ORIGINAL ARTICLE

An isoform of the severe encephalopathy-related *CDKL5* gene, including a novel exon with extremely high sequence conservation, is specifically expressed in brain

Yann Fichou^{1,2}, Juliette Nectoux^{1,2}, Nadia Bahi-Buisson^{2,3}, Jamel Chelly^{1,2,4} and Thierry Bienvenu^{1,2,4}

We report the identification of a novel exon, which is referred to as exon 16b, within the cyclin-dependent kinase (CDK)-like 5 (*CDKL5*) gene that is implicated in the X-linked infantile spasm syndrome and the early-onset seizure variant of Rett syndrome. Interestingly, it is highly conserved in species through evolution, suggesting a potential functional role, but does not display any homology with other referenced sequences. Most importantly, the transcript including this exon is specifically expressed in brain. We suggest that *CDKL5* exon 16b should now be considered in the genetic screening of patients presenting with a *CDKL5*-related disease profile.

Journal of Human Genetics (2011) 56, 52–57; doi:10.1038/jhg.2010.143; published online 2 December 2010

Keywords: alternative splicing; *CDKL5*; conservation; early seizures; Rett syndrome; subcellular localization

INTRODUCTION

The X-linked cyclin-dependent kinase (CDK)-like 5 gene (*CDKL5*, OMIM 300203), formerly known as the *STK9* gene,¹ encodes a serine/threonine kinase that has been associated with the X-linked infantile spasm syndrome (ISSX, OMIM 308350) and the early-onset seizure variant of Rett syndrome (RTT, OMIM 312750).^{2–8} The gene is composed of 20 encoding exons, and so far > 50 mutations in patients with *CDKL5*-related encephalopathy have been reported, affecting quasi-exclusively girls.^{3–21} Patients with *CDKL5* mutations typically present severe intellectual disability and early seizures, microcephaly in ~ 1/3 of cases, as well as RTT-like features such as severe hypotonia, hand stereotypies and sleep disturbances.²² Several patients having symptoms reminiscent of a *CDKL5*-related disease profile have had no mutation identified, in spite of the large screening strategy that is available, suggesting that other unexplored regions and/or genetic loci are involved in the disease. At the protein level, the N-terminal domain is involved in the catalytic activity of the protein, whereas the C-terminal region may regulate its subcellular localization.^{6,23} *CDKL5* phosphorylation targets are currently largely unknown. So far, it has been shown that the protein, *in vitro*, may autophosphorylate, and it phosphorylates the product of the methyl-CpG-binding protein 2 (*MECP2*, OMIM 300005) gene, which is mutated in > 90% of classic RTT patients,²⁴ suggesting a common signaling pathway

between these two proteins;^{6,9,25} and the N-terminal region of the DNA methyltransferase 1 in nuclei.²⁶

In this study, we report the coincidental finding of an additional exon within the *CDKL5* gene and describe its conservative feature in several species. This novel exon displays (1) no homology with other sequences in the human genome and (2) an extremely high level of similarity with orthologous sequences, suggesting a potential functional role. We suggest that this exon, which we referred to as exon 16b, should be screened in patients presenting with *CDKL5*-related disease with no mutations identified in the regions studied so far.

MATERIALS AND METHODS

Sequence alignment and splice site analysis

Both nucleotide and amino-acid orthologous sequences from 24 species were obtained from the Ensembl Genome Browser v57 (<http://www.ensembl.org>) in a Fasta format, and then submitted to a multiple sequence alignment with ClustalW2 (<http://www.ebi.ac.uk/Tools/clustalw2/>) with default parameters. Splice site scores were calculated with three available online tools: Analyzer Splice Tool (<http://ast.bioinfo.tau.ac.il/SpliceSiteFrame.htm>), Human Splicing Finder v2.4 (<http://www.umd.be/HSE/>)²⁷ and the Berkeley Drosophila Genome Project applied to humans (http://www.fruitfly.org/seq_tools/splice.html) with default parameters.

¹Inserm, U1016, Paris, France; ²Institut Cochin, Université Paris Descartes, CNRS (UMR 8104), Paris, France; ³Service de Neuropédiatrie, Hôpital Necker-Enfants Malades, Paris, France and ⁴Assistance Publique-Hôpitaux de Paris, Hôpital Cochin, Laboratoire de Biochimie et Génétique Moléculaire, Paris, France

Correspondence: Dr T Bienvenu, Institut Cochin—Inserm U1016, Faculté de Médecine Cochin, Département de Génétique et Développement—Port-Royal, 24, rue du Faubourg Saint-Jacques, 75014, Paris, France.

