FEBS Letters 375 (1995) 50–52 FEBS 16225

Genomic structure and alternative splicing of the murine *bhk/ctk/ntk* gene

M.A. Ershler[a,b,*], I.M. Samokhvalov[a], A.V. Belyavsky[a,b], J.W.M. Visser[b]

[a]*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov str. 32, 117984, Moscow B-334, Russian Federation*
[b]*New York Blood Center, 310 East 67th Str., New York, NY, USA*

Received 15 September 1995

Abstract Recently, we and others have cloned cDNAs encoding a second member of the Csk family of inhibitory protein kinases, which we termed Bhk [M.A. Ershler et al. (1994) Dokl. Akad. Nauk. 339, 679–683]. In the present study, two new distinct types of *bhk* mRNA were found in addition to the third form described previously. Analysis of the *bhk* genomic structure established that three exons participate in the alternative splicing of *bhk* mRNA.

Key words: Tyrosine protein kinase; Alternative splicing; Genomic structure

1. Introduction

Protein tyrosine kinases (PTKs) play an essential role in intracellular and extracellular signal transduction, which regulates cell proliferation and differentiation. A variety of nonreceptor PTKs has been described to participate in intracellular signal transduction [1]. The most extensively characterized nonreceptor PTKs belong to the Src-superfamily. The activity of Src-related enzymes is repressed by phosphorylation of a conserved carboxyl-terminal tyrosine residue, an event mediated by a cytoplasmic protein kinase termed Csk [2,3]. Csk has no autophosporylation site within its kinase domain and is further distinguished from the Src-family of kinases by the lack of a conserved carboxyl-terminal tyrosine residue and myristoylation site. We and others recently described cDNAs encoding Csk-related PTK [4,5,6,7,8], which was termed variously as Bhk, Ctk or Ntk. *bhk/ctk/ntk* (hereafter termed *bhk*) mRNA was found to be expressed predominantly in brain but also in hematopoietic tissues: thymus, spleen and bone marrow.

In this report we demonstrate the existence of alternatively spliced variants of the *bhk* RNA, and characterize the genomic structure of the gene.

2. Materials and methods

2.1. Cloning of 5′ region of bhk mRNA

The 'rapid amplification of cDNA ends' (RACE) procedure was performed as described in [9]. Brain poly(A)$^+$ RNA was reverse transcribed with Superscript reverse transcriptase (Gibco BRL, Gaithersburg, MD) using random hexamer primers. The product was tailed with dA by terminal transferase (Boehringer Mannheim, Indianapolis, IN) and amplified using the T13 primer (5′-GGGAGGCCCTTTTTTTT-TTTTT) and an antisense primer (5′-CTTGGGGCGAGAGTTC-TCA) corresponding to the 187–205 nucleotide position of the *bhk* clone [4], size selected on agarose gel and cloned into pcDNA II vector (Invitrogene, CA). Positive clones identified by hybridization with the 5′ *bhk* probe were sequenced by the termination method using Sequenase (USB, Cleveland, OH).

2.2. Isolation and mapping of genomic clones

bhk genomic clones were isolated from a mouse (129SVJ strain) λFIXII genomic library (Stratagene, San Diego, CA) using a radiolabelled 1.66 kb Bhk cDNA probe [4]. 10^6 recombinant phages were screened, 16 positive clones identified, and two of them plaque purified. Restriction mapping analysis of the λ recombinant clones was performed by a partial endonuclease digestion combined with indirect end labeling [10] and by complete digestion with several restriction endonu-

