A structural view of PA2G4 isoforms with opposing functions in cancer

Brendan W. Stevenson[1], Michael A. Gorman[2], Jessica Koach[3,4], Belamy B. Cheung[4,5], Glenn M. Marshall[5,6], Michael W. Parker[1,2], and Jessica K. Holien[1,7,8,*]

From the [1]St. Vincent's Institute of Medical Research, Fitzroy, Victoria, Australia, the [2]Bio21 Molecular Science and Biotechnology Institute and the [7]Department of Surgery, University of Melbourne, Parkville, Victoria, Australia, the [3]Department of Pediatrics, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California, USA, the [4]Children's Cancer Institute Australia for Medical Research, Lowy Cancer Research Centre, University of New South Wales, Sydney, Australia, the [5]School of Women's and Children's Health, University of New South Wales, Randwick, New South Wales, Australia, the [6]Kids Cancer Centre, Sydney Children's Hospital, Randwick, New South Wales, Australia, and the [8]School of Science, College of Science, Engineering, and Health, RMIT University, Melbourne, Victoria, Australia

Edited by Alex Toker

The role of proliferation-associated protein 2G4 (PA2G4), alternatively known as ErbB3-binding protein 1 (EBP1), in cancer has become apparent over the past 20 years. PA2G4 expression levels are correlated with prognosis in a range of human cancers, including neuroblastoma, cervical, brain, breast, prostate, pancreatic, hepatocellular, and other tumors. There are two PA2G4 isoforms, PA2G4-p42 and PA2G4-p48, and although both isoforms of PA2G4 regulate cellular growth and differentiation, these isoforms often have opposing roles depending on the context. Therefore, PA2G4 can function either as a contextual tumor suppressor or as an oncogene, depending on the tissue being studied. However, it is unclear how distinct structural features of the two PA2G4 isoforms translate into different functional outcomes. In this review, we examine published structures to identify important structural and functional components of PA2G4 and consider how they may explain its crucial role in the malignant phenotype. We will highlight the lysine-rich regions, protein-protein interaction sites, and post-translational modifications of the two PA2G4 isoforms and relate these to the functional cellular role of PA2G4. These data will enable a better understanding of the function and structure relationship of the two PA2G4 isoforms and highlight the care that will need to be undertaken for those who wish to conduct isoform-specific structure-based drug design campaigns.

PA2G4 discovery and structure

The human *PA2G4* gene was first isolated, characterized, cloned, and mapped by Lamartine *et al.* (1). They reported a 394-amino acid protein that showed strong similarity to the recently isolated murine p38-2G4 protein (2). Because p38-2G4 was identified as a cell cycle–regulated DNA-binding protein, it was deduced that the human counterpart was likely to have the same function. Three years later, a screen biased toward identifying proteins that bind ErbB3 in the absence of tyrosine phosphorylation described two alternate isoforms of human PA2G4, although they were not fully characterized (3). We now understand that the human PA2G4 gene contains three in-frame ATG initiation codons, which can be differentially spliced to produce either a short PA2G4-p42 (340 amino acids (aa)) or long PA2G4-p48 (394 aa) isoform (4) (Fig. 1). Although both isoforms of *PA2G4* regulate cellular growth and differentiation, these isoforms often have opposing roles, depending on the context. The long PA2G4-p48 isoform is transcribed by the first ATG codon and has been shown to promote tumorigenesis. It is the predominant form of PA2G4 found in mammalian cells and is located in both the nucleus and cytoplasm (4). Conversely, the short PA2G4-p42 acts as a tumor suppressor. The PA2G4-p42 isoform is transcribed by the third ATG codon of the *PA2G4* gene, is expressed at relatively low levels in mammalian cells, and has been localized only in the cytoplasm (4). Many published biological studies do not differentiate between these isoforms. Where possible throughout this review article, we have referred to them via their isoform-naming convention (*i.e.* PA2G4-p48 or PA2G4-p42). Where there is ambiguity over which isoform the study was conducted on, or for features that are conserved across both isoforms, we have not added the isoform-naming convention (*i.e.* PA2G4). Furthermore, unless stated, amino acid numbering is based upon the larger PA2G4-p48 numbering.

To date, three protein structures of PA2G4 have been published (Table 1). The first was a 2.5 Å resolution X-ray crystal structure of murine PA2G4-p48 (PDB entry 2V6C) (5). This structure was derived from an N- and C-terminally truncated PA2G4-p48 construct (aa Glu[8]–Ser[360]) and has continuous electron density observed between residues 8 and 360 (5). The second structure was a higher-resolution (1.6 Å) X-ray crystal structure of the human PA2G4-p48 (PDB entry 2Q8K) (6, 7). This was obtained with electron density observed between residues 7 and 362 and revealed a highly conserved structural fold to the murine PA2G4, with a root mean square deviation (RMSD), a measure of structural similarity, across the α-carbons of 0.44 Å. Notably, in this structure, a loop region (aa 283–292) does not fit the electron density well and is missing two amino acids (284 and 285). In 2019, a small 20-amino acid fragment of PA2G4 in complex with HNF4α was also solved
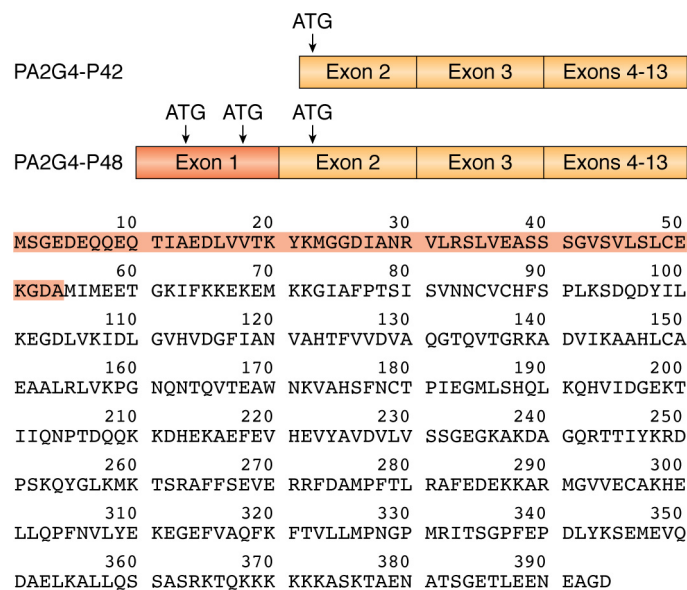
Figure 1. Diagrammatic representation and amino acid sequence of the two PA2G4 isoforms. PA2G4-p42 is translated from the third ATG codon, whereas the long PA2G4-p48 is translated from the first ATG codon. Exon one encodes for the first 54 amino acids, highlighted by the *red box*.

Table 1
Summary of the known PA2G4-p48 structures
For construct length, numbers in parenthesis indicate vector-derived residues. NA, not applicable.

PDB code	2V6C (5)	2Q8K (6, 7)	6SXO (8)
Year	2007	2007	2020
Construct length	(+2)Glu[9]–Ser[360]	(+8)Ser[2]–Asp[394]	Met[1]–Asp[394]
Electron density observed	(+1)Glu[9]–Ser[360]	Gln[7]–Ala[362]	Gln[7]–Ala[381]
Resolution (Å)	2.50	1.60	3.30
Method	X-ray	X-ray	Cryo-EM
Gene organism	*Mus musculus* (murine)	*Homo sapiens*	*H. sapiens*
Strain	BL21 (DE3)	BL21 (DE3)	Rosetta 2
Expression system	*Escherichia coli* BL21 (DE3)	*E. coli* BL21 (DE3)	*E. coli* Rosetta 2 cells
RMSD (to 2V6C) (Å)	NA	0.44	0.86

(PDB entry 6CHT), although only 7 amino acids (AELKALL) were resolved, and we will not discuss this structure further (9). Finally, in 2020 a cryo-EM structure of the human PA2G4-p48, ranging between 3.3 and 8 Å resolution, was elucidated in complex with the human 80S ribosome (PDB entry 6SXO) (8). Interestingly, this is the only structure that has the most C-terminal 33 amino acids resolved. Due to these additional amino acids, unless stated, this structure was utilized in all figures. The cryo-EM structure also showed good alignment to both the murine and human X-ray structures with an RMSD across the α-carbons of 0.86 Å for both.

All of the PA2G4-p48 structures revealed the highly conserved central PA2G4/Ebp1 structure, conserved between mice and humans, and have provided insights into the protein's functionality. These structures revealed that PA2G4 is structurally related to the family of Type II methionine aminopeptidases (Type II MAPs), due to its overall "pita bread" conformation with two α-helices enclosing two antiparallel β-sheets (Fig. 2) (6, 10). MAPs are metalloproteases that act to cleave the first methionine off of a growing polypeptide chain (11). Notably, PA2G4-p48 only has 25% sequence similarity to Type II MAPs, and the differences in the sequences highlight the differences in function. Importantly, although both have a conserved concave binding pocket, residues Asp[262], His[339], and Glu[459] in hMAP2

facilitate the catalytic function of the Type II MAPs. In PA2G4, these residues are replaced by Asn[120], Asp[195], and Lys[320] (Fig. 2). These substitutions alter the binding pocket charge environment, particularly the positive to negative charge swap with a change from Glu[459] to Lys[320] in PA2G4. Furthermore, unlike hMAP2, none of the solved PA2G4 crystal structures show any evidence that zinc ions or methionine could be present in this binding pocket. Lastly, PA2G4 contains additional C-terminal extensions (aa 337-394), helical (aa 205), and loop (aa 63–72 and 128–134) insertions, which differ from Type II MAPs (5). These structural differences may explain the inability of PA2G4 to perform hydrolytic substrate cleavage on a common MAP substrate. Thus, although PA2G4 adopts a hMAP2-like structure, it clearly has a unique functionality.