E-mail: thierry.bienvenu@inserm.fr

Received 28 June 2010; revised 18 October 2010; accepted 22 October 2010; published online 2 December 2010

Patient cohort and genetic screening

A total of 1000 individuals were selected for the screening of *CDKL5* exon 16b. The cohort consisted of 100 normal female individuals and 900 *MECP2*, *CDKL5* and *FOXG1* mutation-negative patients. The patient cohort included nine women with the classical form of Rett syndrome, 21 patients (19 women; 2 men) with the early-onset seizure variant of Rett syndrome and 11 patients (seven women; three men) with the congenital variant of Rett syndrome. The cohort also included 859 patients (800 women; 59 men) with a diagnosis reminiscent of Rett syndrome. This heterogeneous group of patients has been incompletely classified, but presents clinical features of interest according to the selection criteria for this study, including mental retardation, autistic traits, stereotypic hand movements, progressive microcephaly, epileptic encephalopathy and/or severe congenital encephalopathy without recognizable etiology. However, none of them was explained by mutations in the *MECP2*, *CDKL5*, *FOXG1* and *ARX* genes. After informed consent, genomic DNA from these 1000 individuals was extracted from peripheral blood lymphocytes according to standard protocols and served as a template for subsequent amplification of the whole *CDKL5* exon 16b, and as flanking splice sites with 0.5 µm of both forward and reverse primers 5'-AATTAGTTTATTTGTATCTTATATTGC-3' and 5'-CAGTTTTCAGGGCTACCATAC-3'. All amplifications were carried out in a GeneAmp PCR System 9700 (Applied Biosystems, Courtaboeuf, France) in a 96-well-plate format as follows: PCR Amplification Buffer II 1X, 2.0 mM MgCl₂, 250 µm dNTP each and 2.5 U AmpliTaq Gold DNA Polymerase (Applied Biosystems) in a 50 µl reaction volume with 50 ng template DNA. Cycle conditions were as follows: an initial denaturation step at 95 °C for 5 min; followed by 40 cycles of denaturation at 95 °C for 30 s, annealing at 55 °C for 30 s, extension at 72 °C for 30 s; and a final extension step at 72 °C for 2 min. PCR products were purified with the NucleoFast 96 PCR kit (Macherey-Nagel, Hoerd, France) in a microplate format. Sequencing reactions were carried out with the BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems) with 20 ng of purified fragments and 0.5 µm of either forward or reverse primer under the following conditions: 1 min at 96 °C, followed by 25 cycles at 96 °C for 10 s, 50 °C for 5 s and 60 °C for 4 min. Excess of fluorescent dye terminators from cycle sequencing was removed by centrifugation with MultiScreen HV 96-well plates (Millipore, Molsheim, France). Purified sequencing products were finally loaded on the Applied Biosystems 3130 Genetic Analyzer (Applied Biosystems) for sequencing. Output data were analyzed with Sequencher 4.8 (Gene Codes Corporation, Ann Arbor, MI, USA) and chromatograms were visually controlled by two independent operators.

Animals

All animal care and experimental procedures were carried out in accordance with the European Communities Council Directive and approved by the local ethical committee. Experiments were conducted with 6-month-old male C57Bl/6N mice bred in our own animal facility. Animals were kept in a 12 h light/dark cycle, with controlled humidity (60–80%) and temperature (22 ± 1 °C). Food and water were freely available. Cortex, hippocampus, cerebellum and olfactory bulb, as well as heart, liver, lung, kidney and muscle were prepared by standard procedures.

Total RNA extraction and transcript analysis

Total RNA was extracted from human cells and mouse tissues with the RNeasy Mini Kit (Qiagen, Courtaboeuf, France) according to the protocol described by the manufacturer, and converted into first-strand cDNA with the SuperScript II Reverse Transcriptase (Invitrogen, Cergy-Pontoise, France) using random hexamers as described. The human *CDKL5* transcript was amplified with 1 µm each primer (forward: 5'-AGACAACCAGCATTCGATCC-3'; reverse: 5'-TGCAACGTCAGAAGATCAGG-3') in 1X GeneAmp PCR Buffer II, 3.5 mM MgCl₂, 250 µm of each dNTP and 0.03 U µl⁻¹ AmpliTaq Gold DNA Polymerase (Applied Biosystems). PCR amplifications were carried out in a GeneAmp PCR System 9700 (Applied Biosystems) as follows: 95 °C for 5 min for enzyme activation; followed by 35 cycles of 95 °C for 30 s, 60 °C for 30 s, 72 °C for 30 s; and a final extension step at 72 °C for 7 min. PCR products were loaded on a 2% agarose gel for size control.