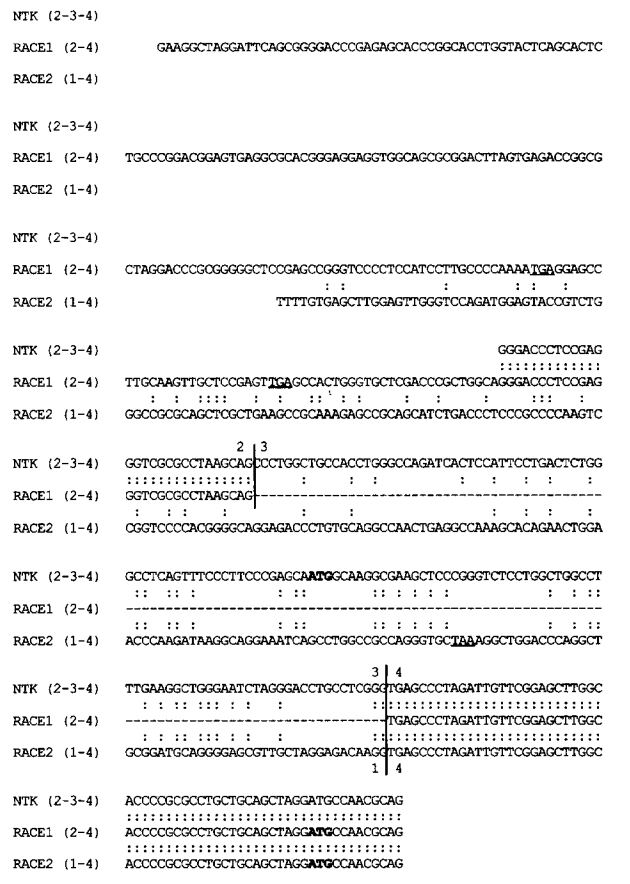

Fig. 1. Nucleotide sequences of alternatively spliced 5′ region of *bhk* mRNA. Exons participating in the formation of the splice variants are shown in parentheses. Dashed lines designate gaps added for optimal alignment. Solid lines indicate the boundary between exons, and numbers indicate the involved exons. The putative translation initiation codons (ATG) are in bold typeface, while the upstream in-frame stop codons for each type of *bhk* cDNA are underlined. Identical residues are marked by double dots. The first 228 bases of *ntk* are shown according to [6].

*Corresponding author at 'a'. Fax: (7) (095) 135-1405.
E-mail: bel@imb.msk.su

Abbreviations: bp, base pair(s); kb, kilobase(s); PTK, protein tyrosine kinase; RACE, rapid amplification of cDNA ends; UTR, 5′ untranslated region.

The nucleotide sequences presented here have been submitted to the EMBL/GenBank database under accession numbers X89723, X89724.

SSDI 0014-5793(95)01154-4

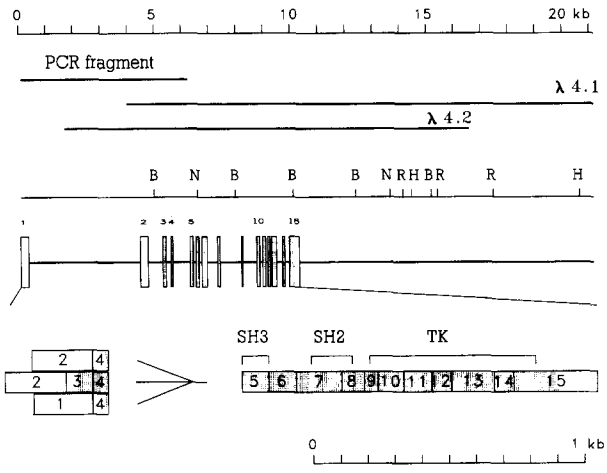

Fig. 2. Genomic and cDNA structure of *bhk*. At the top, the scale is given for genomic DNA, at the bottom for cDNA. Alignments of the lambda genomic clones and the PCR fragment generated from genomic DNA are given in the upper portion. A restriction enzyme map is shown in the middle (B = *Bam*HI; N = *Not*I; H = *Hin*dIII; R = *Eco*RI). Exon/intron structure is indicated in the lower portion; coding regions are shaded; arabic numerals indicate exon numbers. The structure of the alternatively spliced mRNAs is shown at the bottom.

cleases. The exon/intron structure was determined by PCR with primers from different regions of the *bhk* cDNA sequence, followed by size determination by agarose electrophoresis and sequencing of PCR fragments with the aid of the SequiTherm sequencing kit (Epicentre Technologies, Madison, WI).

Amplification of a 6.2 kb PCR fragment, which contains exons 1 to 4, from genomic DNA as template, was performed using the XL-PCR kit (Perkin-Elmer, Branchburg, NJ).

3. Results and discussion

The *bhk* cDNA clone isolated previously had an open reading frame with no stop codons in the 5′ untranslated region (5′-UTR) raising the possibility of the existence of additional coding sequences in the uncloned part of the mRNA. This 5′ region was cloned using a RACE procedure. Twenty-one 5′ region-derived sequences were found to form two groups. Ten sequences aligned in the group RACE1, and eleven in RACE2 (Fig. 1). Neither has additional ATG codons in frame with the previously obtained sequence, but do contain in-frame stop codons (Fig. 1). These sequences encode the same putative polypeptide product and differ only in their 5′ untranslated sequences. The published cDNA sequence of *ntk* [6] is almost completely identical to *bhk* cDNA, but contains a 134 bp insert in the 5′ region. Thus there are at least three variants of *bhk* mRNA which diverge from the same point. These three variants encode two putative peptides. All three sequences were found to be expressed in brain and thymus by RT-PCR procedure (data not shown).