PA2G4 interactions with DNA and role in RNA processing and translation

The ability of PA2G4 to bind nucleic acids is well-documented (12–16). The PA2G4 protein has a predicted amphipathic helical domain (aa 204–246), suggesting that it may directly interact with DNA and/or proteins. PA2G4 has been shown to interact indirectly with E2F1 at E2F1 promoter elements by binding to a complex of nuclear proteins, including retinoblastoma (Rb) protein (17) and Sin3A (18, 19). The
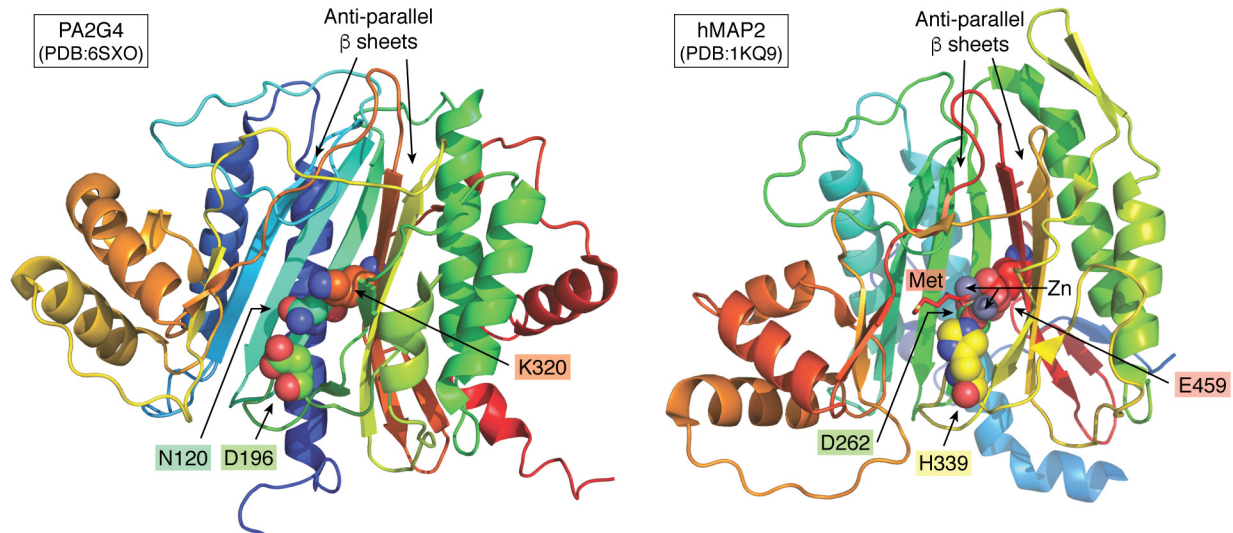
Figure 2. Comparison of PA2G4 and MAP-II structure. Both structures are *rainbow-colored* from the N terminus (*blue*) to the C terminus (*red*). Highlighted are the antiparallel β-sheets and the amino acid differences (PA2G4 Asn[120], Asp[196], and Lys[320] and hMAP-II Asp[262], His[339], and Glu[459]) that are hypothesized to make PA2G4 unable to act like an hMAP-II and catalyze the release of Met via its zinc ions (*Zn*).

activity of Rb and PA2G4 on the E2F1 promoter is regulated by the ErbB-3 ligand heregulin (20). The C-terminal domain of PA2G4 is an important DNA-binding domain. Xia *et al.* (17) showed that this domain (residues 300–372) was necessary for Rb binding and that the binding of PA2G4 to Rb is important in transcriptional repression in Rb+ cells, as PA2G4 mutants lacking the Rb-binding domain could no longer repress transcription of cyclin E. PA2G4 has also been shown to bind to HDAC2 both *in vitro* and *in vivo*, with the ability of PA2G4 to recruit histone deacetylase (HDAC) related to its ability to repress transcription. HDAC inhibitors such as trichostatin A can significantly reduce PA2G4-mediated repression (21). PA2G4 strongly binds protein kinase B (Akt) and suppresses apoptosis. Phosphorylated nuclear but not cytoplasmic Akt interacts with PA2G4 and enhances its antiapoptotic action independent of AKT kinase activity (12). Kim *et al.* (22) determined that PA2G4 also binds to the p53 E3 ligase HDM2, enhancing HDM2–p53 association and thereby promoting p53 polyubiquitination and subsequent degradation to decrease p53 activity. Furthermore, Kim *et al.* (23) proposed that PA2G4 modulates p53 levels by stabilization of the HDM2 protein and by facilitating HDM2 interaction with Akt.

PA2G4 is involved in transcriptional regulation and also seems to participate in the control of protein translation (14, 15). Analysis of the crystal structure of PA2G4 revealed two putative RNA-binding sites; the helix α5 (aa 205–213), which contains positively charged lysine residues often found in RNA interaction sites (6), and an extended positively charged surface patch, also with a lysine-rich sequence ([364]RKTQKKKKKK[373]), situated within the C-terminal extension (aa 338–394) (5, 6). Interestingly, it was originally suggested via sequence analysis that PA2G4 contained an s70-like domain (14) and a dsRNA-binding domain (24) and that these were the RNA-binding sites of PA2G4. Although deletion of these putative domains abrogated RNA binding, the crystal structures have since shown that neither of these domains are present in PA2G4, leading to the conclusion that the loss of function due to deletion of the

residues was likely to have severely disrupted protein folding (5). Furthermore, mutations within the lysine-rich C-terminal extension (aa 364–373) partially inhibit nucleolar localization (14). As the nucleolus is the main site of ribosome biosynthesis, this suggests that the RNA-binding activity of PA2G4 contributes to this function.

PA2G4 protein binds to a number of RNAs, including mRNA (13), rRNA, and ribonucleoprotein complexes (15, 24). PA2G4 seems to bind more strongly to ribosomal 5S RNA when compared with a mixture of t-RNAs or the S-domain RNA of the signal recognition particle, suggesting a direct interaction of PA2G4 with mature ribosomes via the 5S RNA (6). As a member of the ribonucleoprotein complex, which is the association of RNA and protein, PA2G4 contains a dsRNA-binding domain that mediates its interaction with dsRNA (14, 24). Deletion of this domain diminishes its localization to the nucleolus and reduces the formation of ribonucleoprotein complexes (24). In the cytoplasm, PA2G4 associates with mature ribosomes and acts as a cellular inhibitor of eukaryotic initiation factor 2α (eIF2α) phosphorylation by binding the dsRNA-activated kinase, PKR, which renders it incapable of phosphorylating eIF2α (24).

PA2G4 may also play a role in the processing of rRNA precursors and intermediates. It associates with nucleophosmin (B23) in the nucleolus, and ablation of either B23 or PA2G4 reduces ribosome biosynthesis by inhibiting the processing of the 32S intermediate for subsequent maturation of the 28S rRNA (25). This means PA2G4 acts as an endonuclease that cleaves the 38S pre-rRNA to generate the mature 28S rRNA. Fig. 3 summarizes the molecular interactions of PA2G4 in the cytoplasm and nucleus and its key role in cellular processes.

The function of structural regions of PA2G4

PA2G4 contains numerous structural regions (*i.e.* lysine-rich regions, an L*XX*LL motif, and numerous phosphorylation sites). The functions of these regions of PA2G4 have been
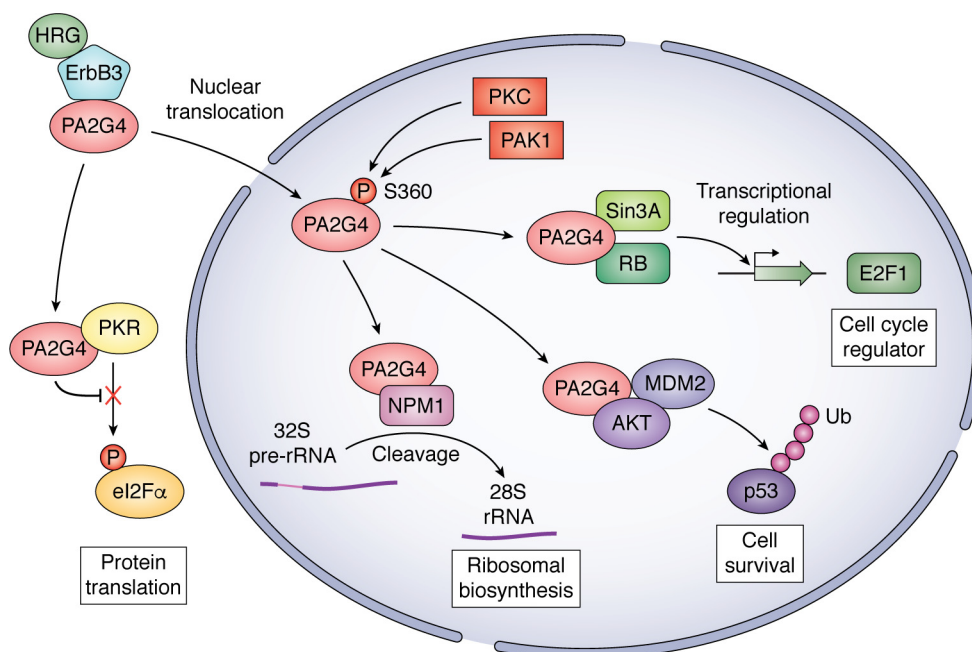
Figure 3. Key molecular interactions and cellular processes of PA2G4. PA2G4 binds to ErbB-3 in the cytoplasm; however, upon stimulation with heregulin (*HRG*), PA2G4 dissociates from ErbB-3 and translocates to the nucleus. PA2G4 is phosphorylated by protein kinase C and p21-activated kinase (PAK1). Phosphorylation of PA2G4 at serine 360 is needed for proper nuclear localization and for binding to ErbB-3 and nuclear Akt. PA2G4 binds to many proteins, including Sin3A and Rb, to enable transcription of E2F1, a cell cycle regulator. PA2G4 regulates cell survival by interacting with Akt and MDM2 to form a nucleoprotein complex to deregulate p53. PA2G4 also plays a role in the processing of rRNA precursors and intermediates by associating with nucleophosmin in the nucleolus and acting as an endonuclease that cleaves the 32S pre-rRNA to generate the mature 28S rRNA. In the cytoplasm, PA2G4 associates with mature ribosomes and acts as a cellular inhibitor of eIF2α (eukaryotic initiation factor) phosphorylation by binding the dsRNA-activated kinase, PKR, which renders it incapable of phosphorylating eIF2α, and is involved in protein translation.

elucidated predominantly via mutagenesis studies. We have summarized these studies and mapped them onto the PA2G4-p48 structure (Fig. 4).

Lysine-rich regions

There are three lysine-rich regions in the PA2G4-p48 structure (Fig. 4): one near the N terminus (Lys[20]–Lys[22]), another on loop 1 (Lys[65]–Lys[72]), and the third at the C terminus (Lys[365]–Lys[376]).

