Oncogene (2013) **32,** 2499–2509

www.nature.com/onc

REVIEW

Emerging roles for the pro-oncogenic anterior gradient-2 in cancer development

E Chevet[1,2], D Fessart[3,4], F Delom[3,4], A Mulot[1,2], B Vojtesek[5], R Hrstka[5], E Murray[5,6], T Gray[6] and T Hupp[6]

Clinical studies have defined the core 'genetic blueprint' of a cancer cell, but this information does not necessarily predict the cancer phenotype. Signalling hubs that mediate such phenotype have been identified largely using OMICS platforms that measure dynamic molecular changes within the cancer cell landscape. The pro-oncogenic protein anterior gradient 2 (AGR2) is a case in point; AGR2 has been shown using a range of expression platforms to be involved in asthma, inflammatory bowel disease, cell transformation, cancer drug resistance and metastatic growth. AGR2 protein is also highly overexpressed in a diverse range of human cancers and can be secreted and detected in extracellular fluids, thus representing a compelling pro-oncogenic signalling intermediate in human cancer. AGR2 belongs to the protein disulphide isomerase family with all the key features of an endoplasmic reticulum-resident protein—this gives clues into how it might function as an oncoprotein through the regulation of protein folding, maturation and secretion that can drive metastatic cell growth. In this review, we will describe the known aspects of AGR2 molecular biology, including gene structure and regulation, emerging protein interaction networks and how its subcellular localization mediates its biological functions. We will finally review the cases of AGR2 expression in human cancers, the pathophysiological consequences of AGR2 overexpression, its potential role as a tumour biomarker that predicts the response to therapy and how the AGR2 pathway might form the basis for drug discovery programmes aimed at targeting protein folding/maturation pathways that mediate secretion and metastasis.

Oncogene (2013) **32,** 2499–2509; doi:10.1038/onc.2012.346; published online 3 September 2012

Keywords: AGR2; PDI; endoplasmic reticulum; secretome; p53; tumour

INTRODUCTION

In recent years, an increasing number of reports have pointed towards intracellular organelles as playing major roles in cancer development and progression. This is particularly true for the endoplasmic reticulum (ER), which has appeared as a major player in early carcinogenesis and in tumour growth.[1,2] The ER is the first compartment of the secretory pathway controlling the cell's calcium fluxes, lipid homeostasis and the biogenesis of transmembrane and secretory proteins.[3] The latter is achieved through the coordinated action of several molecular machines, including protein synthesis and translocation, protein folding, export, degradation and ER stress signalling (Figure 1). Genes/proteins that constitute these molecular machines have been shown to significantly contribute to the development and progression of cancers.[2,4] This is also true for oxidative protein folding components, which are altered in various cancers, thereby leading to cell damages and altered protein secretion that contributes to adaptation to the local tumour micro-environment, tissue remodelling and metastatic cell growth. Major actors in oxidative protein folding comprise the family of protein disulphide isomerases (PDIs; Benham[5]), which is characterized by the presence of the thioredoxin motif CXXC. The roles played by PDIs in cancer remain unclear. Although few somatic mutations were found in those genes,[6] changes in their expression landscape were identified from proteomics screens. Recently, PDI-like proteins have been identified that present incomplete or evolutionarily divergent CXXC motifs,[7] but their molecular functions remain enigmatic. Anterior gradient 2 belongs to this category (AGR2). In the present review, we provide an in-depth description of AGR2 in the context of the AGR family, the previously reported functions of this protein and how the AGR2 pathway might be targeted to develop new therapeutic strategies to inhibit cancer growth.

THE AGR GENE FAMILY

AGR2

The AGR family is composed of three proteins, including AGR1, AGR2 and AGR3. *AGR2* gene was first identified in *Xenopus laevis*. *Xenopus* anterior gradient-1 and -2 (XAG-1 and XAG-2) comprise one class of cement gland-specific genes based on their spatial expression patterns.[8,9] Studies showed that XAG-2 function in specifying the fate of the dorso-anterior ectoderm that forms the cement gland and in inducing the forebrain's fate. During early development, XAG-1 and XAG-2 are expressed within ectoderm-derived organs like the cement gland, which is a mucus-secreting organ located at the extreme anterior end of the frog's head.[10] The cement gland forms a cone of columnar epithelium and is

[1]INSERM U1053, Bordeaux, France; [2]Université Bordeaux-Segalen, Bordeaux, France; [3]Centre de recherche Cardio-Thoracique de Bordeaux, Université Bordeaux-Segalen, Bordeaux, France; [4]INSERM, Centre de recherche Cardio-Thoracique de Bordeaux, Bordeaux, France; [5]Regional Centre for Applied Molecular Oncology, Masaryk Memorial Cancer Institute, Brno, Czech Republic and [6]p53 Signal Transduction Laboratories, Cell Signalling Unit, Institute of Genetics and Molecular Medicine, Cancer Research Centre, University of Edinburgh, Edinburgh, UK. Correspondence: Dr E Chevet, INSERM U1053, Bordeaux, France or Université Bordeaux-Segalen, Bordeaux, France or Dr B Vojtesek, Masaryk Memorial Cancer Institute, Brno, Czech Republic or Professor T Hupp, p53 Signal Transduction Laboratories, Cell Signalling Unit, Institute of Genetics and Molecular Medicine, Cancer Research Centre, University of Edinburgh, Crewe Road South, Edinburgh EH4 2XR, UK.
E-mail: eric.chevet@inserm.fr or vojtesek@mou.cz or ted.hupp@ed.ac.uk
Received 13 April 2012; revised 20 June 2012; accepted 26 June 2012; published online 3 September 2012

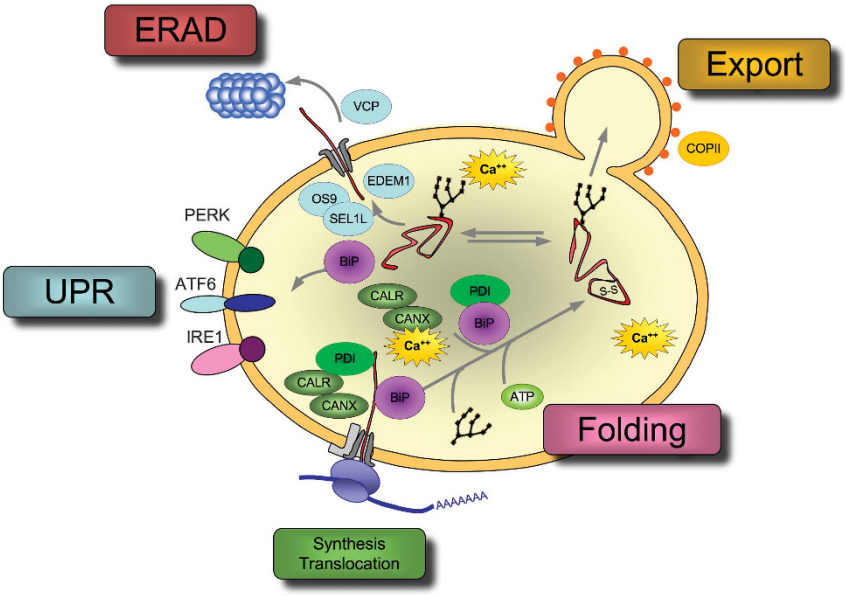

Figure 1. ER molecular machines that control protein homeostasis. Five major molecular machines, namely translation/translocation (green), folding (pink), export (yellow), degradation (ER-associated degradation, red) and ER stress signalling (blue) control ER protein homeostasis.[1]

involved in the attachment of the frog embryo to a solid support.[11] In mouse embryos, *XAG-2* homologous gene (Gob4/MAgr2) was also identified as a highly expressed mRNA biomarker of murine gut.[12] Mouse AGR2 is localized in mucus-secreting cells in these tissues by *in situ* hybridization. The majority of proteomics or transcriptomics screens identified AGR2 as a dominant outlier and hence most research was focused on validating this protein; however, the other homologues like AGR3 or ERp29 are beginning to be identified in other systems.

AGR3

AGR3 was first identified in a proteomic analysis of purified membrane preparations from multiple human breast tumour-derived cell lines in efforts to identify oncogenic membrane signalling proteins. Adam and co-workers[13–15] identified a unique protein BCMP11 that was homologous to AGR2 and whose expression correlates with oestrogen receptor (ER) in human breast cancers. In view of their high degree of sequence identity (71%), BCMP11 was named AGR3. The function of AGR3 is not known, but AGR3 has since been identified to be expressed in ER-negative human ovarian cancers[16] and to mediate cisplatin resistance in tumour xenograft studies. In the latter study, it is interesting to note that many ovarian cancers that express AGR3 can be AGR2-negative, indicating that their expression can be uncoupled. The functions of the two genes might not be redundant, as AGR2 and AGR3 do not similarly colocalize to the ER[16] and AGR3 has generally not been identified in the many OMICS screens in human cancer, suggesting a more critical role for AGR2 in carcinogenesis. However, a recent study using exome sequencing to identify the mutant cancer gene landscape in cancer has indicated that the *AGR3* gene is mutated in approximately 10% of kidney tumours examined,[17] suggesting that AGR3 perturbation can also contribute to cancer progression.

