAG8	Semaphorin 3D (SEMA3D)	2.50	<0.01	14
G3V901	Exostosin-1 (EXT1)	2.45	<0.01	3
F1MA79	Slit homolog 2 (SLIT2)	2.43	0.01	4
D3Z8H1	Thrombospondin-3 (THBS3)	2.41	<0.01	2
G3V8L3	Lamin A (LMNA)	2.39	<0.01	15
E9PTT2	Exostosin-2 (EXT2)	2.34	<0.01	7

P31000	Vimentin (VIM)	2.28	0.01	21
P85972	Vinculin (VCL)	2.26	<0.01	11
F1LN51	Tetraspanin-7 (TSPAN7)	2.24	<0.01	2
F1LMA7	C-type mannose receptor 2 (MRC2)	2.24	<0.01	17
P63090	Pleiotrophin (PTN)	2.17	0.02	5
F1M766	Uncharacterized protein	2.16	<0.01	2
G3V6S3	Calumenin (CALU)	2.11	<0.01	4
Q9JI92	Syntenin-1 (SDCBP)	2.07	0.02	2
F1M9B2	Insulin-like growth factor-binding protein 7 (IGFBP7)	2.06	<0.01	15
F1M335	Peroxidasin (PXDN)	2.04	<0.01	28
D3ZP79	Amyloid beta A4 protein C99 (APP)	2.01	<0.01	17

***Table 5.3 (cont)*. Effects of AGR2 expression on the secretome of clonal Rama37 cells**

Downregulated in AGR2-expressing cells				
UniProt Accession	***Description***	***Fold change down***	***Anova (p)***	***Unique peptides***
P15978	Class I histocompatibility antigen, Non-RT1.A alpha-1 chain (RT1-AW2)	202.32	<0.01	2
D3ZAX4	Tumor necrosis factor receptor superfamily member 27 (EDA2R)	59.94	<0.01	2
D4AAE4	Protease, serine, 22 (PRSS22)	36.87	<0.01	14
Q9WU74	Lipolysis-stimulated lipoprotein receptor (LSR)	34.48	<0.01	9
Q4V8Q2	Kunitz-type protease inhibitor 1 (SPINT1)	32.87	<0.01	3
P21674	Follistatin (FST)	32.53	<0.01	17
P15865	Histone H1.2 (HIST1H1C)	30.39	<0.01	2
F1M199	Agrin (AGRN)	28.97	<0.01	20
D4A9V5	Lysyl oxidase-like 4 (LOXL4)	28.48	<0.01	5
B0BMY8	Histone H3 (H3F3B)	26.46	<0.01	3
G3V6C1	Serpin B5	24.28	<0.01	3
O55162	Ly6/PLAUR domain-containing protein 3 (LYPD3)	22.16	<0.01	2
D4A499	Phospholipase A2 inhibitor and Ly6/PLAUR domain-containing protein (PINLYP)	19.17	<0.01	3
F1LNW5	Beta-nerve growth factor (NGF)	17.08	<0.01	6
F1LRH4	Laminin subunit gamma-2 (LAMC2)	16.39	<0.01	28
P08721	Osteopontin (SPP1)	16.38	<0.01	16
P97523	Hepatocyte growth factor receptor (MET)	16.18	<0.01	5
F1LQP8	Cadherin-6 (CDH6)	15.06	<0.01	3
P62804	Histone H4 (HIST1H4B)	13.06	<0.01	4
F1LS18	Histone H2B (HIST1H2BL)	13.02	<0.01	4
G3V8P4	Receptor-type tyrosine-protein phosphatase F (PTPRF)	12.37	<0.01	2

F8WFK8	Cadherin-related family member 1 (CDHR1)	10.51	<0.01	14
D3ZN05	Laminin subunit alpha-3 (LAMA3)	10.34	<0.01	39
Q6P3V9	60S ribosomal protein L4 (RPL4)	9.48	<0.01	3
G3V667	Integrin, alpha 6 (ITGA6)	9.45	<0.01	2
D3ZI60	Protein RGD1561984	9.41	<0.01	2
D3Z8E2	Semaphorin 3E (SEMA3E)	9.25	<0.01	9
F1MAN8	Laminin subunit alpha-5 (LAMA5)	9.15	<0.01	53
F1LQI2	Roundabout homolog 1 (ROBO1)	8.99	<0.01	2
Q6PDV7	60S ribosomal protein L10 (RPL10)	8.75	<0.01	3
D4A8R6	Cytosolic acyl coenzyme A thioester hydrolase (ACOT7)	8.61	<0.01	4
Q6P686	Osteoclast-stimulating factor 1 (OSTF1)	7.82	<0.01	2
G3V6B5	Serpin B7	7.40	<0.01	17
F1LQ82	E3 ubiquitin-protein ligase (NEDD4)	7.35	<0.01	2
G3V9W9	Protocadherin Fat 1 (FAT1)	7.34	<0.01	56
Q3KR76	Plasminogen activator, urokinase (PLAU)	7.13	<0.01	13
D3ZJR7	Desmocollin-2 (DSC2)	7.07	<0.01	9
G3V7I9	Furin	6.89	<0.01	2
F1LPL4	Receptor-type tyrosine-protein phosphatase kappa (PTPRK)	6.53	<0.01	9
Q4V8H5	Aspartyl aminopeptidase (DNPEP)	6.44	<0.01	4
Q5BJ93	Enolase 1, alpha (ENO1)	6.18	<0.01	15
F1LPI5	Laminin subunit beta-3 (LAMB3)	6.07	<0.01	22
Q9WVH8	Fibulin-5 (FBLN5)	6.06	<0.01	5
P07150	Annexin A1 (ANXA1)	6.04	<0.01	8
Q642A6	von Willebrand factor A domain-containing protein 1 (VWA1)	5.62	<0.01	3
F1LS07	C-X-C motif chemokine 12 (CXCL12)	5.59	<0.01	2
P29524	Plasminogen activator inhibitor 2 type A (SERPINB2)	5.56	<0.01	3
P84100	60S ribosomal protein L19 (RPL19)	5.51	<0.01	2

F1LSD3	Integrin beta (ITGB4)	5.47	<0.01	8
Q5RKH2	Galactokinase 1 (GALK1)	5.25	<0.01	2
D3ZVF0	Calpain-3 (NCL)	5.25	<0.01	5
F8WG74	Peptidyl-prolyl cis-trans isomerase FKBP4, N-terminally processed (FKBP4)	5.22	<0.01	2
D3ZRM9	60S ribosomal protein L13 (RPL13)	5.04	<0.01	3
P83732	60S ribosomal protein L24 (RPL24)	5.03	<0.01	3
P62832	60S ribosomal protein L23 (RPL23)	4.92	<0.01	2
F1LML0	Asparagine--tRNA ligase, cytoplasmic (NARS)	4.79	<0.01	2
Q7M0E3	Destrin (DSTN)	4.79	<0.01	7
F1LRR7	Suppressor of tumorigenicity 14 protein (ST14)	4.74	<0.01	13
D3ZCG9	Integrin alpha 3 variant A (ITGA3)	4.68	<0.01	9
Q5XIF6	Tubulin alpha-4A chain (TUBA4A)	4.68	<0.01	3
D3ZYS7	Ras GTPase-activating protein-binding protein 1 (G3BP1)	4.66	<0.01	2
G3V826	Transketolase (TKT)	4.63	<0.01	17
B0K031	60S ribosomal protein L7 (RPL7)	4.52	<0.01	3
P85968	6-phosphogluconate dehydrogenase, decarboxylating (PGD)	4.38	<0.01	4
G3V786	Aldose reductase-related protein 2 (AKR1B3)	4.33	<0.01	4
P49134	Integrin beta-1 (ITGB1)	4.27	<0.01	9
G3V6A4	Heterogeneous nuclear ribonucleoprotein D (HNRPD)	4.16	0.01	2
P03957	Stromelysin-1 (MMP3)	4.04	<0.01	10
F1LV50	Collagen and calcium-binding EGF domain-containing protein 1 (CCBE1)	3.87	<0.01	3
F1M6X1	Vascular endothelial growth factor A (VEGFA)	3.84	<0.01	3
B5DEN4	L-lactate dehydrogenase (LDHA)	3.82	<0.01	10
D3ZRR8	Semaphorin 3C (SEMA3C)	3.81	<0.01	10
Q9WUQ1	A disintegrin and metalloproteinase with thrombospondin motifs 1 (ADAMTS1)	3.81	<0.01	3
P05369	Farnesyl pyrophosphate synthase (FDPS)	3.80	<0.01	7
P05065	Fructose-bisphosphate aldolase A (ALDOA)	3.72	<0.01	19

Q63083	Nucleobindin-1 (NUCB1)	3.71	<0.01	34
F1LSW7	60S ribosomal protein L14 RPL14)	3.65	<0.01	3
B2RYX0	Nascent polypeptide-associated complex subunit alpha, muscle-specific form (NACA)	3.59	0.01	3
F1LM41	Major vault protein (MVP)	3.57	0.02	2
P34058	Heat shock protein HSP 90-beta (HSP90AB1)	3.56	<0.01	11
Q6P502	T-complex protein 1 subunit gamma (CCT3)	3.55	<0.01	3
P48500	Triosephosphate isomerase (TPI1)	3.52	<0.01	8
Q66HL2	Cortactin (CTTN)	3.52	<0.01	3
P04797	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	3.52	0.01	6
Q8VHC1	Cystatin E/M (CST6)	3.49	<0.01	9
D3ZH86	Uncharacterized protein	3.43	<0.01	2
P16617	Phosphoglycerate kinase 1 (PGK1)	3.43	<0.01	12
D4AC23	T-complex protein 1 subunit eta (CCT7)	3.43	0.02	2
Q4FZY0	EF-hand domain-containing protein D2 (EFHD2)	3.42	<0.01	4
D4AA31	Lysosomal Pro-X carboxypeptidase (PRCP)	3.37	<0.01	5
Q6AYD3	Proliferation-associated 2G4 (PA2G4)	3.34	<0.01	3
Q9ERA7	Mesothelin (MSLN)	3.31	<0.01	10
D3ZTR7	N-acetyllactosaminide alpha-1,3-galactosyltransferase (GGTA1)	3.28	<0.01	2
P97603	Neogenin (NEO1)	3.27	<0.01	7
F1LST1	Fibronectin (FN1)	3.26	<0.01	4
Q6P6V0	Glucose-6-phosphate isomerase (GPI)	3.26	<0.01	2
Q9R1E9	Connective tissue growth factor (CTGF)	3.25	<0.01	21
Q6AXS5	Plasminogen activator inhibitor 1 RNA-binding protein (SERBP1)	3.23	0.01	2
F1LRD5	Isocitrate dehydrogenase [NADP] (IDH1)	3.23	<0.01	2
P08699	Galectin-3 (LGALS3)	3.23	<0.01	2
P62634	Cellular nucleic acid-binding protein (CNBP)	3.22	<0.01	2
P05197	Elongation factor 2 (EEF2)	3.21	<0.01	15

P13084	Nucleophosmin (NPM1)	3.18	<0.01	3
P21531	60S ribosomal protein L3 (RPL3)	3.17	<0.01	2
Q9EPH8	Polyadenylate-binding protein 1 (PABPC1)	3.10	<0.01	4
F1M390	Ezrin (EZR)	3.09	<0.01	14
P63039	60 kDa heat shock protein, mitochondrial (HSPD1)	3.09	<0.01	6
F1M798	Bone morphogenetic protein 1 (BMP1)	3.07	<0.01	20
F1LWC2	Uncharacterized protein	3.06	<0.01	4
F1LQS3	60S ribosomal protein L6 (RPL6)	3.06	<0.01	6
Q3MHS9	Chaperonin containing Tcp1, subunit 6A (Zeta 1) (CCT6A)	3.06	0.01	3
F1LYX9	Desmoglein-2 (DSG2)	3.05	<0.01	4
Q6P3V8	Eukaryotic translation initiation factor 4A1 (EIF4A1)	3.04	0.01	3
D3ZS68	Poly(rC)-binding protein 1 (PCBP1)	3.03	0.01	2
Q1EHB3	A disintegrin and metalloproteinase with thrombospondin motifs 7 (ADAMTS7)	3.03	<0.01	14
G3V7C6	Tubulin beta-4B chain (TUBB2C)	3.00	<0.01	2
P28480	T-complex protein 1 subunit alpha (TCP1)	2.98	0.01	4
Q68FR6	Elongation factor 1-gamma (EEF1G)	2.95	<0.01	4
P19945	60S acidic ribosomal protein P0 (RPLP0)	2.95	0.01	2
P12001	60S ribosomal protein L18 (RPL18)	2.92	<0.01	5
D3ZC88	Elongation factor 1-alpha (EF1A)	2.89	0.01	7
P11030	Acyl-CoA-binding protein (DBI)	2.87	<0.01	2
Q99MZ8	LIM and SH3 domain protein 1 (LASP1)	2.85	<0.01	3
Q6AYQ9	Peptidyl-prolyl cis-trans isomerase (PPIC)	2.82	<0.01	3
B2GV92	Prostaglandin E synthase 3 (PTGES3)	2.81	0.03	2
D3ZEI0	Protein LOC690096	2.79	<0.01	3
F1MAC8	Pyruvate kinase (PKM2)	2.75	<0.01	2
F1LQS4	ATP-citrate synthase (ACLY)	2.73	0.03	3
Q9R0C5	N-acetylgalactosaminyltransferase 7 (GALNT7)	2.68	<0.01	4

G3V8R1	Nucleobindin 2 (NUCB2)	2.66	<0.01	8
P11980	Pyruvate kinase isozymes M1/M2 (PKM2)	2.63	0.01	5
P51635	Alcohol dehydrogenase [NADP(+)] (AKR1A1)	2.61	0.02	2
D4AE41	RNA binding motif protein, X-linked-like-1 (RBMXL1)	2.60	0.02	2
D4ACJ7	Heterogeneous nuclear ribonucleoprotein A1 (HNRNPA1)	2.59	0.01	4
E9PTS0	Heterogeneous nuclear ribonucleoprotein A/B (HNRNPAB)	2.58	<0.01	4
F2Z3S4	60S ribosomal protein L11 (RPL11)	2.57	0.01	2
Q794E4	Heterogeneous nuclear ribonucleoprotein F (HNRNPF)	2.57	0.01	2
P82995	Heat shock protein HSP 90-alpha (HSP90AA1)	2.52	<0.01	7
D3ZTK5	60S ribosomal protein L27 (RPL27)	2.51	<0.01	2
Q6TUG0	DnaJ homolog subfamily B member 11 (DNAJB11)	2.51	<0.01	3
Q9JLZ1	Glutaredoxin-3 (GLRX3)	2.50	0.03	3
P25113	Phosphoglycerate mutase 1 (PGAM1)	2.50	0.01	5
P26772	10 kDa heat shock protein, mitochondrial (HSPE1)	2.49	<0.01	3
F1LM73	Netrin receptor UNC5C (UNC5C)	2.45	<0.01	3
Q3T1J1	Eukaryotic translation initiation factor 5A-1 (EIF5A)	2.44	0.05	2
F1LT35	Protein RGD1564606	2.41	<0.01	3
Q4G097	Cyclic AMP-dependent transcription factor ATF-6 beta (ATF6B)	2.37	<0.01	3
D4ABT3	Protein CDV3 homolog (CDV3)	2.36	<0.01	3
F1LRX5	Spectrin beta chain, non-erythrocytic 1 (SPTBN1)	2.34	0.04	2
F1LSS8	Neuropilin-2 (NRP2)	2.28	<0.01	16
F1MAP4	Multiple epidermal growth factor-like domains protein 10 (MEGF10)	2.28	0.02	3
D3ZC07	Serine/threonine-protein kinase N3 (PKN3)	2.25	<0.01	2
D3ZD13	Filamin-B (FLNB)	2.25	0.02	3
Q63610	Tropomyosin alpha-3 chain (TPM3)	2.23	<0.01	2
P07943	Aldose reductase (AKR1B1)	2.23	0.02	3
P69897	Tubulin beta-5 chain (TUBB5)	2.22	0.03	3

G3V836	Clusterin (CLU)	2.21	<0.01	11
Q7TPB1	T-complex protein 1 subunit delta (CCT4)	2.21	<0.01	2
F1LRV4	Heat shock 70 kDa protein 4 (HSPA4)	2.21	0.02	8
P63245	Guanine nucleotide-binding protein subunit beta-2-like 1 (GNB2L1)	2.21	0.02	5
P63018	Heat shock cognate 71 kDa protein (HSPA8)	2.21	<0.01	11
F1LP72	Elongation factor 1-delta (EEF1D)	2.20	0.04	5
D3ZPJ7	Clathrin light chain A (CLTA)	2.20	<0.01	2
O35814	Stress-induced-phosphoprotein 1 (STIP1)	2.20	0.01	12
A7VJC2	Heterogeneous nuclear ribonucleoproteins A2/B1 (HNRNPA2B1)	2.18	<0.01	6
D3ZW98	Thioredoxin-like protein 1 (TXNL1)	2.17	0.01	3
P62859	40S ribosomal protein S28 (RPS28)	2.12	<0.01	3
P07151	Beta-2-microglobulin (B2M)	2.11	<0.01	4
Q642B0	Glypican 4 (GPC4)	2.04	<0.01	12

5.2.4 Molecular characterisation of secreted AGR2

AGR2 was identified in the CM of WT AGR2-expressing cells in the experiments outlined above, and a number of studies have also reported secretion of AGR2 [301, 388, 426, 440]. Potential evidence that secretion of AGR2 may be of functional importance has been demonstrated in part through a proteomic screen using a ΔKTEL AGR2 mutant, which is highly secreted [438]. Here it was found that ΔKTEL AGR2 appears to have pro-proliferative effects as well as potentially promoting protein synthesis and secretion, but whether this has any relevance to WT AGR2 is uncertain. In addition, one study found that AGR2-containing CM from AGR2-expressing pancreatic cancer cells could promote proliferation of non-expressing cells, a feature that was lost upon AGR2 silencing [440]. However, whether this relates directly to extracellular AGR2 is also not clear. The question still remains however, of whether appearance of AGR2 in the extracellular space is an active process or simply the result of accumulation of intracellular AGR2 from dead and dying cells. To address this question, a stable Rama 37 cell line expressing the same mutant ΔKTEL AGR2 mentioned above (ΔKTEL clone 8), previously created for the present studies to explore the effects of AGR2 localisation on its function (Chapter 3), was used to explore the characteristics of secreted ΔKTEL AGR2.

5.2.4.1 Secreted AGR2 has a higher molecular mass than intracellular AGR2

In Chapter 3, stable cell lines expressing a highly secreted form of AGR2 (ΔKTEL AGR2) were constructed in order to investigate the effects of AGR2 secretion on its function. Very little intracellular ΔKTEL AGR2 protein could be detected in these cells, so to examine whether this was a result of increased secretion of the ΔKTEL mutant of AGR2, the level of intracellular and secreted AGR2 was compared between WT AGR2 clone 2 cells and ΔKTEL AGR2 clone 8 cells (Fig. 3.9). Interestingly, when directly comparing secreted and intracellular AGR2, the secreted protein band was 2-3 kDa larger than the intracellular protein band (Fig. 5.4). This small increase in mass suggests that secreted ΔKTEL AGR2 undergoes some post-translational modification (PTM).

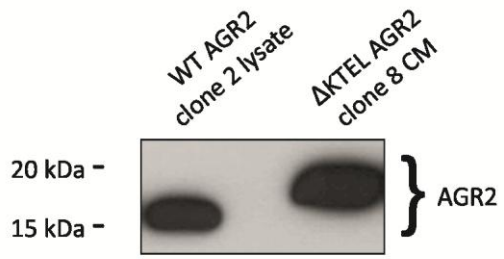


Figure 5.4. ***Extracellular AGR2 has an apparent higher molecular mass than intracellular AGR2*****.** Whole cell lysate from WT AGR2 clone 2 cells and CM from ΔKTEL AGR2 clone 8 cells were collected and probed for the presence of AGR2 by Western blot.

5.2.4.2 Higher molecular weight AGR2 is also secreted under non-serum starved conditions

For reasons outlined in section 5.2.2.1, the detection of proteins of interest in conditioned medium requires incubation of cells in serum-free medium. To ensure that the shift in AGR2 mass observed in secreted protein was not a consequence of incubation of cells in serum-free medium, AGR2 was immunoprecipitated from the CM of WT AGR2 clone 2 cells and ΔKTEL clone 8 cells collected in normal, fully supplemented growth medium (Fig.5.5).

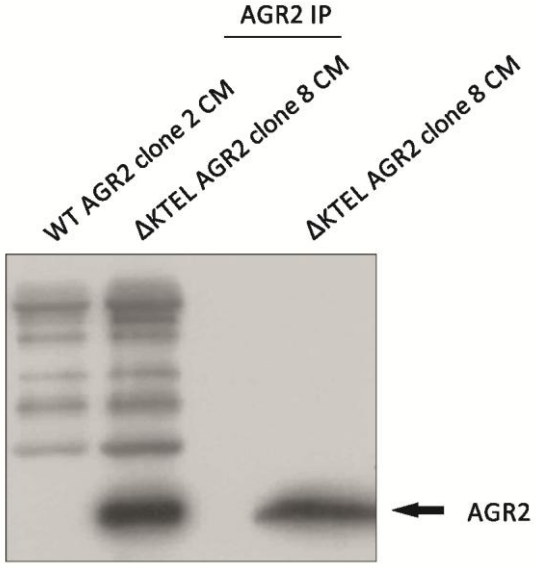


Figure 5.5. ***Higher molecular weight AGR2 is not necessarily induced by serum-starvation*****.** Conditioned medium was collected from WT AGR2 clone 2 cells and ΔKTEL clone 8 cells in fully supplemented growth medium. Secreted AGR2 was immunoprecipitated and compared to higher molecular weight secreted AGR2 from CM of ΔKTEL clone 8 cells in serum-free medium.

Extracellular AGR2 immunoprecipitated from ΔKTEL clone 8 CM had the same mass as AGR2 seen in concentrated ΔKTEL 8 CM collected in serum-free medium, indicating that the increase in mass of AGR2 was not caused by the use of serum-free medium. There was no detectable extracellular AGR2 immunoprecipitated from WT AGR2 clone 2 cells, but this may be the result of relatively low levels of secreted AGR2 from these cells (Table 5.3). The abundance of extracellular/secreted AGR2 from ΔKTEL clone 8 cells makes this cell line an ideal tool for investigating the nature of the modification to AGR2.

5.2.4.3 Determination of post translational modifications in secreted AGR2

The small increase in molecular mass observed with the secreted ΔKTEL AGR2 mutant suggested that it had undergone some post-translational modification (PTM). A bioinformatics-based study of all proteins within the Swiss-prot database found that the two most commonly predicted PTMs (based on prediction by similarity in the Swiss-prot database) were N-linked glycosylation (O-glycosylation being the 8th most common) followed by phosphorylation. For experimentally validated PTMs, phosphorylation was the most common, followed by acetylation and then N-linked glycosylation, with O-linked glycosylation the seventh most common [650]. It is well known that most secreted proteins are glycosylated [651], and phosphorylation of secreted proteins, which can be mediated through secreted kinases [652] is known to occur [653, 654]. Acetylation of secreted proteins is relatively rare [655], therefore, it was reasoned that the most likely cause of the shift in mass of secreted AGR2 was either glycosylation or phosphorylation.

Using several PTM prediction tools, the likelihood of these PTMs in AGR2 was assessed. As expected, there were no acetylation sites predicted for AGR2 [656], but several potential phosphorylation sites were identified, along with several O-linked (but not N-linked) glycosylation sites (Appendix 25 and 26). This prediction data pointed to the shift in mass most likely being caused by a phosphorylation or O-linked glycosylation event.

To test this hypothesis, CM from ΔKTEL clone 8 cells was collected and treated with λ-phosphatase (λPPase) to remove any phosphate groups attached to secreted AGR2 (Fig. 5.6). In order to ensure that λPPase was active in the concentrated CM, the samples were spiked with the multi-phosphorylated protein Mps1 (monopolar spindle 1) and a decrease in its mass monitored as an indication of active λPPase.

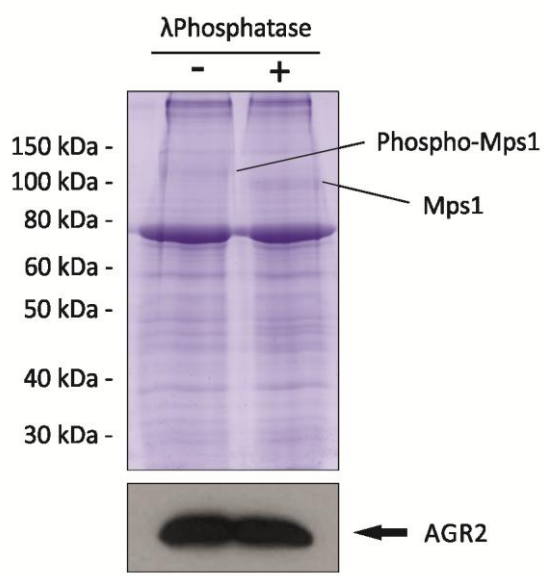


Figure 5.6. ***Secreted AGR2 is not phosphorylated*.** Conditioned medium (CM) from ΔKTEL AGR2 clone 8 cells (containing secreted AGR2) was concentrated by centrifugal filter and treated with λ-phosphatase (λPPase) to dephosphorylate any phosphorylated proteins. To ensure the concentrated CM did not inhibit λPPase activity, the phoshoprotein Mps1 (molecular weight: 97 kDa unphnosphorylated) was added to CM to monitor λPPase activity. A weak lower molecular weight Mps1 band can be observed in the λPPase-positive lane, indicating that the enzyme remains active in concentrated CM. A Western blot was performed on the same samples but loaded in different wells to the Coomassie-stained wells, and for this reason the Western blot lanes and gel lanes do not completely align.

Treatment with λPPase had no effect on the molecular mass of the secreted AGR2 band. There was a decrease of approximately 10 kDa in the molecular mass of the exogenously added Mps1 protein with λPPase treatment, consistent with previous reports of the change in Mps1 molecular mass following dephosphorylation [657]. This suggests that λPPase enzyme was active in the concentrated CM and suggests that the shift in mass for secreted AGR2 is not caused by phosphorylation.

The possibility of AGR2 glycosylation was subsequently investigated, initially through the use of periodic acid-Schiff (PAS) staining. Intracellular and secreted AGR2 was immunoprecipitated from WT AGR2 pool cell lysate and concentrated KTEL clone 8 CM, respectively. The immunoprecipitated proteins were then run on an SDS-PAGE gel and subject to PAS staining (Fig. 5.7). Recombinant AGR2 from *E. coli* and culture medium containing large amounts of the glycosylated protein transferrin were used as negative and positive controls, respectively.

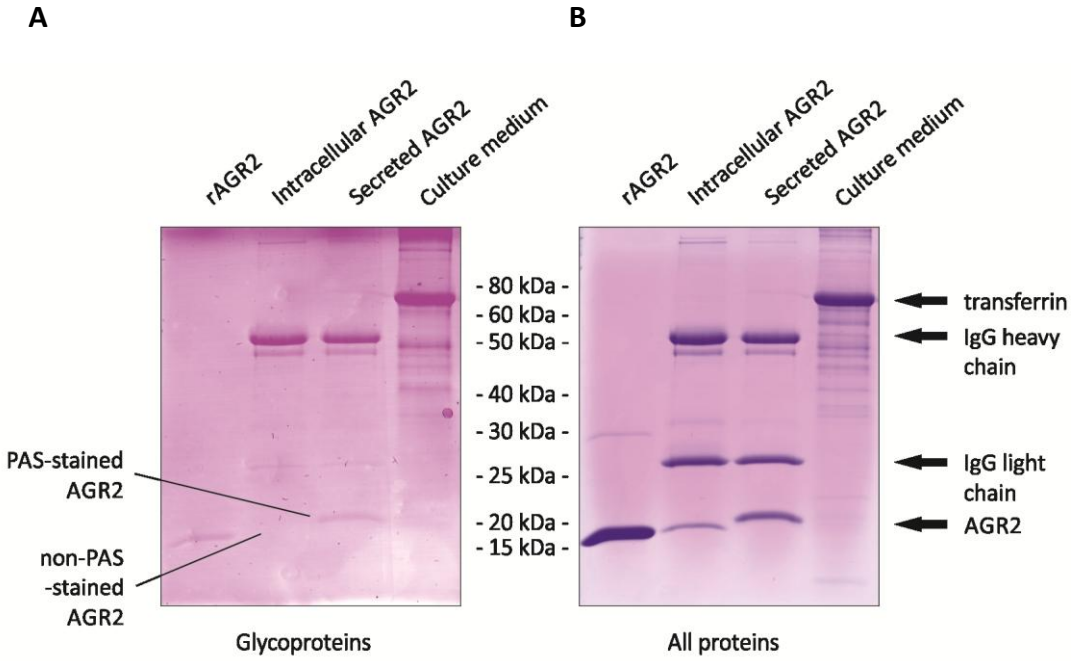


Figure 5.7. ***Secreted AGR2 contains glycan chains*.** Intracellular and secreted AGR2 were immunoprecipitated from WT AGR2 pool cell lysate and ΔKTEL clone 8 conditioned medium (CM), respectively. Proteins were then run on an SDS-PAGE gel and **(A)** stained for the presence of glycoproteins using the periodic acid-Schiff stain method (see Materials and Methods) or **(B)** counterstained with Coomassie stain to show all proteins.

Strong Schiff staining was seen for transferrin from CM (Fig. 5.7A), as well as for the IgG heavy chain in both immunoprecipitated lanes (approx. 50 kDa) resulting from their glycosylation. Secreted AGR2 but not intracellular AGR2 was stained in the PAS reaction, suggesting that secreted AGR2 is indeed glycosylated, although it is possible that the lower amount of intracellular AGR2 present in the gel (Fig. 5.7B) may fall below the detection limit of the Schiff stain. Also, the small amount of Schiff-positive recombinant AGR2 is surprising, but recent reports indicate that *E. coli*-derived proteins can be O-glycosylated [658], and thus the weak staining here may represent some glycosylation of recombinant AGR2.

To provide stronger evidence for the glycosylation of secreted AGR2, lysate from WT AGR2 pool cells and concentrated ΔKTEL clone 8 CM were subjected to enzymatic deglycosylation using a commercially available kit (New England Biolabs). Intracellular AGR2 was unaffected by treatment with the deglycosylation mix, but secreted AGR2 was shifted back down to the mass of intracellular protein, confirming that the protein is glycosylated (Fig. 5.8A). However, as the deglycosylation mix contains glycosidases for both N- and O-

linked glycans, this data only provided evidence for the presence of glycosylation, rather than N- or O-linked glycosylation specifically.

To try to determine whether the AGR2 was N- or O-glycosylated (or indeed both), secreted AGR2 was treated with PNGase F, an N-glycan-specific endoglycosidase that cleaves off the entire N-linked glycan at the Asp-GlcNAc linkage [659]. Once again, to ensure the concentrated ΔKTEL clone 8 CM did not interfere with PNGase F activity, its ability to remove N-glycans from spiked α-acid glycoprotein (α-AG) was tested (Fig. 5.8B).

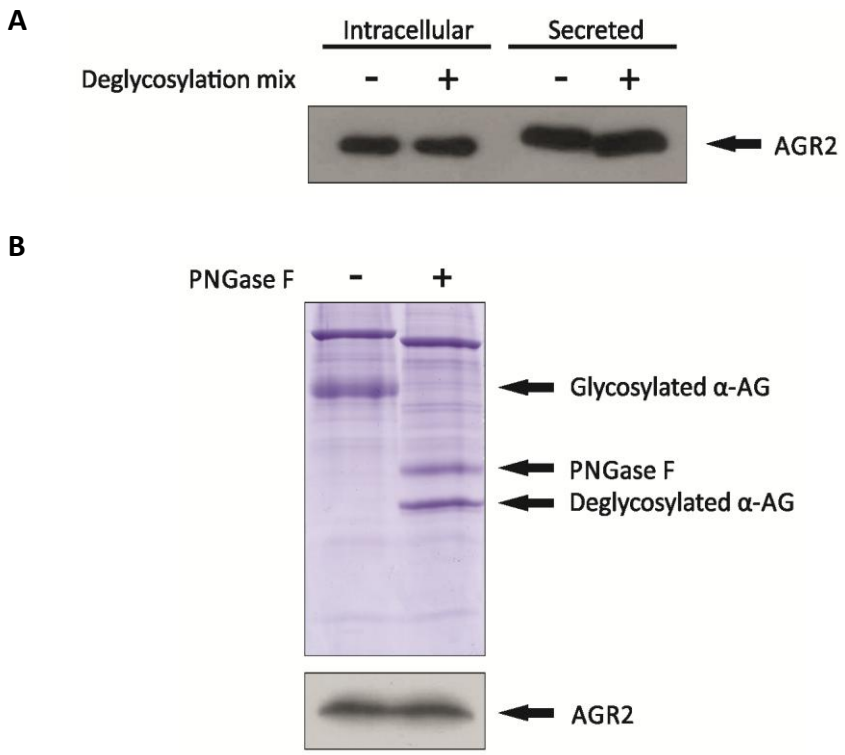


Figure 5.8. ***Secreted AGR2 contains O-linked but not N-linked glycan chains*****. (A)** Intracellular AGR2 from WT AGR2 pool cell lysate and secreted AGR2 from conditioned medium (CM) from ΔKTEL AGR2 clone 8 cells were subjected to enzymatic deglycosylation using a commercial kit containing PNGase F, O-glycosidase, neuraminidase, β1-4 galactosidase and β-N-acetylglucosaminidase. The activities of the enzymes appeared to be unaffected by either lysis buffer or concentrated CM as the deglycosylation of fetuin in these buffers was unaffected (Appendix 27). **(B)** CM from ΔKTEL AGR2 clone 8 cells was subjected to PNGase F treatment. To ensure concentrated CM did not inhibit PNGase F, the N-glycosylated protein α-acid glycoprotein (α-AG) was added to concentrated CM to monitor PNGase activity. α-AG was deglysosylated in concentrated CM (as was transferrin), indicating that the enzyme remains active in concentrated CM.

Treatment with PNGase F produced no change in the mass of secreted AGR2, but the enzyme was active in the concentrated CM as it was able to deglycosylate both α-acid glycoprotein and transferrin. Together, this presents strong evidence that AGR2 is not N-glycosylated, but suggests instead that it is O-glycosylated upon its secretion from the cell.

5.2.4.4 Glycosylated AGR2 appears to be produced from only healthy cells

To find out whether glycosylation of AGR2 also occurred in human cells that naturally overexpress AGR2, the CM was collected from the MCF7A breast cancer cell line. MCF7A cells are particularly susceptible to cell death and lysis upon serum-starvation (personal observation), so this presented an opportunity to test whether AGR2 was indeed secreted in a relatively active process, rather than being released as a result of cell lysis. CM from MCF7A cells in increasingly nutrient-poor media were collected and probed for AGR2 (Fig. 5.9).

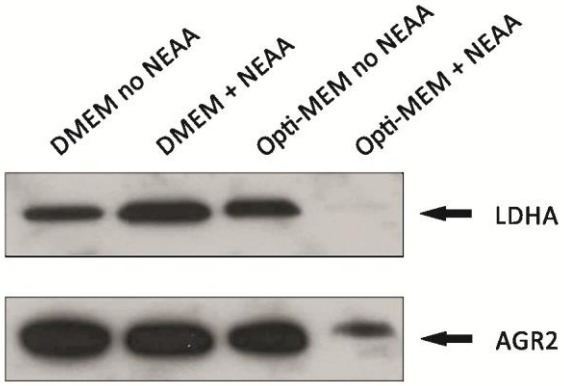


Figure 5.9. ***Glycosylated AGR2 is not released from lysed cells*.** MCF7A cells were incubated in the indicated serum-free media for 24 h. Conditioned medium (CM) was collected, precipitated and probed for the presence of AGR2 by Western blot. The levels of LDHA in these CM were used as a measure of the extent of cell lysis under each serum-free condition. NEAA: non essential amino acids.

The extent of cell lysis was monitored by the appearance of LDHA in the CM. As expected, MCF7A cells appeared to be quite susceptible to cell death upon serum starvation, as demonstrated by the large amount of LDHA found in CM from DMEM-incubated cells (+ and – non-essential amino acids (NEAA)) and also from Opti-MEM-incubated cell without NEAA. There was, however, very little LDHA release from Opti-MEM

+ NEAA-incubated cells, indicating a much lower level of cell lysis in this medium and thus a higher proportion of live cells. This suggests that these cells are very dependent on both exogenous amino acids and growth factors for survival.

AGR2 was present in the CM from cells under all four test conditions. The amount of extracellular AGR2 was approximately equal under all conditions except for Opti-MEM + NEAA-incubated cells, where there was substantially less protein but importantly, the protein showed the characteristic upward shift in mass associated with glycosylation. Coupled with the very low amounts of LDHA released from these cells, the presence of glycosylated AGR2 here suggests that the protein is only glycosylated and secreted from live cells. Likewise, as MCF7A cells produce large amounts of AGR2 [309], the high levels of LDHA and lower mass AGR2 in CM from the DMEM- and Opti-MEM with no NEAA-incubated cells suggests that at least some of this AGR2 has been released as a result of cell lysis.

5.2.5 AGR2 is not a cell surface-resident protein

A number of ER chaperone proteins are known to localise to the cell surface under certain conditions, including GRP78 [447, 660] and calreticulin [661] during tumourigenesis, and PDI during platelet activation and coagulation [446, 662]. Therefore, having established that AGR2 can also be secreted, the binding of AGR2 to the cell surface was investigated.

Three staining methods were compared for their ability to stain specifically for surface proteins. Paraformaldehyde (PFA) fixation with Triton X-100 permeabilisation was used as a control for intracellular protein staining, while PFA fixation without permeabilising and 'live cell IF' [520] were used to assess surface staining. The cell surface protein CD44 was used as a positive control for surface staining, as the antibody used recognised an extracellular epitope of the protein. Likewise, the cytoplasmic protein, MEK1, was used as a monitor of cell permeabilisation, as non-permeabilised cells should not stain for MEK1.

Permeabilised cells stained for AGR2, MEK1 and CD44 as expected (Fig. 5.10A and B). In non-permeabilised cells, the staining pattern of AGR2 was more punctuate and less uniform than in permeabilised cells, but these non-permeabilised cells also stained for MEK1 to the same extent as permeabilised cells. This suggests that the membrane integrity of these cells has been compromised and that they are in fact permeabilised, and so the AGR2 staining seen with these cells could be intracellular protein. However, cells stained

using the live-cell IF technique stained positively for the cell surface marker CD44 but were negative for both MEK1 and AGR2. This indicates that the cells remained impermeable to antibody and that the staining observed is in fact cell surface protein. Furthermore, the data indicate that AGR2 is not detectably resident on the cell surface and so any interactions with cell surface molecules are likely to be at least only transient events.

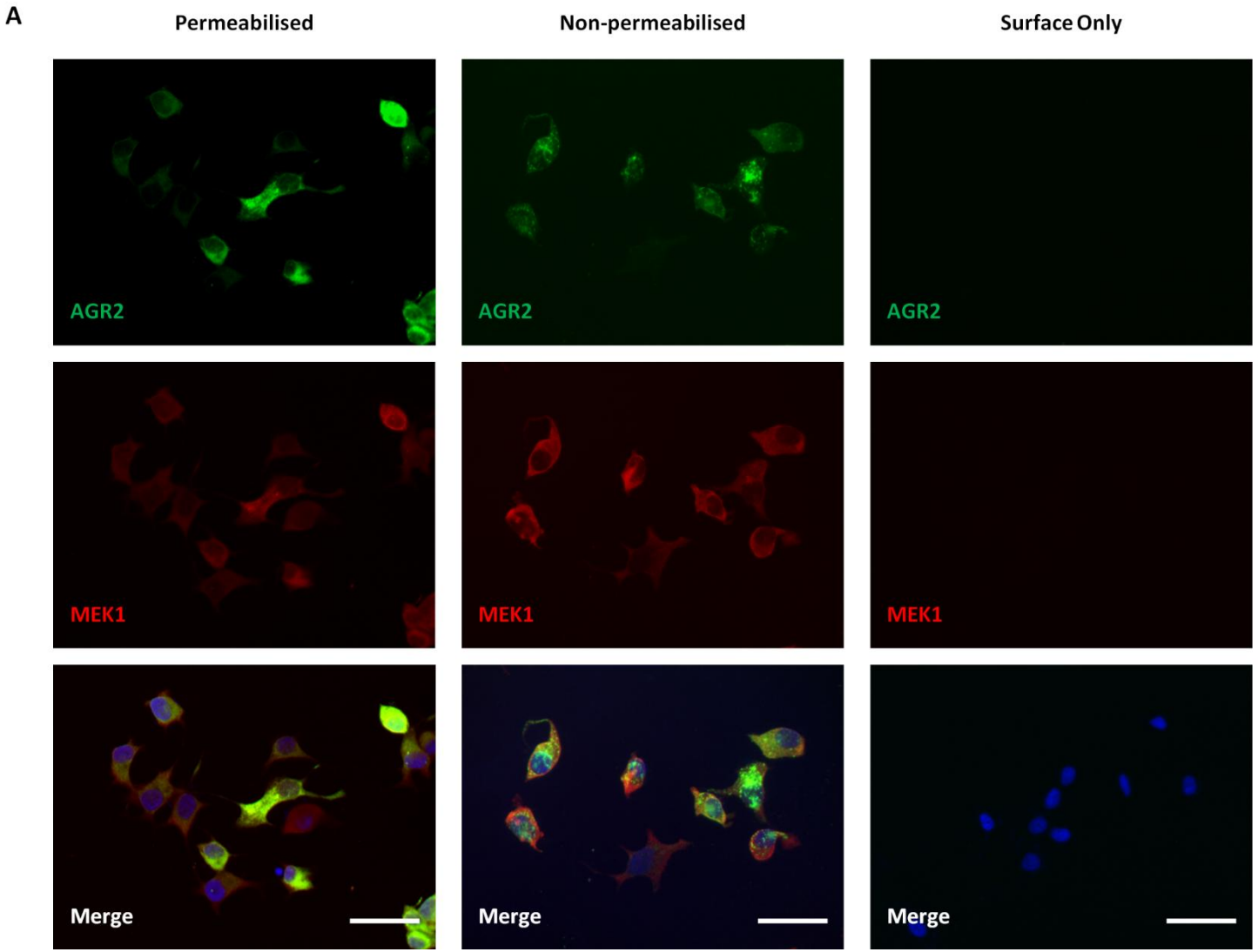


Figure 5.10. ***Comparison of MCF7A staining methods.*** For permeabilised and non-permeabilised staining, cells were fixed with 4 % PFA, permeabilised with 1 % triton in PBS as required, and then stained with primary and secondary antibodies. For surface only staining, cells were incubated on ice with primary and then secondary antibodies, prior to fixation with 4% PFA. **(A)** Cells were co-stained with AGR2 (green) and the cytoplasmic marker MEK1 (red) to monitor cell permeability of all staining techniques. **(B)** Cells were co-stained with AGR2 (red) and the cell surface marker CD44 (green) as a positive control for cell surface staining. Nuclei (blue) were stained with DAPI. Scale bar = 50 µm.

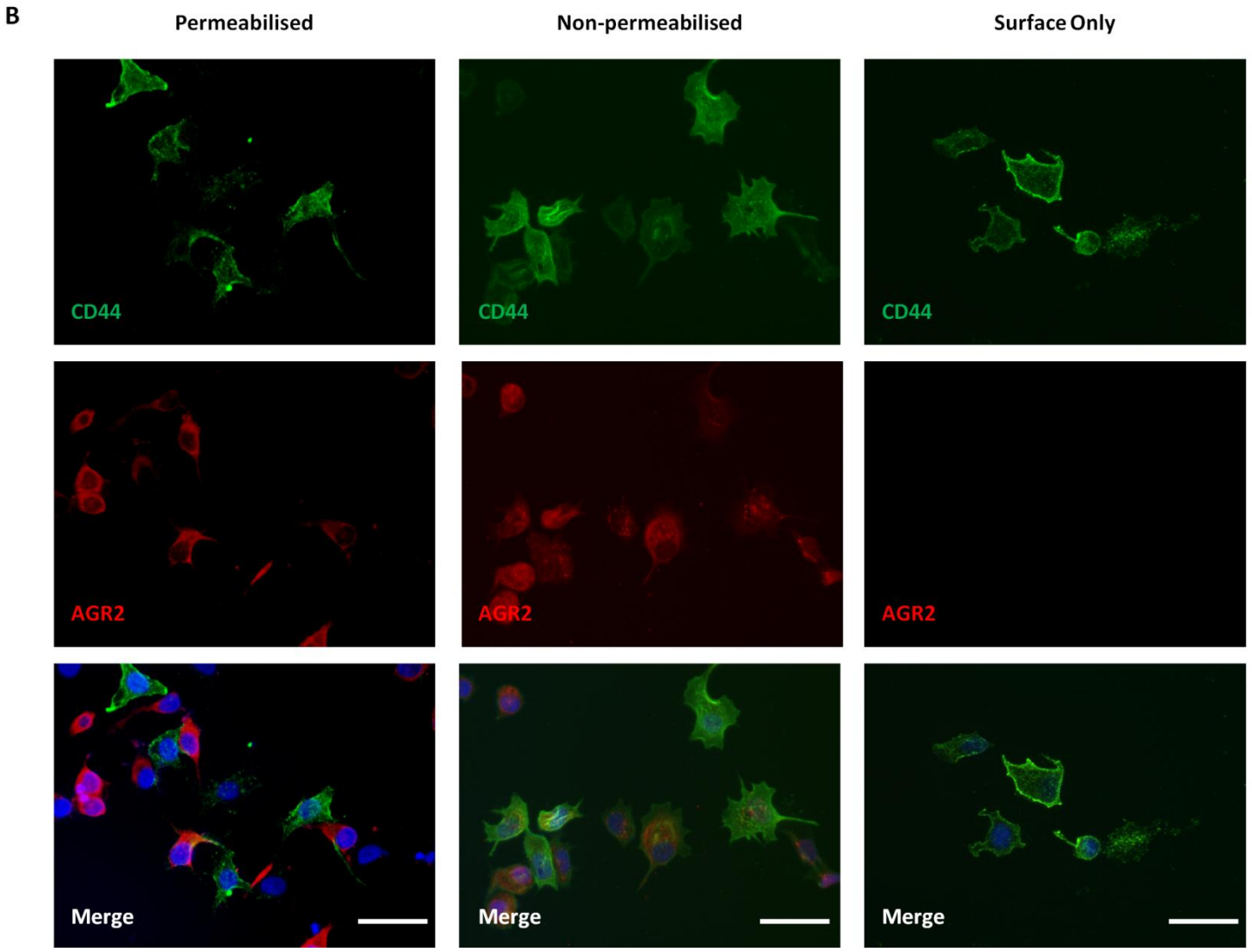


Figure 5.10. ***Comparison of MCF7A staining methods (cont.).*** For permeabilised and non-permeabilised staining, cells were fixed with 4 % PFA, permeabilised with 1 % triton in PBS as required, and then stained with primary and secondary antibodies. For surface only staining, cells were incubated on ice with primary and then secondary antibodies, prior to fixation with 4% PFA. **(A)** Cells were co-stained with AGR2 (green) and the cytoplasmic marker MEK1 (red) to monitor cell permeability of all staining techniques. **(B)** Cells were co-stained with AGR2 (red) and the cell surface marker CD44 (green) as a positive control for cell surface staining. Nuclei (blue) were stained with DAPI. Scale bar = 50 µm.

5.3 Discussion

5.3.1 AGR2-induced changes to the cell proteome

It has been shown here that AGR2 does not appear to have a strong effect on the abundance of cellular proteins, at least in a pooled population of cells. This is perhaps not unexpected, given the ER-localisation of AGR2 and the largely extracellular (or at least post-ER) nature of its substrates reported to date, including mucins and AREG [301, 408, 412, 418, 427, 506], and the lysosomal proteases cathepsins B and D [437]. Of course, there may be influences on signalling pathways through PTMs induced by AGR2, as previously shown with the AGR2-dependent de-phosphorylation of the Hippo pathway co-activator YAP-1 [506], but this is a question for a separate study.

The changes to the intracellular proteome therefore most likely reflect indirect changes induced by AGR2, rather than direct regulation of particular substrate molecules. As mentioned above, the downregulation of CRABP1 and glypican-4 in WT AGR2 pool cells have been reported to promote cell dedifferentiation and migration, through influences on retinoic acid and the Wnt pathway, respectively [626, 629]. The change in level also observed in the extracellular signalling protein WISP2 (Wnt-inducible secreted protein 2), a Wnt activator [663], in AGR2-expressing cells (Table 5.3) may point to a link between AGR2 and Wnt signalling, at least in Rama 37 cells.

The effects of AGR2 on secreted proteins were much more substantial, particularly in the context of clonal cells. It is possible that many of the downregulated proteins in WT AGR2 clonal cells relative to EV clonal cells simply relates to differences in levels of cell lysis between these cell types. This may account, in part, for the differences in the sets of proteins that are differentially expressed between WT AGR2 clonal and pooled cells. Additionally, the differences in cell morphologies between clonal cells also likely have fundamental effects on the cell secretome, given that EMT-like transformations encompass a strong element of altered transcription (reviewed in [210]) and thus would be expected to have more a more fundamental effect on the proteome than the presumably post-translational chaperone activity of AGR2. There were, however, some similarities in the classes of proteins differentially regulated in WT AGR2 clonal and pooled cells, e.g. laminins were downregulated in both cell types, whereas ADAMTS proteins were generally upregulated. Both MEGF10 (multiple epidermal growth factor-like domains protein 10), which appears to be a pro-adhesive protein [664, 665], and glypican-4 were downregulated

in both clonal and pooled cells lines, but as membrane proteins, it is not clear whether differences in abundances of these proteins in the secretome reflect differences in membrane levels or differences in soluble fragments. As a result, the multiple changes to the secretome might reflect a global chaperone function for AGR2, similar to BiP [310], where specific client proteins might be determined simply by the proteins that are produced by a given cell (e.g. as a result of transcriptional changes in EMT). With this is mind however, AGR2 most likely has some specific substrates as well, as overexpression of AGR2 alone is able to induce metastasis [309]. However, these substrates still remain elusive.

In the absence of truly common proteins between clonal and pooled cells, it is important to note that the vast majority of differentially regulated proteins contain at least one disulphide bond, which is perhaps not unexpected given the prevalence of this PTM in secreted proteins [312]. However, given the recorded thiol-dependent interactions between AGR2 and at least one substrate proteins [422], this does provide a means whereby AGR2 might directly interact with these proteins to affect the extracellular abundance. The fact that downregulated proteins also contain disulphides might suggest that AGR2 interaction with these proteins results in their degradation, possibly through ERAD (ER-associated degradation). PDI has been shown to promote ER retention of procollagen I and its subsequent proteasomal degradation, an activity that was lost upon deletion of the PDI ER-retention sequence and thus exit from the ER, suggesting that it is indeed the retention in the ER that promotes the degradation [666]. Similarly, ERp44, which contains a single-cysteine active-site motif like AGR2 [315], mediates retention of incompletely folded proteins in the ER [320, 573], although whether this also results in the activation of ERAD has not been determined. Thus while AGR2 clearly promotes the secretion of mucins [408, 412, 418, 422, 427], it might also downregulate the secretion of other proteins, although how this process might be regulated is unclear.

5.3.2 Relating proteomic changes to cell adhesion

Using IPA analysis, differentially regulated protein in clonal WT AGR2 and EV cells related to cell adhesion were investigated, in order to relate possible proteomic changes to the reduced attachment behaviour of AGR2 positive cells in monolayer culture observed in Chapter 4. Proteins potentially related to ‘decreased adhesion of tumour cells’ are shown in Table 5.4.

Uniprot Accession	*Description*	*Fold Change*	*Literature reports (number of publications)*
Q9ERB4	Versican core protein (VCAN)	30.68	Decreases (3)
P17491	Inhibin beta B (INHBB)	6.85	Decreases (8)
P08699	Galectin-3 (LGALS3)	-3.23	Increases (4)
F1LST1	Fibronectin (FN1)	-3.26	Increases (22)
F1M6X1	Vascular endothelial growth factor A (VEGFA)	-3.84	Increases (7)
P08721	Osteopontin (SPP1)	-16.38	Increases (5)
F1LRH4	Laminin subunit gamma-2 (LAMC2)	-16.39	Increases (6)
G3V6C1	Serpin B5	-24.28	Increases (2)

Table 5.4. ***Differentially expressed secretome proteins associated with the term 'decreased adhesion of tumour cells' by IPA.*** Both up- and down-regulated proteins in the secretome of WT AGR2 versus EV clonal cells were analysed by IPA in relation to possible functions in cell adhesion. Max fold change is expressed relative to EV clone 2 cells. Literature reports refers to the reported effect of the relevant protein on adhesion in the literature, and the number of publications identified by IPA that support this role.

In this way, it appears that both up- and down-regulation of secreted protein in WT AGR2 clonal cells can affect their adhesion. In particular, the downregulation of fibronectin and LAMC2 might explain why monolayer cells were much less adherent than EV clone cells, producing lower amount of matrix for cells to adhere to. In a similar vein, versican has been shown to be an anti-adhesive protein [667, 668], which prevented the adhesion of prostate cancer cells to fibronectin but not laminin [669], suggesting that reduced cell adhesion in WT AGR2 clonal cells might be the result of weakened or inhibited cell-matrix interactions.

IPA was also used to identify changes in proteins that might account for the higher attachment rate of WT AGR2 clonal cells relative to EV clone cells (Chapter 4). Whereas most proteins involved in adhesion identified by IPA related to decreased adhesion (as above), a small group of proteins associated with 'increased adhesion of connective tissue' was also identified (Table 5.5).

Uniprot Accession	*Description*	*Fold Change*	*Literature reports (number of publications)*
P02454	Collagen alpha-1 (I) chain (COL1A1)	24.88	Increases (1)
D3ZE04	Collagen alpha-1 (VII) chain (COL7A1)	13.63	Increases (3)
G3V7F3	WNT1 inducible signaling pathway protein 2 (WISP2)	12.66	Increases (1)
P43145	Adrenomedullin (ADM)	5.66	Increases (5)

Table 5.5. ***Differentially expressed proteins associated with the term 'increased adhesion of connective tissue' by IPA.*** Both up- and down-regulated proteins in the secretome of WT AGR2 clonal cells were analysed by IPA in relation to possible functions in cell adhesion. Max fold change is expressed relative to EV clone 2 cells. Literature reports refers to the reported effect of the relevant protein on adhesion in the literature, and the number of publications identified by IPA that support this role.