The N-terminal region is only present in the PA2G4-p48 isoform. This region has been suggested to be involved in nucleolar localization, as N-terminal deletion mutants of the protein are detected in the cytoplasm and excluded from the nucleus (14). Mutagenesis studies reveal that residues Lys[20]–Lys[22] on PA2G4 are essential for nuclear localization, with K20A/K22A mutants being unable to localize to the nucleus (14). Nuclear localization is a passive mechanism, and generally, the protein signal for nuclear localization is extended basic sequences (26). However, studies have shown that the essential component is a strong positive charge over a specific interaction surface (27). Lys[20]–Lys[22] are located on a solvent-exposed region of the PA2G4 protein. Interestingly, this position is adjacent to the second lysine-rich region, the highly basic loop 1 ([63]IFKKE-KEMK[71]). Due to the structural proximity of loop 1 and Lys[20]–Lys[22], alongside their importance in nuclear localization, it could be speculated that PA2G4 contains a bipartite nucleolar localization signal (6). Nuclear export and nucleolar localization and/or retention may also be regulated by the C terminus. Specifically, alanine mutation of either Arg[364] or Lys[365] resulted
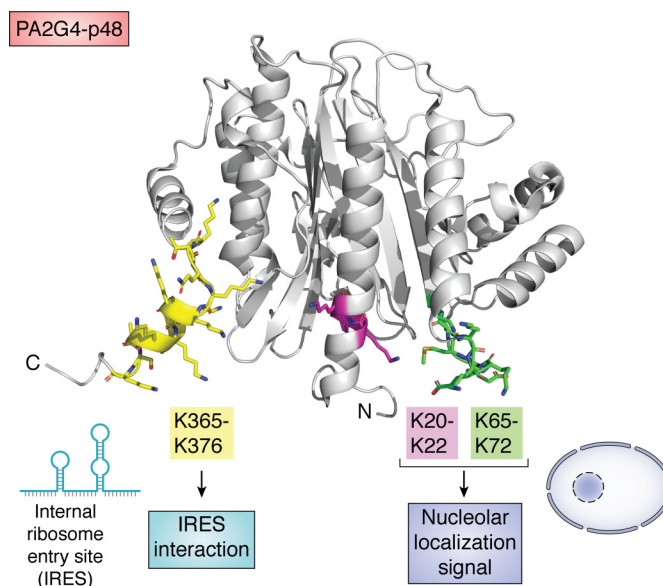

Figure 4. Lysine-rich regions of PA2G4. Shown is the structure of PA2G4-p48 with the three lysine-rich regions highlighted. *Magenta*, is Lys[20]–Lys[22]; *green*, Lys[65]–Lys[72]; *yellow*, Lys[365]–Lys[376]. The C-terminal region is proposed to bind to the IRES, leading to PA2G4's ITAS function. The other two sites are proposed to work together, leading to a bipartite nucleolar localization signal.

in only partial nucleolar localization (14). Lys[22] in the N-terminal region has also recently been found to be important for an interaction with the minor groove of RNA (8).

The C-terminal region of PA2G4 is highly positive, with six continuous lysine residues. Aside from its aforementioned role

in nucleolar localization, it is widely thought to be associated with PA2G4's role as an internal ribosome entry site (IRES)-*trans*-acting factor (ITAFs) (5). Initially, PA2G4 was identified as IRES-*trans*-acting factor (ITAF45) (1, 2), and its ITAF activity has been implicated in the pathogenesis of the neurovirulent Theiler's murine encephalomyelitis virus (28). ITAFs are RNA-binding proteins that have an additional role in the initiation of cap-independent translation (29, 30), which is the first major phase of protein biosynthesis (31). Little is currently known about the role of cellular IRES function; however, ITAFs are believed to be recruited from their cellular functions to assist the folding and/or stabilization of the active form of IRES (29, 30, 32–34). This in turn changes the binding affinity for the IRES to the translation apparatus, promoting the initiation of RNA translation (35). This is one of the two highly regulated translation initiation methods in eukaryotic mRNAs, the other being the cap-dependent initiation mechanism, which occurs in many mRNAs (36). Removal of the C-terminal region (construct containing only aa 1–360) led to PA2G4 being unable to bind to the IRES (5), suggesting an essential role of the C terminus. This was recently elucidated in the cryo-EM structure, where the C terminus of PA2G4 formed an α helix that filled the major groove of the RNA (8). Thus, the authors postulate that the C-terminal region of PA2G4 acts like the well-known arginine-rich motifs that are present in viral transactivation and signal recognition (8, 37, 38). Notably, this C terminus is a feature that distinguishes PA2G4 from the Type II MAPs and may relate to its specific functionality.

LXXLL motif and PA2G4-p42/PAG4-p48-isoform conserved protein-protein interaction sites

The C-terminal region has been shown to be an important protein-protein interaction binding domain, particularly in the repression of transcription. It contains an L*XX*LL motif, commonly known as a nuclear receptor interaction domain (39). Similar to nuclear receptors, the PA2G4 L*XX*LL motif ([354]LKALL[358], Fig. 5) is helical and forms the charge-clamp interaction that neutralizes the helix dipole (40). Notably, for PA2G4, this is likely to be a weaker interaction than the traditional nuclear receptor charge-clamp interaction, due to a slightly different helical architecture and the positively charged residue being positioned upstream of the L*XX*LL motif. This may lead to a greater diversity of binding partners. Monie *et al.* (5) suggests that dissociation of the [354]LKALL[358] helix from the surface of PA2G4 may allow for another L*XX*L motif to bind to the pocket. However, they note that a binding partner for this site has not been identified to date. Interestingly, the recent cryo-EM complex displayed this motif exposed from the rest of the 80S ribosome, enhancing the suggestion that this motif could bind proteins and affect the ITAF role of PA2G4 (8). To date, proteins shown to bind to the region include Rb, Akt, Sin3A, androgen receptor, and histone deacetylase 2 (HDAC2), which functions to repress transcription in cancer cell lines (12, 16, 20, 41, 42).

Our studies suggested that an alternate region of PA2G4-p48 might be a potential protein-protein interaction site, specifically interacting with the oncoprotein MYCN (43). To investi-
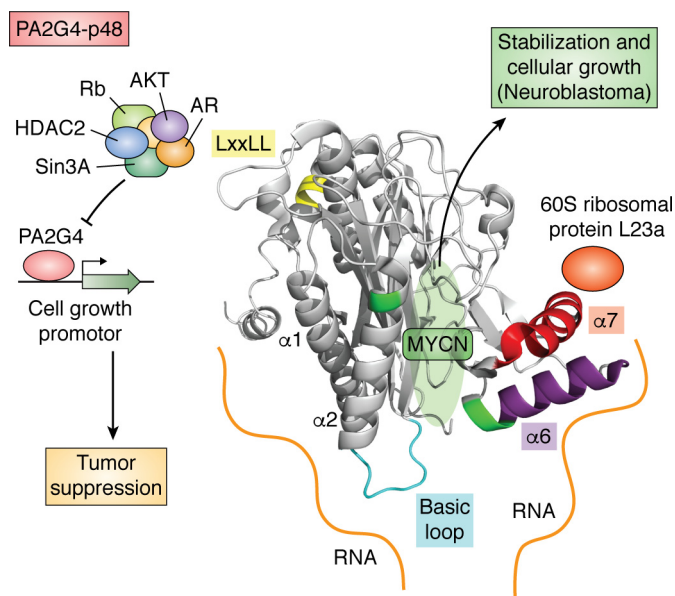

Figure 5. L*XX*LL region and known protein-protein binding sites of PA2G4-p48. Shown in *green* is the proposed position where MYCN binds to PA2G4 and the specific amino acids on α2 (Ser[57]) and α6 (Arg[271] and Arg[272]), which leads to stabilization of MYCN and cellular growth. The α6 helix (*purple*) along with the basic loop (*cyan*) interacts with the rRNA (*orange*), as shown in the cryo-EM structure. This structure also revealed that the α7 helix (*red cartoon*) may also be another potential protein-protein binding site, as it interacts with the 60S ribosomal protein L23a (shown as *red circle*), which is part of the larger 80S ribosome. Interestingly, the L*XX*LL motif (*yellow*) is completely accessible in the cryo-EM structure, further suggesting that this is a potential interaction site. This site is a major protein-protein interaction site and interacts with numerous proteins, including Rb, Akt, Sin3A, AR, and HDAC2. This stops the protein from then interacting with its downstream cell growth promotor, leading to tumor suppression.

gate this, we created a series of GST-MYCN deletions and analyzed their ability to bind a FLAG-tagged PA2G4-p48 (43). We identified a 7-residue sequence (aa 248–254), DHKALST, able to bind to PA2G4-p48 in a dose-dependent manner. Our computational modeling and mutagenesis studies confirmed that Ser[47], Arg[271], and Arg[272] were essential for a protein-protein interaction with MYCN, with alanine mutations resulting in reduced binding (via surface plasmon resonance) and a more stable MYCN (43) (Fig. 5). Interestingly, the PA2G4-p48 groove that we propose MYCN binds to is between two helices, α2 and α6, and positively charged loop 1 described above (Ile[63]–Lys[72]). These regions were all shown to interact with RNA in the human 80S ribosome (Fig. 5). Specifically, loop 1 (aa Ile[63]–Lys[72]) interacts with rRNA via residues Ile[63], Phe[64], Lys[65], and Lys[66], whereas from helix α6, Gln[254], Tyr[255], Leu[257], Lys[258], Arg[263], Phe[266], Ser[267], and Arg[271] were all shown to have direct and specific (mostly π–π and π–cation stacking) interactions with the His[59] loop of the rRNA (8). This lends further evidence that this may be an alternate, yet to be fully explored, protein-protein interface for PA2G4-p48.