For quantitative assessment of *Cdkl5* transcripts in mouse tissues, two couples of primers were designed: one for the total *Cdkl5* transcript and

Table 1 Primers for quantitative RT-PCR analysis of total *Cdkl5* (mCdkl5) and *Cdkl5*-16b (mCdkl5-16b) transcripts in mouse tissue extracts, normalized by *Actb* (mActb) transcript analysis

Name	Primer sequence (5' → 3')
mCdkl5_F	GGACCACTTCCATCTGAGCAA
mCdkl5_R	CGGAGCCCATGGAAACG
mCdkl5-16b_F	AAAAGAAGAAACTCAAACAGTAA
mCdkl5-16b_R	AATTAAGAGAGGGAAGAGGGATT
mActb_F	TGACCTGAAGTACCCATTGA
mActb_R	CCATGTCGTCCAGTTGGTAAC

Abbreviations: F, forward; R, reverse.

another for the specific study of the alternative longer isoform (Table 1). All reactions were carried out in a 96-well plate format in a 7500 Real-Time PCR System with SYBR Green PCR Master Mix (Applied Biosystems), as described by the manufacturer, and with 0.8 µm of each primer. Conditions were 50 °C for 2 min, 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Housekeeping gene *Actb* was used for normalization. The purity of PCR products was determined by a melting-curve analysis, and data were analyzed with Sequence Detection System software v.1.3.1. (Applied Biosystems).

Plasmid construction

The open-reading frame of the *CDKL5* gene was first amplified with Platinum Taq DNA polymerase high fidelity (Invitrogen), with 1X high fidelity PCR buffer, 1.5 mM MgCl₂ and 250 µm each dNTP, using the pEGFP-CDKL5 plasmid (kindly provided by Dr Charlotte Kilstrup-Nielsen, University of Insubria, Busto Arsizio, Italy) as a template with respective sense and antisense primers (0.5 µm each): 5'-CGCGGATCCGCGATGAAGATTCCTAACATTGGTAATGTGATG-3' and 5'-TGCTCTAGAGCACTTGCCCGTCAGTGCCG-3'. *Bam*HI and *Xba*I restriction enzyme sites (bold/underlined in the above sequences) were respectively added in either primer for subsequent cloning experiments. Amplification conditions were as follows: an initial denaturation step at 94 °C for 2 min; followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 60 °C for 30 s, extension at 68 °C for 3 min; and a final extension step at 68 °C for 5 min. After purification with the QIAquick Gel Extraction Kit (Qiagen), the PCR product was cloned into the pCR4-TOPO vector (Invitrogen) with the TOPO TA Cloning Kit for Sequencing (Invitrogen). This plasmid, as well as pcDNA3.1-MECP2_e1-myc-His (kindly provided by Dr Yuzhi Zhang, Hospital for Sick Children, Toronto, Canada), were double-digested by *Bam*HI and *Xba*I restriction enzymes. After gel purification, the *CDKL5* open-reading frame was subcloned into the background vector with the T4 DNA ligase (New England Biolabs, purchased from Ozyme, Saint-Quentin-en-Yvelines, France) in 1X T4 DNA ligase buffer at 4 °C, overnight, to generate the pcDNA3.1-CDKL5-myc-His expression plasmid.

Three overlapping DNA fragments (F1, F2 and F3) were first individually amplified using first-strand cDNA from wild-type human fibroblasts as a template with primers (forward/reverse) F1 (F1_F: 5'-CCCAGGGACAAAGTACCTCA-3'/F1_R: 5'-TCATCTCTGGAGGGAGCTGT-3'), F2 (F2_F: 5'-AGCCTGCAACTCTGTGTCACC-3'/F2_R: 5'-CTTGATGCTTGATTCTCCC-3') and F3 (F3_F: 5'-GGGAGAATCCAAGCATCAAG-3'/F3_R: 5'-TTTGGGTCAACCAGCAGAAG-3'), in 1X PCR buffer, 1.5 mM MgCl₂, 250 µm each dNTP and 0.5 µm each primer with Taq DNA Polymerase (Invitrogen) under the following conditions: initial denaturation at 94 °C for 3 min; 40 cycles of denaturation at 94 °C for 30 s; annealing at 55 °C for 30 s; extension at 72 °C for 30 s; and a final extension at 72 °C for 5 min. All fragments were gel purified as described above. In a second round, equivalent molar ratios of the three purified fragments were mixed together and amplified with primers F1_F and F3_R in the same conditions as described above, except for 2 min of extension time, to generate a 1665-bp fragment that was subsequently gel purified. This amplicon, as well as pcDNA3.1-CDKL5-myc-His, were double-digested with *Eco*RI and *Bbs*I restriction enzymes (incubation, 37 °C, 60 min; inactivation, 65 °C, 20 min), separated on a 0.5% agarose gel and gel purified. Finally, the insert containing

the *CDKL5* exon 16b sequence was subcloned into the open vector in conditions described above to generate the pcDNA3.1-*CDKL5*-16b-myc-His plasmid. After inactivation at 65 °C for 10 min, ligation products were transformed, as described by the manufacturer, in XL10-Gold Ultracompetent Cells (Stratagene, purchased from Agilent Technologies, Massy, France), which were cultured in Luria-Bertani medium supplemented with 100 µg ml⁻¹ kanamycin. Plasmids were extracted and purified with a QIAfilter Plasmid Midi Kit (Qiagen). All PCR products, intermediate and final plasmid constructs, were analyzed by direct sequencing. The size of the expressed, tagged products was confirmed by western blotting analysis (data not shown).

