SHORT COMMUNICATION

Molecular Characterization of a cDNA Encoding Functional Human CLK4 Kinase and Localization to Chromosome 4q35

Jörg Schultz,* Tania Jones,† Peer Bork,* Denise Sheer,† Stephanie Blencke,‡ Silvia Steyrer,‡ Ursula Wellbrock,‡ Dorian Bevec,‡ Axel Ullrich,§ and Christian Wallasch‡,[1]

‡*Axxima Pharmaceuticals AG, Am Klopferspitz 19, D-82152 Martinsried, Germany; *EMBL, Meyerhofstrasse 1, D-69012 Heidelberg, Germany; †Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, United Kingdom; and §Max-Planck-Institut für Biochemie, Am Klopferspitz 18A, D-82152 Martinsried, Germany*

Received August 31, 2000; accepted November 16, 2000

Phosphorylated serine- and arginine-rich (SR) proteins play an important role in the formation of spliceosomes, possibly controlling the regulation of alternative splicing. Enzymes that phosphorylate the SR proteins belong to the family of CDC2/CDC28-like kinases (CLK). Employing nucleotide sequence comparison of human expressed sequence tag sequences to the murine counterpart, we identified, cloned, and recombinantly expressed the human orthologue to the murine CLK4 cDNA. When fused to glutathione *S*-transferase, the catalytically active human CLK4 is able to autophosphorylate and to phosphorylate myelin basic protein, but not histone H2B as a substrate. Inspection of mRNA accumulation demonstrated gene expression in all human tissues, with the most prominent abundance in liver, kidney, brain, and heart. Using fluorescence *in situ* hybridization, the human CLK4 cDNA was localized to band q35 on chromosome 4.

Protein phosphorylation is a key mechanism in the control of many cellular processes including signal transduction or cell cycle regulation. One protein family, characterized by serine- and arginine-rich (SR) members, plays an important role in the formation of spliceosomal complexes and regulation of alternative RNA splicing (2, 7, 9). A prominent feature of these SR proteins is their high degree of phosphorylation, which seems to be necessary for spliceosome assembly (11, 19). SR proteins are enriched in nuclear speckles, nuclear structures sensitive to alterations in cellular transcription activities (3, 12). It is known that SR proteins are phosphorylated by a number of protein kinases, including U1 70-kDa activity (21), SRPK1 (6), topoisomerase I (15), lamin B-receptor kinase (14), SRPK2 (20), and CLK proteins (5, 13). Moreover, CDC2/CDC28-like (CLK) proteins contain an SR protein homologue domain and are shown to autophosphorylate (8, 10, 13). In this study we provide data on the mRNA expression pattern and chromosomal localization of human CLK4 as well as data on the catalytic activity of this new human protein kinase.

An algorithm combining BLAST (1) based searches with a domain identification protocol (16, 17) was used to search for novel human protein kinases in expressed sequence tag (EST) databases (18). Four hundred seventy-four human ESTs that belong to the protein kinase family but did not correspond to any known human protein were identified, representing more than 150 novel proteins (18). Of those, 49 had a possible orthologue in a closely related species, thus allowing prediction of the detailed function of the novel protein. These predicted proteins were searched for members of the CDC2/CDC28-like kinase family, revealing a set of ESTs (EMBL Accession Nos. AA631990, AI039778, AI004959, AI094075, AA449725, H29221, and AI355046) that represented a possible human orthologue of the murine CLK4 protein. The EST clone AA449725 was purchased from the Resource Center/Primary Database (Berlin), and its nucleotide sequence was determined. Sequence alignment to the murine CLK4 cDNA revealed a 91-nt deletion in the human EST clone. To investigate whether the purchased clone represents a cloning artifact, we used specific oligonucleotide primers and performed an RT-PCR with total human kidney RNA as reaction template (sense primer: 5′ GATC-GAATTC-ATG CGG CAT TCC AAA AGA AC; antisense primer: 5′ CAT TTT TTA CGA TTT TCA CT). After amplification, the PCR product was subcloned into the original EST vector pT7T3D-Pac (Pharmacia) using the *Eco*RI site, introduced by the 5′ sense primer before the start codon, and the internal and unique *Sph*I site of the human CLK4 cDNA. Sequence inspection of the cloned PCR

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession No. AF294429.