Last, the cryo-EM structure also showed that helix α7 (Asp[287]–Leu[301]) also has the potential to interact with other proteins, specifically in this structure, 60S ribosomal protein L23a (Fig. 5) (8). To our knowledge, a role for this helix has not been described previously, and its interactions with ribosomal protein L23a are predominantly hydrophobic.
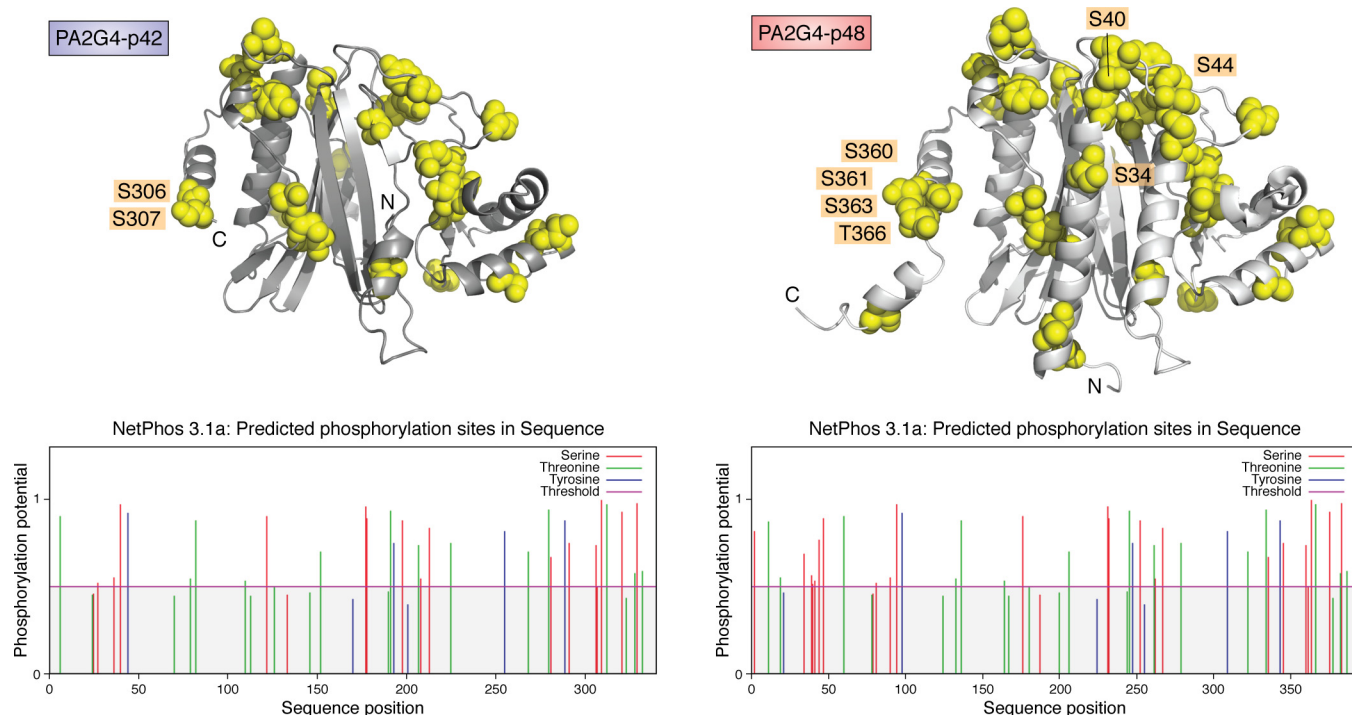
Figure 6. Potential phosphorylation sites. Potential phosphorylation sites as calculated from the NetPhos3.1 server (*yellow spheres*) are mapped onto PA2G4-p42 (*left*) and PA2G4-p48 (*right*). Highlighted are those confirmed as phosphorylation sites and the N and C terminus of both structures. Notably, many of the predicted phosphorylation sites on the p48 isoforms are actually buried behind the N terminus helix. This helix is missing in the P42 isoform; therefore, all potential sites could be post-translationally modified, leading to alternate functional effects.

Post-translational modification sites

Post-translational modifications, including phosphorylation, ubiquitination, sumolation, glycosylation, nitrosylation, methylation, acetylation, lipidation, and proteolysis, are commonly employed in the human proteome to regulate cellular activity. These modifications may lead to some of the functional differences between the PA2G4-p48 and PA2G4-p42 isoforms (described in detail in the next section of this review). Because phosphorylation is the most common type of post-translational modification, we used the NetPhos 3.1 Server to predict Ser, Thr, and Tyr phosphorylation sites based upon the sequence of the PA2G4-p48 and PA2G4-p42 isoforms and subsequently mapped them to their respective structures (Fig. 6). The sequence analysis showed that the PA2G4-p48 isoform contained 42 potential phosphorylation sites, whereas the PA2G4-p42 isoform contained only 33 potential sites. Notably, for the PA2G4-p48 isoform, seven of these are buried, and four are beyond the region that has been crystallized to date, whereas for the PA2G4-p42 isoform, all of these have the potential to be surface-exposed and thus available for phosphorylation. Notably, the lack of PA2G4-p42 experimentally determined structures means this is yet to be confirmed.

Two of these sites, present in both isoforms, are adjacent to each other, Ser[360] and Ser[361], and numerous studies have shown that phosphorylation of these residues is essential for nucleolar localization (4) and/or interactions with the receptor tyrosine-protein kinase erbB-3 (ERBB3), the E3 ubiquitin-protein ligase Bre1 (HUWE1), and Akt (12, 44, 45). Phosphorylation of proteins can modulate the nature or strength of protein-protein interactions by a variety of methods, including

conformationally changing the protein structure and/or changing the binding energy of a complex (reviewed in Ref. 46). Considering that this phosphorylation site is adjacent to the L*XX*LL motif described above, it would be reasonable to deduce that this is a major protein-protein interaction site for numerous proteins. The structures solved thus far suggest that Ser[360] is slightly more exposed and available for phosphorylation than Ser[361], and we have chosen to present the results of the studies thus far based upon the serine the authors of those studies have pointed to. Notably, being adjacent to each other, there may be some ambiguity if only one serine or both are phosphorylated.

The role of Ser[361] in nucleolar localization was elucidated by the use of the PA2G4 mutants S361A and S361D (4). The S361A mutant causes the Ser to be unable to be phosphorylated and leads to the protein remaining in the nucleus regardless of growth factor stimulation. In contrast, the PA2G4 S361D mutant, a mutant mimicking protein kinase C (PKC)-mediated phosphorylation, resulted in no PA2G4 detected in the nucleolus, therefore suggesting that phosphorylation of this residue is essential for nuclear localization. Notably, however, this residue is present in both isoforms; thus, it is unclear why only the larger PA2G4-p48 isoform is found in the nucleus.

The crystal structures indicate that Ser[360] is more solvent-exposed than Ser[361] and thus more likely to be phosphorylated. There are other studies showing the role of Ser[360] phosphorylation, particularly its functional role in protein-protein interactions. Studies have shown that PKC-δ–triggered phosphorylation on Ser[360] is responsible for the association of PA2G4-p48 with nuclear Akt and prevention of apoptosis (12). By using mutants similar to the aforementioned PA2G4 Ser[361]

mutants, Ahn *et al.* (12) showed that PA2G4-p48 S360A, which cannot be phosphorylated, barely binds Akt, whereas PA2G4-p48 S360D, which mimics phosphorylation, binds strongly to nuclear Akt. Using the same S360D mutant, Liu *et al.* (45) noticed that the PA2G4-p42 isoform was strongly polyubiquitinated when compared with WT PA2G4-p42. Even more interesting, they found that the PA2G4-p48 isoform was unable to undergo the same polyubiquitination. After conducting pulldown assays and co-immunoprecipitation, they concluded that the ubiquitination was due to the binding of the E3 ubiquitin-protein ligase, Bre1. Notably, both the PA2G4-p48 and PA2G4-p42 isoforms were able to interact with Bre1; however, only its interaction with the phosphorylated Ser[360] PA2G4-p42 isoform led to polyubiquitination and subsequent degradation. The authors concluded that this difference in function may be due to the location of the isoform (*i.e.* PA2G4-p42 is located predominantly in the cytoplasm, whereas PA2G4-p48 is located in both the cytoplasm and nucleus). Because Bre1 is a large protein (975 aa) and the protein folding of PA2G4 means that the N terminus of PA2G4-p48 is only approximately 20 Å away from this ubiquitination site, could it also be possible that the N terminus of PA2G4-p48 plays a role in inhibiting Bre1's ability to cause ubiquitination? Ser[363] and Thr[366] are located at the C terminus of PA2G4 and have been shown to be phosphorylated by PKC (44). Phosphorylation was shown to be enhanced by the presence of heregulin, but the functional effect of this phosphorylation was not clearly elucidated. With at least four known phosphorylation sites, it is clear that this C-terminal region is of significant importance in controlling the function of PA2G4 either via its ability to interact with other proteins or its ability to impact nuclear localization (due to the close proximity signal described above).

Three serine residues in the N terminus, Ser[34], Ser[40], and Ser[44], have also been confirmed as PA2G4-p48 phosphorylation sites (47, 48). Importantly, these three sites are unique to the PA2G4-p48 isoform and are thought to be partly responsible for the oncogenic role of the PA2G4-p48 isoform (discussed in more detail under "The tumor-promoting role of PA2G4-p48"). Ser[34] was shown to be phosphorylated by cyclin-dependent kinase 2 (CDK2), which led to tumor growth (49). Importantly, this growth could be ablated via a PA2G4-p48 S37A mutation (48). It is likely that this phosphorylation is increasing the affinity for another protein, although which protein(s) this is or how this leads to an increase in tumor growth remains to be determined. PA2G4-p48 Ser[40] and Ser[44] were explored by Wang *et al.* (47), who demonstrated that the PA2G4-p48 sequence [40]SSGVS[44] resembles the putative glycogen synthase kinase 3β (GSK3β) phosphorylation sequence (Ser/Thr-*X*-*X*-*X*-Ser/Thr). An important feature of this sequence is that GSK3β usually requires priming phosphorylation on Ser or Thr at position +4 in order to phosphorylate (50, 51); therefore, they hypothesized that one Ser would be a priming site, which in turn allows the other to be phosphorylated. Subsequent co-immunoprecipitation studies with WT PA2G4-p48 and a S44A mutant suggested that Ser[44] was the priming phosphorylation site for GSK3β recruitment, with subsequent Ser[40] phosphorylation by GSK3β (47). Importantly, phosphorylation of either Ser[40] or Ser[44] is essential for the interaction between PA2G4-

p48 and F-box/WD repeat–containing protein-7 (FBXW7) and subsequent degradation of PA2G4-p48 via the recruitment of ubiquitinating proteins (47). Therefore, these phosphorylation sites are important regulatory sites for PA2G4-p48. The validity for the other phosphorylation sites suggested by NetPhos remains to be experimentally evaluated; however, it would be intriguing to see whether the buried PA2G4-p48 phosphorylation sites, exposed in the PA2G4-p42 isoform, are important for the different functions of the isoforms.