Cell culture, transfection and immunofluorescence

COS-7 cells (ATCC number: CRL-1651) were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum and antibiotics at 37 °C, 5% CO₂ under a humidified atmosphere.

One day before transfection, cells were seeded onto round coverslips and cultured in minimal medium. They were transfected with 800 ng of plasmid pcDNA3.1-*CDKL5*-myc-His or pcDNA3.1-*CDKL5*-16b-myc-His, using Lipofectamine 2000 (Invitrogen) as the transfection reagent, following the manufacturer's instructions. At 48 h after transfection, cells were either treated with 50 nM Leptomycin B (Sigma-Aldrich, Lyon, France) for 4 h, fixed with 100% methanol at -20 °C for 5 min, permeabilized in a phosphate-buffered saline in 0.2% solution of 1X/Triton X-100 and unspecific sites blocked with a 5% non-fat milk solution diluted in 1X phosphate-buffered saline. Cells were then incubated with the primary mouse c-Myc (9E10) sc-40 monoclonal antibody (Santa Cruz Biotechnology, Heidelberg, Germany) (dilution 1:200) and then with a Texas Red-conjugated secondary rabbit anti-mouse IgG (dilution 1:4000). All coverslips were mounted with Vectashield mounting medium with 4',6-diamidino-2-phenylindole (Vector Laboratories, purchased from Abcys, Paris, France) and analyzed with a Leica DMRA2 fluorescence microscope (Leica Microsystems SAS, Nanterre, France). All steps were carried out at room temperature unless otherwise indicated.

RESULTS

A novel highly conserved exon within the *CDKL5* gene

By studying *CDKL5* expression in human fibroblasts, we coincidentally generated two products by RT-PCR (Figure 1a). Investigation of these fragments by direct sequencing revealed the expected identity for the shortest band, corresponding to a 161-bp fragment, including exons 15, 16 and 17 (Figure 1b). Surprisingly, the longest product consisted of an additional 123-bp interval incorporated within the shortest product that actually matches with a genomic sequence between *CDKL5* exons 16 and 17 (Ensembl Genome Browser v57 genomic location: chrX:18,641,999-18,642,121) and is flanked by both an acceptor (AG) and a donor (GT) splice site. Nucleotide alignment of this putative exon and the flanking splice sites in 24

species, accounting for a total of 127 bp (Figure 2a), was carried out and indicated a very high degree of homology. Indeed, a ≥95% sequence similarity was observed in 17/24 species, considering the human sequence as the reference (data not shown). These data suggest an alternative splicing of this yet unreported exon, referred to as exon 16b, which most importantly maintains the phase of the open reading frame of the gene. Acceptor splice site 'strength' assessed by bioinformatics studies (data not shown) indicated a 0.89 average score with three different tools, whereas the donor splice site was not found with Fruitfly (~0.6–0.7 score with Analyzer Splice Tool and Human Splicing Finder). These results suggest that this latter site may be alternatively recognized by the endogenous splicing machinery, and thus contributes to functionally restrict the incorporation of *CDKL5* exon 16b within the transcript to a low level.

Amino acid alignment (41 amino acid residues) also showed high homology (Figure 2b; (95% sequence similarity in 17/24 species), reinforcing the evolutionarily conserved feature of the domain and suggesting its potential functional role. Interestingly considering its conservation, searching for homologous regions by bioinformatics investigation (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>, <http://www.sciencegateway.org/tools/index.html>, <http://www.ebi.ac.uk/Tools/sequence.html>) did not reveal any evidence of similarity with any known domains in other proteins, highlighting the singularity of this sequence.

Genetic studies in controls and *CDKL5*-related patients

To investigate potential mutations in *CDKL5* exon 16b, 1000 control and patient DNAs (including patients with typical and atypical Rett syndrome without mutations in the *MECP2*, *CDKL5* and *FOXG1* genes) were screened by direct sequencing. No sequence variation could be detected within the exon and flanking sequences in either group of individuals.