AGR1

The third member of the AGR family was highlighted in a bioinformatics study[18] that identified AGR2 and AGR3 as highly similar to ERP18/19, also named AGR1, which could be considered the founding gene of the AGR family given its higher similarity to the classic thioredoxin fold containing the CXXC motifs and its presence in invertebrates.[19,20] The function of ERP18 involves its ability to form mixed disulphides with client proteins in the ER;[21] however, ERP18 has not been identified in OMICS screens in human cancer, thus its role in human cancer remains unknown.

Future studies on dissecting the role of the AGR family in biology and disease: genetic models that define CXXC/S motif protein function

A key missing methodology thus far in analysing the AGR-1/2/3 protein family is the lack of powerful animal models to dissect signalling *in vivo* using classic genetics. This is mainly due to the fact that, although the three genes belong to the same family (Figure 2), the *AGR2* and *AGR3* genes emerged during the evolution of chordates and thus far appear confined to vertebrates, whereas AGR1/ERP18 seems to be the founding member of this trio and is found within complex invertebrates, like *Caenorhabditis elegans*. There has been no expansion of the *AGR2* and *AGR3* genes as their appearance in vertebrates, suggests that strict selection pressures exist in maintaining essential functions during evolution by natural selection. Similar selection pressures have apparently maintained the closest orthologues of the AGR2 family, TRX1 and TRX2 (Figure 2). Other non-catalytic PDIs have also been identified as ERp44 (PDIA10) and ERp29, 27 (PDIA9, PDIA8), and as reported for the AGR family their functions remain poorly understood.[7,22,23] Although these proteins can form mixed disulphide bonds with their clients, their functions in the control of cell's secretome still remains equivocal. One research avenue to shed light on the fundamental role of the CXXC/S PDI family would be to use animal models such as *C. elegans* to dissect genetically protein folding pathways. For example, if conditional or constitutive phenotypes could be identified for an AGR1/ERP18 allele in *C. elegans*, then suppressor screens could be initiated to define genetic pathways important to overcome such bottlenecks. In addition, in vertebrates the *AGR2* and *AGR3* gene functions could be similarly dissected using the powerful gene knock-in TALEN technologies in zebrafish and in murine systems to evaluate specific motif mutations on biological signalling *in vivo*. The development of genetic models in vertebrates, for example,

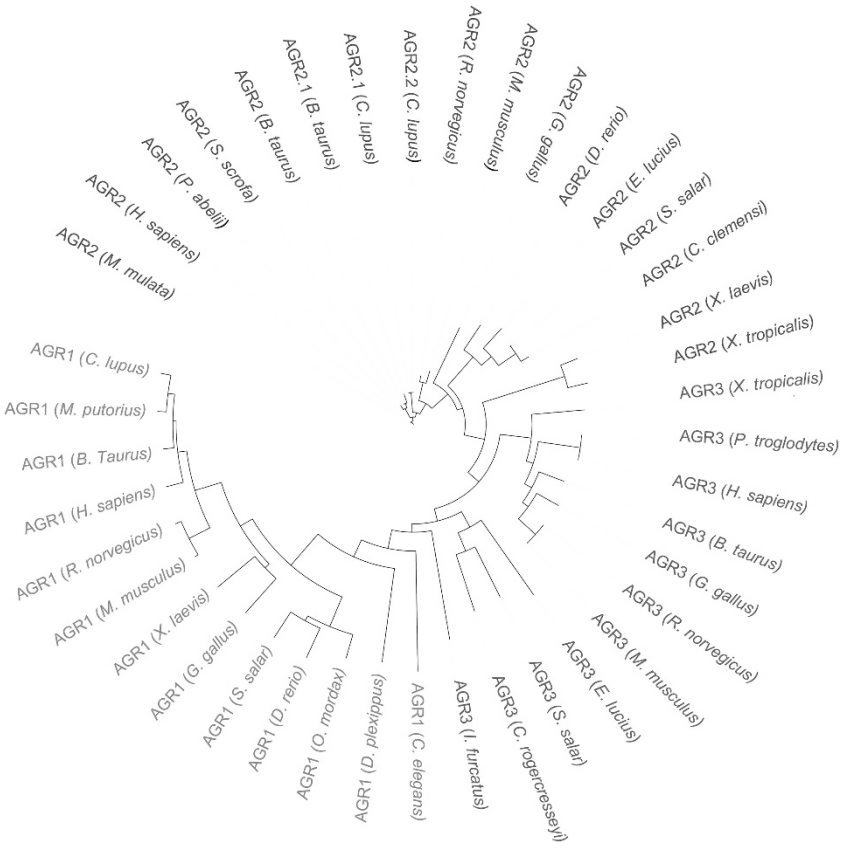

Figure 2. The AGR family tree. Protein sequences from 41 AGR family members were aligned using ClustalW and the corresponding matrix represented as a tree using the iTOL website (http://itol.embl.de/). Blue: AGR2; green: AGR3; red: AGR1/ERP18. A full colour version of this figure is available at the *Oncogene* journal online.

creating a CXXS to CXXC (or SXXS) allele conversion of AGR2, would add significant understanding of fundamental aspect in the CXXC/S family.

AGR2: GENE ARCHITECTURE AND EXPRESSION

Genomic structure

The detailed genomic structure for human AGR2 was determined on the basis of its mRNA sequences and genomic DNA clone by using the Spidey program.[24] The *AGR2* gene spans a region of 50 kb in genomic DNA. Two transcripts were also detected and both mRNA isoforms contained eight exons and seven introns. Both cDNA clones of the two mRNA isoforms were fully sequenced and were found to contain the same open-reading frame. The putative translation start site for AGR2 contained the KOZAK consensus sequence.[24] In their study, they reviewed the public databases and the Celera Discovery Systems to establish a gene model for AGR2, and annotated an additional 5′ exon, confirmed by reverse transcription–polymerase chain reaction in a tissue panel.[25] Moreover, using radiation hybrid analyses and FISH, Petek and co-workers[26] mapped the gene encoding human AGR2 to chromosome 7p21.3.

Developmental regulation of the AGR2 promoter

In situ hybridization and reverse transcription-polymerase chain reaction analyses revealed that *AGR2* mRNA is strongly expressed in endoderm-derived organs that contain mucus-secreting cells. Zheng and co-workers[25] also investigated the promoter activity of the human *AGR2* gene independent of oestrogen responsiveness. Using a luciferase reporter gene construct driven by the *AGR2* promoter, they observed that co-transfection with the forkhead box transcription factors FOXA1 and FOXA2, which have been implicated in maintaining goblet cell function, led to a significantly increased luciferase activity in HEK293 cells. In addition, Zheng and co-workers[25] described that there is a binding site for hepatic nuclear factor 1, which belongs to the same family as FOXA1 and FOXA2, in the *AGR2* promoter region at SNP 07AGRNP53. Furthermore, using a global chromatin immunoprecipitation approach in three tumour cell lines (MCF7, SW480 and Ntera2), Krig and co-workers[27] identified *AGR2* promoter as a binding site MCF7-specific to the oncogene ZNF217. Finally, using the UCSC genome database (http://genome.ucsc.edu), relevant NF-κB and SOX9 binding sites were found on the *AGR2* promoter.

AGR2 promoter regulation by oestrogen

Historically, using the technique of suppression subtractive hybridization as a forerunner to transcriptomics, the Weigel laboratory reported the identification of 29 gene fragments that were expressed in the ER-positive MCF7 breast carcinoma cell line that were absent or minimally expressed in the ER-negative MDA-MB-231 breast carcinoma cell line.[28] One of these gene fragments, DEME2 (GenBank EST accession no. AA506763), exhibited an expression pattern that correlated with the expression of ER in a panel of eight breast carcinoma cell lines. DEME-2 was named hAG-2 (for example, AGR2) based on sequence similarity. AGR2 was significantly upregulated in response to oestradiol. Indeed, transplantation of epithelium and stroma of normal breast tissue

into female nude mouse treated with oestradiol revealed significant upregulation of *AGR2* mRNA upon oestradiol treatment.[29]

Regulation of AGR2 promoter by androgens

Zhang and co-workers[24] have reported that *AGR2* mRNA is androgen inducible. This induction is time and dose dependent, with a maximum fold increase more than a 10-fold in the level of *AGR2* mRNA, and requires the presence of the androgen receptor. Also, this induction was completely abolished by cycloheximide, suggesting that newly synthesized proteins play a role in the induction process. In this study, they also determined the effects of various agents, including protein kinase inhibitors and activators, on *AGR2* induction.