Functional and structural differences between the PA2G4-p42 and PA2G4-p48 isoforms

It has been established that the two PA2G4 isoforms have opposing roles in cancer, and the underlying structural reason for this has been the subject of numerous studies in recent years (reviewed in Ref. 52). The longer p48 isoform of PA2G4 localizes to the cytoplasm and nucleus, suppressing apoptosis and promoting cellular growth. In contrast, the shorter PA2G4-p42 isoform of PA2G4, which lacks the N-terminal 54 amino acids present in PA2G4-p48, is confined predominantly to the cytoplasm, where it acts to promote differentiation and suppress proliferation (4). Monie *et al.* (5) suggested that the PA2G4-p42 isoform was likely to lack one and a half helices from the N terminus of PA2G4-p48. Our modeling studies confirm that PA2G4-p42 is missing the entire helix α1 and half of α2 (Fig. 7A). In agreement with Monie *et al.* (5), we also show that truncation of this N terminus is likely to expose a large hydrophobic surface (Fig. 7B), which, when exposed to solvent, would have a destabilizing effect and may explain the significantly lower levels of PA2G4-p42 expression (4). We extended this analysis and generated electrostatic potential surfaces using Pymolv2.3.2 (Schrodinger, LLC). This confirmed the hydrophobic patch on the PA2G4-p42 isoform (Fig. 7B). Furthermore, due to the removal of the N terminus, we found different electrostatic potential profiles across this face of the protein, with the PA2G4-p42 isoform being more positive and the PA2G4-p48 isoform more negative (Fig. 7B). This may account for alternate protein-binding patterns in cells.

The differing N termini between the isoforms has also been suggested to be responsible for different interactions with protein binding partners. It has been shown that PA2G4-p42 but not PA2G4-p48 is able to bind to ErbB3 (4). Mechanistically, this is thought to be due to the N-terminal 54-residue extension of PA2G4-p48, which blocks/masks the ErbB3-binding motif on the PA2G4-p42 isoform, preventing this interaction from occurring (4). This is confirmed by our modeling studies and others, which show a large hydrophobic surface on the PA2G4-p42 isoform, postulated to be the binding pocket of ErbB3. Conversely, the larger N-terminal domain of PA2G4-p48 has been shown to be essential for other protein interactions. Specifically, the HDM2 protein has been shown to bind tightly to the PA2G4-p48 isoform, but not an N-terminal domain mutant (aa 48–394) that mirrors the PA2G4-p42 isoform (22). As discussed above, this truncation also removes Lys[20] and Lys[22], which are shown to be essential for nucleolar localization (14).

This minor structural difference is an intriguing site to target for small-molecule inhibition. Drugs that interact only with the
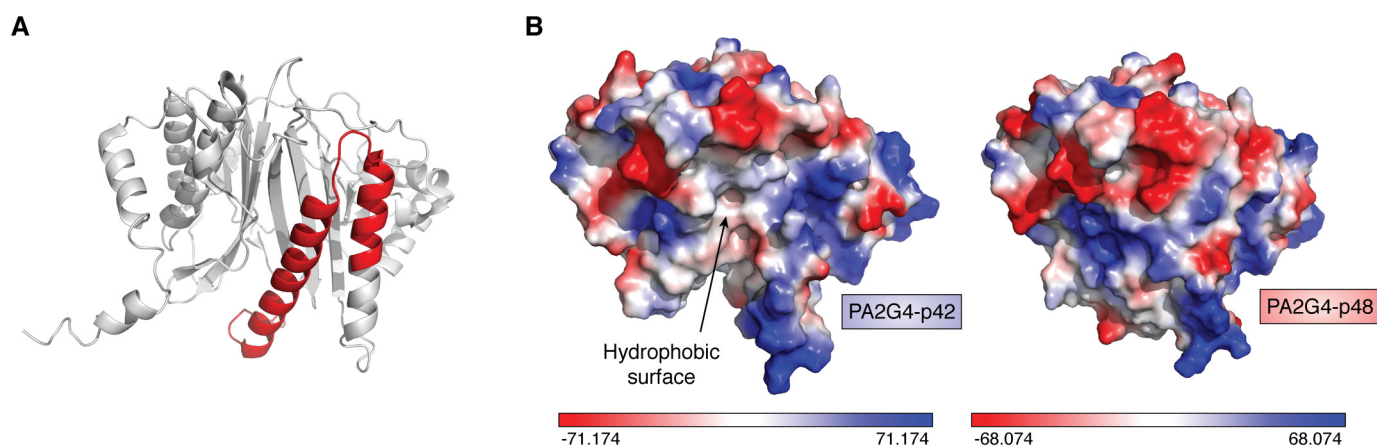
Figure 7. Differences between the p42 and p48 isoforms of PA2G4. Our models showed that the truncation involves the removal of the first helix and part of the second (*A*, highlighted in *red*). This results in the exposure of a hydrophobic surface (*white surface* shown by an *arrow*; *B, left*), which is proposed as a binding surface for ErbB3. Notably, this also changes the electrostatic potential on this surface of the PA2G4 protein, where the large negatively charged region of PA2G4-p48 (shown in *red*; *B, right*) is now hydrophobic (shown in *white* and highlighted by an *arrow*; *B, left*), which may result in their different functional roles.

PA2G4-p48 isoform would have the potential to inhibit the oncogenic role of PA2G4-p48, while having little effect on the tumor-suppressive role of PA2G4-p42. In our study, we used structure-based drug design methods to target a binding site boarded by helix $\alpha2$ (Ser[44]–Lys[62]) (43). We hypothesized that the known PA2G4 inhibitor WS6 (53) bound to the same groove on PA2G4 that MYCN binds to, a groove boarded by helix $\alpha2$ (43). Further biological assessment of WS6 confirmed that it had an inhibitory effect on PA2G4-p48's oncogenic role in malignant neuroblastoma cells *in vitro* and *in vivo* (43). Due to the lack of helix $\alpha2$ in PA2G4-p42, we now hypothesize that this compound would not be able to bind to the shorter PA2G4-p42 isoform effectively and therefore should have no effect on this isoform's function. However, further experiments are needed to support our hypothesis.

Another interesting drug discovery approach would be to selectively repress PA2G4-p48, while also enhancing PA2G4-p42 via a protein partner. Insights into this approach may be obtained from one of the most interesting cases of a PA2G4-binding protein, FBXW7. FBXW7 has been characterized as a general tumor suppressor as it functions as a substrate recognition subunit of the SCF (SKP1/CUL1/F-box protein) E3 ubiquitin ligase complex (54), which in turn regulates a network of proteins with central roles in cell division, growth, and differentiation (55, 56). FBXW7 has been shown to bind to both the PA2G4-p42 and PA2G4-p48 isoforms but to elicit different responses in tumor development (47). Using mutagenesis and immunoprecipitation, the authors showed that only the WD40 domain of FBXW7 interacts with the N-terminal tail of PA2G4-p48, specifically involving the Ser[40] and Ser[44] residues discussed above being phosphorylated by GSK3β, leading to ubiquitination and proteasomal degradation of the p48 isoform (47). Interestingly, this interaction formed a negative regulation loop where PA2G4-p48 binds to FBXW7. This result masks FBXW7's nuclear localization signal, thereby altering its cellular location and subsequent ubiquitination functions on other oncoproteins. However, although it is missing the N terminus, the PA2G4-p42 isoform was still able to bind to FBXW7 (47).

Through extensive mutagenesis and immunoprecipitation, the authors showed a differing interaction profile between FBXW7 and the PA2G4-p42 isoform, with the shorter isoform interacting with both the F-box and the WD40 domains of FBXW7. Because F-box domains act to recruit ubiquitin-conjugating enzymes, they postulated that the PA2G4-p42 isoform acts as a scaffolding protein for FBXW7, stabilizing and enhancing its tumor-suppressing role (47). Although mirroring the endogenous interactions of FBXW1 with both PA2G4 isoforms via therapeutics would be difficult, it is intriguing to speculate about other proteins that may be interacting with both isoforms but eliciting alternate roles. The alternate roles of the isoforms in tumor development are further discussed below.

Tumor-suppressive role of PA2G4-p42

The shorter PA2G4-p42 isoform is generally found at much lower levels in the cell (4). Although recently there is some evidence that the longer isoform may also be ubiquitinated (47), early studies suggested that the shorter isoform was the only isoform of PA2G4 able to be ubiquitinated (45). Regardless, there is strong evidence that the PA2G4-p42 isoform alone is ubiquitinated by the E3 ubiquitin-protein ligase, Bre1. This E3 binds to the phosphorylated PA2G4-p42 isoform, specifically requiring phosphorylation of residue Ser[360], promoting its polyubiquitination and degradation (45). This process is thought to be enhanced in cancer cells, which in turn reduces any tumor-suppressive actions of the PA2G4-p42 isoform. A series of studies have been conducted to elucidate the cellular pathways underlying the tumor suppressor function of PA2G4-p42. Specifically, these studies have shown that PA2G4 binds to proteins including androgen receptor (AR), phosphatidylinositol 3-kinase (PI3K) (49), Erb-3 receptor (57), and Rb (17), preventing them from binding to their specific cell growth regulators. The PA2G4-p42 isoform has been shown to be tumor-suppressive in prostate, colon, breast, non-small-cell lung carcinoma, and salivary adenoid cystic carcinoma cells and/or models (16, 20, 41, 42, 48, 58).

For prostate cancer, overexpression of PA2G4-p42 in LNCaP androgen-sensitive human prostate adenocarcinoma cells produced a decrease in tumorigenic growth both *in vitro* and in xenograft models (16). Similarly, siRNA knockdown of PA2G4-p42-Ebp1 in LNCaP cells increased AR activation (42). Overexpression of the PA2G4-p42 isoform has been shown to significantly down-regulate six genes in the AR signaling pathway; AR, prostate-specific antigen (PSA; kallikrein 3), kallikrein 2, POV-1, TMPRSS2, and prostate-derived factor (16). As described above, the C-terminal region of PA2G4 has been shown to directly bind to the AR (41) and Sin3A (16), a known co-repressor of AR. Interestingly, although this C terminus is conserved between the PA2G4-p42 and PA2G4-p48 isoforms, only the p42 isoform has been shown to bind to the AR. Therefore, it would not be unreasonable to postulate that the PA2G4-p42 isoform acts as part of a larger complex involving AR, Sin3A, and HDAC, all of which bind to and repress PSA gene transcription (16). The significant tumor-suppressive role seen by the PA2G4-p42 isoform in prostate cancer, may be a result of PA2G4-p42 not only binding to and suppressing the AR, but also via down-regulation of PI3K activity (49). PA2G4-p42 interacts with the p85 subunit of PI3K, encouraging the recruitment of the chaperone-E3 ligase complex HSP70/CHIP, which in turn leads to ubiquitination and degradation of PI3K (49). The central role of the PI3K growth-signaling pathway in numerous cancers (59), the ability of the PA2G4-p42 isoform to suppress the PI3K growth-signaling pathway and subsequently reduce Akt activation, may explain the general tumor-suppressive role of the PA2G4-p42 isoform.