[1] To whom correspondence should be addressed. Telephone: +49-89-740 165-19. Fax: +49-89-740 165-20. E-mail: wallasch@axxima.com.

doi:10.1006/geno.2000.6447
0888-7543/01 $35.00

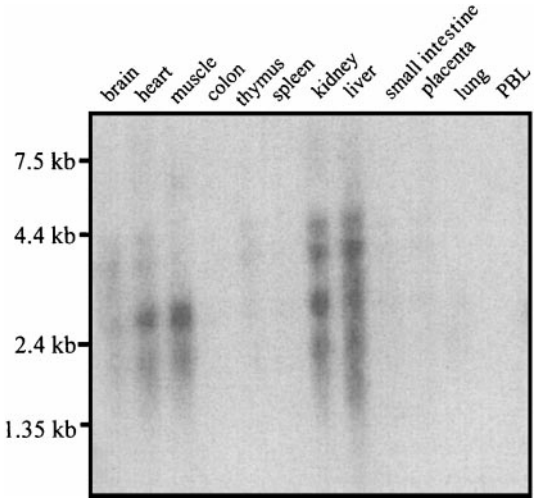

FIG. 1. Northern blot of 1 μg of poly(A)$^+$ RNA from various human tissues hybridized with the radiolabeled CLK4-specific oligonucleotide probe (CLK4: 5′-CAC AGT AGT CAT TCC TGT ATT CGT CAA C-3′). Molecular size markers in kilobases are shown on the right.

product determined the full-length human sequence of CLK4, including the missing 91 nt absent in the EST clone. This cDNA encodes the complete protein sequence (481 amino acids) and contains a 322-bp-long 3′untranslated region followed by a poly(A) tail (data not shown). The deduced protein was estimated to be 54 kDa and have a p*I* of 9.06. Comparison of the published protein sequence of the murine CLK4 (13) with the deduced amino acid sequence from the human CLK4 cDNA revealed 97% identity, strongly indicating that the human CLK4 sequence also has functional similarity to the enzyme present in the mouse.

Northern blot hybridization experiments using highly pure poly(A)$^+$ RNA from different human tissues (12-Lane MTN blot, Clontech) and human multiple tissue expression arrays (MTE, Clontech) were performed using the complete human CLK4 cDNA sequence as [α-^{32}P]ATP-labeled probe. When inspecting the human multiple tissue expression (MTE) array, which contains RNA from 68 different tissues and 8 cell lines, we detected CLK4 mRNA signals in all cases (data not shown). With the 12-Lane MTN blot, the strongest signals were obtained from liver, kidney, heart, muscle, and brain samples (Fig. 1). Furthermore, accumulation of four different mRNA species, ranging between ~2.0 and ~4.5 kb, was detected. Interestingly, in kidney and in liver samples, all four types are almost evenly accumulated, whereas in heart and muscle samples, the most prominent accumulation was detected for the ~2.8-kb CLK4 mRNA species.

To determine the chromosomal localization of human CLK4, the full-length cDNA was used as probe for FISH to human chromosome spreads as described (4). As shown in Fig. 2, specific hybridization signal was clearly detected on band q35 of chromosome 4.