Whilst changes to collagen molecules are probably related to longer term adhesion effects, adrenomedullin has been shown to stimulate an enhanced rate of adhesion in prostate cancer cell lines [670], and thus might play a role in cell attachment in Rama 37 cells. It is also interesting that these molecules were identified as promoting adhesion of connective tissue (and thus presumably mesenchymal-type cells), which would be consistent with the change in morphology observed in WT AGR2 clonal cells. It will be interesting to determine if these proteins directly affect the adhesion behaviour of WT AGR2 clonal cells and thus may warrant further investigation as possible substrates of AGR2.

5.3.3 AGR2-influenced regulatory and interaction networks

In order to identify potential regulatory pathways that might explain some of the changes observed in AGR2-expressing cells, the secretome data were analysed by IPA. Upstream analysis revealed that two proteins upregulated in WT AGR2 clone cells, IGF2 (insulin-like growth factor 2) and WISP2, were themselves upstream regulators of several proteins differentially expressed between AGR2-positive and -negative cells: IGFBP2, IGFBP7, LAMB3, LAMC2, CTGF, Serpine 1, VEGFA and fibronectin (Fig. 5.11). WISP2 is a regulator of the majority of these proteins and may provide clues to a possible mechanism whereby AGR2 induces pro-metastatic effects. Aside from the favourable downregulation of laminin proteins observed here (in the context of promoting metastasis), secreted WISP2

has been shown to activate the canonical Wnt pathway, causing partial dedifferentiation of adipocytes and promoting the assumption of a myofibroblast phenotype, associated with increased migration [663, 671]. WISP2 overexpression has been reported in both pre-neoplastic and neoplastic breast lesions [672] but equally, studies have shown that it is a protective factor against EMT [673, 674]. These differences in WISP2 function might reflect the influence of the tissue microenvironment on Wnt signalling (which can be a source for Wnt ligands) [675, 676], relative to one Wnt ligand in isolation. Accordingly, it is important to note that connective tissue growth factor (CTGF), which was downregulated in CM from WT AGR2 clone cells and is also regulated by WISP2 (Fig. 5.11), modulates Wnt pathway signalling [677] and thus could play a role in determining the outcome of WISP2 signalling.

Given its role in EMT, differentiation and migration [41, 678], influences on the Wnt pathway through AGR2 provide an attractive mechanism whereby AGR2 might promote the acquisition of a migratory and invasive phenotype. AGR2 has been associated with the Wnt pathway through its repression of the Sox9 transcription factor in gastric mucous cells, which itself represses several Wnt ligands [408], although no direct action of AGR2 on the Wnt pathway was explored. It will be interesting therefore to investigate changes in intracellular Wnt pathway proteins in WT AGR2 clonal cells, to determine whether Wnt signalling might account for the differential regulation of other proteins in these cells, as well as phenotypic alterations, thus potentially placing an AGR2-Wnt axis at the heart of AGR2-induced metastasis. It is also interesting that Wnt ligands tend to be cysteine-rich proteins, where correct disulphide bond formation is crucial for function, at least in the case of Wnt3a [679]. This indirectly suggests a role for PDI proteins in the regulation of the Wnt pathway through the generation of active Wnt ligands. However, given the morphological change in WT AGR2 clonal cells, it remains to be determined whether WISP2 expression is directly governed by AGR2, or whether it changes as a result of (or indeed a cause of) the change in morphology.

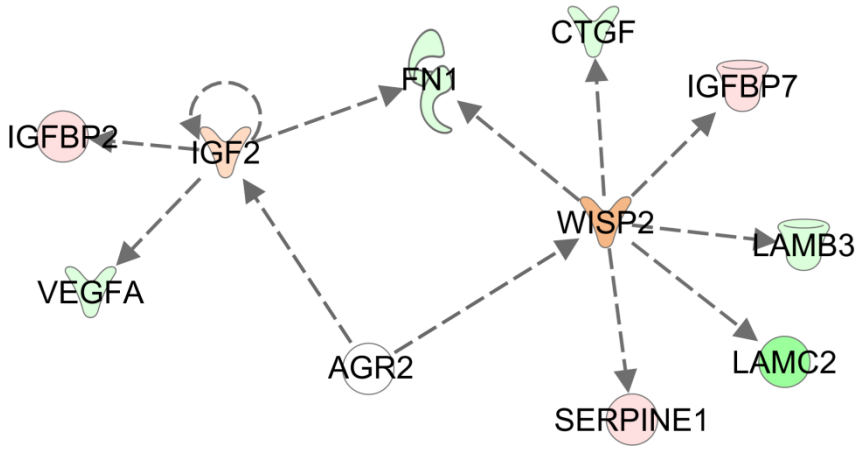


Figure 5.11. ***Top regulatory network of AGR2-influenced proteins in clonal cells.*** IGF2 and WISP2 were differentially expressed in AGR2-positive and negative clonal cells lines, and amongst these differentially expressed proteins, both were identified by Ingenuity Pathway Analysis as the top regulatory proteins. Red shading indicates upregulation in AGR2 clonal cells and green shading a downregulation. The intensity of shading reflects the intensity of change.

The contribution of IGF2 to the regulation of proteins differentially expressed in WT AGR2 clonal cells may be less significant given the few proteins it appears to regulate in these cells (Fig. 5.11) but again, as only secreted proteins are represented here, it is unclear what effects IGF2 might have intracellularly. Nevertheless, overexpression of IGF2 is associated with poorer disease prognosis and is seen in many cancers (reviewed in [680]), and the increased abundance of IGF2 and the IGF-binding proteins (IGFBP-2, and -5, Table 5.3) could promote tumour progression through anti-apoptotic and anti-adhesive effects [681, 682]**.** Overall, it might be interesting to investigate the roles of Wnt and IGF signalling pathways as regulators of AGR2-mediated tumourigenesis.

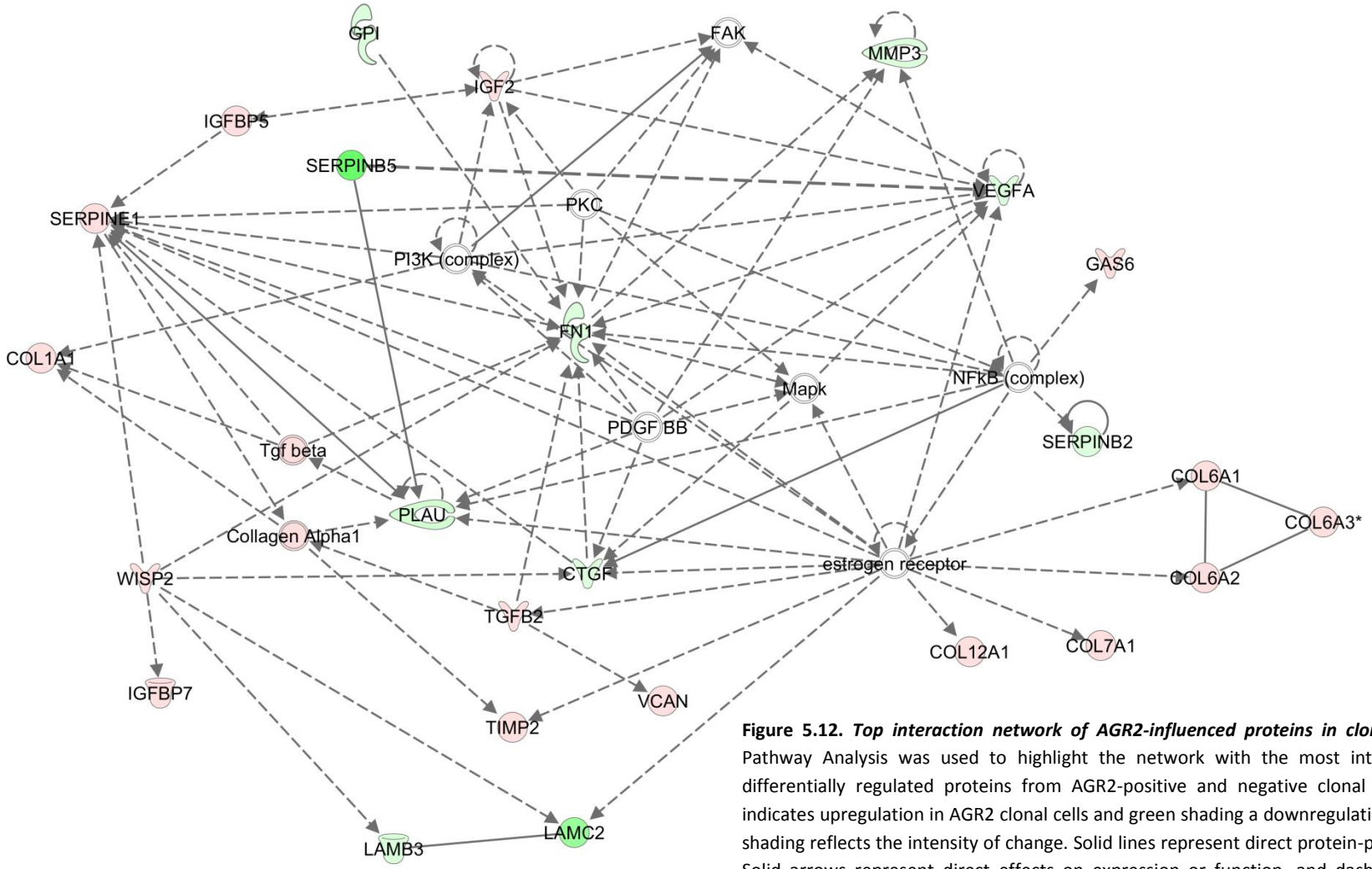


Figure 5.12. ***Top interaction network of AGR2-influenced proteins in clonal cells*****.** Ingenuity Pathway Analysis was used to highlight the network with the most interactions involving differentially regulated proteins from AGR2-positive and negative clonal cells. Red shading indicates upregulation in AGR2 clonal cells and green shading a downregulation. The intensity of shading reflects the intensity of change. Solid lines represent direct protein-protein interactions. Solid arrows represent direct effects on expression or function, and dashed arrows indirect effects.

WISP2 and IGF2 interactions are seen in the context of a broader landscape in the top interaction network (based on published literature) for secreted proteins from the clonal cells tested shown in Fig. 5.12. This figure displays both protein-protein interactions and influences on protein expression/function. This network seems to highlight the multitude of factors in the CM of WT AGR2 clonal cells that converge on modulating the expression of fibronectin. As discussed above, downregulation of fibronectin might partly explain why these cells were so easily detachable in monolayers (Chapter 4), as there is a relatively lower amount of fibronectin matrix for these cells to adhere to relative to EV clone cells. In a wider context, downregulation of fibronectin has been observed in some cancers, where downregulation seems to correlate with a more metastatic phenotype [683-685]. Thus AGR2-associated downregulation of fibronectin could contribute to cell adhesion and metastasis.

The network in Fig. 5.12 also highlights the effects of WISP2, as well as a number of other proteins, on two major components of the plasmin system, resulting in upregulation of serpine 1 (also known as plasminogen activator inhibitor-1 (PAI-1)),and downregulation of urokinase-type plasminogen activator (uPA/PLAU). This points to an overall downregulation of the plasmin system, which is somewhat unexpected given the association of increased plasmin activity with matrix degradation and cancer dissemination (reviewed in [686]). However, in the context of cell adhesion, increased levels of PAI-1 have been shown to promote integrin internalisation in a uPA and uPA receptor (uPAR)-dependent manner, which leads to increased cell detachment [687, 688]. This was found to occur on a variety of matrix proteins, including fibronectin, and thus the combined decrease in secreted fibronectin and anti-adhesive effects of increased PAI-1 in WT AGR2 clonal cells probably contribute to reduced cell adhesion. Again in a wider context, increased PAI-1 might increase tumour cell dissemination by promoting tumour angiogenesis [689, 690], and also through promoting fibrin formation and release of pro-coagulation factors from the endothelium [691, 692], leading to thrombus formation that promotes arrest of circulating tumour cells in the vasculature and aids in their extravasation [693, 694].

Thus several regulatory and effector proteins have been identified in the secretome of WT AGR2 clonal cells that might contribute to both the observed adhesion behaviour of these cells, and might also contribute to the pro-metastatic activity associated with AGR2.

5.3.4 Glycosylation of AGR2

It has been demonstrated here for the first time that AGR2 becomes glycosylated upon secretion as mutant protein from Rama 37 rat cells and as wild type protein secreted from human MCF7A cells, and that the glycosylation is not N-linked, but probably O-linked. Whilst it is possible that the KTEL deletion could be wholly or partially responsible for this glycosylation in Rama 37 cells, this seems unlikely given the observation of a similar shift in mass of secreted WT AGR2 from MCF7A cells. Furthermore, the predicted glycosylation site(s) are situated at the N-terminus of the protein (Appendix 26), away from the C-terminal KTEL sequence (Fig. 3.1). Thus the KTEL mutant probably highlights that glycosylation is a post-ER event, rather than causing an ectopic glycosylation event.

WT AGR2 is localised to the ER (Chapter 3) and several other ER resident chaperone proteins have been reported to be secreted or cell surface-bound [446, 447, 660-662, 695-697], and at least two of these (GRP78 and ERp44) are believed to be also O-glycosylated [447, 695-697]. However, while this indicates that glycosylation and secretion of ER proteins is not unique to AGR2, the consequences of these events in the context of AGR2 are largely unknown.

As O-glycosylation occurs within the Golgi [359], it is possible that secretion and glycosylation of ER proteins occurs due to lack of retrieval back to the ER through the C-terminal ER-retention sequence. This implies some sort of 'escape' from the ER. Ni and colleagues [698] suggest that, at least in the case of GRP78, this may occur through an overwhelming of the KDEL receptor system by overexpression of GRP78 protein, such that the retrieval system is simply not able to recover all molecules due to the their sheer number, and this is possibly the case in the AGR2-overproducing MCF7A cells. Alternatively, AGR2 may be secreted through interaction with a transport protein, in a manner akin to GRP78 where the proteins MTJ-1 and Par-4 appear to be required for its transportation to the cell membrane [699, 700]. A recent report showing AGR2 secretion into the gastrointestinal mucus also noted that C81S AGR2 was secreted from transfected CHO cells whereas WT AGR2 was not, possibly suggesting that Cys81 plays a role in ER retention and that masking it, either through mutation or substrate binding, could lead to AGR2 secretion [426]. The exact mechanism involved in AGR2 secretion will require further investigation.

It is also difficult to predict the role of secreted AGR2 when the role of intracellular AGR2 is still unclear, although data from Chapter 4 suggests that secreted AGR2 may still play a role in cell adhesion. Similarly, the effects of the possible O-glycosylation of AGR2 are not clear. The predicted glycosylation sites all lie within the flexible N-terminal region of

AGR2, between Thr23 and Ser45 (Appendix 26) and O-glycosylation is indeed preferentially found in unstructured protein regions [701]. The addition of glycan chains appears to enhance thermodynamic stability of proteins, by favouring a more folded protein state through steric hindrance of the glycan chain [702]. Reports and simulations have indeed suggested that glycosylation causes a loss of structural flexibility [703, 704] (again probably through steric hindrance), and one study has shown that N-glycosylation of a flexible loop region of porcine peptidase increases its structural stability [705]. Therefore, glycosylation of secreted AGR2 may serve to stabilise this flexible region of AGR2, effectively making it less flexible, but with unknown consequences for the activity of AGR2. It is interesting however that loss of flexibility in an unstructured region of the RNA chaperone α-crystallin resulted in reduction of its chaperone activity [337], possibly suggesting that glycosylated AGR2 may have compromised or altered substrate binding specificity.

Despite several reports that have noted that AGR2 is associated with the cell membrane [302, 443-445] and that it is also present on the cell surface [437], the present study could not find any evidence of membrane-bound AGR2. It is important to note however that most of the previously mentioned studies used immunohistochemical methods and did not necessarily observe extracellular AGR2 directly, possibly indicating that it was resident on the cytoplasmic side of the membrane. However, the study by Dumartin and colleagues [437] noted cell surface AGR2 expression based on data from both flow cytometry and immunofluorescence (IF) imaging. It is important to note that their reports of cell surface AGR2 using IF were based on cells fixed in PFA and then stained for AGR2, which has been demonstrated here to still stain for intracellular proteins. This 'cell surface' AGR2 may therefore be intracellular AGR2. The observation of cell surface AGR2 from flow cytometry is harder to explain, but due to the necessity of trypsinising adherent cells for flow cytometry analysis, it is possible that AGR2 antibody may have penetrated the cell during sample preparation and given a false positive reading of cell surface AGR2, as several reports have shown that trypsinisation affects membrane integrity and cell permeability [706, 707]. Several studies have employed the staining method used in the present study to demonstrate cell surface expression of proteins [708-710] and so the lack of membrane staining with AGR2 seen with three different AGR2 antibodies (personal observation) strongly suggests that secreted AGR2 does not have prolonged interaction with the cell surface, at least of MCF7A cells. Whether extracellular AGR2 behaves differently in different cell types (as in [437]) remains a possibility.

Overall, exploring the role of this secreted form of AGR2 presents an exciting new direction in the biology of AGR2, and it will be particularly interesting to examine the

biophysical characteristics of secreted AGR2, along with any potential interactions with cell surface proteins such as dystroglycan and C4.4a that have been previously reported in the yeast two hybrid system [389].

5.4 Conclusions

In this chapter, it has been demonstrated that AGR2 expression results in a number of changes to both the intracellular and extracellular proteome of Rama 37 cells, and these changes are more numerous between AGR2-positive and –negative clonal cells. The secretion of a glycosylated form of AGR2 (probably O-glycosylated) has also been described, although the consequences of this glycosylation remain presently unclear.

Chapter 6

General Discussion

The work presented here examined the role of the endoplasmic reticulum (ER) protein AGR2 in cell adhesion. AGR2 is overexpressed in a number of cancer types, and determining the role of AGR2 in cancer progression could be fruitful in identifying new drug targets in AGR2-positive cancers, given the association of AGR2 expression with reduced patient survival [309, 443, 457-462, 467].

AGR2 was demonstrated to be associated with an elongated cell morphology in clonal cell Rama 37 rat mammary cells, which in turn led to decreased adhesion of the cells in confluent monolayers. On the other hand, AGR2 expression was associated with an increase in the rate of attachment of single or clumps of cells. It was also demonstrated that both intracellular and extracellular AGR2 appear to affect cell adhesion.

Through NMR structural studies and mutagenesis, the dimerisation interface of AGR2 was identified, along with the key residues that mediate this dimerisation. Although

all the forces governing this association are not entirely clear, these experiments have permitted the generation of a monomeric form of AGR2, enabling the study of the role of dimerisation in AGR2 activity, as well as several other potentially important domains. Through these experiments, the underlying features defining the adhesion phenotype of Rama 37 cells appears to be cell morphology, but the AGR2 active site Cys residue, as well as AGR2 dimerisation, appear to be the most important features that promote AGR2-associated morphology changes. Additionally, through the use of a highly secreted mutant form of AGR2 (ΔKTEL AGR2), it has been demonstrated here that AGR2 becomes glycosylated upon its secretion from Rama 37 cells, and WT AGR2 from human MCF7A breast cancer cells shows a similar change in mass upon secretion. While this glycosylation is probably O-linked in nature, the significance of its attachment to secreted AGR2 remains unclear.

Finally, a proteomics approach was utilised to identify changes to the proteome of AGR2-expressing cells that might account for the role of AGR2 in adhesion and metastasis. These experiments revealed that AGR2-expressing clonal cells secrete decreased levels of ECM proteins relative to non-AGR2-expressing cells, which might explain why these cells are generally less adherent when grown as a monolayer. The reason for the increased rate of attachment of clonal AGR2 cells is not as evident, although at least one protein (adrenomedullin) was identified that might promote this increased attachment. Overall, an association between AGR2, cell elongation and changes to cell adhesion have been identified, along with changes in the secretome of AGR2-expressing cells that might contribute to these characteristics. Although proteome changes were identified in relation to the effects observed on cell adhesion, these observed changes might also offer an insight into how AGR2 induces metastasis.

6.1 A role for AGR2 in metastasis

In addition to the downregulation of basement membrane proteins mentioned above, and the upregulation of matrix-degrading protease such as MMP19 and ADAMTS8 (Table 5.3), which together would presumably promote metastasis by disrupting cell-matrix interactions, two classes of proteins that might contribute to AGR2-driven metastasis have emerged from the proteomics screen presented in Chapter 5.

6.1.1 Immunomodulatory proteins

AGR2-expressing cells showed elevated levels of the CXCL1 and CXCL2 chemokines, and AGR2-expressing clonal cells had decreased levels of the CXCL12 chemokine (Tables 5.2 and 5.3). Both CXCL1 and CXCL2 are principally neutrophil chemoattractants [711-713] and CXCL12 can act as both a neutrophil chemoattractant [714], but also as an anti-inflammatory factor [715]. This suggests that neutrophils might be recruited to AGR2-expressing tumours. These tumour-associated neutrophils (TANs) can contribute to metastasis through the promotion of matrix degradation and angiogenesis. TANs secrete several proteases (collagenase-IV, heparanase and neutrophil elastase) that can degrade the basement membrane and surrounding extracellular matrix, facilitating the dissemination of tumour cells, whilst at the same time releasing pro-angiogenic factors, including VEGF, CXCL1 and CXCL6 (reviewed in [184]). Thus TANs can act to couple ECM degradation with neovascularisation, promoting metastasis.

Aside from the release of chemotactic cytokines, there was also a strong reduction in levels of the histocompatibility antigen RT1-AW2 in clonal WT AGR2-expressing cells (Table 5.3). This may point to an AGR2-mediated programme of evasion of immune-surveillance mechanisms, a mechanism that is known to be exploited by neoplastic cells [716-718]. This is further supported by the downregulation of CXCL12 seen here in AGR2-expressing cells, as this molecule has been shown to act as an anti-inflammatory cytokine and is associated with increased metastasis of breast cancer cells, possibly through relieving an autocrine loop of CXCL12 that allows chemotaxis of breast cancer cells towards CXCL12-secreting sites such as the bone marrow [715, 719, 720]. Overall, this association of immunomodulatory protein release from AGR2-expressing cells might provide a novel pathway for AGR2-mediated tumour progression, promoting neovascularisation and tumour cell dissemination, whilst at the same time preventing tumour cell recognition by the host immune system.

6.1.2 Collagens

There was a marked increase in the abundance of several collagens in AGR2-expressing cells clonal cells (Table 5.3), where there was increased abundance of all three chains of collagen VI (COL6A1, COL6A2 and COL6A3), as well as two collagen I chains (COL1A and COL2A), COL3A1, COL7A1 and COL12A1. There was also a strong upregulation

of procollagen C-endopeptidase enhancer 1 (PCOLCE), a secreted protein required for the processing and maturation of newly-secreted collagen molecules [721] and ADAMTS2, which is involved in the maturation of procollagens I, II and III [722]. This might be important given that a signature pattern of expression of COL1A1, COL1A2, COL3A1, COL6A3 is associated with EMT [723].

In particular, the upregulation of all three subunits of collagen VI in clonal AGR2 cells in the present study suggests that it may play a particularly important role in AGR2-mediated tumour progression. Collagen VI induces EMT in breast cancer and has been shown to promote tumour angiogenesis both *in vitro* and *in vivo* [724, 725]. Collagen VI expression within tumours is often associated with the vasculature or invasive front (reviewed in [726]), and *Col6a1*$^{-/-}$ mice show defects in vasculature formation as a result of lost survival and sprouting of endothelial cells, with a resultant slowing of tumour growth [725]. This suggests that collagen VI plays an important role in tumour angiogenesis. It is also important to highlight the overlapping and complementary actions of possible TAN recruitment and collagen VI secretion by AGR2-expressing cells, as the new vessels stimulated by TANs should be stabilised by the release of collagen VI, thus establishing vasculature that can support tumour cell dissemination and metastasis.

6.1.3 A model for AGR2-driven metastasis

With the above proteomic changes to AGR2-expressing cells in mind, coupled with changes in adhesion and morphology (Chapter 4), a possible mechanism of AGR2-driven metastasis is emerging (Fig. 6.1). Through overexpression of AGR2, a subset of epithelial cells undergoes an EMT-type transformation. This is evidenced not only from the correlation between AGR2 expression and elongated cell morphology (Chapter 4), but also that clonal AGR2-expressing cells upregulated the expression of several proteins related to EMT, including TGFβ, neural cell adhesion molecule (NCAM) and vimentin (Table 5.3), although further experimentation will be required to confirm that EMT is indeed induced in these cells. These cells are, as a result, less adherent to the cell matrix and are more migratory (Chapter 4). At the same time, the decreased output of basement membrane proteins from AGR2-expressing cells, as well as increased secretion of matrix degrading enzymes (Chapter 5), causes fundamental changes to the surrounding matrix, which presumably allow these cells to break away from the main tumour mass. Release of chemokines could at the same time promote the recruitment of neutrophils, resulting in

further matrix breakdown and the stimulation of angiogenesis. The release of collagens from AGR2-expressing cells can then serve to stabilise the newly formed vasculature, in this way presenting a possible pathway for the more migratory AGR2 cells to disseminate to secondary sites. These cells have been shown in the present study to be resistant to anoikis and also exhibit a faster rate of attachment as single cell suspensions or clumps of cells in suspension, characteristics which should be favourable to promote survival of these cells in the bloodstream and promote their adhesion to the endothelium as secondary sites.

It is also interesting to note that normal breast tissue undergoes profound architectural changes during puberty, pregnancy and lactation, where ECM must be remodelled and cells must migrate through this remodelled ECM to form new breast alveoli (reviewed in [505]). It is also noteworthy that this process is coupled to new blood vessel formation in these tissues [727] and hematopoietic cells play an important role in branching and elongation of these new ducts [728]. Furthermore, AGR2 has been shown to facilitate lobuloalveolar development in a mouse model [456] and its expression is observed in normal breast tissue [390, 497]. Thus it might be that AGR2 plays a physiological role in breast tissue remodelling, but over activation of this pathway through overexpression of AGR2 leads to tumourigenesis and metastasis.

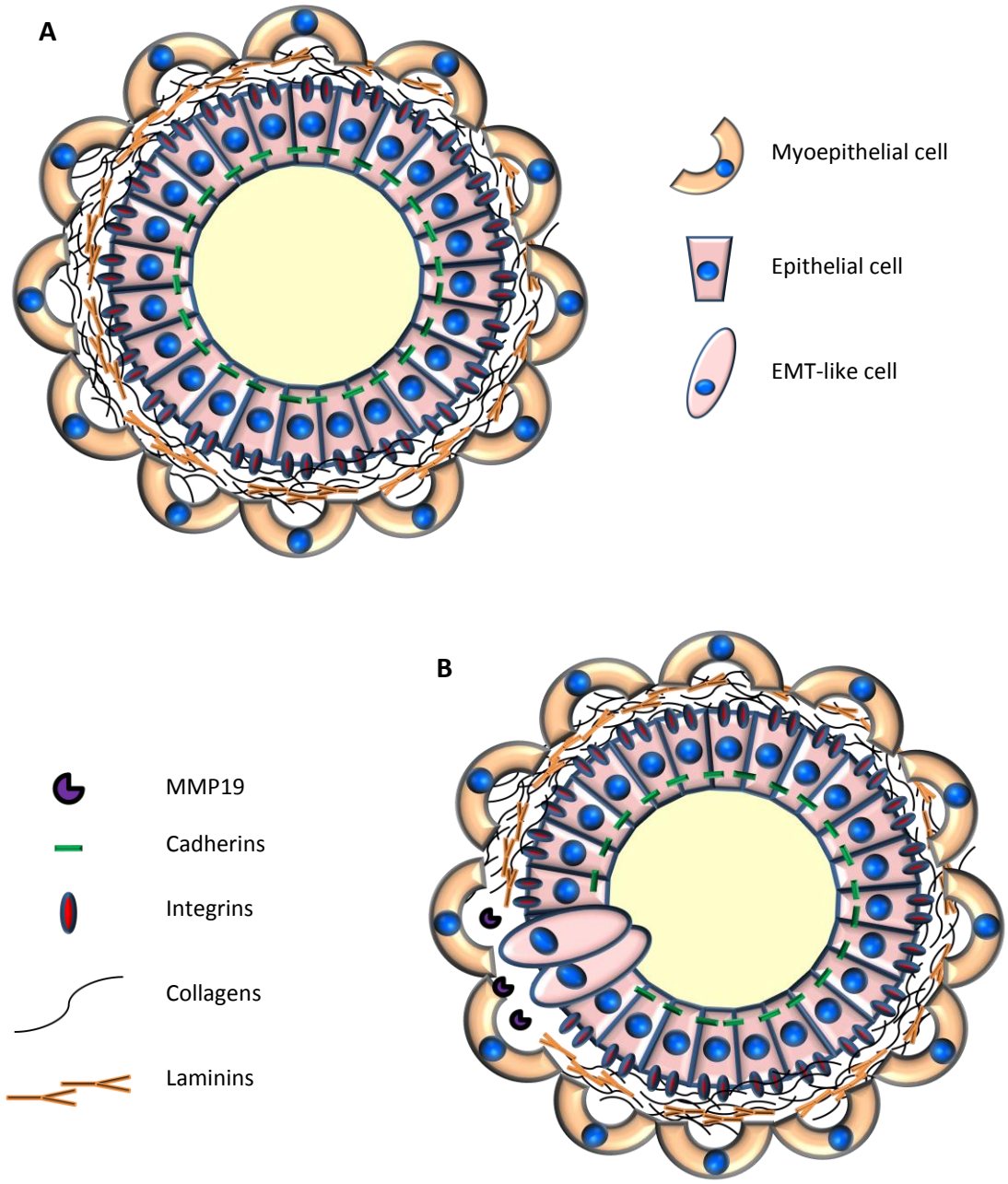


Figure 6.1 (continued overleaf). ***Possible mechanism of AGR2-induced metastasis in breast tissue.*** **(A)** In the existing breast duct, epithelial cells surround the ductal lumen, and themselves are surrounded by myoepithelial cells and extracellular matrix. **(B)** Overexpression of AGR2 promotes an EMT-type transformation of some epithelial cells, at the same time leading to production of matrix-degrading enzymes such as MMP19. The transformed cells also downregulate the production of basement membrane proteins, leading to further matrix remodelling. **(C)** These cells become less adherent to the surrounding matrix and more migratory, also releasing cytokines to recruit neutrophils that aid in matrix remodelling and angiogenesis. These new blood vessels are stabilised by the release of collagen IV from AGR2-expressing cells, leading to eventual intravasation of these cells and travel to secondary sites.

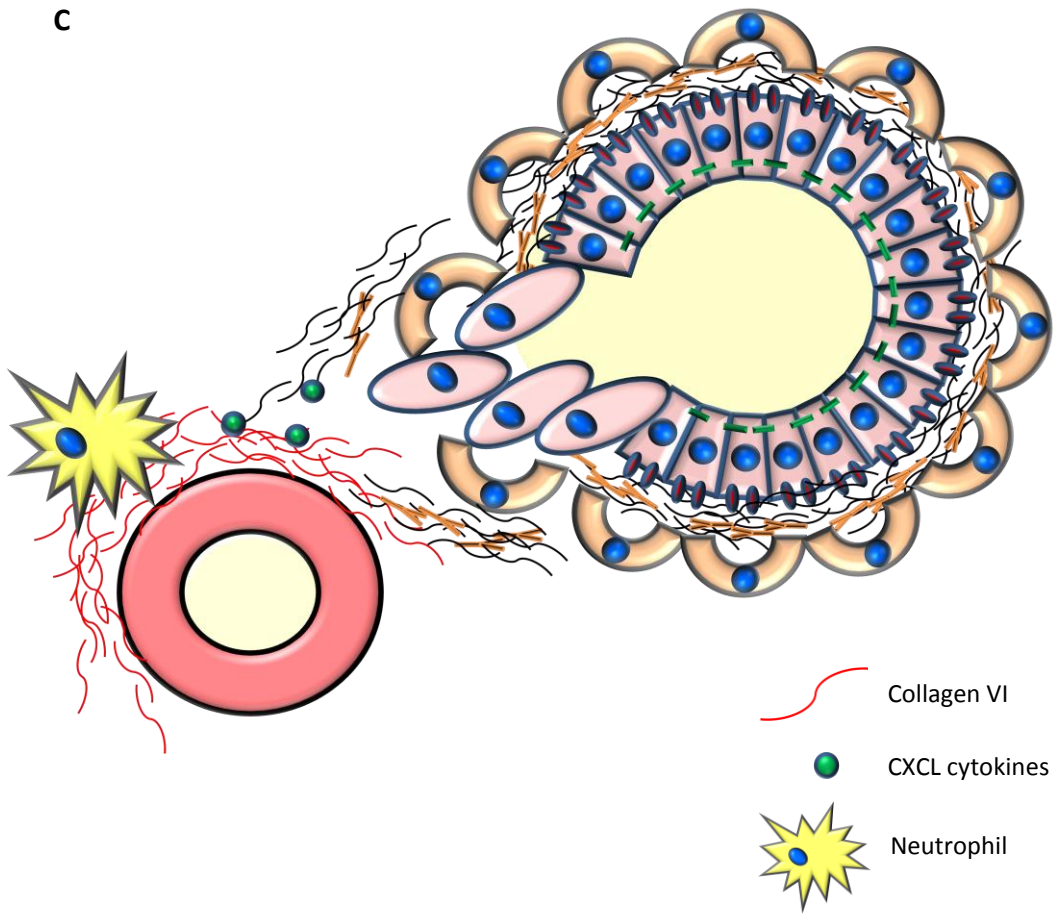


Figure 6.1 (cont.). ***Possible mechanism of AGR2-induced metastasis in breast tissue.*** **(A)** In the existing breast duct, epithelial cells surround the ductal lumen, and themselves are surrounded by myoepithelial cells and extracellular matrix. **(B)** Overexpression of AGR2 promotes an EMT-type transformation of some epithelial cells, at the same time leading to production of matrix-degrading enzymes such as MMP19. The transformed cells also downregulate the production of basement membrane proteins, leading to further matrix remodelling. **(C)** These cells become less adherent to the surrounding matrix and more migratory, also releasing cytokines to recruit neutrophils that aid in matrix remodelling and angiogenesis. These new blood vessels are stabilised by the release of collagen IV from AGR2-expressing cells, leading to eventual intravasation of these cells and travel to secondary sites.

6.2 Future work

A number of unanswered question still remain surrounding the role of AGR2 in adhesion and metastasis presented in the present study. It is still unclear whether AGR2 is responsible for the EMT-type transformation observed in the highest expressing WT AGR2 Rama 37 cells. Despite several attempts, AGR2 could not be knocked down by shRNA to investigate whether cells reverted or at least partially reverted to a more epithelial

morphology upon loss of AGR2. MET is thought to promote the colonisation of secondary tumours (reviewed in [264]) and so it is not unreasonable to assume that loss of AGR2 expression could allow cells to revert back to an epithelial phenotype. Indeed, knockdown of AGR2 in the PC3 prostate cancer cell line promoted these cells to revert from a fibroblast-type morphology to a more epithelial morphology [515], suggesting that AGR2-induced changes to cell morphology are at least partially reversible.

In relation to morphology changes in AGR2-expressing cells, it is also not entirely clear which proteins differentially expressed in clonal WT AGR2-expressing cells are mediated directly by AGR2, and which are downstream effects of the change in morphology. Determining the substrates for AGR2 will therefore also be important for the future development of any therapeutic agents. For these reasons, achieving knockdown of AGR2 will therefore be a priority of future work.

One other area of important future work will be to determine the extent to which the metastatic ability of cell lines expressing mutant AGR2 is affected by each mutation, if at all. Mutations resulting in loss of AGR2 dimerisation or the Cys81 thiol appeared to completely ablate the change to an elongated morphology associated with WT AGR2 expression, potentially suggesting that these are the most important features driving AGR2 activity. Through the use of a previously employed rat model of breast cancer progression [309], it should be possible to shed light on functionally important regions of AGR2 in promoting metastasis, which might lead to the identification of a novel therapeutic target within AGR2.

Several new avenues of investigation have also been opened up by the present report. The discovery of a secreted and glycosylated form of AGR2 should also be an interesting line of future investigation. Although it is not the only report of a secreted and glycosylated ER chaperone [446-449, 695-697], it is the first to demonstrate this facet in AGR2 biology. The main focus of future studies should revolve around the nature of this glycosylation, its consequences on AGR2 structure and the consequence of this form of AGR2 on cellular activity.

Several differentially expressed proteins in AGR2-expressing cells were associated with immunomodulatory functions and angiogenesis. A handful of previous studies have reported the association of AGR2 with angiogenesis [301, 414], but, as far as the present author is aware, none so far on the possible immunomodulatory effects of AGR2 (aside from using AGR2 in targeted immune therapy [729, 730]). Thus exploring the role of AGR2 in these processes in the context of tumour progression could yield important information in the hunt for anti-AGR2 therapies.

6.3 Conclusion

Although the exact function of AGR2 still remains elusive, the data presented in this thesis reveal several possible ways that it may be involved in both normal development and also in the generation of metastases. A framework has been established for identifying AGR2-induced changes to the cell proteome (in particular the secretome) which may be productive in better pinpointing the actions of AGR2, due to the importance of secreted factors in the metastatic process and the identification of a number of differentially-expressed secreted proteins in AGR2-expressing cells. Important insights have been gained into the biophysical properties of AGR2, which may serve to better understand its regulation of cellular processes. Overall, a clearer picture of the role of AGR2 in metastasis is emerging, which might eventually lead to the generation of anti-AGR2 therapeutic agents to treat AGR2-positive cancers.

Appendices

Appendix 1 (overleaf). ***Publication: Metastasis-Promoting Anterior Gradient 2 Protein Has a Dimeric Thioredoxin Fold Structure and a Role in Cell Adhesion.*** Pryank Patel, Christopher Clarke, Dong Liu Barraclough, Thomas Adam Jowitt, Philip Spencer Rudland, Roger Barraclough and Lu-Yun Lian

jmb Article

Metastasis-Promoting Anterior Gradient 2 Protein Has a Dimeric Thioredoxin Fold Structure and a Role in Cell Adhesion

Pryank Patel[1], Christopher Clarke[1], Dong Liu Barraclough[2], Thomas Adam Jowitt[3], Philip Spencer Rudland[1], Roger Barraclough[1] and Lu-Yun Lian[1]

1 - Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK
2 - Institute of Ageing and Chronic Disease, University of Liverpool, Daulby Street, Liverpool L69 3GA, UK
3 - Biomolecular Analysis Core Facility, Faculty of Life Sciences, University of Manchester, Carys Bannister Building, Dover Street, Manchester M13 9PL, UK

Correspondence to Lu-Yun Lian and Roger Barraclough: lu-yun.lian@liverpool.ac.uk; brb@liverpool.ac.uk
http://dx.doi.org/10.1016/j.jmb.2012.12.009
Edited by Y. Shi

Abstract

Anterior gradient 2 (AGR2) is a normal endoplasmic reticulum protein that has two important abnormal functions, amphibian limb regeneration and human cancer metastasis promotion. These normal intracellular and abnormal extracellular roles can be attributed to the multidomain structure of AGR2. The NMR structure shows that AGR2 consists of an unstructured N-terminal region followed by a thioredoxin fold. The protein exists in monomer–dimer equilibrium with a K_d of 8.83 μM, and intermolecular salt bridges involving E60 and K64 within the folded domain serve to stabilize the dimer. The unstructured region is primarily responsible for the ability of AGR2 to promote cell adhesion, while dimerization is less important for this activity. The structural data of AGR2 show a separation between potential catalytic redox activity and adhesion function within the context of metastasis and development.

Introduction

Anterior gradient 2 (AGR2) is a major development-regulating protein that specifies the formation of embryonic ectoderm of tadpoles[1] and induces the remarkable effect of limb regeneration in salamanders.[2] In mammals, AGR2 is produced in a restricted number of normal tissues, most notably in the secretory goblet cells of the intestine,[3] where it is associated with normal processes of secretion of mucin,[4] a large cysteine-rich glycoprotein that protects the inner lining of the intestine.

AGR2's rise to prominence among proteins that control development, however, has been its role in abnormal progression of disseminated disease, in particular, cancer. In oncogenesis, AGR2 is upregu-lated in the primary tumors of breast,[5,6] lung,[7] ovarian,[8] oesophageal,[9] pancreatic[10] and prostate cancers.[11] Its elevated level in breast and prostate cancer is associated with markedly reduced survival of patients.[11–14]

AGR2 has been found in circulating tumor cells from breast, colon and prostate cancer patients,[15] in the circulating blood of patients with ovarian cancer[16] and in ectopic lesions of the non-malignant condition of endometriosis,[17] hence suggesting a potential role for AGR2 in cancer dissemination. Overexpression of AGR2 in a benign mammary cell system induced a metastatic dissemination to other organs, particularly the lungs, when transplanted into the mammary glands of syngeneic rats.[6] In this metastatic model, a change in cell activity induced by AGR2 was an enhanced rate of cell adhesion.[6]

AGR2 is distantly related to the large protein disulfide isomerase (PDI) family. Its primary sequence contains a signal sequence between residues 1 and 20, which is required for import into the endoplasmic reticulum but is not found in the mature protein. AGR2 also contains an atypical C-terminal endoplasmic reticulum retention sequence, KTEL.[18]

AGR2 has a single CXXS motif that is likely to be responsible for its normal role by forming mixed disulfides with the intestinal mucin MUC2[4], mucins MUC1[19] and MUC5AC.[20] A closely related protein, AGR3,[21] of unknown function also contains a 21-amino-acid signal sequence for endoplasmic reticulum entry and is about 72% identical with amino acids 41–175 of AGR2. However, at the N-terminus, outside the signal peptide region, AGR3 has a shorter segment corresponding to AGR2 amino acids 21-40 (Fig. 1).

We describe here the first structure of AGR2 (without the 20-amino-acid signal peptide) and show that the N-terminal amino acid region 21–40 is unfolded and that the C-terminal domain adopts the thioredoxin fold. The N-terminal region is primarily responsible for the cell adhesion properties of AGR2, whereas the folded domain forms a dimer through specific intermolecular salt bridges. AGR2, therefore, appears to have a bifunctional role in promoting cell adhesion and in PDI-like reactions.

Results

Biophysical characterization of AGR2

$AGR2_{21–175}$ and the truncated $AGR2_{41–175}$ (Fig. 1) have a molecular mass of 30.6±0.1 kDa and 26.0 ±0.1 kDa, respectively, rather than the corresponding theoretical molecular mass of 17.8 kDa and 16.1 kDa, suggesting that both proteins exist in equilibrium between monomers and dimers (Fig. 2a). The determined molecular mass suggests that the region 21–40 makes only minimal, if any, contribu-tion to the dimerization.

The heterogeneous nature of AGR2 is not unique; the closely related protein AGR3, which shares 65% amino acid sequence identity with AGR2, is a mixture of species with a molecular mass ranging from 31.5 kDa to 26.3 kDa (Fig. 2a). Mutation of the only cysteine residue in $AGR2_{41–175}$, C81 to serine, had a minimal effect on its dimeric state (Fig. 2b) and treatment with DTNB [5,5′-dithiobis-(2-nitrobenzoic acid)] showed that this cysteine residue exists in the reduced state, thereby eliminating an intersubunit disulfide bridge. The monomer–dimer equilibrium of $AGR2_{41–175}$ was investigated using equilibrium analytical ultracentrifugation (Fig. 2d). The results fit well to a monomer–dimer distribution model, with a random distribution of residuals. The association constant ($\log_{10}$ Ka) was 5.054, indicating that the concentration for dissociation is approximately 8.83 µM. Therefore, at the 0.5- to 1-mM concentra-tions used for the structural studies, the protein is predominantly a dimer.

The $^{1}H–^{15}N$ heteronuclear single quantum co-herence (HSQC) spectrum of $AGR2_{21–175}$ contains a mixture of well-dispersed and intense poorly dispersed resonances, whereas the $AGR2_{41–175}$

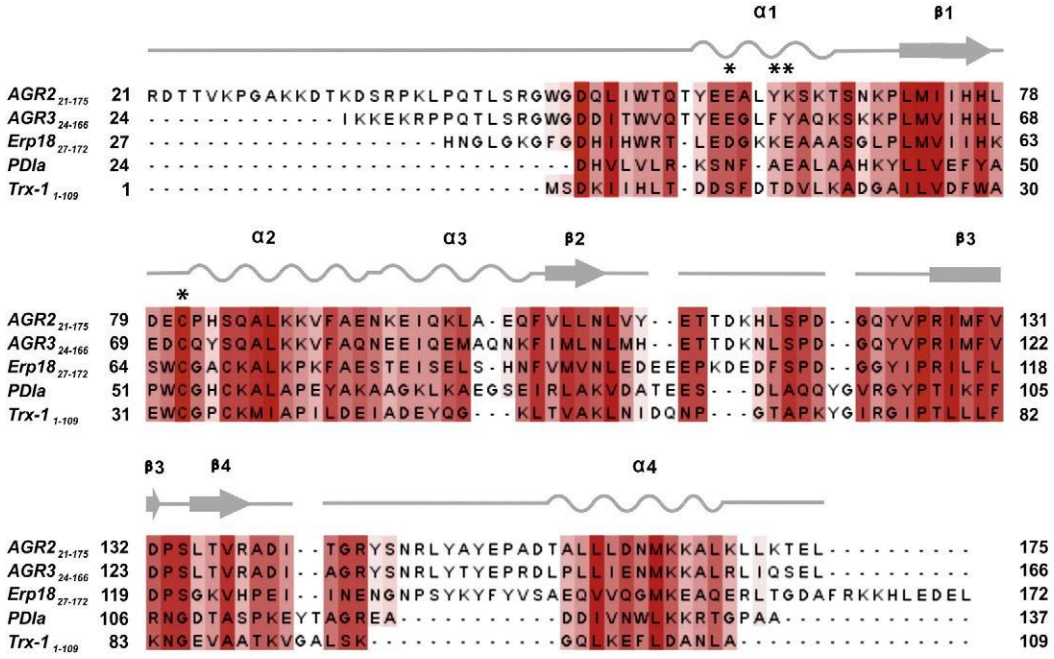


Fig. 1. Sequence of AGR2 and AGR3. Sequence alignment of $AGR2_{21–175}$, $AGR3_{24–166}$, Erp18 and thioredoxin A. The sequences are shown without their N-terminal signal peptides. Conserved residues are shaded in red. The asterisks (*) indicate the residues that were mutated during these studies. Secondary structure elements of AGR2 determined in this paper are shown above the sequence.

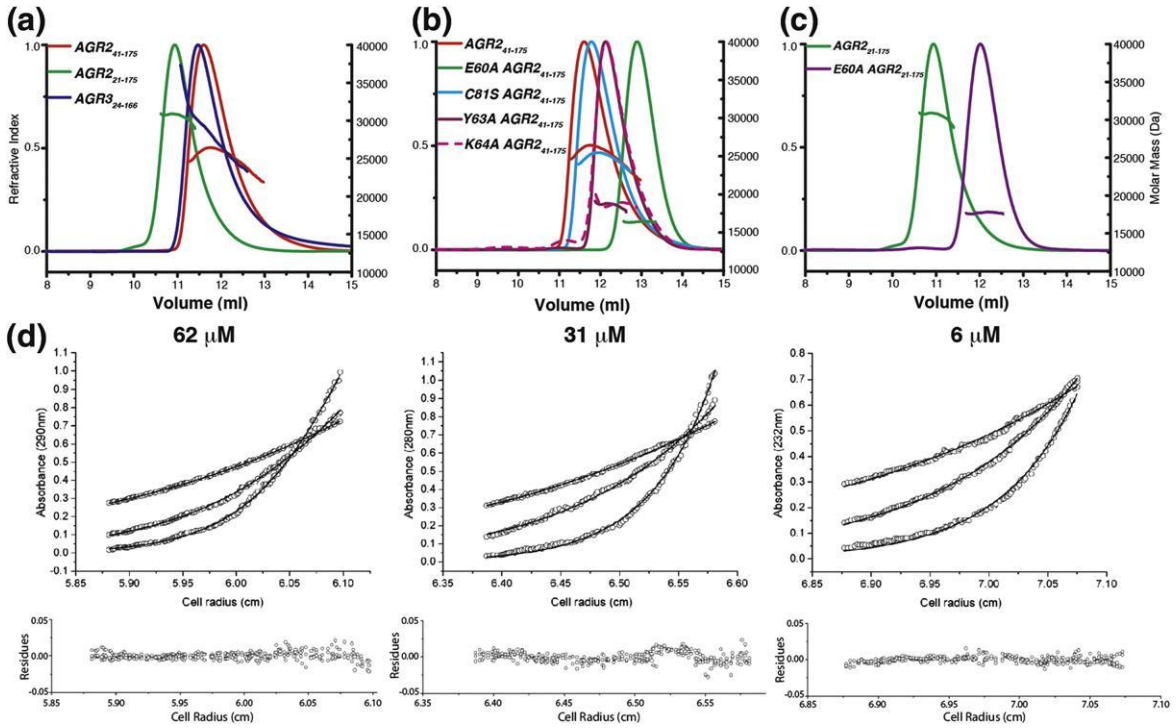


Fig. 2. MALLS and analytical ultracentrifugation characterization of AGR2 and AGR3. (a) MALLS analysis of wild-type AGR2 $_{21–175}$, AGR2$_{41–175}$ and AGR3$_{24–166}$. The major isolated peak for AGR2$_{41–175}$ has a mean molecular mass of 26.0± 0.1 kDa, 31.1±0.6 kDa for AGR2$_{21–175}$ and 30.6±0.1 kDa for AGR3$_{24–166}$, rather than the expected molecular mass of, respectively, 16.1 kDa, 17.8 kDa and 16.9 kDa (for monomers) or 32.2 kDa, 35.6 kDa and 33.8 kDa (for dimers), confirming monomer–dimer equilibrium. (b) MALLS analyses of AGR2$_{41–175}$ and its mutants E60A, C81S, Y63A and K64A confirm that mutation of the E60 residue abolishes dimerization whereas mutation Y63 or K64A disrupts, rather than abolishes, dimerization. The minimal effects of the C81S mutation indicate that dimerization is not due to intersubunit disulfide linkages. (c) MALLS analysis of AGR2$_{21–175}$ and E60A AGR2$_{21–175}$ confirms that residue E60 plays an important part in AGR2 dimerization. (d) Fitting of AGR2$_{41–175}$ equilibrium analytical ultracentrifuge data to a monomer–dimer equilibrium at 62 μM, 31 μM and 6 μM. The sample was in 20 mM sodium phosphate (pH 6.5) with 150 mM NaCl at 20 °C. The results fit well to a monomer–dimer distribution model with a random distribution of residuals to give a dissociation constant of 8.83 μM under these conditions.

spectrum consists of predominantly well-dispersed cross-peaks. This suggests that the first 20 amino acids of AGR2$_{21–\ 175}$ are largely unstructured. The absence of peak doubling in the ^{1}H–^{15}N HSQC spectrum of AGR2$_{41–175}$ indicates that the dimer is symmetrical (Fig. 3a). The monomer–dimer equilibrium significantly reduced the NMR spectral quality through chemical exchange line-broadening and made the resonance assignments and structure determination more challenging. However, since there were no suitable proteins against which modeling could be undertaken to identify the residues responsible for dimerization, it was not possible to convert the dimeric state to a monomer through mutagenesis until after the structure of the dimer was determined. Besides, there were also no previous data that identified the specific roles, if any, of the structured and unstructured domains of AGR2. Hence, NMR studies were pursued with conditions under which the protein was predominantly dimeric.

NMR studies of AGR2$_{41–175}$ homodimer

To obtain good quality data, a combination of high-level deuterium labeling of the protein, selective ^{13}C-methyl labeling and transverse relaxation optimized spectroscopy-based NMR experiments was required to yield complete backbone resonances (^{15}N, ^{1}H, ^{13}C$^{\alpha}$, ^{13}C$^{\beta}$ and ^{13}C′) and the important nuclear Overhauser enhancement (NOE) data re-quired for structure determination. Leucines, iso-leucines and valines account for nearly 24% of the AGR2$_{41–\ 175}$ polypeptide sequence, and these side-chain resonances were readily assigned using the three-dimensional (3D) (H)CCH$_3$–TOCSY experi-ments (Fig. 3b) [22] on {[U- ^{15} N, ^{13} C, ^{2} H], Ile $^{\delta 1}$ [^{13}CH$_3$],Leu,Val[^{13}CH$_3$]}AGR2$_{41–175}$. The residue-

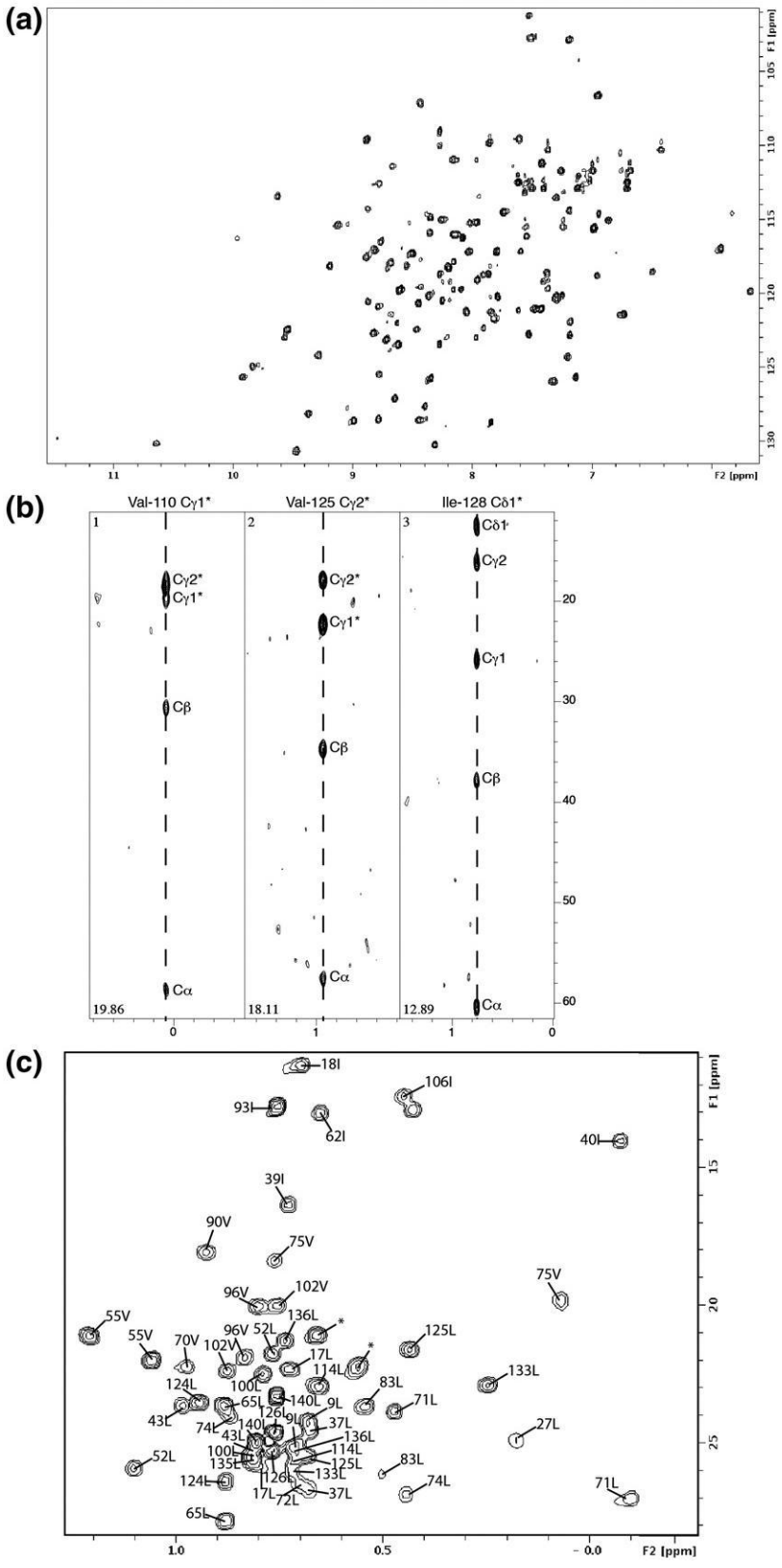


Fig. 3. Spectra of AGR2$_{41–175}$. (a) ^{15}N–^{1}H HSQC spectrum of {[U-^{15}N,^{13}C, ^{2}H]}AGR2$_{41–175}$. (b) Strips from the 3D (H)CCH$_3$–TOCSY experiments using {[U-^{15}N,^{13}C,^{2}H], Ile$^{\delta1}$[^{13}CH$_3$],Leu, Val[^{13}CH$_3$]}AGR2$_{41–175}$ showing the side-chain assignments of Val110, Val125 and Ile128. (c) The assigned methyl region of the ^{1}H – ^{13}C HSQC spectrum of {[U-^{15}N,^{13}C,^{2}H], Ile$^{\delta1}$[^{13}CH$_3$],Leu, Val[^{13}CH$_3$]}AGR2 $_{41–175}$. Asterisks (*) represent peaks with multiple or ambiguous assignments.

specific assignments (Fig. 3c) enabled the NOE spectroscopy (NOESY) spectrum to be assigned and the generation of the key distance restraints.