The *Cdkl5*-16b transcript is expressed in mouse brain tissues

The anatomical distribution of the *Cdkl5*-16b transcript isoform was assessed in various brain sections of both fetal and adult wild-type mice by *in situ* hybridization, but no staining could be revealed in any structures (data not shown). Northern blot analysis also failed at identifying the transcript in multiple mouse tissues. *Cdkl5*-16b expression was then further assessed in several mouse tissues by quantitative RT-PCR. Results indicated that, as observed in human fibroblasts, the alternative *Cdkl5* isoform including exon 16b is detected in mouse tissues, and specifically in all regions studied in the brain at various ratios (hippocampus ≈ cortex < cerebellum < olfactory bulb), whereas it is not found in other organs (Figure 3). This result may suggest the potential functional importance of this isoform in the brain.

Subcellular localization of *CDKL5* isoforms

Human *CDKL5* and *CDKL5*-16b isoforms overexpressed as myc-tagged fusion proteins in COS-7 cells showed similar patterns. Typically, both fusion proteins are very predominantly found in the cytosolic compartment of transfected cells (Figures 4b, c, h and i; >90% cells). Immunofluorescence studies with markers of cellular organelles and structures (that is, endoplasmic reticulum, Golgi apparatus, microtubules and microfilaments) did not indicate any specific localization of the fusion proteins (data not shown). The remaining transfected cells show a nuclear staining of both fusion proteins (data not shown). These results slightly differ from those obtained by our group^{12,22} and others^{23,25} that predominantly show a nuclear staining, and suggest that the localization of overexpressed *CDKL5* proteins may depend on the cell type, and potentially on the

Figure 1 An alternatively spliced exon 16b between exons 16 and 17 within the *CDKL5* gene. (a) *CDKL5* transcript analysis revealed two PCR products. (b) Schematic representation of the amplified PCR products. Black arrows: forward and reverse PCR primers. Exon 16b: 123 bp. Black arrows indicate forward and reverse relative primer positions.

Figure 2 Phylogenetic molecular analysis of *CDKL5* exon 16b in 24 species by ClustalW2. (a) Nucleotide alignment of *CDKL5* exon 16b and flanking splice sites (127 bp). Arrows indicate the AG acceptor and the GT donor splice sites. (b) Amino acid alignment of the putative peptide translated from exon 16b (41 amino acids).

passage number and/or the cell cycle, but is not affected by the reporter tag, at least in our conditions.

Rusconi *et al.*²³ demonstrated that an active nuclear export mediates the dynamic subcellular localization of CDKL5. To further explore the dynamics of our isoform, transfected cells were treated with Leptomycin B to inhibit this nuclear export system. As previously described with CDKL5 (see ref. 23), we observed that the CDKL5-16b protein isoform is sequestered within the nucleus in ~100% of cells (Figures 4e, f, k and l), thus suggesting that the exon 16b-translated amino acid sequence does not affect the nuclear export of the protein mediated by the receptor CRM1/Exportin 1.

DISCUSSION

This study reports the identification of an additional 123-bp exon between exons 16 and 17 of the X-linked *CDKL5* gene, which we

referred to as *CDKL5* exon 16b. This sequence shows a remarkably high degree of similarity between species, suggesting a functional role that has been maintained through evolution. It is potentially translated in a supplemental 41 amino acid peptide, which is supposed to be included within the protein sequence between threonine 792 and valine 793, for a final long isoform containing 1071 amino acids, instead of the 1030 residues in the 'classical' form of the polypeptide. Curiously, and unexpectedly considering its evolutionarily conserved feature, this sequence does not show homology with any other referenced domain. Our efforts toward the identification of the full-length protein, by generating an antibody to specifically identify it in tissues, were unfortunately not successful, neither by western blotting analysis nor by immunofluorescence studies in human fibroblasts or mouse tissues. This observation does not unequivocally mean that the full-length protein is not expressed in these tissues, but may reflect its

low amount that cannot be identified by the method. Indeed, at this step of the study, there is no evidence of the existence of the 1071 full-length protein isoform. However, one of the most important findings in favor of the expression of the long *CDKL5* isoform *in vivo* comes from our observation that the transcript containing exon 16b is detected in all brain regions, suggesting that this longer isoform may have a specific function in these tissues, reinforced by the fact that patients with mutations in *CDKL5* suffer from severe encephalopathy. Although our trials to investigate the expression pattern of isoform 16b by both *in situ* hybridization and northern blot analysis did not succeed, these experiments highlight the combination of two limitations in this particular case: (1) the amount of *Cdkl5*-16b transcript that is supposed to be low and (2) the size of probes,

which are required to be 123 bp in length to be specific and are thus not optimal in terms of design.

CDKL5 exon 16b was then sequenced in numerous control individuals and patients with a *CDKL5* disease profile with unidentified mutations. No variant could be found in >1900 X chromosomes, suggesting that mutations, if existing, are not frequent within this exon. The absence of sequence variations in this exon 16b in more than 1000 individuals also reinforces the high sequence conservation in this genomic region. Another hypothesis might be that a mutation in exon 16b would result in a slightly different phenotype, thus implicating that the cohort of patients selected for the study was not correctly chosen and that the role/function of this protein is to be understood to explore the phenotype induced by potential mutations in exon 16b, particularly at the neurological level.