The tumor-suppressive pathways for the PA2G4-p42 isoform in breast cancer have also been well-studied. PA2G4 was first discovered as an ErbB-3 receptor-binding protein due to its ability to interact with the cytoplasmic tail of ErB3 (3). ErbB3 belongs to the ErbB family of receptor tyrosine kinases, which modulate intracellular signaling cascades and are implicated in breast cancer oncogenesis. When this receptor is stimulated with the ErbB-3/4 ligand heregulin in AU565 breast cancer cells, the PA2G4-p42 isoform dissociates from the ErbB-3 receptor and is translocated from the cytosol to the nucleus, where it undergoes phosphorylation (3, 4, 44, 57). PA2G4 also reduces the protein levels of another Erb family member, ErbB2, also known as HER2 (19, 60). In another example of its multifunctional nature, PA2G4 both increases ubiquitination of ErbB2 and suppresses intracellular ErbB2 mRNA promotors, either directly or via a complex of proteins (60). Interestingly, this study showed no significant difference in the expression levels of either PA2G4 isoform, and cells transfected with both PA2G4 isoforms show an increase in tamoxifen sensitivity (60). Does this suggest that the Erb family has a similar profile as FBWX1 described above and interacts with the alternate isoforms of PA2G4 differently? More biochemical experiments will be needed to explore this; however, it is clear that through multiple interactions with the Erb family, PA2G4-p42 is postulated to act as a tumor suppressor.

An additional tumor-suppressive role for PA2G4-p42 in breast cancer is via its interaction with the Rb protein, which acts as a tumor suppressor and is often down-regulated in solid tumors, including breast cancer (61). The C-terminal residues 300–372 of the PA2G4-p42 isoform bind directly to the Rb protein (17). The interaction between the PA2G4-p42 motif and Rb leads to the inhibition of the E2F1 transcription factor and a decrease in cyclin E levels, two proteins whose activity promotes oncogenesis (41, 62). This was shown experimentally by Zhang *et al.* (41), who demonstrated that PA2G4-p42 can inhibit the transcriptional activity in PA2G4-p42 isoform–transfected MCF-7 cells. In addition, PA2G4-p42 inhibits the transcription of other E2F1-regulated promoters, such as the c-MYC and cyclin D1. Notably, ectopic expression of PA2G4-p42 in AU565 cells inhibits proliferation and induces cellular differentiation, in a similar manner to treatment with the Erb-3/4 ligand, heregulin (described above) (57). The inhibition by PA2G4 of the E2F-regulated promotors is enhanced by heregulin treatment (20). This suggests a synergistic role between PA2G4 and heregulin with at least two alternate pathways in breast cancer, where PA2G4-p42 plays a role as an intracellular signaling protein with the ability to interact with transcriptional modulators. These experiments in the breast cancer setting have also elucidated that PA2G4-p42 is unlikely to repress transcription alone. Rather, its transcriptional repression activity seems to be partially dependent on recruitment of HDAC activity, specifically HDAC2, via the C-terminal 293–372 aa motif of PA2G4-p42 (41). As in the prostate cancer example above, these residues are conserved between the PA2G4-p48 and PA2G4-p42 isoforms; therefore, the exposed hydrophobic face and differential phosphorylation profiles are likely to be the differential features. Regardless, there is strong evidence that the expression levels of PA2G4 could potentially be used to guide clinical treatment decisions.

Although the tumor suppressor role of the p42 isoform has been established in breast and prostate cancer, a recent study found that PA2G4-p42 also acts as an oncoprotein, and both PA2G4 isoforms increase the malignant phenotype of neuroblastoma cells *in vitro* (43). Further studies are therefore required to explore the mechanism underlying the apparent contextual role of PA2G4-p42 cancer.

The tumor-promoting role of PA2G4-p48

The long PA2G4-p48 isoform has numerous roles in cancer, including inhibiting the tumor-suppressing activity of p53 (22), binding to mRNA and promoting up-regulation of antiapoptotic proteins (13), directly binding to and stabilizing the MYCN oncoprotein in a positive forward feedback expression loop (43), and stabilizing proteins by down-regulating their degradation via the ubiquitination pathway (47). The PA2G4-p48 isoform has been shown to directly interact with Bcl-2 protein mRNA, promoting the expression of this antiapoptotic protein (13). The phosphorylated PA2G4-p48 isoform interacts with nuclear Akt and prevents apoptotic cell death (12). Furthermore, knockdown of PA2G4-p48 increases the levels of endogenous p53 in glioblastoma cells. The PA2G4-p48 isoform also directly binds HDM2, which promotes HDM2-dependent p53 ubiquitination (22). Combining this with the evidence that the PA2G4-p48 isoform is highly expressed in more aggressive cancers (22, 63) suggests that the reduction in p53 levels may

be regulated by the levels of PA2G4-p48 as a mechanism of oncogenesis.

Knockdown of PA2G4-p48 in muscle stem cells and loss-of-expression studies in mice and chick embryos indicate a role in regulating proliferation and differentiation (64). Consistent with a role in proliferation, PA2G4-p48 knockout mice are 30% smaller compared with WT littermates and show cellular hallmarks associated with growth retardation (65). Cellular analysis from these mice demonstrated a decrease in mouse embryonic fibroblast growth. Biochemical analysis revealed a decrease in insulin-like growth factor 1 (IGF1) and a paradoxical increase in MAPK signaling transduction pathway activity, which the authors deduce to be related to differing cellular backgrounds. These findings were supported by an observed increase in expression of Seveneless homolog 1 (SOS), an upstream activator of the Ras-MAPK signaling cascade, and suggest a complex role of PA2G4-48 in the MAPK signaling pathway. The MAPK signaling pathway has a long-established overall role to sustain cell survival, proliferation, and differentiation in eukaryotic cells, and as a consequence, the MAPK signaling pathway is a promising therapeutic target in multiple cancers (66). Therefore, it is not surprising that high expression of PA2G4-p48 has also been implicated in numerous cancers.

Our studies showed that in neuroblastoma cells, PA2G4-p48 targeted siRNA knockdown reduced MYCN protein levels (43). This decrease was normalized in the presence of the proteasomal inhibitor MG132, suggesting that PA2G4-p48 protects MYCN from proteasomal-mediated degradation. Furthermore, loss of PA2G4-p48 appears to sensitize neuroblastoma cells to 13-*cis*-retinoic acid treatment, suggesting that PA2G4-p48 may contribute to the therapy resistance observed in MYCN-amplified neuroblastoma cells (43). For these reasons, PA2G4-p48 is an attractive target for therapeutic inhibition of the MYCN signaling pathway in neuroblastoma and other MYCN-amplified cancers. Our studies revealed that PA2G4 was able to bind directly to MYCN, leading to an oncogenic forward feedback loop of high expression levels of PA2G4-p48 and MYCN proteins (43).

PA2G4-p48 has also been implicated in the tumorigenesis of oral squamous cell carcinoma (OSCC). As demonstrated by Mei *et al*. (67), PA2G4-p48 expression in OSCC cells is linked with that of podoplanin, a mucin-type protein. Podoplanin is both a prognostic and diagnostic marker in various cancers (68). PA2G4-p48 been shown to bind to a DNA sequence in the podoplanin gene promoter, suggesting a role of PA2G4-p48 in podoplanin transcriptional regulation. In line with these findings, down-regulation of PA2G4-p48 results in a reduction in malignant phenotypes, including invasion and anchorage-independent growth (67). In contrast to these findings, ectopic expression of full-length PA2G4 cDNA in OSCC cell lines results in decreased growth and proliferation (69). Similar findings have been demonstrated in salivary adenoid cystic carcinoma (58). It remains unclear in these studies which isoform of PA2G4 is modulating cellular function. Thus, the role of PA2G4 in oral cancers requires further investigation.

Finally, the role of PA2G4-p48 in acute myeloid leukemia is an emerging area of research. A recent study has demonstrated that peripheral blood samples from patients with acute myeloid leukemia demonstrate increased PA2G4-p48 expression, with an associated increase in the proliferating cell nuclear antigen oncoprotein. In these cell lines, PA2G4-p48 has been shown to bind to RNA polymerase I to increase the synthesis of rRNA (70).

Notably, aside from the aforementioned cancers, poor clinical outcomes have been associated with high PA2G4-p48 expression in patients with neuroblastoma (43), glioblastoma (22), cervical cancer (71), prostate cancer (72), and breast cancer (63).

Conclusions

The proliferation-associated protein 2G4 (PA2G4) has two distinct isoforms, which function as a contextual tumor suppressor or oncoprotein, depending on the tissue type being studied. Although this review has bought together the structural and functional studies regarding PA2G4, much is still to be elucidated. There is only one significant structural difference between the two isoforms (*i.e.* the p42 is missing the most N-terminal 54 amino acids); however, they have opposing functional roles in the context of cancer. Some of this can be explained by the exposure of a hydrophobic surface pocket in the PA2G4-p42 isoform, which is likely to provide a protein-protein interaction site that is unique for this shorter isoform. This exposed surface also contains alternate potential phosphorylation sites, although to our knowledge none of these have been experimentally verified. Another reason for the alternate functional role may be the location of the protein, with the PA2G4-p48 isoform containing an N-terminal nucleolar localization signal. However, it is more likely that the additional phosphorylation sites in this N-terminal region (namely Ser^{34}, Ser^{40}, and Ser^{44}), which are known to bind to specific proteins, play a larger role in its oncogenic function.

As PA2G4 has a distinct role in cancer progression, it may serve as a novel therapeutic target. However, because the two isoforms have opposing functional roles, care needs to be taken to enable targeting of the oncogenic isoform only. In this review, we have highlighted the multiple protein-protein interaction sites that may be targeted by small-molecule therapeutics (73). In designing these inhibitors, it will be important to focus on sites essential for the oncogenic role of PA2G4, such as its interaction site with MYCN, which we postulate also binds to WS6, a PA2G4 inhibitor (43). An alternate approach might be to focus on protein partners, such as FBXW1, which interacts with both isoforms but elicits alternate functional effects in the cell (47). Notably, both of the aforementioned studies were published in the last 3 years; thus, it is likely that we have yet to discover the full extent of protein partners.

There are still some confounding results, especially relating to the ability of PA2G4-p42 to act as an onco-suppressor by interacting with other proteins via its C-terminal region, a region conserved in the larger PA2G4-p48 isoform. Potentially, this highlights a competitive binding nature, although additional biochemical studies are needed that can link the structure-function of the alternate PA2G4 isoforms. Although not discussed in this review, there is also the potential to use genetically engineered mouse models to express the different PA2G4

isoforms and assess its biological impact in normal and cancer cells. Using PA2G4-p48 knockout mice, one could investigate whether depletion of PA2G4-p48 can delay or prevent tumor growth when crossed with murine cancer models (*i.e.* TH-MYCN mice to study neuroblastoma).