The monomer subunit structure was calculated using 595 inter-proton distance restraints obtained from methyl-methyl, methyl-HN and HN-HN NOE's; 95 residual dipolar coupling (RDC) restraints measured in two different alignment media [Pf1 bacteriophage and 5% (v/v) pentaethylene glycol monododecyl ether:hexanol in a 0.96:1 ratio]; 179 TALOS-derived backbone torsion angles and 37 hydrogen bonds. The 10 calculated structures with the lowest total free energy and the least number of violations have a backbone RMSD of 0.42 Å for the structured regions (Table 1 and Fig. 4a). It is interesting to note that, despite the low number of NOE restraints, it is possible to obtain a high-resolution structure. This is due to the fact that the distance restraints that are important to "fold" the protein are available, since they often involve amino acid residues with methyl groups.

Dimerization of AGR2$_{41–175}$ occurs via the antiparallel arrangement of the α1 helix

Although it was possible to produce heterodimers composed of unlabeled and $^{13}C/^{15}N$ AGR2$_{41–175}$ (Fig. S1), no intermolecular NOEs were detected using the inherently insensitive $^{13}C/^{15}N$-filtered NOE experiments. We, therefore, used an RDC-assisted modeling method, which was reported to be suitable for the determination of the solution structures of symmetrical dimers,[23] to determine the structure of the AGR2$_{41–175}$ dimer. In this approach, RDCs are first used to obtain the orientation of the symmetry axis relative to the monomer structure. As the symmetry axis adopts the same orientation relative to the monomers, irrespective of alignment media, two non-degenerate alignment tensors should share a common axis, enabling the orientation of the symmetry axis to be identified. AGR2$_{41–175}$ dimer was aligned in Pf1 bacteriophage and 5% (v/v) pentaethylene glycol monododecyl ether:hexanol in a 0.96:1 ratio; the RDCs were measured and the orientation of the symmetry axis was determined as described in the experimental procedure (Fig. 4c and d). Homodimer models in agreement with these symmetry restraints were then constructed using the grid-search algorithm, and these models were evaluated based on the quality of the interface between the subunits and the agreement with the experimental data.

The correlation between the measured and back-calculated dimer RDCs shows a good fit between the two values, with a Q value of 0.018 for data acquired in both phage and polyethylene glycol (PEG). The AGR2$_{41–175}$ dimer model structure obtained (Fig. 5a) shows that the interface is made up of the α1 helix, with a buried area of approximately 812 Å^2. Across the representative models of the ensemble, the ionic interaction between E60 of one subunit and K64 of the other (labeled as K64*) is absolutely conserved with intact electrostatic interactions between the two side-chain carboxylate oxygens of E60 and the side-chain ammonium hydrogen of K64 (Fig. 5b). Both these residues are located in the α1 helix, and hence the E60–K64* electrostatic interaction can only be achieved through an antiparallel arrangement of the α1 helix (Fig. 5b and c). Y63–Y63* hydrophobic interactions are also possible from the model. Further evidence that the α1 helix forms the dimer interface was obtained from amide solvent protec-tion data obtained using paramagnetic relaxation agents (Fig. 5a and Fig. S2).

Table 1. Structure statistics for AGR2$_{41–175}$ and E60A AGR2$_{41–175}$

	AGR2$_{41–175}$	E60A AGR2$_{41–175}$
NMR distance and dihedral constraints		
Distance constraints		
Total NOE	595	2364
Ambiguous	22	253
Intra-residue	73	993
Inter-residue	500	1118
Sequential (|i −j|=1)	218	572
Medium range (|i −j| b 4)	142	314
Long range (|i −j| N 5)	141	232
Hydrogen bonds	37	40
Total dihedral angle restraints		
φ	90	82
ψ	89	80
Total RDCs	95	57
NH phage	43	57
NH PEG[a]	52	—[a]
RDC Q-factor	0.018	0.101
Average pairwise RMSD** (Å)		
Heavy[b]	1.31	1.55
Backbone[c]	0.42	0.59

[a] NH PEG RDCs were not collected.

[b] RMSD calculated with CNS.

[c] Ten structures were superimposed using the backbone atoms of residues in secondary structure elements of AGR2$_{41–175}$ (58– 67, 72–77, 82–93, 95–103, 105–108, 127–132, 135–138 and 157–168). The corresponding secondary structures are α1, β1, α2, α3, β2, β3, β4 and α4.

Mutagenesis of residues responsible for dimerization

The importance of E60, Y63 and K64 in forming the intermolecular interactions was tested by mutat-ing each to an alanine followed by characterization using multi-angle laser light scattering (MALLS) (Fig. 2b). E60A AGR2$_{41 – 175}$ is a monomer (molec-ular mass, 16.4 kDa). The K64A (molecular mass, 18.7 kDa) mutant exhibits disrupted dimerization, thus confirming the importance of the E60–K64 interaction in maintaining the dimer structure. The molecular mass of Y63A AGR2$_{41– 175}$ (molecular mass, 18.9 kDa) is similar to that of the K64A

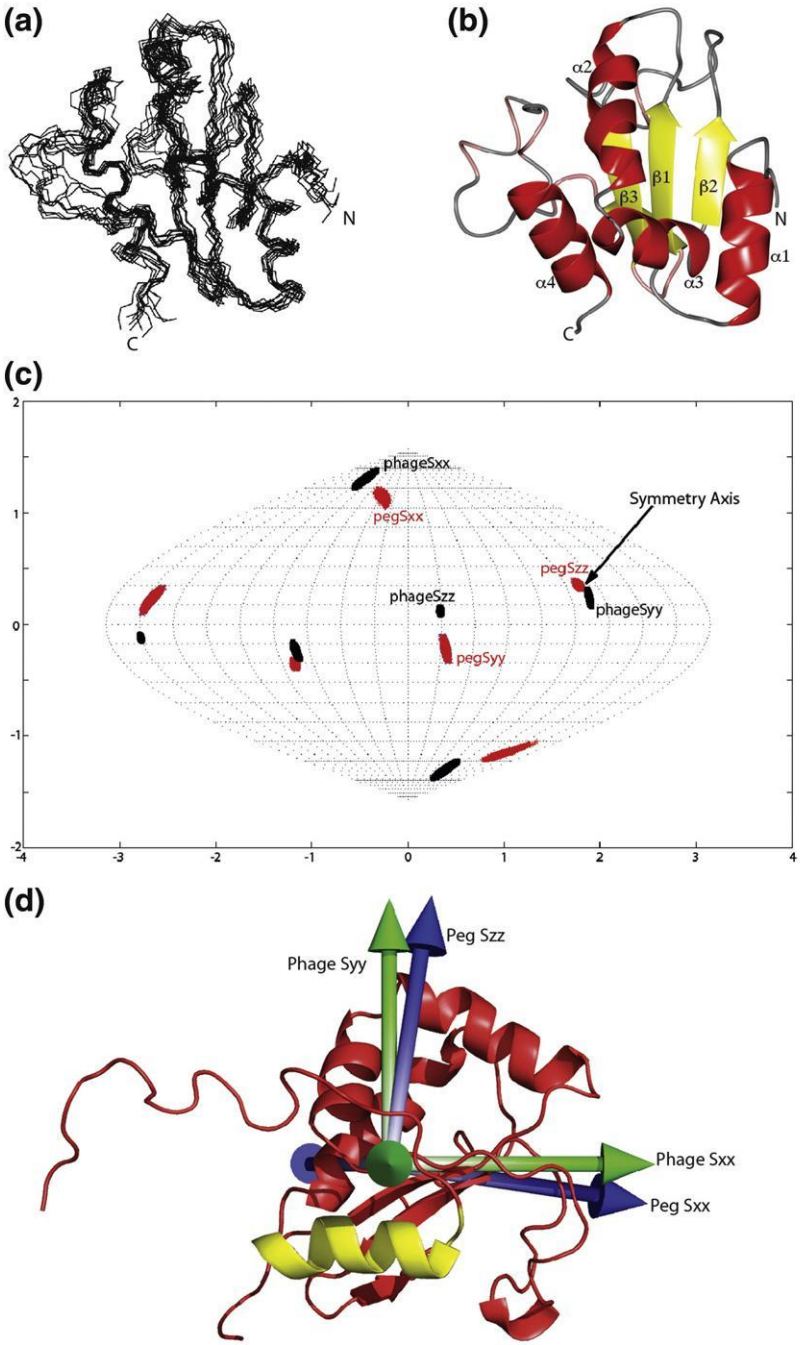


Fig. 4. Overall structure of AGR2 monomer subunit and directions of alignment axis for of $AGR2_{41–175}$.

(a) Superposition of the 10 lowest-energy conformers of monomer subunit of $AGR2_{41–175}$. (b) Ribbon representation of the subunit structure of $AGR2_{41–175}$ with the secondary structures marked. (c) Sauson-Flamsteed plot showing orientations of the principle align-ment tensor solutions calculated from RDCs of $AGR2_{41–175}$ in Pf1 bacteriophage and 5% pentaethylene glycol monododecyl ether:hex-anol in a 0.96:1 ratio. The overlap of the PEG Sxx and bacteriophage Syy alignment tensor axes indi-cates the C2 symmetry axis of $AGR2_{41–175}$. (d) The order tensor directions plotted onto the molecu-lar frame of $AGR2_{41–175}$ to show the symmetry axis obtained from the two alignment media.

mutant. The MALLS analysis of $AGR2_{21–175}$ and E60A $AGR2_{21–175}$ further shows that AGR2 dimer-ization occurs independently of residues 21–40 (Fig. 2c), since both proteins dimerize irrespective of whether or not these N-terminal residues are present. The intersubunit E60–K64* salt bridge is confirmed by the effects of salt concentration on the stability of the dimer. A comparison of the chemical shifts between the monomeric E60A $AGR2_{41–175}$ and the wild-type $AGR2_{41–175}$ protein shows that the residues whose chemical shifts are most affected are from the α1 helix (Fig. S3).

Structure determination of the E60A $AGR2_{41-175}$ monomer

E60A $AGR2_{41–175}$ is monomeric in solution. The structure of E60A $AGR2_{41–175}$, determined using standard methods, yielded 10 low-energy and least restraint violation structures (Table 1 and Fig. 6a and b). The all-atom RMSD between the mutant and one subunit of the wild-type dimer is 2.82 Å (Fig. 6c), confirming the structure of the latter, which was obtained using a significantly reduced number of structural restraints. E60A $AGR2_{41–175}$ had a

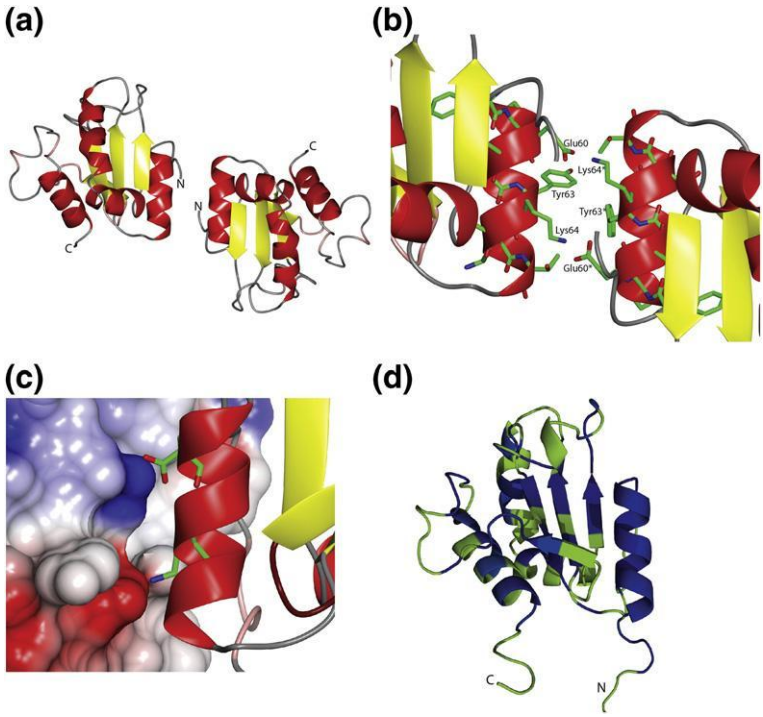


Fig. 5. Structural analysis of $AGR2_{41–175}$ homodimer. (a) Homo-dimeric model of $AGR2_{41–175}$ show-ing the antiparallel arrangement of the α1 helix. (b) Dimeric interface, with the residues comprising the α1 helix represented as sticks. The intersubunit salt bridge between the carboxylate group of Glu60 and the ammonium group of Lys64 forms the major interaction between the subunits, with the Y63–Y63 hydrophobic interaction having some influence on dimer stability. (c) Complementary charged sur-faces at the dimer interface with surface representation for one and ribbons for the other subunit. (d) Ribbon representation of $AGR2_{41–175}$ mapped according to paramagnetic relaxation effect induced by the presence of Gd(III)-DTPA. Re-gions of high solvent protection are colored blue and regions of low solvent protection or no data avail-able are colored green. Regions are considered to have high solvent protection if the peak intensity ratio, calculated from ^{15}N HSQC spectra of $AGR2_{41–175}$ collected in the presence and absence of 2 mM Gd(III)-DTPA, is above 0.65 and are considered to have low solvent protection if the peak intensity ratio is below 0.65.

tendency to aggregate over time, degrading the overall quality of the NMR data, especially those involving residues that form the dimer interface in the wild-type protein. This explains why the structure of the monomer is of lower quality than that of the dimer. The solution-state characteristics of the monomeric protein suggest that AGR2 requires a binding partner (either another protein or itself) to achieve structural stability.

Overall structure of AGR2

Each $AGR2_{41–175}$ monomer subunit consists of a four-stranded β-sheet core surrounded by four α-helices: α1 (residues 58–67), β1 (residues 72–77), α2 (residues 82–93), α3 (residues 95–103), β2 (residues 105–108), β3 (residues 127–132), β4 (residues 135–138) and α4 (residues 157–168) (Fig. 4b). This α-β-α-α-β-β-β-α arrangement is similar to a thioredoxin fold. The active-site CPHS motif is located at the N-terminal end of the α3 helix, consistent with other known structures of thioredoxin domains.[24] The α3 and α4 helices are bent with respect to each other at an angle near to 90°. A short 3_{10} helix is present in the region between the β2 and β3 strands, which is otherwise intrinsically disor-dered. Another 19-residue region of intrinsic disorder occurs between the β4 and α4 structural elements; within this region is an insertion consisting of residues 150–156, which is unique to AGR2, AGR3 [Protein Data Bank (PDB) accession number 3PH9] and Erp18,[25] when compared to other PDI family functional domains.

The structure of AGR2 possesses features typically observed in members of the PDI family. The conserved proline (P126 in AGR2) at the start of β3, like other PDI family members, is in the cis conformation (Fig. 6d). This conserved cis proline is located close to the potential catalytic active site and has been shown to be important for preserving the structural integrity of the active site for both substrate binding and catalysis in other PDI mem-bers. Another feature of thioredoxin domains, also found in the $AGR2_{41–175}$ structure, is the broken helix immediately following the CXXC(S) motif. However, rather than a proline being present in this kink, as found in many PDI domains, the corresponding residues in AGR2 are Asn94 and Lys95. In the case of the closely related AGR3 (GI: 66774045), the equivalent residues are Asn84/Glu85.

Similar to human PDI, Erp44 and Erp18, a hydrophobic patch formed by residues including A86, V90, I92, Y124, V125, I128, Y150, A151, Y152 and L161 can be found surrounding the CPHS motif (Fig. 6e) of AGR2. As has been shown or postulated for other PDI proteins, this large hydrophobic patch could be involved in binding target proteins.

We have detected neither isomerase nor glutar-edoxin activity using biochemical enzymatic assays and standard substrates, although mixed disulfide

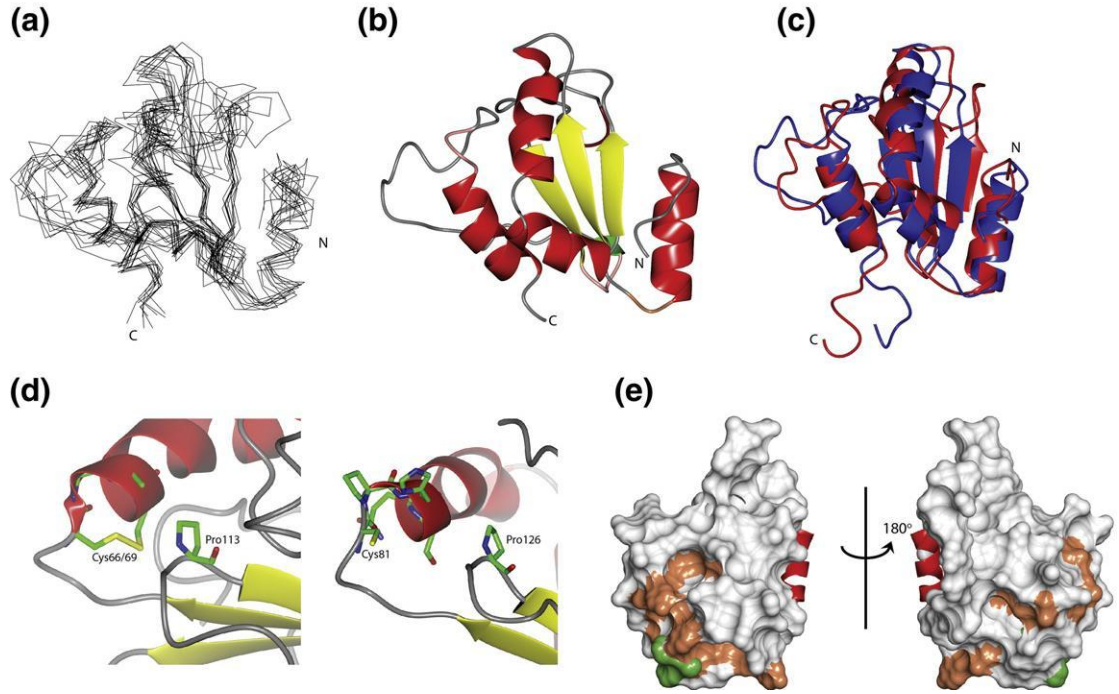


Fig. 6. Structure of E60A $AGR2_{41–175}$. (a) Superposition of the 10 lowest-energy conformers of monomer subunit of E60A $AGR2_{41–175}$. Structures were calculated with a total of 2257 NOE-derived distance restraints, 162 dihedral angles, 40 hydrogen bonds and 57 RDC restraints. (b) Ribbon representation of the structure of E60A $AGR2_{41–175}$ with the secondary structures indicated. (c) Superposition of lowest-energy conformers of $AGR2_{41–175}$ and E60A AGR2 with an all-atom RMSD of 2.82 Å between the two structures. (d) Catalytic active-site conformation in PDIs. This is well formed around the CXXC motif of Erp18 (PDB ID: 1SEN) (left) in comparison with the poorer conformation of CPHS in AGR2 (right). The cis proline is a structural requirement for correct function in PDI members. (e) Hydrophobic patch (brown) around the CPHS motif (green) shows its location is distant from the dimerization helix α1 (red). The orientations of the two structures differ by 180°.

formation between AGR2 and MUC2 has been reported.[4] In PDIs, there has been considerable attention paid to the buried glutamate and lysine salt bridge within the proximity of the cysteines in the CXXC motif; in human PDI (GI: 110815912), these are formed between Glu49 and Lys83. In AGR2, the corresponding residues, I75 and N108, do not form salt bridges, similar to the structure of Erp44 where the equivalent residues are Asn23 and Arg60.[26] In addition, the highly conserved Arg120 in human PDI has been shown, through mutation analysis, to have a direct effect on catalytic activity. This residue occurs in an unstructured region of the protein, with its side chain able to adopt a continuum of conformations ranging from totally solvent exposed to buried conformations, in which it is able to form other hydrogen bonding networks. The movement of the Arg side chain, in and out of the active site, affects the pK_as of both active-site cysteines, and this modulation is important in the thiol-disulfide catalytic cycle. In AGR2, this conserved arginine (Arg128) also adopts a range of conformations[27] and is likewise expected to play a role in modulating the pK_a of the active-site cysteine during this cycle.

Amino acids 21–40, not dimerization, determine the effect of AGR2 on the rate of cell adhesion

$AGR2_{21–175}$, when coated on plastic substrata, significantly enhances the rate of adhesion of rat mammary tumor cells[6] by 2.4-fold over uncoated substrata (adhesion efficiencies, 45.0 ±1.4% and 18.6 ±1.3%, respectively; P b 0.001, Student's t test) (Fig. 7). In contrast, $AGR2_{41–175}$, which lacks the 21– 40 region, not only failed to enhance the rate of cell adhesion but also, like the negative control protein bovine serum albumin, exhibited a significantly reduced rate relative to uncoated wells (adhesion efficiencies: $AGR2_{41–175}$, 10.4±0.2%; bovine serum albumin, 3.6±1.5%; P =0.001 and P b 0.001, respec-tively, compared to uncoated plates, Student's t test) (Fig. 7). The importance of amino acids 21–40 of AGR2 in enhancing the rate of cell adhesion was further demonstrated by the closely related protein $AGR3_{24–166}$, in which segment 21–40 of AGR2 is replaced by 9 amino acids in AGR3, of which 7 are identical with or similar to those found in AGR2. However, unlike $AGR2_{21–175}$, $AGR3_{24–166}$ exhibited similar adhesion ability (adhesion efficiency, 12.9 ±

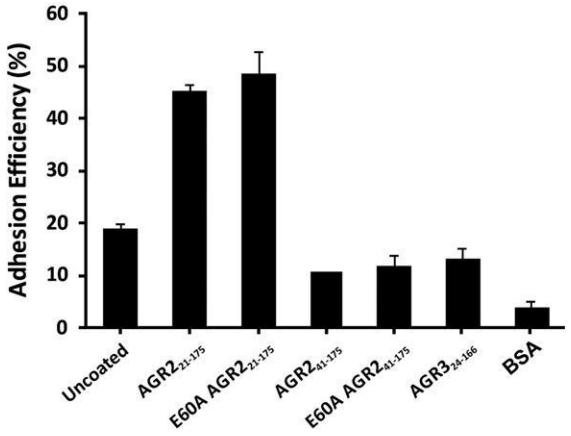


Fig. 7. AGR2 residues 21–40 are necessary for promoting cell adhesion. 24-well tissue culture plates with the indicated proteins or with coating buffer only (uncoated). Following a 60-min incubation with Rama 37 cells, adhering cells were counted and adhesion efficien-cies were calculated as the percentage of starting cells that adhered to the plate. Each condition was carried out in duplicate and the data represent the mean±standard error of the duplicates from six independent experiments for uncoated and $AGR2_{21–175}$ or three independent experi-ments for other conditions. The adhesion efficiencies of dimer $AGR2_{21–175}$ and monomer E60 $AGR2_{21–175}$ are similar, showing that dimerization is not important for the adhesion properties of AGR2. The effect of removal of the N-terminal region, 21–40, shows the importance of this region for adhesion.

2.4%; P =0.43, Student's t test) to the truncated $AGR2_{41–175}$ (Fig. 7). This suggests that the role in cell adhesion is specific for the N-terminal sequence 21–40 of AGR2. However, further investigation showed that a synthetic peptide consisting of this polybasic amino acid sequence 21–40 of AGR2 alone was unable to promote an enhanced rate of adhesion of the cells to the substratum (Fig. S4). It would appear that the folded domain must also be present and covalently linked to the 21–40 peptides to increase the rate of cell adhesion. However, an SXXS mutant of $AGR2_{21–175}$, which might be expected to abolish any potential PDI or disulfide-forming activity by AGR2, exhibited the same enhancement of adhesion as the non-mutant protein (Fig. S5), suggesting that these activities are not associated with the adhesion-enhancing activity of AGR2.

The monomeric mutant proteins E60A $AGR2_{21–175}$ and E60A $AGR2_{41–175}$ exhibited rates of cell adhe-sion (48.3±4.3% and 11.5±0.3%, respectively) that were not significantly different from the corresponding native dimeric proteins, $AGR2_{21–175}$ and $AGR2_{41–175}$ (P =0.58 and P =0.74, respectively, Student's t test) (Fig. 7). This shows that monomeric and dimeric forms have similar cell adhesion properties.

Discussion

NMR structure of AGR2

The AGR2 dimer and monomer structures described here are examples, which clearly demonstrate that the quality rather than the number of NOEs is important for obtaining a high-resolution structure, provided that NOEs are supplemented with RDC restraints.

The experiments to confirm the dimer structure—mutagenesis and solvent protection—are neces-sary, since the method used to obtain this structure is a modeling approach, assisted by experimental restraints.[23] One of the main problems with this method is reproducibility, although other factors are also important. The method relies on two assump-tions. The first one is that the multimer in question possesses C2 symmetry around the symmetry axis; for the vast majority of biological homodimers, this is a safe assumption. The second assumption is that the calculated starting subunit structure is good enough to produce an accurate representation of the oligomeric complex. The initial AGR2 subunit struc-ture was calculated with limited NOEs due to the isotope labeling schemes adopted, which were necessary to overcome the line-broadening effects and allow good data for resonance assignments and to provide the crucial NOEs to determine the protein fold. An added advantage of the labeling scheme used is that there is little possibility of misinterpreting intersubunit NOEs as intrasubunit NOEs, or vice versa. Furthermore, intersubunit NOEs are unlikely to be observed in the NOESY spectra because of the nature of the labels introduced. In theory, these sparse NOEs could have led to very poor structures; in the event, however, when supplemented with high-quality RDCs, good structures of the monomer subunits are obtained, giving confidence to the final dimer structure.

Dimerization of AGR2

The RDCs are also crucial for calculating the dimer structure. The rigid-body docking method is followed by minimization at the interface, back-calculation of the RDCs of the homodimer and then analysis of the quality of the interface. It is, however, not difficult to imagine situations where flexible or other regions in a structure could interfere with the docking of the subunits, resulting in these dimer structures being rejected as plausible structures. One other condition that is required for this method to work is that, in cases where there might be a mixture of species in equilibrium, the oligomer being calculated is the dominant contributor to the RDC data being collected.[23] In the case of AGR2, which has been shown to have a dissociation constant of 8.8 μM, the

vast majority of the protein in solution is in a homodimer form; thus, this condition is satisfied.

Only a limited number of PDIs exist as dimers; yeast PDI, which contains multiple active and inactive thioredoxin domains, forms dimers by interactions involving the inactive thioredoxin b′ domain.[28] However, this domain is not present in AGR2. Erp29, which contains a single thioredoxin fold with a single CLPC active site, forms a tight dimer with an extensive dimer interface consisting of multiple electrostatic and hydrophobic interactions involving residues toward its N-terminal region.[29] Erp16/18 also resembles AGR2 and AGR3 in possessing only one thioredoxin fold; although originally thought to be a dimer,[30] this was not confirmed in the NMR structure, which showed Erp16/18 to be a monomer.[25] That Erp16/18 is a monomer is not surprising from our structure of AGR2; both lysine and tyrosine residues corre-sponding to positions 64 and 63, respectively, in AGR2, that are identified in the present paper to have a role in AGR2 dimerization are absent in Erp16/18.

A comparison of the amino acid sequences of AGR2 from 21 species (Fig. S6) shows that amino acid 63 is conserved as Y in amphibians, marsupials and mammals and as a highly conservative substitution to F63 in fish and birds. E60 is conserved in all 21 species from zebrafish to humans, and K64 is conserved in amphibians, marsupials and mammals, but as R in mouse. However, in fish, K64 is hydrophobic W (or L in one case), suggesting that fish AGR2 either does not form dimers or does not form dimers through the same salt bridge, relying perhaps on the hydrophobic interactions involving F or Y at position 63. A comparison of the amino acid sequence of AGR3 from 15 species (Fig. S7) shows that the glutamic acid, which corresponds to position 60 in AGR2, is conserved in 14 out of 15 species, including fish, except as D in Northern Pike; however, the position in AGR3 corresponding to position K64 of AGR2 is not conserved at all, even in human AGR3. This result is consistent with the MALLS experiments on human AGR3, which showed self-association of AGR3 but, in contrast to human AGR2, behaved heterogeneously, with molecular mass varying from 23 kDa to 37 kDa. Thus, AGR3 behaves differently from AGR2 with regard to oligomer formation.

A role of AGR2 dimerization has yet to be defined, in common with some better-characterized members of the PDI family.[25,28,29] The calculated dissociation constant of dimerization in the present paper suggests that AGR2 is dimeric in its normal location of the endoplasmic reticulum, where high concentrations will be maintained. Since one normal client protein of AGR2, the cysteine-rich MUC2,[4] is multimeric, this dimerization may enable AGR2 to present multiple sites for MUC2 binding. It is notable that the dimerization interface of AGR2 is located at the region far from the hydrophobic patch around the catalytic CPHS motif (Fig. 6e), suggesting that this patch is free to interact with other client proteins. Hydrophobic interactions have been reported to be important for the interaction between PDI and a large substrate, creatine kinase. Further-more, at higher concentrations of PDI, the formation of PDI oligomers was associated with a chaperonin activity, rather than with protein folding activity.[31] A similar change in function from folding to chaperone has been reported to be associated with the concentration-dependent oligomerization of the Arabidopsis thioredoxin-like protein, AtTDX.[32] Thus, the dimerization of AGR2 may also play a role in its normal function.

Normal and abnormal roles of AGR2

The results here are consistent with AGR2 exhibiting normal intracellular and abnormal extracellular roles. In normal cells, intracellular AGR2 forms mixed disulfides with mucins, MUC1, MUC2 and MUC5AC,[4,19,20] and thereby contributes to their secretion. In tumor cells, however, $AGR2_{21–175}$ can stimulate cell adhesion activity.[6] The results here are also consistent with reports that AGR2 is not only located on the cell surface in cancer cells using immunofluorescence techniques[33] and se-creted from pancreatic cancer cells,[10] but also found in the blood and possibly urine of, respec-tively, ovarian[16] and prostate[34,35] cancer patients. AGR2, however, is not unique among PDI-related proteins by acting extracellularly at the cell mem-brane. Other proteins that contain thioredoxin folds have been found to play important roles on the cell surface because many cell surface proteins contain disulfide linkages.[36,37] PDI and related endoplas-mic reticulum proteins can influence, through their catalytic activity on the cell surface, the behavior of diverse receptors,[38] proteases[39] and particularly cell adhesion molecules such as integrins.[37] PDI itself catalyses disulfide exchange on integrins to enhance platelet adhesion[40]; blocking cell surface PDI with function-blocking PDI-specific antibodies specifically reduces fibrinogen-mediated platelet aggregation.[41] The unstructured region, amino acids 21–40, of AGR2 is required for the cell-adhesion-promoting activity of AGR2. Thus, it is likely that the unstructured region serves to attach the AGR2 protein to the substratum, allowing the structured part of the protein to mediate cell adhesion, either by interacting, in a dimerization-independent manner, with one of its proposed cell surface targets, for example, dystroglycan or C4.4a,[5] or alternatively, through a specific thior-edoxin activity, to interact with the cell surface. The observation that the SXXS mutation does not reduce the adhesion-promoting effect of AGR2

suggests that the adhesion-promoting effect of AGR2 may be mediated by binding to one of its partner proteins, further highlighting an abnormal extracellular role for AGR2.

However, how AGR2 functions at the tumor cell surface to induce cell adhesion in cancer development and metastasis [6] and to affect patient survival[12–14] is largely unknown. The elucidation of the 3D structure of AGR2 has revealed a separation between potential catalytic redox activity and adhesion function within the context of metastasis and development.

Materials and Methods

Plasmids for expression of proteins

Wild-type cDNA encoding human $AGR2_{21–175}$ (residues 21–175) (GI: 68012759) was cloned into the pTWIN1 vector (NEB), and the $AGR2_{41–175}$ construct was gener-ated by subcloning the relevant part of $AGR2_{21–175}$ cDNA into pET151-D vector (Invitrogen). The E60A $AGR2_{21–175}$, E60A $AGR2_{41–175}$, C81S $AGR2_{41–175}$, K64A $AGR2_{41–175}$ and Y63A $AGR2_{41–175}$ mutations were generated using a QuikChange site-directed mutagenesis kit (Stratagene), in accordance with the manufacturer's instructions. Details of the primers are given in Supplementary Data.

Protein expression and purification

Unlabeled $AGR2_{21–175}$ and $AGR2_{41–175}$, including all mutants, were expressed in LB medium. [U-^{15}N], [U-^{15}N, ^{13}C]$AGR2_{41–175}$ and E60 $AGR2_{41–175}$, [U-^{15}N, ^{13}C, ^{2}H] $AGR2_{41–175}$ and {[U-^{15}N, ^{12}C, ^{2}H], Ile $\delta1$[^{13}CH_3],Leu,Val[^{13}CH_3]}$AGR2_{41–175}$ were prepared for these studies. For [U-^{15}N] and [U-^{15}N, ^{13}C]proteins, expressions were carried out in M9 medium containing 1 g/L ^{15}NH_4Cl with either 4 g/L ^{12}C-glucose or ^{13}C-glucose. For the production of [U-^{15}N, ^{13}C, ^{2}H]$AGR2_{41–175}$, Silantes OD2 medium was used (Silantes GmbH). The bacteria were grown to an OD_{600} of 0.5–0.6 at 37 °C, and protein expression was induced with 1 mM IPTG overnight at 18 °C.

Wild-type {[U-^{15}N, ^{13}C, ^{2}H], Ile $\delta1$[^{13}CH_3],Leu,Val[^{13}CH_3]} $AGR2_{41–175}$ was prepared using a modified protocol to that described by Tugarinov et al .[42] A 5-mL culture of LB medium was inoculated with a freshly transformed colony of BL21 Star (DE3) cells and grown at 37 °C to an OD_{600} of 0.7. The cells were centrifuged at 1200g at room temperature, resuspended in 20 mL of unlabeled M9/H_2O medium to a starting OD_{600} of 0.1 and incubated to OD_{600} of 0.6. The cells were again centrifuged at 1200g, resuspended in 100 mL M9/D_2O medium, prepared with [^{2}H, ^{13}C]glucose and ^{15}NH_4Cl to a starting OD_{600} of 0.1 and grown to an OD_{600} of 0.5. The culture was diluted to 200 mL with M9/D_2O medium and grown to an OD_{600} of 0.5. The culture was diluted to 1 L using the same M9/D_2O medium and grown to an OD_{600} of 0.25, after which 2-keto-3-d_2-1,2,3,4-^{13}C-butyrate and 2-keto-3-methyl-^{13}C-d_1-4-^{13}C-butyrate were added to final concentrations of 70 mg/L and 120 mg/L, respectively. The culture was incubated further for 1 h, and protein expression was induced with 1 mM IPTG for 7–8 h at 30 °C. The process of cell growth took 30 h from start to completion.

Cells were harvested and sonicated in 20 mM Tris–HCl (pH 7.5), 500 mM NaCl and 20 mM imidazole; the soluble fraction was obtained by centrifugation at 27,000g for 45 min. The supernatant was applied to a charged HisTrap FF 5-mL affinity column (GE Healthcare); washed with 20 mM Tris–HCl (pH 7.5), 500 mM NaCl, 20 mM imidaz-ole and eluted in 20 mM Tris (pH 7.5), 500 mM NaCl with a linear imidazole gradient from 20 mM to 500 mM. The eluted protein was cleaved with recombinant tobacco etch virus protease overnight at 4 °C and further purified by gel filtration on a Superdex-75 column (GE Healthcare). Protein concentrations were determined from the absor-bance at 280 nm.

All wild-type $AGR2_{41–175}$ samples for NMR were prepared in 20 mM 4-morpholineethanesulfonic acid (pH 6.5), 100 mM NaCl, 2 mM DTT, 0.02% (w/v) NaN_3 and 10% (v/v) D_2O or 99.9% D_2O, at a final protein concentration of 1 mM. E60A $AGR2_{41–175}$ samples were prepared in 20 mM phosphate buffer (pH 6.0), 150 mM NaCl, 3 mM DTT, 0.02% (w/v) NaN_3 and 10% (v/v) D_2O or 99.9% D_2O, at a final protein concentration of 0.3–1.2 mM.

Mixed-labeled samples of $AGR2_{41–175}$ were made by mixing equal concentrations of labeled and unlabeled wild-type $AGR2_{41–175}$ at room temperature for between 1 h and 120 h and with gentle heating up to 40 °C for between 2 min and 15 min. Mild denaturing conditions were also used in which hexahistidine-tagged ^{13}C/^{15}N-labeled $AGR2_{41–175}$ and untagged unlabeled $AGR2_{41–175}$ were incubated in 2 M urea overnight. The mixture was then thoroughly dialyzed into non-denaturing buffer over a period of 48 h. Nickel-affinity chromatography was then used to isolate hexahistidine-tagged protein and eluted with an imidazole gradient. Two peaks from the elution gradient were observed, one corresponding to mixed-labeled $AGR2_{41–175}$ and one corresponding to purely labeled wild-type $AGR2_{41–175}$.

Sedimentation equilibrium

Purified $AGR2_{41–175}$ was buffer exchanged into 20 mM sodium phosphate (pH 6.5) with 150 mM NaCl, on a Superdex-75 10/300 (GE Healthcare) column. The eluted fractions were pooled and used for the dilution series. The final concentrations were 62 μM, 31 μM and 6 μM, which were loaded into 6-well Epon-filled centerpieces with quartz windows. The samples were centrifuged at 15,000 rpm, 22,000 rpm and 30,000 rpm for 15 h at 20 °C performing scans at 290 nm, 280 nm and 232 nm for the three concentrations and analyzed using the Sedfit/ Sedphat suite of programs.[43]

Size-exclusion chromatography MALLS

Proteins were chromatographed on Superdex-75 300 mm×10 mm gel-filtration columns (GE Healthcare Life Sciences) equilibrated in 50 mM phosphate and 50 mM NaCl (pH 6.8) at 0.71 mL/min. Elution was monitored by a Wyatt EOS 18-angle laser photometer (Wyatt Technology, Santa Barbara, CA), an Optilab rEX refractive index detector and a Jasco UV-2077 Plus UV/Vis spectrophotometer (Jasco, Easton, MD); these were

coupled to a quasi elastic light-scattering detector for simultaneous measurement of hydrodynamic radius. Molar mass measurements were performed using both Astra 5.3.2.16 software (Wyatt Technology) and the "three detector method". Values of mass and hydrodynamic radius are expressed as mean±standard error.

NMR spectroscopy

All spectra were acquired at 298 K on Bruker Avance III 600- and 800-MHz spectrometers. For $AGR2_{41–175}$, sequence-specific backbone resonance assignment was obtained using transverse relaxation optimized spectros-copy-based multidimensional heteronuclear NMR exper-iments [HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO]. Leucine, isoleucine and valine side-chain methyl group assignments were obtained from a 3D (H)CCH_3–TOCSY experiment with 22.8 ms mixing time.[22] For E60A $AGR2_{41–175}$, sequence-specific back-bone resonance assignment was obtained using stan-dard NMR experiments [HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, HBHANH, HBHA (CO)NH]. Side-chain assignment was obtained using a 3D HCCH–TOCSY experiment. NOEs were derived from 3D ^{15}N- and ^{13}C-edited NOESY–HSQC and 3D HMQC-NOESY-HMQC experiments with 200 ms mixing times. Paramagnetic surface mapping studies were carried out in 0.5–12 mM gadolinium complexed with diethylenetria-mine pentaacetic acid (Gd-DTPA) and 5–30 mM TEMPO, as paramagnetic relaxation agents. One-bond ^{1}H–^{15}N RDCs were obtained from [^{1}H,^{15}N] in-phase/ antiphase HSQC spectra. Values of the axial and rhombic components of the molecular alignment tensor for RDCs were obtained using the program PALES.[44]

RDC measurements

For dimeric $AGR2_{41–175}$, RDCs were collected in two separate media: Pf1 bacteriophage (21 mg/mL) (Asla Biotech) and 5% (v/v) pentaethylene glycol monododecyl ether:hexanol in a 0.96:1 ratio. Protein concentrations in the range 0.3–0.6 mM were used in these experiments. The final values of the axial and rhombic components in the refinement stage were, respectively, −18.95 Hz and 0.40 Hz. For E60A $AGR2_{41–175}$, RDCs were collected in Pf1 bacteriophage medium and the final values of the axial and rhombic components in the refinement stage were, respectively, 14.05 Hz and 0.46 Hz.

NMR assignments and structure calculations

All NMR spectra were processed with TopSpin (Bruker) and analyzed using the CCPN Analysis Package. $AGR2_{41–175}$ and E60A $AGR2_{41–175}$ structure calculations were carried out using the CNS.[45] The structures of $AGR2_{41–175}$ monomer subunit and E60A $AGR2_{41–175}$ monomer were carried out using the CNS program with the IUPAC PARALLHDGv5.3 and TOPALLHDGv5.3 parameter sets. Backbone torsion angles were derived from analysis of C^{α}, C^{β} and C′ chemical shifts using the TALOS+ program.[46] The final set of calculated structures was obtained from water-refinement calculations in XPLOR-NIH.

Dimer modeling

An RDC-assisted modeling method, suitable for the structure determination of symmetrical dimers,[23] was used to orientate the monomeric subunits to obtain the dimeric structure. The method relies on the fact that an aligned protein has one of the principle axes of the alignment tensor orientated in parallel with the symmetry axis of the multimer. Hence, when the alignment tensors of a protein in two or more different alignment media are measured, the common alignment tensor is indica-tive of the symmetry axis. We measured the RDCs of $AGR2_{41–175}$ in two media, Pf1 bacteriophage (21 mg/ mL) (Asla Biotech) and 5% (v/v) pentaethylene glycol monododecyl ether:hexanol. The REDCAT[47] program and the monomeric structure of $AGR2_{41–175}$ were used to determine the alignment tensors in each medium. Directions of the axes were plotted on a Sauson-Flamsteed plot; the overlap of the PEG Sxx and bacteriophage Syy alignment tensor axes indicated the C2 symmetry axis of $AGR2_{41–175}$. Homodimer models in agreement with this symmetry restraint were then constructed using a grid-search algorithm, and these models were evaluated based on the quality of the interface between the subunits and the agreement with the experimental data.

The VMD software package[48] was used to generate models of the homodimer. The lowest-energy structure of the ensemble of calculated monomeric subunit $AGR2_{41–175}$ structure was selected and rotated 180° about the derived symmetry axis to generate the orientation of the second subunit of the $AGR2_{41–175}$ homodimer. The grid search was performed using 230 steps in the x- and z-axes (1 step equivalent to 1 Å), generating a total of 3002 dimeric models with different x and z grid points. Generated models were rejected if intermolecular backbone atoms were b 4 Å or if the closest intermolecular atomic distance was N 2 Å, in accordance with Ref. 23. Relaxation of the side chains at the interface was performed with 500 ps molecular dynamics simulation and 100 steps of energy minimi-zation using the NAMD program[49] and the CHARMM22 force field software. The calculated models were then evaluated according to the correlation between mea-sured and simulated RDC values using PALES and a residue-pairing score for each model.[50] Rigid-body minimization was performed on the 20 best dimer structures selected from the residue-pairing and RDC correlation scores; 75.4% of the non-proline and non-glycine residues are in the most favored regions of the Ramachandran plot, 22.2% are in the additionally allowed regions, 0.8% are in the generously allowed regions and 1.6% are in the disallowed regions.

Biochemical analyses

Assay of free sulfhydryl groups was carried out using the Ellman's assay, which is a colorimetric assay based on the reaction between free thiol groups and DTNB.[51] AGR2 (3.5 μM) and DTNB (4 mg/mL) in 0.1 M phosphate buffer (pH 8.0) (total volume of 2.2 mL) were incubated for 5 min, and the absorbance was measured at 412 nm. Bovine RNase A and C81S $AGR2_{41–175}$ were used as negative controls in these assays.

Cell adhesion assays

Cell adhesion assays were performed as described by Liu et al .[6] 24-well tissue culture plates with 1 nmol, or a specified concentration of AGR2 protein, by air drying at 37 °C. Before use, coated wells were washed once with phosphate-buffered saline and left to air dry at room temperature. Cells were grown to 60–70% confluence, trypsinized and counted using a Coulter counter (Beckman Coulter). The cells were then resuspended to 2×10^5 cells/mL, recounted and 1 mL of cell suspension added to each well. Following a 60-min incubation at 37 °C, cells were washed three times with phosphate-buffered saline, 5×10^{-4}% (w/v) trypsin was added and the number of released cells was determined. Adhesion efficiency was calculated as the percentage of the total number of cells added to the well that had adhered during 60 min. Error bars represent the standard errors of the means from three independent experiments carried out with duplicate wells.

Accession numbers

Accession codes are as follows: wild-type $AGR2_{41–175}$, PDB ID: 2LNS; E60A $AGR2_{41–175}$, PDB ID: 2LNT; wild-type $AGR2_{41–175}$ and E60A $AGR2_{41–175}$ chemical shifts, BioMagResBank codes 18178 and 18179, respectively.

Acknowledgements

We thank Dr. Marie Phelan for helpful discussions and the University of Liverpool for its support of the NMR Center for Structural Biology. We also thank Xu Wang (Department of Chemistry and Biochemistry, University of Alberta) for providing the scripts for homodimer structure calculation and helpful discussions. This work was supported by the North West Cancer Research Fund (Grant No. C758 to B.R.B., D.L.B., L.Y.-L. and P.S.R.) and a Medical Research Council studentship to C.C.

Author Contributions. Conceived and designed the experiments: B.R.B., P.S.R., D.L.B. and L.Y.-L. Performed the experiments: P.P., C.C., D.L.B. and T.A.J. Analyzed the data: P.P., T.J., C.C. and L.Y.-L. Wrote the paper: P.P., L.Y.-L., P.S.R. and B.R.B.

Conflict of Interest Statement. The authors declare that there are no conflicts of interests.

Supplementary Data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2012.12.009

Received 28 June 2012; Received in revised form 13 December 2012; Accepted 15 December 2012
Available online 26 December 2012

Keywords:
adhesion;
AGR2;
dimer;
NMR
thioredoxin

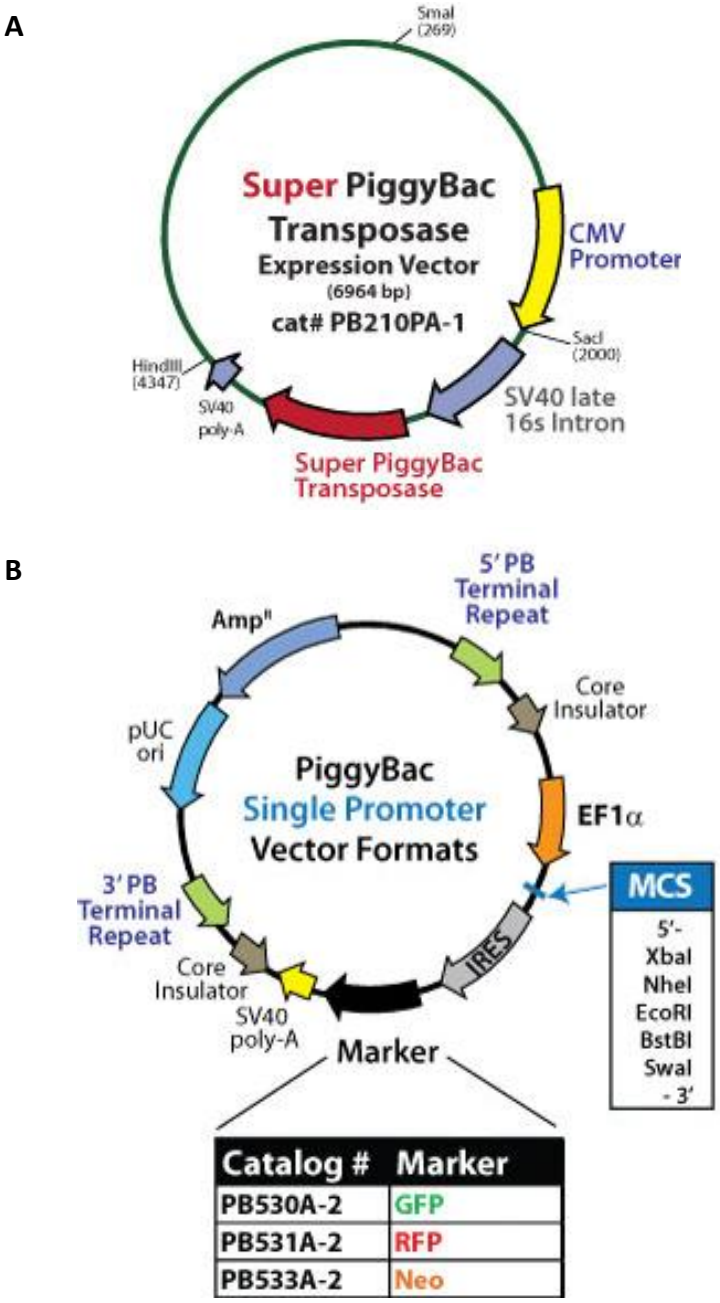
Abbreviations used:
AGR2, anterior gradient 2; PDI, protein disulfide isomerase; HMQC, heteronuclear multiple quantum coherence; HSQC, heteronuclear single quantum coherence; NOE, nuclear Overhauser enhancement; 3D, three-dimensional; NOESY, NOE spectroscopy; RDC, residual dipolar coupling; PEG, polyethylene glycol; MALLS, multi-angle laser light scattering; PDB, Protein Data Bank.