Regulation of the AGR2 promoter by (patho)physiological stress conditions

When the breast cancer cell line MDA-MB-231 was subjected to serum depletion alone or combined with oxygen depletion to mimic conditions reflecting tumour cell-associated stress, Zweitzig and co-workers[30] found that *AGR2* mRNA is induced about fivefold. Specific inhibitors for ERK1/2, JNK, p38 and PI3K were tested in this model and only ERK1/2 inhibition was associated with a block of *AGR2* mRNA induction.[30] It was demonstrated that tunicamycin, an inhibitor of N-linked glycosylation, promoted the induction of *AGR2* mRNA expression.[24] *AGR2* mRNA was indeed induced upon ER stress using various inducers such as tunicamycin, the reducing agent dithiothreitol or the sarcoplasmic/endoplasmic reticulum calcium pump inhibitor thapsigargin.[31] Moreover, ER-stress-mediated upregulation of *AGR2* mRNA mainly resulted from activation of the ATF6 and IRE1 arms of the unfolded protein response. This was demonstrated using small interfering RNA-mediated silencing of PERK, ATF6 or IRE1 and measuring *AGR2* mRNA expression upon ER stress, thereby suggesting the presence of an ERSE-like element on its promoter.[31] Thus far, the role of ATF6 and IRE1 was not investigated on *AGR1* or *AGR3* mRNA expression.

Future research areas on AGR family gene expression: biochemical dissection of the AGR transcriptional machine

Although clinical studies have shown that AGR2 and AGR3 are expressed *in vivo* within human cancers by both hormone-dependent and hormone-independent pathways, we know very little about the architecture of the dynamic multi-protein complexes that drive these promoters. This is a very compelling area of research, as AGR2 is induced by the anti-oestrogen Tamoxifen and thus provides a mechanism to account for induced or acquired resistance to this drug.[32] AGR2 is the only compelling Tamoxifen-induced gene with a pedigree in diverse biological systems, including colitis, limb regeneration, metastatic growth and p53 inhibition. Thus, identifying the components that co-ordinate the Tamoxifen signal provide potential targets for sensitizing AGR2-positive Tamoxifen-resistant cancers. A more fundamental question also remains provocative—as the *AGR2* and *AGR3* genes are contiguous on the same chromosome, but are often expressed in a mutually exclusive manner;[16] what signalling pathways and multi-protein complex assemblies mediate this selective promoter usage *in vivo*?

CORE BIOCHEMICAL FUNCTIONS OF AGR2

The ER landscape of AGR2: intracellular biochemical function within the ER

One of the most conspicuous motifs in AGR2 is the thioredoxin fold that was initially identified by Persson and co-workers.[18] Although the thioredoxin activity of AGR2 has never been demonstrated either *in vitro* or *in vivo*, the potential implication of this motif in AGR2 biology has been suggested in some instances.[16,33] It is curious that the apparent 'founding' gene in this family (ERP18) contains the classic 'thioredoxin fold (W-CXXC-K)', but has not been an outlier OMICS biomarker in the many screens reviewed in this article. Rather, AGR2 contains the divergent thioredoxin fold E/D-CXXS-Q that has been implicated in disease development. Thus, the role of the CXXS-containing PDIs in vertebrate biology will likely be a key focus in understanding the protein folding function of this family and how this effects disease progression. In the case of ERP18, synthetic mutants containing a mutation of the CXXC motif to CXXS resulting in the covalent trapping of dithiothreitol-sensitive covalent intermediates[21] and similar approaches might be used to trap physiological substrates in other CXXS-containing PDIs. AGR2 contains putative CXXS TX-domain motifs also found in two other PDIs/ERps, Eug1p and ERP44.[22,34]

Functional aspects for the CXXS motif proteins have been suggested, such as interactions with intermediates in redox reactions during folding[22] and retrograde transport in the secretory pathway.[35] The CXXS motif possesses lower activity associated with disulphide bond reorganization, but may, however, contribute to isomerization of already existing disulphide bridges and possibly perform specialized functions in the ER.[35,36] A cysteine residue within this domain forms mixed disulphide bonds with mucin-2, indicating a direct role for AGR2 in mucin processing.[33] Mice lacking AGR2 were viable, but were highly susceptible to colitis, indicating a critical role for AGR2 in protection from disease.[33] Another study reported that AGR2$^{-/-}$ intestine has also decreased goblet cell mucin-2, dramatic expansion of the Paneth cell compartment, abnormal Paneth cell localization, elevated ER stress, severe terminal ileitis and colitis.[37] Another study also highlighted the ability of allergens to induce AGR2 and mucin-5 secretion; this secretion of mucin-5AC/B is attenuated in AGR2$^{-/-}$ cells.[38] The role of AGR2 in the development of airway goblet cells was confirmed and was demonstrated to be under the control of the transcription factors Foxp1 and Foxp4, most likely to achieve mucus secretion.[39] Finally, ER functions of AGR2 were shown to contribute to this compartment's homeostasis through the signalling of the unfolded protein response[31] (Figure 5).

The intracellular and extracellular landscape of AGR2

The 'degenerate ER retrieval motif' of AGR2: AGR2 protein harbours a canonical cleavable signal sequence for targeting into the secretory pathway in addition to the thioredoxin fold, and a non-optimal ER retention motif at its C terminus (KTEL) (Figure 3). This is likely to have a significant impact on how AGR2 protein is trafficked to different compartments and functions in cells. In addition, the ability of AGR2 to mediate protein folding of pro-metastatic secretory molecules is also likely to be regulated by its residence time in the ER. Although the *Xenopus* homologue AGR2 protein (XAG-2) was shown to be secreted,[10] it appears that hAGR2 can be found in various cellular compartments, including the ER,[31] the nucleus,[40] the cell surface and the extracellular milieu.[41]The majority of PDIs/ERPs harbours a typical H/KDEL ER retrieval signal and localizes to the ER. Both AGR2 and AGR3 ER retrieval signals, and although atypical, were able to bind to KDEL

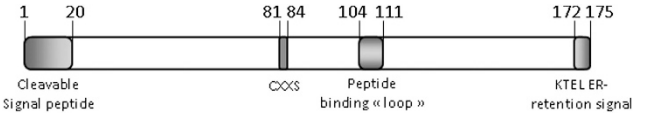

Figure 3. AGR2 protein. Primary structure of the AGR2 protein. The identified functional domains and amino acids implicated in regulation of its function are indicated by grey shaded boxes.

receptors to be retrieved in the ER.[42] It was also shown that AGR2 can escape the ER retrieval machinery and be secreted to play an autocrine/paracrine role.[41,43] Subcellular fractionation approaches revealed that AGR2 can also localize to the nucleus.[44]

Altogether, these observations lead to the speculation that the divergence of the highly conserved KTEL sequence of AGR2 from the canonical KDEL may lower the affinity of AGR2 to the KDEL receptor protein and allow AGR2 to exit the ER at a higher rate when overexpressed. Although mainly localised in the ER, AGR2 was reported to play other intracellular roles. For instance, AGR2 was shown to regulate p53 signalling,[45] induce the EGFR ligand amphiregulin (AREG)[46] and interact with the AAA$^+$ protein Reptin,[47] thereby suggesting regulatory functions in gene expression. This part of AGR2 biology remains to be further studied to understand (1) how AGR2 can localize to the cytosol and (2) what could be its true functions in this compartment (Figure 5). However, these results may help to explain the observations that show that AGR2's ability to induce the expression of either CDX2 (transcription factor) or AREG is lost when either the C terminus is deleted or mutated to KDEL.[40] Taken together, these data suggest that AGR2's tumour-promoting ability could be due to the combined action of its activity within the cell and that achieved as a secreted protein.

Because of the similarity between AGR2 with the XAG proteins, it may be speculated that AGR2 may be involved in the proliferation of mammalian tissues. Whereas AGR2 is currently not embedded in a known pathway, AGR2 can mediate metastasis in animal models. Indeed, when an expression vector for AGR2 cDNA was transfected into benign non-metastatic rat mammary tumour cells (Rama37), metastases occurred in the lungs of animals receiving the AGR2-transfected cells in 77–92% of animals with primary tumours compared with no metastases in the control groups. AGR2-transfected cells exhibited enhanced rates of adhesion to a plastic substratum.[48] This observation was recently substantiated by a study showing that AGR2 was a cell surface antigen-promoting tumour cell dissemination through the activation of cathepsins B and D.[41] The extracellular role of AGR2 has been clearly illustrated by nAG, a new member of the anterior gradient family of proteins that was shown to be important in the regeneration of amputated limbs in amphibians.[49] Artificial reintroduction of nAG by electroporation is sufficient to rescue regeneration in denervated salamander limbs. It is suggested that nAG is secreted by the nerve sheath cells and then acts directly on limb blasternal cells. In this study, a potential nAG-binding partner Prod1 was identified by yeast two hybrid, although the protein–protein interactions between these two proteins have not been validated in cell systems. This is a glycosylphosphatidylinisotol-anchored protein of the Ly6 superfamily as is the C4.4 protein identified in the yeast two-hybrid screen carried out against human AGR2 and that has been proposed to serve as AGR2 receptor,[49] although this has not been validated at the molecular level. The mechanisms whereby AGR2 can stimulate tissue regrowth might be due to its regulation of the adhesion and migration pathways required to remodel cells into a new environment. Altogether, evidences for AGR2 binding to cell surface proteins and for its presence in the extracellular milieu suggest a role for this protein outside the cell (Figure 5).