With this information in mind, it is clear that there is still much to be elucidated regarding the subtle structural changes that lead to the differing functional effect of the PA2G4 isoforms. The recent cryo-EM structure highlights the benefit of obtaining structures of complexes where these subtle binding differences may be seen. Because both PA2G4 isoforms have therapeutic potential across the field of oncology, any future PA2G4 structures, particularly those in complex with binding partners, will be an important step in understanding the biology of this protein.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: aa, amino acids; PDB, Protein Data Bank; RMSD, root mean square deviation; MAP, methionine aminopeptidase; eIF2α, eukaryotic initiation factor 2α; IRES, internal ribosome entry site; ITAF, IRES-*trans*-acting factor; AR, androgen receptor; PI3K, phosphatidylinositol 3-kinase; Rb, retinoblastoma; PSA, prostate-specific antigen; MAPK, mitogen-activated protein kinase.

References

1. Lamartine, J., Seri, M., Cinti, R., Heitzmann, F., Creaven, M., Radomski, N., Jost, E., Lenoir, G. M., Romeo, G., and Sylla, B. S. (1997) Molecular cloning and mapping of a human cDNA (PA2G4) that encodes a protein highly homologous to the mouse cell cycle protein p38-2G4. *Cytogenet. Cell Genet.* **78**, 31–35 CrossRef Medline

2. Radomski, N., and Jost, E. (1995) Molecular cloning of a murine cDNA encoding a novel protein, p38-2G4, which varies with the cell cycle. *Exp. Cell Res.* **220**, 434–445 CrossRef Medline

3. Yoo, J. Y., Wang, X. W., Rishi, A. K., Lessor, T., Xia, X. M., Gustafson, T. A., and Hamburger, A. W. (2000) Interaction of the PA2G4 (EBP1) protein with ErbB-3 and regulation of this binding by heregulin. *Br. J. Cancer* **82**, 683–690 CrossRef Medline

4. Liu, Z., Ahn, J. Y., Liu, X., and Ye, K. (2006) Ebp1 isoforms distinctively regulate cell survival and differentiation. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10917–10922 CrossRef Medline

5. Monie, T. P., Perrin, A. J., Birtley, J. R., Sweeney, T. R., Karakasiliotis, I., Chaudhry, Y., Roberts, L. O., Matthews, S., Goodfellow, I. G., and Curry, S. (2007) Structural insights into the transcriptional and translational roles of Ebp1. *EMBO J.* **26**, 3936–3944 CrossRef Medline

6. Kowalinski, E., Bange, G., Bradatsch, B., Hurt, E., Wild, K., and Sinning, I. (2007) The crystal structure of Ebp1 reveals a methionine aminopeptidase fold as binding platform for multiple interactions. *FEBS Lett.* **581**, 4450–4454 CrossRef Medline

7. Kowalinski, E., Bange, G., Wild, K., and Sinning, I. (2007) Expression, purification, crystallization and preliminary crystallographic analysis of the proliferation-associated protein Ebp1. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **63**, 768–770 CrossRef Medline

8. Wild, K., Aleksić, M., Lapouge, K., Juaire, K. D., Flemming, D., Pfeffer, S., and Sinning, I. (2020) MetAP-like Ebp1 occupies the human ribosomal tunnel exit and recruits flexible rRNA expansion segments. *Nat. Commun.* **11**, 776 CrossRef Medline

9. Han, E. H., Singh, P., Lee, I. K., Urrutia, R., and Chi, Y. I. (2019) ErbB3-binding protein 1 (EBP1) represses HNF4α-mediated transcription and insulin secretion in pancreatic β-cells. *J. Biol. Chem.* **294**, 13983–13994 CrossRef Medline

10. Bazan, J. F., Weaver, L. H., Roderick, S. L., Huber, R., and Matthews, B. W. (1994) Sequence and structure comparison suggest that methionine aminopeptidase, prolidase, aminopeptidase P, and creatinase share a common fold. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 2473–2477 CrossRef Medline

11. Ball, L. A., and Kaesberg, P. (1973) Cleavage of the N-terminal formylmethionine residue from a bacteriophage coat protein *in vitro*. *J. Mol. Biol.* **79**, 531–537 CrossRef Medline

12. Ahn, J. Y., Liu, X., Liu, Z., Pereira, L., Cheng, D., Peng, J., Wade, P. A., Hamburger, A. W., and Ye, K. (2006) Nuclear Akt associates with PKC-phosphorylated Ebp1, preventing DNA fragmentation by inhibition of caspase-activated DNase. *EMBO J.* **25**, 2083–2095 CrossRef Medline

13. Bose, S. K., Sengupta, T. K., Bandyopadhyay, S., and Spicer, E. K. (2006) Identification of Ebp1 as a component of cytoplasmic *bcl-2* mRNP (messenger ribonucleoprotein particle) complexes. *Biochem. J.* **396**, 99–107 CrossRef Medline

14. Squatrito, M., Mancino, M., Donzelli, M., Areces, L. B., and Draetta, G. F. (2004) EBP1 is a nucleolar growth-regulating protein that is part of pre-ribosomal ribonucleoprotein complexes. *Oncogene* **23**, 4454–4465 CrossRef Medline

15. Pilipenko, E. V., Pestova, T. V., Kolupaeva, V. G., Khitrina, E. V., Poperechnaya, A. N., Agol, V. I., and Hellen, C. U. (2000) A cell cycle-dependent protein serves as a template-specific translation initiation factor. *Genes Dev.* **14**, 2028–2045 Medline

16. 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**, 9890–9895 CrossRef Medline

17. Xia, X., Cheng, A., Lessor, T., Zhang, Y., and Hamburger, A. W. (2001) Ebp1, an ErbB-3 binding protein, interacts with Rb and affects Rb transcriptional regulation. *J. Cell. Physiol.* **187**, 209–217 CrossRef Medline

18. Zhang, Y., Akinmade, D., and Hamburger, A. W. (2005) The ErbB3 binding protein Ebp1 interacts with Sin3A to repress E2F1 and AR-mediated transcription. *Nucleic Acids Res.* **33**, 6024–6033 CrossRef Medline

19. Ghosh, A., Awasthi, S., and Hamburger, A. W. (2013) ErbB3-binding protein EBP1 decreases ErbB2 levels via a transcriptional mechanism. *Oncol. Rep.* **29**, 1161–1166 CrossRef Medline

20. Zhang, Y., and Hamburger, A. W. (2004) Heregulin regulates the ability of the ErbB3-binding protein Ebp1 to bind E2F promoter elements and repress E2F-mediated transcription. *J. Biol. Chem.* **279**, 26126–26133 CrossRef Medline

21. Zhang, Y., Woodford, N., Xia, X., and Hamburger, A. W. (2003) Repression of E2F1-mediated transcription by the ErbB3 binding protein Ebp1 involves histone deacetylases. *Nucleic Acids Res.* **31**, 2168–2177 CrossRef Medline

22. Kim, C. K., Nguyen, T. L., Joo, K. M., Nam, D. H., Park, J., Lee, K. H., Cho, S. W., and Ahn, J. Y. (2010) Negative regulation of p53 by the long isoform

ASBMB

of ErbB3 binding protein Ebp1 in brain tumors. *Cancer Res.* **70,** 9730–9741 CrossRef Medline

23. Kim, C. K., Lee, S. B., Nguyen, T. L., Lee, K. H., Um, S. H., Kim, J., and Ahn, J. Y. (2012) Long isoform of ErbB3 binding protein, p48, mediates protein kinase B/Akt-dependent HDM2 stabilization and nuclear localization. *Exp. Cell Res.* **318,** 136–143 CrossRef Medline

24. Squatrito, M., Mancino, M., Sala, L., and Draetta, G. F. (2006) Ebp1 is a dsRNA-binding protein associated with ribosomes that modulates eIF2α phosphorylation. *Biochem. Biophys. Res. Commun.* **344,** 859–868 CrossRef Medline

25. Okada, M., Jang, S. W., and Ye, K. (2007) Ebp1 association with nucleophosmin/B23 is essential for regulating cell proliferation and suppressing apoptosis. *J. Biol. Chem.* **282,** 36744–36754 CrossRef Medline

26. Fujiwara, T., Suzuki, S., Kanno, M., Sugiyama, H., Takahashi, H., and Tanaka, J. (2006) Mapping a nucleolar targeting sequence of an RNA binding nucleolar protein, Nop25. *Exp. Cell Res.* **312,** 1703–1712 CrossRef Medline

27. Martin, R. M., Ter-Avetisyan, G., Herce, H. D., Ludwig, A. K., Lättig-Tünnemann, G., and Cardoso, M. C. (2015) Principles of protein targeting to the nucleolus. *Nucleus* **6,** 314–325 CrossRef Medline

28. Gromeier, M., Alexander, L., and Wimmer, E. (1996) Internal ribosomal entry site substitution eliminates neurovirulence in intergeneric poliovirus recombinants. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 2370–2375 CrossRef Medline

29. Hellen, C. U., and Sarnow, P. (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev.* **15,** 1593–1612 CrossRef Medline

30. Komar, A. A., and Hatzoglou, M. (2005) Internal ribosome entry sites in cellular mRNAs: mystery of their existence. *J. Biol. Chem.* **280,** 23425–23428 CrossRef Medline

31. Schmeing, T. M., and Ramakrishnan, V. (2009) What recent ribosome structures have revealed about the mechanism of translation. *Nature* **461,** 1234–1242 CrossRef Medline

32. Spriggs, K. A., Bushell, M., Mitchell, S. A., and Willis, A. E. (2005) Internal ribosome entry segment-mediated translation during apoptosis: the role of IRES-*trans*-acting factors. *Cell Death Differ.* **12,** 585–591 CrossRef Medline

33. Lewis, S. M., and Holcik, M. (2008) For IRES *trans*-acting factors, it is all about location. *Oncogene* **27,** 1033–1035 CrossRef Medline

34. Stoneley, M., and Willis, A. E. (2004) Cellular internal ribosome entry segments: structures, trans-acting factors and regulation of gene expression. *Oncogene* **23,** 3200–3207 CrossRef Medline

35. Komar, A. A., and Hatzoglou, M. (2011) Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. *Cell Cycle* **10,** 229–240 CrossRef Medline

36. Pacheco, A., and Martinez-Salas, E. (2010) Insights into the biology of IRES elements through riboproteomic approaches. *J. Biomed. Biotechnol.* **2010,** 458927 CrossRef Medline