References

1. Aberger, F., Weidinger, G., Grunz, H. & Richter, K. (1998). Anterior specification of embryonic ectoderm: the role of the Xenopus cement gland-specific gene XAG-2. Mech. Dev. 72, 115–130.
2. Kumar, A., Godwin, J. W., Gates, P. B., Garza-Garcia, A. A. & Brockes, J. P. (2007). Molecular basis for the nerve dependence of limb regeneration in an adult vertebrate. Science, 318, 772–777.
3. Komiya, T., Tanigawa, Y. & Hirohashi, S. (1999). Cloning of the gene gob-4, which is expressed in intestinal goblet cells in mice. Biochim. Biophys. Acta, 1444, 434–438.
4. Park, S. W., Zhen, G., Verhaeghe, C., Nakagami, Y., Nguyenvu, L. T., Barczak, A. J. et al. (2009). The protein disulfide isomerase AGR2 is essential for production of intestinal mucus. Proc. Natl Acad. Sci. USA, 106, 6950–6955.
5. Fletcher, G. C., Patel, S., Tyson, K., Adam, P. J., Schenker, M., Loader, J. A. et al. (2003). hAG-2 and hAG-3, human homologues of genes involved in differentiation, are associated with oestrogen receptor-positive breast tumours and interact with metastasis gene C4.4a and dystroglycan. Br. J. Cancer, 88, 579–585.
6. Liu, D., Rudland, P. S., Sibson, D. R., Platt-Higgins, A. & Barraclough, R. (2005). Human homologue of cement gland protein, a novel metastasis inducer associated with breast carcinomas. Cancer Res. 65, 3796–3805.
7. Fritzsche, F. R., Dahl, E., Dankof, A., Burkhardt, M., Pahl, S., Petersen, I. et al. (2007). Expression of AGR2 in non small cell lung cancer. Histol. Histopathol. 22, 703–708.
8. Park, K., Chung, Y. J., So, H., Kim, K., Park, J., Oh, M. et al. (2011). AGR2, a mucinous ovarian cancer marker, promotes cell proliferation and migration. Exp. Mol. Med. 43, 91–100.
9. Pohler, E., Craig, A. L., Cotton, J., Lawrie, L., Dillon, J. F., Ross, P. et al. (2004). The Barrett's antigen anterior

gradient-2 silences the p53 transcriptional response to DNA damage. Mol. Cell. Proteomics, 3, 534–547.
10. Ramachandran, V., Arumugam, T., Wang, H. & Logsdon, C. D. (2008). Anterior gradient 2 is expressed and secreted during the development of pancreatic cancer and promotes cancer cell survival. Cancer Res. 68, 7811–7818.
11. Zhang, Y., Forootan, S. S., Liu, D., Barraclough, R., Foster, C. S., Rudland, P. S. & Ke, Y. (2007). Increased expression of anterior gradient-2 is signif-icantly associated with poor survival of prostate cancer patients. Prostate Cancer Prostatic Dis. 10, 293–300.
12. Barraclough, D. L., Platt-Higgins, A., de Silva Rudland, S., Barraclough, R., Winstanley, J., West, C. R. & Rudland, P. S. (2009). The metastasis-associated anterior gradient 2 protein is correlated with poor survival of breast cancer patients. Am. J. Pathol. 175, 1848–1857.
13. Hrstka, R., Nenutil, R., Fourtouna, A., Maslon, M. M., Naughton, C., Langdon, S. et al. (2010). The pro-metastatic protein anterior gradient-2 predicts poor prognosis in tamoxifen-treated breast cancers. Onco-gene, 29, 4838–4847.
14. Innes, H. E., Liu, D., Barraclough, R., Davies, M. P., O'Neill, P. A., Platt-Higgins, A. et al. (2006). Signif-icance of the metastasis-inducing protein AGR2 for outcome in hormonally treated breast cancer patients. Br. J. Cancer, 94, 1057–1065.
15. Smirnov, D. A., Zweitzig, D. R., Foulk, B. W., Miller, M. C., Doyle, G. V., Pienta, K. J. et al. (2005). Global gene expression profiling of circulating tumor cells. Cancer Res. 65, 4993–4997.
16. Edgell, T. A., Barraclough, D. L., Rajic, A., Dhulia, J., Lewis, K. J., Armes, J. E. et al. (2010). Increased plasma concentrations of anterior gradient 2 protein are positively associated with ovarian cancer. Clin. Sci. 118, 717–725.
17. Hapangama, D. K., Raju, R. S., Valentijn, A. J., Barraclough, D., Hart, A., Turner, M. A. et al. (2012). Aberrant expression of metastasis-inducing proteins in ectopic and matched eutopic endometrium of women with endometriosis: implications for the pathogenesis of endometriosis. Hum. Reprod. 27, 394–407.
18. Raykhel, I., Alanen, H., Salo, K., Jurvansuu, J., Nguyen, V. D., Latva-Ranta, M. & Ruddock, L. (2007). A molecular specificity code for the three mammalian KDEL receptors. J. Cell Biol. 179, 1193–1204.
19. Norris, A. M., Gore, A., Balboni, A., Young, A., Longnecker, D. S. & Korc, M. (2012). AGR2 is a SMAD4-suppressible gene that modulates MUC1 levels and promotes the initiation and progression of pancreatic intraepithelial neoplasia. Oncogene, http: //dx.doi.org/10.1038/onc.2012.394.
20. 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. Am. J. Respir. Cell Mol. Biol. 47, 178–185.
21. Persson, S., Rosenquist, M., Knoblach, B., Khosravi-Far, R., Sommarin, M. & Michalak, M. (2005). Diversity of the protein disulfide isomerase family: identification of breast tumor induced Hag2 and Hag3 as novel members of the protein family. Mol. Phylogenet. Evol. 36, 734–740.
22. Uhrin, D., Uhrinova, S., Leadbeater, C., Nairn, J., Price, N. C. & Barlow, P. N. (2000). 3D HCCH-TOCSY for resonance assignment of methyl-containing side chains in ^{13}C-labeled proteins. J. Magn. Reson. 142, 288–293.
23. Wang, X., Bansal, S., Jiang, M. & Prestegard, J. H. (2008). RDC-assisted modeling of symmetric protein homo-oligomers. Protein Sci. 17, 899–907.
24. Kozlov, G., Maattanen, P., Thomas, D. Y. & Gehring, K. (2010). A structural overview of the PDI family of proteins. FEBS J. 277, 3924–3936.
25. Rowe, M. L., Ruddock, L. W., Kelly, G., Schmidt, J. M., Williamson, R. A. & Howard, M. J. (2009). Solution structure and dynamics of ERp18, a small endoplas-mic reticulum resident oxidoreductase. Biochemistry, 48, 4596–4606.
26. Wang, L., Vavassori, S., Li, S., Ke, H., Anelli, T., Degano, M. et al. (2008). Crystal structure of human ERp44 shows a dynamic functional modulation by its carboxy-terminal tail. EMBO Rep. 9, 642–647.
27. Lappi, A. K., Lensink, M. F., Alanen, H. I., Salo, K. E., Lobell, M., Juffer, A. H. & Ruddock, L. W. (2004). A conserved arginine plays a role in the catalytic cycle of the protein disulphide isomerases. J. Mol. Biol. 335, 283–295.
28. Wallis, A. K., Sidhu, A., Byrne, L. J., Howard, M. J., Ruddock, L. W., Williamson, R. A. & Freedman, R. B. (2009). The ligand-binding b′ domain of human protein disulphide-isomerase mediates homodimerization. Protein Sci. 18, 2569–2577.
29. Barak, N. N., Neumann, P., Sevvana, M., Schutkowski, M., Naumann, K., Malesevic, M. et al. (2009). Crystal structure and functional analysis of the protein disulfide isomerase-related protein ERp29. J. Mol. Biol. 385, 1630–1642.
30. Jeong, W., Lee, D. Y., Park, S. & Rhee, S. G. (2008). ERp16, an endoplasmic reticulum-resident thiol-disul-fide oxidoreductase: biochemical properties and role in apoptosis induced by endoplasmic reticulum stress. J. Biol. Chem. 283, 25557–25566.
31. Zhao, T. J., Ou, W. B., Xie, Q., Liu, Y., Yan, Y. B. & Zhou, H. M. (2005). Catalysis of creatine kinase refolding by protein disulfide isomerase involves disulfide cross-link and dimer to tetramer switch. J. Biol. Chem. 280, 13470–13476.
32. Lee, J. R., Lee, S. S., Jang, H. H., Lee, Y. M., Park, J. H., Park, S. C. et al. (2009). Heat-shock dependent oligomeric status alters the function of a plant-specific thioredoxin-like protein AtTDX. Proc. Natl Acad. Sci. USA, 106, 5978–5983.
33. Dumartin, L., Whiteman, H. J., Weeks, M. E., Hariharan, D., Dmitrovic, B., Iacobuzio-Donahue, C. A. et al. (2011). AGR2 is a novel surface antigen that promotes the dissemination of pancreatic cancer cells through regulation of cathepsins B and D. Cancer Res. 71, 7091–7102.
34. Wayner, E. A., Quek, S. I., Ahmad, R., Ho, M. E., Loprieno, M. A., Zhou, Y. et al. (2011). Development of an ELISA to detect the secreted prostate cancer biomarker AGR2 in voided urine. Prostate, 72, 1023–1034.

35. Bu, H., Bormann, S., Schäfer, G., Horninger, W., Massoner, P., Neeb, A. et al. (2010). The anterior gradient 2 (AGR2) gene is overexpressed in prostate cancer and may be useful as a urine sediment marker for prostate cancer detection. Prostate, 71, 575–587.
36. Turano, C., Coppari, S., Altieri, F. & Ferraro, A. (2002). Proteins of the PDI family: unpredicted non-ER locations and functions. J. Cell. Physiol. 193, 154–163.
37. Jordan, P. A. & Gibbins, J. M. (2006). Extracellular disulfide exchange and the regulation of cellular function. Antioxid. Redox Signaling, 8, 312–324.
38. Couderc, B., Prats, H., Bayard, F. & Amalric, F. (1991). Potential oncogenic effects of basic fibroblast growth factor requires cooperation between CUG and AUG-initiated forms. Cell Regul. 2, 709–718.
39. Willems, S. H., Tape, C. J., Stanley, P. L., Taylor, N. A., Mills, I. G., Neal, D. E. et al. (2010). Thiol isomerases negatively regulate the cellular shedding activity of ADAM17. Biochem. J. 428, 439–450.
40. Lahav, J., Wijnen, E. M., Hess, O., Hamaia, S. W., Griffiths, D., Makris, M. et al. (2003). Enzymatically catalyzed disulfide exchange is required for platelet adhesion to collagen via integrin α2β1. Blood, 102, 2085–2092.
41. Lahav, J., Jurk, K., Hess, O., Barnes, M. J., Farndale, R. W., Luboshitz, J. & Kehrel, B. E. (2002). Sustained integrin ligation involves extracellular free sulfhydryls and enzymatically catalyzed disulfide exchange. Blood, 100, 2472–2478.
42. Tugarinov, V., Kanelis, V. & Kay, L. E. (2006). Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. Nat. Protoc. 1, 749–754.
43. Vistica, J., Dam, J., Balbo, A., Yikilmaz, E., Mariuzza, R. A., Rouault, T. A. & Schuck, P. (2004). Sedimen-tation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. Anal. Biochem. 326, 234–256.
44. Zweckstetter, M. (2008). NMR: prediction of molecular alignment from structure using the PALES software. Nat. Protoc. 3, 679–690.
45. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, L., Gros, P., Grosse-Kunstleve, R. W. et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr., Sect. D: Biol. Crystallogr. 54, 905–921.
46. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. (2009). TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J. Biomol. NMR, 44, 213–223.
47. Valafar, H. & Prestegard, J. H. (2004). REDCAT: a residual dipolar coupling analysis tool. J. Magn. Reson. 167, 228–241.
48. Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graphics, 14, 27–28.
49. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E. et al. (2005). Scalable molecular dynamics with NAMD. J. Comput. Chem. 26, 1781–1802.
50. Moont, G., Gabb, H. A. & Sternberg, M. J. (1999). Use of pair potentials across protein interfaces in screen-ing predicted docked complexes. Proteins, 35, 364–373.
51. Habeeb, A. F. (1972). Reaction of protein sulfhydryl groups with Ellman's reagent. Methods Enzymol. 25, 457–464.

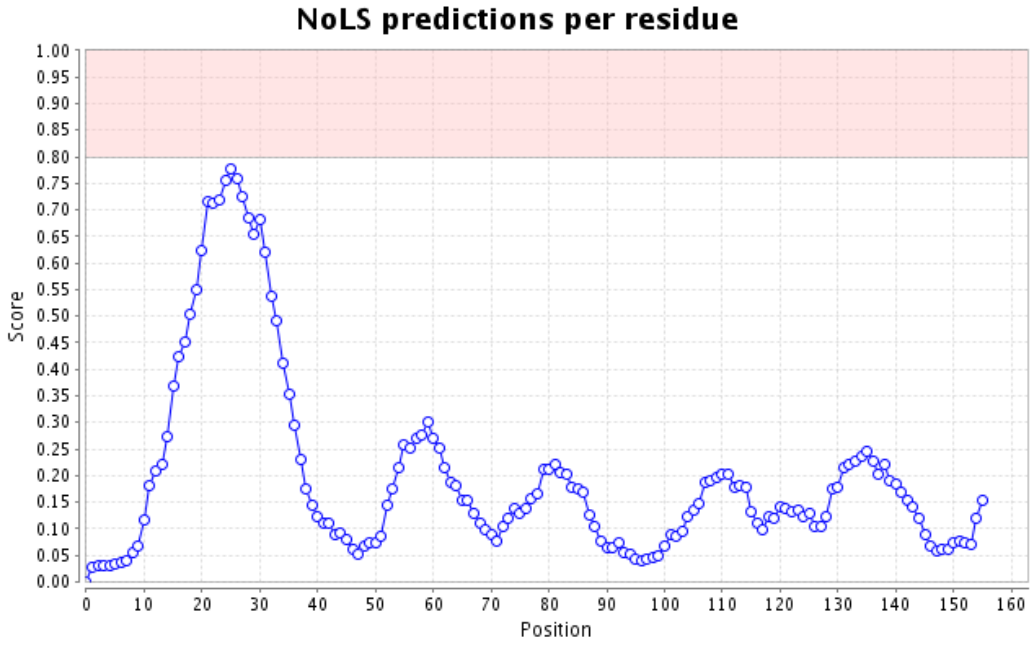


Appendix 2. ***PiggyBac transposon vector maps***. **(A)** Transposase expression vector. **(B)** Gene of interest expression vector. Note that the vector used in these studies contained the neomycin resistance marker (PB533A-2). Reproduced from Systems Biosciences [739].

NoLS predictions for AGR2

(these predictions are based on sequence and Jpred predictions)

This protein has no predicted NoLS.



Appendix 3. ***Prediction of AGR2 nucleolar localisation sequences (NoLS) by NoD server [731].*** Full length AGR2 primary protein sequence was submitted to http://www.compbio.dundee.ac.uk/www-nod/ for NoLS prediction. A score above 0.8 represent a candidate NoLS.

cNLS Mapper Result

Predicted NLSs in query sequence
MEKIPVSAFLLLVALSYTLARDTTVKPGAKKDTKDSRPKLPQTLSRGWGD 50 QLIWTQTYEEALYKSKTSNKPLMIIHHLDECPHSQALKKVFAENKEIQKL 100 AEQFVLLNLVYETTDKHLSPDGQYVPRIMFVDPSLTVRADITGRYSNRLY 150 AYEPADTALLLDNMKKALKLLKTEL 175

Predicted monopartite NLS		
Pos.	Sequence	Score

Predicted bipartite NLS		
Pos.	Sequence	Score
138	RADITGRYSNRLYAYEPADTALLLDNMKKALKL	5.7
144	RYSNRLYAYEPADTALLLDNMKKALKLLKTE	5.1

Appendix 4. ***Prediction of AGR2 nuclear localisation sequences (NLS) by cNLS Mapper server [575].*** Full length AGR2 primary protein sequence was submitted to http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi for NLS prediction. The predicted NLS is highlighted in red. Green fluorescent protein tagged with an NLS with a score above 8 localises exclusively to the nucleus, with a score of 3-5 to the nucleus and cytoplasm and with a score of 1-2 to the cytoplasm.

Pdb file name: 2LNS.pdb

Protein-Protein hydrophobic Interaction

POS	RES	CHA	POS	RES	CHA
53	ILE	A	53	ILE	B
63	TYR	A	63	TYR	B

NO PROTEIN-PROTEIN DISULPHIDE BRIDGES ARE FOUND

NO PROTEIN-PROTEIN HYDROGEN BONDS ARE FOUND

NO PROTEIN-PROTEIN MAIN CHAIN-SIDE CHAIN HYDROGEN BONDS ARE FOUND

Protein-protein Side chain-Side chain hydrogen bonds

POS	CHAIN	RES	ATOM	POS	CHAIN	RES	ATOM	MO	Dd-a	Dh-a	A(d-H-N)	A(a-O=C)
54	A	TRP	NE1	50	B	ASP	OD2	-	2.94	2.32	125.15	999.99
64	A	LYS	NZ	60	B	GLU	OE1	-	2.24	9.99	999.99	999.99
64	A	LYS	NZ	60	B	GLU	OE2	-	2.23	9.99	999.99	999.99
54	B	TRP	NE1	50	A	ASP	OD2	-	2.78	2.03	139.32	999.99
64	B	LYS	NZ	60	A	GLU	OE1	-	2.39	9.99	999.99	999.99

Appendix 5 (continued overleaf). ***Measurement of protein interactions by Protein Interaction Calculator (PIC) server [732].*** The AGR2 $_{41\text{-}175}$ structure file (2LNS.pdb) was submitted to http://pic.mbu.iisc.ernet.in/. POS: position. RES: residue. CHA: chain. DIST: distance. ANGL: angle. Dd-a: distance between donor and acceptor (Å). Dh-a: distance between donor hydrogen and acceptor (Å). A (d-H-N): angle between donor and acceptor amine group (degrees). A value of 1 (999.99) indicates the angle is not defined. A (d-O=C): angle between donor and acceptor carbonyl group (degrees). A value of 999.99 indicates that the angle is not defined.

64 B LYS NZ 60 A GLU OE2 - 2.33 9.99 999.99 999.99

Protein-Protein Ionic Interaction

POS RES CHA POS RES CHA

60 GLU A 64 LYS B

64 LYS A 60 GLU B

Protein-Protein aromatic-aromatic Interaction

POS RES CHA POS RES CHA DIST ANGL

63 TYR A 63 TYR B 6.54 72.55

NO PROTEIN-PROTEIN AROMATIC-SULPHUR INTERACTIONS ARE FOUND

Protein-Protein Cation-pi Interaction

POS RES CHA POS RES CHA DIST ANGL

54 TRP A 46 ARG B 5.36 68.02

54 TRP B 46 ARG A 5.06 65.58

Appendix 5 (cont.). ***Measurement of protein interactions by Protein Interaction Calculator (PIC) server [732] (cont.).*** The AGR2 $_{41\text{-}175}$ structure file (2LNS.pdb) was submitted to http://pic.mbu.iisc.ernet.in/. POS: position. RES: residue. CHA: chain. DIST: distance. ANGL: angle. Dd-a: distance between donor and acceptor (Angstrom (Å)). Dh-a: distance between donor hydrogen and acceptor (Å). A (d-H-N): angle between donor and acceptor amine group (degrees). A value of 1 (999.99) indicates the angle is not defined. A (d-O=C): angle between donor and acceptor carbonyl group (degrees). A value of 999.99 indicates that the angle is not defined.

Welcome to the CaPTURE program. Copyright 1999 Justin P. Gallivan.

********** Report for the protein pdb2lns.ent **********

This protein has 280 amino acids of which:

24	(8.6%) are Lysine
10	(3.6%) are Arginine
8	(2.9%) are Phenylalanine
14	(5.0%) are Tyrosine
4	(1.4%) are Tryptophan

This protein has 2 chains. They are:

Chain A with 140 residues.

Chain B with 140 residues.

The protein has an apparent molecular weight of 32427 D

********** Cation-Pi Summary **********

Number of ARG/PHE interacting pairs: 0

Number of energetically significant ARG/PHE cation-pi interactions: 0

Number of ARG/TYR interacting pairs: 6

Number of energetically significant ARG/TYR cation-pi interactions: 0

Number of ARG/TRP interacting pairs: 2

Number of energetically significant ARG/TRP cation-pi interactions: 0

Number of LYS/PHE interacting pairs: 2

Number of energetically significant LYS/PHE cation-pi interactions: 0

Number of LYS/TYR interacting pairs: 2

Number of energetically significant LYS/TYR cation-pi interactions: 0

Number of LYS/TRP interacting pairs: 2

Number of energetically significant LYS/TRP cation-pi interactions: 0

* There is a total of 0 energetically significant cation-pi interactions.

* There are 0 cation-pi interactions with E(es) <= -2.0 kcal/mol

* There are 0 cation-pi interactions with E(es) between -2.0 and -1.0 kcal/mol and with E(vdw) <= -1.0 kcal/mol

Appendix 6. ***Measurement of cation-π interaction in AGR2 structure by CaPTURE server [565].*** The AGR2 $_{41\text{-}175}$ structure file (2LNS.pdb) was submitted to http://capture.caltech.edu/.

HBPLUS Hydrogen Bond Calculator v 3.2
(c) I McDonald, D Naylor, D Jones and J Thornton 1993 All Rights Reserved.
Citing HBPLUS in publications that use these results is condition of use.
2LNS <- Brookhaven Code "2LNS.pdb" <- PDB file

<---DONOR--->				<-ACCEPTOR-->		atom			^						^		H-
c	i						cat	<-CA-CA->		^		H-A-AA		^		H-	
h	n	atom	resd	res		DA	||	num		DHA	H-A	angle	D-A-AA		Bond		
n	s	type	num	typ		dist	DA	aas	dist	angle	dist		angle		num		
A0046-ARG	N	A0043-THR	O	2.56	MM	3	4.69	133.3	1.79	109.7	125.5		1				
A0046-ARG	NH2	A0050-ASP	OD2	2.43	SS	4	7.62	129.3	1.67	146.5	129.9		2				
A0052-LEU	N	A0049-GLY	O	2.83	MM	3	4.58	111.0	2.33	98.6	117.3		3				
A0051-GLN	NE2	A0050-ASP	OD2	2.62	SS	1	3.74	152.0	1.72	104.0	101.4		4				
B0054-TRP	NE1	A0050-ASP	OD2	2.78	SS	-1	7.48	147.2	1.90	92.4	98.2		5				
A0055-THR	N	A0053-ILE	O	3.10	MM	2	5.83	130.7	2.37	106.2	92.4		6				
A0057-THR	N	A0056-GLN	OE1	2.68	MS	1	3.74	159.0	1.74	114.2	106.7		7				
A0056-GLN	NE2	A0108-ASN	O	2.86	SM	52	6.00	145.5	2.00	160.6	159.2		8				
A0061-ALA	N	A0058-TYR	O	2.46	MM	3	5.00	115.3	1.88	128.4	147.4		9				
A0062-LEU	N	A0058-TYR	O	3.10	MM	4	6.40	152.8	2.19	138.6	146.6		10				
A0059-GLU	N	A0059-GLU	OE2	2.48	MS	0	-1.00	131.2	1.72	97.8	91.4		11				
A0062-LEU	N	A0059-GLU	O	2.52	MM	3	4.90	110.2	2.00	112.9	132.9		12				
A0063-TYR	N	A0059-GLU	O	2.69	MM	4	5.92	162.4	1.74	153.8	159.4		13				
A0064-LYS	N	A0060-GLU	O	2.88	MM	4	6.16	168.1	1.91	171.1	167.7		14				
B0064-LYS	NZ	A0060-GLU	OE2	2.33	SS	-1	9.06	115.9	1.68	115.4	92.5		15				
A0065-SER	N	A0061-ALA	O	3.10	MM	4	6.16	160.2	2.15	145.2	151.4		16				
A0066-LYS	N	A0062-LEU	O	2.69	MM	4	5.74	166.6	1.72	157.1	161.9		17				
A0067-THR	N	A0063-TYR	O	3.04	MM	4	6.32	156.0	2.11	158.7	165.2		18				
A0067-THR	OG1	A0064-LYS	O	2.49	SM	3	5.29	136.0	1.70	141.1	138.4		19				
A0064-LYS	NZ	A0067-THR	OG1	2.89	SS	3	5.29	111.4	2.34	95.2	103.2		20				
A0068-SER	N	A0065-SER	O	2.86	MM	3	5.57	153.1	1.94	122.3	130.7		21				
A0069-ASN	N	A0066-LYS	O	2.59	MM	3	4.36	141.2	1.75	101.4	114.9		22				
A0066-LYS	NZ	A0069-ASN	OD1	2.44	SS	3	4.36	121.1	1.74	111.1	130.7		23				
A0066-LYS	NZ	A0133-PRO	O	2.63	SM	67	7.21	134.2	1.80	148.5	160.1		24				
A0069-ASN	ND2	A0069-ASN	O	2.55	SM	0	-1.00	116.1	1.96	127.1	113.7		25				
A0072-LEU	N	A0131-VAL	O	2.74	MM	59	4.47	162.6	1.78	141.8	145.7		26				
A0131-VAL	N	A0072-LEU	O	2.60	MM	59	4.47	138.5	1.79	138.2	145.4		27				
A0074-ILE	N	A0129-MET	O	2.82	MM	55	5.20	173.7	1.84	159.1	157.5		28				
A0129-MET	N	A0074-ILE	O	3.01	MM	55	5.20	143.5	2.16	142.7	135.1		29				
A0076-HIS	N	A0127-ARG	O	2.90	MM	51	5.00	140.0	2.08	145.4	135.5		30				
A0127-ARG	N	A0076-HIS	O	3.34	MM	51	5.00	155.1	2.42	158.2	153.9		31				
A0076-HIS	ND1	A0125-VAL	O	3.43	SM	49	7.14	159.4	2.49	148.6	142.8		32				
A0076-HIS	NE2	A0109-LEU	O	3.02	SM	33	6.40	123.8	2.36	137.9	141.3		33				
A0077-HIS	NE2	A0108-ASN	OD1	2.48	SS	31	5.66	126.1	1.77	158.1	156.2		34				
A0084-SER	N	A0081-CYS	O	3.02	MM	3	6.24	130.9	2.29	153.1	153.9		35				
A0086-ALA	N	A0083-HIS	O	2.70	MM	3	5.29	146.7	1.82	115.7	127.0		36				
A0087-LEU	N	A0084-SER	O	2.90	MM	3	5.39	105.2	2.48	117.3	133.0		37				
A0088-LYS	N	A0084-SER	O	2.93	MM	4	6.08	179.3	1.94	163.9	164.1		38				
A0089-LYS	NZ	A0085-GLN	O	2.65	SM	4	6.24	145.3	1.73	109.0	97.7		39				
A0089-LYS	N	A0086-ALA	O	2.86	MM	3	5.29	146.4	1.99	109.3	120.2		40				
A0091-PHE	N	A0087-LEU	O	2.81	MM	4	5.66	146.7	1.94	160.8	169.2		41				

Appendix 7 (continued overleaf). ***Measurement of hydrogen bonding in the AGR2 structure by HBPLUS software [733].*** The AGR2 $_{41\text{-}175}$ structure file (2LNS.pdb) was used for hydrogen bond measurements. chn: chain. ins: insert. red/res: residue. D: donor. A: acceptor. AA: acceptor antecedent (next residue in chain after acceptor). cat: category. M: main chain. S: side-chain. CA: alpha carbon. num aas: distance (in amino acids)between donor and acceptor groups (-1 indicates this does not apply). dist: distance. DHA angle: angle formed at the hydrogen of donor and acceptor. H-A-AA: smaller angle at the acceptor formed by the hydrogen and an acceptor antecedent. D-A-AA: smaller angle at the acceptor formed by the donor and an acceptor antecedent. H: hydrogen. num: number. All distances are in Angstroms (Å) and all angles in degrees. Inter-chain hydrogen bonds are highlighted in red.

A0092-ALA N	A0089-LYS O	3.18	MM	3	6.08	153.2	2.27	124.2	128.0	42
A0093-GLU N	A0090-VAL O	2.75	MM	3	5.20	129.7	2.02	112.9	127.6	43
A0094-ASN N	A0090-VAL O	2.69	MM	4	5.74	166.6	1.73	167.8	169.6	44
A0094-ASN ND2	A0090-VAL O	2.79	SM	4	5.74	161.2	1.85	110.1	105.5	45
A0095-LYS N	A0093-GLU O	2.55	MM	2	5.48	136.5	1.75	116.6	101.8	46
A0097-ILE N	A0094-ASN O	2.68	MM	3	5.57	126.0	1.99	133.7	145.6	47
A0157-THR OG1	A0094-ASN OD1	2.72	SS	63	7.48	144.3	1.87	150.4	144.2	48
A0099-LYS N	A0095-LYS O	2.92	MM	4	6.00	164.7	1.96	163.0	162.7	49
A0100-LEU N	A0096-GLU O	3.30	MM	4	6.56	141.0	2.47	144.7	154.3	50
A0100-LEU N	A0097-ILE O	2.85	MM	3	5.29	133.7	2.08	111.5	124.9	51
A0101-ALA N	A0097-ILE O	2.75	MM	4	5.92	148.8	1.86	146.0	155.3	52
A0101-ALA N	A0098-GLN O	2.43	MM	3	4.47	110.8	1.90	108.8	130.9	53
A0102-GLU N	A0098-GLN O	2.68	MM	4	5.57	158.0	1.74	162.4	158.9	54
A0103-GLN NE2	A0099-LYS O	2.61	SM	4	6.24	149.0	1.72	129.6	129.7	55
A0113-THR OG1	A0111-TYR O	2.77	SM	2	6.00	165.7	1.82	158.9	160.8	56
A0115-ASP N	A0113-THR O	2.66	MM	2	5.66	121.1	2.02	120.0	104.3	57
A0118-LEU N	A0115-ASP O	2.73	MM	3	5.48	161.1	1.78	132.9	138.5	58
A0127-ARG N	A0125-VAL O	2.43	MM	2	5.57	105.4	1.97	124.5	112.1	59
A0152-TYR OH	A0126-PRO O	2.70	SM	26	7.28	135.3	1.93	165.8	166.0	60
A0129-MET N	A0127-ARG O	2.99	MM	2	5.66	114.6	2.44	111.7	95.4	61
A0127-ARG NH2	A0152-TYR OH	2.31	SS	25	9.85	108.2	1.79	164.2	171.1	62
A0132-ASP N	A0136-THR O	2.84	MM	4	5.00	142.8	2.00	145.9	137.3	63
A0135-LEU N	A0132-ASP O	2.56	MM	3	5.29	142.6	1.72	140.2	144.0	64
A0136-THR OG1	A0132-ASP OD1	3.01	SS	4	5.00	148.4	2.14	109.2	99.6	65
A0138-ARG NH1	A0132-ASP OD2	3.28	SS	6	6.32	163.3	2.30	133.2	135.8	66
A0138-ARG N	A0136-THR O	2.62	MM	2	5.66	122.7	1.96	119.8	104.5	67
A0139-ALA N	A0137-VAL O	2.73	MM	2	5.48	140.9	1.89	109.0	95.9	68
A0141-ILE N	A0138-ARG O	3.02	MM	3	5.74	133.6	2.26	126.8	130.2	69
A0163-ASN ND2	A0141-ILE O	2.95	SM	22	7.75	152.6	2.04	150.2	142.8	70
A0166-LYS NZ	A0142-THR O	3.05	SM	24	8.83	147.0	2.12	178.0	168.3	71
A0148-ARG N	A0145-TYR O	2.74	MM	3	5.48	170.4	1.77	137.1	139.3	72
A0151-ALA N	A0148-ARG O	2.90	MM	3	5.83	150.3	2.01	139.9	145.2	73
A0148-ARG NE	A0156-ASP OD1	2.30	SS	8	7.68	127.4	1.55	147.2	164.9	74
A0153-GLU N	A0151-ALA O	2.38	MM	2	5.57	105.6	1.92	124.0	114.2	75
A0156-ASP N	A0153-GLU O	2.87	MM	3	4.69	172.6	1.89	114.0	114.9	76
A0159-LEU N	A0156-ASP O	2.66	MM	3	5.74	125.0	1.97	147.4	157.3	77
A0160-LEU N	A0156-ASP O	3.25	MM	4	6.32	160.4	2.31	139.6	142.8	78
A0161-LEU N	A0157-THR O	2.65	MM	4	5.48	161.7	1.70	146.7	152.2	79
A0162-ASP N	A0158-ALA O	2.94	MM	4	6.08	167.4	1.97	147.1	151.0	80
A0163-ASN N	A0159-LEU O	2.68	MM	4	5.66	170.5	1.71	161.1	164.5	81
A0164-MET N	A0160-LEU O	3.09	MM	4	6.24	169.4	2.12	150.6	150.8	82
A0165-LYS N	A0161-LEU O	2.67	MM	4	5.66	163.0	1.72	164.1	167.7	83
A0166-LYS N	A0162-ASP O	2.55	MM	4	5.57	122.8	1.88	139.8	157.9	84
A0166-LYS N	A0163-ASN O	2.66	MM	3	4.69	123.4	1.99	98.4	116.3	85
A0167-ALA N	A0163-ASN O	2.48	MM	4	5.74	136.1	1.68	152.4	167.8	86
A0167-ALA N	A0164-MET O	2.88	MM	3	4.69	121.2	2.25	94.2	110.6	87
A0168-LEU N	A0164-MET O	3.05	MM	4	6.24	163.6	2.10	165.0	166.0	88
A0169-LYS NZ	A0165-LYS O	2.24	SM	4	6.63	102.6	1.76	137.3	126.3	89
A0169-LYS N	A0166-LYS O	3.13	MM	3	5.00	149.6	2.24	102.9	106.9	90
B0040-THR OG1	B0039-PHE O	2.65	SM	1	3.74	141.9	1.83	105.6	93.9	91
B0040-THR OG1	B0040-THR O	2.60	SM	0	-1.00	122.1	1.96	97.0	90.8	92
B0046-ARG N	B0043-THR O	2.74	MM	3	4.69	138.9	1.93	105.6	117.9	93
B0047-GLY N	B0045-SER O	2.58	MM	2	5.48	139.1	1.76	115.4	101.1	94
B0046-ARG NH1	B0050-ASP OD1	2.42	SS	4	7.00	125.5	1.69	110.5	112.3	95
B0046-ARG NH2	B0050-ASP OD2	2.54	SS	4	7.00	148.4	1.63	110.7	101.3	96
B0050-ASP N	B0047-GLY O	2.70	MM	3	4.36	147.4	1.82	117.1	116.4	97
B0051-GLN NE2	B0050-ASP O	2.50	SM	1	3.74	132.6	1.73	95.3	91.2	98
B0051-GLN NE2	B0050-ASP OD2	2.57	SS	1	3.74	142.4	1.72	124.9	117.2	99
B0055-THR N	B0053-ILE O	3.07	MM	2	5.74	132.0	2.32	106.4	92.6	100
B0108-ASN ND2	B0055-THR O	3.30	SM	53	7.35	140.8	2.48	172.6	162.3	101
B0057-THR N	B0056-GLN OE1	2.66	MS	1	3.74	152.7	1.76	117.8	108.5	102
B0056-GLN NE2	B0108-ASN O	2.66	SM	52	6.00	150.9	1.76	160.0	157.9	103
B0061-ALA N	B0057-THR O	3.36	MM	4	6.56	166.0	2.40	149.1	153.2	104

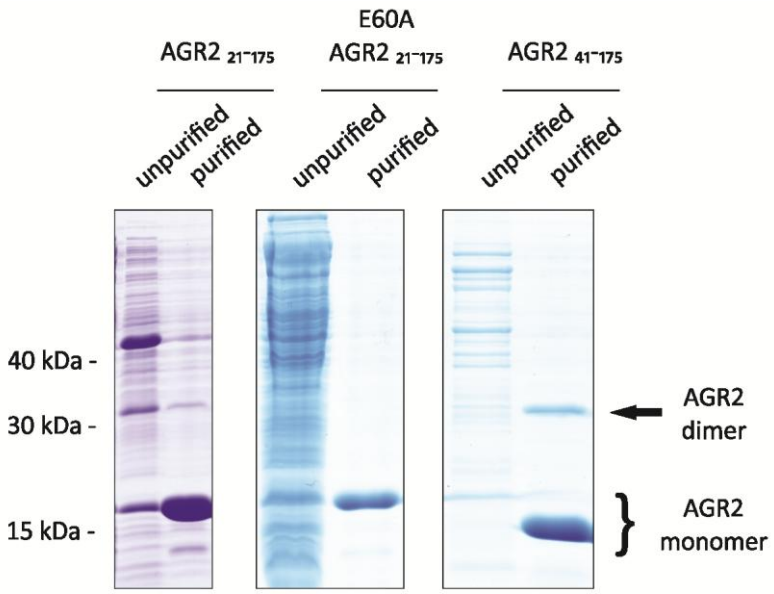
Appendix 7. ***Measurement of hydrogen bonding in the AGR2 structure by HBPLUS software (cont.).***

B0061-ALA N	B0058-TYR O	2.53	MM	3	5.00	110.0	2.02	123.4	142.6	105
B0062-LEU N	B0058-TYR O	2.96	MM	4	6.24	154.1	2.04	142.8	150.1	106
B0062-LEU N	B0059-GLU O	2.54	MM	3	4.80	110.4	2.03	106.9	127.3	107
B0063-TYR N	B0059-GLU O	2.77	MM	4	6.08	157.8	1.83	158.3	165.5	108
B0064-LYS N	B0060-GLU O	2.88	MM	4	6.16	174.3	1.89	172.7	170.8	109
B0065-SER N	B0061-ALA O	3.29	MM	4	6.40	158.8	2.36	144.1	149.8	110
B0066-LYS N	B0062-LEU O	2.70	MM	4	5.83	155.6	1.78	147.0	155.4	111
B0067-THR N	B0063-TYR O	2.85	MM	4	6.08	166.4	1.88	175.0	172.3	112
B0068-SER N	B0065-SER O	2.67	MM	3	5.20	145.4	1.81	115.8	127.8	113
B0069-ASN N	B0065-SER O	3.00	MM	4	6.24	116.4	2.44	138.6	153.0	114
B0065-SER OG	B0070-LYS O	3.00	SM	5	6.16	143.8	2.17	161.9	167.1	115
B0069-ASN N	B0066-LYS O	2.65	MM	3	4.69	146.7	1.78	108.2	119.9	116
B0069-ASN ND2	B0069-ASN O	2.62	SM	0	-1.00	135.8	1.83	122.1	107.8	117
B0072-LEU N	B0131-VAL O	2.81	MM	59	4.58	168.2	1.84	141.7	144.3	118
B0131-VAL N	B0072-LEU O	2.58	MM	59	4.58	138.5	1.77	141.0	147.8	119
B0074-ILE N	B0129-MET O	2.82	MM	55	5.29	169.5	1.84	164.0	160.8	120
B0129-MET N	B0074-ILE O	3.05	MM	55	5.29	137.4	2.25	144.2	135.8	121
B0076-HIS N	B0127-ARG O	2.94	MM	51	5.00	139.2	2.12	143.8	133.4	122
B0127-ARG N	B0076-HIS O	3.34	MM	51	5.00	150.2	2.46	154.2	152.7	123
B0077-HIS NE2	B0108-ASN OD1	2.47	SS	31	5.66	124.3	1.78	154.9	154.7	124
B0084-SER OG	B0078-LEU O	2.83	SM	6	6.48	104.0	2.44	145.8	163.9	125
B0084-SER OG	B0079-ASP OD1	3.19	SS	5	5.74	168.8	2.23	163.8	160.5	126
B0084-SER N	B0081-CYS O	2.85	MM	3	6.08	135.1	2.07	156.1	156.6	127
B0086-ALA N	B0083-HIS O	2.71	MM	3	5.20	143.7	1.86	112.2	124.5	128
B0087-LEU N	B0084-SER O	2.92	MM	3	5.48	106.0	2.49	120.8	135.9	129
B0088-LYS N	B0084-SER O	2.97	MM	4	6.16	178.3	1.99	162.7	162.5	130
B0089-LYS NZ	B0085-GLN O	2.60	SM	4	6.24	144.3	1.68	109.6	97.9	131
B0089-LYS N	B0086-ALA O	2.88	MM	3	5.29	147.4	2.00	109.5	120.0	132
B0091-PHE N	B0087-LEU O	2.83	MM	4	5.66	147.4	1.96	161.5	169.3	133
B0092-ALA N	B0089-LYS O	3.16	MM	3	6.08	153.3	2.25	125.3	129.0	134
B0093-GLU N	B0090-VAL O	2.76	MM	3	5.20	129.5	2.03	113.0	127.7	135
B0094-ASN N	B0090-VAL O	2.69	MM	4	5.74	166.7	1.73	167.8	169.5	136
B0094-ASN ND2	B0090-VAL O	2.79	SM	4	5.74	161.1	1.85	110.2	105.5	137
B0095-LYS N	B0093-GLU O	2.55	MM	2	5.48	136.5	1.75	116.7	101.9	138
B0097-ILE N	B0094-ASN O	2.68	MM	3	5.57	126.1	1.99	133.7	145.6	139
B0157-THR OG1	B0094-ASN OD1	2.72	SS	63	7.48	143.6	1.88	150.6	144.2	140
B0099-LYS N	B0095-LYS O	2.91	MM	4	6.00	165.0	1.96	163.1	162.8	141
B0100-LEU N	B0096-GLU O	3.30	MM	4	6.56	140.7	2.48	144.5	154.3	142
B0100-LEU N	B0097-ILE O	2.85	MM	3	5.29	133.8	2.08	111.4	124.8	143
B0101-ALA N	B0097-ILE O	2.75	MM	4	5.92	148.7	1.87	146.0	155.3	144
B0101-ALA N	B0098-GLN O	2.44	MM	3	4.47	111.0	1.91	108.8	130.9	145
B0102-GLU N	B0098-GLN O	2.68	MM	4	5.57	157.8	1.74	162.6	159.0	146
B0103-GLN NE2	B0099-LYS O	2.60	SM	4	6.24	147.7	1.72	130.5	130.2	147
B0113-THR OG1	B0111-TYR O	2.78	SM	2	6.00	165.8	1.83	159.4	160.4	148
B0115-ASP N	B0113-THR O	2.69	MM	2	5.66	122.8	2.03	118.7	103.0	149
B0118-LEU N	B0115-ASP O	2.73	MM	3	5.48	161.9	1.78	133.4	138.9	150
B0116-LYS N	B0115-ASP OD1	2.76	MS	1	3.74	108.6	2.29	115.2	110.6	151
B0127-ARG N	B0125-VAL O	2.52	MM	2	5.66	98.4	2.18	123.3	112.3	152
B0129-MET N	B0127-ARG O	2.87	MM	2	5.57	118.3	2.27	113.0	96.5	153
B0127-ARG NH2	B0152-TYR OH	2.31	SS	25	9.85	91.9	2.05	127.7	140.9	154
B0132-ASP N	B0136-THR O	2.84	MM	4	5.10	136.1	2.05	153.2	140.1	155
B0135-LEU N	B0132-ASP O	2.55	MM	3	5.39	134.0	1.77	140.1	148.2	156
B0138-ARG NH1	B0132-ASP OD2	3.29	SS	6	6.16	148.5	2.39	101.9	110.6	157
B0138-ARG N	B0136-THR O	2.64	MM	2	5.66	117.9	2.03	120.5	106.0	158
B0139-ALA N	B0137-VAL O	2.80	MM	2	5.48	143.9	1.94	104.5	93.1	159
B0141-ILE N	B0138-ARG O	2.97	MM	3	5.74	135.6	2.19	128.7	132.9	160
B0163-ASN ND2	B0141-ILE O	3.00	SM	22	7.81	154.3	2.08	145.9	139.2	161
B0166-LYS NZ	B0142-THR O	3.02	SM	24	8.77	145.5	2.10	177.8	169.3	162
B0142-THR OG1	B0146-SER OG	2.64	SS	4	5.57	150.7	1.76	95.1	90.5	163
B0148-ARG N	B0145-TYR O	2.71	MM	3	5.39	166.8	1.75	132.1	136.1	164
B0151-ALA N	B0148-ARG O	3.06	MM	3	5.92	141.7	2.23	133.3	140.0	165
B0148-ARG NE	B0156-ASP OD1	2.35	SS	8	7.68	133.0	1.55	150.6	166.5	166
B0153-GLU N	B0151-ALA O	2.39	MM	2	5.57	110.1	1.87	124.1	112.9	167

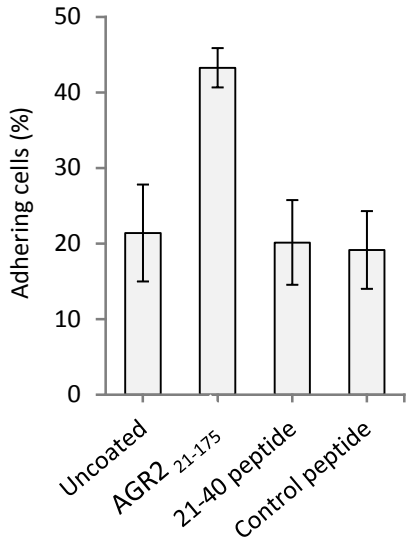
Appendix 7. ***Measurement of hydrogen bonding in the AGR2 structure by HBPLUS software (cont.).***

B0156-ASP N	B0153-GLU O	2.90	MM	3	4.80	172.9	1.93	113.5	114.6	168
B0159-LEU N	B0156-ASP O	2.67	MM	3	5.74	123.7	2.00	146.7	157.1	169
B0160-LEU N	B0156-ASP O	3.24	MM	4	6.24	161.0	2.29	140.2	143.2	170
B0161-LEU N	B0157-THR O	2.64	MM	4	5.48	161.4	1.70	146.1	151.7	171
B0162-ASP N	B0158-ALA O	2.94	MM	4	6.08	168.4	1.97	148.7	152.1	172
B0163-ASN N	B0159-LEU O	2.68	MM	4	5.66	169.6	1.71	160.1	163.8	173
B0164-MET N	B0160-LEU O	3.09	MM	4	6.24	169.7	2.12	150.9	151.1	174
B0165-LYS N	B0161-LEU O	2.67	MM	4	5.66	163.4	1.72	164.2	167.7	175
B0166-LYS N	B0162-ASP O	2.54	MM	4	5.57	122.7	1.87	139.9	158.0	176
B0166-LYS N	B0163-ASN O	2.66	MM	3	4.69	123.6	1.99	98.4	116.3	177
B0167-ALA N	B0163-ASN O	2.49	MM	4	5.74	137.2	1.68	152.6	167.6	178
B0167-ALA N	B0164-MET O	2.89	MM	3	4.80	120.4	2.27	93.9	110.4	179
B0168-LEU N	B0164-MET O	3.05	MM	4	6.24	163.1	2.10	164.3	165.7	180
B0169-LYS NZ	B0165-LYS O	2.24	SM	4	6.63	102.7	1.76	137.3	126.3	181
B0169-LYS N	B0166-LYS O	3.10	MM	3	5.00	149.1	2.22	104.3	107.9	182

Appendix 7. ***Measurement of hydrogen bonding in the AGR2 structure by HBPLUS software (cont.).***

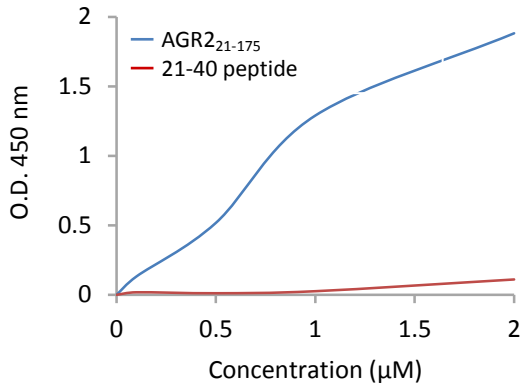


Appendix 8. ***SDS-resistant dimers in recombinant AGR2 purification.*** Proteins were purified as described in Materials and Methods and the unpurified and purified fractions runs on SDS-PAGE gels. $AGR2_{21\text{-}175}$ has an expected molecular weight of 17.8 kDA for a monomer and 35.6 kDa for a dimer. $AGR2_{41\text{-}175}$ has an expected molecular weight of 16.1 kDa for a monomer and 32.2 kDa for a dimer. The band seen at ~45 kDa in $AGR2_{21\text{-}175}$ is spill over from the unpurified lane, whilst the band at ~10 kDa is a low molecular weight contaminant. Notice the slight increase in dimer band size in $AGR2_{41\text{-}175}$ compared to $AGR2_{21\text{-}175}$ and the lack of dimer band in E60A $AGR2_{21\text{-}175}$.

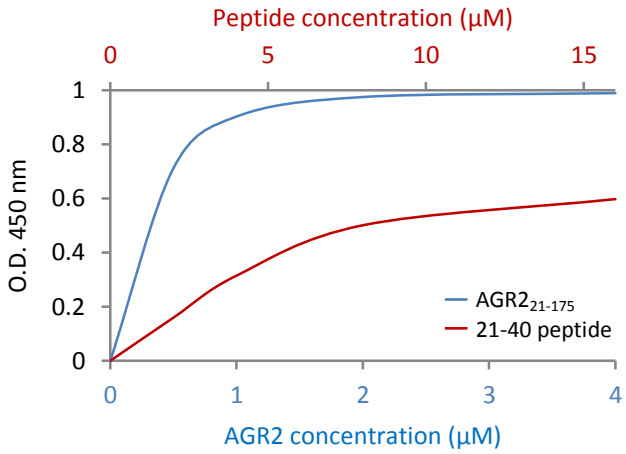


Appendix 9. ***Peptide coating does not affect rate of cell adhesion.*** Wells were coated overnight with buffer containing equimolar concentrations of either $AGR2_{21\text{-}175}$, a peptide corresponding to residues 21-40 of AGR2 (sequence RDTTVKPGAKKDTKDSRPKL) or an unrelated peptide with a similar number of positively charged residues (sequence TADAMNREVSSLKNKLRRGD). Rama 37 cells were then allowed to adhere to the coated wells for 60 min and the number of bound cells was counted using an automated cell counter. The number of adhering cells is expressed as a percentage of the number added to each well. The means from duplicate wells from 3 independent experiments are shown, ±SE.

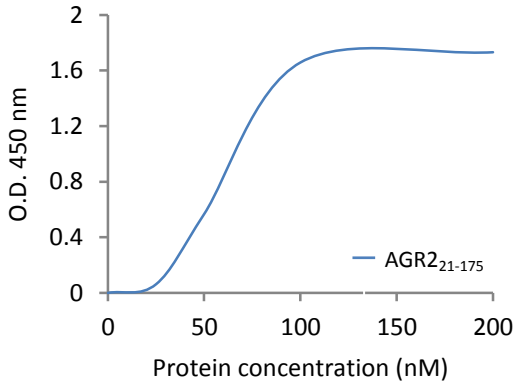
A



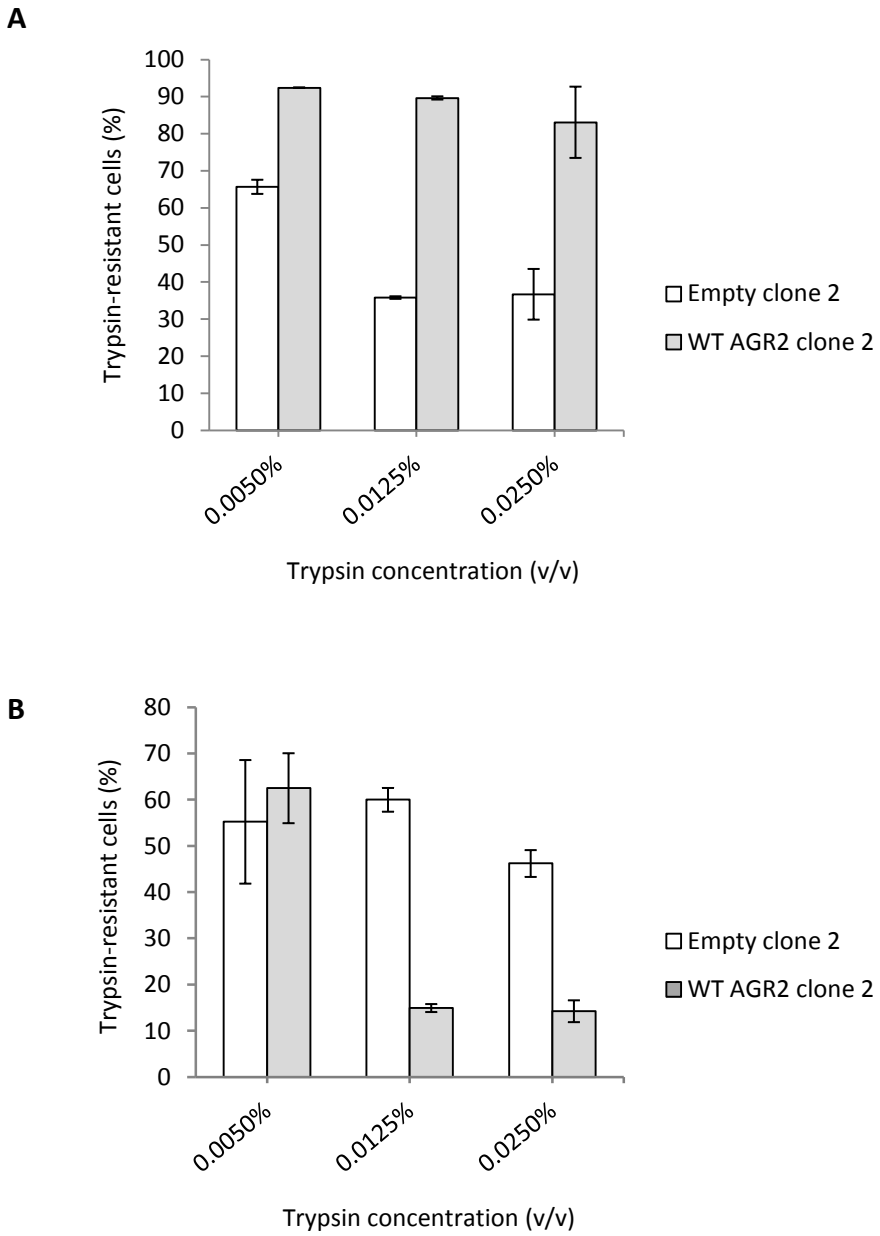
B



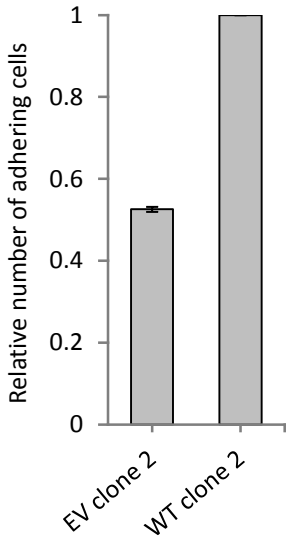
C



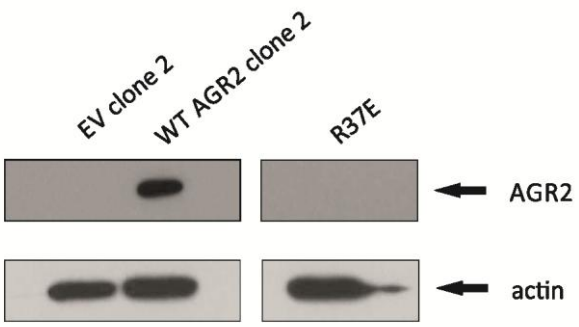
Appendix 10. ***Calibration of AGR2 21-40 peptide coating.*** Wells were coated overnight with the indicated concentrations of AGR2$_{21-175}$ or 21-40 peptide and then subjected to an ELISA assay to measure the relative amounts of protein/peptide bound. **(A)** ELISA of coating in PBS buffer (pH 7.4). **(B)** ELISA of coating in carbonate buffer (pH 10.6). The bottom x-axis (blue) indicates the amount of AGR2 $_{21-175}$ coated and the top x-axis (red) indicates the amount of peptide coated. **(C)** Representative calibration curve of AGR2$_{21-175}$ binding used to determine the amount of peptide required to achieve equal binding.



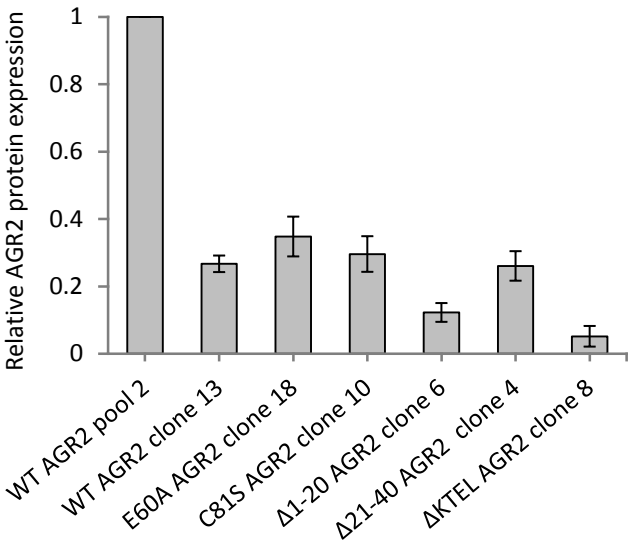
Appendix 11. ***Optimisation of trypsin detachment assay*****. (A)** Wild type (WT) AGR2-expressing and empty vector (EV) control cells were allowed to adhere for 30 min to cell culture plates. Attached cells were then treated with the indicated concentrations of trypsin for 5 min and the remaining attached cells were counted with an automated cell counter. The percentage of trypsin-resistant cells represents the proportion of applied cells attaching in 30 min that remained adherent after treatment with the indicated trypsin concentrations. The means from duplicate wells from a single experiment are shown, ±SD. **(B)** Cells were grown overnight to 70-80 % confluence and then trypsinised in the indicated trypsin solutions for 5 min. The remaining adhering cells were counted with an automated cell counter. The number of trypsin-resistant cells is expressed as a percentage of the total number of cells in each well. The means from duplicate wells from a single experiment are shown, ±SD.



