

The dyslexia-associated gene *KIAA0319* encodes highly N- and O-glycosylated plasma membrane and secreted isoforms

Antonio Velayos-Baeza, Claudio Toma[†], Silvia Paracchini and Anthony P. Monaco*

The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK

Received August 30, 2007; Revised and Accepted December 4, 2007

The *KIAA0319* gene has been recently associated with developmental dyslexia and shown to be involved in neuronal migration. The deduced *KIAA0319* protein contains several polycystic kidney disease (PKD) domains which may mediate the interaction between neurons and glial fibres during neuronal migration. We have previously reported the presence of several alternative splicing variants, some of which are predicted to affect the deduced protein. In this study, we over-expressed constructs containing the main form (A) and two alternative variants (B and C) of *KIAA0319*. We show that the full-length *KIAA0319* (A) is a type I plasma membrane protein, a topology consistent with its proposed function in neuronal migration. The oligomeric status of *KIAA0319* is mainly dimeric, and this condition depends on the cysteine-rich regions of the protein, especially the transmembrane (TM) domain and surrounding sequence. *KIAA0319* is highly glycosylated in different mammalian cell lines. The central region including the PKD domains is N-glycosylated. Furthermore, a short fragment N-terminal to the PKD domains contains mucin-type O-glycosylation. The two alternative isoforms are soluble proteins lacking the TM domain and, interestingly, only isoform B is secreted. *KIAA0319*-deletion proteins lacking the TM domain were also secreted. These results suggest that *KIAA0319* could be involved not only in cell–cell interactions, but also in signalling.

INTRODUCTION

Developmental dyslexia (DD), a common cognitive disorder affecting around 5% of school children, is a specific and significant impairment in reading ability despite normal intelligence, adequate environment and educational opportunities (1). DD has a strong genetic component, with nine dyslexia susceptibility (*DYX*) loci currently reported (2). Several genes have been proposed as candidates: *DYX1C1* for *DYX1* locus (3), *KIAA0319* and *DCDC2* for *DYX2* locus (4–8), *ROBO1* for *DYX5* locus (9) (10–12 for recent reviews) and *MRPL19* and *C2ORF3* for *DYX3* locus (13). Abnormalities in the development of the brain are increasingly reported in DD and, interestingly, most of the candidate genes reported so far, including *KIAA0319*, participate in brain development (14).

Positive linkage has been consistently reported in a number of studies on the *DYX2* locus, where the *KIAA0319* gene appears as a particularly strong candidate due to the association

with reading abilities found in several independent samples (12). We have recently reported that the expression of the *KIAA0319* gene is reduced (15) on chromosomes carrying a specific haplotype associated with DD (4). Moreover, *KIAA0319* is required for neuronal migration in the developing cerebral neocortex (15). The protein encoded by *KIAA0319* is predicted to have a signal peptide (SP), a motif at N-terminus with seven cysteines (MANSC), five polycystic kidney disease (PKD) domains and a single transmembrane (TM) domain (16). The MANSC domain is present in animal membrane and extracellular proteins, and has been postulated to bind other proteins (17). The PKD domains have been implicated in cell–cell adhesion processes (18). These data led us to suggest a possible role of the *KIAA0319* PKD domains in the adhesion between neurons and glial fibres during neuronal migration (15).

We have undertaken a functional analysis of *KIAA0319* to understand the molecular mechanisms this protein may exploit

*To whom correspondence should be addressed: Tel: +44 1865287502; Fax: +44 1865287650; Email: anthony.monaco@well.ox.ac.uk.

[†]Present address: Department of Biology, University of Bologna, Bologna, Italy.

during brain development and the effects that alteration of its expression levels could have in this process. In this context, we have shown the presence of several alternative splicing variants in the *KIAA0319* gene, both in humans and rodents (16). The predominant form is the full-length transcript (or variant A), including 21 exons. Forms lacking exons 19 or 19+20 (variants B and C, respectively), or variations in the 5'-UTR sequence, were detected in both fetal and embryonic human brain samples. The predicted KIAA0319 isoforms B and C would lack the TM domain, encoded by exon 19, and therefore would also change their sub-cellular localization. These non-membranal isoforms could potentially have different functions than the full-length protein.

Here we present the first experimental data on the KIAA0319 protein. We show that the main isoform of KIAA0319 is a dimeric protein with plasma membrane (PM) localization. This localization changes to secretion or endoplasmic reticulum (ER) retention for the TM domain-lacking isoforms B and C, respectively. We have also studied the effect that the different domains have on the dimerization and in the sub-cellular localization of this protein using a deletion approach. We present evidence that KIAA0319 is both N- and O-glycosylated and map the region involved in O-glycosylation. This basic characterization provides the tools needed to advance in the functional analysis of the KIAA0319 protein which, in turn, would allow us to gain new insights on the biological mechanisms underlying the reading process and, more in general, of cognition.

RESULTS

Characterization of KIAA0319 protein over-expressed in human cell lines

Variants A, B and C of the KIAA0319 protein (Fig. 1 and Table 1) were over-expressed in HEK293T and/or MRC5 cell lines. The expected size of the KIAA0319 protein, without the SP (first 20 residues), is 1052 amino acid and 115.8 kDa for the full-length (KA), 958 amino acid/104.9 kDa for variant B (KB) and 991 amino acid/108.4 kDa for variant C (KC); tagging of the protein adds 28 amino acid/3.33 kDa with myc+His, and 261 amino acid/29.38 kDa with EGFP. The over-expressed proteins were used to check the activity by immunofluorescence (IF) and western blotting (WB) of the specific antibodies against the K-PKD deletion protein (rat polyclonal K1r and K2r) and against peptides [A+B] (rabbit polyclonal R1 and R2) of the KIAA0319 protein (see *Materials and Methods* and Fig. 1). These antisera and the affinity-purified anti-peptide A showed the same pattern, which was also identical to that detected with the antibody against the tag: Fig. 2A–C shows the detection by IF of KC protein with anti-myc antibody and antiserum R2 as an example. Affinity-purified anti-peptide B did not show immunoreactivity (data not shown). The C-terminal tagging did not modify the expression or the distribution of the protein: the same pattern was detected in tagged and in non-tagged proteins (Supplementary Materials, Fig. S1).

KIAA0319 is a type I protein that localizes in the PM. We checked the sub-cellular localization of the over-expressed

protein in transiently transfected cells (Fig. 2). The KB form, lacking the TM domain, was detected in the ER (Fig. 2D–F); the same result was obtained for the KC form, also lacking the TM domain (data not shown). The full-length (KA) protein was detected in the PM, with some co-localization with ER and *trans*-Golgi network (TGN) markers (Fig. 2G–L). To confirm the PM localization of KA, as well as its topology, IF was performed under conditions in which the antibodies cannot enter the cell (without permeabilization). Therefore, only antigens exposed to the extracellular part of the cell can be recognized. Transfected cells were fixed and, in a detergent-free environment, subjected to detection with both specific antiserum R2 and anti-myc antibody. Under these conditions, KA was detected with R2 antiserum but not with anti-myc (Fig. 2M–O). Under the usual conditions (permeabilization), both antibodies detect the protein (Fig. 2P–R). These results indicate that the N-terminus of the protein is extracellular and the C-terminus is the cytoplasmic domain of the protein. No signal was obtained for KB or KC under non-permeabilization conditions (data not shown).

KIAA0319 is a glycosylated dimer. Protein lysates from transiently transfected cells were separated by SDS–PAGE and subjected to WB with anti-tag and/or anti-KIAA0319 antibodies. A single band was detected for KB and KC under reducing conditions, whereas four bands were apparent for KA, being the two large bands approximately twice the size of the small ones (Fig. 3A, left panel). Analysis under non-reducing conditions (Fig. 3A, right) showed an increased signal of the two larger bands of KA while signal for the two smaller bands was diminished. New bands were detected for KB and KC in these conditions, the strongest one with an apparent size of >270 kDa. These results suggest that KIAA0319 forms mainly dimers. To further check this possibility we performed co-immunoprecipitation (co-IP) experiments with protein lysates from cells expressing two differently tagged KA proteins. Figure 3B shows the results obtained with two full-length KIAA0319 proteins tagged with myc+His (Am) and with EGFP (AG): AG is detected after IP with anti-His antibody (lane 4), and Am is detected after IP with anti-GFP antibody (lane 9), in cells co-expressing both proteins, but not in the controls expressing only one of them (lanes 2 and 11). These results indicate that these differently tagged proteins interact, and support the formation of dimers by KIAA0319.

We used the specific antisera R2 and K1r, raised against the human KIAA0319 protein, to try the detection of the endogenous protein. As protein sources we used two human brain lysates: total protein lysate from adult cerebral cortex (AL) and the membrane fraction of the protein lysate from fetal frontal cortex (FM). Figure 3C shows a band of ~150–160 kDa detected with both antisera. A band of ~300 kDa is also clearly detected with K1r antiserum. These results are consistent with those obtained with the over-expressed protein (Fig. 3A).