Several tyrosine kinase mRNAs, for example *lck* [11] and *igfr-II* [12], are known to have alternatively spliced 5′-UTR, although the function of these regions is not clear. Possible functions of the 5′ UTR include the modulation of mRNA

Table 1
Exon/intron lengths and boundaries of the mouse *bhk* gene

Intron	Exon					Intron		
Splice acceptor			size bp			Splice donor	Size kb	
		No.	mBHK	hCSK			mBHK	hCSK
		1	>252	>68	GGAGACAAG	**gt**gagtngg	4.2	>6.4
		2	>257	–	CCTAAGCAG	**gt**gagcgtg	0.6	–
					L P R			
cacacac**ag**	C CCT GGC TGC	3	134	–	CTG CCT CGG	**gt**aatgatc	0.17	–
	[V S P]				M P T			
ccttcac**ag**	GTG AGC CCT	4	60	80	ATG CCA ACG	**gt**gagtgtg	0.75	0.3
	Q R W				A C E			
ccttccc**ag**	CAG CGC TGG	5	**114**	**114**	GCC TGT GAG	**gt**gagaggg	0.05	0.1
	D K S				L M P			
cctgtac**ag**	GAC AAG AGC	6	116	113	CTC ATG CC	**gt**gagtngc	0.08	0.33
	W F H				M V E			
atcctac**ag**	A TGG TTT CAT	7	**220**	**220**	ATG GTG GAG	**gt**gacgtgt	0.35	0.92
	H Y T				L A K			
cctccac**ag**	CAC TAC ACC	8	**94**	**94**	CTC GCC AAG G	**gt**atgagag	0.7	0.18
	A G W				E F G			
tctctgc**ag**	CT GGC TGG	9	**66**	**66**	GAG TTT GGA G	**gt**gaggagg	0.5	0.08
	A V L				V M T			
cctccac**ag**	CC GTC CTA	10	**100**	**100**	GTG ATG AC	**gt**gagtgtt	0.08	0.09
	K L Q				V S K			
accccc**ag**	G AAG CTG CAG	11	85	91	GTG AGC AAG	**gt**gtgcagg	0.08	0.42
	G N L				F A L			
cttaaac**ag**	GGC AAC CTG	12	**74**	**74**	TTT GCT CT	**gt**aagtgac	0.09	0.1
	H V A				K N G			
tgtctct**ag**	T CAT GTT GCT	13	**196**	**196**	AAA AAC GGG	**gt**gagcagc	0.28	0.1
	R F S				P K M			
gcccaac**ag**	CGG TTC TCC	14	**87**	**87**	CCC AAG ATG	**gt**ggtgagc	0.09	0.25
	S L K							
tgtccac**ag**	TCG CTA AAG	15	>355	866				

Sequences of exons are shown in uppercase letters, introns in lower case. The deduced amino acid sequence of *Bhk* is given above the nucleotide sequence. The amino acids of exon 4 are in brackets because only one of the alternatively spliced variants can be translated. The intron sequences adjacent to the exon and the length of exons which coincide in *bhk* and *csk* are in bold letters.

degradation [13] and regulation of translation efficiency [14]. It is therefore possible that the 5′ UTR of the *bhk* RNA can be involved in the regulation of its translatability and/or stability. It has also been reported that the Lck PTK associates with the cytoplasmic tail of CD4 and CD8 through its 32 amino-terminal amino acid residues [15]. By analogy, the alternative amino-terminal regions of the Bhk protein may play a role in anchoring to membrane proteins or in substrate specificity of Bhk.

To further analyze the mechanism of formation of different *bhk* RNA forms, we isolated the genomic copy of *bhk*. The corresponding restriction map and exon/intron structure is shown in Fig. 2. The gene consists of 13 coding exons and 2 noncoding exons distributed over 10.5 kb. The sizes of the exons range from 60 bp to 355 bp (Table 1). The coding exons are distributed on 5.0 kb. The catalytic domain of *Bhk* is encoded by seven exons, 9–15, while the SH3 domain is encoded by a single exon 5, and the SH2 domain is encoded by two exons, 6 and 7. The ATG codons, initiating translation in different alternatively spliced variants, are located in exons 3 and 4 (see Fig. 1). The sequence of the exon/intron splice junctions conform to the consensus for donor and acceptor site [16] (Table 1).