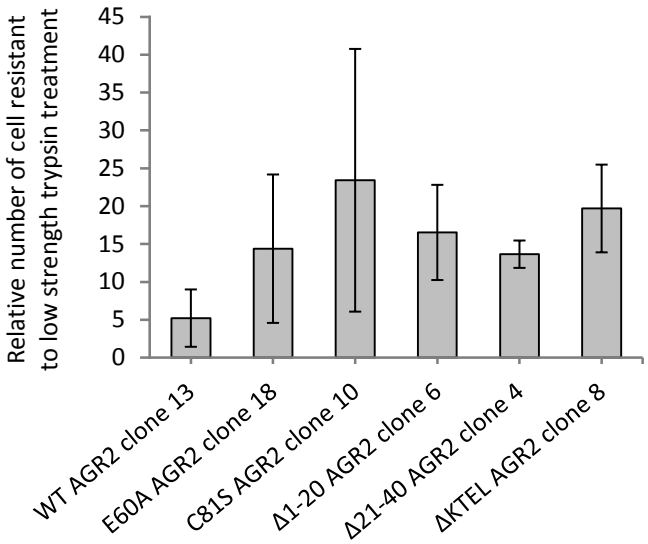
Appendix 12. ***Clonal AGR2-expressing cell attach more strongly than control cells.*** Cells in suspension were added to wells and briefly centrifuged at 10 x *g* to ensure all cells were in contact with the plate surface. After a subsequent 10 min incubation, plates were sealed and centrifuged upside down for 10 min at 300 x *g*. Remaining adhering cells were quantified using crystal violet staining as described in Materials and Methods. The mean staining relative to WT AGR2 clone 2 from 3 independent experiments carried out using duplicate wells is shown, ±SE.



Appendix 13*. R37E cells are elongated but do not express AGR2.* Cells were grown to 70-80 % confluence, harvested and probed for AGR2 by Western blot. The R37E lane is from a different part of the same blot.



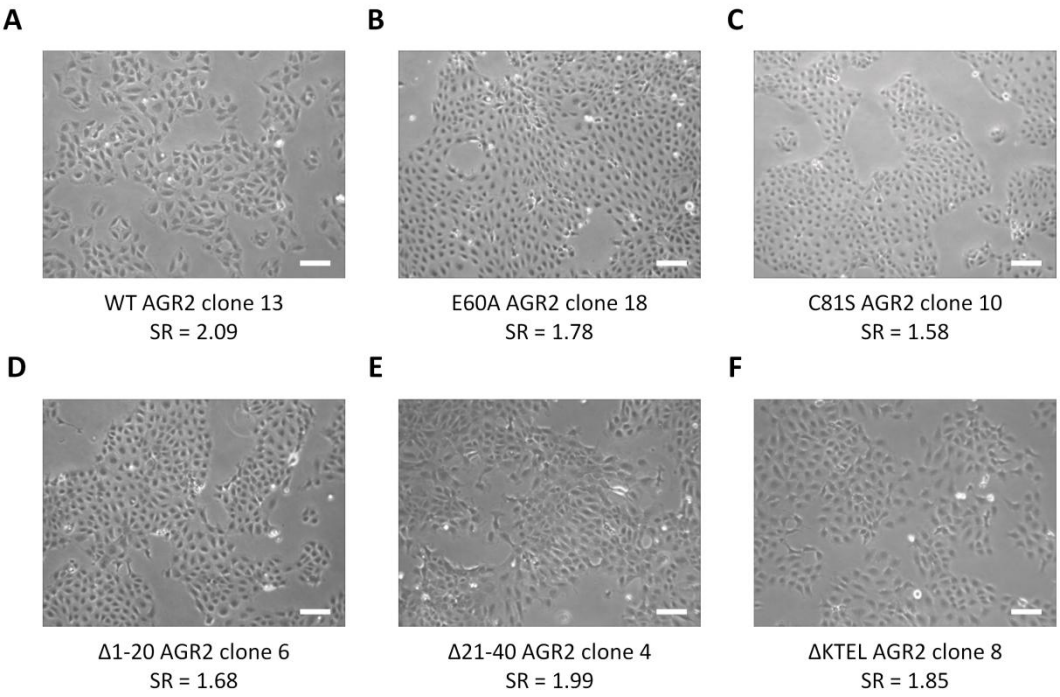
Appendix 14. ***Quantification of steady-state protein levels of mutant AGR2 cell lines.*** Mean steady-state protein levels from three sets of whole cell lysates were measured by Western blot, and AGR2 abundance was measured relative to the levels in WT AGR2 pool 2 cells. Data are shown ±SE.



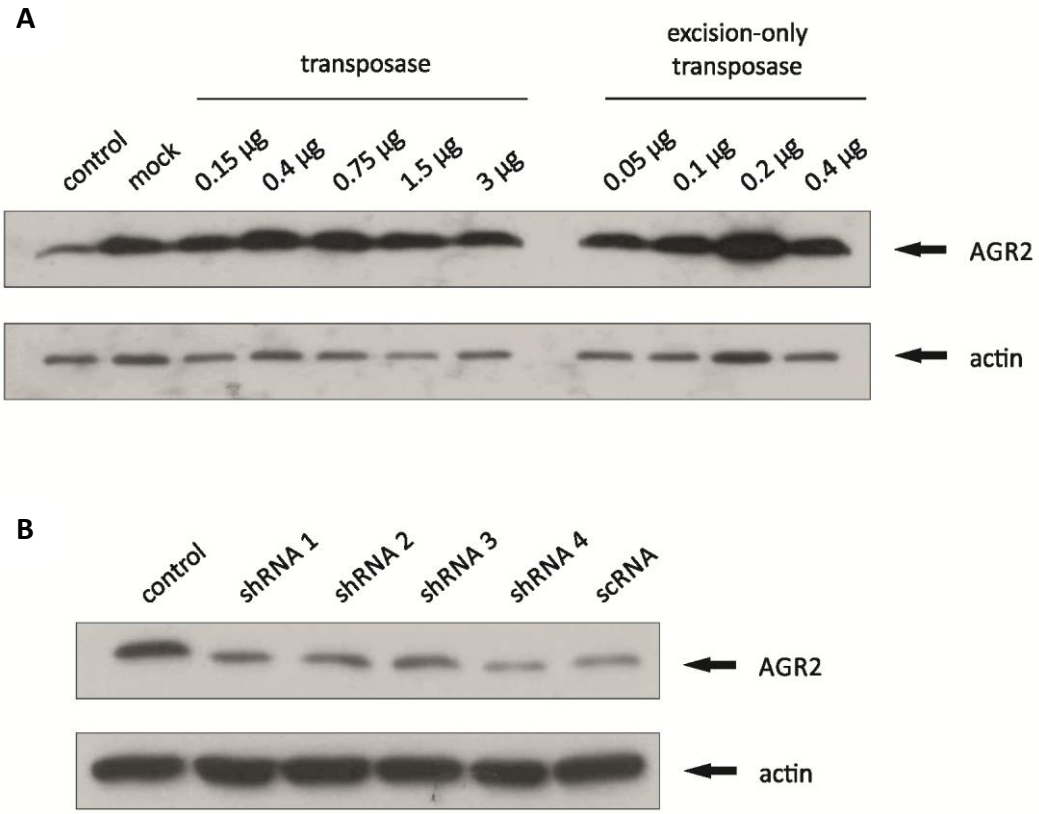
Appendix 15. ***Monolayer attachment properties of mutant-AGR2- expressing clonal cell lines.*** The indicated cells were grown overnight to 70-80 % confluence and then trypsinised in a weak trypsin solution for 5 min. The remaining adhering cells were counted with an automated cell counter. The number of trypsin-resistant cells is expressed as a percentage of the total number of cells in each well. The proportion of trypsin-resistant cells is expressed relative to the number of trypsin-resistant WT clone 2 cells in individual experiments. The means from duplicate wells from 3 independent experiments are shown, ±SE.

Cell	SR
EV clone 1	1.82
EV clone 2	1.78
EV clone 3	1.94
WT AGR2 clone 2	2.48
WT AGR2 clone 5	2.53
WT AGR2 clone 13	2.09
E60A AGR2 clone 1	1.78
E60A AGR2 clone 18	1.78
E60A AGR2 clone 19	1.81
C81S AGR2 clone 7	1.74
C81S AGR2 clone 10	1.58
C81S AGR2 clone 15	1.70
Δ1-20 AGR2 clone 6	1.68
Δ1-20 AGR2 clone 8	1.93
Δ1-20 AGR2 clone 11	1.84
Δ21-40 AGR2 clone 3	2.13
Δ21-40 AGR2 clone 4	1.99
Δ21-40 AGR2 clone 8	1.72
ΔKTEL AGR2 clone 5	2.06
ΔKTEL AGR2 clone 8	1.85
ΔKTEL AGR2 clone 20	1.70

Appendix 16. ***Mean shape ratio (SR) values for clonal cell lines.*** The SR values for at least 150 cells for each cell line were measured as described in Materials and Methods, and the mean values are presented here.

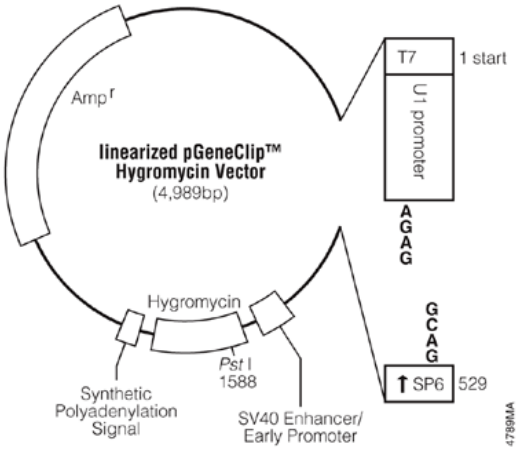


Appendix 17. ***Morphology and SR values of comparable AGR2-expressing cell lines.*** Representative images and mean SR values for AGR2-expressing clonal cell lines expressing similar steady-state proteins levels of AGR2. The SR values for at least 200 cells for each cell line were measured as described in Materials and Methods, and the mean values are presented here. **(A)** WT AGR2 clone 13, **(B)** E60A AGR2 clone 18, **(C)** C81S AGR2 clone 10, **(D)** Δ1-20 AGR2 clone 6, **(E)** Δ21-40 AGR2 clone 4 and **(F)** ΔKTEL AGR2 clone 8. SR: shape ratio.



Appendix 18. ***Knockdown of AGR2 in WT AGR2 clone 2 cells.*** **(A)** WT AGR2 clone 2 cells were grown for 2 weeks in selection-free medium and then transfected with the indicated amount of PiggyBac transposase vector or with PiggyBac excision-only transposase vector, as described by the manufacturer (Systems Biosciences). Cells were passaged 3 times, harvested and AGR2 protein was measured by Western blot. **(B)** WT AGR2 clone 2 cells were transfected with 2 µg of the indicated short hairpin RNA (shRNA) or scrambled RNA (scRNA, see Appendix 19 for more details). After selection in hygromycin medium for approximately 14 days, cells were harvested and AGR2 protein was measured by Western blot. Although there appears to be some AGR2 knockdown with shRNA 4, the equivalent knockdown in scRNA-transfected cells indicates that this potential effect is not AGR2 shRNA-specific.

A



B

shRNA1: GAGCCGATATCACTGGAAGAT

shRNA2: GCAGTTTGTCCTCCTCAATCT

shRNA3: AAGCAACAAACCCTTGATGAT

shRNA4: AGACATATGAAGAAGCTCTAT

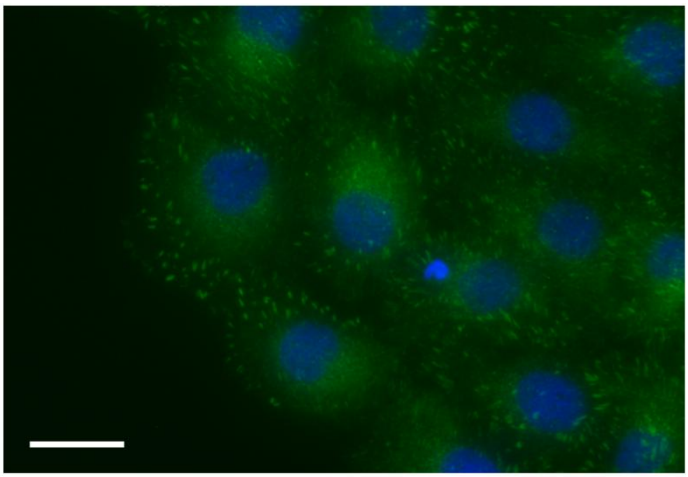
scRNA : GGAATCTCATTCGATGCATAC

C

ATGGAGAAAATTCCAGTGTCAGCATTCTTGCTCCTTGTGGCCCTCTCCTACACTCTGGCCAGAGATACCACAGT
CAAACCTGGAGCCAAAAAGGACACAAAGGACTCTCGACCCAAACTGCCCCAGACCCTCTCCAGAGGTTGGGG
TGACCAACTCATCTGGACTCAGACATATGAAGAAGCTCTATATAAATCCAAGACAAGCAACAAACCCTTGATG
ATTATTCATCACTTGGATGAGTGCCCACACAGTCAAGCTTTAAAGAAAGTGTTTGCTGAAAATAAAGAAATCCA
GAAATTGGCAGAGCAGTTTGTCCTCCTCAATCTGGTTTATGAAACAACTGACAAACACCTTTCTCCTGATGGCC
AGTATGTCCCCAGGATTATGTTTGTTGACCCATCTCTGACAGTTAGAGCCGATATCACTGGAAGATATTCAAAT
CGTCTCTATGCTTACGAACCTGCAGATACAGCTCTGTTGCTTGACAACATGAAGAAAGCTCTCAAGTTGCTGAA
GACTGAATTGTAA

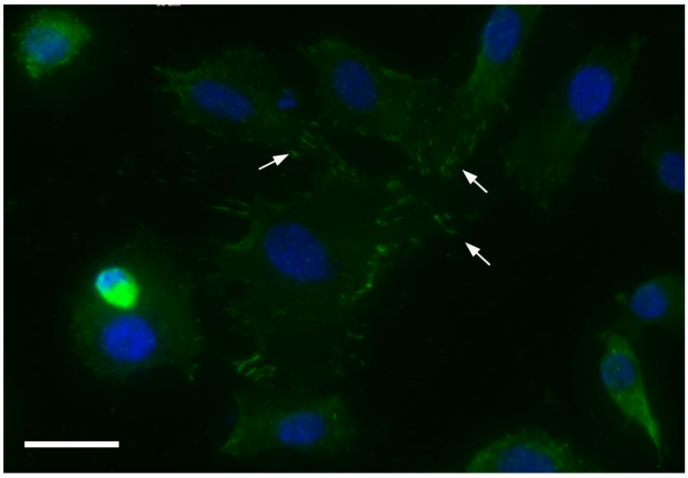
Appendix 19. ***shRNA constructs used for AGR2 knockdown.*** **(A)** Vector map of shRNA plasmid. **(B)** shRNA sequences generated by Qiagen for AGR2 knockdown. **(C)** Position of targeted sequences of shRNAs (underlined) within the AGR2 coding sequence.

A



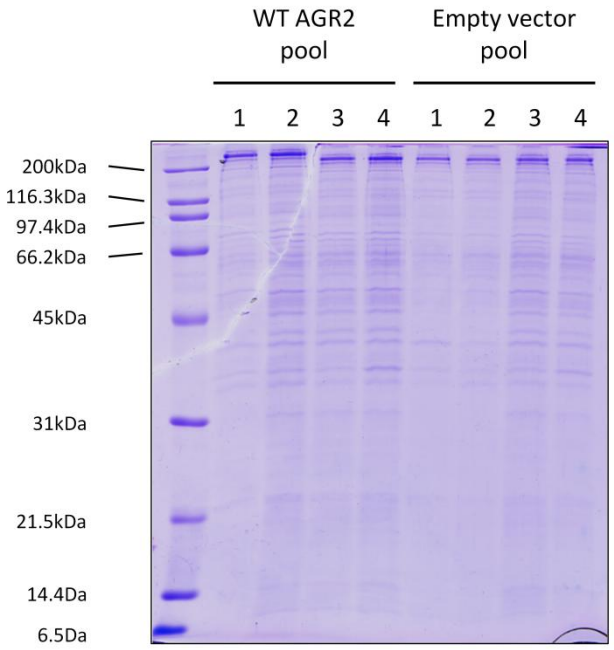
EV clone 2

B



WT AGR2 clone 2

Appendix 20. ***WT AGR2 clonal cells display fewer focal adhesions than EV clonal cells.*** Cells were seeded onto fibronectin-coated CultureSlides and grown overnight. Cells were then fixed and permeabilised, and focal adhesions were visualised through paxillin staining. There was a much greater concentration of punctate paxillin staining in EV clone cells **(A)**, whereas in WT AGR2 clone 2 cells **(B)**, staining was more limited to the edges of the cells (white arrows). The focal adhesions in WT AGR2 clone 2 cells were also more elongated than those in EV clone 2 cells, indicative of higher traction forces travelling through these focal adhesions [108, 110] and thus suggesting that these cells are more motile than EV clone 2 cells. Images kindly provided by Richard Smith, Institute of Integrative Biology, University of Liverpool. Scalebar: 25 µm.



Appendix 21. ***Normalisation of conditioned medium to cell number.*** CM was collected from 4 replicate plates of WT AGR2 pool and EV pool cells in serum-free medium. After CM collection, the number of cells in these plates was counted using an automated cell counter. Protein was captured onto Strataclean beads (see Materials and Methods) from 1 mL of CM from the replicate with the lowest cell count (out of the 8 total replicates). For other replicates, protein was captured from proportionally less than 1 mL CM, based on cell count, e.g. for a sample with a cell count 1.25 times greater than the lowest cell count, protein was captured from 1.25 times less CM (i.e. 800 µL in this case). The captured proteins were then run on an SDS-PAGE gel and stained with Coomassie brilliant blue. Note that even between replicates from the same cells, the levels of secreted proteins were not equal as denoted by the different protein staining intensities in the relevant lanes. This suggests that normalising CM samples to cell number is poorly reproducible. Numbers 1-4 denote the biological replicates. Gel image provided by Dr Deborah Simpson, Centre for Proteome Research, Institute of Integrative Biology, University of Liverpool.

Phosphorylation sites predicted: Ser: 5 Thr: 5 Tyr: 2

Serine predictions

Name	Pos	Context	Score	Pred
Sequence	7	KIPVSAFLL	0.016	.
Sequence	16	LVALSYTLA	0.120	.
Sequence	36	DTKDSRPKL	0.870	*S*
Sequence	45	PQTLSRGWG	0.190	.
Sequence	65	ALYKSKTSN	0.748	*S*
Sequence	68	KSKTSNKPL	0.971	*S*
Sequence	84	ECPHSQALK	0.194	.
Sequence	119	DKHLSPDGQ	0.990	*S*
Sequence	134	FVDPSLTVR	0.019	.
Sequence	146	TGRYSNRLY	0.913	*S*

Threonine predictions

Name	Pos	Context	Score	Pred
Sequence	18	ALSYTLARD	0.012	.
Sequence	23	LARDTTVKP	0.963	*T*
Sequence	24	ARDTTVKPG	0.976	*T*
Sequence	33	AKKDTKDSR	0.366	.
Sequence	43	KLPQTLSRG	0.179	.
Sequence	55	QLIWTQTYE	0.013	.
Sequence	57	IWTQTYEEA	0.961	*T*
Sequence	67	YKSKTSNKP	0.211	.
Sequence	113	LVYETTDKH	0.088	.
Sequence	114	VYETTDKHL	0.842	*T*
Sequence	136	DPSLTVRAD	0.919	*T*
Sequence	142	RADITGRYS	0.428	.
Sequence	157	EPADTALLL	0.248	.
Sequence	173	KLLKTEL--	0.009	.

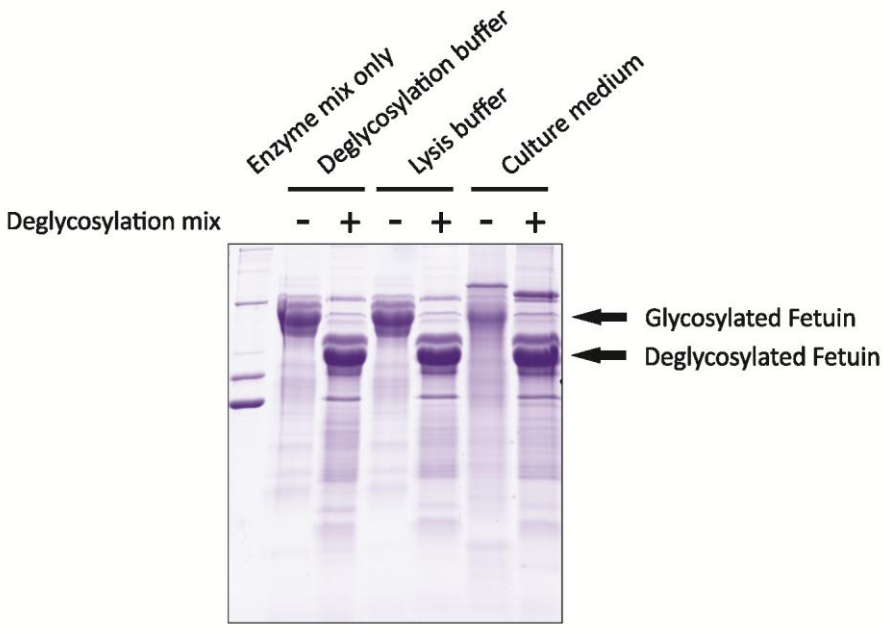
Tyrosine predictions

Name	Pos	Context	Score	Pred
Sequence	17	VALSYTLAR	0.040	.
Sequence	58	WTQTYEEAL	0.395	.
Sequence	63	EEALYKSKT	0.970	*Y*
Sequence	111	LNLVYETTD	0.277	.
Sequence	124	PDGQYVPRI	0.405	.
Sequence	145	ITGRYSNRL	0.099	.
Sequence	150	SNRLYAYEP	0.738	*Y*
Sequence	152	RLYAYEPAD	0.054	.

Appendix 22. ***Prediction of AGR2 phosphorylation sites by the NetPhos2.0 server*** [734]. The full length AGR2 sequence was submitted to http://www.cbs.dtu.dk/services/NetPhos/ for phosphorylation prediction. Residues flanked with asterisks (*x*) are predicted to be phosphorylated.

Residue Number	*Residue*	*Score*	*O-glycosylation Prediction*
7	Ser	0.078	-
16	Ser	0.173	-
18	Thr	0.327	-
23	Thr	0.639	Positive
24	Thr	0.775	Positive
33	Thr	0.863	Positive
36	Ser	0.943	Positive
43	Thr	0.860	Positive
45	Ser	0.782	Positive
55	Thr	0.281	-
57	Thr	0.163	-
65	Ser	0.142	-
67	Thr	0.046	-
68	Ser	0.051	-
84	Ser	0.084	-
113	Thr	0.080	-
114	Thr	0.045	-
119	Ser	0.211	-
134	Ser	0.136	-
136	Thr	0.134	-
142	Thr	0.168	-
146	Ser	0.347	-
157	Thr	0.110	-
173	Thr	0.010	-

Appendix 23. ***Prediction of AGR2 O-glycosylation sites by the NetOGlyc 4.0 server*** [735]. The full length AGR2 sequence was submitted to http://www.cbs.dtu.dk/services/NetOGlyc/ for O-glycosylation prediction.



Appendix 24. ***Control reactions for enzyme deglycosylation kit***. To ensure that the deglycosylation enzymes were not inhibited by elements of cell lysis buffer or concentrated conditioned medium, the extent of fetuin deglycosylationin the supplied buffer was compared to that in lysis buffer and concentrated culture medium, using the procedure described in Materials and Methods. Fetuin is both N- and O-glycosylated, and was equally well deglycosylated under all three reaction conditions.

References

1. Hanahan, D. and Weinberg, R.A. (2011) *Hallmarks of cancer: the next generation.* Cell. 144 (5), p. 646-74.
2. Jemal, A., Bray, F., Center, M.M., Ferlay, J., Ward, E., and Forman, D. (2011) *Global cancer statistics.* CA Cancer J Clin. 61 (2), p. 69-90.
3. Sakamoto, S. and Kyprianou, N. (2010) *Targeting anoikis resistance in prostate cancer metastasis.* Mol Aspects Med. 31 (2), p. 205-14.
4. Fidler, I.J. (2002) *Critical determinants of metastasis.* Semin Cancer Biol. 12 (2), p. 89-96.
5. Steeg, P.S. (2006) *Tumor metastasis: mechanistic insights and clinical challenges.* Nat Med. 12 (8), p. 895-904.
6. Knights, A.J., Funnell, A.P., Crossley, M., and Pearson, R.C. (2012) *Holding Tight: Cell Junctions and Cancer Spread.* Trends Cancer Res. 8, p. 61-69.
7. Dusek, R.L. and Attardi, L.D. (2011) *Desmosomes: new perpetrators in tumour suppression.* Nat Rev Cancer. 11 (5), p. 317-23.
8. Gooding, J.M., Yap, K.L., and Ikura, M. (2004) *The cadherin-catenin complex as a focal point of cell adhesion and signalling: new insights from three-dimensional structures.* Bioessays. 26 (5), p. 497-511.
9. Menke, A. and Giehl, K. (2012) *Regulation of adherens junctions by Rho GTPases and p120-catenin.* Arch Biochem Biophys. 524 (1), p. 48-55.
10. Pokutta, S., Herrenknecht, K., Kemler, R., and Engel, J. (1994) *Conformational changes of the recombinant extracellular domain of E-cadherin upon calcium binding.* Eur J Biochem. 223 (3), p. 1019-26.
11. Davis, M.A., Ireton, R.C., and Reynolds, A.B. (2003) *A core function for p120-catenin in cadherin turnover.* J Cell Biol. 163 (3), p. 525-34.
12. Maeda, M., Johnson, E., Mandal, S.H., Lawson, K.R., Keim, S.A., Svoboda, R.A., Caplan, S., Wahl, J.K., 3rd, Wheelock, M.J., and Johnson, K.R. (2006) *Expression of inappropriate cadherins by epithelial tumor cells promotes endocytosis and degradation of E-cadherin via competition for p120(ctn).* Oncogene. 25 (33), p. 4595-604.

13. Xiao, K., Allison, D.F., Buckley, K.M., Kottke, M.D., Vincent, P.A., Faundez, V., and Kowalczyk, A.P. (2003) *Cellular levels of p120 catenin function as a set point for cadherin expression levels in microvascular endothelial cells.* J Cell Biol. 163 (3), p. 535-45.

14. Ferber, A., Yaen, C., Sarmiento, E., and Martinez, J. (2002) *An octapeptide in the juxtamembrane domain of VE-cadherin is important for p120ctn binding and cell proliferation.* Exp Cell Res. 274 (1), p. 35-44.

15. Thoreson, M.A., Anastasiadis, P.Z., Daniel, J.M., Ireton, R.C., Wheelock, M.J., Johnson, K.R., Hummingbird, D.K., and Reynolds, A.B. (2000) *Selective uncoupling of p120(ctn) from E-cadherin disrupts strong adhesion.* J Cell Biol. 148 (1), p. 189-202.

16. Noren, N.K., Liu, B.P., Burridge, K., and Kreft, B. (2000) *p120 catenin regulates the actin cytoskeleton via Rho family GTPases.* J Cell Biol. 150 (3), p. 567-80.

17. Lilien, J., Balsamo, J., Arregui, C., and Xu, G. (2002) *Turn-off, drop-out: functional state switching of cadherins.* Dev Dyn. 224 (1), p. 18-29.

18. Perez-Moreno, M. and Fuchs, E. (2006) *Catenins: keeping cells from getting their signals crossed.* Dev Cell. 11 (5), p. 601-12.

19. Nelson, W.J. and Nusse, R. (2004) *Convergence of Wnt, beta-catenin, and cadherin pathways.* Science. 303 (5663), p. 1483-7.

20. Aberle, H., Butz, S., Stappert, J., Weissig, H., Kemler, R., and Hoschuetzky, H. (1994) *Assembly of the cadherin-catenin complex in vitro with recombinant proteins.* J Cell Sci. 107 (Pt 12), p. 3655-63.

21. Rimm, D.L., Koslov, E.R., Kebriaei, P., Cianci, C.D., and Morrow, J.S. (1995) *Alpha 1(E)-catenin is an actin-binding and -bundling protein mediating the attachment of F-actin to the membrane adhesion complex.* Proc Natl Acad Sci U S A. 92 (19), p. 8813-7.

22. Yang, C. and Svitkina, T. (2011) *Filopodia initiation: focus on the Arp2/3 complex and formins.* Cell Adh Migr. 5 (5), p. 402-8.

23. Garrod, D. and Chidgey, M. (2008) *Desmosome structure, composition and function.* Biochim Biophys Acta. 1778 (3), p. 572-87.

24. Kundu, S.T., Gosavi, P., Khapare, N., Patel, R., Hosing, A.S., Maru, G.B., Ingle, A., Decaprio, J.A., and Dalal, S.N. (2008) *Plakophilin3 downregulation leads to a decrease in cell adhesion and promotes metastasis.* Int J Cancer. 123 (10), p. 2303-14.

25. Chidgey, M.A., Clarke, J.P., and Garrod, D.R. (1996) *Expression of full-length desmosomol glycoproteins (desmocollins) is not sufficient to confer strong adhesion on transfected L929 cells.* J Invest Dermatol. 106 (4), p. 689-95.

26. Chitaev, N.A. and Troyanovsky, S.M. (1997) *Direct Ca2+-dependent heterophilic interaction between desmosomal cadherins, desmoglein and desmocollin, contributes to cell-cell adhesion.* J Cell Biol. 138 (1), p. 193-201.

27. Marcozzi, C., Burdett, I.D., Buxton, R.S., and Magee, A.I. (1998) *Coexpression of both types of desmosomal cadherin and plakoglobin confers strong intercellular adhesion.* J Cell Sci. 111 (Pt 4), p. 495-509.

28. McGrath, J.A., McMillan, J.R., Shemanko, C.S., Runswick, S.K., Leigh, I.M., Lane, E.B., Garrod, D.R., and Eady, R.A. (1997) *Mutations in the plakophilin 1 gene result in ectodermal dysplasia/skin fragility syndrome.* Nat Genet. 17 (2), p. 240-4.

29. Lewis, J.E., Wahl, J.K., 3rd, Sass, K.M., Jensen, P.J., Johnson, K.R., and Wheelock, M.J. (1997) *Cross-talk between adherens junctions and desmosomes depends on plakoglobin.* J Cell Biol. 136 (4), p. 919-34.

30. Ruiz, P., Brinkmann, V., Ledermann, B., Behrend, M., Grund, C., Thalhammer, C., Vogel, F., Birchmeier, C., Gunthert, U., Franke, W.W., and Birchmeier, W. (1996) *Targeted mutation of plakoglobin in mice reveals essential functions of desmosomes in the embryonic heart.* J Cell Biol. 135 (1), p. 215-25.

31. Zhurinsky, J., Shtutman, M., and Ben-Ze'ev, A. (2000) *Plakoglobin and beta-catenin: protein interactions, regulation and biological roles.* J Cell Sci. 113 (Pt 18), p. 3127-39.

32. Hartsock, A. and Nelson, W.J. (2008) *Adherens and tight junctions: structure, function and connections to the actin cytoskeleton.* Biochim Biophys Acta. 1778 (3), p. 660-9.

33. Tang, V.W. and Goodenough, D.A. (2003) *Paracellular ion channel at the tight junction.* Biophys J. 84 (3), p. 1660-73.

34. Fanning, A.S., Jameson, B.J., Jesaitis, L.A., and Anderson, J.M. (1998) *The tight junction protein ZO-1 establishes a link between the transmembrane protein occludin and the actin cytoskeleton.* J Biol Chem. 273 (45), p. 29745-53.

35. Furuse, M., Itoh, M., Hirase, T., Nagafuchi, A., Yonemura, S., and Tsukita, S. (1994) *Direct association of occludin with ZO-1 and its possible involvement in the localization of occludin at tight junctions.* J Cell Biol. 127 (6 Pt 1), p. 1617-26.

36. Itoh, M., Morita, K., and Tsukita, S. (1999) *Characterization of ZO-2 as a MAGUK family member associated with tight as well as adherens junctions with a binding affinity to occludin and alpha catenin.* J Biol Chem. 274 (9), p. 5981-6.

37. Wittchen, E.S., Haskins, J., and Stevenson, B.R. (2000) *Exogenous expression of the amino-terminal half of the tight junction protein ZO-3 perturbs junctional complex assembly.* J Cell Biol. 151 (4), p. 825-36.

38. Paris, L., Tonutti, L., Vannini, C., and Bazzoni, G. (2008) *Structural organization of the tight junctions.* Biochim Biophys Acta. 1778 (3), p. 646-59.

39. Wittchen, E.S., Haskins, J., and Stevenson, B.R. (2003) *NZO-3 expression causes global changes to actin cytoskeleton in Madin-Darby canine kidney cells: linking a tight junction protein to Rho GTPases.* Mol Biol Cell. 14 (5), p. 1757-68.

40. Howard, S., Deroo, T., Fujita, Y., and Itasaki, N. (2011) *A positive role of cadherin in Wnt/beta-catenin signalling during epithelial-mesenchymal transition.* PLoS One. 6 (8), p. e23899.

41. Teo, J.L. and Kahn, M. (2010) *The Wnt signaling pathway in cellular proliferation and differentiation: A tale of two coactivators.* Adv Drug Deliv Rev. 62 (12), p. 1149-55.

42. Chun, M.G. and Hanahan, D. (2010) *Genetic deletion of the desmosomal component desmoplakin promotes tumor microinvasion in a mouse model of pancreatic neuroendocrine carcinogenesis.* PLoS Genet. 6 (9), p. e1001120.

43. Yang, L., Chen, Y., Cui, T., Knosel, T., Zhang, Q., Albring, K.F., Huber, O., and Petersen, I. (2012) *Desmoplakin acts as a tumor suppressor by inhibition of the Wnt/beta-catenin signaling pathway in human lung cancer.* Carcinogenesis. 33 (10), p. 1863-70.

44. Tobioka, H., Isomura, H., Kokai, Y., Tokunaga, Y., Yamaguchi, J., and Sawada, N. (2004) *Occludin expression decreases with the progression of human endometrial carcinoma.* Hum Pathol. 35 (2), p. 159-64.

45. Orban, E., Szabo, E., Lotz, G., Kupcsulik, P., Paska, C., Schaff, Z., and Kiss, A. (2008) *Different expression of occludin and ZO-1 in primary and metastatic liver tumors.* Pathol Oncol Res. 14 (3), p. 299-306.

46. Tokes, A.M., Kulka, J., Paku, S., Szik, A., Paska, C., Novak, P.K., Szilak, L., Kiss, A., Bogi, K., and Schaff, Z. (2005) *Claudin-1, -3 and -4 proteins and mRNA expression in benign and malignant breast lesions: a research study.* Breast Cancer Res. 7 (2), p. R296-305.

47. Kominsky, S.L., Argani, P., Korz, D., Evron, E., Raman, V., Garrett, E., Rein, A., Sauter, G., Kallioniemi, O.P., and Sukumar, S. (2003) *Loss of the tight junction protein claudin-7 correlates with histological grade in both ductal carcinoma in situ and invasive ductal carcinoma of the breast.* Oncogene. 22 (13), p. 2021-33.

48. Brennan, K., Offiah, G., McSherry, E.A., and Hopkins, A.M. (2009) *Tight junctions: a barrier to the initiation and progression of breast cancer?* J Biomed Biotechnol. 2010, p. 460607.

49. Berrier, A.L. and Yamada, K.M. (2007) *Cell-matrix adhesion.* J Cell Physiol. 213 (3), p. 565-73.

50. Calderwood, D.A., Campbell, I.D., and Critchley, D.R. (2013) *Talins and kindlins: partners in integrin-mediated adhesion.* Nat Rev Mol Cell Biol. 14 (8), p. 503-17.

51. Ganguly, K.K., Pal, S., Moulik, S., and Chatterjee, A. (2013) *Integrins and metastasis.* Cell Adh Migr. 7 (3), p. 251-61.

52. Zaidel-Bar, R., Itzkovitz, S., Ma'ayan, A., Iyengar, R., and Geiger, B. (2007) *Functional atlas of the integrin adhesome.* Nat Cell Biol. 9 (8), p. 858-67.

53. Hynes, R.O. (2002) *Integrins: bidirectional, allosteric signaling machines.* Cell. 110 (6), p. 673-87.

54. Humphries, M.J. (2000) *Integrin structure.* Biochem Soc Trans. 28 (4), p. 311-39.

55. Campbell, I.D. and Humphries, M.J. (2011) *Integrin structure, activation, and interactions.* Cold Spring Harb Perspect Biol. 3 (3).

56. Takagi, J., Strokovich, K., Springer, T.A., and Walz, T. (2003) *Structure of integrin alpha5beta1 in complex with fibronectin.* EMBO J. 22 (18), p. 4607-15.

57. Bouvard, D., Brakebusch, C., Gustafsson, E., Aszodi, A., Bengtsson, T., Berna, A., and Fassler, R. (2001) *Functional consequences of integrin gene mutations in mice.* Circ Res. 89 (3), p. 211-23.

58. Dong, X., Mi, L.Z., Zhu, J., Wang, W., Hu, P., Luo, B.H., and Springer, T.A. (2012) *alpha(V)beta(3) integrin crystal structures and their functional implications.* Biochemistry. 51 (44), p. 8814-28.

59. Anthis, N.J. and Campbell, I.D. (2011) *The tail of integrin activation.* Trends Biochem Sci. 36 (4), p. 191-8.

60. Shattil, S.J., Kim, C., and Ginsberg, M.H. (2010) *The final steps of integrin activation: the end game.* Nat Rev Mol Cell Biol. 11 (4), p. 288-300.

61. Ye, F., Kim, C., and Ginsberg, M.H. (2011) *Molecular mechanism of inside-out integrin regulation.* J Thromb Haemost. 9 Suppl 1, p. 20-5.

62. Burridge, K. and Connell, L. (1983) *A new protein of adhesion plaques and ruffling membranes.* J Cell Biol. 97 (2), p. 359-67.

63. Petrich, B.G. (2009) *Talin-dependent integrin signalling in vivo.* Thromb Haemost. 101 (6), p. 1020-4.

64. Calderwood, D.A., Zent, R., Grant, R., Rees, D.J., Hynes, R.O., and Ginsberg, M.H. (1999) *The Talin head domain binds to integrin beta subunit cytoplasmic tails and regulates integrin activation.* J Biol Chem. 274 (40), p. 28071-4.

65. Legate, K.R. and Fassler, R. (2009) *Mechanisms that regulate adaptor binding to beta-integrin cytoplasmic tails.* J Cell Sci. 122 (Pt 2), p. 187-98.

66. Wegener, K.L., Partridge, A.W., Han, J., Pickford, A.R., Liddington, R.C., Ginsberg, M.H., and Campbell, I.D. (2007) *Structural basis of integrin activation by talin.* Cell. 128 (1), p. 171-82.

67. Wegener, K.L. and Campbell, I.D. (2008) *Transmembrane and cytoplasmic domains in integrin activation and protein-protein interactions (review).* Mol Membr Biol. 25 (5), p. 376-87.

68. Anthis, N.J., Wegener, K.L., Ye, F., Kim, C., Goult, B.T., Lowe, E.D., Vakonakis, I., Bate, N., Critchley, D.R., Ginsberg, M.H., and Campbell, I.D. (2009) *The structure of an integrin/talin complex reveals the basis of inside-out signal transduction.* EMBO J. 28 (22), p. 3623-32.

69. Goksoy, E., Ma, Y.Q., Wang, X., Kong, X., Perera, D., Plow, E.F., and Qin, J. (2008) *Structural basis for the autoinhibition of talin in regulating integrin activation.* Mol Cell. 31 (1), p. 124-33.

70. Goult, B.T., Bate, N., Anthis, N.J., Wegener, K.L., Gingras, A.R., Patel, B., Barsukov, I.L., Campbell, I.D., Roberts, G.C., and Critchley, D.R. (2009) *The structure of an interdomain complex that regulates talin activity.* J Biol Chem. 284 (22), p. 15097-106.

71. Song, X., Yang, J., Hirbawi, J., Ye, S., Perera, H.D., Goksoy, E., Dwivedi, P., Plow, E.F., Zhang, R., and Qin, J. (2012) *A novel membrane-dependent on/off switch mechanism of talin FERM domain at sites of cell adhesion.* Cell Res. 22 (11), p. 1533-45.

72. Legate, K.R., Takahashi, S., Bonakdar, N., Fabry, B., Boettiger, D., Zent, R., and Fassler, R. (2011) *Integrin adhesion and force coupling are independently regulated by localized PtdIns(4,5)2 synthesis.* EMBO J. 30 (22), p. 4539-53.

73. Wynne, J.P., Wu, J., Su, W., Mor, A., Patsoukis, N., Boussiotis, V.A., Hubbard, S.R., and Philips, M.R. (2012) *Rap1-interacting adapter molecule (RIAM) associates with the plasma membrane via a proximity detector.* J Cell Biol. 199 (2), p. 317-30.

74. Lee, H.S., Lim, C.J., Puzon-McLaughlin, W., Shattil, S.J., and Ginsberg, M.H. (2009) *RIAM activates integrins by linking talin to ras GTPase membrane-targeting sequences.* J Biol Chem. 284 (8), p. 5119-27.

75. Watanabe, N., Bodin, L., Pandey, M., Krause, M., Coughlin, S., Boussiotis, V.A., Ginsberg, M.H., and Shattil, S.J. (2008) *Mechanisms and consequences of agonist-induced talin recruitment to platelet integrin alphaIIbbeta3.* J Cell Biol. 181 (7), p. 1211-22.

76. Lawson, C., Lim, S.T., Uryu, S., Chen, X.L., Calderwood, D.A., and Schlaepfer, D.D. (2012) *FAK promotes recruitment of talin to nascent adhesions to control cell motility.* J Cell Biol. 196 (2), p. 223-32.

77. Miller, N.L., Lawson, C., Chen, X.L., Lim, S.T., and Schlaepfer, D.D. (2012) *Rgnef (p190RhoGEF) knockout inhibits RhoA activity, focal adhesion establishment, and cell motility downstream of integrins.* PLoS One. 7 (5), p. e37830.

78. Goult, B.T., Bouaouina, M., Harburger, D.S., Bate, N., Patel, B., Anthis, N.J., Campbell, I.D., Calderwood, D.A., Barsukov, I.L., Roberts, G.C., and Critchley, D.R. (2009) *The structure of the N-terminus of kindlin-1: a domain important for alphaiibbeta3 integrin activation.* J Mol Biol. 394 (5), p. 944-56.

79. Ma, Y.Q., Qin, J., Wu, C., and Plow, E.F. (2008) *Kindlin-2 (Mig-2): a co-activator of beta3 integrins.* J Cell Biol. 181 (3), p. 439-46.

80. Harburger, D.S., Bouaouina, M., and Calderwood, D.A. (2009) *Kindlin-1 and -2 directly bind the C-terminal region of beta integrin cytoplasmic tails and exert integrin-specific activation effects.* J Biol Chem. 284 (17), p. 11485-97.

81. Moser, M., Nieswandt, B., Ussar, S., Pozgajova, M., and Fassler, R. (2008) *Kindlin-3 is essential for integrin activation and platelet aggregation.* Nat Med. 14 (3), p. 325-30.

82. Yates, L.A., Fuzery, A.K., Bonet, R., Campbell, I.D., and Gilbert, R.J. (2012) *Biophysical analysis of Kindlin-3 reveals an elongated conformation and maps integrin binding to the membrane-distal beta-subunit NPXY motif.* J Biol Chem. 287 (45), p. 37715-31.

83. Bledzka, K., Liu, J., Xu, Z., Perera, H.D., Yadav, S.P., Bialkowska, K., Qin, J., Ma, Y.Q., and Plow, E.F. (2012) *Spatial coordination of kindlin-2 with talin head domain in interaction with integrin beta cytoplasmic tails.* J Biol Chem. 287 (29), p. 24585-94.

84. Kahner, B.N., Kato, H., Banno, A., Ginsberg, M.H., Shattil, S.J., and Ye, F. (2012) *Kindlins, integrin activation and the regulation of talin recruitment to alphaIIbbeta3.* PLoS One. 7 (3), p. e34056.

85. Malinin, N.L., Zhang, L., Choi, J., Ciocea, A., Razorenova, O., Ma, Y.Q., Podrez, E.A., Tosi, M., Lennon, D.P., Caplan, A.I., Shurin, S.B., Plow, E.F., and Byzova, T.V. (2009) *A point mutation in KINDLIN3 ablates activation of three integrin subfamilies in humans.* Nat Med. 15 (3), p. 313-8.

86. Qu, H., Tu, Y., Shi, X., Larjava, H., Saleem, M.A., Shattil, S.J., Fukuda, K., Qin, J., Kretzler, M., and Wu, C. (2011) *Kindlin-2 regulates podocyte adhesion and fibronectin matrix deposition through interactions with phosphoinositides and integrins.* J Cell Sci. 124 (Pt 6), p. 879-91.

87. Montanez, E., Ussar, S., Schifferer, M., Bosl, M., Zent, R., Moser, M., and Fassler, R. (2008) *Kindlin-2 controls bidirectional signaling of integrins.* Genes Dev. 22 (10), p. 1325-30.

88. Bouaouina, M., Goult, B.T., Huet-Calderwood, C., Bate, N., Brahme, N.N., Barsukov, I.L., Critchley, D.R., and Calderwood, D.A. (2012) *A conserved lipid-binding loop in the kindlin FERM F1 domain is required for kindlin-mediated alphaIIbbeta3 integrin coactivation.* J Biol Chem. 287 (10), p. 6979-90.

89. Liu, J., Fukuda, K., Xu, Z., Ma, Y.Q., Hirbawi, J., Mao, X., Wu, C., Plow, E.F., and Qin, J. (2011) *Structural basis of phosphoinositide binding to kindlin-2 protein pleckstrin homology domain in regulating integrin activation.* J Biol Chem. 286 (50), p. 43334-42.

90. Perera, H.D., Ma, Y.Q., Yang, J., Hirbawi, J., Plow, E.F., and Qin, J. (2011) *Membrane binding of the N-terminal ubiquitin-like domain of kindlin-2 is crucial for its regulation of integrin activation.* Structure. 19 (11), p. 1664-71.

91. Liu, W., Draheim, K.M., Zhang, R., Calderwood, D.A., and Boggon, T.J. (2013) *Mechanism for KRIT1 release of ICAP1-mediated suppression of integrin activation.* Mol Cell. 49 (4), p. 719-29.

92. Kiema, T., Lad, Y., Jiang, P., Oxley, C.L., Baldassarre, M., Wegener, K.L., Campbell, I.D., Ylanne, J., and Calderwood, D.A. (2006) *The molecular basis of filamin binding to integrins and competition with talin.* Mol Cell. 21 (3), p. 337-47.

93. Yang, X., Pursell, B., Lu, S., Chang, T.K., and Mercurio, A.M. (2009) *Regulation of beta 4-integrin expression by epigenetic modifications in the mammary gland and during the epithelial-to-mesenchymal transition.* J Cell Sci. 122 (Pt 14), p. 2473-80.

94. Gimond, C., van Der Flier, A., van Delft, S., Brakebusch, C., Kuikman, I., Collard, J.G., Fassler, R., and Sonnenberg, A. (1999) *Induction of cell scattering by expression of beta1 integrins in beta1-deficient epithelial cells requires activation of members of the rho family of GTPases and downregulation of cadherin and catenin function.* J Cell Biol. 147 (6), p. 1325-40.

95. Koenig, A., Mueller, C., Hasel, C., Adler, G., and Menke, A. (2006) *Collagen type I induces disruption of E-cadherin-mediated cell-cell contacts and promotes proliferation of pancreatic carcinoma cells.* Cancer Res. 66 (9), p. 4662-71.

96. Maschler, S., Wirl, G., Spring, H., Bredow, D.V., Sordat, I., Beug, H., and Reichmann, E. (2005) *Tumor cell invasiveness correlates with changes in integrin expression and localization.* Oncogene. 24 (12), p. 2032-41.

97. Mise, N., Savai, R., Yu, H., Schwarz, J., Kaminski, N., and Eickelberg, O. (2012) *Zyxin is a transforming growth factor-beta (TGF-beta)/Smad3 target gene that regulates lung cancer cell motility via integrin alpha5beta1.* J Biol Chem. 287 (37), p. 31393-405.

98. Breuss, J.M., Gallo, J., DeLisser, H.M., Klimanskaya, I.V., Folkesson, H.G., Pittet, J.F., Nishimura, S.L., Aldape, K., Landers, D.V., Carpenter, W., and et al. (1995) *Expression of the beta 6 integrin subunit in development, neoplasia and tissue repair suggests a role in epithelial remodeling.* J Cell Sci. 108 (Pt 6), p. 2241-51.

99. Munger, J.S., Huang, X., Kawakatsu, H., Griffiths, M.J., Dalton, S.L., Wu, J., Pittet, J.F., Kaminski, N., Garat, C., Matthay, M.A., Rifkin, D.B., and Sheppard, D. (1999) *The integrin alpha v beta 6 binds and activates latent TGF beta 1: a mechanism for regulating pulmonary inflammation and fibrosis.* Cell. 96 (3), p. 319-28.

100. Mu, D., Cambier, S., Fjellbirkeland, L., Baron, J.L., Munger, J.S., Kawakatsu, H., Sheppard, D., Broaddus, V.C., and Nishimura, S.L. (2002) *The integrin alpha(v)beta8 mediates epithelial homeostasis through MT1-MMP-dependent activation of TGF-beta1.* J Cell Biol. 157 (3), p. 493-507.

101. Annes, J.P., Munger, J.S., and Rifkin, D.B. (2003) *Making sense of latent TGFbeta activation.* J Cell Sci. 116 (Pt 2), p. 217-24.

102. Mercurio, A.M. and Rabinovitz, I. (2001) *Towards a mechanistic understanding of tumor invasion--lessons from the alpha6beta 4 integrin.* Semin Cancer Biol. 11 (2), p. 129-41.

103. Mainiero, F., Murgia, C., Wary, K.K., Curatola, A.M., Pepe, A., Blumemberg, M., Westwick, J.K., Der, C.J., and Giancotti, F.G. (1997) *The coupling of alpha6beta4*

integrin to Ras-MAP kinase pathways mediated by Shc controls keratinocyte proliferation. EMBO J. 16 (9), p. 2365-75.

104. Trusolino, L., Bertotti, A., and Comoglio, P.M. (2001) *A signaling adapter function for alpha6beta4 integrin in the control of HGF-dependent invasive growth.* Cell. 107 (5), p. 643-54.

105. Shaw, L.M., Rabinovitz, I., Wang, H.H., Toker, A., and Mercurio, A.M. (1997) *Activation of phosphoinositide 3-OH kinase by the alpha6beta4 integrin promotes carcinoma invasion.* Cell. 91 (7), p. 949-60.

106. Weaver, V.M., Lelievre, S., Lakins, J.N., Chrenek, M.A., Jones, J.C., Giancotti, F., Werb, Z., and Bissell, M.J. (2002) *beta4 integrin-dependent formation of polarized three-dimensional architecture confers resistance to apoptosis in normal and malignant mammary epithelium.* Cancer Cell. 2 (3), p. 205-16.

107. Zahir, N., Lakins, J.N., Russell, A., Ming, W., Chatterjee, C., Rozenberg, G.I., Marinkovich, M.P., and Weaver, V.M. (2003) *Autocrine laminin-5 ligates alpha6beta4 integrin and activates RAC and NFkappaB to mediate anchorage-independent survival of mammary tumors.* J Cell Biol. 163 (6), p. 1397-407.

108. Alexandrova, A.Y., Arnold, K., Schaub, S., Vasiliev, J.M., Meister, J.J., Bershadsky, A.D., and Verkhovsky, A.B. (2008) *Comparative dynamics of retrograde actin flow and focal adhesions: formation of nascent adhesions triggers transition from fast to slow flow.* PLoS One. 3 (9), p. e3234.

109. Ridley, A.J. (2011) *Life at the leading edge.* Cell. 145 (7), p. 1012-22.

110. Gardel, M.L., Sabass, B., Ji, L., Danuser, G., Schwarz, U.S., and Waterman, C.M. (2008) *Traction stress in focal adhesions correlates biphasically with actin retrograde flow speed.* J Cell Biol. 183 (6), p. 999-1005.

111. Ponti, A., Machacek, M., Gupton, S.L., Waterman-Storer, C.M., and Danuser, G. (2004) *Two distinct actin networks drive the protrusion of migrating cells.* Science. 305 (5691), p. 1782-6.

112. Lin, C.H., Espreafico, E.M., Mooseker, M.S., and Forscher, P. (1997) *Myosin drives retrograde F-actin flow in neuronal growth cones.* Biol Bull. 192 (1), p. 183-5.

113. Choi, C.K., Vicente-Manzanares, M., Zareno, J., Whitmore, L.A., Mogilner, A., and Horwitz, A.R. (2008) *Actin and alpha-actinin orchestrate the assembly and maturation of nascent adhesions in a myosin II motor-independent manner.* Nat Cell Biol. 10 (9), p. 1039-50.

114. Nishizaka, T., Shi, Q., and Sheetz, M.P. (2000) *Position-dependent linkages of fibronectin- integrin-cytoskeleton.* Proc Natl Acad Sci U S A. 97 (2), p. 692-7.

115. Guo, W.H. and Wang, Y.L. (2007) *Retrograde fluxes of focal adhesion proteins in response to cell migration and mechanical signals.* Mol Biol Cell. 18 (11), p. 4519-27.

116. Wolfenson, H., Bershadsky, A., Henis, Y.I., and Geiger, B. (2011) *Actomyosin-generated tension controls the molecular kinetics of focal adhesions.* J Cell Sci. 124 (Pt 9), p. 1425-32.

117. Ji, L., Lim, J., and Danuser, G. (2008) *Fluctuations of intracellular forces during cell protrusion.* Nat Cell Biol. 10 (12), p. 1393-400.

118. Smith, M.L., Gourdon, D., Little, W.C., Kubow, K.E., Eguiluz, R.A., Luna-Morris, S., and Vogel, V. (2007) *Force-induced unfolding of fibronectin in the extracellular matrix of living cells.* PLoS Biol. 5 (10), p. e268.

119. Jiang, G., Giannone, G., Critchley, D.R., Fukumoto, E., and Sheetz, M.P. (2003) *Two-piconewton slip bond between fibronectin and the cytoskeleton depends on talin.* Nature. 424 (6946), p. 334-7.

120. Ziegler, W.H., Gingras, A.R., Critchley, D.R., and Emsley, J. (2008) *Integrin connections to the cytoskeleton through talin and vinculin.* Biochem Soc Trans. 36 (Pt 2), p. 235-9.

121. del Rio, A., Perez-Jimenez, R., Liu, R., Roca-Cusachs, P., Fernandez, J.M., and Sheetz, M.P. (2009) *Stretching single talin rod molecules activates vinculin binding.* Science. 323 (5914), p. 638-41.

122. Lee, S.E., Kamm, R.D., and Mofrad, M.R. (2007) *Force-induced activation of talin and its possible role in focal adhesion mechanotransduction.* J Biomech. 40 (9), p. 2096-106.