From the WB results, the apparent sizes of the protein bands were larger than expected for all three isoforms. KIAA0319 has several N-glycosylation consensus sites (Fig. 1A) so it was probable that this increase in size was due to glycosylation. The same protein lysates were compared after treatment with or without peptide: N-glycosylidase F (PNGase F), an

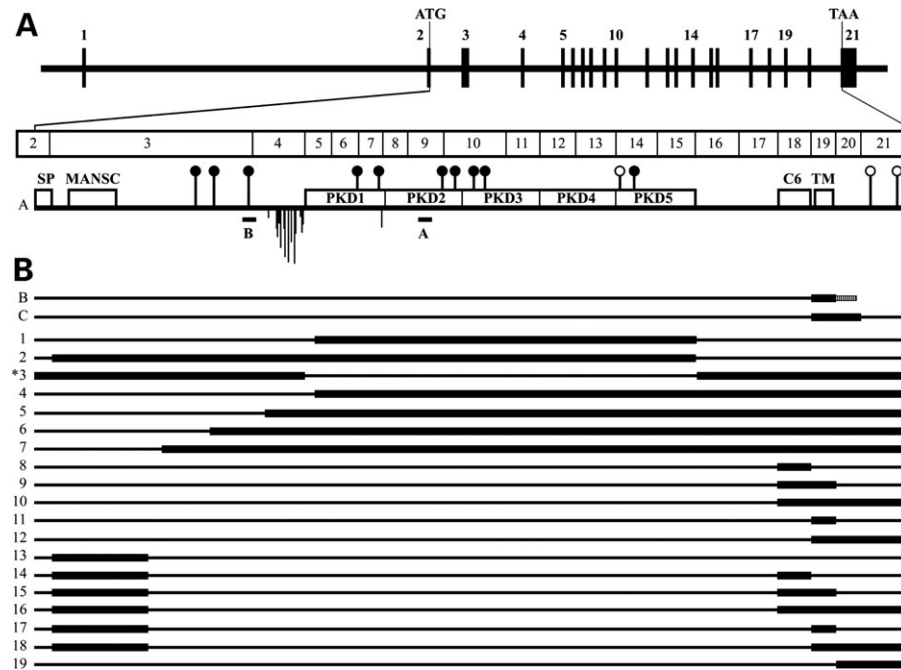


Figure 1. The human KIAA0319 protein isoforms and deletion-proteins. (A) Scheme of the *KIAA0319* gene (genomic DNA, top; cDNA, middle) and its encoded full-length protein (isoform A, bottom). Regions or domains detected by computational analysis (16) are shown: signal peptide (SP), MANSC domain, PKD (1–5) domains, C6 domain (region of the protein that includes 6 cysteine residues), and TM domain. Circles represent N-Glycosylation consensus sites where glycosylation is either predicted (black) or not (white); vertical lines below the main bar represent predicted O-glycosylation according to NetOGlyc software (<http://www.cbs.dtu.dk/services/NetOGlyc>) (only values above the 0.5 threshold are represented, and the highest value is 0.72). The two peptides (A and B) used to raise polyclonal antibodies are also indicated. (B) Graphical representation of alternative isoforms B and C, and the deletion proteins (1–19) used in this work, all expressed in vector pcDNA4-TO-mycHis, except for '3', in pHlsec (*). Details for each protein are shown in Table 1. The deleted regions of the protein are represented by large filled bars. Semi-filled area at the C-terminus of isoform B represents a sequence different from isoform A (Table 1).

enzyme which removes nearly all N-glycosylation. A shift in size was observed in all cases (Fig. 3D), and indicates that KIAA0319 is N-glycosylated. Thus, the four bands detected for the full-length protein probably correspond to ER-glycosylated monomer (~150–170 kDa), fully glycosylated monomer (~180–210 kDa), ER-glycosylated dimer (~280–320 kDa) and fully glycosylated dimer (~350–400 kDa). To check this hypothesis, PM proteins were isolated following cell surface biotinylation. Figure 3E shows that only the largest of the two bands (monomeric forms) present in the lysate (L) and the flow through (FT) is detected after purification of biotinylated proteins (E). This result agrees with the previous hypothesis as only the fully (Golgi)-glycosylated form would reach the PM.

Variant B, but not variant C, is secreted. The absence of exon 19 in both alternative KIAA0319 variants (B and C) means that these two proteins lack the TM domain and, therefore, they would enter the secretory pathway as soluble (non-membranal) proteins. In order to determine whether KB and KC are secreted to the extracellular medium or retained in the ER, culture medium of cells expressing His-tagged proteins (KA, KB or KC) was subjected to His-purification. The recovered proteins were analysed by WB with anti-myc antibody (Fig. 3F). As expected, the membrane-attached protein KA was not detected. The soluble protein KC was also undetectable. Only KB was observed, with an apparent size considerably larger than that found for the same protein

purified from cell lysates (used as a control). This can be explained as a result of the extra glycosylation added in the post-ER compartments. The same result was obtained on four repeat experiments and is not associated with low expression of KC isoform as all the proteins were detected in the cell lysates as shown in Fig. 3A (see also Supplementary Materials, Fig. S2).

Domain characterization by deletion analysis

A series of deletion constructs (Fig. 1B, Table 1) were created to analyse the effect of removing specific domains or regions within the KIAA0319 protein on its localization, structure or post-translational modifications. The normal reading frame was always preserved. For example, when deleting the protein region encoded by exon 19 (Kd19, construct #11), the exon sequence was removed and the reading frame conserved by introduction of C in the cDNA, leading to the substitution of the corresponding protein region by an alanine residue. All the deletion constructs were myc+His tagged at the C-terminus unless specified otherwise. From hereon, these deletion proteins/constructs will be referred to as shown in Table 1, with both protein name and construct number (Kd19 [11]); this number also identifies the constructs in Fig. 1B.

The TM domain is the only region needed for PM localization. The localization of the KIAA0319 isoforms described above depends on the protein entering the secretory

Table 1. KIAA0319 constructs and proteins

| # ^a | Plasmid | Protein ^b | Del amino acid ^c | New ^d | No. of amino acid | Tagging amino acid ^e | Size (kDa) ^f |
|----------------|-----------------|----------------------|-----------------------------|------------------|-------------------|---------------------------------|-------------------------|
| A | pcD4KAG | KAGFP | — | — | 1333 | 1073–1333 | 145.20 |
| A | pcD4KAm | KA | — | — | 1100 | 1073–1100 | 119.15 |
| B | pcD4KBm | KB | 953–1072 | 953–978 | 1006 | 979–1006 | 108.22 |
| C | pcD4KCm | KC | 953–1013 | — | 1039 | 1012–1039 | 111.76 |
| 1 | pcD4Kmd5-15 | Kd5-15 | 345–812 | T | 633 | 606–633 | 68.58 |
| 2 | pcD4Kmd3-15 | Kd3-15 | 23–810 | — | 312 | 285–312 | 33.74 |
| 3 | pHLK-PKD | K-PKD | 1–331 815–1072 | newSP+ETG — | 520 | 515–520 | 53.41 |
| 4 | pcD4Kmd5-21 | Kd5-21 | 346–1072 | — | 372 | 346–372 | 38.54 |
| 5 | pcD4Kmd4b-21 | Kd4b-21 | 283–1072 | — | 310 | 283–310 | 32.29 |
| 6 | pcD4Kmd3c-21 | Kd3c-21 | 216–1072 | — | 243 | 216–243 | 25.15 |
| 7 | pcD4Kmd3b-21 | Kd3b-21 | 156–1072 | — | 183 | 156–183 | 18.68 |
| 8 | pcD4Kmd18 | Kd18 | 912–952 | — | 1059 | 1032–1059 | 114.45 |
| 9 | pcD4Kmd18-19 | Kd18-19 | 912–983 | A | 1029 | 1002–1029 | 110.95 |
| 10 | pcD4Kmd18-21 | Kd18-21 | 913–1072 | — | 938 | 913–938 | 100.22 |
| 11 | pcD4Kmd19 | Kd19 | 953–983 | A | 1070 | 1043–1070 | 115.66 |
| 12 | pcD4Kmd19-21 | Kd19-21 | 953–1072 | — | 978 | 953–978 | 104.87 |
| 13 | pcD4Kmd3a | Kd3a | 23–134 | — | 988 | 961–988 | 106.58 |
| 14 | pcD4Kmd3ad18 | Kd3ad18 | 23–134 912–952 | — — | 947 | 920–947 | 101.88 |
| 15 | pcD4Kmd3ad18-19 | Kd3ad18-19 | 23–134 912–983 | — A | 917 | 890–917 | 98.38 |
| 16 | pcD4Kmd3ad18-21 | Kd3ad18-21 | 23–134 913–1072 | — — | 826 | 801–826 | 87.64 |
| 17 | pcD4Kmd3ad19 | Kd3ad19 | 23–134 953–983 | — A | 958 | 931–958 | 103.08 |
| 18 | pcD4Kmd3ad19-21 | Kd3ad19-21 | 23–134 953–1072 | — — | 866 | 841–866 | 92.29 |
| 19 | pcD4Kmd20-21 | Kd20-21 | 984–1072 | — | 1009 | 984–1009 | 108.43 |

^aConstruct as shown in Fig. 1.
^bProtein name used in the text.
^cResidues deleted from the full-length protein.
^dNew residues introduced in place of the deleted protein fragment, excluding the tagging sequence. For KIAA0319 isoform B (KB), the fragment 953–1072 from the full-length protein is substituted by 26 alternative residues (DKKGLKSGKKQSTPSWITWMNRKEWN), encoded by exon 20, due to frameshift. K-PKD retains the sequence ETG after cleavage of the SP derived from pHLsec (MGILPSPGMPALLSLVSLLSVLLMGCVA↓ETG).
^eResidues derived from the vector sequence including the C-terminal tag. All of them have the myc (bold)+His tag ([L/I/F][E]SRGPFE**QKLISEEDL**NMHTG HHHHHH) except KAGFP (22 amino acid +EGFP) and K-PKD (6×His tag).
^fSize of the deduced protein, without the SP (20 amino acid/1.94 kDa, except for K-PKD: 28 amino acid/2.78 kDa).