The finding of this additional exon in *CDKL5* parallels the discovery of exon 1 in the *MECP2* gene, the gene that is mutated in ~90% of patients with classical Rett syndrome. Until 2004 and the co-discovery of exon 1 within this latter gene,^{28,29} only three exons had been reported. Following this finding, we and others have contributed to finding pathogenic mutations in this domain,^{24,29–35} illustrating the relevance of studying the novel *CDKL5* exon 16b not only in genetic diagnosis but also to provide insights into the understanding of its function.

In 2006, Bertani *et al.*²⁵ reported a punctate staining of *CDKL5* within the nucleus in the mouse fibroblast NIH/3T3 cell line over-expressing the gene, a pattern that was later confirmed in similar experiments in the human epithelial HeLa cell line by Rusconi *et al.*²³ In this study, we also investigated the subcellular localization of both isoforms (*CDKL5* and *CDKL5*-16b) by transfecting the plasmid constructs in COS-7 cells. These constructs overexpressed in cells showed a common cellular pattern with a predominant cytosolic staining that slightly differs from previous reports,^{12,22,23,25} indicating (1) that the additional peptide encoded by exon 16b does not contain any functional signal that would mediate the protein localization and (2) that *CDKL5* localization may be cell-type dependent. *CDKL5* is a

Figure 3 Relative expression of *Cdkl5*-16b transcript to total *Cdkl5* transcript in adult mouse tissues. AU, arbitrary unit. *Cdkl5* isoforms are normalized to the mouse *Actb* gene. *R*=relative ratio between *Cdkl5*-16b and total *Cdkl5*, expressed in arbitrary unit.

Figure 4 Human *CDKL5* isoforms overexpressed in COS-7 cells with or without treatment with Leptomycin B ($\pm$ LMB). *CDKL5*-myc-His and *CDKL5*-16b-myc-His were immunolabeled with a monoclonal anti-c-Myc antibody (b, e, h, k). Fixed cells were mounted with Vectashield medium with 4',6-diamidino-2-phenylindole (a, d, g, j). The lower row shows the merged images (c, f, i, l).

serine/threonine kinase that shares common protein motifs and properties with CDKs. In terms of subcellular localization, CDKs are mostly nuclear. However, although CDK1 is nuclear in mitosis as a heterodimer with cyclin B1 (CDK1/cyclin B1), it is localized within the cytoplasm in interphase, suggesting that it may also have a role in this compartment.³⁶ In mouse astrocytes, Cdk6 is predominantly found in the cytoplasm.³⁷ These examples illustrate well the cell/tissue-specific dynamic features of this class of molecules. As study done by Rusconi *et al.*²³ highlighted the dynamic shuttling of CDKL5 between the nucleus and the cytoplasm, a similar experiment was carried out with the CDKL5-16b isoform and revealed that the nuclear export of both proteins is mediated by the receptor CRM1/Exportin 1.

In summary, we found a novel, highly conserved exon within the *CDKL5* gene, referred to as exon 16b. Very interestingly, transcript analysis revealed that the exon 16b-containing mRNA isoform is specifically found in the brain, and we suggest that the molecular analysis of this new exon should now be taken into consideration for the genetic diagnosis of patients presenting with *CDKL5*-related disease. Further experiments to investigate the subcellular localization of the protein confirmed the dynamic intracellular behavior of the protein. Although the precise mechanisms of CDKL5 regulation in cells, as well as the cellular/molecular factors regulating this mechanism, are largely unknown, this work may help at understanding the pathophysiology of *CDKL5*-related disease.

ACKNOWLEDGEMENTS

We thank Dr Evelyn Bloch-Gallego and Stéphanie Backer for their technical expertise in *in situ* hybridization, and Yoann Saillour for his technical assistance with the 7500 Real-Time PCR System. This work was supported by the Institut National de la Santé et de la Recherche Médicale (ANR E-RARE EuroRETT Network).