37. Grotwinkel, J. T., Wild, K., Segnitz, B., and Sinning, I. (2014) SRP RNA remodeling by SRP68 explains its role in protein translocation. *Science* **344,** 101–104 CrossRef Medline

38. Battiste, J. L., Mao, H., Rao, N. S., Tan, R., Muhandiram, D. R., Kay, L. E., Frankel, A. D., and Williamson, J. R. (1996) α helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science* **273,** 1547–1551 CrossRef Medline

39. Heery, D. M., Kalkhoven, E., Hoare, S., and Parker, M. G. (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387,** 733–736 CrossRef Medline

40. Plevin, M. J., Mills, M. M., and Ikura, M. (2005) The LxxLL motif: a multifunctional binding sequence in transcriptional regulation. *Trends Biochem. Sci.* **30,** 66–69 CrossRef Medline

41. Zhang, Y., Fondell, J. D., Wang, Q., Xia, X., Cheng, A., Lu, M. L., and Hamburger, A. W. (2002) Repression of androgen receptor mediated transcription by the ErbB-3 binding protein, Ebp1. *Oncogene* **21,** 5609–5618 CrossRef Medline

42. Zhang, Y., and Hamburger, A. W. (2005) Specificity and heregulin regulation of Ebp1 (ErbB3 binding protein 1) mediated repression of androgen receptor signalling. *Br. J. Cancer* **92,** 140–146 CrossRef Medline

43. Koach, J., Holien, J. K., Massudi, H., Carter, D. R., Ciampa, O. C., Herath, M., Lim, T., Seneviratne, J. A., Milazzo, G., Murray, J. E., McCarroll, J. A., Liu, B., Mayoh, C., Keenan, B., Stevenson, B. W., *et al.* (2019) Drugging MYCN oncogenic signaling through the MYCN-PA2G4 binding interface. *Cancer Res.* **79,** 5652–5667 CrossRef Medline

44. Lessor, T. J., and Hamburger, A. W. (2001) Regulation of the ErbB3 binding protein Ebp1 by protein kinase C. *Mol. Cell. Endocrinol.* **175,** 185–191 CrossRef Medline

45. Liu, Z., Oh, S. M., Okada, M., Liu, X., Cheng, D., Peng, J., Brat, D. J., Sun, S. Y., Zhou, W., Gu, W., and Ye, K. (2009) Human BRE1 is an E3 ubiquitin ligase for Ebp1 tumor suppressor. *Mol. Biol. Cell* **20,** 757–768 CrossRef Medline

46. Nishi, H., Hashimoto, K., and Panchenko, A. R. (2011) Phosphorylation in protein-protein binding: effect on stability and function. *Structure* **19,** 1807–1815 CrossRef Medline

47. Wang, Y., Zhang, P., Wang, Y., Zhan, P., Liu, C., Mao, J. H., and Wei, G. (2017) Distinct interactions of EBP1 isoforms with FBXW7 elicits different functions in cancer. *Cancer Res.* **77,** 1983–1996 CrossRef Medline

48. Ko, H. R., Kim, C. K., and Ahn, J. Y. (2015) Phosphorylation of the N-terminal domain of p48 Ebp1 by CDK2 is required for tumorigenic function of p48. *Mol. Carcinog.* **54,** 1283–1291 CrossRef Medline

49. Ko, H. R., Kim, C. K., Lee, S. B., Song, J., Lee, K. H., Kim, K. K., Park, K. W., Cho, S. W., and Ahn, J. Y. (2014) P42 Ebp1 regulates the proteasomal degradation of the p85 regulatory subunit of PI3K by recruiting a chaperone-E3 ligase complex HSP70/CHIP. *Cell Death Dis.* **5,** e1131 CrossRef Medline

50. Nash, P., Tang, X., Orlicky, S., Chen, Q., Gertler, F. B., Mendenhall, M. D., Sicheri, F., Pawson, T., and Tyers, M. (2001) Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication. *Nature* **414,** 514–521 CrossRef Medline

51. Luan, Y., and Wang, P. (2014) FBW7-mediated ubiquitination and degradation of KLF5. *World J. Biol. Chem.* **5,** 216–223 CrossRef Medline

52. Ko, H. R., Chang, Y. S., Park, W. S., and Ahn, J. Y. (2016) Opposing roles of the two isoforms of ErbB3 binding protein 1 in human cancer cells. *Int. J. Cancer* **139,** 1202–1208 CrossRef Medline

53. Shen, W., Tremblay, M. S., Deshmukh, V. A., Wang, W., Filippi, C. M., Harb, G., Zhang, Y. Q., Kamireddy, A., Baaten, J. E., Jin, Q., Wu, T., Swoboda, J. G., Cho, C. Y., Li, J., Laffitte, B. A., *et al.* (2013) Small-molecule inducer of β cell proliferation identified by high-throughput screening. *J. Am. Chem. Soc.* **135,** 1669–1672 CrossRef Medline

54. Skowyra, D., Craig, K. L., Tyers, M., Elledge, S. J., and Harper, J. W. (1997) F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. *Cell* **91,** 209–219 CrossRef Medline

55. Davis, R. J., Welcker, M., and Clurman, B. E. (2014) Tumor suppression by the Fbw7 ubiquitin ligase: mechanisms and opportunities. *Cancer Cell* **26,** 455–464 CrossRef Medline

56. Welcker, M., and Clurman, B. E. (2008) FBW7 ubiquitin ligase: a tumour suppressor at the crossroads of cell division, growth and differentiation. *Nat. Rev. Cancer* **8,** 83–93 CrossRef Medline

57. Lessor, T. J., Yoo, J. Y., Xia, X., Woodford, N., and Hamburger, A. W. (2000) Ectopic expression of the ErbB-3 binding protein ebp1 inhibits growth and induces differentiation of human breast cancer cell lines. *J. Cell. Physiol.* **183,** 321–329 CrossRef Medline

58. Yu, Y., Chen, W., Zhang, Y., Hamburger, A. W., Pan, H., and Zhang, Z. (2007) Suppression of salivary adenoid cystic carcinoma growth and metastasis by ErbB3 binding protein Ebp1 gene transfer. *Int. J. Cancer* **120,** 1909–1913 CrossRef Medline

59. Fruman, D. A., and Rommel, C. (2014) PI3K and cancer: lessons, challenges and opportunities. *Nat. Rev. Drug Discov.* **13,** 140–156 CrossRef Medline

60. Lu, Y., Zhou, H., Chen, W., Zhang, Y., and Hamburger, A. W. (2011) The ErbB3 binding protein EBP1 regulates ErbB2 protein levels and tamoxifen sensitivity in breast cancer cells. *Breast Cancer Res. Treat.* **126,** 27–36 CrossRef Medline

61. Patel, V. L., Cytryn, K. N., Shortliffe, E. H., and Safran, C. (2000) The collaborative health care team: the role of individual and group expertise. *Teach Learn. Med.* **12,** 117–132 CrossRef Medline

62. Nevins, J. R. (2001) The Rb/E2F pathway and cancer. *Hum. Mol. Genet.* **10,** 699–703 CrossRef Medline

ASBMB

J. Biol. Chem. (2020) 295(47) 16100–16112 **16111**

63. Ou, K., Kesuma, D., Ganesan, K., Yu, K., Soon, S. Y., Lee, S. Y., Goh, X. P., Hooi, M., Chen, W., Jikuya, H., Ichikawa, T., Kuyama, H., Matsuo, E., Nishimura, O., and Tan, P. (2006) Quantitative profiling of drug-associated proteomic alterations by combined 2-nitrobenzenesulfenyl chloride (NBS) isotope labeling and 2DE/MS identification. *J. Proteome Res.* **5,** 2194–2206 CrossRef Medline

64. Figeac, N., Serralbo, O., Marcelle, C., and Zammit, P. S. (2014) ErbB3 binding protein-1 (Ebp1) controls proliferation and myogenic differentiation of muscle stem cells. *Dev. Biol.* **386,** 135–151 CrossRef Medline

65. Zhang, Y., Lu, Y., Zhou, H., Lee, M., Liu, Z., Hassel, B. A., and Hamburger, A. W. (2008) Alterations in cell growth and signaling in ErbB3 binding protein-1 (Ebp1) deficient mice. *BMC Cell Biol.* **9,** 69 CrossRef Medline

66. Braicu, C., Buse, M., Busuioc, C., Drula, R., Gulei, D., Raduly, L., Rusu, A., Irimie, A., Atanasov, A. G., Slaby, O., Ionescu, C., and Berindan-Neagoe, I. (2019) A comprehensive review on MAPK: a promising therapeutic target in cancer. *Cancers (Basel)* **11,** 1618 CrossRef Medline

67. Mei, Y., Zhang, P., Zuo, H., Clark, D., Xia, R., Li, J., Liu, Z., and Mao, L. (2014) Ebp1 activates podoplanin expression and contributes to oral tumorigenesis. *Oncogene* **33,** 3839–3850 CrossRef Medline

68. Quintanilla, M., Montero-Montero, L., Renart, J., and Martín-Villar, E. (2019) Podoplanin in inflammation and cancer. *Int. J. Mol. Sci.* **20,** 707 CrossRef Medline

69. Zhou, X., Chen, W., Zhang, Y., Sun, J., Wang, Q., and Yu, Y. (2010) Potential therapeutic strategy for oral squamous cell carcinoma by ErbB3-binding protein 1 gene transfer. *J. Cancer Res. Clin. Oncol.* **136,** 891–896 CrossRef Medline

70. Nguyen Le, X. T., Zhu, L., Lee, Y., Ta, L., and Mitchell, B. S. (2016) Expression and role of the ErbB3-binding protein 1 in acute myelogenous leukemic cells. *Clin. Cancer Res.* **22,** 3320–3327 CrossRef Medline

71. Liu, L., Li, X. D., Chen, H. Y., Cui, J. S., and Xu, D. Y. (2015) Significance of Ebp1 and p53 protein expression in cervical cancer. *Genet. Mol. Res.* **14,** 11860–11866 CrossRef Medline

72. Gannon, P. O., Koumakpayi, I. H., Le Page, C., Karakiewicz, P. I., Mes-Masson, A. M., and Saad, F. (2008) Ebp1 expression in benign and malignant prostate. *Cancer Cell Int.* **8,** 18 CrossRef Medline

73. Nero, T. L., Morton, C. J., Holien, J. K., Wielens, J., and Parker, M. W. (2014) Oncogenic protein interfaces: small molecules, big challenges. *Nat. Rev. Cancer* **14,** 248–262 CrossRef Medline

ASBMB