pathway: removal of the SP leads to cytoplasmic localization (data not shown). Once a protein enters the secretory pathway, its final location is often determined by signals present in its sequence. In order to check if the progression of KIAA0319 along this pathway depends on specific regions or domains, HEK293T cells were transfected with different deletions plasmids and processed for IF (Fig. 4). When the five PKD domains in the central part of the N-terminus of the KIAA0319 protein were removed by deletion of exons 5–15 in the cDNA (Kd5–15 [1]), no difference in the sub-cellular localization was detected when compared with the full-length protein (Fig. 4A and B). Deletion of the 5' end of exon 3 (Kd3a [13]) or exon 18 (Kd18 [8]) was used to remove the MANSC or C6 domains, respectively. These Cys-rich domain-deletion proteins could be found in the PM (Fig. 4D and F) although the number of cells expressing proteins with a clear PM pattern was lower than for KA or Kd5–15. Deletion of both MANSC and C6 domains (Kd3ad18 [14], Fig. 4E), or MANSC to PKD domains (Kd3–15 [2], Fig. 4C), produced a similar result. Finally, when the cytoplasmic domain was removed by deletion of exons 20 and 21 (Kd20–21 [19]) a normal PM pattern was also obtained (Fig. 4G). These

results suggest none of these regions is essential for the protein to advance along the secretory pathway, and that only the TM domain is needed for PM localization of the KIAA0319 protein, providing retention into the membrane.

Proteins without the TM domain are secreted. A number of constructs in which exon 19 has been removed were obtained (Fig. 1B). All these deletion proteins localized to the ER (Fig. 4H–L) as it happened for KB and KC (Fig. 2A–F). In some cases the ER showed an aberrant morphology triggered by accumulation of over-expressed, and probably missfolded, proteins.

Cell lysates and proteins purified from the culture media (His-purification) were analysed by WB in both reducing and non-reducing conditions (Fig. 5A and Supplementary Materials, Fig. S2). All these proteins were detected approximately at the same level in cell lysates, mainly in their monomeric form in reducing conditions (Fig. 5A, top left panel) while a ladder-like pattern was obtained in non-reducing conditions (Fig. 5A, top right panel), the number of bands depending on the number of Cys residues present in the protein. All proteins without the TM domain were detected in the culture

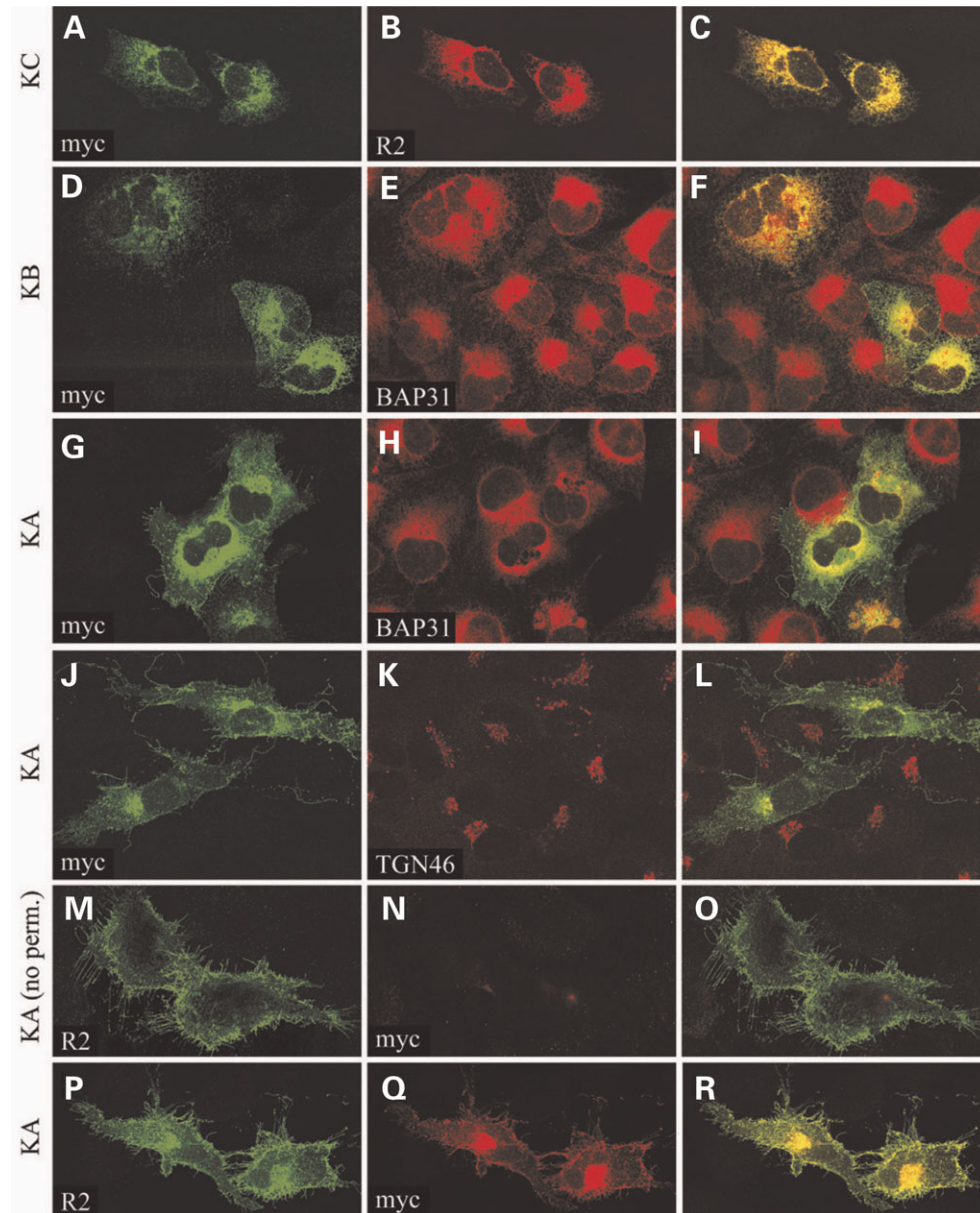


Figure 2. Sub-cellular localization of KIAA0319. C-terminal myc + His-tagged KIAA0319 isoforms C (KC, A–C), B (KB, D–F) or A (KA, G–R) were over-expressed in MRC5 cells and detected with monoclonal anti-myc (A, D, G, J, N and Q) and/or polyclonal anti-KIAA0319 R2 antiserum (B, M and P). Anti-BAP31 (E and H) and anti-TGN46 (K) were used as ER and TGN sub-cellular markers, respectively. Merge pictures are shown on the right column. (A–C) Specific detection of KIAA0319 protein by R2 antiserum. (D–F) ER localization of KB. (G–I) KA reaches the plasma membrane. (J–L) Partial co-localization of KA with TGN marker. (M–R) Antiserum R2, but not anti-myc antibody, detects KIAA0319 in no-permeabilization conditions; both antibodies give overlapping signals under permeabilization conditions (P–R).

medium after His-purification, but the secretion levels were different (Fig. 5A, lower panels). Secretion of these proteins to the culture medium was also detected directly, without His-purification, as shown in Fig. 5B. The lower number of Cys residues in the protein usually corresponded to the higher level of secretion.

The cysteine-containing domains are involved in dimerization. In order to check if the formation of dimers is dependent on Cys residues we compared the bands detected in different conditions between deletion-proteins lacking one or several

Cys-containing regions (MANSC domain, eight residues; C6 domain, six residues; TM domain, two residues; or C-terminus, three residues) (Fig. 5; Supplementary Materials, Figs S1B and S2). In the presence of the TM domain, the deletion of MANSC, C6 or both (top panel, Fig. 5A, lanes 1 and 2; Supplementary Materials, Fig. S2, lanes 4, 9 and 10) do not seem to affect the formation of dimers (compare with Fig. 3A and lane 1 in Supplementary Materials, Fig. S2) suggesting that the Cys residues in the TM domain and/or the cytoplasmic domain might be those involved in dimerization.