- Montini, E., Andolfi, G., Caruso, A., Buchner, G., Walpole, S. M., Mariani, M. *et al.* Identification and characterization of a novel serine-threonine kinase gene from the Xp22 region. *Genomics* **51**, 427–433 (1998).
- Kalscheuer, V. M., Tao, J., Donnelly, A., Hollway, G., Schwinger, E., Kübart, S. *et al.* Disruption of the serine/threonine kinase 9 gene causes severe X-linked infantile spasms and mental retardation. *Am. J. Hum. Genet.* **72**, 1401–1411 (2003).
- Weaving, L. S., Christodoulou, J., Williamson, S. L., Friend, K. L., McKenzie, O. L., Archer, H. *et al.* Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *Am. J. Hum. Genet.* **75**, 1079–1093 (2004).
- Tao, J., Van Esch, H., Hagedorn-Greife, M., Hoffmann, K., Moser, B., Raynaud, M. *et al.* Mutations in the X-linked cyclin-dependent kinase-like 5 (CDKL5/STK9) gene are associated with severe neurodevelopmental retardation. *Am. J. Hum. Genet.* **75**, 1149–1154 (2004).
- Evans, J. C., Archer, H. L., Colley, J. P., Ravn, K., Nielsen, J. B., Kerr, A. *et al.* Early onset seizures and Rett-like features associated with mutations in CDKL5. *Eur. J. Hum. Genet.* **13**, 1113–1120 (2005).
- Lin, C., Franco, B. & Rosner, M. R. CDKL5/Stk9 kinase inactivation is associated with neuronal developmental disorders. *Hum. Mol. Genet.* **14**, 3775–3786 (2005).
- Scala, E., Ariani, F., Mari, F., Caselli, R., Pescucci, C., Longo, I. *et al.* CDKL5/STK9 is mutated in Rett syndrome variant with infantile spasms. *J. Med. Genet.* **42**, 103–107 (2005).
- Archer, H. L., Evans, J., Edwards, S., Colley, J., Newbury-Ecob, R., O'Callaghan, F. *et al.* CDKL5 mutations cause infantile spasms, early onset seizures, and severe mental retardation in female patients. *J. Med. Genet.* **43**, 729–734 (2006).
- Mari, F., Azimonti, S., Bertani, I., Bolognese, F., Colombo, E., Caselli, R. *et al.* CDKL5 belongs to the same molecular pathway of MeCP2 and it is responsible for the early-onset seizure variant of Rett syndrome. *Hum. Mol. Genet.* **14**, 1935–1946 (2005).
- Nectoux, J., Heron, D., Tallot, M., Chelly, J. & Bienvenu, T. Maternal origin of a novel C-terminal truncation mutation in CDKL5 causing a severe atypical form of Rett syndrome. *Clin. Genet.* **70**, 29–33 (2006).
- Van Esch, H., Jansen, A., Bauters, M., Froyen, G. & Fryns, J. P. Encephalopathy and bilateral cataract in a boy with an interstitial deletion of Xp22 comprising the CDKL5 and NHS genes. *Am. J. Med. Genet. A* **143A**, 364–369 (2007).
- Rosas-Vargas, H., Bahi-Buisson, N., Philippe, C., Nectoux, J., Girard, B., N'Guyen Morel, M. A. *et al.* Impairment of CDKL5 nuclear localisation as a cause for severe infantile encephalopathy. *J. Med. Genet.* **45**, 172–178 (2008).
- Elia, M., Falco, M., Ferri, R., Spalletta, A., Bottitta, M., Calabrese, G. *et al.* CDKL5 mutations in boys with severe encephalopathy and early-onset intractable epilepsy. *Neurology* **71**, 997–999 (2008).
- Erez, A., Patel, A. J., Wang, X., Xia, Z., Bhatt, S. S., Craigen, W. *et al.* Alu-specific microhomology-mediated deletions in CDKL5 in females with early-onset seizure disorder. *Neurogenetics* **10**, 363–369 (2009).
- Fichou, Y., Bieth, E., Bahi-Buisson, N., Nectoux, J., Girard, B., Chelly, J. *et al.* Re: CDKL5 mutations in boys with severe encephalopathy and early-onset intractable epilepsy. *Neurology* **73**, 77–78 (2009).
- Sartori, S., Di Rosa, G., Polli, R., Bettella, E., Tricomi, G., Tortorella, G. *et al.* A novel CDKL5 mutation in a 47,XXY boy with the early-onset seizure variant of Rett syndrome. *Am. J. Med. Genet. A* **149A**, 232–236 (2009).
- Russo, S., Marchi, M., Cogliati, F., Bonati, M. T., Pintaudi, M., Veneselli, E. *et al.* Novel mutations in the CDKL5 gene, predicted effects and associated phenotypes. *Neurogenetics* **10**, 241–250 (2009).
- Sprovieri, T., Conforti, F. L., Fiumara, A., Mazzei, R., Ungaro, C., Citrigno, L. *et al.* A novel mutation in the X-linked cyclin-dependent kinase-like 5 (CDKL5) gene associated with a severe Rett phenotype. *Am. J. Med. Genet. A* **149A**, 722–725 (2009).