The molecular chaperone landscape of AGR2-emerging protein–protein interactions

The 'substrate-binding loop' of AGR2. Although AGR2 protein forms mixed disulphides with the mucin family and has a specific interaction with the AAA$^+$ protein Reptin, we are only beginning to learn about the functional motifs in the protein that regulate protein–protein interactions. The founder gene *AGR1* (ERP18 (PDB code 2K8V)) shows no evidence of being monomeric, while our unpublished data demonstrate that AGR2 forms homodimers in solution and in cells indicating that biochemical models of AGR2 function will likely incorporate a dimeric structure to protein–protein interactions. Yeast two-hybrid and peptide combinatorial libraries have been used to build on the AGR2 interactome. The yeast two-hybrid screen identified the listed proteins, as given in Table 1; only one of these has been well validated, namely the AAA$^+$ ATPase protein Reptin.[47] The AGR2:Reptin protein–protein interactions were validated using biochemical approaches and subsequently the interaction site for Reptin on AGR2 was fine-mapped to a peptide interface, comprising the proposed substrate-binding loop on AGR2, within amino acids F104 to Y111. Mutations in AGR2 at the junction codons of the loop at 104 and 111 attenuate Reptin binding to the protein.[47] Small peptides from the AGR2 substrate-binding loop are sufficient to reconstitute the ADP-responsive binding of Reptin (Maslon *et al.*[47] and unpublished data), thus identifying the first potential protein–protein interaction assay to develop drug leads for inhibiting the Reptin–AGR2 complex. Whether the substrate-binding loop is a dominant interaction site for the many AGR2-interacting proteins will be one key aim of future work on its mechanism of action and on developing possible strategies to therapeutically target this class of protein–protein interactions. Indeed, phosphosite plus annotated a phosphorylation site at Y111, suggesting that phosphorylation at this motif could regulate protein–protein interactions of AGR2.

Table 1. AGR2-interacting proteins identified in Y2H (acquired from *Hybrigenics*) or published co-immunoprecipitation studies

Gene name	*Function*	*Method*	*Reference*
ARHGAP29	GTPase activating protein (Rho)	Y2H	Unpublished data
CKAP2	Cytoskeletal-linked protein involved in mitosis	Y2H	Unpublished data
CHD6	Chromatin remodelling factor	Y2H	Unpublished data
DAG1	Links the cytoskeleton and the extracellular matrix	Y2H	Fletcher *et al.*[50]
GPSM2	Regulates G-protein activation	Y2H	Unpublished data
HECTD1	E3 ubiquitin ligase	Y2H	Unpublished data
HIVEP1	DNA binding protein	Y2H	Unpublished data
NRIP1	Binds hormone-dependent receptors	Y2H	Unpublished data
NRXN3	Controls adhesion and receptor signalling	Y2H	Unpublished data
RUVBL2	AAA$^+$ ATPase—DNA repair and transcription	Y2H	Maslon *et al.*[47]
LYPD3	Regulates cell migration	Y2H	Unpublished data
KDELR	KDEL receptors	CoIP	Raykhel *et al.*[42]
C4.4	Metastasis-associated GPI-anchored protein	Y2H	Fletcher *et al.*[50]
MUC2	Mucin 2	CoIP	Park *et al.*[33]
PROD1	Axotolt homolog for human CD59	Y2H	Kumar *et al.*[49]

Abbreviations: AGR2, anterior gradient 2; GPI, glycosylphosphatidylinisotol.

The 'consensus peptide-binding activity' of AGR2

Some molecular chaperones like heat-shock protein 70 can harbour canonical consensus binding motifs[51] that direct the class of interactions targeted by these proteins. There is no *a priori* reason as to why PDI/chaperones like AGR2 would have an intrinsic and specific peptide-binding function that directs its potential protein folding activity in the ER, other than its ability to interact with reduced and oxidized cysteines through its thioredoxin fold. However, when AGR2 was presented to a combinatorial peptide library, the synthetic AGR2 protein expressed in *Escherichia coli* exhibited an intrinsic affinity for peptides containing the pentapeptide consensus (T/S)–X–(I/V/M)–(Y/W/F)–(Y/W/F).[16] The minimal pentapeptide TXIYY was shown to bind specifically to the AGR2 in crude lysates and the site of interaction was later mapped to the same C-terminal region on AGR2 to which Reptin binds.[52] However, the structural basis for this specific protein–protein interaction remains unsolved and whether the peptide forms a 1:1 or 1:2 stoichiometry with the AGR2 dimer remains undefined.

An intrinsic peptide docking function of a target protein can also be exploited to develop synthetic 'peptide mimetics' to determine whether a protein is potentially 'drugable' *in vivo*. This approach has been most exploited with the MDM2 and MDM4 oncoproteins.[53–55] In the case of AGR2, synthetic EGFP–peptide fusion protein expression constructs transfected into cells can 'stabilize' AGR2 protein and stimulate the activity of the p53 tumour suppressor,[44] a pathway that is negatively regulated by AGR2.[45] Such peptides synergize with ultraviolet stress to induce the nuclear import of p53 and the synthesis of the consensus peptide-binding motif to the cell-membrane-permeable delivery sequence[45,56] similarly induced the nuclear translocation of p53. These data highlight the potential drugability of the AGR2 pathway with respect to stimulation of the p53 tumour suppressor.

Future research areas on protein structure and function in cells

Structure–function analyses of PDIs have largely focused on the role of the thioredoxin fold in mediating protein folding landscapes. However, whether these proteins have consensus docking sites on their cargo proteins and whether they might be drugable have not been addressed, except perhaps for the 2,4-dinitrochlorobenzene-mediated inhibition of TRX1.[57] Although various PDIs, like thioredoxins, have been implicated in radiation resistance or drug resistance, the interacting proteins are essentially undefined and the role of disulphide motifs not well understood in cell systems. This forms a compelling area of future work for AGR2: to exploit its functional domains (leader sequence, substrate-binding loop, dimeric nature, thioredoxin fold, ER retention site and TXIYY peptide consensus binding domain) to identify its global interactome using proteomic technologies and how these protein interactions towards specific motifs might direct AG2 to function in the ER, outwith the ER, and when shuttled extracellularly. In addition, the development of a panel of protein-interaction mutants in these same motifs would create a toolbox for this emerging field that can be used to develop goalposts for understanding AGR2 functions in a range of biological situations.

AGR2 EXPRESSION IN CANCERS

AGR2 exhibits the basic features of a pro-oncogenic protein

To evaluate the role of AGR2 in cancer (Figure 4), the over-expression and suppression of AGR2 in cancer cell lines has been used to determine whether it has growth-suppressive or pro-growth functions. Clonogenic assays were used to demonstrate that AGR2 can enhance cancer cell survival, rather than inhibit cell growth.[32] Deletion of the last 10 amino acids of AGR2, harbouring the ER retention site of AGR2, prevented clonogenic growth

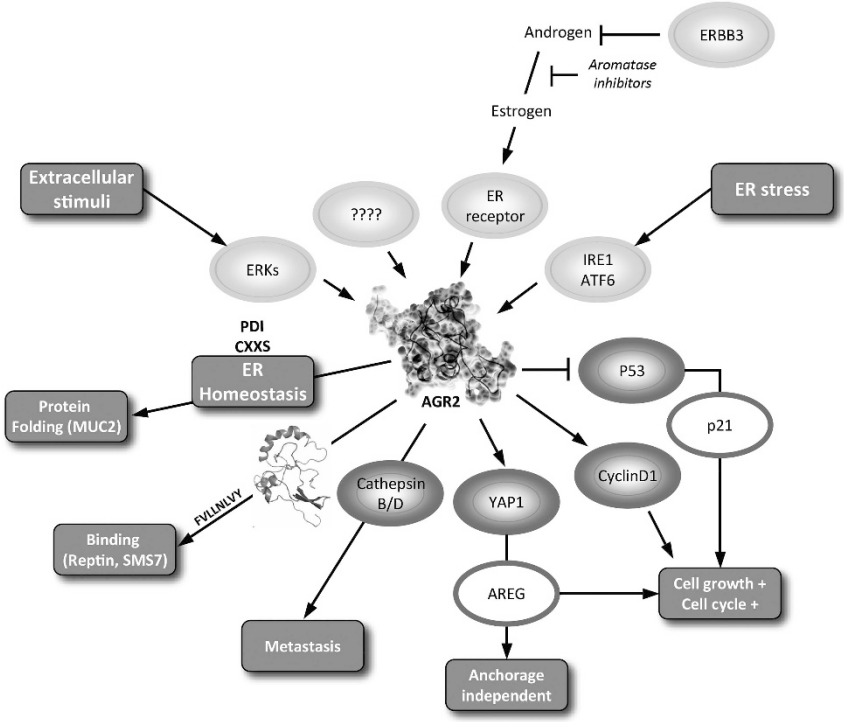

Figure 4. AGR2 biological pathways. AGR2 is presented as the central point of this picture (structure derived from ERP18[31]). Pathway intermediates known to regulate AGR2 expression are indicated in light grey ovals. AGR2-dependent functions are indicated in dark grey ovals. Potential (not experimentally demonstrated) intermediates are indicated as empty ovals. Physiological/pathophysiological inputs or outputs are represented as grey boxes.