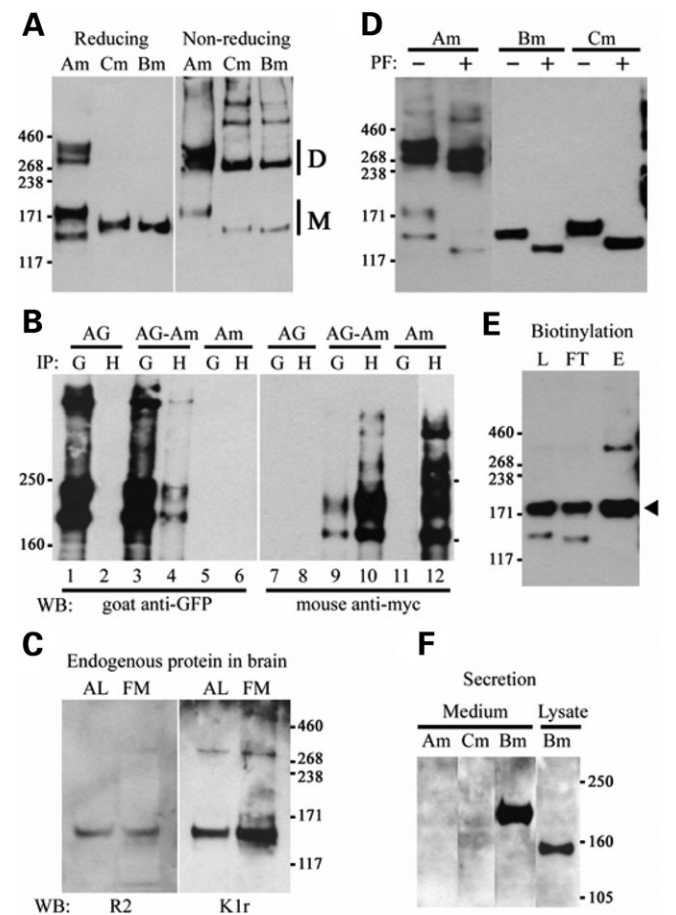


Figure 3. Analysis of KIAA0319 protein by WB. Untagged full-length KIAA0319 protein (E), C-terminal myc + His-tagged isoforms A (Am), B (Bm) and C (Cm), and C-terminal EGFP-tagged isoform A (AG) were over-expressed in HEK293T cells and processed for WB analysis with specific anti-serum R2, unless otherwise specified. (A) Comparison between reducing and non-reducing conditions; M, monomeric forms; D, dimeric forms. (B) Co-IP of differently-tagged KIAA0319 protein in co-transfected cells (AG-Am), but not in control single-transfected cells (AG, Am), with rabbit anti-GFP (G) and rabbit anti-His (H) antibodies indicates homodimerization of KIAA0319. (C) Detection of the KIAA0319 endogenous protein in human brain with specific antisera R2 and K1r, in reducing conditions, after Adult cerebral cortex tissue lysate (AL) and Fetal frontal lobe membrane protein lysate (FM). (D) Treatment with PNGase F (PF) shows N-glycosylation of KIAA0319. (E) Purification of plasma membrane, fully-glycosylated, KIAA0319 protein (arrowhead) after biotinylation; L, lysate; FT, flow-through; E, elute. (F) Purification of His-tagged proteins from conditioned medium shows that only isoform B is secreted; the size difference with the control cell lysate-purified protein indicates that the secreted protein has been fully glycosylated.

As described above, deletion of exon 19 removes the TM domain, and also four Cys residues, two of them included in the TM domain. Analysis of these soluble/secreted proteins under non-reducing conditions (Fig. 5A, right, and B) shows that Cys residues are needed for dimerization/multimerization: Kd3ad18-21 [16] (Fig. 5A, lane 5; B, lane 10) or K-PKD [3] (Fig. 5B, lane 12) have no Cys residues and only the monomeric bands are detected, independently of gel running conditions, while dimerization is detected in the single Cys-containing protein Kd3ad18-19 [15] (Fig. 5A, lane 4; B, lane 9). The presence of a higher number of Cys residues in

these proteins leads to multimerization in the cell lysates but, interestingly, not in the secreted proteins. Thus, the presence of the C6 domain alone (Kd3ad19 [17], Kd3ad19-21 [18]), the MANSC domain alone (Kd18-19 [9], Kd18-21 [10]) or both (Kd19 [11], KB, KC, Kd19-21 [21]), not only triggers dimerization but multimerization of the soluble proteins from the cell lysate (Fig. 5A, top right panel, lanes 3 and 6; data shown only for C6-containing proteins) while only the monomeric and dimeric forms are detected in the secreted proteins (Fig. 5A, bottom right, lanes 3 and 6; B, all lanes apart from 1, 9, 10 and 12). All the secreted proteins are detected mainly in monomeric and dimeric forms; higher bands can be detected after overexposure but at a much lower level. These results support the hypothesis that KIAA0319 exists mainly as a dimer. In all cases, independently of the level of secretion for each protein, the dimeric form is the main band detected when the C6 domain is present (Fig. 5B, lanes 2, 3, 7, 8 and 11), suggesting involvement of the C6 domain in dimerization. The same does not appear to be the case for the MANSC domain: the monomeric form is more abundant than the dimer when this domain is present but C6 is deleted (Fig. 5B, lanes 5 and 6). Identical results, obtained in a different experiment, are shown in Supplementary Materials, Fig. S2.

We checked our hypothesis of Cys residues involvement in dimerization by an alternative approach, mutagenesis of those residues instead of deletion of the domains containing them. Figure 6A shows the distribution of these residues in the KIAA0319 protein and the change introduced in each position. The only Cys residue that was not changed was C20, included in the SP. In most cases, Cys was mutated to Ala; Ser, Val and Leu were used in some cases for practical reasons (generation of specific restriction sites in the cDNA for easy identification of the different expression constructs); C1068 was changed to Tyr because this is the residue present in the mouse and rat KIAA0319 proteins. The analysis of the Cys-mutant proteins is shown in Figure 6B. In reducing conditions, the bands corresponding to dimeric forms disappear when the Cys residues in the TM-cytoplasmic domain are mutated (lanes M3, M5, M6 and M7), suggesting that these residues are responsible for the 'not-fully reduction' detected by WB. The dimeric forms can be detected in non-reducing conditions in all the mutant proteins except for M7, where all the Cys residues have been mutated. The detected signal is particularly strong when the Cys residues in the TM-cytoplasmic domain are not changed (lanes M0, M1, M2 and M4). When these residues have been mutated (M3), additional mutation of the Cys residues in the C6 domain (M6) decreases the level of dimerization while changes in the MANSC domain (M5) do not seem to affect this process. These results support those described for the deletion proteins. The relative strength of the signals from the ER-glycosylated and the fully-glycosylated bands (around 160 and 200 kDa, respectively, for the monomeric forms) varies between the different proteins. While mutations only in the TM-cytoplasmic domain (M3) do not seem to have any effect (compare with M0), alteration of the Cys residues in either MANSC (M1) or C6 (M2) domains lead to a reduction of the fully-glycosylated forms. These results suggest that the folding of the protein is affected in these cases and the missfolded proteins would

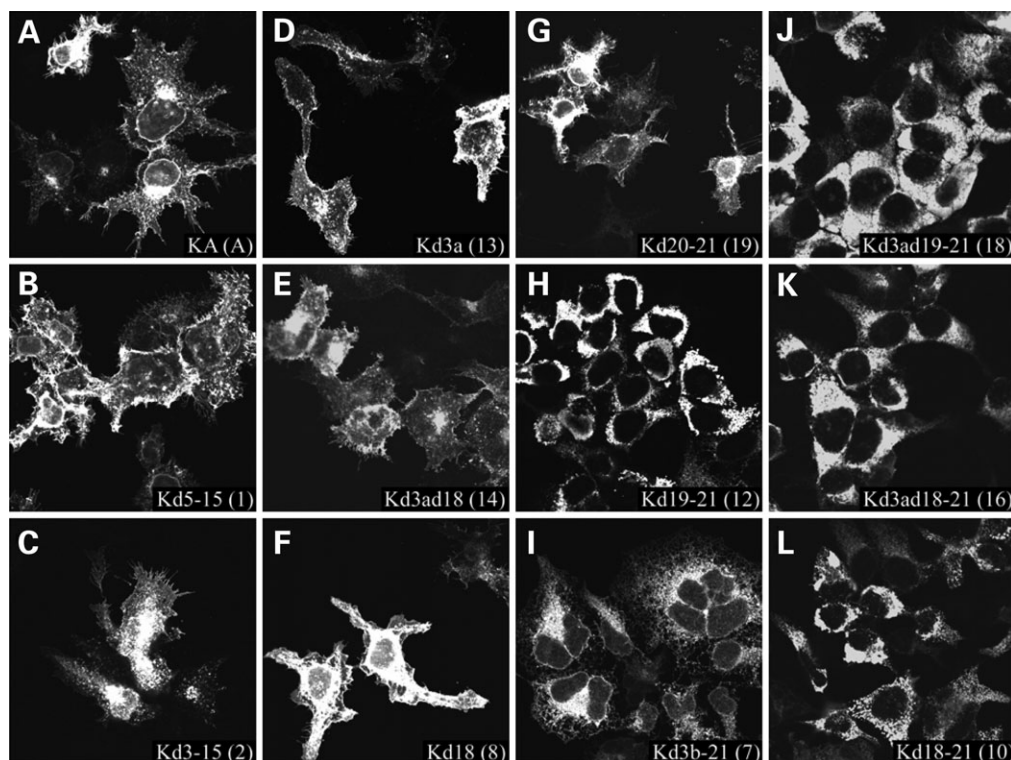


Figure 4. Sub-cellular localization of deletion-proteins. C-terminal myc + His-tagged proteins were over-expressed in HEK293T cells, processed for IF and detected with monoclonal anti-myc antibody. Identification of each protein is shown at the bottom right corner according to Table 1 and Fig.1.

accumulate in the ER. Similar results were detected by IF, where the PM localization of the mutant proteins was very reduced for MANSC and/or C6 Cys-mutants (data not shown).

Glycosylation of the KIAA0319 protein

As described above, KIAA0319 is N-glycosylated. However, after treatment with PNGase F to remove N-glycosylation, there were still four bands detected for the full-length protein (Fig. 3D) and not just the two expected, deglycosylated, bands (monomer and dimer). This result suggested that post-translational modifications other than N-glycosylation were present. In order to further characterize these modifications, several deletion constructs (2–7 in Fig. 1B and Table 1) were used to narrow down the regions involved. These deletion constructs were transfected into HEK293T cells and the expressed proteins analysed by WB. When the proteins lack the TM domain (all but Kd3-15 [2]), the ‘cell lysates’ mostly represent those proteins retained in the ER, carrying only modifications that are added in that compartment, and the ‘culture medium’ represents the proteins that have been secreted, carrying all the modifications added throughout the secretory pathway.