- Saletti, V., Canafoglia, L., Cambiaso, P., Russo, S., Marchi, M. & Riva, D. A CDKL5 mutated child with precocious puberty. *Am. J. Med. Genet. A* **149A**, 1046–1051 (2009).
- Bahi-Buisson, N., Girard, B., Gautier, A., Nectoux, J., Fichou, Y., Saillour, Y. *et al.* Epileptic encephalopathy in a girl with an interstitial deletion of Xp22 comprising promoter and exon 1 of the CDKL5 gene. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **153B**, 202–207 (2010).
- Psoni, S., Willems, P. J., Kanavakis, E., Mavrou, A., Frissyr, H., Traeger-Synodinos, J. *et al.* A novel p.Arg970X mutation in the last exon of the CDKL5 gene resulting in late-onset seizure disorder. *Eur. J. Paediatr. Neurol.* **14**, 188–191 (2010).
- Bahi-Buisson, N., Nectoux, J., Rosas-Vargas, H., Milh, M., Boddaert, N., Girard, B. *et al.* Key clinical features to identify girls with CDKL5 mutations. *Brain* **131**, 2647–2661 (2008).
- Rusconi, L., Salvatoni, L., Giudici, L., Bertani, I., Kilstrup-Nielsen, C., Broccoli, V. *et al.* CDKL5 expression is modulated during neuronal development and its subcellular distribution is tightly regulated by the C-terminal tail. *J. Biol. Chem.* **283**, 30101–30111 (2008).
- Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U. & Zoghbi, H. I. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **23**, 185–188 (1999).
- Bertani, I., Rusconi, L., Bolognese, F., Forlani, G., Conca, B., De Monte, L. *et al.* Functional consequences of mutations in CDKL5, an X-linked gene involved in infantile spasms and mental retardation. *J. Biol. Chem.* **281**, 32048–32056 (2006).
- Kameshita, I., Sekiguchi, M., Hamasaki, D., Sugiyama, Y., Hatano, N., Suetake, I. *et al.* Cyclin-dependent kinase-like 5 binds and phosphorylates DNA methyltransferase 1. *Biochem. Biophys. Res. Commun.* **377**, 1162–1167 (2008).
- Desmet, F. O., Hamroun, D., Lalande, M., Collod-Bérout, G., Claustres, M. & Bérout, C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **37**, e67 (2009).
- Kriacounis, S. & Bird, A. The major form of MeCP2 has a novel N-terminus generated by alternative splicing. *Nucleic Acids Res.* **32**, 1818–1823 (2004).
- Mnatzakanian, G. N., Lohi, H., Munteanu, I., Alfred, S. E., Yamada, T., MacLeod, P. J. *et al.* (2004) A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nat. Genet.* **36**, 339–341.
- Ravn, K., Nielsen, J. B. & Schwartz, M. Mutations found within exon 1 of MECP2 in Danish patients with Rett syndrome. *Clin. Genet.* **67**, 532–533 (2005).
- Bartholdi, D., Klein, A., Weissert, M., Koenig, N., Baumer, A., Boltshauser, E. *et al.* Clinical profiles of four patients with Rett syndrome carrying a novel exon 1 mutation or genomic rearrangement in the MECP2 gene. *Clin. Genet.* **69**, 319–326 (2006).
- Chunshu, Y., Endoh, K., Soutome, M., Kawamura, R. & Kubota, T. A patient with classic Rett syndrome with a novel mutation in MECP2 exon 1. *Clin. Genet.* **70**, 530–531 (2006).
- Quenard, A., Yilmaz, S., Fontaine, H., Bienvenu, T., Moncla, A., des Portes, V. *et al.* Deleterious mutations in exon 1 of MECP2 in Rett syndrome. *Eur. J. Med. Genet.* **49**, 313–322 (2006).
- Fichou, Y., Nectoux, J., Bahi-Buisson, N., Rosas-Vargas, H., Girard, B., Chelly, J. *et al.* The first missense mutation causing Rett syndrome specifically affecting the MeCP2_e1 isoform. *Neurogenetics* **10**, 127–133 (2009).
- Saunders, C. J., Minassian, B. E., Chow, E. W., Zhao, W. & Vincent, J. B. Novel exon 1 mutations in MECP2 implicate isoform MeCP2_e1 in classical Rett syndrome. *Am. J. Med. Genet. A* **149A**, 1019–1023 (2009).
- Moore, J. D., Yang, J., Truant, R. & Kornbluth, S. Nuclear import of Cdk/cyclin complexes: identification of distinct mechanisms for import of Cdk2/cyclin E and Cdc2/cyclin B1. *J. Cell. Biol.* **144**, 213–224 (1999).
- Ericson, K. K., Krull, D., Slomiany, P. & Grossel, M. J. Expression of cyclin-dependent kinase 6, but not cyclin-dependent kinase 4, alters morphology of cultured mouse astrocytes. *Mol. Cancer Res.* **9**, 654–664 (2003).