stimulation, highlighting a role for this motif in AGR2 survival signalling. Further, AGR2 silencing was shown to inhibit proliferation, invasion and survival *in vitro* in pancreatic[43] and breast cancer cell lines.[58] AGR2 is also upstream of and stimulates key cancer-signalling pathways, such as cyclin D1, c-Myc, p-Src and survivin,[58] based on using both short interfering RNA and ectopic expression to manipulate AGR2 levels. Moreover, conditioned media from cells silenced in AGR2 have a reduced ability to stimulate proliferation of pancreatic cancer cells.[43] AGR2 has also been described as an oncogene[59] supporting a role in cellular transformation and adenocarcinoma growth. In the premalignant Barrett oesophagus and oesophageal cancer models, AGR2 overexpression induces colony formation and transformation.[45,59] Conversely, short interfering RNA- or short hairpin RNA-mediated AGR2 knockdown inhibits colony and subcutaneous growth in oesophageal and pancreatic cancer models.[43,59]

AGR2 as a mediator of Tamoxifen drug resistance in human breast cancers

As discussed above, the *AGR2* promoter responds to oestrogen in normal tissue as well as some cancer cell lines. Coincidentally, AGR2 protein is overexpressed in ER-positive human cancers. The anti-oestrogen Tamoxifen can be used successfully to treat human breast cancers; however, intrinsic resistance and/or acquired resistance remains a significant problem. It has been suggested that some genes paradoxically act as agonists of Tamoxifen and might mediate intrinsic resistance. As such, identification of oestrogen-responsive genes induced by Tamoxifen, yet play a survival role, has been one key aim in the field of drug resistance.[60] Three nearly simultaneous observations pointed towards AGR2 as being one of these elusive genes that might mediate the resistance of breast cancer to Tamoxifen. Firstly, was the observation that in cancer patients treated with the oestrogen suppressor Letrazole, the *AGR2* gene (not *AGR3*) was one of the top suppressed genes in biopsies from patients post-treatment that respond well to the drug, suggesting that resistance to anti-oestrogens might be due to failure to suppress *AGR2*.[32] Second, in a 'deep' proteomics screen using the label-free approach PaCIFIC to search for dominant proteins induced by Tamoxifen, AGR2 was the second most induced protein,[61] and xenograft studies demonstrated that the *AG2* gene can mediate cisplatin resistance.[61] Third, an analysis of *AGR2* gene expression in relation to prognostic markers in ER-positive breast cancers indicated that high *AGR2* expression was linked significantly to poor prognosis.[32]

At the molecular level, it was also demonstrated that this induction of AGR2 by Tamoxifen was a direct effect of ERα activation using chromatin immunoprecipitation. Elevated ERα was bound to the *AGR2* promoter in the presence of oestradiol or Tamoxifen and the *AGR2* promoter (-175 to $+35$) was activated in a luciferase reporter assay in the presence of oestradiol or Tamoxifen.[32] Understanding the molecular details of why Tamoxifen, selectively, induces AGR2 expression while acting as an anti-oestrogen for the majority of oestrogen-responsive genes might identify novel drug targets that sensitize Tamoxifen-resistant/AGR2-positive cancers.

A dominant role for AGR2 in oesophageal cancer progression

Oesophageal cancer progression is one of the minority of cancer types known to evolve through progressive changes in tissue architecture, from metaplasia to dysplasia, and adenocarcinoma. Oesophageal adenocarcinoma differs from other cancer types in that the environmental stress of bile acid reflux plays an apparently important role in disease progression and that p53 tumour suppressor gene mutation can occur very early in the carcinogenic sequence.[62] The early selection pressure for p53 mutation in oesophageal cancer occurs during the replacement of squamous epithelium with metaplastic epithelium, also called 'Barrett's oesophagus'. To identify potentially novel p53 inhibitors in this cancer 'intermediate', a clinical proteomics screen had been set up in a proliferative disease (Barrett's epithelium) and identified AGR2, which was validated as a potent inhibitor of p53-dependent transcription and a growth-promoting proto-oncogene.[45] Subsequently, it was shown that overexpression of AGR2 is maintained in oeosphageal cancer tissue, as the large majority of adenocarcinomas express AGR2 as defined using immunohistochemical methods.[46] *AGR2* mRNA expression is also a dominant feature of a recently identified murine model of Barrett's oesophageal epithelium induced through deletion of the squamous stem cell progenitor p63.[63] As oesophageal cancers are presumably oestrogen-independent, developing therapeutic strategies to inhibit the AGR2 pathway (if not the PDI machine) might prove to be different than those used to inhibit its Tamoxifen-resistance activity in breast cancers.

Uncoupling of AGR2 and AGR3 expression in human ovarian cancers

Despite the link between the oestrogen receptor and expression of AGR2 in breast cancers, oestrogen-independent expression can be observed in other human cancers.[46] It is notable that AGR2, and not AGR3, is the gene/protein that was found by various OMICs platforms to be expressed by oestrogen.[64] Follow-up studies confirmed that AGR3 is overexpressed in breast tumours and that AGR3 is co-expressed with AGR2 in breast cancer tissue with a strong positive correlation with ERα status.[50] However, AGR3 was not co-expressed to a high degree with AGR2 in prostate cancers, indicating that AGR2 and AGR3 expression can be uncoupled. The expression of both AGR2 and AGR3 was evaluated in complex hormone-independent human cancer to identify a suitable clinical model in which to begin to study AGR3/AGR2 function. Studies found that AGR2 and AGR3 are overexpressed in four different subtypes of primary human ovarian cancer, but although the expression of AGR3 and AGR2 can be co-incident in mucinous ovarian cancers, they are uncoupled in the other three types of primary ovarian cancers.[16] These later data suggest that distinct stresses can selectively induce either AGR2 or AGR3 *in vivo*.

Selective expression of AGR2 in liver cancers

AGR2 is mainly expressed in the normal biliary tree in both fetal and adult normal liver. More particularly, the tall epithelial cells covering the large bile ducts as well as gallbladder epithelial cells showed strong AGR2 staining.[65] Moreover, hepatocellular carcinoma tumours do not show any significant AGR2 staining, whereas fibrolamellar carcinoma did.[65,66] Both hilar and extrahepatic cholangiocarcinoma reveal positive AGR2 staining. In contrast, only 50% of the intrahepatic cholangiocarcinoma analysed display strong AGR2 staining.[65] This result was further investigated and led to the correlation of AGR2 staining with mucus production (Lepreux *et al*.[65] and Bioulac-Sage, Balabaud and Chevet, unpublished results). These results can be put in perspective of those obtained using $AGR2^{-/-}$ mice[33] and might provide a selective advantage to mucus-producing tumour cells. This observation might also provide some information on the cellular origin of intrahepatic cholangiocarcinoma, which still remains poorly defined.[67]

AGR2 and metastasis

Many cancer studies revealed that overexpression or suppression of AGR2, in different model systems, can affect cell proliferation, invasion and survival *in vitro*, metastasis and tumour growth

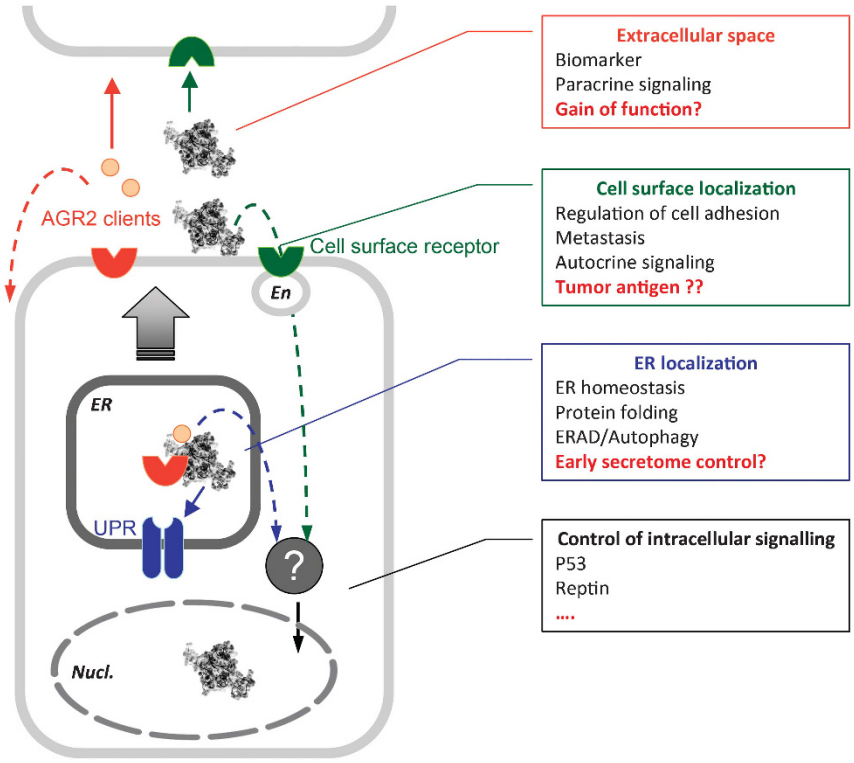

Figure 5. Localization and functions of AGR2. The different location of AGR2 are reported and correlated with its reported functions in these compartments.