Post-translational modifications are concentrated in the central region of KIAA0319. Figure 7A shows the analysis of the N-glycosylation by enzymatic treatment with PNGase F (able to hydrolyse nearly all types of N-glycan chains from proteins) and endoglycosidase H (endoH) (able to cleave only high mannose structures and hybrid structures,

leaving the first N-acetylglucosamine (GlcNAc) residue attached to the protein; Golgi-modified glycoproteins are usually resistant to endoH treatment). We split the KIAA0319 protein into three fragments (relative to PKD domains): N-terminal (Kd5-21 [4]), central (PKD domains) (K-PKD [3]) and C-terminal (Kd3-15 [2]). The C-terminal fragment did not show any difference in size after PNGase F treatment and its apparent size was as expected (33.74 kDa) (Fig. 7A, top panel), suggesting that this region lacks post-translational modifications that affect protein size. The central region including all the PKD domains presents a clear size shift after enzymatic treatment, reaching then an apparent size similar to that expected for the unmodified protein (53.41 kDa) in both the ER-retained and the secreted proteins (Fig. 7A, second panel). This suggests that the PKD region is modified only by N-glycosylation, the contribution of such a modification to the protein final size being of ~10–15 kDa (compare lanes 2 and 4 with lanes 3, 5 and 6). Finally, the N-terminus (Fig. 7A, third panel) also shows a shift in size after removal of the N-glycosylation (compare lanes 1 with 3 and 5, or 2 and 4 with 6). The size of the N-deglycosylated protein from cell lysates (lanes 3 and 5) is similar to that expected for this protein (38.54 kDa), but the secreted deglycosylated protein (lane 6) shows a much higher size, ~75–80 kDa. This indicates that N-glycosylation contributes with 10–15 kDa to the final size of this fragment while post-translational modifications other than N-glycosylation represent ~35 kDa. Analyses of deletion proteins containing smaller N-terminal regions (Kd4b-21 [5], Kd3c-21 [6] and Kd3b-21 [7], with a expected size of 32.29,

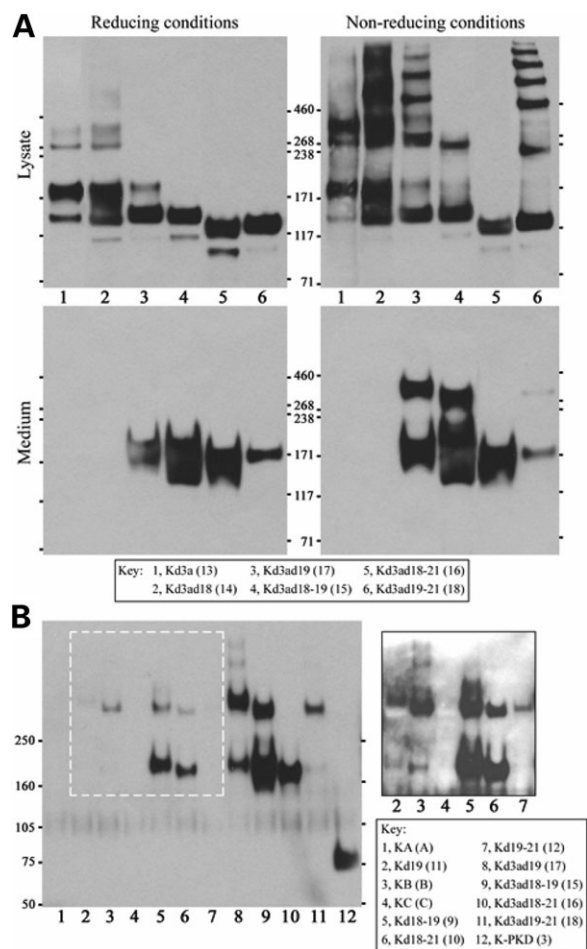


Figure 5. Analysis of secreted deletion-proteins. KIAA0319 deletion proteins (Table 1 and Fig. 1) with C-terminal myc + His tags (6×His tag for P-PKD) were over-expressed in HEK293T cells. (A) Comparison of secreted, His-purified, proteins (medium) with the proteins present in the cell lysates, both in reducing and non-reducing conditions; detection with monoclonal anti-myc antibody. (B) Direct analysis under non-reducing conditions of secreted proteins present in the conditioned medium (no His-purification); detection with polyclonal anti-His antibody. The inset shows a higher exposure of the marked region of the blot.

25.15 and 18.68 kDa, respectively; Fig. 7A, three bottom panels) showed that Kd3b-21 [7], containing the very N-terminus of the protein, has no modifications (bottom panel); Kd3c-21 [6] presents only a size shift of 4–5 kDa due to N-glycosylation (compare lanes 2 with 5 and 6, second panel from bottom); and Kd4b-21 [5] (third panel from bottom) presents a pattern similar to the one described for Kd5-21 [4] but with smaller size shifts, ~15 kDa being non-N-glycosylation (compare lanes 5 and 6) versus around 35 kDa for Kd5-21 [4]. These results indicate that the non-N-glycosylation modifications affecting the size of KIAA0319 are localized just N-terminal to the PKD domains.

KIAA0319 is O-glycosylated. The region N-terminal to the PKD domains in KIAA0319 described above was predicted to be O-glycosylated by *in silico* analysis (Fig. 1A). In order to check this possibility, the deletion proteins K-PKD [3], Kd5-21 [4] and Kd4b-21 [5] were over-expressed in the

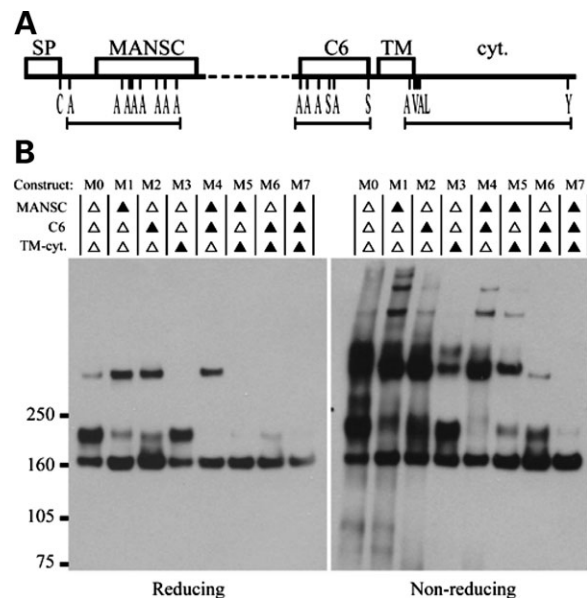


Figure 6. Effect of the Cys residues in dimerization. (A) Representation of the KIAA0319 full-length protein termini showing the distribution of the Cys residues (positions 20, 25, 57, 61, 62, 67, 77, 82, 88, 913, 917, 923, 930, 932, 952, 977, 979, 980, 981 and 1068). These residues, except for C20, were changed to Ala (A), Ser (S), Val (V), Leu (L) or Tyr (Y) as indicated, and included in groups (C25 to C88; C913 to C952; C977 to C1068) in several expression constructs with myc+His C-terminal tags. (B) Analysis of Cys mutation proteins under reduced and non-reducing conditions; detection with monoclonal anti-myc antibody. The state of those residues is indicated by open triangles (no mutation) and black triangles (all the Cys residues are mutated) for the MANSC, C6, and TM-cytoplasmic domains.

Chinese hamster ovary (CHO)-LecR mutant cell line Lec3.8.2.1, deficient in both N- and O-glycosylation (19). Figure 7B shows the results from similar experiments to those shown in Fig. 7A, but using the LecR mutant and the CHO-K1 control cell line. The results obtained with the cell lysates are the same in control and mutant cell lines (lanes 1 to 3 versus lanes 4–6). However, the secreted proteins show clearly different sizes due to the deficiency in the glycosylation pathways of CHO-LecR (lanes 7–9 versus lanes 10–12). In the case of Kd5-21 [4] and Kd4b-21 [5], these differences remain even after removal of the N-glycosylation with PNGase F (compare lanes 9 and 12 in central and bottom panels) suggesting that KIAA0319 is O-glycosylated. To further check this possibility, enzymatic treatment with neuraminidase, O-glycosydase and PNGase F were performed on these two secreted proteins in CHO-K1 cell line (Fig. 7C). Neuraminidase treatment will remove terminal neuraminic acids from both N- and O-glycosidic-bound oligosaccharide chains; O-glycosidase will release the disaccharide Galβ (1–3)GalNAc from O-glycans bound to Ser or Thr residues if it is not substituted with neuraminic acid. Incubation with neuraminidase leads to size shift even after PNGase F treatment, an indication of neuraminic acid substitution in the O-glycans of this protein (compare lanes 1 and 7, and 2 and 8 in Fig. 7C). Incubation with O-glycosidase, with or without PNGase F treatment, makes no difference with the control (lanes 5 and 6 versus lanes 7 and 8) when it is used alone but it does when samples are previously treated with neuraminidase (lanes 3 and 4). A similar result was obtained for Kd5-21