To investigate whether human CLK4 is a fully active kinase CLK4, wildtype (wt) cDNA was subcloned (*Eco*RI/*Xho*I sites) into the pGEX-5X1 vector (Pharmacia), generating an in-frame glutathione *S*-transferase (GST)–CLK4-wt fusion construct. In addition, using the QuickChange Site-Directed Mutagenesis Kit (Stratagene), a dominant-negative mutant of the GST–CLK4-wt (GST–CLK4-K189R) was constructed by mutating the lysine at amino acid position 189 into an arginine. This mutation should ultimately lead to the generation of a catalytically inactive, negative form of the wildtype kinase as shown previously for murine CLK4 and human CLK2 (10, 13). After sequence confirmation, the plasmids pGEX5X1 CLK4wt and pGEX-5X1 CLK4 K189R were transformed into *Escherichia coli* strain DH5α, and recombinant human CLK4 molecules were produced and purified according to the purification protocol for pGEX fusion proteins by Pharmacia. Equal amounts of the fusion proteins were used in an *in vitro* kinase assay. The kinase assays were carried out for 30 min at 30°C in a volume of 50 μl {1 μg of fusion protein, 5 μl of 10× kinase buffer (200 mM Hepes, pH 7.5; 10 mM $MgCl_2$; 10 mM DTT; 2 mM sodium orthovanadate), 10 μl of 50 μM ATP, 5 μl of the appropriate substrate [5 mg/ml MBP; 5 mg/ml histone 2B (H2B)], and H_2O to a final volume of 50 μl, complemented with 1 μCi of [γ-^{32}P]ATP per sample}. We clearly demonstrate phosphorylation of myelin basic protein (MBP) as test substrate for GST–CLK4-wt protein. In sharp contrast, however, when H2B was used as a potential substrate, no phosphorylation signals were detected. In both cases, the mutant form GST–CLK4-K189R was catalytically inactive, confirming the relevance of the lysine 189 for the protein′s catalytic activity. In addition, autophosphorylation of the GST–CLK4-wt fusion protein (having a calculated molecular mass of ~83 kDa), was observed (Fig. 3). The signal seen at lower molecular mass (~47 kDa) most probably originates from autophosphorylation of a cleavage product of the GST–CLK4-wt fusion protein. As ex-

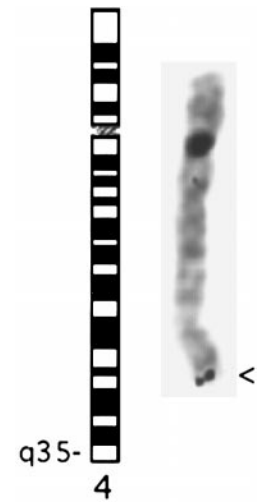

FIG. 2. FISH demonstrating localization of CLK4 cDNA probe to human chromosome band 4q35. FITC signal on G-banded chromosome is shown, and the position of the signal is shown on the corresponding idiogram.

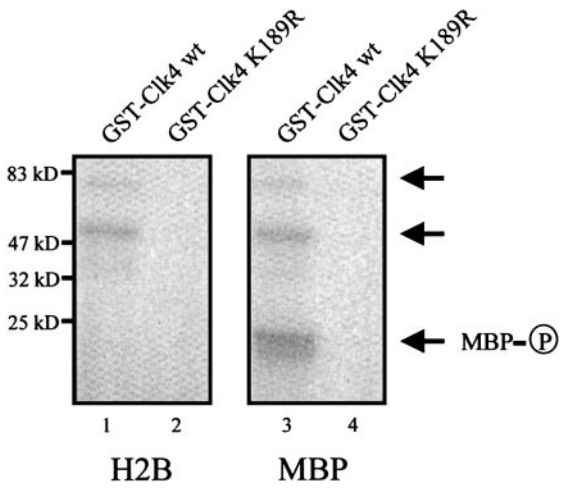

FIG. 3. *In vitro* phosphorylation assay. Autophosphorylation of the GST–CLK4-wt protein and phosphorylation of MBP but not H2B substrate can be detected. The GST–CLK4-K189R mutant is inactive. Molecular mass markers in kilodaltons are shown on the right.

pected, the fusion protein harboring the predicted catalytically inactive form CLK4-K189R did not autophosphorylate.

In summary, we describe the molecular cloning, recombinant expression, catalytic activity, and chromosomal localization of an additional CLK protein, human CLK4, which is by sequence comparison the human orthologue of the murine CLK4 protein. The knowledge of the chromosomal localization of the human CLK4 gene provides a valuable genetic marker for further studies and understanding of the regulation of cellular splicing reactions.