in vitro.[43,59] However, the molecular mechanisms at the origin of these phenotypes are not yet clearly understood, although it was recently shown that AGR2 is present at the cell surface of pancreatic cancer cells,[41] a phenomenon similar to that previously reported in tumour immunity.[68,69] As demonstrated for other ER-resident proteins,[2] we have recently shown that the increased expression of AGR2 can enhance ER folding capacity, allowing cancer cells to cope with increased protein production and secretion.[31] Genomic analysis of AGR2-stable cells by cDNA microarray revealed that AGR2 overexpression upregulates the expression of genes involved in cell proliferation, invasion and angiogenesis,[70] which are very important for tumour progression and metastasis. Conversely, the genes involved in the negative regulation of cell proliferation, adhesion and death are downregulated.[70] Moreover, the overexpression of AGR2 in grafted cells results in greater propensity to form lung metastases when propagated as xenografts in nude mice,[48] showing that AGR2 has an influence on the *in vivo* tumour biology. In breast cancer models, overexpression of AGR2 failed to alter tumour formation *in vivo* or growth rate *in vitro*, but rather, reduced cell adhesion and increased the numbers of metastases.[48] A similar observation was obtained in head and neck squamous cell carcinoma cells, in which AGR2 silencing (also observed upon CD147 signalling) reduced cell proliferation migration and invasion,[71] thus CD147 could also represent an AGR2 client in this model. Thus, these results suggest that AGR2 does not only promote cell migration but can also control the cell adhesion rates of detached cells, which may correlate with the ability of metastatic cells to colonize distal sites. The differences in whether migration or adhesion dominates through an AGR2 signal are likely due to the fact that different cancer cell models were used in distinct experiments and that each cell type exploits distinct pathways that mediate cancer survival. AGR2 involvement in metastasis is also demonstrated in prostate adenocarcinoma cells in which an enhanced invasive behavior of cell-overexpressing AGR2 is shown.[72] AGR2 promotes migration and invasion of prostate cancer cells, consistent with previous results in other tumour types, but at the opposite, it attenuates cell growth.[73] In this study, overexpression of AGR2 led to a reduction of colony formation and cell proliferation of prostate cells, possibly brought by a cell cycle arrest.[73] Thus, AGR2 could regulate epithelial–mesenchymal transition (EMT) in development and cancer. EMT is a developmental programme used in cell differentiation during embryogenesis, but parts of this genetic programme are believed to be reactivated during metastasis to transform malignant epithelial cells into motile and invasive mesenchymal-like cells.[74,75] It is believed that EMT is a transient state and that the process is reversed once the cells form a metastatic lesion. Similar observations have also been reported for SIP, a transcription factor that control EMT. SIP1 induced an invasion phenotype while at the same time attenuating cell cycle progression.[76] Nevertheless, the involvement of AGR2 in EMT needs to be further investigated.

Cell signalling mechanisms of AGR2 in cancer

As yet, the cellular mechanism by which AGR2 promotes growth are poorly understood, but there have been a number of observations that are beginning to shed light on the role it plays in cell signalling networks. One pathway that AGR2 has been implicated in is the EGFR pathway. Recently, Dong *et al*.[46] have shown that AGR2 induced expression of AREG, a growth-promoting EGFR ligand. This study functionally linked AGR2 and AREG and supported a significant role for AGR2 in lung adenocarcinomas and the regulation of cell growth. As a result of AGR2 expression, AREG may stimulate the EGFR signalling pathway and may be responsible for the increased cell proliferation and anchorage-independent growth observed in transformed cells.[43,59] It was also shown that this was a specific induction of AREG, there was no induction of other known EGFR

ligands detected and that activation resulted in increased phosphorylation of both EGFR itself and AKT downstream of the receptor. AREG expression has also been shown to be regulated by the Hippo pathway, which serves to regulate cell proliferation and apoptosis, and functions in regulating organ size.[77] Repression of the Hippo pathway results in YAP dephosphorylation followed by transport to the nucleus, where it inhibits apoptosis and promotes cell division. Nuclear YAP protein is associated with neoplasia and has been observed in lung, colon, ovarian and breast adenocarcinoma.[46,78] Another study showed that the Erb3 binding protein 1 could inhibit the expression of AGR2 in prostate cancer cells. The mechanism by which this occurs is yet to be determined, but the Erb3 binding protein 1 can negatively regulate the androgen receptor and so this could be one possible mode.

A key clinical model used to evaluate cancer progression mechanisms is oesophageal adenocarcinoma where selection pressures are being placed on the survival of cells with either wild type-p53 or mutant p53, early in carcinogenesis. The discovery of AGR2 as a dominant overexpressed protein in oesophageal metaplasia and its validation as a p53 inhibitor suggests that AGR2 pathway might be a key mechanism to silence p53 signalling.[45] The introduction of the *AGR2* gene into cancer cell models can:[45] (i) enhance cell survival in a clonogenic assay, similar to the p53 mutant HIS175 allele; (ii) p53 transcriptional activity was reduced when co-transfected with AGR2; (iii) phosphorylation at p53-activating phosphorylation sites (Ser15 and Ser392) upon exposure to ultraviolet was attenuated in cells overexpressing AGR2, which suggest that AGR2's ability to act as a survival factor may be linked to its activity as a p53 kinase inhibitor; and (iv) the introduction of the AGR2 to HCT116 colon cancer cells resulted in the redistribution of p53 from the nucleus to the cytoplasm in ultraviolet-irradiated cells.[44] In agreement with this, targeting AGR2 with short interfering RNA or the use of the TXIYY peptide aptamers induced a redistribution of p53 protein from the cytosol to the nucleus,[44] suggesting that the AGR2 pathway might be targeted therapeutically to stimulate the p53 pathway. The role of AGR2 as an oncogenic wild type-p53 inhibitor cannot, at present, be reconciled with the role of AGR2 as a pro-oncogenic migration/invasion factor, as the cell models used in these latter studies have mutant p53.

Future research areas on challenging the function of AGR2 in cancers

Murine transgenic technologies have revolutionized our understanding of gene–gene interactions with proteins such as RAS, APC, p53, PTEN and SMAD4. The vast majority of this research is based on genes identified from human genetic studies. AGR2 represents a gene identified from OMIC approaches and it has not yet been evaluated in the many oncogenic murine transgenic systems that exist. The AGR2 field will need to integrate the function of AGR2 as an ER-resident chaperone into the genetic matrix identified from key oncogenic nodes to link protein folding/secretome functions of AGR2 to cancer cell survival. Further, the use of specific mutant alleles of AGR2 in such murine cancer models will also define the interactome of the protein that is linked to other oncogenic pathways. For example, we would expect AGR2-overproducing transgenes to put selective pressures on silencing p53 and we might therefore expect that tumours that evolve from $AGR2^{+}$ background might not have selected for p53 mutations to the same extent as AGR2-negative cancers.

QUESTIONS AND FUTURE PERSPECTIVES

In conclusion, AGR2 can stimulate cancer cell proliferation, invasion and survival *in vitro*, resistance to chemotherapy using murine xenografts, metastasis and tumour growth *in vitro*, is relatively unique as a PDI in that it is known to be overexpressed in a large range of human cancers types, and its expression levels can be used to predict patient prognosis in a number of cancers. In an attempt to better explain the function of AGR2 in cancer, we have proposed a fundamental, core role for AGR2 is in regulating the ER capacity to adapt. This ER buffering capacity of AGR2 can in turn have an impact on the nature of the cell secretome under physiological or pathological conditions, as well as affect intracellular transcription networks such as p53 that are known to sense the unfolded protein response. Consequently, shifts in this equilibrium could contribute to the development and progression of human cancers. Although the biological manifestations of AGR2 as a proto-oncogenic protein are now evident, the molecular mechanisms of how it functions and is regulated, nevertheless, remains poorly defined. On the basis of the unique properties of AGR2 we have reviewed here, a number of compelling basic and clinical research questions remain, including (i) why does AGR2 have a 'weak' or non-canonical ER-retention sequence and does this play a role in novel intracellular trafficking outwith the ER that drives cancer growth (Figure 5); (ii) what significance does the changes in the cellular redox balance of a cancer cell have on a PDI-like AGR2 with a single cysteine in its thioredoxin fold (CXXS); (iii) what is the role of the specific 'peptide-binding' function of AGR2 in cargo binding/trafficking and does this provide a peptido-mimetic, therapeutic intervention strategy; (iv) is the AGR2 pathway, including its upstream regulators, its cofactor-interacting proteins or its dimerization interface 'drugable'; (v) is the 'secretome' of AGR2 that mediates its core functions in cell migration, growth or adhesion 'drugable'; and (vi) will transgenic or genetic models that overproduce missense mutant alleles of AGR2 shed light on its ER signalling and pro-oncogenic functions? Answers to these fundamental questions would shed light on the role of the ER protein folding quality control in human cancer cell growth and provide an understanding of how the secretome provides adaptive advantages to the pro-metastatic cancer cell.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We apologize to authors whose work has not been cited here owing to space limitations. This work was supported by grants from INSERM, INCa, Ligue contre le Cancer, Association pour la Recherche sur le Cancer (EC), Ligue contre le Cancer and a Chaire mixte Inserm—Univ. Bordeaux (FD), GACR number P301/10/1615 (RH), IGA MZ CR number NT/13794-4/2012 (BV) and by the European Regional Development Fund (RECAMO CZ.1.05/2.1.00/03.0101) (BV, RH, EM), and Cancer Research UK, MRC and BBSRC (TH).