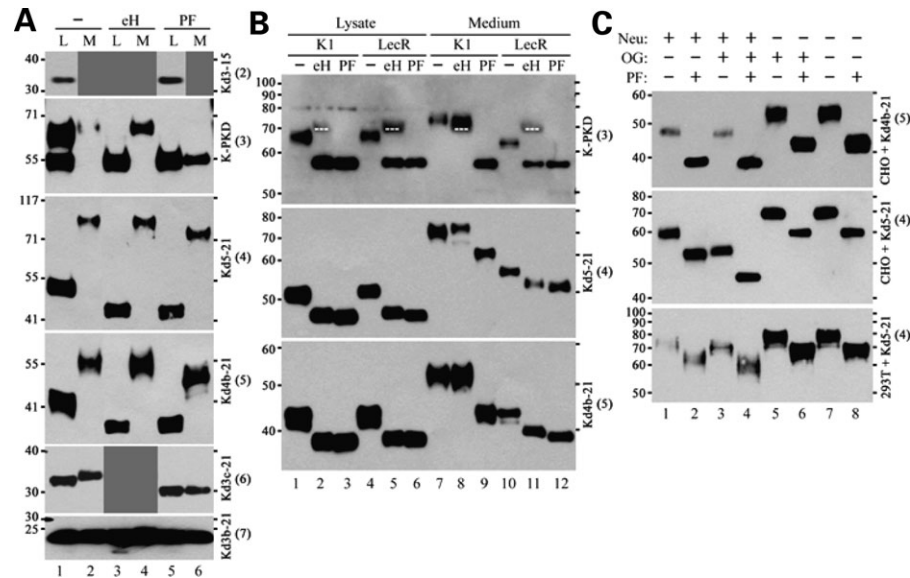


Figure 7. Glycosylation analysis of KIAA0319. C-terminal myc + His tagged (6×His for P-PKD) deletion proteins (Table 1 and Fig. 1) were over-expressed in mammalian cell lines and treated with endoH (eH), PNGase F (PF), Neuraminidase (Neu) and/or O-glycosylase (OG) as indicated; detection with monoclonal anti-myc antibody unless indicated otherwise. (A) N-glycosylation in proteins expressed in HEK293T cells; L, cell lysates; M, His-purified proteins from conditioned medium; P-PKD detected with rabbit anti-His antibody. (B) Glycosylation in proteins from cell lysates and from conditioned medium (His-purified) after over-expression in CHO-K1 (control) and CHO-LecR (glycosylation-deficient) cell lines; P-PKD detected with rabbit R2 antiserum. The bands crossed with white dotted lines in eH lanes in the top panel are endoH protein (cross-reaction detected after film overexposure). (C) Enzymatic detection of O-glycosylation in His-purified secreted proteins after overexpression in CHO-K1 (top and medium) and HEK293T (bottom) cell lines.

[4] in HEK293T cells (Fig. 7C, bottom panel) although the protein is ~10 kDa bigger than in CHO cells (compare central and bottom panels). This indicates that the degree of modification in N- and/or O-glycans is different in these two cell lines.

DISCUSSION

Research focused on the genetics of dyslexia has yielded promising results in the last few years. The identification of several candidate genes has opened the way for the analysis at the molecular level of the causal mechanisms of this condition. Except for the recently reported candidates for *DYX3* locus (13), where no functional data are still available, all the candidate genes participate in brain development through processes such as neuronal migration or axon growth (7,9,15,20). The presence of specific domains and/or the homology with other known proteins has allowed the proposal of possible mechanisms by which these proteins could be involved in developmental processes (reviewed in 14). However, for most of these proteins, there are no experimental data to support these hypotheses.

We show here the first experimental data for the characterization of the KIAA0319 protein. These data corroborate some of the predicted features of this protein: single TM domain, PM localization, C-terminal cytoplasmic domain and N-terminal extracellular domain. This type I topology is necessary for the proposed function of KIAA0319 as a protein involved in the interaction between neurons and glial fibres during neuronal migration (15,16). Such an interaction would most probably be mediated by the PKD domains

present in the central region of KIAA0319. PKD domains have been shown to form homophilic interactions and to mediate cell–cell adhesion in polycystin-1 (18,21). Apart from the full-length (variant A) protein, we also analysed the proteins encoded by variants B (KB) and C (KC) (16). Both isoforms lack the TM domain encoded by exon 19 and, as expected, their localization is not membrane-associated. The over-expressed proteins are clearly detected in the ER but, interestingly, KB is secreted while KC is not (at least at detectable levels). The putative implications of these results point towards a functional spectrum of KIAA0319 wider than previously thought. The soluble, secreted KB isoform might be able to act as a signalling protein where the MANSC and/or PKD domains could mediate the interaction with receptor proteins in other cells. Alternatively, or maybe in addition to the signalling function, this secreted form could act as a regulator/competitor in the postulated cell–cell interactions mediated by the membrane-attached full-length KIAA0319. The functional role of KC, however, may be completely different. The fact that, despite lacking the TM domain, this isoform is not secreted suggests that the mRNA might be involved in transcript regulation of the *KIAA0319* gene; another possibility is that KC may participate in KIAA0319 isoform-heterologous dimerization. Similar examples of membrane proteins with one or several soluble isoforms are known, and these soluble forms can modulate the function of the membrane-bound form, lose that function or even acquire new roles (22).

All the deletion proteins generated in this work lacking the TM domain were secreted, even when the ‘cytoplasmic domain’, encoded by exons 20+21, was present. This

indicates that there must be a conformational difference in KC that prevents the protein from exiting the ER, or a sequence acting as an ER-retention signal. The secreted condition of these deletion proteins provided us with a useful tool to check the role of different regions of the protein. When a protein is secreted it has passed the ER quality control, suggesting that it has been folded properly. In contrast, proteins without the right folding state are retained in the ER (23,24). These secreted KIAA0319-deletion proteins were usually detected in both monomeric and dimeric conformation, while the same proteins retained in the cell could be detected also as multimers. Formation of dimers depended on the presence of Cys-rich regions in the proteins, and were detected in denaturing conditions, suggesting that this conformation is mediated by Cys residues. This was confirmed by site directed mutagenesis. The Cys residues in the TM domain seem to have a particularly important role in this process. The region here called C6, encoded by exon 18, also showed a particular effect in dimer formation. This suggests that several Cys residues in these domains are probably involved in disulphide bond formation between the two molecules of the KIAA0319 dimer. The C6 region shows a close similarity with the epidermal growth factor-like module (EGF domain) (25,26), although it does not fit the proposed consensus sequence and, therefore, was not detected as such in the domain and motif searches. The EGF domain forms three disulphide bridges between the six Cys residues of its sequence. However, the disulphide bridge pattern in KIAA0319 might be different, involving Cys residues from different molecules. In any case, some of the functions of KIAA0319 would most probably require a dimeric conformation. In experiments performed in polycystin-1 (18), the authors suggest that the PKD domains might mediate homodimerization on the same membrane as well as *trans* interactions between opposing molecules at the cell-cell contact site. However, these experiments do not indicate that the dimeric conformation is necessary for the adhesion/cell-cell interaction function.

The deletion proteins in which the TM domain was not removed showed a PM localization. However, compared with the full-length protein, this pattern was weaker when Cys-rich regions were removed, possibly due to a lower rate of properly folded protein. These results indicate that only the TM domain is needed for PM localization, although the presence of some other regions might be important to obtain proper folding and conformation that will then facilitate the exit from the ER. Alteration of the Cys residues in the MANSC and the C6 domains lead to an increase in misfolded protein, suggesting an important role of these residues in the structure of these domains. The same does not seem to be true for the residues in the TM and cytoplasmic domains, as their mutation only affects the dimerization process but does not lead to an increased ER retention of the protein.

We also showed that KIAA0319 is both N- and O-glycosylated. Glycosylation is a very important post-translational modification, with roles in developmental processes (27). N-glycosylation can be involved in a wide range of molecular and cellular processes, such as protein folding, stability, trafficking or cell adhesion (28,29). Similarly, O-glycosylation has been found to function in protein structure and stability, immunity, receptor-mediated

signalling, non-specific protein interactions, modulation of the activity of enzymes and signalling molecules, and protein expression and processing (30). The most abundant form of O-linked glycosylation is the 'mucin-type', characterized by *N*-acetylgalactosamine (GalNAc) attached to the hydroxyl group of Ser/Thr side chains, which is restricted to proteins that pass through the Golgi compartment (31,32). The level of substitution in the glycan chains varies between different cell types and at different times as a function of the enzymes from the glycosylation-pathways that are available. We analysed the glycosylation pattern of KIAA0319 in the human cell line HEK293T and also in CHO cells, commonly used for this kind of analysis due to the different LecR mutants available (33). Glycoproteins are synthesized with severely truncated N- and O-linked carbohydrates in the CHO Lec3.2.8.1. cell line (33), the predicted *N*-glycans being (Man)₅(GlcNAc)₂ and the predicted *O*-glycans (mucin-type) remaining as a single GalNAc residue; other *O*-glycans would also lack neuraminic acids. The single GalNAc residue expected in each O-glycosylated site in the LecR cell line would contribute with 0.221 kDa to the molecular weight of the protein; the difference in size (~8–10 kDa) between the cell lysate and the secreted Kd5-21 [4] protein in this cell line after N-deglycosylation (lanes 6 and 12 in Fig. 7B, central panel) suggests the presence of a high number of mucin-type O-glycosylation sites and/or other kind of modification [probably other type(s) of O-glycosylation] in the region of the protein encoded by exon 4, just N-terminal to the PKD domains. Unlike for N-glycosylation, no specific motif has been identified as an O-glycosylation target (31), but the mucin-type *O*-glycans typically occur in segments of membrane-associated or secreted polypeptides that are rich in Ser, Thr and Pro (27). The KIAA0319 protein region modified by O-glycosylation meets these criteria, with these three amino acids accounting for ~50% of the residues (S, 18.46%; T, 10.77%; P, 14.62% for region 216–345, present in deletion protein Kd5-21 [4] but not in the non-O-glycosylated Kd3c-21 [6], and S, 20.31%; T, 15.62%; P, 20.31% for the 268–331 region, encoded by exon 4). The results obtained with the CHO cell lines provide evidence that KIAA0319 is O-glycosylated and indicate that this glycosylation is mainly of the mucin-type. This was further confirmed by the fact that most of the O-linked carbohydrates were removed after *O*-glycosidase treatment, which also proves that this modification is of the mucin-type with core 1 structure (31). The results obtained in the human HEK293T cell line indicate that the glycosylation pattern is more complex than in CHO cells, and it is likely that such a pattern is even more complex in the brain-expressed protein. We used two different specific antibodies against the PKD-domains region to detect the endogenous protein from human brain lysates. Two bands were detected at similar positions in the blots as the ER-glycosylated monomeric and dimeric forms of the over-expressed proteins, suggesting that they correspond to the endogenous KIAA0319. However, we did not detect any difference between reducing and non-reducing conditions, or after PNGaseF treatment (not shown). These results could be explained by a masking effect due to post-ER modification of the protein, most probably O-glycosylation, which could interfere in the detection by the mentioned antisera.