REFERENCES

1. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25:** 3389–3402.
2. Blencowe, B. J., Issner, R., Kim, J., McCaw, P., and Sharp, P. A. (1995). New proteins related to the Ser-Arg family of splicing factors. *RNA* **8:** 852–865.
3. Bregman, D. B., Du, L., van der Zee, S., and Warren, S. L. (1995). Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. *J. Cell. Biol.* **129:** 287–298.
4. Cherif, D., Julier, C., Delattre, O., Derre, J., Lathrop, G. M., and Berger, R. (1990). Simultaneous localization of cosmids and chromosome R-banding by fluorescence microscopy: Application to regional mapping of human chromosome 11. *Proc. Natl. Acad. Sci. USA* **87:** 6639–6643.
5. Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J. L., Bell, J. C., and Duncan, P. I. (1996). The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *EMBO J.* **15:** 265–275.
6. Gui, J. F., Lane, W. S., and Fu, X. D. (1994). A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature* **369:** 678–682.
7. Harada, K., Martin, S. S., and Frankel, A. D. (1996). Selection of RNA-binding peptides *in vivo*. *Nature* **380:** 175–179.
8. Howell, B. W., Afar, D. E., Lew, J., Douville, E. M., Icely, P. L., Gray, D. A., and Bell, J. C. (1991). STY, a tyrosine-phosphorylating enzyme with sequence homology to serine/threonine kinases. *Mol. Cell. Biol.* **11:** 568–572.
9. Krämer, A. (1996). The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu. Rev. Biochem.* **65:** 367–409.
10. Lee, K., Du, C., Horn, M., and Rabinow, L. (1996). Activity and autophosphorylation of LAMMER protein kinases. *J. Biol. Chem.* **271:** 27299–27303.
11. Mermoud, J. E., Cohen, P. T., and Lamond, A. I. (1994). Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. *EMBO J.* **13:** 5679–5688.
12. Misteli, T., Caceres, J. F., Clement, J. Q., Krainer, A. R., Wilkinson, M. F., and Spector, D. L. (1998). Serine phosphorylation of SR proteins is required for their recruitment to sites of transcription *in vivo*. *J. Cell. Biol.* **143:** 297–307.
13. Nayler, O., Stamm, S., and Ullrich, A. (1997). Characterization and comparison of four serine- and arginine-rich (SR) protein kinases. *Biochem. J.* **326:** 693–700.
14. Nikolakaki, E., Simos, G., Georgatos, S. D., and Giannakouros, T. (1996). A nuclear envelope-associated kinase phosphorylates arginine–serine motifs and modulates interactions between the lamin B receptor and other nuclear proteins. *J. Biol. Chem.* **271:** 8365–8372.
15. Rossi, F., Labourier, E., Forne, T., Divita, G., Derancourt, J., Riou, J. F., Antoine, E., Cathala, G., Brunel, C., and Tazi, J. (1996). Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. *Nature* **381:** 80–82.
16. Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998). SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc. Natl. Acad. Sci. USA* **95:** 5857–5864.
17. Schultz, J., Copley, R. R., Doerks, T., Ponting, C. P., and Bork, P. (2000). SMART: A web-based tool for the study of genetically mobile domains. *Nucleic Acids Res.* **28:** 231–234.
18. Schultz, J., Doerks, T., Ponting, C. P., Copley, R. R., and Bork, P. (2000). More than 1,000 putative new human signalling proteins revealed by EST data mining. *Nat. Genet.* **25:** 201–204.
19. Tazi, J., Kornstadt, U., Rossi, F., Jeanteur, P., Cathala, G., Brunel, C., and Luhrmann, R. (1993). Thiophosphorylation of U1-70K protein inhibits pre-mRNA splicing. *Nature* **363:** 283–286.
20. Wang, H. Y., Lin, W., Dyck, J. A., Yeakley, J. M., Songyang, Z., Cantley, L. C., and Fu, X. D. (1998). SRPK2: A differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells. *J. Cell. Biol.* **140:** 737–750.
21. Woppmann, A., Will, C. L., Kornstadt, U., Zuo, P., Manley, J. L., and Luhrmann, R. (1993). Identification of an snRNP-associated kinase activity that phosphorylates arginine/serine rich domains typical of splicing factors. *Nucleic Acids Res.* **21:** 2815–2822.