REFERENCES

1 Ma Y, Hendershot LM. The role of the unfolded protein response in tumour development: friend or foe? *Nat Rev Cancer* 2004; **4**: 966–977.
2 Moenner M, Pluquet O, Bouchecareilh M, Chevet E. Integrated endoplasmic reticulum stress responses in cancerdisclaimer. *Cancer Res* 2007; **67**: 10631–10634.
3 Chevet E, Cameron PH, Pelletier MF, Thomas DY, Bergeron JJ. The endoplasmic reticulum: integration of protein folding, quality control, signaling and degradation. *Curr Opin Struct Biol* 2001; **11**: 120–124.
4 Ni M, Lee AS. ER chaperones in mammalian development and human diseases. *FEBS Lett* 2007; **581**: 3641–3651.
5 Benham AM. The protein disulfide isomerase family: key players in health and disease. *Antioxid Redox Signal* 2011; **16**: 781–789.
6 Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad IB *et al*. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat Genet* 2012; **44**: 694–698.
7 Kozlov G, Maattanen P, Thomas DY, Gehring K. A structural overview of the PDI family of proteins. *FEBS J* 2010; **277**: 3924–3936.

8 Bradley L, Wainstock D, Sive H. Positive and negative signals modulate formation of the *Xenopus* cement gland. *Development* 1996; **122**: 2739–2750.
9 Sive H, Bradley L. A sticky problem: the *Xenopus* cement gland as a paradigm for anteroposterior patterning. *Dev Dyn* 1996; **205**: 265–280.
10 Aberger F, Weidinger G, Grunz H, Richter K. Anterior specification of embryonic ectoderm: the role of the *Xenopus* cement gland-specific gene XAG-2. *Mech Dev* 1998; **72**: 115–130.
11 Sive HL, Hattori K, Weintraub H. Progressive determination during formation of the anteroposterior axis in *Xenopus* laevis. *Cell* 1989; **58**: 171–180.
12 Komiya T, Tanigawa Y, Hirohashi S. Cloning of the gene gob-4, which is expressed in intestinal goblet cells in mice. *Biochim Biophys Acta* 1999; **1444**: 434–438.
13 Adam GC, Sorensen EJ, Cravatt BF. Trifunctional chemical probes for the consolidated detection and identification of enzyme activities from complex proteomes. *Mol Cell Proteomics* 2002; **1**: 828–835.
14 Adam GC, Sorensen EJ, Cravatt BF. Proteomic profiling of mechanistically distinct enzyme classes using a common chemotype. *Nat Biotechnol* 2002; **20**: 805–809.
15 Adam PJ, Boyd R, Tyson KL, Fletcher GC, Stamps A, Hudson L *et al.* Comprehensive proteomic analysis of breast cancer cell membranes reveals unique proteins with potential roles in clinical cancer. *J Biol Chem* 2003; **278**: 6482–6489.
16 Gray TA, Maclaine NJ, Michie CO, Bouchalova P, Murray E, Howie J *et al.* Anterior Gradient-3: a novel biomarker for ovarian cancer that mediates cisplatin resistance in xenograft models. *J Immunol Methods* 2012; **378**: 20–32.
17 Xu X, Hou Y, Yin X, Bao L, Tang A, Song L *et al.* Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* 2012; **148**: 886–895.
18 Persson S, Rosenquist M, Knoblach B, Khosravi-Far R, Sommarin M, Michalak M. 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* 2005; **36**: 734–740.
19 Alanen HI, Williamson RA, Howard MJ, Lappi AK, Jantti HP, Rautio SM *et al.* Functional characterization of ERp18, a new endoplasmic reticulum-located thioredoxin superfamily member. *J Biol Chem* 2003; **278**: 28912–28920.
20 Knoblach B, Keller BO, Groenendyk J, Aldred S, Zheng J, Lemire BD *et al.* ERp19 and ERp46, new members of the thioredoxin family of endoplasmic reticulum proteins. *Mol Cell Proteomics* 2003; **2**: 1104–1019.
21 Jessop CE, Watkins RH, Simmons JJ, Tasab M, Bulleid NJ. Protein disulphide isomerase family members show distinct substrate specificity: P5 is targeted to BiP client proteins. *J Cell Sci* 2009; **122**: 4287–4295.
22 Anelli T, Alessio M, Mezghrani A, Simmen T, Talamo F, Bachi A *et al.* ERp44, a novel endoplasmic reticulum folding assistant of the thioredoxin family. *EMBO J* 2002; **21**: 835–844.
23 Bambang IF, Lee YK, Richardson DR, Zhang D. Endoplasmic reticulum protein 29 regulates epithelial cell integrity during the mesenchymal–epithelial transition in breast cancer cells. *Oncogene* 2013; **32**: 1240–1251.
24 Zhang JS, Gong A, Cheville JC, Smith DI, Young CY. AGR2an androgen-inducible secretory protein overexpressed in prostate cancer. *Genes Chromosomes Cancer* 2005; **43**: 249–259.
25 Zheng W, Rosenstiel P, Huse K, Sina C, Valentonyte R, Mah N *et al.* Evaluation of AGR2 and AGR3 as candidate genes for inflammatory bowel disease. *Genes Immun* 2006; **7**: 11–18.
26 Petek E, Windpassinger C, Egger H, Kroisel PM, Wagner K. Localization of the human anterior gradient-2 gene (AGR2) to chromosome band 7p21.3 by radiation hybrid mapping and fluorescence *in situ* hybridisation. *Cytogenet Cell Genet* 2000; **89**: 141–142.
27 Krig SR, Jin VX, Bieda MC, O'Geen H, Yaswen P, Green R *et al.* Identification of genes directly regulated by the oncogene ZNF217 using chromatin immunoprecipitation (ChIP)-chip assays. *J Biol Chem* 2007; **282**: 9703–9712.
28 Thompson DA, McPherson LA, Carmeci C, deConinck EC, Weigel RJ. Identification of two estrogen receptor transcripts with novel 5′ exons isolated from a MCF7 cDNA library. *J Steroid Biochem Mol Biol* 1997; **62**: 143–153.
29 Wilson CL, Sims AH, Howell A, Miller CJ, Clarke RB. Effects of oestrogen on gene expression in epithelium and stroma of normal human breast tissue. *Endocr Relat Cancer* 2006; **13**: 617–628.
30 Zweitzig DR, Smirnov DA, Connelly MC, Terstappen LW, O'Hara SM, Moran E. Physiological stress induces the metastasis marker AGR2 in breast cancer cells. *Mol Cell Biochem* 2007; **306**: 255–260.
31 Higa A, Mulot A, Delom F, Bouchecareilh M, Nguyen DT, Boismenu D *et al.* Role of pro-oncogenic protein disulfide isomerase (PDI) family member anterior gradient 2 (AGR2) in the control of endoplasmic reticulum homeostasis. *J Biol Chem* 2011; **286**: 44855–44868.
32 Hrstka R, Nenutil R, Fourtouna A, Maslon MM, Naughton C, Langdon S *et al.* The pro-metastatic protein anterior gradient-2 predicts poor prognosis in tamoxifen-treated breast cancers. *Oncogene* 2010; **29**: 4838–4847.
33 Park SW, Zhen G, Verhaeghe C, Nakagami Y, Nguyenvu LT, Barczak AJ *et al.* The protein disulfide isomerase AGR2 is essential for production of intestinal mucus. *Proc Natl Acad Sci USA* 2009; **106**: 6950–6955.
34 Tachibana C, Stevens TH. The yeast EUG1 gene encodes an endoplasmic reticulum protein that is functionally related to protein disulfide isomerase. *Mol Cell Biol* 1992; **12**: 4601–4611.
35 Norgaard P, Winther JR. Mutation of yeast Eug1p CXXS active sites to CXXC results in a dramatic increase in protein disulphide isomerase activity. *Biochem J* 2001; **358**: 269–274.
36 Wunderlich M, Otto A, Maskos K, Mucke M, Seckler R, Glockshuber R. Efficient catalysis of disulfide formation during protein folding with a single active-site cysteine. *J Mol Biol* 1995; **247**: 28–33.
37 Zhao F, Edwards R, Dizon D, Afrasiabi K, Mastroianni JR, Geyfman M *et al.* Disruption of Paneth and goblet cell homeostasis and increased endoplasmic reticulum stress in Agr2 $-/-$ mice. *Dev Biol* 2010; **338**: 270–279.
38 Schroeder BW, Verhaeghe C, Park SW, Nguyenvu LT, Huang X, Zhen G *et al.* AGR2 is induced in asthma and promotes allergen-induced mucin overproduction. *Am J Respir Cell Mol Biol* 2012; **47**: 178–185.
39 Li S, Wang Y, Zhang Y, Lu MM, Demayo FJ, Dekker JD *et al.* Foxp1/4 control epithelial cell fate during lung development and regeneration through regulation of anterior gradient 2. *Development* 2012; **139**: 2500–2509.
40 Gupta A, Dong A, Lowe AW. AGR2 gene function requires a unique endoplasmic reticulum localization motif. *J Biol Chem* 2012; **287**: 4773–4782.
41 Dumartin L, Whiteman HJ, Weeks ME, Hariharan D, Dmitrovic B, Iacobuzio-Donahue CA *et al.* AGR2 is a novel surface antigen that promotes the dissemination of pancreatic cancer cells through regulation of cathepsins B and D. *Cancer Res* 2011; **71**: 7091–7102.
42 Raykhel I, Alanen H, Salo K, Jurvansuu J, Nguyen VD, Latva-Ranta M *et al.* A molecular specificity code for the three mammalian KDEL receptors. *J Cell Biol* 2007; **179**: 1193–1204.
43 Ramachandran V, Arumugam T, Wang H, Logsdon CD. Anterior gradient 2 is expressed and secreted during the development of pancreatic cancer and promotes cancer cell survival. *Cancer Res* 2008; **68**: 7811–7818.
44 Fourtouna A, Murray E, Nicholson J, Maslon M, Pang L, Dryden D *et al.* The anterior gradient-2 Pathway as a model for developing peptide-aptamer anticancer drug leads that stimulate p53 function. *Current Chemical Biology* 2009; **3**: 124–137.
45 Pohler E, Craig AL, Cotton J, Lawrie L, Dillon JF, Ross P *et al.* The Barrett's antigen anterior gradient-2 silences the p53 transcriptional response to DNA damage. *Mol Cell Proteomics* 2004; **3**: 534–547.
46 Dong A, Gupta A, Pai RK, Tun M, Lowe AW. The human adenocarcinoma-associated gene, AGR2, induces expression of amphiregulin through Hippo pathway co-activator YAP1 activation. *J Biol Chem* 2011; **286**: 18301–18310.
47 Maslon MM, Hrstka R, Vojtesek B, Hupp TR. A divergent substrate-binding loop within the pro-oncogenic protein anterior gradient-2 forms a docking site for Reptin. *J Mol Biol* 2010; **404**: 418–438.
48 Liu D, Rudland PS, Sibson DR, Platt-Higgins A, Barraclough R. Human homologue of cement gland protein, a novel metastasis inducer associated with breast carcinomas. *Cancer Res* 2005; **65**: 3796–3805.
49 Kumar A, Godwin JW, Gates PB, Garza-Garcia AA, Brockes JP. Molecular basis for the nerve dependence of limb regeneration in an adult vertebrate. *Science* 2007; **318**: 772–777.
50 Fletcher GC, Patel S, Tyson K, Adam PJ, Schenker M, Loader JA *et al.* 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* 2003; **88**: 579–85.
51 Takenaka IM, Leung SM, McAndrew SJ, Brown JP, Hightower LE. Hsc70-binding peptides selected from a phage display peptide library that resemble organellar targeting sequences. *J Biol Chem* 1995; **270**: 19839–19844.
52 Clarke DJ, Murray E, Hupp T, Mackay CL, Langridge-Smith PR. Mapping a noncovalent protein–peptide interface by top-down FTICR mass spectrometry using electron capture dissociation. *J Am Soc Mass Spectrom* 2011; **22**: 1432–1440.
53 Fraser JA, Vojtesek B, Hupp TR. A novel p53 phosphorylation site within the MDM2 ubiquitination signal: I. phosphorylation at SER269 *in vivo* is linked to inactivation of p53 function. *J Biol Chem* 2010; **285**: 37762–37672.
54 Worrall EG, Worrall L, Blackburn E, Walkinshaw M, Hupp TR. The effects of phosphomimetic lid mutation on the thermostability of the N-terminal domain of MDM2. *J Mol Biol* 2010; **398**: 414–428.
55 Worrall EG, Wawrzynow B, Worrall L, Walkinshaw M, Ball KL, Hupp TR. Regulation of the E3 ubiquitin ligase activity of MDM2 by an N-terminal pseudo-substrate motif. *J Chem Biol* 2009; **2**: 113–129.
56 Fischer PM, Zhelev NZ, Wang S, Melville JE, Fahraeus R, Lane DP. Structure-activity relationship of truncated and substituted analogues of the intracellular delivery vector Penetratin. *J Pept Res* 2000; **55**: 163–172.