123. Ren, X.D., Kiosses, W.B., Sieg, D.J., Otey, C.A., Schlaepfer, D.D., and Schwartz, M.A. (2000) *Focal adhesion kinase suppresses Rho activity to promote focal adhesion turnover.* J Cell Sci. 113 (Pt 20), p. 3673-8.

124. Humphries, J.D., Wang, P., Streuli, C., Geiger, B., Humphries, M.J., and Ballestrem, C. (2007) *Vinculin controls focal adhesion formation by direct interactions with talin and actin.* J Cell Biol. 179 (5), p. 1043-57.

125. Giannone, G., Jiang, G., Sutton, D.H., Critchley, D.R., and Sheetz, M.P. (2003) *Talin1 is critical for force-dependent reinforcement of initial integrin-cytoskeleton bonds but not tyrosine kinase activation.* J Cell Biol. 163 (2), p. 409-19.

126. Mierke, C.T., Kollmannsberger, P., Zitterbart, D.P., Smith, J., Fabry, B., and Goldmann, W.H. (2008) *Mechano-coupling and regulation of contractility by the vinculin tail domain.* Biophys J. 94 (2), p. 661-70.

127. Wang, P., Ballestrem, C., and Streuli, C.H. (2011) *The C terminus of talin links integrins to cell cycle progression.* J Cell Biol. 195 (3), p. 499-513.

128. Margadant, F., Chew, L.L., Hu, X., Yu, H., Bate, N., Zhang, X., and Sheetz, M. (2011) *Mechanotransduction in vivo by repeated talin stretch-relaxation events depends upon vinculin.* PLoS Biol. 9 (12), p. e1001223.

129. Friedland, J.C., Lee, M.H., and Boettiger, D. (2009) *Mechanically activated integrin switch controls alpha5beta1 function.* Science. 323 (5914), p. 642-4.

130. Yu, C.H., Law, J.B., Suryana, M., Low, H.Y., and Sheetz, M.P. (2011) *Early integrin binding to Arg-Gly-Asp peptide activates actin polymerization and contractile movement that stimulates outward translocation.* Proc Natl Acad Sci U S A. 108 (51), p. 20585-90.

131. Walcott, S. and Sun, S.X. (2010) *A mechanical model of actin stress fiber formation and substrate elasticity sensing in adherent cells.* Proc Natl Acad Sci U S A. 107 (17), p. 7757-62.

132. Endlich, N., Otey, C.A., Kriz, W., and Endlich, K. (2007) *Movement of stress fibers away from focal adhesions identifies focal adhesions as sites of stress fiber assembly in stationary cells.* Cell Motil Cytoskeleton. 64 (12), p. 966-76.

133. Hotulainen, P. and Lappalainen, P. (2006) *Stress fibers are generated by two distinct actin assembly mechanisms in motile cells.* J Cell Biol. 173 (3), p. 383-94.

134. Cramer, L.P., Siebert, M., and Mitchison, T.J. (1997) *Identification of novel graded polarity actin filament bundles in locomoting heart fibroblasts: implications for the generation of motile force.* J Cell Biol. 136 (6), p. 1287-305.

135. Soranno, T. and Bell, E. (1982) *Cytostructural dynamics of spreading and translocating cells.* J Cell Biol. 95 (1), p. 127-36.

136. Wang, Y.L. (1984) *Reorganization of actin filament bundles in living fibroblasts.* J Cell Biol. 99 (4 Pt 1), p. 1478-85.

137. Heath, J.P. (1983) *Behaviour and structure of the leading lamella in moving fibroblasts. I. Occurrence and centripetal movement of arc-shaped microfilament bundles beneath the dorsal cell surface.* J Cell Sci. 60, p. 331-54.

138. Ridley, A.J. and Hall, A. (1992) *The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors.* Cell. 70 (3), p. 389-99.

139. Pellegrin, S. and Mellor, H. (2007) *Actin stress fibres.* J Cell Sci. 120 (Pt 20), p. 3491-9.

140. Rajagopalan, P., Marganski, W.A., Brown, X.Q., and Wong, J.Y. (2004) *Direct comparison of the spread area, contractility, and migration of balb/c 3T3 fibroblasts adhered to fibronectin- and RGD-modified substrata.* Biophys J. 87 (4), p. 2818-27.

141. Beningo, K.A., Dembo, M., Kaverina, I., Small, J.V., and Wang, Y.L. (2001) *Nascent focal adhesions are responsible for the generation of strong propulsive forces in migrating fibroblasts.* J Cell Biol. 153 (4), p. 881-8.

142. Yoshigi, M., Hoffman, L.M., Jensen, C.C., Yost, H.J., and Beckerle, M.C. (2005) *Mechanical force mobilizes zyxin from focal adhesions to actin filaments and regulates cytoskeletal reinforcement.* J Cell Biol. 171 (2), p. 209-15.

143. Stricker, J., Aratyn-Schaus, Y., Oakes, P.W., and Gardel, M.L. (2011) *Spatiotemporal constraints on the force-dependent growth of focal adhesions.* Biophys J. 100 (12), p. 2883-93.

144. Bach, C.T., Creed, S., Zhong, J., Mahmassani, M., Schevzov, G., Stehn, J., Cowell, L.N., Naumanen, P., Lappalainen, P., Gunning, P.W., and O'Neill, G.M. (2009) *Tropomyosin isoform expression regulates the transition of adhesions to determine cell speed and direction.* Mol Cell Biol. 29 (6), p. 1506-14.

145. Huttenlocher, A. and Horwitz, A.R. (2011) *Integrins in cell migration.* Cold Spring Harb Perspect Biol. 3 (9), p. a005074.

146. Webb, D.J., Donais, K., Whitmore, L.A., Thomas, S.M., Turner, C.E., Parsons, J.T., and Horwitz, A.F. (2004) *FAK-Src signalling through paxillin, ERK and MLCK regulates adhesion disassembly.* Nat Cell Biol. 6 (2), p. 154-61.

147. Laukaitis, C.M., Webb, D.J., Donais, K., and Horwitz, A.F. (2001) *Differential dynamics of alpha 5 integrin, paxillin, and alpha-actinin during formation and disassembly of adhesions in migrating cells.* J Cell Biol. 153 (7), p. 1427-40.

148. Katz, B.Z., Zamir, E., Bershadsky, A., Kam, Z., Yamada, K.M., and Geiger, B. (2000) *Physical state of the extracellular matrix regulates the structure and molecular composition of cell-matrix adhesions.* Mol Biol Cell. 11 (3), p. 1047-60.

149. Zamir, E., Katz, B.Z., Aota, S., Yamada, K.M., Geiger, B., and Kam, Z. (1999) *Molecular diversity of cell-matrix adhesions.* J Cell Sci. 112 (Pt 11), p. 1655-69.

150. Shemesh, T., Verkhovsky, A.B., Svitkina, T.M., Bershadsky, A.D., and Kozlov, M.M. (2009) *Role of focal adhesions and mechanical stresses in the formation and progression of the lamellipodium-lamellum interface [corrected].* Biophys J. 97 (5), p. 1254-64.

151. Wehrle-Haller, B. (2012) *Assembly and disassembly of cell matrix adhesions.* Curr Opin Cell Biol. 24 (5), p. 569-81.

152. Ezratty, E.J., Bertaux, C., Marcantonio, E.E., and Gundersen, G.G. (2009) *Clathrin mediates integrin endocytosis for focal adhesion disassembly in migrating cells.* J Cell Biol. 187 (5), p. 733-47.

153. Kaverina, I., Krylyshkina, O., and Small, J.V. (1999) *Microtubule targeting of substrate contacts promotes their relaxation and dissociation.* J Cell Biol. 146 (5), p. 1033-44.

154. Ezratty, E.J., Partridge, M.A., and Gundersen, G.G. (2005) *Microtubule-induced focal adhesion disassembly is mediated by dynamin and focal adhesion kinase.* Nat Cell Biol. 7 (6), p. 581-90.

155. Owen, K.A., Pixley, F.J., Thomas, K.S., Vicente-Manzanares, M., Ray, B.J., Horwitz, A.F., Parsons, J.T., Beggs, H.E., Stanley, E.R., and Bouton, A.H. (2007) *Regulation of lamellipodial persistence, adhesion turnover, and motility in macrophages by focal adhesion kinase.* J Cell Biol. 179 (6), p. 1275-87.

156. Klemke, R.L., Cai, S., Giannini, A.L., Gallagher, P.J., de Lanerolle, P., and Cheresh, D.A. (1997) *Regulation of cell motility by mitogen-activated protein kinase.* J Cell Biol. 137 (2), p. 481-92.

157. Somlyo, A.P. and Somlyo, A.V. (2003) *Ca2+ sensitivity of smooth muscle and nonmuscle myosin II: modulated by G proteins, kinases, and myosin phosphatase.* Physiol Rev. 83 (4), p. 1325-58.

158. Kuo, J.C., Han, X., Hsiao, C.T., Yates, J.R., 3rd, and Waterman, C.M. (2011) *Analysis of the myosin-II-responsive focal adhesion proteome reveals a role for beta-Pix in negative regulation of focal adhesion maturation.* Nat Cell Biol. 13 (4), p. 383-93.

159. Chan, K.T., Bennin, D.A., and Huttenlocher, A. (2010) *Regulation of adhesion dynamics by calpain-mediated proteolysis of focal adhesion kinase (FAK).* J Biol Chem. 285 (15), p. 11418-26.

160. Bate, N., Gingras, A.R., Bachir, A., Horwitz, R., Ye, F., Patel, B., Goult, B.T., and Critchley, D.R. (2012) *Talin contains a C-terminal calpain2 cleavage site important in focal adhesion dynamics.* PLoS One. 7 (4), p. e34461.

161. Bridgewater, R.E., Norman, J.C., and Caswell, P.T. (2012) *Integrin trafficking at a glance.* J Cell Sci. 125 (Pt 16), p. 3695-701.

162. White, D.P., Caswell, P.T., and Norman, J.C. (2007) *alpha v beta3 and alpha5beta1 integrin recycling pathways dictate downstream Rho kinase signaling to regulate persistent cell migration.* J Cell Biol. 177 (3), p. 515-25.

163. Semenza, G.L. (2003) *Targeting HIF-1 for cancer therapy.* Nat Rev Cancer. 3 (10), p. 721-32.

164. Semenza, G.L. (2001) *Regulation of hypoxia-induced angiogenesis: a chaperone escorts VEGF to the dance.* J Clin Invest. 108 (1), p. 39-40.

165. Ferrara, N., Gerber, H.P., and LeCouter, J. (2003) *The biology of VEGF and its receptors.* Nat Med. 9 (6), p. 669-76.

166. Bergers, G., Brekken, R., McMahon, G., Vu, T.H., Itoh, T., Tamaki, K., Tanzawa, K., Thorpe, P., Itohara, S., Werb, Z., and Hanahan, D. (2000) *Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis.* Nat Cell Biol. 2 (10), p. 737-44.

167. Grothey, A. and Galanis, E. (2009) *Targeting angiogenesis: progress with anti-VEGF treatment with large molecules.* Nat Rev Clin Oncol. 6 (9), p. 507-18.

168. Grothey, A. and Ellis, L.M. (2008) *Targeting angiogenesis driven by vascular endothelial growth factors using antibody-based therapies.* Cancer J. 14 (3), p. 170-7.

169. Lamalice, L., Le Boeuf, F., and Huot, J. (2007) *Endothelial cell migration during angiogenesis.* Circ Res. 100 (6), p. 782-94.

170. Gonzalez-Moreno, O., Lecanda, J., Green, J.E., Segura, V., Catena, R., Serrano, D., and Calvo, A. (2010) *VEGF elicits epithelial-mesenchymal transition (EMT) in prostate intraepithelial neoplasia (PIN)-like cells via an autocrine loop.* Exp Cell Res. 316 (4), p. 554-67.

171. Sakurai, T. and Kudo, M. (2011) *Signaling pathways governing tumor angiogenesis.* Oncology. 81 Suppl 1, p. 24-9.

172. Cao, Y., Cao, R., and Hedlund, E.M. (2008) *Regulation of tumor angiogenesis and metastasis by FGF and PDGF signaling pathways.* J Mol Med (Berl). 86 (7), p. 785-9.

173. Andrae, J., Gallini, R., and Betsholtz, C. (2008) *Role of platelet-derived growth factors in physiology and medicine.* Genes Dev. 22 (10), p. 1276-312.

174. Lindahl, P., Johansson, B.R., Leveen, P., and Betsholtz, C. (1997) *Pericyte loss and microaneurysm formation in PDGF-B-deficient mice.* Science. 277 (5323), p. 242-5.

175. Augustin, H.G., Koh, G.Y., Thurston, G., and Alitalo, K. (2009) *Control of vascular morphogenesis and homeostasis through the angiopoietin-Tie system.* Nat Rev Mol Cell Biol. 10 (3), p. 165-77.

176. Mantovani, A. and Sica, A. (2010) *Macrophages, innate immunity and cancer: balance, tolerance, and diversity.* Curr Opin Immunol. 22 (2), p. 231-7.

177. DeNardo, D.G., Barreto, J.B., Andreu, P., Vasquez, L., Tawfik, D., Kolhatkar, N., and Coussens, L.M. (2009) *CD4(+) T cells regulate pulmonary metastasis of mammary*

carcinomas by enhancing protumor properties of macrophages. Cancer Cell. 16 (2), p. 91-102.

178. Carmeliet, P. (2003) *Angiogenesis in health and disease.* Nat Med. 9 (6), p. 653-60.

179. Thurston, G. and Kitajewski, J. (2008) *VEGF and Delta-Notch: interacting signalling pathways in tumour angiogenesis.* Br J Cancer. 99 (8), p. 1204-9.

180. Belperio, J.A., Keane, M.P., Arenberg, D.A., Addison, C.L., Ehlert, J.E., Burdick, M.D., and Strieter, R.M. (2000) *CXC chemokines in angiogenesis.* J Leukoc Biol. 68 (1), p. 1-8.

181. Biswas, S.K. and Mantovani, A. (2010) *Macrophage plasticity and interaction with lymphocyte subsets: cancer as a paradigm.* Nat Immunol. 11 (10), p. 889-96.

182. Lin, E.Y., Li, J.F., Gnatovskiy, L., Deng, Y., Zhu, L., Grzesik, D.A., Qian, H., Xue, X.N., and Pollard, J.W. (2006) *Macrophages regulate the angiogenic switch in a mouse model of breast cancer.* Cancer Res. 66 (23), p. 11238-46.

183. Quatromoni, J.G. and Eruslanov, E. (2012) *Tumor-associated macrophages: function, phenotype, and link to prognosis in human lung cancer.* Am J Transl Res. 4 (4), p. 376-89.

184. Fridlender, Z.G. and Albelda, S.M. (2012) *Tumor-associated neutrophils: friend or foe?* Carcinogenesis. 33 (5), p. 949-55.

185. Xing, F., Saidou, J., and Watabe, K. (2010) *Cancer associated fibroblasts (CAFs) in tumor microenvironment.* Front Biosci (Landmark Ed). 15, p. 166-79.

186. Newman, A.C., Nakatsu, M.N., Chou, W., Gershon, P.D., and Hughes, C.C. (2011) *The requirement for fibroblasts in angiogenesis: fibroblast-derived matrix proteins are essential for endothelial cell lumen formation.* Mol Biol Cell. 22 (20), p. 3791-800.

187. Hashizume, H., Baluk, P., Morikawa, S., McLean, J.W., Thurston, G., Roberge, S., Jain, R.K., and McDonald, D.M. (2000) *Openings between defective endothelial cells explain tumor vessel leakiness.* Am J Pathol. 156 (4), p. 1363-80.

188. Dvorak, H.F., Nagy, J.A., Feng, D., Brown, L.F., and Dvorak, A.M. (1999) *Vascular permeability factor/vascular endothelial growth factor and the significance of microvascular hyperpermeability in angiogenesis.* Curr Top Microbiol Immunol. 237, p. 97-132.

189. Weis, S.M. and Cheresh, D.A. (2011) *alphaV integrins in angiogenesis and cancer.* Cold Spring Harb Perspect Med. 1 (1), p. a006478.

190. Miles, F.L., Pruitt, F.L., van Golen, K.L., and Cooper, C.R. (2008) *Stepping out of the flow: capillary extravasation in cancer metastasis.* Clin Exp Metastasis. 25 (4), p. 305-24.

191. Hoeben, A., Landuyt, B., Highley, M.S., Wildiers, H., Van Oosterom, A.T., and De Bruijn, E.A. (2004) *Vascular endothelial growth factor and angiogenesis.* Pharmacol Rev. 56 (4), p. 549-80.

192. Drabsch, Y. and ten Dijke, P. (2011) *TGF-beta signaling in breast cancer cell invasion and bone metastasis.* J Mammary Gland Biol Neoplasia. 16 (2), p. 97-108.

193. Gligorijevic, B., Wyckoff, J., Yamaguchi, H., Wang, Y., Roussos, E.T., and Condeelis, J. (2012) *N-WASP-mediated invadopodium formation is involved in intravasation and lung metastasis of mammary tumors.* J Cell Sci. 125 (Pt 3), p. 724-34.

194. Chabottaux, V., Ricaud, S., Host, L., Blacher, S., Paye, A., Thiry, M., Garofalakis, A., Pestourie, C., Gombert, K., Bruyere, F., Lewandowsky, D., Tavitian, B., Foidart, J.M., Duconge, F., and Noel, A. (2009) *Membrane-type 4 matrix metalloproteinase (MT4-MMP) induces lung metastasis by alteration of primary breast tumour vascular architecture.* J Cell Mol Med. 13 (9B), p. 4002-13.

195. Frohlich, C., Klitgaard, M., Noer, J.B., Kotzsch, A., Nehammer, C., Kronqvist, P., Berthelsen, J., Blobel, C., Kveiborg, M., Albrechtsen, R., and Wewer, U.M. (2013) *ADAM12 is expressed in the tumour vasculature and mediates ectodomain shedding of several membrane-anchored endothelial proteins.* Biochem J. 452 (1), p. 97-109.

196. Sonoshita, M., Aoki, M., Fuwa, H., Aoki, K., Hosogi, H., Sakai, Y., Hashida, H., Takabayashi, A., Sasaki, M., Robine, S., Itoh, K., Yoshioka, K., Kakizaki, F., Kitamura, T., Oshima, M., and Taketo, M.M. (2011) *Suppression of colon cancer metastasis by Aes through inhibition of Notch signaling.* Cancer Cell. 19 (1), p. 125-37.

197. Gil-Bernabe, A.M., Lucotti, S., and Muschel, R.J. (2013) *Coagulation and metastasis: what does the experimental literature tell us?* Br J Haematol. 162 (4), p. 433-41.

198. Camerer, E., Qazi, A.A., Duong, D.N., Cornelissen, I., Advincula, R., and Coughlin, S.R. (2004) *Platelets, protease-activated receptors, and fibrinogen in hematogenous metastasis.* Blood. 104 (2), p. 397-401.

199. Stoletov, K., Kato, H., Zardouzian, E., Kelber, J., Yang, J., Shattil, S., and Klemke, R. (2010) *Visualizing extravasation dynamics of metastatic tumor cells.* J Cell Sci. 123 (Pt 13), p. 2332-41.

200. Ito, S., Nakanishi, H., Ikehara, Y., Kato, T., Kasai, Y., Ito, K., Akiyama, S., Nakao, A., and Tatematsu, M. (2001) *Real-time observation of micrometastasis formation in*

the living mouse liver using a green fluorescent protein gene-tagged rat tongue carcinoma cell line. Int J Cancer. 93 (2), p. 212-7.

201. Kienast, Y., von Baumgarten, L., Fuhrmann, M., Klinkert, W.E., Goldbrunner, R., Herms, J., and Winkler, F. (2010) *Real-time imaging reveals the single steps of brain metastasis formation.* Nat Med. 16 (1), p. 116-22.
202. Strell, C. and Entschladen, F. (2008) *Extravasation of leukocytes in comparison to tumor cells.* Cell Commun Signal. 6, p. 10.
203. Shirure, V.S., Reynolds, N.M., and Burdick, M.M. (2012) *Mac-2 binding protein is a novel E-selectin ligand expressed by breast cancer cells.* PLoS One. 7 (9), p. e44529.
204. Qi, J., Chen, N., Wang, J., and Siu, C.H. (2005) *Transendothelial migration of melanoma cells involves N-cadherin-mediated adhesion and activation of the beta-catenin signaling pathway.* Mol Biol Cell. 16 (9), p. 4386-97.
205. Orian-Rousseau, V. (2010) *CD44, a therapeutic target for metastasising tumours.* Eur J Cancer. 46 (7), p. 1271-7.
206. Zen, K., Liu, D.Q., Guo, Y.L., Wang, C., Shan, J., Fang, M., Zhang, C.Y., and Liu, Y. (2008) *CD44v4 is a major E-selectin ligand that mediates breast cancer cell transendothelial migration.* PLoS One. 3 (3), p. e1826.
207. Reymond, N., d'Agua, B.B., and Ridley, A.J. (2013) *Crossing the endothelial barrier during metastasis.* Nat Rev Cancer. 13 (12), p. 858-70.
208. Thiery, J.P., Acloque, H., Huang, R.Y., and Nieto, M.A. (2009) *Epithelial-mesenchymal transitions in development and disease.* Cell. 139 (5), p. 871-90.
209. Thiery, J.P. and Sleeman, J.P. (2006) *Complex networks orchestrate epithelial-mesenchymal transitions.* Nat Rev Mol Cell Biol. 7 (2), p. 131-42.
210. Lamouille, S., Xu, J., and Derynck, R. (2014) *Molecular mechanisms of epithelial-mesenchymal transition.* Nat Rev Mol Cell Biol. 15 (3), p. 178-96.
211. Kalluri, R. and Weinberg, R.A. (2009) *The basics of epithelial-mesenchymal transition.* J Clin Invest. 119 (6), p. 1420-8.
212. St Johnston, D. and Ahringer, J. (2010) *Cell polarity in eggs and epithelia: parallels and diversity.* Cell. 141 (5), p. 757-74.
213. Huang, R.Y., Guilford, P., and Thiery, J.P. (2012) *Early events in cell adhesion and polarity during epithelial-mesenchymal transition.* J Cell Sci. 125 (Pt 19), p. 4417-22.
214. Yilmaz, M. and Christofori, G. (2009) *EMT, the cytoskeleton, and cancer cell invasion.* Cancer Metastasis Rev. 28 (1-2), p. 15-33.
215. Niehrs, C. (2012) *The complex world of WNT receptor signalling.* Nat Rev Mol Cell Biol. 13 (12), p. 767-79.

216. Kourtidis, A., Ngok, S.P., and Anastasiadis, P.Z. (2013) *p120 catenin: an essential regulator of cadherin stability, adhesion-induced signaling, and cancer progression.* Prog Mol Biol Transl Sci. 116, p. 409-32.

217. Yilmaz, M. and Christofori, G. (2010) *Mechanisms of motility in metastasizing cells.* Mol Cancer Res. 8 (5), p. 629-42.

218. Wheelock, M.J., Shintani, Y., Maeda, M., Fukumoto, Y., and Johnson, K.R. (2008) *Cadherin switching.* J Cell Sci. 121 (Pt 6), p. 727-35.

219. Theveneau, E. and Mayor, R. (2012) *Cadherins in collective cell migration of mesenchymal cells.* Curr Opin Cell Biol. 24 (5), p. 677-84.

220. Cavallaro, U. and Christofori, G. (2004) *Cell adhesion and signalling by cadherins and Ig-CAMs in cancer.* Nat Rev Cancer. 4 (2), p. 118-32.

221. Lehembre, F., Yilmaz, M., Wicki, A., Schomber, T., Strittmatter, K., Ziegler, D., Kren, A., Went, P., Derksen, P.W., Berns, A., Jonkers, J., and Christofori, G. (2008) *NCAM-induced focal adhesion assembly: a functional switch upon loss of E-cadherin.* EMBO J. 27 (19), p. 2603-15.

222. Toivola, D.M., Tao, G.Z., Habtezion, A., Liao, J., and Omary, M.B. (2005) *Cellular integrity plus: organelle-related and protein-targeting functions of intermediate filaments.* Trends Cell Biol. 15 (11), p. 608-17.

223. Mendez, M.G., Kojima, S., and Goldman, R.D. (2010) *Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition.* FASEB J. 24 (6), p. 1838-51.

224. Feng, X.H. and Derynck, R. (2005) *Specificity and versatility in tgf-beta signaling through Smads.* Annu Rev Cell Dev Biol. 21, p. 659-93.

225. Chen, D., Zhao, M., and Mundy, G.R. (2004) *Bone morphogenetic proteins.* Growth Factors. 22 (4), p. 233-41.

226. Ikushima, H. and Miyazono, K. (2010) *TGFbeta signalling: a complex web in cancer progression.* Nat Rev Cancer. 10 (6), p. 415-24.

227. Morita, T., Mayanagi, T., and Sobue, K. (2007) *Dual roles of myocardin-related transcription factors in epithelial mesenchymal transition via slug induction and actin remodeling.* J Cell Biol. 179 (5), p. 1027-42.

228. Hoot, K.E., Lighthall, J., Han, G., Lu, S.L., Li, A., Ju, W., Kulesz-Martin, M., Bottinger, E., and Wang, X.J. (2008) *Keratinocyte-specific Smad2 ablation results in increased epithelial-mesenchymal transition during skin cancer formation and progression.* J Clin Invest. 118 (8), p. 2722-32.

229. Shirakihara, T., Saitoh, M., and Miyazono, K. (2007) *Differential regulation of epithelial and mesenchymal markers by deltaEF1 proteins in epithelial mesenchymal transition induced by TGF-beta.* Mol Biol Cell. 18 (9), p. 3533-44.

230. Kang, Y., Chen, C.R., and Massague, J. (2003) *A self-enabling TGFbeta response coupled to stress signaling: Smad engages stress response factor ATF3 for Id1 repression in epithelial cells.* Mol Cell. 11 (4), p. 915-26.

231. Yang, Y.C., Piek, E., Zavadil, J., Liang, D., Xie, D., Heyer, J., Pavlidis, P., Kucherlapati, R., Roberts, A.B., and Bottinger, E.P. (2003) *Hierarchical model of gene regulation by transforming growth factor beta.* Proc Natl Acad Sci U S A. 100 (18), p. 10269-74.

232. Zavadil, J., Bitzer, M., Liang, D., Yang, Y.C., Massimi, A., Kneitz, S., Piek, E., and Bottinger, E.P. (2001) *Genetic programs of epithelial cell plasticity directed by transforming growth factor-beta.* Proc Natl Acad Sci U S A. 98 (12), p. 6686-91.

233. Kaimori, A., Potter, J., Kaimori, J.Y., Wang, C., Mezey, E., and Koteish, A. (2007) *Transforming growth factor-beta1 induces an epithelial-to-mesenchymal transition state in mouse hepatocytes in vitro.* J Biol Chem. 282 (30), p. 22089-101.

234. Derynck, R. and Zhang, Y.E. (2003) *Smad-dependent and Smad-independent pathways in TGF-beta family signalling.* Nature. 425 (6958), p. 577-84.

235. Bakin, A.V., Tomlinson, A.K., Bhowmick, N.A., Moses, H.L., and Arteaga, C.L. (2000) *Phosphatidylinositol 3-kinase function is required for transforming growth factor beta-mediated epithelial to mesenchymal transition and cell migration.* J Biol Chem. 275 (47), p. 36803-10.

236. Xie, L., Law, B.K., Chytil, A.M., Brown, K.A., Aakre, M.E., and Moses, H.L. (2004) *Activation of the Erk pathway is required for TGF-beta1-induced EMT in vitro.* Neoplasia. 6 (5), p. 603-10.

237. Yu, L., Hebert, M.C., and Zhang, Y.E. (2002) *TGF-beta receptor-activated p38 MAP kinase mediates Smad-independent TGF-beta responses.* EMBO J. 21 (14), p. 3749-59.

238. Marchetti, A., Colletti, M., Cozzolino, A.M., Steindler, C., Lunadei, M., Mancone, C., and Tripodi, M. (2008) *ERK5/MAPK is activated by TGFbeta in hepatocytes and required for the GSK-3beta-mediated Snail protein stabilization.* Cell Signal. 20 (11), p. 2113-8.

239. Grande, M., Franzen, A., Karlsson, J.O., Ericson, L.E., Heldin, N.E., and Nilsson, M. (2002) *Transforming growth factor-beta and epidermal growth factor synergistically stimulate epithelial to mesenchymal transition (EMT) through a MEK-*

dependent mechanism in primary cultured pig thyrocytes. J Cell Sci. 115 (Pt 22), p. 4227-36.

240. Uttamsingh, S., Bao, X., Nguyen, K.T., Bhanot, M., Gong, J., Chan, J.L., Liu, F., Chu, T.T., and Wang, L.H. (2008) *Synergistic effect between EGF and TGF-beta1 in inducing oncogenic properties of intestinal epithelial cells.* Oncogene. 27 (18), p. 2626-34.

241. Lamouille, S., Connolly, E., Smyth, J.W., Akhurst, R.J., and Derynck, R. (2012) *TGF-beta-induced activation of mTOR complex 2 drives epithelial-mesenchymal transition and cell invasion.* J Cell Sci. 125 (Pt 5), p. 1259-73.

242. Etienne-Manneville, S. and Hall, A. (2002) *Rho GTPases in cell biology.* Nature. 420 (6916), p. 629-35.

243. Terry, S., Nie, M., Matter, K., and Balda, M.S. (2010) *Rho signaling and tight junction functions.* Physiology (Bethesda). 25 (1), p. 16-26.

244. Ozdamar, B., Bose, R., Barrios-Rodiles, M., Wang, H.R., Zhang, Y., and Wrana, J.L. (2005) *Regulation of the polarity protein Par6 by TGFbeta receptors controls epithelial cell plasticity.* Science. 307 (5715), p. 1603-9.

245. Bhowmick, N.A., Ghiassi, M., Bakin, A., Aakre, M., Lundquist, C.A., Engel, M.E., Arteaga, C.L., and Moses, H.L. (2001) *Transforming growth factor-beta1 mediates epithelial to mesenchymal transdifferentiation through a RhoA-dependent mechanism.* Mol Biol Cell. 12 (1), p. 27-36.

246. Wu, M.Y. and Hill, C.S. (2009) *Tgf-beta superfamily signaling in embryonic development and homeostasis.* Dev Cell. 16 (3), p. 329-43.

247. LeBrun, J.-J. (2012) *The dual role of TGF in human cancer: from tumor suppression to cancer metastasis.* ISRN Molecular Biology 2012, p. 1-28.

248. Derynck, R., Akhurst, R.J., and Balmain, A. (2001) *TGF-beta signaling in tumor suppression and cancer progression.* Nat Genet. 29 (2), p. 117-29.

249. Funaba, M., Zimmerman, C.M., and Mathews, L.S. (2002) *Modulation of Smad2-mediated signaling by extracellular signal-regulated kinase.* J Biol Chem. 277 (44), p. 41361-8.

250. Gulhati, P., Bowen, K.A., Liu, J., Stevens, P.D., Rychahou, P.G., Chen, M., Lee, E.Y., Weiss, H.L., O'Connor, K.L., Gao, T., and Evers, B.M. (2011) *mTORC1 and mTORC2 regulate EMT, motility, and metastasis of colorectal cancer via RhoA and Rac1 signaling pathways.* Cancer Res. 71 (9), p. 3246-56.

251. Billottet, C., Tuefferd, M., Gentien, D., Rapinat, A., Thiery, J.P., Broet, P., and Jouanneau, J. (2008) *Modulation of several waves of gene expression during FGF-1*

induced epithelial-mesenchymal transition of carcinoma cells. J Cell Biochem. 104 (3), p. 826-39.

252. Valles, A.M., Boyer, B., Tarone, G., and Thiery, J.P. (1996) *Alpha 2 beta 1 integrin is required for the collagen and FGF-1 induced cell dispersion in a rat bladder carcinoma cell line.* Cell Adhes Commun. 4 (3), p. 187-99.

253. Savagner, P., Yamada, K.M., and Thiery, J.P. (1997) *The zinc-finger protein slug causes desmosome dissociation, an initial and necessary step for growth factor-induced epithelial-mesenchymal transition.* J Cell Biol. 137 (6), p. 1403-19.

254. Yang, A.D., Camp, E.R., Fan, F., Shen, L., Gray, M.J., Liu, W., Somcio, R., Bauer, T.W., Wu, Y., Hicklin, D.J., and Ellis, L.M. (2006) *Vascular endothelial growth factor receptor-1 activation mediates epithelial to mesenchymal transition in human pancreatic carcinoma cells.* Cancer Res. 66 (1), p. 46-51.

255. Wanami, L.S., Chen, H.Y., Peiro, S., Garcia de Herreros, A., and Bachelder, R.E. (2008) *Vascular endothelial growth factor-A stimulates Snail expression in breast tumor cells: implications for tumor progression.* Exp Cell Res. 314 (13), p. 2448-53.

256. Peinado, H., Marin, F., Cubillo, E., Stark, H.J., Fusenig, N., Nieto, M.A., and Cano, A. (2004) *Snail and E47 repressors of E-cadherin induce distinct invasive and angiogenic properties in vivo.* J Cell Sci. 117 (Pt 13), p. 2827-39.

257. Liu, P., Wakamiya, M., Shea, M.J., Albrecht, U., Behringer, R.R., and Bradley, A. (1999) *Requirement for Wnt3 in vertebrate axis formation.* Nat Genet. 22 (4), p. 361-5.

258. Zhou, B.P., Deng, J., Xia, W., Xu, J., Li, Y.M., Gunduz, M., and Hung, M.C. (2004) *Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition.* Nat Cell Biol. 6 (10), p. 931-40.

259. Xie, M., Zhang, L., He, C.S., Xu, F., Liu, J.L., Hu, Z.H., Zhao, L.P., and Tian, Y. (2012) *Activation of Notch-1 enhances epithelial-mesenchymal transition in gefitinib-acquired resistant lung cancer cells.* J Cell Biochem. 113 (5), p. 1501-13.

260. Niessen, K., Fu, Y., Chang, L., Hoodless, P.A., McFadden, D., and Karsan, A. (2008) *Slug is a direct Notch target required for initiation of cardiac cushion cellularization.* J Cell Biol. 182 (2), p. 315-25.

261. De Craene, B. and Berx, G. (2013) *Regulatory networks defining EMT during cancer initiation and progression.* Nat Rev Cancer. 13 (2), p. 97-110.

262. Peinado, H., Olmeda, D., and Cano, A. (2007) *Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype?* Nat Rev Cancer. 7 (6), p. 415-28.

263. Zhang, J. and Ma, L. (2012) *MicroRNA control of epithelial-mesenchymal transition and metastasis.* Cancer Metastasis Rev. 31 (3-4), p. 653-62.

264. Yao, D., Dai, C., and Peng, S. (2011) *Mechanism of the mesenchymal-epithelial transition and its relationship with metastatic tumor formation.* Mol Cancer Res. 9 (12), p. 1608-20.

265. Chaffer, C.L., Brennan, J.P., Slavin, J.L., Blick, T., Thompson, E.W., and Williams, E.D. (2006) *Mesenchymal-to-epithelial transition facilitates bladder cancer metastasis: role of fibroblast growth factor receptor-2.* Cancer Res. 66 (23), p. 11271-8.

266. Oltean, S., Febbo, P.G., and Garcia-Blanco, M.A. (2008) *Dunning rat prostate adenocarcinomas and alternative splicing reporters: powerful tools to study epithelial plasticity in prostate tumors in vivo.* Clin Exp Metastasis. 25 (6), p. 611-9.

267. Oltean, S., Sorg, B.S., Albrecht, T., Bonano, V.I., Brazas, R.M., Dewhirst, M.W., and Garcia-Blanco, M.A. (2006) *Alternative inclusion of fibroblast growth factor receptor 2 exon IIIc in Dunning prostate tumors reveals unexpected epithelial mesenchymal plasticity.* Proc Natl Acad Sci U S A. 103 (38), p. 14116-21.

268. Dykxhoorn, D.M., Wu, Y., Xie, H., Yu, F., Lal, A., Petrocca, F., Martinvalet, D., Song, E., Lim, B., and Lieberman, J. (2009) *miR-200 enhances mouse breast cancer cell colonization to form distant metastases.* PLoS One. 4 (9), p. e7181.

269. Tsuji, T., Ibaragi, S., and Hu, G.F. (2009) *Epithelial-mesenchymal transition and cell cooperativity in metastasis.* Cancer Res. 69 (18), p. 7135-9.

270. Shipitsin, M., Campbell, L.L., Argani, P., Weremowicz, S., Bloushtain-Qimron, N., Yao, J., Nikolskaya, T., Serebryiskaya, T., Beroukhim, R., Hu, M., Halushka, M.K., Sukumar, S., Parker, L.M., Anderson, K.S., Harris, L.N., Garber, J.E., Richardson, A.L., Schnitt, S.J., Nikolsky, Y., Gelman, R.S., and Polyak, K. (2007) *Molecular definition of breast tumor heterogeneity.* Cancer Cell. 11 (3), p. 259-73.

271. Frisch, S.M. (1997) *The epithelial cell default-phenotype hypothesis and its implications for cancer.* Bioessays. 19 (8), p. 705-9.

272. Yates, C.C., Shepard, C.R., Stolz, D.B., and Wells, A. (2007) *Co-culturing human prostate carcinoma cells with hepatocytes leads to increased expression of E-cadherin.* Br J Cancer. 96 (8), p. 1246-52.

273. Meng, F. and Wu, G. (2012) *The rejuvenated scenario of epithelial-mesenchymal transition (EMT) and cancer metastasis.* Cancer Metastasis Rev. 31 (3-4), p. 455-67.

274. Paget, S. (1989) *The distribution of secondary growths in cancer of the breast. 1889.* Cancer Metastasis Rev. 8 (2), p. 98-101.

275. Wells, A., Yates, C., and Shepard, C.R. (2008) *E-cadherin as an indicator of mesenchymal to epithelial reverting transitions during the metastatic seeding of disseminated carcinomas.* Clin Exp Metastasis. 25 (6), p. 621-8.

276. Chiarugi, P. and Giannoni, E. (2008) *Anoikis: a necessary death program for anchorage-dependent cells.* Biochem Pharmacol. 76 (11), p. 1352-64.

277. Kim, Y.N., Koo, K.H., Sung, J.Y., Yun, U.J., and Kim, H. (2012) *Anoikis resistance: an essential prerequisite for tumor metastasis.* Int J Cell Biol. 2012, p. 306879.

278. Coates, J.M., Galante, J.M., and Bold, R.J. (2010) *Cancer therapy beyond apoptosis: autophagy and anoikis as mechanisms of cell death.* J Surg Res. 164 (2), p. 301-8.

279. Tan, K., Goldstein, D., Crowe, P., and Yang, J.L. (2013) *Uncovering a key to the process of metastasis in human cancers: a review of critical regulators of anoikis.* J Cancer Res Clin Oncol. 139 (11), p. 1795-805.

280. Wong, W.W. and Puthalakath, H. (2008) *Bcl-2 family proteins: the sentinels of the mitochondrial apoptosis pathway.* IUBMB Life. 60 (6), p. 390-7.

281. Shigemasa, K., Katoh, O., Shiroyama, Y., Mihara, S., Mukai, K., Nagai, N., and Ohama, K. (2002) *Increased MCL-1 expression is associated with poor prognosis in ovarian carcinomas.* Jpn J Cancer Res. 93 (5), p. 542-50.

282. Watanabe, J., Kushihata, F., Honda, K., Sugita, A., Tateishi, N., Mominoki, K., Matsuda, S., and Kobayashi, N. (2004) *Prognostic significance of Bcl-xL in human hepatocellular carcinoma.* Surgery. 135 (6), p. 604-12.

283. Chunhacha, P., Pongrakhananon, V., Rojanasakul, Y., and Chanvorachote, P. (2012) *Caveolin-1 regulates Mcl-1 stability and anoikis in lung carcinoma cells.* Am J Physiol Cell Physiol. 302 (9), p. C1284-92.

284. Li, Z., Zhao, J., Du, Y., Park, H.R., Sun, S.Y., Bernal-Mizrachi, L., Aitken, A., Khuri, F.R., and Fu, H. (2008) *Down-regulation of 14-3-3zeta suppresses anchorage-independent growth of lung cancer cells through anoikis activation.* Proc Natl Acad Sci U S A. 105 (1), p. 162-7.

285. Danes, C.G., Wyszomierski, S.L., Lu, J., Neal, C.L., Yang, W., and Yu, D. (2008) *14-3-3 zeta down-regulates p53 in mammary epithelial cells and confers luminal filling.* Cancer Res. 68 (6), p. 1760-7.

286. Neal, C.L. and Yu, D. (2009) *14-3-3zeta as a prognostic marker and therapeutic target for cancer.* Expert Opin Ther Targets. 14 (12), p. 1343-54.

287. Matta, A., Siu, K.W., and Ralhan, R. (2012) *14-3-3 zeta as novel molecular target for cancer therapy.* Expert Opin Ther Targets. 16 (5), p. 515-23.

288. Ulukaya, E., Acilan, C., and Yilmaz, Y. (2011) *Apoptosis: why and how does it occur in biology?* Cell Biochem Funct. 29 (6), p. 468-80.

289. Aoki, K., Kurooka, M., Chen, J.J., Petryniak, J., Nabel, E.G., and Nabel, G.J. (2001) *Extracellular matrix interacts with soluble CD95L: retention and enhancement of cytotoxicity.* Nat Immunol. 2 (4), p. 333-7.

290. Mawji, I.A., Simpson, C.D., Hurren, R., Gronda, M., Williams, M.A., Filmus, J., Jonkman, J., Da Costa, R.S., Wilson, B.C., Thomas, M.P., Reed, J.C., Glinsky, G.V., and Schimmer, A.D. (2007) *Critical role for Fas-associated death domain-like interleukin-1-converting enzyme-like inhibitory protein in anoikis resistance and distant tumor formation.* J Natl Cancer Inst. 99 (10), p. 811-22.

291. Chen, H.X., Liu, Y.J., Zhou, X.D., and Luo, R.Y. (2005) *Expression of cellular FLICE/caspase-8 inhibitory protein is associated with malignant potential in endometrial carcinoma.* Int J Gynecol Cancer. 15 (4), p. 663-70.

292. Safa, A.R. and Pollok, K.E. (2011) *Targeting the Anti-Apoptotic Protein c-FLIP for Cancer Therapy.* Cancers (Basel). 3 (2), p. 1639-71.

293. Brunquell, C., Biliran, H., Jennings, S., Ireland, S.K., Chen, R., and Ruoslahti, E. (2012) *TLE1 is an anoikis regulator and is downregulated by Bit1 in breast cancer cells.* Mol Cancer Res. 10 (11), p. 1482-95.

294. Jan, Y., Matter, M., Pai, J.T., Chen, Y.L., Pilch, J., Komatsu, M., Ong, E., Fukuda, M., and Ruoslahti, E. (2004) *A mitochondrial protein, Bit1, mediates apoptosis regulated by integrins and Groucho/TLE corepressors.* Cell. 116 (5), p. 751-62.

295. Jenning, S., Pham, T., Ireland, S.K., Ruoslahti, E., and Biliran, H. (2013) *Bit1 in anoikis resistance and tumor metastasis.* Cancer Lett. 333 (2), p. 147-51.

296. Karmali, P.P., Brunquell, C., Tram, H., Ireland, S.K., Ruoslahti, E., and Biliran, H. (2011) *Metastasis of tumor cells is enhanced by downregulation of Bit1.* PLoS One. 6 (8), p. e23840.

297. Guadamillas, M.C., Cerezo, A., and Del Pozo, M.A. (2011) *Overcoming anoikis--pathways to anchorage-independent growth in cancer.* J Cell Sci. 124 (Pt 19), p. 3189-97.

298. Nylandsted, J., Rohde, M., Brand, K., Bastholm, L., Elling, F., and Jaattela, M. (2000) *Selective depletion of heat shock protein 70 (Hsp70) activates a tumor-specific death program that is independent of caspases and bypasses Bcl-2.* Proc Natl Acad Sci U S A. 97 (14), p. 7871-6.

299. Jolly, C. and Morimoto, R.I. (2000) *Role of the heat shock response and molecular chaperones in oncogenesis and cell death.* J Natl Cancer Inst. 92 (19), p. 1564-72.

300. Van Drie, J.H. (2011) *Protein folding, protein homeostasis, and cancer.* Chin J Cancer. 30 (2), p. 124-37.

301. Park, K., Chung, Y.J., So, H., Kim, K., Park, J., Oh, M., Jo, M., Choi, K., Lee, E.J., Choi, Y.L., Song, S.Y., Bae, D.S., Kim, B.G., and Lee, J.H. (2011) *AGR2, a mucinous ovarian cancer marker, promotes cell proliferation and migration.* Exp Mol Med. 43 (2), p. 91-100.

302. McKinney, K.Q., Lee, J.G., Sindram, D., Russo, M.W., Han, D.K., Bonkovsky, H.L., and Hwang, S.I. (2012) *Identification of differentially expressed proteins from primary versus metastatic pancreatic cancer cells using subcellular proteomics.* Cancer Genomics Proteomics. 9 (5), p. 257-63.

303. Inoue, M., Hiyama, K., Nakabayashi, K., Morii, E., Minami, M., Sawabata, N., Shintani, Y., Nakagiri, T., Susaki, Y., Maeda, J., Higashiyama, M., Okami, J., Yoshida, Y., Ding, J., Otomo, Y., and Okumura, M. (2012) *An accurate and rapid detection of lymph node metastasis in non-small cell lung cancer patients based on one-step nucleic acid amplification assay.* Lung Cancer. 78 (3), p. 212-8.

304. Yu, H., Zhao, J., Lin, L., Zhang, Y., Zhong, F., Liu, Y., Yu, Y., Shen, H., Han, M., He, F., and Yang, P. (2012) *Proteomic study explores AGR2 as pro-metastatic protein in HCC.* Mol Biosyst. 8 (10), p. 2710-8.

305. Zhang, Y., Ali, T.Z., Zhou, H., D'Souza, D.R., Lu, Y., Jaffe, J., Liu, Z., Passaniti, A., and Hamburger, A.W. (2010) *ErbB3 binding protein 1 represses metastasis-promoting gene anterior gradient protein 2 in prostate cancer.* Cancer Res. 70 (1), p. 240-8.

306. Smirnov, D.A., Zweitzig, D.R., Foulk, B.W., Miller, M.C., Doyle, G.V., Pienta, K.J., Meropol, N.J., Weiner, L.M., Cohen, S.J., Moreno, J.G., Connelly, M.C., Terstappen, L.W., and O'Hara, S.M. (2005) *Global gene expression profiling of circulating tumor cells.* Cancer Res. 65 (12), p. 4993-7.

307. Sweeny, L., Liu, Z., Bush, B.D., Hartman, Y., Zhou, T., and Rosenthal, E.L. (2012) *CD147 and AGR2 expression promote cellular proliferation and metastasis of head and neck squamous cell carcinoma.* Exp Cell Res. 318 (14), p. 1788-98.

308. Li, Y., Lu, J., Peng, Z., Tan, G., Liu, N., Huang, D., Zhang, Z., Duan, C., Tang, X., and Tang, F. (2014) *N,N'-dinitrosopiperazine-mediated AGR2 is involved in metastasis of nasopharyngeal carcinoma.* PLoS One. 9 (4), p. e92081.

309. Liu, D., Rudland, P.S., Sibson, D.R., Platt-Higgins, A., and Barraclough, R. (2005) *Human homologue of cement gland protein, a novel metastasis inducer associated with breast carcinomas.* Cancer Res. 65 (9), p. 3796-805.

310. Gidalevitz, T., Stevens, F., and Argon, Y. (2013) *Orchestration of secretory protein folding by ER chaperones.* Biochim Biophys Acta. 1833 (11), p. 2410-24.

311. Hatahet, F. and Ruddock, L.W. (2009) *Protein disulfide isomerase: a critical evaluation of its function in disulfide bond formation.* Antioxid Redox Signal. 11 (11), p. 2807-50.

312. Appenzeller-Herzog, C. and Ellgaard, L. (2008) *The human PDI family: versatility packed into a single fold.* Biochim Biophys Acta. 1783 (4), p. 535-48.

313. Creighton, T.E. (1997) *Protein folding coupled to disulphide bond formation.* Biol Chem. 378 (8), p. 731-44.

314. Goldberger, R.F., Epstein, C.J., and Anfinsen, C.B. (1964) *Purification and Properties of a Microsomal Enzyme System Catalyzing the Reactivation of Reduced Ribonuclease and Lysozyme.* J Biol Chem. 239, p. 1406-10.

315. Kozlov, G., Maattanen, P., Thomas, D.Y., and Gehring, K. (2010) *A structural overview of the PDI family of proteins.* FEBS J. 277 (19), p. 3924-36.

316. Kozlov, G., Maattanen, P., Schrag, J.D., Pollock, S., Cygler, M., Nagar, B., Thomas, D.Y., and Gehring, K. (2006) *Crystal structure of the bb' domains of the protein disulfide isomerase ERp57.* Structure. 14 (8), p. 1331-9.

317. Walker, K.W. and Gilbert, H.F. (1997) *Scanning and escape during protein-disulfide isomerase-assisted protein folding.* J Biol Chem. 272 (14), p. 8845-8.

318. van Lith, M., Hartigan, N., Hatch, J., and Benham, A.M. (2005) *PDILT, a divergent testis-specific protein disulfide isomerase with a non-classical SXXC motif that engages in disulfide-dependent interactions in the endoplasmic reticulum.* J Biol Chem. 280 (2), p. 1376-83.

319. Sugiura, Y., Araki, K., Iemura, S., Natsume, T., Hoseki, J., and Nagata, K. (2010) *Novel thioredoxin-related transmembrane protein TMX4 has reductase activity.* J Biol Chem. 285 (10), p. 7135-42.

320. Anelli, T., Alessio, M., Mezghrani, A., Simmen, T., Talamo, F., Bachi, A., and Sitia, R. (2002) *ERp44, a novel endoplasmic reticulum folding assistant of the thioredoxin family.* EMBO J. 21 (4), p. 835-44.

321. Darby, N.J. and Creighton, T.E. (1995) *Functional properties of the individual thioredoxin-like domains of protein disulfide isomerase.* Biochemistry. 34 (37), p. 11725-35.

322. Frickel, E.M., Frei, P., Bouvier, M., Stafford, W.F., Helenius, A., Glockshuber, R., and Ellgaard, L. (2004) *ERp57 is a multifunctional thiol-disulfide oxidoreductase.* J Biol Chem. 279 (18), p. 18277-87.

323. Alanen, H.I., Williamson, R.A., Howard, M.J., Hatahet, F.S., Salo, K.E., Kauppila, A., Kellokumpu, S., and Ruddock, L.W. (2006) *ERp27, a new non-catalytic endoplasmic reticulum-located human protein disulfide isomerase family member, interacts with ERp57.* J Biol Chem. 281 (44), p. 33727-38.

324. Barak, N.N., Neumann, P., Sevvana, M., Schutkowski, M., Naumann, K., Malesevic, M., Reichardt, H., Fischer, G., Stubbs, M.T., and Ferrari, D.M. (2009) *Crystal structure and functional analysis of the protein disulfide isomerase-related protein ERp29.* J Mol Biol. 385 (5), p. 1630-42.

325. Tompa, P. (2002) *Intrinsically unstructured proteins.* Trends Biochem Sci. 27 (10), p. 527-33.

326. Dyson, H.J. and Wright, P.E. (2005) *Intrinsically unstructured proteins and their functions.* Nat Rev Mol Cell Biol. 6 (3), p. 197-208.

327. Vucetic, S., Brown, C.J., Dunker, A.K., and Obradovic, Z. (2003) *Flavors of protein disorder.* Proteins. 52 (4), p. 573-84.

328. Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F., and Jones, D.T. (2004) *Prediction and functional analysis of native disorder in proteins from the three kingdoms of life.* J Mol Biol. 337 (3), p. 635-45.

329. Iakoucheva, L.M., Radivojac, P., Brown, C.J., O'Connor, T.R., Sikes, J.G., Obradovic, Z., and Dunker, A.K. (2004) *The importance of intrinsic disorder for protein phosphorylation.* Nucleic Acids Res. 32 (3), p. 1037-49.

330. Namba, K. (2001) *Roles of partly unfolded conformations in macromolecular self-assembly.* Genes Cells. 6 (1), p. 1-12.

331. Fuxreiter, M., Tompa, P., Simon, I., Uversky, V.N., Hansen, J.C., and Asturias, F.J. (2008) *Malleable machines take shape in eukaryotic transcriptional regulation.* Nat Chem Biol. 4 (12), p. 728-37.

332. Babu, M.M., van der Lee, R., de Groot, N.S., and Gsponer, J. (2011) *Intrinsically disordered proteins: regulation and disease.* Curr Opin Struct Biol. 21 (3), p. 432-40.

333. Shoemaker, B.A., Portman, J.J., and Wolynes, P.G. (2000) *Speeding molecular recognition by using the folding funnel: the fly-casting mechanism.* Proc Natl Acad Sci U S A. 97 (16), p. 8868-73.

334. Tompa, P. and Csermely, P. (2004) *The role of structural disorder in the function of RNA and protein chaperones.* FASEB J. 18 (11), p. 1169-75.

335. Fu, X., Li, W., Mao, Q., and Chang, Z. (2003) *Disulfide bonds convert small heat shock protein Hsp16.3 from a chaperone to a non-chaperone: implications for the*

evolution of cysteine in molecular chaperones. Biochem Biophys Res Commun. 308 (3), p. 627-35.

336. Bhattacharyya, J. and Das, K.P. (1999) *Molecular chaperone-like properties of an unfolded protein, alpha(s)-casein.* J Biol Chem. 274 (22), p. 15505-9.

337. Smulders, R., Carver, J.A., Lindner, R.A., van Boekel, M.A., Bloemendal, H., and de Jong, W.W. (1996) *Immobilization of the C-terminal extension of bovine alphaA-crystallin reduces chaperone-like activity.* J Biol Chem. 271 (46), p. 29060-6.

338. Vacic, V., Markwick, P.R., Oldfield, C.J., Zhao, X., Haynes, C., Uversky, V.N., and Iakoucheva, L.M. (2012) *Disease-associated mutations disrupt functionally important regions of intrinsic protein disorder.* PLoS Comput Biol. 8 (10), p. e1002709.

339. Barlowe, C.K. and Miller, E.A. (2013) *Secretory protein biogenesis and traffic in the early secretory pathway.* Genetics. 193 (2), p. 383-410.

340. Ellgaard, L., Molinari, M., and Helenius, A. (1999) *Setting the standards: quality control in the secretory pathway.* Science. 286 (5446), p. 1882-8.

341. Semenza, J.C., Hardwick, K.G., Dean, N., and Pelham, H.R. (1990) *ERD2, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway.* Cell. 61 (7), p. 1349-57.