Over-expressed protein in CHO-LecR and in HEK293S-GnTI-, a human cell line deficient in GlcNAc-T1 activity shown to present *N*-glycans at the (Man)₅(GlcNAc)₂ stage (34), was detected in the PM (data not shown), suggesting that full glycosylation is not necessary for KIAA0319 to reach PM localization. The function that glycosylation may have in the KIAA0319 protein is still unknown, but participation in cell–cell interaction or in cell adhesion is plausible.

The basic characterization of the KIAA0319 protein that we report here is the first step of the road towards the final purpose of understanding the role that the different isoforms of KIAA0319 may have in brain development and how small changes in gene expression can affect this process. There are still many questions that remain opened. What is the specific function(s) of the different domains in this protein? Do the functions of KIAA0319 change depending on the oligomeric status of the protein? What is the role of the glycosylation detected in the central region of the protein? We are currently addressing several of these questions and will hopefully be able to fill in some of the still many gaps in the understanding of the molecular mechanisms involved in DD.

MATERIALS AND METHODS

Plasmids

Variants A (full-length), B (no exon 19) and C (no exons 19+20) from human *KIAA0319* gene have been described previously (16). Standard PCR and molecular cloning techniques were used to generate mammalian expression plasmids for these three forms as well as for a number of deletion and mutant constructs as described in Table 1, Figs 1 and 6, using the pcDNA4-TO-mycHis vector (Invitrogen). The EGFP-tag was obtained from the pEGFP-N1 plasmid (Clontech). Primer sequences and cloning conditions are available upon request. All constructs use the KIAA0319 SP sequence except for pHLK-PKD, obtained in the expression vector pHLsec (35), which uses the μ -phosphatase secretion leader sequence.

Antibodies

Custom polyclonal KIAA0319-specific antisera R1 and R2 were obtained from Eurogentec Ltd after immunization of two rabbits with peptides [A+B] (residues 471–485, C+EEINGPFIEEKTSVD, and 255–269, C+QLQEQQSSNS GKEVL, of human KIAA0319 protein, respectively). Polyclonal antibodies K1r and K2r were obtained from Harlan UK Ltd after immunization of two rats with the deletion protein K-PKD (Table 1), expanding the whole PKD domain region of KIAA0319. Primary anti-tag antibodies were: mouse monoclonal 9E10 against myc tag, rabbit polyclonal against 6-His tag (Bethyl Laboratories), rabbit polyclonal anti-GFP (Molecular Probes) and goat polyclonal anti-GFP (abcam). Mouse monoclonal A1/182 antibody against BAP31 (36), kindly provided by Dr H.P. Hauri (Basel, Switzerland), was used as an ER marker. Sheep polyclonal anti-TGN46 was kindly provided by Dr S. Ponnambalam (Leeds, UK), and used as a TGN marker. Secondary antibodies were: goat

anti-mouse IgG-HRP and goat anti-rabbit IgG-HRP (BioRad); bovine anti-goat IgG-HRP (Santa Cruz Biotechnology); rabbit anti-rat IgG-HRP (Sigma) and Alexa-Fluor 488 (green) goat anti-mouse and Alexa-Fluor 594 (red) goat anti-rabbit (Molecular Probes).

Cells, cell culture and transfection

The cell lines used in this study were Human embryonic kidney 293T, MRC-5 SV2 (SV40-transformed human fetal lung fibroblasts), CHO-K1 and the glycosylation-deficient (LecR) CHO-Lec3.2.8.1. (19). The CHO Lec3.2.8.1. cell line carries mutations affecting UDP-GlcNAc 2-epimerase [*lec3*, involved in the synthesis of neuraminic acid (Neu5Ac)], CMP-Neu5Ac Golgi transporter (*lec2*, needed for transport of Neu5Ac into the Golgi), UDP-Gal Golgi transporter (*lec8*, needed for the transport of UDP-Gal into the Golgi) and GlcNAc-T1 (*lec1*, needed for the addition of the first GlcNAc residue to the (Man)₅(GlcNAc)₂ backbone of *N*-glycans in the Golgi) (33). HEK293T and MRC5 cells were grown in Dulbecco's modified Eagle medium (Sigma). CHO cell lines were grown in Nutrient mixture F-12 HAM (Sigma). The media were supplemented with 10% fetal bovine serum, 2 mM L-Glutamine and 100 U/ml penicillin/streptomycin, and the cells grown at 37°C with 5% CO₂. Cells were transfected using GeneJuice (Novagen) as transfection reagent and processed 36–48 h post-transfection.

Immunofluorescence microscopy

Cells were plated onto coverslips (coated with poly-L-lysine for HEK293T cells) and transfected the following day as described above. Medium was removed after 36–48 h, and the cells washed with phosphate-buffered saline (PBS), fixed for 30 min with 3% paraformaldehyde in PBS at room temperature and washed five times with PBS. For permeabilization, fixed cells were incubated for 10 min with IF buffer [0.1% Triton X-100 in Tris-buffered saline (TBS)]. Cells were incubated for 1 h with primary antibodies diluted in IF blocking buffer (0.1% Triton X-100 in TBS+10 mg/ml fish gelatine + 10% normal goat serum) and washed five times with PBS. After incubation with Alexa Fluor-conjugated secondary antibodies (diluted in blocking buffer), coverslips were washed and mounted with Fluoromount-G (Southern Biotech) and examined using a Zeiss LSM510 META Confocal Imaging System. In 'no-permeabilization' conditions, Triton X-100 was omitted from all buffers.

Protein preparation, immunoprecipitation, biotinylation and western blot analysis

Cells were harvested at 1000 rpm for 5 min at 4°C in a standard microcentrifuge, washed once with PBS, resuspended in lysis buffer (50 mM Tris–HCl, pH = 8, 150 mM NaCl, 1% Triton X-100) containing protease inhibitor cocktail (Sigma) and incubated on ice for 20–30 min. Cell lysates (post-nuclear supernatant) were obtained by centrifugation at 4000 rpm for 10 min at 4°C. His-tagged proteins were purified using His-Select nickel affinity gel (Sigma) according to the manufacturer's instructions. Human normal adult cerebral cortex

tissue lysate and human fetal frontal lobe membrane protein lysate (ProSci Inc) were used for the detection of the endogenous protein. For IP, cell lysates were incubated with Protein G (Sigma) for 2–3 h at 4°C and cleared at 14 000 rpm for 5 min. Protein G Sepharose 4 Fast Flow (GE Healthcare) beads were washed thrice with PBS, resuspended in IP buffer (50 mM Tris–HCl, pH=7.8, 150 mM NaCl, 1 mM EDTA) containing protease inhibitor cocktail, incubated with primary antibody for 3–4 h at room temperature, and washed thrice with IP buffer. Antibody-coated Protein G Sepharose beads were added to the pre-cleared cell lysates, incubated overnight at 4°C, washed five times with IP buffer and resuspended in the same buffer. For protein biotinylation, the Pinpoint cell surface protein isolation kit (Pierce) was used according to manufacturer's recommendations. Proteins were prepared with NuPAGE LDS sample buffer, with or without reducing agent, and subjected to denaturing electrophoresis using NuPAGE Novex 3–8% Tris-Acetate or NuPAGE Novex 10% Bis-Tris gels (Invitrogen) following the manufacturer's instructions. After electrophoresis, gels were transferred onto Invitrolon PVDF membranes (Invitrogen), checked by Ponceau-staining and then destained and processed for WB using standard conditions. Immunoreactive bands were visualized with ECL Plus Western Blot Detection Reagent (GE Healthcare).

Enzymatic analysis of glycosylation

Prior to electrophoresis, cell lysates and/or His-purified proteins were treated with PNGase F, recombinant endoglycosidase H (New England Biolabs), *Clostridium perfringens* Acylneuraminyl hydrolase (Neuraminidase) and/or *O*-glycopeptide endo- β -galactosyl-*N*-acetyl- α -galactosaminohydrolase (*O*-glycosidase) (Roche), according to manufacturer's instructions. For the experiment shown in Fig. 7C, His-purified proteins (eluted in 50 mM sodium phosphate, pH=8, 300 mM NaCl, 250 mM imidazole) were diluted 1:5 in 50 mM sodium phosphate, pH=7.5, and sequentially treated at 37°C, with or without enzyme, with neuraminidase for 2 h, *O*-glycosidase for 2 h and (after denaturation at 100°C for 10 min in denaturation buffer+NP40) PNGase F for 1 h.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

We thank Dr Radu Aricescu for the pHLsec vector and for his help with glycosylation-deficient cell lines, and Dr Jo Nettleship and Nahid Rahman for the purification of protein K-PKD used in antibody production. We also thank Drs Zoe Holloway and Clotilde Lévecque for helpful discussions and comments on the manuscript.