To gain insight into the evolution of the tyrosine kinase gene family, the genomic structure of *bhk* was compared with that of *csk* [17], which belongs to the same subclass of tyrosine kinases. A striking similarity in the exon/intron structure between the two genes was found. Out of 15 exons in *bhk* only two exons, 2 and 3, have no analogs in *csk* (Table 1). The first three exons of the *bhk* gene form three alternatively spliced cDNAs that encode two putative peptides, whereas alternative splicing has not been demonstrated for *csk*. Of the remaining 13 exons, three exons contain untranslated regions, 1, 4 and 15 (see Fig. 2). Ten exons have exactly the same boundaries; of these ten exons, eight have exactly the same size in both genes and the other two coding exons, 6 and 11, differ in length by exact multiples of three nucleotides (Table 1). Analysis of the genomic organization of *bhk* strongly supports the assumption that Bhk and Csk belong to the same class of PTKs and may arise from duplication of the same ancestor gene.

In conclusion, we established the complete genomic structure of the murine *bhk* gene. A partial genomic structure of *bhk* was reported by earlier by Hamaguchi et al. [18]. Their map lacks exons 1 and 3 in our classification. When this manuscript was prepared for publication, existence of alternative forms of *bhk* mRNA was reported by Kaneko et al. [19].

Acknowledgements: We thank Dmitrii Ovchinnikov for help in the cloning of the genomic copy, Anna Krotkova for skillful technical assistance and Dr. Gertrude Pfaffenbach for critical reading of the manuscript. This study was supported by Grant 07-13-065 from the Netherlands Organization for the Advancement of Pure Research (NWO) and INTAS Grant 93-3338.

References

[1] Bolen, J.B. (1993) Oncogene 8, 2025–2031.
[2] Cooper, J.A. and Howell, B. (1993) Cell 73, 1051–1054.
[3] Nada, S., Okada, M., MacAukey, A., Cooper, J.A. and Nakagawa, H. (1991) Nature 351, 69–72.
[4] Ershler M.A., Krivtsov A.V., Krotkova A.V., Belyavsky, A.V. and Visser J.W.M. (1994) Dokl. Akad. Nauk. 339, 679–683.
[5] Klages, S., Adam, D., Class, K., Fargnoli, J., Bolen, J.B. and Penhallow, R.C. (1994) Proc. Natl. Acad. Sci. USA 91, 2597–2601.
[6] Chow, L.M., Jarvis, C., Hu, Q., Nye, S.H., Gervais, F.G., Veillette, A. and Matis, L.A. (1994) Proc. Natl. Acad. Sci. USA 91, 4975–4979.
[7] Sakano, S., Iwama, A., Inazawa, J., Ariyama, T., Ohno, M. and Suda, T. (1994) Oncogene 9,1155–1161.
[8] Bennett, B.D., Cowley, S., Jiang,S., London, R., Deng, B., Grabarek, J., Groopman, J.E., Goeddel, D.V. and Avraham, H. (1994) J. Biol. Chem. 269, 1068–1074.
[9] Harvey, R.J. and Darlison, M.G. (1991) Nucleic Acids Res. 19, 4002.
[10] Rackwitz, H.R., Zehetner, G., Frischauf, A.M. and Lehrach, H. (1984) Gene 30, 195–200
[11] Garvin, A.M., Pawar, S., Marth, J.D. and Perlmutter, R.M. (1988) Mol. Cell. Biol. 8, 3058–3064.
[12] Nielsen, F.C. (1992) Prog. Growth Factor Res. 4, 257–290.
[13] Roy, N., Laflamme, G. and Raymond, V. (1992) Nucleic Acids Res. 20, 5753–5762.
[14] Gabrielsson, B.G., Carmignac, D.F., Flavell, D.M. and Robinson, I.C. (1995) Endocrinology 136, 209–217.
[15] Turner, J.M., Brodsky, M.H., Irving, B.A., Levin, S.D., Perlmutter, R.M. and Littman, D.R. (1990) Cell 60, 755–765.
[16] Mount, S.M. (1982) Nucleic Acids Res. 10, 459–472.
[17] Brauninger, A., Karn, T., Strebhardt, K. and Rubsamen-Waigmann, H. (1993) Oncogene 8, 1365–1369.
[18] Hamaguchi, I., Iwama, A., Yamaguchi, N., Sakano, S., Matsuda, Y. and Suda, T. (1994) Oncogene 9, 3371–3374.
[19] Koneko, Y., Nonoguchi, K., Fukuyama, H., Takano, S., Higashitsuji, H., Nishiyama, H., Takenawa, J., Nakayama, H. and Fujita, J. (1995) Oncogene 10, 945–952.