342. Raykhel, I., Alanen, H., Salo, K., Jurvansuu, J., Nguyen, V.D., Latva-Ranta, M., and Ruddock, L. (2007) *A molecular specificity code for the three mammalian KDEL receptors.* J Cell Biol. 179 (6), p. 1193-204.

343. Cosson, P. and Letourneur, F. (1994) *Coatomer interaction with di-lysine endoplasmic reticulum retention motifs.* Science. 263 (5153), p. 1629-31.

344. Wilson, D.W., Lewis, M.J., and Pelham, H.R. (1993) *pH-dependent binding of KDEL to its receptor in vitro.* J Biol Chem. 268 (10), p. 7465-8.

345. Ellgaard, L. and Helenius, A. (2003) *Quality control in the endoplasmic reticulum.* Nat Rev Mol Cell Biol. 4 (3), p. 181-91.

346. An, H.J., Gip, P., Kim, J., Wu, S., Park, K.W., McVaugh, C.T., Schaffer, D.V., Bertozzi, C.R., and Lebrilla, C.B. (2012) *Extensive determination of glycan heterogeneity reveals an unusual abundance of high mannose glycans in enriched plasma membranes of human embryonic stem cells.* Mol Cell Proteomics. 11 (4), p. M111 010660.

347. Stanley, P. (2011) *Golgi glycosylation.* Cold Spring Harb Perspect Biol. 3 (4).

348. Moremen, K.W., Trimble, R.B., and Herscovics, A. (1994) *Glycosidases of the asparagine-linked oligosaccharide processing pathway.* Glycobiology. 4 (2), p. 113-25.

349. Roth, J. (2002) *Protein N-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions.* Chem Rev. 102 (2), p. 285-303.

350. Rabouille, C., Hui, N., Hunte, F., Kieckbusch, R., Berger, E.G., Warren, G., and Nilsson, T. (1995) *Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides.* J Cell Sci. 108 (Pt 4), p. 1617-27.

351. Tu, L. and Banfield, D.K. (2010) *Localization of Golgi-resident glycosyltransferases.* Cell Mol Life Sci. 67 (1), p. 29-41.

352. Moremen, K.W., Tiemeyer, M., and Nairn, A.V. (2012) *Vertebrate protein glycosylation: diversity, synthesis and function.* Nat Rev Mol Cell Biol. 13 (7), p. 448-62.

353. Nilsson, T., Pypaert, M., Hoe, M.H., Slusarewicz, P., Berger, E.G., and Warren, G. (1993) *Overlapping distribution of two glycosyltransferases in the Golgi apparatus of HeLa cells.* J Cell Biol. 120 (1), p. 5-13.

354. Velasco, A., Hendricks, L., Moremen, K.W., Tulsiani, D.R., Touster, O., and Farquhar, M.G. (1993) *Cell type-dependent variations in the subcellular distribution of alpha-mannosidase I and II.* J Cell Biol. 122 (1), p. 39-51.

355. Reynders, E., Foulquier, F., Annaert, W., and Matthijs, G. (2011) *How Golgi glycosylation meets and needs trafficking: the case of the COG complex.* Glycobiology. 21 (7), p. 853-63.

356. Paroutis, P., Touret, N., and Grinstein, S. (2004) *The pH of the secretory pathway: measurement, determinants, and regulation.* Physiology (Bethesda). 19, p. 207-15.

357. Maeda, Y. and Kinoshita, T. (2010) *The acidic environment of the Golgi is critical for glycosylation and transport.* Methods Enzymol. 480, p. 495-510.

358. Axelsson, M.A., Karlsson, N.G., Steel, D.M., Ouwendijk, J., Nilsson, T., and Hansson, G.C. (2001) *Neutralization of pH in the Golgi apparatus causes redistribution of glycosyltransferases and changes in the O-glycosylation of mucins.* Glycobiology. 11 (8), p. 633-44.

359. Bennett, E.P., Mandel, U., Clausen, H., Gerken, T.A., Fritz, T.A., and Tabak, L.A. (2012) *Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family.* Glycobiology. 22 (6), p. 736-56.

360. Bergstrom, K.S. and Xia, L. (2013) *Mucin-type O-glycans and their roles in intestinal homeostasis.* Glycobiology. 23 (9), p. 1026-37.

361. Schweizer, A., Clausen, H., van Meer, G., and Hauri, H.P. (1994) *Localization of O-glycan initiation, sphingomyelin synthesis, and glucosylceramide synthesis in Vero cells with respect to the endoplasmic reticulum-Golgi intermediate compartment.* J Biol Chem. 269 (6), p. 4035-41.

362. Gill, D.J., Chia, J., Senewiratne, J., and Bard, F. (2010) *Regulation of O-glycosylation through Golgi-to-ER relocation of initiation enzymes.* J Cell Biol. 189 (5), p. 843-58.

363. Gill, D.J., Tham, K.M., Chia, J., Wang, S.C., Steentoft, C., Clausen, H., Bard-Chapeau, E.A., and Bard, F.A. (2013) *Initiation of GalNAc-type O-glycosylation in the endoplasmic reticulum promotes cancer cell invasiveness.* Proc Natl Acad Sci U S A. 110 (34), p. E3152-61.

364. Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., and Brunak, S. (2004) *Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence.* Proteomics. 4 (6), p. 1633-49.

365. De Matteis, M.A. and Luini, A. (2008) *Exiting the Golgi complex.* Nat Rev Mol Cell Biol. 9 (4), p. 273-84.

366. Heider, M.R. and Munson, M. (2012) *Exorcising the exocyst complex.* Traffic. 13 (7), p. 898-907.

367. Sugihara, K., Asano, S., Tanaka, K., Iwamatsu, A., Okawa, K., and Ohta, Y. (2002) *The exocyst complex binds the small GTPase RalA to mediate filopodia formation.* Nat Cell Biol. 4 (1), p. 73-8.

368. Rosse, C., Hatzoglou, A., Parrini, M.C., White, M.A., Chavrier, P., and Camonis, J. (2006) *RalB mobilizes the exocyst to drive cell migration.* Mol Cell Biol. 26 (2), p. 727-34.

369. Zuo, X., Zhang, J., Zhang, Y., Hsu, S.C., Zhou, D., and Guo, W. (2006) *Exo70 interacts with the Arp2/3 complex and regulates cell migration.* Nat Cell Biol. 8 (12), p. 1383-8.

370. Grindstaff, K.K., Yeaman, C., Anandasabapathy, N., Hsu, S.C., Rodriguez-Boulan, E., Scheller, R.H., and Nelson, W.J. (1998) *Sec6/8 complex is recruited to cell-cell contacts and specifies transport vesicle delivery to the basal-lateral membrane in epithelial cells.* Cell. 93 (5), p. 731-40.

371. Yeaman, C., Grindstaff, K.K., and Nelson, W.J. (2004) *Mechanism of recruiting Sec6/8 (exocyst) complex to the apical junctional complex during polarization of epithelial cells.* J Cell Sci. 117 (Pt 4), p. 559-70.

372. Langevin, J., Morgan, M.J., Sibarita, J.B., Aresta, S., Murthy, M., Schwarz, T., Camonis, J., and Bellaiche, Y. (2005) *Drosophila exocyst components Sec5, Sec6, and Sec15 regulate DE-Cadherin trafficking from recycling endosomes to the plasma membrane.* Dev Cell. 9 (3), p. 365-76.

373. Spiczka, K.S. and Yeaman, C. (2008) *Ral-regulated interaction between Sec5 and paxillin targets Exocyst to focal complexes during cell migration.* J Cell Sci. 121 (Pt 17), p. 2880-91.

374. Thapa, N., Sun, Y., Schramp, M., Choi, S., Ling, K., and Anderson, R.A. (2012) *Phosphoinositide signaling regulates the exocyst complex and polarized integrin trafficking in directionally migrating cells.* Dev Cell. 22 (1), p. 116-30.

375. Thompson, D.A. and 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.* Biochem Biophys Res Commun. 251 (1), p. 111-6.

376. Komiya, T., Tanigawa, Y., and Hirohashi, S. (1999) *Cloning of the gene gob-4, which is expressed in intestinal goblet cells in mice.* Biochim Biophys Acta. 1444 (3), p. 434-8.

377. Aberger, F., Weidinger, G., Grunz, H., and Richter, K. (1998) *Anterior specification of embryonic ectoderm: the role of the Xenopus cement gland-specific gene XAG-2.* Mech Dev. 72 (1-2), p. 115-30.

378. Persson, S., Rosenquist, M., Knoblach, B., Khosravi-Far, R., Sommarin, M., and Michalak, M. (2005) *Diversity of the protein disulfide isomerase family: identification of breast tumor induced Hag2 and Hag3 as novel members of the protein family.* Mol Phylogenet Evol. 36 (3), p. 734-40.

379. Rowe, M.L., Ruddock, L.W., Kelly, G., Schmidt, J.M., Williamson, R.A., and Howard, M.J. (2009) *Solution structure and dynamics of ERp18, a small endoplasmic reticulum resident oxidoreductase.* Biochemistry. 48 (21), p. 4596-606.

380. Alanen, H.I., Williamson, R.A., Howard, M.J., Lappi, A.K., Jantti, H.P., Rautio, S.M., Kellokumpu, S., and Ruddock, L.W. (2003) *Functional characterization of ERp18, a new endoplasmic reticulum-located thioredoxin superfamily member.* J Biol Chem. 278 (31), p. 28912-20.

381. Adam, G.C., Sorensen, E.J., and Cravatt, B.F. (2002) *Trifunctional chemical probes for the consolidated detection and identification of enzyme activities from complex proteomes.* Mol Cell Proteomics. 1 (10), p. 828-35.

382. Adam, G.C., Sorensen, E.J., and Cravatt, B.F. (2002) *Proteomic profiling of mechanistically distinct enzyme classes using a common chemotype.* Nat Biotechnol. 20 (8), p. 805-9.

383. Adam, P.J., Boyd, R., Tyson, K.L., Fletcher, G.C., Stamps, A., Hudson, L., Poyser, H.R., Redpath, N., Griffiths, M., Steers, G., Harris, A.L., Patel, S., Berry, J., Loader, J.A., Townsend, R.R., Daviet, L., Legrain, P., Parekh, R., and Terrett, J.A. (2003) *Comprehensive proteomic analysis of breast cancer cell membranes reveals unique proteins with potential roles in clinical cancer.* J Biol Chem. 278 (8), p. 6482-9.

384. Zheng, W., Rosenstiel, P., Huse, K., Sina, C., Valentonyte, R., Mah, N., Zeitlmann, L., Grosse, J., Ruf, N., Nurnberg, P., Costello, C.M., Onnie, C., Mathew, C., Platzer, M., Schreiber, S., and Hampe, J. (2006) *Evaluation of AGR2 and AGR3 as candidate genes for inflammatory bowel disease.* Genes Immun. 7 (1), p. 11-8.

385. Gray, T.A., MacLaine, N.J., Michie, C.O., Bouchalova, P., Murray, E., Howie, J., Hrstka, R., Maslon, M.M., Nenutil, R., Vojtesek, B., Langdon, S., Hayward, L., Gourley, C., and Hupp, T.R. (2012) *Anterior Gradient-3: a novel biomarker for ovarian cancer that mediates cisplatin resistance in xenograft models.* J Immunol Methods. 378 (1-2), p. 20-32.

386. King, E.R., Tung, C.S., Tsang, Y.T., Zu, Z., Lok, G.T., Deavers, M.T., Malpica, A., Wolf, J.K., Lu, K.H., Birrer, M.J., Mok, S.C., Gershenson, D.M., and Wong, K.K. (2011) *The anterior gradient homolog 3 (AGR3) gene is associated with differentiation and survival in ovarian cancer.* Am J Surg Pathol. 35 (6), p. 904-12.

387. Petek, E., Windpassinger, C., Egger, H., Kroisel, P.M., and Wagner, K. (2000) *Localization of the human anterior gradient-2 gene (AGR2) to chromosome band 7p21.3 by radiation hybrid mapping and fluorescencein situ hybridisation.* Cytogenet Cell Genet. 89 (3-4), p. 141-2.

388. Zhang, J.S., Gong, A., Cheville, J.C., Smith, D.I., and Young, C.Y. (2005) *AGR2, an androgen-inducible secretory protein overexpressed in prostate cancer.* Genes Chromosomes Cancer. 43 (3), p. 249-59.

389. Fletcher, G.C., Patel, S., Tyson, K., Adam, P.J., Schenker, M., Loader, J.A., Daviet, L., Legrain, P., Parekh, R., Harris, A.L., and Terrett, J.A. (2003) *hAG-2 and hAG-3, human homologues of genes involved in differentiation, are associated with oestrogen receptor-positive breast tumours and interact with metastasis gene C4.4a and dystroglycan.* Br J Cancer. 88 (4), p. 579-85.

390. Wilson, C.L., Sims, A.H., Howell, A., Miller, C.J., and Clarke, R.B. (2006) *Effects of oestrogen on gene expression in epithelium and stroma of normal human breast tissue.* Endocr Relat Cancer. 13 (2), p. 617-28.

391. Bu, H., Schweiger, M.R., Manke, T., Wunderlich, A., Timmermann, B., Kerick, M., Pasqualini, L., Shehu, E., Fuchsberger, C., Cato, A.C., and Klocker, H. (2013) *Anterior gradient 2 and 3--two prototype androgen-responsive genes transcriptionally upregulated by androgens and by oestrogens in prostate cancer cells.* FEBS J. 280 (5), p. 1249-66.

392. Li, S., Wang, Y., Zhang, Y., Lu, M.M., DeMayo, F.J., Dekker, J.D., Tucker, P.W., and Morrisey, E.E. (2012) *Foxp1/4 control epithelial cell fate during lung development and regeneration through regulation of anterior gradient 2.* Development. 139 (14), p. 2500-9.

393. Wright, T.M., Wardell, S.E., Jasper, J.S., Stice, J.P., Safi, R., Nelson, E.R., and McDonnell, D.P. (2014) *Delineation of a FOXA1/ERalpha/AGR2 regulatory loop that is dysregulated in endocrine therapy-resistant breast cancer.* Mol Cancer Res.

394. Wang, D.H., Tiwari, A., Kim, M.E., Clemons, N.J., Regmi, N.L., Hodges, W.A., Berman, D.M., Montgomery, E.A., Watkins, D.N., Zhang, X., Zhang, Q., Jie, C., Spechler, S.J., and Souza, R.F. (2014) *Hedgehog signaling regulates FOXA2 in esophageal embryogenesis and Barrett's metaplasia.* J Clin Invest. 124 (9), p. 3767-80.

395. Friedman, J.R. and Kaestner, K.H. (2006) *The Foxa family of transcription factors in development and metabolism.* Cell Mol Life Sci. 63 (19-20), p. 2317-28.

396. Weinstein, D.C., Ruiz i Altaba, A., Chen, W.S., Hoodless, P., Prezioso, V.R., Jessell, T.M., and Darnell, J.E., Jr. (1994) *The winged-helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo.* Cell. 78 (4), p. 575-88.

397. Shih, D.Q., Navas, M.A., Kuwajima, S., Duncan, S.A., and Stoffel, M. (1999) *Impaired glucose homeostasis and neonatal mortality in hepatocyte nuclear factor 3alpha-deficient mice.* Proc Natl Acad Sci U S A. 96 (18), p. 10152-7.

398. Lu, M.M., Li, S., Yang, H., and Morrisey, E.E. (2002) *Foxp4: a novel member of the Foxp subfamily of winged-helix genes co-expressed with Foxp1 and Foxp2 in pulmonary and gut tissues.* Mech Dev. 119 Suppl 1, p. S197-202.

399. Shu, W., Yang, H., Zhang, L., Lu, M.M., and Morrisey, E.E. (2001) *Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lung and act as transcriptional repressors.* J Biol Chem. 276 (29), p. 27488-97.

400. Shu, W., Lu, M.M., Zhang, Y., Tucker, P.W., Zhou, D., and Morrisey, E.E. (2007) *Foxp2 and Foxp1 cooperatively regulate lung and esophagus development.* Development. 134 (10), p. 1991-2000.

401. Wang, B., Weidenfeld, J., Lu, M.M., Maika, S., Kuziel, W.A., Morrisey, E.E., and Tucker, P.W. (2004) *Foxp1 regulates cardiac outflow tract, endocardial cushion morphogenesis and myocyte proliferation and maturation.* Development. 131 (18), p. 4477-87.

402. Li, S., Weidenfeld, J., and Morrisey, E.E. (2004) *Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions.* Mol Cell Biol. 24 (2), p. 809-22.

403. Lepreux, S., Bioulac-Sage, P., and Chevet, E. (2011) *Differential expression of the anterior gradient protein-2 is a conserved feature during morphogenesis and carcinogenesis of the biliary tree.* Liver Int. 31 (3), p. 322-8.

404. Shigekawa, T., Ijichi, N., Ikeda, K., Horie-Inoue, K., Shimizu, C., Saji, S., Aogi, K., Tsuda, H., Osaki, A., Saeki, T., and Inoue, S. (2011) *FOXP1, an estrogen-inducible transcription factor, modulates cell proliferation in breast cancer cells and 5-year recurrence-free survival of patients with tamoxifen-treated breast cancer.* Horm Cancer. 2 (5), p. 286-97.

405. Hurtado, A., Holmes, K.A., Ross-Innes, C.S., Schmidt, D., and Carroll, J.S. (2011) *FOXA1 is a key determinant of estrogen receptor function and endocrine response.* Nat Genet. 43 (1), p. 27-33.

406. Tang, C.H., Lai, Y.R., Chen, Y.C., Li, C.H., Lu, Y.F., Chen, H.Y., Lien, H.W., Yang, C.H., Huang, C.J., Wang, C.Y., Kao, C.F., and Hwang, S.P. (2014) *Expression of zebrafish anterior gradient 2 in the semicircular canals and supporting cells of otic vesicle sensory patches is regulated by Sox10.* Biochim Biophys Acta. 1839 (6), p. 425-37.

407. Chevet, E., Fessart, D., Delom, F., Mulot, A., Vojtesek, B., Hrstka, R., Murray, E., Gray, T., and Hupp, T. (2013) *Emerging roles for the pro-oncogenic anterior gradient-2 in cancer development.* Oncogene. 32 (20), p. 2499-509.

408. Zhao, F., Edwards, R., Dizon, D., Afrasiabi, K., Mastroianni, J.R., Geyfman, M., Ouellette, A.J., Andersen, B., and Lipkin, S.M. (2010) *Disruption of Paneth and goblet cell homeostasis and increased endoplasmic reticulum stress in Agr2-/- mice.* Dev Biol. 338 (2), p. 270-9.

409. Gupta, A., Wodziak, D., Tun, M., Bouley, D.M., and Lowe, A.W. (2013) *Loss of anterior gradient 2 (Agr2) expression results in hyperplasia and defective lineage maturation in the murine stomach.* J Biol Chem. 288 (6), p. 4321-33.

410. Seymour, P.A., Freude, K.K., Tran, M.N., Mayes, E.E., Jensen, J., Kist, R., Scherer, G., and Sander, M. (2007) *SOX9 is required for maintenance of the pancreatic progenitor cell pool.* Proc Natl Acad Sci U S A. 104 (6), p. 1865-70.

411. Krig, S.R., Jin, V.X., Bieda, M.C., O'Geen, H., Yaswen, P., Green, R., and Farnham, P.J. (2007) *Identification of genes directly regulated by the oncogene ZNF217 using chromatin immunoprecipitation (ChIP)-chip assays.* J Biol Chem. 282 (13), p. 9703-12.

412. Norris, A.M., Gore, A., Balboni, A., Young, A., Longnecker, D.S., and Korc, M. (2013) *AGR2 is a SMAD4-suppressible gene that modulates MUC1 levels and promotes the initiation and progression of pancreatic intraepithelial neoplasia.* Oncogene. 32 (33), p. 3867-76.

413. Zweitzig, D.R., Smirnov, D.A., Connelly, M.C., Terstappen, L.W., O'Hara, S.M., and Moran, E. (2007) *Physiological stress induces the metastasis marker AGR2 in breast cancer cells.* Mol Cell Biochem. 306 (1-2), p. 255-60.

414. Hong, X.Y., Wang, J., and Li, Z. (2013) *AGR2 expression is regulated by HIF-1 and contributes to growth and angiogenesis of glioblastoma.* Cell Biochem Biophys. 67 (3), p. 1487-95.

415. Di Maro, G., Salerno, P., Unger, K., Orlandella, F.M., Monaco, M., Chiappetta, G., Thomas, G., Oczko-Wojciechowska, M., Masullo, M., Jarzab, B., Santoro, M., and Salvatore, G. (2014) *Anterior gradient protein 2 promotes survival, migration and invasion of papillary thyroid carcinoma cells.* Mol Cancer. 13, p. 160.

416. Ryu, J., Park, S.G., Lee, P.Y., Cho, S., Lee do, H., Kim, G.H., Kim, J.H., and Park, B.C. (2013) *Dimerization of pro-oncogenic protein Anterior Gradient 2 is required for the interaction with BiP/GRP78.* Biochem Biophys Res Commun. 430 (2), p. 610-5.

417. Higa, A., Mulot, A., Delom, F., Bouchecareilh, M., Nguyen, D.T., Boismenu, D., Wise, M.J., and Chevet, E. (2011) *Role of pro-oncogenic protein disulfide isomerase (PDI) family member anterior gradient 2 (AGR2) in the control of endoplasmic reticulum homeostasis.* J Biol Chem. 286 (52), p. 44855-68.

418. Schroeder, B.W., Verhaeghe, C., Park, S.W., Nguyenvu, L.T., Huang, X., Zhen, G., and Erle, D.J. (2012) *AGR2 is induced in asthma and promotes allergen-induced mucin overproduction.* Am J Respir Cell Mol Biol. 47 (2), p. 178-85.

419. Groome, M., Lindsay, J., Ross, P.E., Cotton, J.P., Hupp, T.R., and Dillon, J.F. (2008) *Use of oesophageal stress response proteins as potential biomarkers in the screening for Barrett's oesophagus.* Eur J Gastroenterol Hepatol. 20 (10), p. 961-5.

420. Chakrabarti, A., Chen, A.W., and Varner, J.D. (2011) *A review of the mammalian unfolded protein response.* Biotechnol Bioeng. 108 (12), p. 2777-93.

421. Patel, P., Clarke, C., Barraclough, D.L., Jowitt, T.A., Rudland, P.S., Barraclough, R., and Lian, L.Y. (2013) *Metastasis-promoting anterior gradient 2 protein has a dimeric thioredoxin fold structure and a role in cell adhesion.* J Mol Biol. 425 (5), p. 929-43.

422. Park, S.W., Zhen, G., Verhaeghe, C., Nakagami, Y., Nguyenvu, L.T., Barczak, A.J., Killeen, N., and Erle, D.J. (2009) *The protein disulfide isomerase AGR2 is essential for production of intestinal mucus.* Proc Natl Acad Sci U S A. 106 (17), p. 6950-5.

423. Podolsky, D.K., Lynch-Devaney, K., Stow, J.L., Oates, P., Murgue, B., De-Beaumont, M., Sands, B.E., and Mahida, Y.R. (1993) *Identification of human intestinal trefoil factor. Goblet cell-specific expression of a peptide targeted for apical secretion.* J Biol Chem. 268 (16), p. 12230.

424. Chen, Y.C., Lu, Y.F., Li, I.C., and Hwang, S.P. (2012) *Zebrafish Agr2 is required for terminal differentiation of intestinal goblet cells.* PLoS One. 7 (4), p. e34408.

425. Chang, J., Chance, M.R., Nicholas, C., Ahmed, N., Guilmeau, S., Flandez, M., Wang, D., Byun, D.S., Nasser, S., Albanese, J.M., Corner, G.A., Heerdt, B.G., Wilson, A.J., Augenlicht, L.H., and Mariadason, J.M. (2008) *Proteomic changes during intestinal cell maturation in vivo.* J Proteomics. 71 (5), p. 530-46.

426. Bergstrom, J.H., Berg, K.A., Rodriguez-Pineiro, A.M., Stecher, B., Johansson, M.E., and Hansson, G.C. (2014) *AGR2, an endoplasmic reticulum protein, is secreted into the gastrointestinal mucus.* PLoS One. 9 (8), p. e104186.

427. Zhou, M., Chen, H.L., Cheng, S., Mei, L., Zhang, H.L., Xie, M., Xiong, W.N., and Xu, Y.J. (2013) *Effect of dexamethasone on expression of AGR2 protein in asthmatic mice.* J Huazhong Univ Sci Technolog Med Sci. 33 (1), p. 33-6.

428. Cserni, G. (2014) *Reversed polarity of the glandular epithelial cells in micropapillary carcinoma of the large intestine and the EMA/MUC1 immunostain.* Pathology. 46 (6), p. 527-32.

429. Mather, I.H., Jack, L.J., Madara, P.J., and Johnson, V.G. (2001) *The distribution of MUC1, an apical membrane glycoprotein, in mammary epithelial cells at the resolution of the electron microscope: implications for the mechanism of milk secretion.* Cell Tissue Res. 304 (1), p. 91-101.

430. Wesseling, J., van der Valk, S.W., Vos, H.L., Sonnenberg, A., and Hilkens, J. (1995) *Episialin (MUC1) overexpression inhibits integrin-mediated cell adhesion to extracellular matrix components.* J Cell Biol. 129 (1), p. 255-65.

431. Ligtenberg, M.J., Kruijshaar, L., Buijs, F., van Meijer, M., Litvinov, S.V., and Hilkens, J. (1992) *Cell-associated episialin is a complex containing two proteins derived from a common precursor.* J Biol Chem. 267 (9), p. 6171-7.

432. Singh, R. and Bandyopadhyay, D. (2007) *MUC1: a target molecule for cancer therapy.* Cancer Biol Ther. 6 (4), p. 481-6.

433. Brychtova, V., Hermanova, M., Karasek, P., Lenz, J., Selingerova, I., Vojtesek, B., Kala, Z., and Hrstka, R. (2014) *Anterior gradient 2 and mucin 4 expression mirrors tumor cell differentiation in pancreatic adenocarcinomas, but aberrant anterior gradient 2 expression predicts worse patient outcome in poorly differentiated tumors.* Pancreas. 43 (1), p. 75-81.

434. Komatsu, M., Tatum, L., Altman, N.H., Carothers Carraway, C.A., and Carraway, K.L. (2000) *Potentiation of metastasis by cell surface sialomucin complex (rat MUC4), a multifunctional anti-adhesive glycoprotein.* Int J Cancer. 87 (4), p. 480-6.

435. Singh, A.P., Moniaux, N., Chauhan, S.C., Meza, J.L., and Batra, S.K. (2004) *Inhibition of MUC4 expression suppresses pancreatic tumor cell growth and metastasis.* Cancer Res. 64 (2), p. 622-30.

436. Carraway, K.L., Theodoropoulos, G., Kozloski, G.A., and Carothers Carraway, C.A. (2009) *Muc4/MUC4 functions and regulation in cancer.* Future Oncol. 5 (10), p. 1631-40.

437. Dumartin, L., Whiteman, H.J., Weeks, M.E., Hariharan, D., Dmitrovic, B., Iacobuzio-Donahue, C.A., Brentnall, T.A., Bronner, M.P., Feakins, R.M., Timms, J.F., Brennan, C., Lemoine, N.R., and Crnogorac-Jurcevic, T. (2011) *AGR2 is a novel surface antigen that promotes the dissemination of pancreatic cancer cells through regulation of cathepsins B and D.* Cancer Res. 71 (22), p. 7091-102.

438. Gray, T.A., Alsamman, K., Murray, E., Sims, A.H., and Hupp, T.R. (2014) *Engineering a synthetic cell panel to identify signalling components reprogrammed by the cell growth regulator anterior gradient-2.* Mol Biosyst. 10 (6), p. 1409-25.

439. Gupta, A., Dong, A., and Lowe, A.W. (2012) *AGR2 gene function requires a unique endoplasmic reticulum localization motif.* J Biol Chem. 287 (7), p. 4773-82.

440. Ramachandran, V., Arumugam, T., Wang, H., and Logsdon, C.D. (2008) *Anterior gradient 2 is expressed and secreted during the development of pancreatic cancer and promotes cancer cell survival.* Cancer Res. 68 (19), p. 7811-8.

441. Johansson, M.E., Phillipson, M., Petersson, J., Velcich, A., Holm, L., and Hansson, G.C. (2008) *The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria.* Proc Natl Acad Sci U S A. 105 (39), p. 15064-9.

442. Johansson, M.E., Thomsson, K.A., and Hansson, G.C. (2009) *Proteomic analyses of the two mucus layers of the colon barrier reveal that their main component, the Muc2 mucin, is strongly bound to the Fcgbp protein.* J Proteome Res. 8 (7), p. 3549-57.

443. Innes, H.E., Liu, D., Barraclough, R., Davies, M.P., O'Neill, P.A., Platt-Higgins, A., de Silva Rudland, S., Sibson, D.R., and Rudland, P.S. (2006) *Significance of the metastasis-inducing protein AGR2 for outcome in hormonally treated breast cancer patients.* Br J Cancer. 94 (7), p. 1057-65.

444. Maresh, E.L., Mah, V., Alavi, M., Horvath, S., Bagryanova, L., Liebeskind, E.S., Knutzen, L.A., Zhou, Y., Chia, D., Liu, A.Y., and Goodglick, L. (2010) *Differential expression of anterior gradient gene AGR2 in prostate cancer.* BMC Cancer. 10, p. 680.

445. Pohler, E., Craig, A.L., Cotton, J., Lawrie, L., Dillon, J.F., Ross, P., Kernohan, N., and Hupp, T.R. (2004) *The Barrett's antigen anterior gradient-2 silences the p53 transcriptional response to DNA damage.* Mol Cell Proteomics. 3 (6), p. 534-47.

446. Lahav, J., Wijnen, E.M., Hess, O., Hamaia, S.W., Griffiths, D., Makris, M., Knight, C.G., Essex, D.W., and Farndale, R.W. (2003) *Enzymatically catalyzed disulfide exchange is required for platelet adhesion to collagen via integrin alpha2beta1.* Blood. 102 (6), p. 2085-92.

447. Jakobsen, C.G., Rasmussen, N., Laenkholm, A.V., and Ditzel, H.J. (2007) *Phage display derived human monoclonal antibodies isolated by binding to the surface of live primary breast cancer cells recognize GRP78.* Cancer Res. 67 (19), p. 9507-17.

448. Somogyi, E., Petersson, U., Hultenby, K., and Wendel, M. (2003) *Calreticulin--an endoplasmic reticulum protein with calcium-binding activity is also found in the extracellular matrix.* Matrix Biol. 22 (2), p. 179-91.

449. Kern, J., Untergasser, G., Zenzmaier, C., Sarg, B., Gastl, G., Gunsilius, E., and Steurer, M. (2009) *GRP-78 secreted by tumor cells blocks the antiangiogenic activity of bortezomib.* Blood. 114 (18), p. 3960-7.

450. Kumar, A., Godwin, J.W., Gates, P.B., Garza-Garcia, A.A., and Brockes, J.P. (2007) *Molecular basis for the nerve dependence of limb regeneration in an adult vertebrate.* Science. 318 (5851), p. 772-7.

451. da Silva, S.M., Gates, P.B., and Brockes, J.P. (2002) *The newt ortholog of CD59 is implicated in proximodistal identity during amphibian limb regeneration.* Dev Cell. 3 (4), p. 547-55.

452. Garza-Garcia, A., Harris, R., Esposito, D., Gates, P.B., and Driscoll, P.C. (2009) *Solution structure and phylogenetics of Prod1, a member of the three-finger protein superfamily implicated in salamander limb regeneration.* PLoS One. 4 (9), p. e7123.

453. Jacobsen, B. and Ploug, M. (2008) *The urokinase receptor and its structural homologue C4.4A in human cancer: expression, prognosis and pharmacological inhibition.* Curr Med Chem. 15 (25), p. 2559-73.

454. Han, R. and Campbell, K.P. (2007) *Dysferlin and muscle membrane repair.* Curr Opin Cell Biol. 19 (4), p. 409-16.

455. Tu, S.W., Bugde, A., Luby-Phelps, K., and Cobb, M.H. (2011) *WNK1 is required for mitosis and abscission.* Proc Natl Acad Sci U S A. 108 (4), p. 1385-90.

456. Verma, S., Salmans, M.L., Geyfman, M., Wang, H., Yu, Z., Lu, Z., Zhao, F., Lipkin, S.M., and Andersen, B. (2012) *The estrogen-responsive Agr2 gene regulates mammary epithelial proliferation and facilitates lobuloalveolar development.* Dev Biol. 369 (2), p. 249-60.

457. Barraclough, D.L., Platt-Higgins, A., de Silva Rudland, S., Barraclough, R., Winstanley, J., West, C.R., and Rudland, P.S. (2009) *The metastasis-associated anterior gradient 2 protein is correlated with poor survival of breast cancer patients.* Am J Pathol. 175 (5), p. 1848-57.

458. Barraclough, D.L., Sewart, S., Rudland, P.S., Shoker, B.S., Sibson, D.R., Barraclough, R., and Davies, M.P. (2010) *Microarray analysis of suppression subtracted hybridisation libraries identifies genes associated with breast cancer progression.* Cell Oncol. 32 (1-2), p. 87-99.

459. Hrstka, R., Nenutil, R., Fourtouna, A., Maslon, M.M., Naughton, C., Langdon, S., Murray, E., Larionov, A., Petrakova, K., Muller, P., Dixon, M.J., Hupp, T.R., and Vojtesek, B. (2010) *The pro-metastatic protein anterior gradient-2 predicts poor prognosis in tamoxifen-treated breast cancers.* Oncogene. 29 (34), p. 4838-47.

460. Hrstka, R., Murray, E., Brychtova, V., Fabian, P., Hupp, T.R., and Vojtesek, B. (2013) *Identification of an AKT-dependent signalling pathway that mediates tamoxifen-dependent induction of the pro-metastatic protein anterior gradient-2.* Cancer Lett. 333 (2), p. 187-93.

461. da Costa, A., Lenze, D., Hummel, M., Kohn, B., Gruber, A.D., and Klopfleisch, R. (2012) *Identification of six potential markers for the detection of circulating canine mammary tumour cells in the peripheral blood identified by microarray analysis.* J Comp Pathol. 146 (2-3), p. 143-51.

462. Wu, Z.S., Wu, Q., Ding, X.D., Wang, H.Q., Shen, Y.X., and Fang, S.Y. (2008) *[Expression of a novel metastasis-inducing protein human anterior gradient-2 (AGR2) in breast cancer and its clinical and prognostic significance].* Zhonghua Bing Li Xue Za Zhi. 37 (2), p. 109-13.

463. Hengel, S.M., Murray, E., Langdon, S., Hayward, L., O'Donoghue, J., Panchaud, A., Hupp, T., and Goodlett, D.R. (2011) *Data-independent proteomic screen identifies novel tamoxifen agonist that mediates drug resistance.* J Proteome Res. 10 (10), p. 4567-78.

464. Mackay, A., Urruticoechea, A., Dixon, J.M., Dexter, T., Fenwick, K., Ashworth, A., Drury, S., Larionov, A., Young, O., White, S., Miller, W.R., Evans, D.B., and Dowsett, M. (2007) *Molecular response to aromatase inhibitor treatment in primary breast cancer.* Breast Cancer Res. 9 (3), p. R37.

465. Wayner, E.A., Quek, S.I., Ahmad, R., Ho, M.E., Loprieno, M.A., Zhou, Y., Ellis, W.J., True, L.D., and Liu, A.Y. (2012) *Development of an ELISA to detect the secreted prostate cancer biomarker AGR2 in voided urine.* Prostate. 72 (9), p. 1023-34.

466. Kovalev, L.I., Shishkin, S.S., Khasigov, P.Z., Dzeranov, N.K., Kazachenko, A.V., Toropygin, I., and Mamykina, S.V. (2006) *[Identification of AGR2 protein, a novel potential cancer marker, using proteomics technologies].* Prikl Biokhim Mikrobiol. 42 (4), p. 480-4.

467. Zhang, Y., Forootan, S.S., Liu, D., Barraclough, R., Foster, C.S., Rudland, P.S., and Ke, Y. (2007) *Increased expression of anterior gradient-2 is significantly associated with poor survival of prostate cancer patients.* Prostate Cancer Prostatic Dis. 10 (3), p. 293-300.

468. Celma, A., Servian, P., Planas, J., Placer, J., Quilez, M.T., Arbos, M.A., de Torres, I., and Morote, J. (2014) *Clinical significance of proliferative inflammatory atrophy in prostate biopsy.* Actas Urol Esp. 38 (2), p. 122-6.

469. Kani, K., Malihi, P.D., Jiang, Y., Wang, H., Wang, Y., Ruderman, D.L., Agus, D.B., Mallick, P., and Gross, M.E. (2013) *Anterior gradient 2 (AGR2): blood-based biomarker elevated in metastatic prostate cancer associated with the neuroendocrine phenotype.* Prostate. 73 (3), p. 306-15.

470. Hu, Z., Gu, Y., Han, B., Zhang, J., Li, Z., Tian, K., Young, C.Y., and Yuan, H. (2012) *Knockdown of AGR2 induces cellular senescence in prostate cancer cells.* Carcinogenesis. 33 (6), p. 1178-86.

471. Zhang, Y., Linn, D., Liu, Z., Melamed, J., Tavora, F., Young, C.Y., Burger, A.M., and Hamburger, A.W. (2008) *EBP1, an ErbB3-binding protein, is decreased in prostate cancer and implicated in hormone resistance.* Mol Cancer Ther. 7 (10), p. 3176-86.

472. Zhang, Y., Wang, X.W., Jelovac, D., Nakanishi, T., Yu, M.H., Akinmade, D., Goloubeva, O., Ross, D.D., Brodie, A., and Hamburger, A.W. (2005) *The ErbB3-binding protein Ebp1 suppresses androgen receptor-mediated gene transcription and tumorigenesis of prostate cancer cells.* Proc Natl Acad Sci U S A. 102 (28), p. 9890-5.

473. Sung, H.Y., Choi, E.N., Lyu, D., Park, A.K., Ju, W., and Ahn, J.H. (2014) *Aberrant hypomethylation-mediated AGR2 overexpression induces an aggressive phenotype in ovarian cancer cells.* Oncol Rep. 32 (2), p. 815-20.

474. Darb-Esfahani, S., Fritzsche, F., Kristiansen, G., Weichert, W., Sehouli, J., Braicu, I., Dietel, M., and Denkert, C. (2012) *Anterior gradient protein 2 (AGR2) is an independent prognostic factor in ovarian high-grade serous carcinoma.* Virchows Arch. 461 (2), p. 109-16.

475. Edgell, T.A., Barraclough, D.L., Rajic, A., Dhulia, J., Lewis, K.J., Armes, J.E., Barraclough, R., Rudland, P.S., Rice, G.E., and Autelitano, D.J. (2010) *Increased plasma concentrations of anterior gradient 2 protein are positively associated with ovarian cancer.* Clin Sci (Lond). 118 (12), p. 717-25.

476. Rice, G.E., Edgell, T.A., and Autelitano, D.J. (2010) *Evaluation of midkine and anterior gradient 2 in a multimarker panel for the detection of ovarian cancer.* J Exp Clin Cancer Res. 29, p. 62.

477. Konduri, S. and Schwarz, R.E. (2007) *Estrogen receptor beta/alpha ratio predicts response of pancreatic cancer cells to estrogens and phytoestrogens.* J Surg Res. 140 (1), p. 55-66.

478. Riener, M.O., Pilarsky, C., Gerhardt, J., Grutzmann, R., Fritzsche, F.R., Bahra, M., Weichert, W., and Kristiansen, G. (2009) *Prognostic significance of AGR2 in pancreatic ductal adenocarcinoma.* Histol Histopathol. 24 (9), p. 1121-8.

479. Makawita, S., Smith, C., Batruch, I., Zheng, Y., Ruckert, F., Grutzmann, R., Pilarsky, C., Gallinger, S., and Diamandis, E.P. (2011) *Integrated proteomic profiling of cell line conditioned media and pancreatic juice for the identification of pancreatic cancer biomarkers.* Mol Cell Proteomics. 10 (10), p. M111 008599.

480. Chen, R., Pan, S., Duan, X., Nelson, B.H., Sahota, R.A., de Rham, S., Kozarek, R.A., McIntosh, M., and Brentnall, T.A. (2010) *Elevated level of anterior gradient-2 in*

pancreatic juice from patients with pre-malignant pancreatic neoplasia. Mol Cancer. 9, p. 149.

481. Wang, Z., Hao, Y., and Lowe, A.W. (2008) *The adenocarcinoma-associated antigen, AGR2, promotes tumor growth, cell migration, and cellular transformation.* Cancer Res. 68 (2), p. 492-7.

482. Vanderlaag, K.E., Hudak, S., Bald, L., Fayadat-Dilman, L., Sathe, M., Grein, J., and Janatpour, M.J. (2010) *Anterior gradient-2 plays a critical role in breast cancer cell growth and survival by modulating cyclin D1, estrogen receptor-alpha and survivin.* Breast Cancer Res. 12 (3), p. R32.

483. Wang, X., Ouyang, H., Yamamoto, Y., Kumar, P.A., Wei, T.S., Dagher, R., Vincent, M., Lu, X., Bellizzi, A.M., Ho, K.Y., Crum, C.P., Xian, W., and McKeon, F. (2011) *Residual embryonic cells as precursors of a Barrett's-like metaplasia.* Cell. 145 (7), p. 1023-35.

484. Pizzi, M., Fassan, M., Realdon, S., Balistreri, M., Battaglia, G., Giacometti, C., Zaninotto, G., Zagonel, V., De Boni, M., and Rugge, M. (2012) *Anterior gradient 2 profiling in Barrett columnar epithelia and adenocarcinoma.* Hum Pathol. 43 (11), p. 1839-44.

485. Nancarrow, D.J., Clouston, A.D., Smithers, B.M., Gotley, D.C., Drew, P.A., Watson, D.I., Tyagi, S., Hayward, N.K., and Whiteman, D.C. (2011) *Whole genome expression array profiling highlights differences in mucosal defense genes in Barrett's esophagus and esophageal adenocarcinoma.* PLoS One. 6 (7), p. e22513.

486. Varghese, S., Lao-Sirieix, P., and Fitzgerald, R.C. (2012) *Identification and clinical implementation of biomarkers for Barrett's esophagus.* Gastroenterology. 142 (3), p. 435-441 e2.

487. Pizzi, M., Fassan, M., Balistreri, M., Galligioni, A., Rea, F., and Rugge, M. (2012) *Anterior gradient 2 overexpression in lung adenocarcinoma.* Appl Immunohistochem Mol Morphol. 20 (1), p. 31-6.

488. Chung, K., Nishiyama, N., Yamano, S., Komatsu, H., Hanada, S., Wei, M., Wanibuchi, H., Suehiro, S., and Kakehashi, A. (2011) *Serum AGR2 as an early diagnostic and postoperative prognostic biomarker of human lung adenocarcinoma.* Cancer Biomark. 10 (2), p. 101-7.

489. Huang, J., Wang, L., Jiang, M., Chen, Q., Jiang, Z., and Feng, H. (2014) *AGR2-mediated lung adenocarcinoma metastasis novel mechanism network through repression with interferon coupling cytoskeleton to steroid metabolism-dependent humoral immune response.* Cell Immunol. 290 (1), p. 102-6.

490. Chung, K., Nishiyama, N., Wanibuchi, H., Yamano, S., Hanada, S., Wei, M., Suehiro, S., and Kakehashi, A. (2012) *AGR2 as a potential biomarker of human lung adenocarcinoma.* Osaka City Med J. 58 (1), p. 13-24.

491. Fritzsche, F.R., Dahl, E., Dankof, A., Burkhardt, M., Pahl, S., Petersen, I., Dietel, M., and Kristiansen, G. (2007) *Expression of AGR2 in non small cell lung cancer.* Histol Histopathol. 22 (7), p. 703-8.

492. Vivekanandan, P., Micchelli, S.T., and Torbenson, M. (2009) *Anterior gradient-2 is overexpressed by fibrolamellar carcinomas.* Hum Pathol. 40 (3), p. 293-9.

493. Riener, M.O., Thiesler, T., Hellerbrand, C., Amann, T., Cathomas, G., Fritzsche, F.R., Dahl, E., Bahra, M., Weichert, W., Terracciano, L., and Kristiansen, G. (2014) *Loss of anterior gradient-2 expression is an independent prognostic factor in colorectal carcinomas.* Eur J Cancer. 50 (10), p. 1722-30.

494. Kim, H.S., Kang, S.H., Park, C.H., Yang, W.I., Jeung, H.C., Chung, H.C., Roh, J.K., Ahn, J.B., Kim, N.K., Min, B.S., and Rha, S.Y. (2011) *Genome-wide molecular characterization of mucinous colorectal adenocarcinoma using cDNA microarray analysis.* Oncol Rep. 25 (3), p. 717-27.

495. Valladares-Ayerbes, M., Blanco-Calvo, M., Reboredo, M., Lorenzo-Patino, M.J., Iglesias-Diaz, P., Haz, M., Diaz-Prado, S., Medina, V., Santamarina, I., Pertega, S., Figueroa, A., and Anton-Aparicio, L.M. (2012) *Evaluation of the Adenocarcinoma-Associated Gene AGR2 and the Intestinal Stem Cell Marker LGR5 as Biomarkers in Colorectal Cancer.* Int J Mol Sci. 13 (4), p. 4367-87.

496. Valladares-Ayerbes, M., Diaz-Prado, S., Reboredo, M., Medina, V., Iglesias-Diaz, P., Lorenzo-Patino, M.J., Campelo, R.G., Haz, M., Santamarina, I., and Anton-Aparicio, L.M. (2008) *Bioinformatics approach to mRNA markers discovery for detection of circulating tumor cells in patients with gastrointestinal cancer.* Cancer Detect Prev. 32 (3), p. 236-50.

497. Fritzsche, F.R., Dahl, E., Pahl, S., Burkhardt, M., Luo, J., Mayordomo, E., Gansukh, T., Dankof, A., Knuechel, R., Denkert, C., Winzer, K.J., Dietel, M., and Kristiansen, G. (2006) *Prognostic relevance of AGR2 expression in breast cancer.* Clin Cancer Res. 12 (6), p. 1728-34.

498. Ho, M.E., Quek, S.I., True, L.D., Morrissey, C., Corey, E., Vessella, R.L., Dumpit, R., Nelson, P.S., Maresh, E.L., Mah, V., Alavi, M., Kim, S.R., Bagryanova, L., Horvath, S., Chia, D., Goodglick, L., and Liu, A.Y. (2013) *Prostate cancer cell phenotypes based on AGR2 and CD10 expression.* Mod Pathol. 26 (6), p. 849-59.

499. Putti, T.C., El-Rehim, D.M., Rakha, E.A., Paish, C.E., Lee, A.H., Pinder, S.E., and Ellis, I.O. (2005) *Estrogen receptor-negative breast carcinomas: a review of morphology and immunophenotypical analysis.* Mod Pathol. 18 (1), p. 26-35.

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

501. Kim, S.J., Kim, D.H., Kang, D., and Kim, J.H. (2014) *Expression of anterior gradient 2 is decreased with the progression of human biliary tract cancer.* Tohoku J Exp Med. 234 (1), p. 83-8.

502. Shi, T., Gao, Y., Quek, S.I., Fillmore, T.L., Nicora, C.D., Su, D., Zhao, R., Kagan, J., Srivastava, S., Rodland, K.D., Liu, T., Smith, R.D., Chan, D.W., Camp, D.G., 2nd, Liu, A.Y., and Qian, W.J. (2014) *A highly sensitive targeted mass spectrometric assay for quantification of AGR2 protein in human urine and serum.* J Proteome Res. 13 (2), p. 875-82.

503. Bu, H., Bormann, S., Schafer, G., Horninger, W., Massoner, P., Neeb, A., Lakshmanan, V.K., Maddalo, D., Nestl, A., Sultmann, H., Cato, A.C., and Klocker, 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.* Prostate. 71 (6), p. 575-87.

504. DiMaio, M.A., Kwok, S., Montgomery, K.D., Lowe, A.W., and Pai, R.K. (2012) *Immunohistochemical panel for distinguishing esophageal adenocarcinoma from squamous cell carcinoma: a combination of p63, cytokeratin 5/6, MUC5AC, and anterior gradient homolog 2 allows optimal subtyping.* Hum Pathol. 43 (11), p. 1799-807.

505. Macias, H. and Hinck, L. (2012) *Mammary gland development.* Wiley Interdiscip Rev Dev Biol. 1 (4), p. 533-57.

506. Dong, A., Gupta, A., Pai, R.K., Tun, M., and Lowe, A.W. (2011) *The Human Adenocarcinoma-associated Gene, AGR2, Induces Expression of Amphiregulin through Hippo Pathway Co-activator YAP1 Activation.* J Biol Chem. 286 (20), p. 18301-10.

507. Zhao, B., Li, L., and Guan, K.L. (2010) *Hippo signaling at a glance.* J Cell Sci. 123 (Pt 23), p. 4001-6.

508. Busser, B., Sancey, L., Brambilla, E., Coll, J.L., and Hurbin, A. (2011) *The multiple roles of amphiregulin in human cancer.* Biochim Biophys Acta. 1816 (2), p. 119-31.

509. Berquin, I.M., Dziubinski, M.L., Nolan, G.P., and Ethier, S.P. (2001) *A functional screen for genes inducing epidermal growth factor autonomy of human mammary epithelial cells confirms the role of amphiregulin.* Oncogene. 20 (30), p. 4019-28.

510. Berasain, C., Garcia-Trevijano, E.R., Castillo, J., Erroba, E., Santamaria, M., Lee, D.C., Prieto, J., and Avila, M.A. (2005) *Novel role for amphiregulin in protection from liver injury.* J Biol Chem. 280 (19), p. 19012-20.

511. Hurbin, A., Coll, J.L., Dubrez-Daloz, L., Mari, B., Auberger, P., Brambilla, C., and Favrot, M.C. (2005) *Cooperation of amphiregulin and insulin-like growth factor-1 inhibits Bax- and Bad-mediated apoptosis via a protein kinase C-dependent pathway in non-small cell lung cancer cells.* J Biol Chem. 280 (20), p. 19757-67.

512. Motoyama, A.B., Hynes, N.E., and Lane, H.A. (2002) *The efficacy of ErbB receptor-targeted anticancer therapeutics is influenced by the availability of epidermal growth factor-related peptides.* Cancer Res. 62 (11), p. 3151-8.

513. Ferrer-Soler, L., Vazquez-Martin, A., Brunet, J., Menendez, J.A., De Llorens, R., and Colomer, R. (2007) *An update of the mechanisms of resistance to EGFR-tyrosine kinase inhibitors in breast cancer: Gefitinib (Iressa) -induced changes in the expression and nucleo-cytoplasmic trafficking of HER-ligands (Review).* Int J Mol Med. 20 (1), p. 3-10.

514. Busser, B., Sancey, L., Josserand, V., Niang, C., Khochbin, S., Favrot, M.C., Coll, J.L., and Hurbin, A. (2010) *Amphiregulin promotes resistance to gefitinib in nonsmall cell lung cancer cells by regulating Ku70 acetylation.* Mol Ther. 18 (3), p. 536-43.

515. Chanda, D., Lee, J.H., Sawant, A., Hensel, J.A., Isayeva, T., Reilly, S.D., Siegal, G.P., Smith, C., Grizzle, W., Singh, R., and Ponnazhagan, S. (2014) *Anterior gradient protein-2 is a regulator of cellular adhesion in prostate cancer.* PLoS One. 9 (2), p. e89940.

516. Dunnington, D.J., Hughes, C.M., Monaghan, P., and Rudland, P.S. (1983) *Phenotypic instability of rat mammary tumor epithelial cells.* J Natl Cancer Inst. 71 (6), p. 1227-40.

517. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012) *NIH Image to ImageJ: 25 years of image analysis.* Nat Methods. 9 (7), p. 671-5.

518. Chu, L., Tempelman, A., Miller, C.a., and Hammer, D. (1994) *Centrifugation assay of IgE-mediated cell adhesion to antigen-coated gels.* AIChE Journal. 40 (4), p. 692-703.

519. Koo, L.Y., Irvine, D.J., Mayes, A.M., Lauffenburger, D.A., and Griffith, L.G. (2002) *Co-regulation of cell adhesion by nanoscale RGD organization and mechanical stimulus.* J Cell Sci. 115 (Pt 7), p. 1423-33.

520. Zhuang, H. and Matsunami, H. (2008) *Evaluating cell-surface expression and measuring activation of mammalian odorant receptors in heterologous cells.* Nat Protoc. 3 (9), p. 1402-13.

521. Mufson, R.A. and Gesner, T.G. (1987) *Binding and internalization of recombinant human erythropoietin in murine erythroid precursor cells.* Blood. 69 (5), p. 1485-90.

522. Yeliseev, A., Zoubak, L., and Gawrisch, K. (2007) *Use of dual affinity tags for expression and purification of functional peripheral cannabinoid receptor.* Protein Expr Purif. 53 (1), p. 153-63.

523. Holmgren, A. (1979) *Thioredoxin catalyzes the reduction of insulin disulfides by dithiothreitol and dihydrolipoamide.* J Biol Chem. 254 (19), p. 9627-32.

524. Holmgren, A. (1979) *Glutathione-dependent synthesis of deoxyribonucleotides. Purification and characterization of glutaredoxin from Escherichia coli.* J Biol Chem. 254 (9), p. 3664-71.

525. Nandakumar, M.P., Shen, J., Raman, B., and Marten, M.R. (2003) *Solubilization of trichloroacetic acid (TCA) precipitated microbial proteins via naOH for two-dimensional electrophoresis.* J Proteome Res. 2 (1), p. 89-93.

526. Akagi, M., Yokozaki, H., Kitadai, Y., Ito, R., Yasui, W., Haruma, K., Kajiyama, G., and Tahara, E. (1995) *Expression of amphiregulin in human gastric cancer cell lines.* Cancer. 75 (6 Suppl), p. 1460-6.

527. Amanchy, R., Kalume, D.E., Iwahori, A., Zhong, J., and Pandey, A. (2005) *Phosphoproteome analysis of HeLa cells using stable isotope labeling with amino acids in cell culture (SILAC).* J Proteome Res. 4 (5), p. 1661-71.

528. Ambolet-Camoit, A., Bui, L.C., Pierre, S., Chevallier, A., Marchand, A., Coumoul, X., Garlatti, M., Andreau, K., Barouki, R., and Aggerbeck, M. (2010) *2,3,7,8-tetrachlorodibenzo-p-dioxin counteracts the p53 response to a genotoxicant by upregulating expression of the metastasis marker agr2 in the hepatocarcinoma cell line HepG2.* Toxicol Sci. 115 (2), p. 501-12.

529. Ambort, D., van der Post, S., Johansson, M.E., Mackenzie, J., Thomsson, E., Krengel, U., and Hansson, G.C. (2011) *Function of the CysD domain of the gel-forming MUC2 mucin.* Biochem J. 436 (1), p. 61-70.