Conflict of Interest statement. None declared.

FUNDING

Funding to pay Open Access publication charges for this article was provided by The Wellcome Trust.

REFERENCES

- Habib, M. (2000) The neurological basis of developmental dyslexia: an overview and working hypothesis. *Brain*, **123**, 2373–2399.
- Williams, J. and O'Donovan, M.C. (2006) The genetics of developmental dyslexia. *Eur. J. Hum. Genet.*, **14**, 681–689.
- Taipale, M., Kaminen, N., Nopola-Hemmi, J., Haltia, T., Myllyluoma, B., Lyytinen, H., Muller, K., Kaaranen, M., Lindsberg, P.J., Hannula-Jouppi, K. *et al.* (2003) A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain. *Proc. Natl Acad. Sci. USA*, **100**, 11553–11558.
- Francks, C., Paracchini, S., Smith, S.D., Richardson, A.J., Scerri, T.S., Cardon, L.R., Marlow, A.J., MacPhie, I.L., Walter, J., Pennington, B.F. *et al.* (2004) A 77-kilobase region of chromosome 6p22.2 is associated with dyslexia in families from the United Kingdom and from the United States. *Am. J. Hum. Genet.*, **75**, 1046–1058.
- Cope, N., Harold, D., Hill, G., Moskvina, V., Stevenson, J., Holmans, P., Owen, M.J., O'Donovan, M.C. and Williams, J. (2005) Strong evidence that *KIAA0319* on chromosome 6p is a susceptibility gene for developmental dyslexia. *Am. J. Hum. Genet.*, **76**, 581–591.
- Harold, D., Paracchini, S., Scerri, T., Dennis, M., Cope, N., Hill, G., Moskvina, V., Walter, J., Richardson, A.J., Owen, M.J. *et al.* (2006) Further evidence that the *KIAA0319* gene confers susceptibility to developmental dyslexia. *Mol. Psychiatry*, **11**, 1085–1091.
- Meng, H., Smith, S.D., Hager, K., Held, M., Liu, J., Olson, R.K., Pennington, B.F., DeFries, J.C., Gelernter, J., O'Reilly-Pol, T. *et al.* (2005) *DCDC2* is associated with reading disability and modulates neuronal development in the brain. *Proc. Natl Acad. Sci. USA*, **102**, 17053–17058.
- Schumacher, J., Anthoni, H., Dahdouh, F., Konig, I.R., Hillmer, A.M., Kluck, N., Manthey, M., Plume, E., Warnke, A., Remschmidt, H. *et al.* (2006) Strong genetic evidence of *DCDC2* as a susceptibility gene for dyslexia. *Am. J. Hum. Genet.*, **78**, 52–62.
- Hannula-Jouppi, K., Kaminen-Ahola, N., Taipale, M., Eklund, R., Nopola-Hemmi, J., Kaariainen, H. and Kere, J. (2005) The axon guidance receptor gene *ROBO1* is a candidate gene for developmental dyslexia. *PLoS Genet.*, **1**, e50.
- Fisher, S.E. and Francks, C. (2006) Genes, cognition and dyslexia: learning to read the genome. *Trends Cogn. Sci.*, **10**, 250–257.
- McGrath, L.M., Smith, S.D. and Pennington, B.F. (2006) Breakthroughs in the search for dyslexia candidate genes. *Trends Mol. Med.*, **12**, 333–341.
- Paracchini, S., Scerri, T. and Monaco, A.P. (2007) The genetic lexicon of dyslexia. *Annu. Rev. Genomics Hum. Genet.*, **8**, 57–79.
- Anthoni, H., Zucchelli, M., Mattson, H., Muller-Myhsok, B., Fransson, I., Schumacher, J., Massinen, S., Onkamo, P., Warnke, A., Griesemann, H. *et al.* (2007) A locus on 2p12 containing the co-regulated *MRPL19* and *C2ORF3* genes is associated to dyslexia. *Hum. Mol. Genet.*, **16**, 667–677.
- Galaburda, A.M., LoTurco, J., Ramus, F., Fitch, R.H. and Rosen, G.D. (2006) From genes to behavior in developmental dyslexia. *Nat. Neurosci.*, **9**, 1213–1217.
- Paracchini, S., Thomas, A., Castro, S., Lai, C., Paramasivam, M., Wang, Y., Keating, B.J., Taylor, J.M., Hacking, D.F., Scerri, T. *et al.* (2006) The chromosome 6p22 haplotype associated with dyslexia reduces the expression of *KIAA0319*, a novel gene involved in neuronal migration. *Hum. Mol. Genet.*, **15**, 1659–1666.
- Velayos-Baeza, A., Toma, C., da Roza, S., Paracchini, S. and Monaco, A.P. (2007) Alternative splicing in the dyslexia-associated gene *KIAA0319*. *Mamm. Genome*, **18**, 627–634.
- Guo, J., Chen, S., Huang, C., Chen, L., Studholme, D.J., Zhao, S. and Yu, L. (2004) MANSC: a seven-cysteine-containing domain present in animal membrane and extracellular proteins. *Trends Biochem. Sci.*, **29**, 172–174.
- Ibraghimov-Beskrovnaya, O., Bukanov, N.O., Donohue, L.C., Dackowski, W.R., Klinger, K.W. and Landes, G.M. (2000) Strong homophilic interactions of the Ig-like domains of polycystin-1, the protein

- product of an autosomal dominant polycystic kidney disease gene, *PKD1*. *Hum. Mol. Genet.*, **9**, 1641–1649.
19. Stanley, P. (1989) Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity. *Mol. Cell. Biol.*, **9**, 377–383.
 20. Wang, Y., Paramasivam, M., Thomas, A., Bai, J., Kaminen-Ahola, N., Kere, J., Voskuil, J., Rosen, G.D., Galaburda, A.M. and Loturco, J.J. (2006) *DYX1C1* functions in neuronal migration in developing neocortex. *Neuroscience*, **143**, 515–522.
 21. Streets, A.J., Newby, L.J., O'Hare, M.J., Bukanov, N.O., Ibraghimov-Beskrovnaya, O. and Ong, A.C. (2003) Functional analysis of *PKD1* transgenic lines reveals a direct role for polycystin-1 in mediating cell-cell adhesion. *J. Am. Soc. Nephrol.*, **14**, 1804–1815.
 22. Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A. and Soreq, H. (2005) Function of alternative splicing. *Gene*, **344**, 1–20.
 23. Trombetta, E.S. and Parodi, A.J. (2003) Quality control and protein folding in the secretory pathway. *Annu. Rev. Cell Dev. Biol.*, **19**, 649–676.
 24. Ellgaard, L. and Helenius, A. (2003) Quality control in the endoplasmic reticulum. *Nat. Rev. Mol. Cell. Biol.*, **4**, 181–191.
 25. Campbell, I.D. and Bork, P. (1993) Epidermal growth factor-like modules. *Curr. Opin. Struct. Biol.*, **3**, 385–392.
 26. Sanderson, M.P., Dempsey, P.J. and Dunbar, A.J. (2006) Control of ErbB signaling through metalloprotease mediated ectodomain shedding of EGF-like factors. *Growth Factors*, **24**, 121–136.
 27. Haltiwanger, R.S. and Lowe, J.B. (2004) Role of glycosylation in development. *Annu. Rev. Biochem.*, **73**, 491–537.
 28. Ohtsubo, K. and Marth, J.D. (2006) Glycosylation in cellular mechanisms of health and disease. *Cell*, **126**, 855–867.
 29. Freeze, H.H. (2006) Genetic defects in the human glycome. *Nat. Rev. Genet.*, **7**, 537–551.
 30. Wopereis, S., Lefeber, D.J., Morava, E. and Wevers, R.A. (2006) Mechanisms in protein O-glycan biosynthesis and clinical and molecular aspects of protein O-glycan biosynthesis defects: a review. *Clin. Chem.*, **52**, 574–600.
 31. Hanisch, F.G. (2001) O-glycosylation of the mucin type. *Biol. Chem.*, **382**, 143–149.
 32. Hang, H.C. and Bertozzi, C.R. (2005) The chemistry and biology of mucin-type O-linked glycosylation. *Bioorg. Med. Chem.*, **13**, 5021–5034.
 33. Patnaik, S.K. and Stanley, P. (2006) Lectin-resistant CHO glycosylation mutants. *Methods Enzymol.*, **416**, 159–182.
 34. Reeves, P.J., Callewaert, N., Contreras, R. and Khorana, H.G. (2002) Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible *N*-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl Acad. Sci. USA*, **99**, 13419–13424.
 35. Aricescu, A.R., Lu, W. and Jones, E.Y. (2006) A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. D Biol. Crystallogr.*, **62**, 1243–1250.
 36. Klumperman, J., Schweizer, A., Clausen, H., Tang, B.L., Hong, W., Oorschot, V. and Hauri, H.P. (1998) The recycling pathway of protein ERGIC-53 and dynamics of the ER-Golgi intermediate compartment. *J. Cell Sci.*, **111**, 3411–3425.