57 Tian C, Gao P, Zheng Y, Yue W, Wang X, Jin H *et al.* Redox status of thioredoxin-1 (TRX1) determines the sensitivity of human liver carcinoma cells (HepG2) to arsenic trioxide-induced cell death. *Cell Res* 2008; **18**: 458–471.
58 Vanderlaag KE, Hudak S, Bald L, Fayadat-Dilman L, Sathe M, Grein J *et al.* 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* 2010; **12**: R32.
59 Wang Z, Hao Y, Lowe AW. The adenocarcinoma-associated antigen, AGR2, promotes tumor growth, cell migration, and cellular transformation. *Cancer Res* 2008; **68**: 492–497.
60 Vendrell JA, Robertson KE, Ravel P, Bray SE, Bajard A, Purdie CA *et al.* A candidate molecular signature associated with tamoxifen failure in primary breast cancer. *Breast Cancer Res* 2008; **10**: R88.
61 Hengel SM, Murray E, Langdon S, Hayward L, O'Donoghue J, Panchaud A *et al.* Data-independent proteomic screen identifies novel tamoxifen agonist that mediates drug resistance. *J Proteome Res* 2011; **10**: 4567–4578.
62 Varghese S, Lao-Sirieix P, Fitzgerald RC. Identification and clinical implementation of biomarkers for Barrett's esophagus. *Gastroenterology* 2012; **142**: 435–441.
63 Wang X, Ouyang H, Yamamoto Y, Kumar PA, Wei TS, Dagher R *et al.* Residual embryonic cells as precursors of a Barrett's-like metaplasia. *Cell* 2011; **145**: 1023–1035.
64 Thompson DA, Weigel RJ. 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* 1998; **251**: 111–1116.
65 Lepreux S, Bioulac-Sage P, Chevet E. Differential expression of the anterior gradient protein-2 is a conserved feature during morphogenesis and carcinogenesis of the biliary tree. *Liver Int* 2011; **31**: 322–328.
66 Vivekanandan P, Micchelli ST, Torbenson M. Anterior gradient-2 is overexpressed by fibrolamellar carcinomas. *Hum Pathol* 2009; **40**: 293–299.
67 Komuta M, Govaere O, Vandecaveye V, Akiba J, Van Steenbergen W, Verslype C *et al.* Histological diversity in cholangiocellular carcinoma reflects the different cholangiocyte phenotypes. *Hepatology* 2012; **55**: 1876–1888.
68 Obeid M, Tesniere A, Ghiringhelli F, Fimia GM, Apetoh L, Perfettini JL *et al.* Calreticulin exposure dictates the immunogenicity of cancer cell death. *Nat Med* 2007; **13**: 54–61.
69 Panaretakis T, Joza N, Modjtahedi N, Tesniere A, Vitale I, Durchschlag M *et al.* The co-translocation of ERp57 and calreticulin determines the immunogenicity of cell death. *Cell Death Differ* 2008; **15**: 1499–1509.
70 Park K, Chung YJ, So H, Kim K, Park J, Oh M *et al.* AGR2, a mucinous ovarian cancer marker, promotes cell proliferation and migration. *Exp Mol Med* 2011; **43**: 91–100.
71 Sweeny L, Liu Z, Bush BD, Hartman Y, Zhou T, Rosenthal EL. CD147 and AGR2 expression promote cellular proliferation and metastasis of head and neck squamous cell carcinoma. *Exp Cell Res* 2012; **318**: 1788–1798.
72 Zhang Y, Ali TZ, Zhou H, D'Souza DR, Lu Y, Jaffe J *et al.* ErbB3 binding protein 1 represses metastasis-promoting gene anterior gradient protein 2 in prostate cancer. *Cancer Res* 2010; **70**: 240–248.
73 Bu H, Bormann S, Schafer G, Horninger W, Massoner P, Neeb A *et al.* 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* 2011; **71**: 575–587.
74 Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY *et al.* The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 2008; **133**: 704–715.
75 Thiery JP. Epithelial–mesenchymal transitions in development and pathologies. *Curr Opin Cell Biol* 2003; **15**: 740–746.
76 Mejlvang J, Kriajevska M, Vandewalle C, Chernova T, Sayan AE, Berx G *et al.* Direct repression of cyclin D1 by SIP1 attenuates cell cycle progression in cells undergoing an epithelial mesenchymal transition. *Mol Biol Cell* 2007; **18**: 4615–4624.
77 Zeng Q, Hong W. The emerging role of the hippo pathway in cell contact inhibition, organ size control, and cancer development in mammals. *Cancer Cell* 2008; **13**: 188–192.
78 Steinhardt AA, Gayyed MF, Klein AP, Dong J, Maitra A, Pan D *et al.* Expression of Yes-associated protein in common solid tumors. *Hum Pathol* 2008; **39**: 1582–1589.