530. Arboleda, M.J., Lyons, J.F., Kabbinavar, F.F., Bray, M.R., Snow, B.E., Ayala, R., Danino, M., Karlan, B.Y., and Slamon, D.J. (2003) *Overexpression of AKT2/protein*

kinase Bbeta leads to up-regulation of beta1 integrins, increased invasion, and metastasis of human breast and ovarian cancer cells. Cancer Res. 63 (1), p. 196-206.

531. Bai, Z. and Gust, R. (2009) *Breast cancer, estrogen receptor and ligands.* Arch Pharm (Weinheim). 342 (3), p. 133-49.

532. Barraclough, R., Dawson, K.J., and Rudland, P.S. (1984) *Elongated cells derived from rat mammary cuboidal epithelial cell lines resemble cultured mesenchymal cells in their pattern of protein synthesis.* Biochem Biophys Res Commun. 120 (2), p. 351-8.

533. Bass, R., Ruddock, L.W., Klappa, P., and Freedman, R.B. (2004) *A major fraction of endoplasmic reticulum-located glutathione is present as mixed disulfides with protein.* J Biol Chem. 279 (7), p. 5257-62.

534. Benham, A.M., Cabibbo, A., Fassio, A., Bulleid, N., Sitia, R., and Braakman, I. (2000) *The CXXCXXC motif determines the folding, structure and stability of human Ero1-Lalpha.* EMBO J. 19 (17), p. 4493-502.

535. Berx, G. and van Roy, F. (2009) *Involvement of members of the cadherin superfamily in cancer.* Cold Spring Harb Perspect Biol. 1 (6), p. a003129.

536. Zhang, Z. and Henzel, W.J. (2004) *Signal peptide prediction based on analysis of experimentally verified cleavage sites.* Protein Sci. 13 (10), p. 2819-24.

537. Rainey-Barger, E.K., Mkrtchian, S., and Tsai, B. (2007) *Dimerization of ERp29, a PDI-like protein, is essential for its diverse functions.* Mol Biol Cell. 18 (4), p. 1253-60.

538. Wang, X., Bansal, S., Jiang, M., and Prestegard, J.H. (2008) *RDC-assisted modeling of symmetric protein homo-oligomers.* Protein Sci. 17 (5), p. 899-907.

539. Maslon, M.M., Hrstka, R., Vojtesek, B., and Hupp, T.R. (2010) *A divergent substrate-binding loop within the pro-oncogenic protein anterior gradient-2 forms a docking site for Reptin.* J Mol Biol. 404 (3), p. 418-38.

540. Jensen, M.R., Ruigrok, R.W., and Blackledge, M. (2013) *Describing intrinsically disordered proteins at atomic resolution by NMR.* Curr Opin Struct Biol. 23 (3), p. 426-35.

541. Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011) *SignalP 4.0: discriminating signal peptides from transmembrane regions.* Nat Methods. 8 (10), p. 785-6.

542. Harada, M., Murakami, H., Okawa, A., Okimoto, N., Hiraoka, S., Nakahara, T., Akasaka, R., Shiraishi, Y., Futatsugi, N., Mizutani-Koseki, Y., Kuroiwa, A., Shirouzu, M., Yokoyama, S., Taiji, M., Iseki, S., Ornitz, D.M., and Koseki, H. (2009) *FGF9 monomer-dimer equilibrium regulates extracellular matrix affinity and tissue diffusion.* Nat Genet. 41 (3), p. 289-98.

543. Schnitzel, W., Monschein, U., and Besemer, J. (1994) *Monomer-dimer equilibria of interleukin-8 and neutrophil-activating peptide 2. Evidence for IL-8 binding as a dimer and oligomer to IL-8 receptor B.* J Leukoc Biol. 55 (6), p. 763-70.

544. Thier, S.O. (1986) *Potassium physiology.* Am J Med. 80 (4A), p. 3-7.

545. Wollman, E.E., d'Auriol, L., Rimsky, L., Shaw, A., Jacquot, J.P., Wingfield, P., Graber, P., Dessarps, F., Robin, P., Galibert, F., and et al. (1988) *Cloning and expression of a cDNA for human thioredoxin.* J Biol Chem. 263 (30), p. 15506-12.

546. Ibbetson, A.L. and Freedman, R.B. (1976) *Thiol-protein disulphide oxidoreductases. Assay of microsomal membrane-bound glutathione-insulin transhydrogenase and comparison with protein disulphide-isomerase.* Biochem J. 159 (2), p. 377-84.

547. Sun, Q.A., Su, D., Novoselov, S.V., Carlson, B.A., Hatfield, D.L., and Gladyshev, V.N. (2005) *Reaction mechanism and regulation of mammalian thioredoxin/glutathione reductase.* Biochemistry. 44 (44), p. 14528-37.

548. Lillig, C.H. and Berndt, C. (2013) *Glutaredoxins in thiol/disulfide exchange.* Antioxid Redox Signal. 18 (13), p. 1654-65.

549. Choi, K.H., Basma, H., Singh, J., and Cheng, P.W. (2005) *Activation of CMV promoter-controlled glycosyltransferase and beta -galactosidase glycogenes by butyrate, tricostatin A, and 5-aza-2'-deoxycytidine.* Glycoconj J. 22 (1-2), p. 63-9.

550. Palmer, T.D., Rosman, G.J., Osborne, W.R., and Miller, A.D. (1991) *Genetically modified skin fibroblasts persist long after transplantation but gradually inactivate introduced genes.* Proc Natl Acad Sci U S A. 88 (4), p. 1330-4.

551. Qin, J.Y., Zhang, L., Clift, K.L., Hulur, I., Xiang, A.P., Ren, B.Z., and Lahn, B.T. (2010) *Systematic comparison of constitutive promoters and the doxycycline-inducible promoter.* PLoS One. 5 (5), p. e10611.

552. Teschendorf, C., Warrington, K.H., Jr., Siemann, D.W., and Muzyczka, N. (2002) *Comparison of the EF-1 alpha and the CMV promoter for engineering stable tumor cell lines using recombinant adeno-associated virus.* Anticancer Res. 22 (6A), p. 3325-30.

553. Wilson, M.H., Coates, C.J., and George, A.L., Jr. (2007) *PiggyBac transposon-mediated gene transfer in human cells.* Mol Ther. 15 (1), p. 139-45.

554. Rajewska, M., Wegrzyn, K., and Konieczny, I. (2012) *AT-rich region and repeated sequences - the essential elements of replication origins of bacterial replicons.* FEMS Microbiol Rev. 36 (2), p. 408-34.

555. Lambert, N. and Freedman, R.B. (1985) *The latency of rat liver microsomal protein disulphide-isomerase.* Biochem J. 228 (3), p. 635-45.

556. Linstedt, A.D. and Hauri, H.P. (1993) *Giantin, a novel conserved Golgi membrane protein containing a cytoplasmic domain of at least 350 kDa.* Mol Biol Cell. 4 (7), p. 679-93.

557. Zheng, C.F. and Guan, K.L. (1994) *Cytoplasmic localization of the mitogen-activated protein kinase activator MEK.* J Biol Chem. 269 (31), p. 19947-52.

558. Gray, T.A., Murray, E., Nowicki, M.W., Remnant, L., Scherl, A., Muller, P., Vojtesek, B., and 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 Sci. 22 (9), p. 1266-78.

559. Solovyov, A. and Gilbert, H.F. (2004) *Zinc-dependent dimerization of the folding catalyst, protein disulfide isomerase.* Protein Sci. 13 (7), p. 1902-7.

560. Tian, G., Kober, F.X., Lewandrowski, U., Sickmann, A., Lennarz, W.J., and Schindelin, H. (2008) *The catalytic activity of protein-disulfide isomerase requires a conformationally flexible molecule.* J Biol Chem. 283 (48), p. 33630-40.

561. Baker, K.M., Chakravarthi, S., Langton, K.P., Sheppard, A.M., Lu, H., and Bulleid, N.J. (2008) *Low reduction potential of Ero1alpha regulatory disulphides ensures tight control of substrate oxidation.* EMBO J. 27 (22), p. 2988-97.

562. Dias-Gunasekara, S., Gubbens, J., van Lith, M., Dunne, C., Williams, J.A., Kataky, R., Scoones, D., Lapthorn, A., Bulleid, N.J., and Benham, A.M. (2005) *Tissue-specific expression and dimerization of the endoplasmic reticulum oxidoreductase Ero1beta.* J Biol Chem. 280 (38), p. 33066-75.

563. Kpadeh, Z.Z., Jameson-Lee, M., Yeh, A.J., Chertihin, O., Shumilin, I.A., Dey, R., Day, S.R., and Hoffman, P.S. (2013) *Disulfide bond oxidoreductase DsbA2 of Legionella pneumophila exhibits protein disulfide isomerase activity.* J Bacteriol. 195 (8), p. 1825-33.

564. McGaughey, G.B., Gagne, M., and Rappe, A.K. (1998) *pi-Stacking interactions. Alive and well in proteins.* J Biol Chem. 273 (25), p. 15458-63.

565. Gallivan, J.P. and Dougherty, D.A. (1999) *Cation-pi interactions in structural biology.* Proc Natl Acad Sci U S A. 96 (17), p. 9459-64.

566. Crowley, P.B. and Golovin, A. (2005) *Cation-pi interactions in protein-protein interfaces.* Proteins. 59 (2), p. 231-9.

567. Gallivan JP, D.D. (2000) *A computational study of cation-pi interactions versus salt bridges in aqueous media: implications for protein engineering.* J Am Chem Soc 122, p. 870-874.

568. Pace, C.N., Horn, G., Hebert, E.J., Bechert, J., Shaw, K., Urbanikova, L., Scholtz, J.M., and Sevcik, J. (2001) *Tyrosine hydrogen bonds make a large contribution to protein stability.* J Mol Biol. 312 (2), p. 393-404.

569. Inokuchi, Y., Kobayashi, Y., Ito, T., and Ebata, T. (2007) *Conformation of L-tyrosine studied by fluorescence-detected UV-UV and IR-UV double-resonance spectroscopy.* J Phys Chem A. 111 (17), p. 3209-15.

570. Gentile, F., Amodeo, P., Febbraio, F., Picaro, F., Motta, A., Formisano, S., and Nucci, R. (2002) *SDS-resistant active and thermostable dimers are obtained from the dissociation of homotetrameric beta-glycosidase from hyperthermophilic Sulfolobus solfataricus in SDS. Stabilizing role of the A-C intermonomeric interface.* J Biol Chem. 277 (46), p. 44050-60.

571. Wei, P.C., Hsieh, Y.H., Su, M.I., Jiang, X., Hsu, P.H., Lo, W.T., Weng, J.Y., Jeng, Y.M., Wang, J.M., Chen, P.L., Chang, Y.C., Lee, K.F., Tsai, M.D., Shew, J.Y., and Lee, W.H. (2012) *Loss of the oxidative stress sensor NPGPx compromises GRP78 chaperone activity and induces systemic disease.* Mol Cell. 48 (5), p. 747-59.

572. Zhang, S., Wang, G., Fernig, D.G., Rudland, P.S., Webb, S.E., Barraclough, R., and Martin-Fernandez, M. (2005) *Interaction of metastasis-inducing S100A4 protein in vivo by fluorescence lifetime imaging microscopy.* Eur Biophys J. 34 (1), p. 19-27.

573. Anelli, T., Alessio, M., Bachi, A., Bergamelli, L., Bertoli, G., Camerini, S., Mezghrani, A., Ruffato, E., Simmen, T., and Sitia, R. (2003) *Thiol-mediated protein retention in the endoplasmic reticulum: the role of ERp44.* EMBO J. 22 (19), p. 5015-22.

574. Isidoro, C., Maggioni, C., Demoz, M., Pizzagalli, A., Fra, A.M., and Sitia, R. (1996) *Exposed thiols confer localization in the endoplasmic reticulum by retention rather than retrieval.* J Biol Chem. 271 (42), p. 26138-42.

575. Kosugi, S., Hasebe, M., Tomita, M., and Yanagawa, H. (2009) *Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs.* Proc Natl Acad Sci U S A. 106 (25), p. 10171-6.

576. Goldfarb, D.S., Corbett, A.H., Mason, D.A., Harreman, M.T., and Adam, S.A. (2004) *Importin alpha: a multipurpose nuclear-transport receptor.* Trends Cell Biol. 14 (9), p. 505-14.

577. Huber, O., Menard, L., Haurie, V., Nicou, A., Taras, D., and Rosenbaum, J. (2008) *Pontin and reptin, two related ATPases with multiple roles in cancer.* Cancer Res. 68 (17), p. 6873-6.

578. Bendas, G. and Borsig, L. (2012) *Cancer cell adhesion and metastasis: selectins, integrins, and the inhibitory potential of heparins.* Int J Cell Biol. 2012, p. 676731.

579. Tamm, I., Wang, Y., Sausville, E., Scudiero, D.A., Vigna, N., Oltersdorf, T., and Reed, J.C. (1998) *IAP-family protein survivin inhibits caspase activity and apoptosis induced by Fas (CD95), Bax, caspases, and anticancer drugs.* Cancer Res. 58 (23), p. 5315-20.

580. Feng, J., Yang, S., Xu, L., Tian, H., Sun, L., and Tang, X. (2007) *Role of caspase-3 inhibitor in induced anoikis of mesenchymal stem cells in vitro.* J Huazhong Univ Sci Technolog Med Sci. 27 (2), p. 183-5.

581. Carre, A.a.L., V., *Cell adhesion to polystyrene substrates: Relevance of interfacial free energy*, in *Contact Angle, Wettability and Adhesion*, K. Mittal, Editor. 2008, VSP. p. 253-267.

582. Mazia, D., Schatten, G., and Sale, W. (1975) *Adhesion of cells to surfaces coated with polylysine. Applications to electron microscopy.* J Cell Biol. 66 (1), p. 198-200.

583. Wang, L.H. (2004) *Molecular signaling regulating anchorage-independent growth of cancer cells.* Mt Sinai J Med. 71 (6), p. 361-7.

584. Higginbotham, J.N., Demory Beckler, M., Gephart, J.D., Franklin, J.L., Bogatcheva, G., Kremers, G.J., Piston, D.W., Ayers, G.D., McConnell, R.E., Tyska, M.J., and Coffey, R.J. (2011) *Amphiregulin exosomes increase cancer cell invasion.* Curr Biol. 21 (9), p. 779-86.

585. Baillo, A., Giroux, C., and Ethier, S.P. (2011) *Knock-down of amphiregulin inhibits cellular invasion in inflammatory breast cancer.* J Cell Physiol. 226 (10), p. 2691-701.

586. Xia, H., Mao, Q., Paulson, H.L., and Davidson, B.L. (2002) *siRNA-mediated gene silencing in vitro and in vivo.* Nat Biotechnol. 20 (10), p. 1006-10.

587. Das, D.S., Wadhwa, N., Kunj, N., Sarda, K., Pradhan, B.S., and Majumdar, S.S. (2013) *Dickkopf homolog 3 (DKK3) plays a crucial role upstream of WNT/beta-CATENIN signaling for Sertoli cell mediated regulation of spermatogenesis.* PLoS One. 8 (5), p. e63603.

588. Antecol, M.H. and Mukherjee, B.B. (1982) *Effects of 12-O-tetradecanoylphorbol-13-acetate on fibroblasts from individuals genetically predisposed to cancer.* Cancer Res. 42 (9), p. 3870-9.

589. Prag, S., Lepekhin, E.A., Kolkova, K., Hartmann-Petersen, R., Kawa, A., Walmod, P.S., Belman, V., Gallagher, H.C., Berezin, V., Bock, E., and Pedersen, N. (2002) *NCAM regulates cell motility.* J Cell Sci. 115 (Pt 2), p. 283-92.

590. Zimmerman, E., Geiger, B., and Addadi, L. (2002) *Initial stages of cell-matrix adhesion can be mediated and modulated by cell-surface hyaluronan.* Biophys J. 82 (4), p. 1848-57.

591. Cohen, M., Klein, E., Geiger, B., and Addadi, L. (2003) *Organization and adhesive properties of the hyaluronan pericellular coat of chondrocytes and epithelial cells.* Biophys J. 85 (3), p. 1996-2005.

592. Cohen, M., Kam, Z., Addadi, L., and Geiger, B. (2006) *Dynamic study of the transition from hyaluronan- to integrin-mediated adhesion in chondrocytes.* EMBO J. 25 (2), p. 302-11.

593. Nandi, A., Estess, P., and Siegelman, M. (2004) *Bimolecular complex between rolling and firm adhesion receptors required for cell arrest; CD44 association with VLA-4 in T cell extravasation.* Immunity. 20 (4), p. 455-65.

594. DeGrendele, H.C., Estess, P., Picker, L.J., and Siegelman, M.H. (1996) *CD44 and its ligand hyaluronate mediate rolling under physiologic flow: a novel lymphocyte-endothelial cell primary adhesion pathway.* J Exp Med. 183 (3), p. 1119-30.

595. Johnson, P., Maiti, A., Brown, K.L., and Li, R. (2000) *A role for the cell adhesion molecule CD44 and sulfation in leukocyte-endothelial cell adhesion during an inflammatory response?* Biochem Pharmacol. 59 (5), p. 455-65.

596. Yoshihara, S., Kon, A., Kudo, D., Nakazawa, H., Kakizaki, I., Sasaki, M., Endo, M., and Takagaki, K. (2005) *A hyaluronan synthase suppressor, 4-methylumbelliferone, inhibits liver metastasis of melanoma cells.* FEBS Lett. 579 (12), p. 2722-6.

597. Arai, E., Nishida, Y., Wasa, J., Urakawa, H., Zhuo, L., Kimata, K., Kozawa, E., Futamura, N., and Ishiguro, N. (2011) *Inhibition of hyaluronan retention by 4-methylumbelliferone suppresses osteosarcoma cells in vitro and lung metastasis in vivo.* Br J Cancer. 105 (12), p. 1839-49.

598. Carlsson, R., Engvall, E., Freeman, A., and Ruoslahti, E. (1981) *Laminin and fibronectin in cell adhesion: enhanced adhesion of cells from regenerating liver to laminin.* Proc Natl Acad Sci U S A. 78 (4), p. 2403-6.

599. Sottile, J., Hocking, D.C., and Swiatek, P.J. (1998) *Fibronectin matrix assembly enhances adhesion-dependent cell growth.* J Cell Sci. 111 (Pt 19), p. 2933-43.

600. Israelachvili, J., *Intermolecular and Surface Forces*. 1992, Academic Press. p. 238.

601. Sive, H.L., Hattori, K., and Weintraub, H. (1989) *Progressive determination during formation of the anteroposterior axis in Xenopus laevis.* Cell. 58 (1), p. 171-80.

602. Garza-Garcia, A.A., Driscoll, P.C., and Brockes, J.P. (2010) *Evidence for the local evolution of mechanisms underlying limb regeneration in salamanders.* Integr Comp Biol. 50 (4), p. 528-35.

603. Chaurasia, P., Aguirre-Ghiso, J.A., Liang, O.D., Gardsvoll, H., Ploug, M., and Ossowski, L. (2006) *A region in urokinase plasminogen receptor domain III*

controlling a functional association with alpha5beta1 integrin and tumor growth. J Biol Chem. 281 (21), p. 14852-63.

604. Wei, Y., Waltz, D.A., Rao, N., Drummond, R.J., Rosenberg, S., and Chapman, H.A. (1994) *Identification of the urokinase receptor as an adhesion receptor for vitronectin.* J Biol Chem. 269 (51), p. 32380-8.

605. Degryse, B., Resnati, M., Czekay, R.P., Loskutoff, D.J., and Blasi, F. (2005) *Domain 2 of the urokinase receptor contains an integrin-interacting epitope with intrinsic signaling activity: generation of a new integrin inhibitor.* J Biol Chem. 280 (26), p. 24792-803.

606. Deckert, M., Kubar, J., and Bernard, A. (1992) *CD58 and CD59 molecules exhibit potentializing effects in T cell adhesion and activation.* J Immunol. 148 (3), p. 672-7.

607. Deckert, M., Kubar, J., Zoccola, D., Bernard-Pomier, G., Angelisova, P., Horejsi, V., and Bernard, A. (1992) *CD59 molecule: a second ligand for CD2 in T cell adhesion.* Eur J Immunol. 22 (11), p. 2943-7.

608. Barresi, R. and Campbell, K.P. (2006) *Dystroglycan: from biosynthesis to pathogenesis of human disease.* J Cell Sci. 119 (Pt 2), p. 199-207.

609. Yoshikawa, H.Y., Kawano, T., Matsuda, T., Kidoaki, S., and Tanaka, M. (2013) *Morphology and adhesion strength of myoblast cells on photocurable gelatin under native and non-native micromechanical environments.* J Phys Chem B. 117 (15), p. 4081-8.

610. Powers, M.J., Rodriguez, R.E., and Griffith, L.G. (1997) *Cell-substratum adhesion strength as a determinant of hepatocyte aggregate morphology.* Biotechnol Bioeng. 53 (4), p. 415-26.

611. Palecek, S.P., Loftus, J.C., Ginsberg, M.H., Lauffenburger, D.A., and Horwitz, A.F. (1997) *Integrin-ligand binding properties govern cell migration speed through cell-substratum adhesiveness.* Nature. 385 (6616), p. 537-40.

612. Nagano, M., Hoshino, D., Koshikawa, N., Akizawa, T., and Seiki, M. (2012) *Turnover of focal adhesions and cancer cell migration.* Int J Cell Biol. 2012, p. 310616.

613. Kimura, M., Furuichi, M., Yamamoto, M., Kumasaka, T., Mizuno, H., Miyano, M., and Yamaguchi, I. (2002) *The flexible C-terminal region of Aspergillus terreus blasticidin S deaminase: identification of its functional roles with deletion enzymes.* Biochem Biophys Res Commun. 290 (1), p. 421-6.

614. Hollingsworth, M.A. and Swanson, B.J. (2004) *Mucins in cancer: protection and control of the cell surface.* Nat Rev Cancer. 4 (1), p. 45-60.

615. Wreesmann, V.B., Sieczka, E.M., Socci, N.D., Hezel, M., Belbin, T.J., Childs, G., Patel, S.G., Patel, K.N., Tallini, G., Prystowsky, M., Shaha, A.R., Kraus, D., Shah, J.P., Rao, P.H., Ghossein, R., and Singh, B. (2004) *Genome-wide profiling of papillary thyroid cancer identifies MUC1 as an independent prognostic marker.* Cancer Res. 64 (11), p. 3780-9.

616. Khodarev, N.N., Pitroda, S.P., Beckett, M.A., MacDermed, D.M., Huang, L., Kufe, D.W., and Weichselbaum, R.R. (2009) *MUC1-induced transcriptional programs associated with tumorigenesis predict outcome in breast and lung cancer.* Cancer Res. 69 (7), p. 2833-7.

617. Lapointe, J., Li, C., Higgins, J.P., van de Rijn, M., Bair, E., Montgomery, K., Ferrari, M., Egevad, L., Rayford, W., Bergerheim, U., Ekman, P., DeMarzo, A.M., Tibshirani, R., Botstein, D., Brown, P.O., Brooks, J.D., and Pollack, J.R. (2004) *Gene expression profiling identifies clinically relevant subtypes of prostate cancer.* Proc Natl Acad Sci U S A. 101 (3), p. 811-6.

618. Rajabi, H., Alam, M., Takahashi, H., Kharbanda, A., Guha, M., Ahmad, R., and Kufe, D. (2014) *MUC1-C oncoprotein activates the ZEB1/miR-200c regulatory loop and epithelial-mesenchymal transition.* Oncogene. 33 (13), p. 1680-9.

619. Roy, L.D., Sahraei, M., Subramani, D.B., Besmer, D., Nath, S., Tinder, T.L., Bajaj, E., Shanmugam, K., Lee, Y.Y., Hwang, S.I., Gendler, S.J., and Mukherjee, P. (2011) *MUC1 enhances invasiveness of pancreatic cancer cells by inducing epithelial to mesenchymal transition.* Oncogene. 30 (12), p. 1449-59.

620. Gonzalez-Gronow, M., Cuchacovich, M., Llanos, C., Urzua, C., Gawdi, G., and Pizzo, S.V. (2006) *Prostate cancer cell proliferation in vitro is modulated by antibodies against glucose-regulated protein 78 isolated from patient serum.* Cancer Res. 66 (23), p. 11424-31.

621. Kypreou, K.P., Kavvadas, P., Karamessinis, P., Peroulis, M., Alberti, A., Sideras, P., Psarras, S., Capetanaki, Y., Politis, P.K., and Charonis, A.S. (2008) *Altered expression of calreticulin during the development of fibrosis.* Proteomics. 8 (12), p. 2407-19.

622. Hance, M.W., Dole, K., Gopal, U., Bohonowych, J.E., Jezierska-Drutel, A., Neumann, C.A., Liu, H., Garraway, I.P., and Isaacs, J.S. (2012) *Secreted Hsp90 is a novel regulator of the epithelial to mesenchymal transition (EMT) in prostate cancer.* J Biol Chem. 287 (45), p. 37732-44.

623. Singh, A.P., Chaturvedi, P., and Batra, S.K. (2007) *Emerging roles of MUC4 in cancer: a novel target for diagnosis and therapy.* Cancer Res. 67 (2), p. 433-6.

624. Phillips, M.M., Sheaff, M.T., and Szlosarek, P.W. (2013) *Targeting arginine-dependent cancers with arginine-degrading enzymes: opportunities and challenges.* Cancer Res Treat. 45 (4), p. 251-62.

625. Tang, X.H. and Gudas, L.J. (2011) *Retinoids, retinoic acid receptors, and cancer.* Annu Rev Pathol. 6, p. 345-64.

626. Budhu, A.S. and Noy, N. (2002) *Direct channeling of retinoic acid between cellular retinoic acid-binding protein II and retinoic acid receptor sensitizes mammary carcinoma cells to retinoic acid-induced growth arrest.* Mol Cell Biol. 22 (8), p. 2632-41.

627. Friedmann-Morvinski, D. and Verma, I.M. (2014) *Dedifferentiation and reprogramming: origins of cancer stem cells.* EMBO Rep. 15 (3), p. 244-53.

628. Ussar, S., Bezy, O., Bluher, M., and Kahn, C.R. (2012) *Glypican-4 enhances insulin signaling via interaction with the insulin receptor and serves as a novel adipokine.* Diabetes. 61 (9), p. 2289-98.

629. Komiya, Y. and Habas, R. (2008) *Wnt signal transduction pathways.* Organogenesis. 4 (2), p. 68-75.

630. Ohkawara, B., Yamamoto, T.S., Tada, M., and Ueno, N. (2003) *Role of glypican 4 in the regulation of convergent extension movements during gastrulation in Xenopus laevis.* Development. 130 (10), p. 2129-38.

631. Topczewski, J., Sepich, D.S., Myers, D.C., Walker, C., Amores, A., Lele, Z., Hammerschmidt, M., Postlethwait, J., and Solnica-Krezel, L. (2001) *The zebrafish glypican knypek controls cell polarity during gastrulation movements of convergent extension.* Dev Cell. 1 (2), p. 251-64.

632. Chevallet, M., Diemer, H., Van Dorssealer, A., Villiers, C., and Rabilloud, T. (2007) *Toward a better analysis of secreted proteins: the example of the myeloid cells secretome.* Proteomics. 7 (11), p. 1757-70.

633. Ahram, M., Adkins, J.N., Auberry, D.L., Wunschel, D.S., and Springer, D.L. (2005) *A proteomic approach to characterize protein shedding.* Proteomics. 5 (1), p. 123-31.

634. Cattaneo, M., Lotti, L.V., Martino, S., Cardano, M., Orlandi, R., Mariani-Costantini, R., and Biunno, I. (2009) *Functional characterization of two secreted SEL1L isoforms capable of exporting unassembled substrate.* J Biol Chem. 284 (17), p. 11405-15.

635. Holla, O.L., Cameron, J., Berge, K.E., Ranheim, T., and Leren, T.P. (2007) *Degradation of the LDL receptors by PCSK9 is not mediated by a secreted protein acted upon by PCSK9 extracellularly.* BMC Cell Biol. 8, p. 9.

636. Lonka-Nevalaita, L., Lume, M., Leppanen, S., Jokitalo, E., Peranen, J., and Saarma, M. (2010) *Characterization of the intracellular localization, processing, and secretion of two glial cell line-derived neurotrophic factor splice isoforms.* J Neurosci. 30 (34), p. 11403-13.

637. Yousef, H., Conboy, M.J., Li, J., Zeiderman, M., Vazin, T., Schlesinger, C., Schaffer, D.V., and Conboy, I.M. (2013) *hESC-secreted proteins can be enriched for multiple regenerative therapies by heparin-binding.* Aging (Albany NY). 5 (5), p. 357-72.

638. Nakamura, E., Abreu-e-Lima, P., Awakura, Y., Inoue, T., Kamoto, T., Ogawa, O., Kotani, H., Manabe, T., Zhang, G.J., Kondo, K., Nose, V., and Kaelin, W.G., Jr. (2006) *Clusterin is a secreted marker for a hypoxia-inducible factor-independent function of the von Hippel-Lindau tumor suppressor protein.* Am J Pathol. 168 (2), p. 574-84.

639. Shibata, K., Kikkawa, F., Nawa, A., Suganuma, N., and Hamaguchi, M. (1997) *Fibronectin secretion from human peritoneal tissue induces Mr 92,000 type IV collagenase expression and invasion in ovarian cancer cell lines.* Cancer Res. 57 (23), p. 5416-20.

640. Kowalczyk, A.P., Tulloh, R.H., and McKeown-Longo, P.J. (1990) *Polarized fibronectin secretion and localized matrix assembly sites correlate with subendothelial matrix formation.* Blood. 75 (12), p. 2335-42.

641. Lott JA, N.E., *Lactate dehydrogenase*, in *Clinical Enzymology, a Case-oriented Approach*, W.P. Lott JA, Editor. 1987, Year Book Medical Pub. p. 213-214.

642. Glick, J.H., Jr. (1969) *Serum lactate dehydrogenase isoenzyme and total lactate dehydrogenase values in health and disease, and clinical evaluation of these tests by means of discriminant analysis.* Am J Clin Pathol. 52 (3), p. 320-8.

643. Guo, C., Buranych, A., Sarkar, D., Fisher, P.B., and Wang, X.Y. (2013) *The role of tumor-associated macrophages in tumor vascularization.* Vasc Cell. 5 (1), p. 20.

644. Soria, B., Roche, E., Berna, G., Leon-Quinto, T., Reig, J.A., and Martin, F. (2000) *Insulin-secreting cells derived from embryonic stem cells normalize glycemia in streptozotocin-induced diabetic mice.* Diabetes. 49 (2), p. 157-62.

645. Lee, M.S., Igawa, T., Yuan, T.C., Zhang, X.Q., Lin, F.F., and Lin, M.F. (2003) *ErbB-2 signaling is involved in regulating PSA secretion in androgen-independent human prostate cancer LNCaP C-81 cells.* Oncogene. 22 (5), p. 781-96.

646. Islam, S.M., Shinmyo, Y., Okafuji, T., Su, Y., Naser, I.B., Ahmed, G., Zhang, S., Chen, S., Ohta, K., Kiyonari, H., Abe, T., Tanaka, S., Nishinakamura, R., Terashima, T., Kitamura, T., and Tanaka, H. (2009) *Draxin, a repulsive guidance protein for spinal cord and forebrain commissures.* Science. 323 (5912), p. 388-93.

647. Harburg, G.C. and Hinck, L. (2011) *Navigating breast cancer: axon guidance molecules as breast cancer tumor suppressors and oncogenes.* J Mammary Gland Biol Neoplasia. 16 (3), p. 257-70.

648. Spencer, S.C. and Fabre, J.W. (1987) *Identification in rat liver and serum of water-soluble class I MHC molecules possibly homologous to the murine Q10 gene product.* J Exp Med. 165 (6), p. 1595-608.

649. Salih, H.R., Rammensee, H.G., and Steinle, A. (2002) *Cutting edge: down-regulation of MICA on human tumors by proteolytic shedding.* J Immunol. 169 (8), p. 4098-102.

650. Khoury, G.A., Baliban, R.C., and Floudas, C.A. (2011) *Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database.* Sci Rep. 1.

651. Bosques, C.J., Raguram, S., and Sasisekharan, R. (2006) *The sweet side of biomarker discovery.* Nat Biotechnol. 24 (9), p. 1100-1.

652. Tagliabracci, V.S., Engel, J.L., Wen, J., Wiley, S.E., Worby, C.A., Kinch, L.N., Xiao, J., Grishin, N.V., and Dixon, J.E. (2012) *Secreted kinase phosphorylates extracellular proteins that regulate biomineralization.* Science. 336 (6085), p. 1150-3.

653. Zhou, W., Ross, M.M., Tessitore, A., Ornstein, D., Vanmeter, A., Liotta, L.A., and Petricoin, E.F., 3rd. (2009) *An initial characterization of the serum phosphoproteome.* J Proteome Res. 8 (12), p. 5523-31.

654. Carrascal, M., Gay, M., Ovelleiro, D., Casas, V., Gelpi, E., and Abian, J. (2010) *Characterization of the human plasma phosphoproteome using linear ion trap mass spectrometry and multiple search engines.* J Proteome Res. 9 (2), p. 876-84.

655. Yang, Y., Rao, R., Shen, J., Tang, Y., Fiskus, W., Nechtman, J., Atadja, P., and Bhalla, K. (2008) *Role of acetylation and extracellular location of heat shock protein 90alpha in tumor cell invasion.* Cancer Res. 68 (12), p. 4833-42.

656. Hou, T., Zheng, G., Zhang, P., Jia, J., Li, J., Xie, L., Wei, C., and Li, Y. (2014) *LAceP: lysine acetylation site prediction using logistic regression classifiers.* PLoS One. 9 (2), p. e89575.

657. Zhao, Y. and Chen, R.H. (2006) *Mps1 phosphorylation by MAP kinase is required for kinetochore localization of spindle-checkpoint proteins.* Curr Biol. 16 (17), p. 1764-9.

658. Dell, A., Galadari, A., Sastre, F., and Hitchen, P. (2010) *Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes.* Int J Microbiol. 2010, p. 148178.

659. Maley, F., Trimble, R.B., Tarentino, A.L., and Plummer, T.H., Jr. (1989) *Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases.* Anal Biochem. 180 (2), p. 195-204.

660. Liu, Y., Steiniger, S.C., Kim, Y., Kaufmann, G.F., Felding-Habermann, B., and Janda, K.D. (2007) *Mechanistic studies of a peptidic GRP78 ligand for cancer cell-specific drug delivery.* Mol Pharm. 4 (3), p. 435-47.

661. Obeid, M., Panaretakis, T., Joza, N., Tufi, R., Tesniere, A., van Endert, P., Zitvogel, L., and Kroemer, G. (2007) *Calreticulin exposure is required for the immunogenicity of gamma-irradiation and UVC light-induced apoptosis.* Cell Death Differ. 14 (10), p. 1848-50.

662. Popescu, N.I., Lupu, C., and Lupu, F. (2010) *Extracellular protein disulfide isomerase regulates coagulation on endothelial cells through modulation of phosphatidylserine exposure.* Blood. 116 (6), p. 993-1001.

663. Grunberg, J.R., Hammarstedt, A., Hedjazifar, S., and Smith, U. (2014) *The Novel Secreted Adipokine WNT1-inducible Signaling Pathway Protein 2 (WISP2) Is a Mesenchymal Cell Activator of Canonical WNT.* J Biol Chem. 289 (10), p. 6899-907.

664. Suzuki, E. and Nakayama, M. (2007) *The mammalian Ced-1 ortholog MEGF10/KIAA1780 displays a novel adhesion pattern.* Exp Cell Res. 313 (11), p. 2451-64.

665. Suzuki, E. and Nakayama, M. (2007) *MEGF10 is a mammalian ortholog of CED-1 that interacts with clathrin assembly protein complex 2 medium chain and induces large vacuole formation.* Exp Cell Res. 313 (17), p. 3729-42.

666. Ko, M.K. and Kay, E.P. (2004) *PDI-mediated ER retention and proteasomal degradation of procollagen I in corneal endothelial cells.* Exp Cell Res. 295 (1), p. 25-35.

667. Yamagata, M. and Kimata, K. (1994) *Repression of a malignant cell-substratum adhesion phenotype by inhibiting the production of the anti-adhesive proteoglycan, PG-M/versican.* J Cell Sci. 107 (Pt 9), p. 2581-90.

668. Touab, M., Villena, J., Barranco, C., Arumi-Uria, M., and Bassols, A. (2002) *Versican is differentially expressed in human melanoma and may play a role in tumor development.* Am J Pathol. 160 (2), p. 549-57.

669. Sakko, A.J., Ricciardelli, C., Mayne, K., Suwiwat, S., LeBaron, R.G., Marshall, V.R., Tilley, W.D., and Horsfall, D.J. (2003) *Modulation of prostate cancer cell attachment to matrix by versican.* Cancer Res. 63 (16), p. 4786-91.

670. Oulidi, A., Bokhobza, A., Gkika, D., Vanden Abeele, F., Lehen'kyi, V., Ouafik, L., Mauroy, B., and Prevarskaya, N. (2013) *TRPV2 mediates adrenomedullin stimulation of prostate and urothelial cancer cell adhesion, migration and invasion.* PLoS One. 8 (5), p. e64885.

671. Hammarstedt, A., Hedjazifar, S., Jenndahl, L., Gogg, S., Grunberg, J., Gustafson, B., Klimcakova, E., Stich, V., Langin, D., Laakso, M., and Smith, U. (2013) *WISP2 regulates preadipocyte commitment and PPARgamma activation by BMP4.* Proc Natl Acad Sci U S A. 110 (7), p. 2563-8.

672. Banerjee, S., Saxena, N., Sengupta, K., Tawfik, O., Mayo, M.S., and Banerjee, S.K. (2003) *WISP-2 gene in human breast cancer: estrogen and progesterone inducible expression and regulation of tumor cell proliferation.* Neoplasia. 5 (1), p. 63-73.

673. Fritah, A., Saucier, C., De Wever, O., Bracke, M., Bieche, I., Lidereau, R., Gespach, C., Drouot, S., Redeuilh, G., and Sabbah, M. (2008) *Role of WISP-2/CCN5 in the maintenance of a differentiated and noninvasive phenotype in human breast cancer cells.* Mol Cell Biol. 28 (3), p. 1114-23.

674. Ferrand, N., Gnanapragasam, A., Dorothee, G., Redeuilh, G., Larsen, A.K., and Sabbah, M. (2014) *Loss of WISP2/CCN5 in estrogen-dependent MCF7 human breast cancer cells promotes a stem-like cell phenotype.* PLoS One. 9 (2), p. e87878.

675. Huang, D. and Du, X. (2008) *Crosstalk between tumor cells and microenvironment via Wnt pathway in colorectal cancer dissemination.* World J Gastroenterol. 14 (12), p. 1823-7.

676. Luga, V. and Wrana, J.L. (2013) *Tumor-stroma interaction: Revealing fibroblast-secreted exosomes as potent regulators of Wnt-planar cell polarity signaling in cancer metastasis.* Cancer Res. 73 (23), p. 6843-7.

677. Mercurio, S., Latinkic, B., Itasaki, N., Krumlauf, R., and Smith, J.C. (2004) *Connective-tissue growth factor modulates WNT signalling and interacts with the WNT receptor complex.* Development. 131 (9), p. 2137-47.

678. Glackin, C.A. (2014) *Targeting the Twist and Wnt signaling pathways in metastatic breast cancer.* Maturitas. 79 (1), p. 48-51.

679. MacDonald, B.T., Hien, A., Zhang, X., Iranloye, O., Virshup, D.M., Waterman, M.L., and He, X. (2014) *Disulfide bond requirements for active Wnt ligands.* J Biol Chem. 289 (26), p. 18122-36.

680. Livingstone, C. (2013) *IGF2 and cancer.* Endocr Relat Cancer. 20 (6), p. R321-39.

681. Akkiprik, M., Feng, Y., Wang, H., Chen, K., Hu, L., Sahin, A., Krishnamurthy, S., Ozer, A., Hao, X., and Zhang, W. (2008) *Multifunctional roles of insulin-like growth factor binding protein 5 in breast cancer.* Breast Cancer Res. 10 (4), p. 212.

682. Hoeflich, A., Reisinger, R., Lahm, H., Kiess, W., Blum, W.F., Kolb, H.J., Weber, M.M., and Wolf, E. (2001) *Insulin-like growth factor-binding protein 2 in tumorigenesis: protector or promoter?* Cancer Res. 61 (24), p. 8601-10.

683. Werbajh, S.E., Urtreger, A.J., Puricelli, L.I., de Lustig, E.S., Bal de Kier Joffe, E., and Kornblihtt, A.R. (1998) *Downregulation of fibronectin transcription in highly metastatic adenocarcinoma cells.* FEBS Lett. 440 (3), p. 277-81.

684. Jia, D., Entersz, I., Butler, C., and Foty, R.A. (2012) *Fibronectin matrix-mediated cohesion suppresses invasion of prostate cancer cells.* BMC Cancer. 12, p. 94.

685. Urtreger, A.J., Werbajh, S.E., Verrecchia, F., Mauviel, A., Puricelli, L.I., Kornblihtt, A.R., and Bal de Kier Joffe, E.D. (2006) *Fibronectin is distinctly downregulated in murine mammary adenocarcinoma cells with high metastatic potential.* Oncol Rep. 16 (6), p. 1403-10.

686. Dano, K., Behrendt, N., Hoyer-Hansen, G., Johnsen, M., Lund, L.R., Ploug, M., and Romer, J. (2005) *Plasminogen activation and cancer.* Thromb Haemost. 93 (4), p. 676-81.

687. Czekay, R.P., Aertgeerts, K., Curriden, S.A., and Loskutoff, D.J. (2003) *Plasminogen activator inhibitor-1 detaches cells from extracellular matrices by inactivating integrins.* J Cell Biol. 160 (5), p. 781-91.

688. Czekay, R.P. and Loskutoff, D.J. (2009) *Plasminogen activator inhibitors regulate cell adhesion through a uPAR-dependent mechanism.* J Cell Physiol. 220 (3), p. 655-63.

689. Devy, L., Blacher, S., Grignet-Debrus, C., Bajou, K., Masson, V., Gerard, R.D., Gils, A., Carmeliet, G., Carmeliet, P., Declerck, P.J., Noel, A., and Foidart, J.M. (2002) *The pro- or antiangiogenic effect of plasminogen activator inhibitor 1 is dose dependent.* FASEB J. 16 (2), p. 147-54.

690. McMahon, G.A., Petitclerc, E., Stefansson, S., Smith, E., Wong, M.K., Westrick, R.J., Ginsburg, D., Brooks, P.C., and Lawrence, D.A. (2001) *Plasminogen activator inhibitor-1 regulates tumor growth and angiogenesis.* J Biol Chem. 276 (36), p. 33964-8.

691. Brodsky, S.V., Malinowski, K., Golightly, M., Jesty, J., and Goligorsky, M.S. (2002) *Plasminogen activator inhibitor-1 promotes formation of endothelial microparticles with procoagulant potential.* Circulation. 106 (18), p. 2372-8.

692. Hewett, P.W. and Murray, C. (1996) *Modulation of human endothelial cell procoagulant activity in tumour models in vitro.* Int J Cancer. 66 (6), p. 784-9.

693. Im, J.H., Fu, W., Wang, H., Bhatia, S.K., Hammer, D.A., Kowalska, M.A., and Muschel, R.J. (2004) *Coagulation facilitates tumor cell spreading in the pulmonary vasculature during early metastatic colony formation.* Cancer Res. 64 (23), p. 8613-9.

694. Langer, F., Amirkhosravi, A., Ingersoll, S.B., Walker, J.M., Spath, B., Eifrig, B., Bokemeyer, C., and Francis, J.L. (2006) *Experimental metastasis and primary tumor growth in mice with hemophilia A.* J Thromb Haemost. 4 (5), p. 1056-62.

695. Rauschert, N., Brandlein, S., Holzinger, E., Hensel, F., Muller-Hermelink, H.K., and Vollmers, H.P. (2008) *A new tumor-specific variant of GRP78 as target for antibody-based therapy.* Lab Invest. 88 (4), p. 375-86.

696. Sannino, S., Anelli, T., Cortini, M., Masui, S., Degano, M., Fagioli, C., Inaba, K., and Sitia, R. (2014) *Progressive quality control of secretory proteins in the early secretory compartment by ERp44.* J Cell Sci.

697. Zhang, Y., Liu, R., Ni, M., Gill, P., and Lee, A.S. (2010) *Cell surface relocalization of the endoplasmic reticulum chaperone and unfolded protein response regulator GRP78/BiP.* J Biol Chem. 285 (20), p. 15065-75.

698. Ni, M., Zhang, Y., and Lee, A.S. (2011) *Beyond the endoplasmic reticulum: atypical GRP78 in cell viability, signalling and therapeutic targeting.* Biochem J. 434 (2), p. 181-8.

699. Burikhanov, R., Zhao, Y., Goswami, A., Qiu, S., Schwarze, S.R., and Rangnekar, V.M. (2009) *The tumor suppressor Par-4 activates an extrinsic pathway for apoptosis.* Cell. 138 (2), p. 377-88.

700. Misra, U.K., Gonzalez-Gronow, M., Gawdi, G., and Pizzo, S.V. (2005) *The role of MTJ-1 in cell surface translocation of GRP78, a receptor for alpha 2-macroglobulin-dependent signaling.* J Immunol. 174 (4), p. 2092-7.

701. Sargeant, D.P., Gryk, M.R., Maciejewski, M.W., Thapar, V., Kundeti, V., Rajasekaran, S., Romero, P., Dunker, K., Li, S.C., Kaneko, T., and Schiller, M.R. (2012) *Secondary structure, a missing component of sequence-based minimotif definitions.* PLoS One. 7 (12), p. e49957.

702. Shental-Bechor, D. and Levy, Y. (2008) *Effect of glycosylation on protein folding: a close look at thermodynamic stabilization.* Proc Natl Acad Sci U S A. 105 (24), p. 8256-61.

703. Rudd, P.M., Joao, H.C., Coghill, E., Fiten, P., Saunders, M.R., Opdenakker, G., and Dwek, R.A. (1994) *Glycoforms modify the dynamic stability and functional activity of an enzyme.* Biochemistry. 33 (1), p. 17-22.

704. Srimathi, S. and Jayaraman, G. (2005) *Effect of glycosylation on the catalytic and conformational stability of homologous alpha-amylases.* Protein J. 24 (2), p. 79-88.

705. Yoshimasu, M.A., Tanaka, T., Ahn, J.K., and Yada, R.Y. (2004) *Effect of N-linked glycosylation on the aspartic proteinase porcine pepsin expressed from Pichia pastoris.* Glycobiology. 14 (5), p. 417-29.

706. Batista, U., Garvas, M., Nemec, M., Schara, M., Veranic, P., and Koklic, T. (2010) *Effects of different detachment procedures on viability, nitroxide reduction kinetics and plasma membrane heterogeneity of V-79 cells.* Cell Biol Int. 34 (6), p. 663-8.

707. Lamb, J.F. and Ogden, P.H. (1987) *Transient changes in permeability in HeLa and L cells during detachment from a substrate.* Q J Exp Physiol. 72 (2), p. 189-99.

708. Brown, M., Stafford, L.J., Onisk, D., Joaquim, T., Tobb, A., Goldman, L., Fancy, D., Stave, J., and Chambers, R. (2013) *Snorkel: an epitope tagging system for measuring the surface expression of membrane proteins.* PLoS One. 8 (9), p. e73255.

709. Dey, S., Zhan, S., and Matsunami, H. (2011) *Assaying surface expression of chemosensory receptors in heterologous cells.* J Vis Exp. (48).

710. Shepard, B.D., Natarajan, N., Protzko, R.J., Acres, O.W., and Pluznick, J.L. (2013) *A cleavable N-terminal signal peptide promotes widespread olfactory receptor surface expression in HEK293T cells.* PLoS One. 8 (7), p. e68758.

711. Koch, A.E., Kunkel, S.L., Shah, M.R., Hosaka, S., Halloran, M.M., Haines, G.K., Burdick, M.D., Pope, R.M., and Strieter, R.M. (1995) *Growth-related gene product alpha. A chemotactic cytokine for neutrophils in rheumatoid arthritis.* J Immunol. 155 (7), p. 3660-6.

712. Rollins, B.J. (1997) *Chemokines.* Blood. 90 (3), p. 909-28.

713. Bacon, K.B. and Oppenheim, J.J. (1998) *Chemokines in disease models and pathogenesis.* Cytokine Growth Factor Rev. 9 (2), p. 167-73.

714. Delano, M.J., Kelly-Scumpia, K.M., Thayer, T.C., Winfield, R.D., Scumpia, P.O., Cuenca, A.G., Harrington, P.B., O'Malley, K.A., Warner, E., Gabrilovich, S., Mathews, C.E., Laface, D., Heyworth, P.G., Ramphal, R., Strieter, R.M., Moldawer, L.L., and Efron, P.A. (2012) *Neutrophil mobilization from the bone marrow during polymicrobial sepsis is dependent on CXCL12 signaling.* J Immunol. 187 (2), p. 911-8.

715. Meiron, M., Zohar, Y., Anunu, R., Wildbaum, G., and Karin, N. (2008) *CXCL12 (SDF-1alpha) suppresses ongoing experimental autoimmune encephalomyelitis by selecting antigen-specific regulatory T cells.* J Exp Med. 205 (11), p. 2643-55.

716. Cabrera, T., Angustias Fernandez, M., Sierra, A., Garrido, A., Herruzo, A., Escobedo, A., Fabra, A., and Garrido, F. (1996) *High frequency of altered HLA class I phenotypes in invasive breast carcinomas.* Hum Immunol. 50 (2), p. 127-34.

717. Palmisano, G.L., Pistillo, M.P., Capanni, P., Pera, C., Nicolo, G., Salvi, S., Perdelli, L., Pasciucco, G., and Ferrara, G.B. (2001) *Investigation of HLA class I downregulation in breast cancer by RT-PCR.* Hum Immunol. 62 (2), p. 133-9.

718. Bubenik, J. (2004) *MHC class I down-regulation: tumour escape from immune surveillance? (review).* Int J Oncol. 25 (2), p. 487-91.

719. Mirisola, V., Zuccarino, A., Bachmeier, B.E., Sormani, M.P., Falter, J., Nerlich, A., and Pfeffer, U. (2009) *CXCL12/SDF1 expression by breast cancers is an independent prognostic marker of disease-free and overall survival.* Eur J Cancer. 45 (14), p. 2579-87.

720. Moll, N.M. and Ransohoff, R.M. (2010) *CXCL12 and CXCR4 in bone marrow physiology.* Expert Rev Hematol. 3 (3), p. 315-22.

721. Raz, V., Sterrenburg, E., Routledge, S., Venema, A., van der Sluijs, B.M., Trollet, C., Dickson, G., van Engelen, B.G., van der Maarel, S.M., and Antoniou, M.N. (2013) *Nuclear entrapment and extracellular depletion of PCOLCE is associated with muscle degeneration in oculopharyngeal muscular dystrophy.* BMC Neurol. 13, p. 70.

722. Lu, P., Takai, K., Weaver, V.M., and Werb, Z. (2011) *Extracellular matrix degradation and remodeling in development and disease.* Cold Spring Harb Perspect Biol. 3 (12).

723. Cheng, W.a.A., D. (2012) *A subset of co-expressed genes in Slug-based cancer mesenchymal transition signature remains coexpressed in normal samples in a tissue-specific manner.* Available from Nature Precedings <http://hdl.handle.net/10101/npre.2012.6813.1>.

724. Park, J., Morley, T.S., and Scherer, P.E. (2012) *Inhibition of endotrophin, a cleavage product of collagen VI, confers cisplatin sensitivity to tumours.* EMBO Mol Med. 5 (6), p. 935-48.

725. You, W.K., Bonaldo, P., and Stallcup, W.B. (2012) *Collagen VI ablation retards brain tumor progression due to deficits in assembly of the vascular basal lamina.* Am J Pathol. 180 (3), p. 1145-58.

726. Chen, P., Cescon, M., and Bonaldo, P. (2013) *Collagen VI in cancer and its biological mechanisms.* Trends Mol Med. 19 (7), p. 410-7.

727. Djonov, V., Andres, A.C., and Ziemiecki, A. (2001) *Vascular remodelling during the normal and malignant life cycle of the mammary gland.* Microsc Res Tech. 52 (2), p. 182-9.

728. Gouon-Evans, V., Lin, E.Y., and Pollard, J.W. (2002) *Requirement of macrophages and eosinophils and their cytokines/chemokines for mammary gland development.* Breast Cancer Res. 4 (4), p. 155-64.

729. Lee, H.J., Hong, C.Y., Jin, C.J., Kim, M.H., Lee, Y.K., Nguyen-Pham, T.N., Lee, H., Park, B.C., Chung, I.J., Kim, H.J., and Lee, J.J. (2012) *Identification of novel HLA-A*0201-restricted epitopes from anterior gradient-2 as a tumor-associated antigen against colorectal cancer.* Cell Mol Immunol. 9 (2), p. 175-83.

730. Lee, H.J., Hong, C.Y., Kim, M.H., Lee, Y.K., Nguyen-Pham, T.N., Park, B.C., Yang, D.H., Chung, I.J., Kim, H.J., and Lee, J.J. (2012) *In vitro induction of anterior gradient-2-specific cytotoxic T lymphocytes by dendritic cells transduced with recombinant adenoviruses as a potential therapy for colorectal cancer.* Exp Mol Med. 44 (1), p. 60-7.

731. Scott, M.S., Boisvert, F.M., McDowall, M.D., Lamond, A.I., and Barton, G.J. (2010) *Characterization and prediction of protein nucleolar localization sequences.* Nucleic Acids Res. 38 (21), p. 7388-99.

732. Tina, K.G., Bhadra, R., and Srinivasan, N. (2007) *PIC: Protein Interactions Calculator.* Nucleic Acids Res. 35 (Web Server issue), p. W473-6.

733. McDonald, I.K. and Thornton, J.M. (1994) *Satisfying hydrogen bonding potential in proteins.* J Mol Biol. 238 (5), p. 777-93.

734. Blom, N., Gammeltoft, S., and Brunak, S. (1999) *Sequence and structure-based prediction of eukaryotic protein phosphorylation sites.* J Mol Biol. 294 (5), p. 1351-62.

735. Steentoft, C., Vakhrushev, S.Y., Joshi, H.J., Kong, Y., Vester-Christensen, M.B., Schjoldager, K.T., Lavrsen, K., Dabelsteen, S., Pedersen, N.B., Marcos-Silva, L., Gupta, R., Bennett, E.P., Mandel, U., Brunak, S., Wandall, H.H., Levery, S.B., and Clausen, H. (2013) *Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology.* EMBO J. 32 (10), p. 1478-88.

Websites

www.lifetechnologies.com/oligoperfect/

www.neb.com

web.expasy.org/protparam/

www.systembio.com