SHORT COMMUNICATION

Akiko Shibui · Takeshi Tsunoda · Naohiko Seki
Yutaka Suzuki · Sumio Sugano · Kazuo Sugane

Cloning, expression analysis, and chromosomal localization of a novel butyrophilin-like receptor

Received: January 22, 1999 / Accepted: March 16, 1999

Abstract We isolated a cDNA clone which shows a similarity with human butyrophilin from a human colon mucosa cDNA library. The cDNA is 1964 bases long, with one open reading frame encoding a protein of 433 amino acids. The deduced amino acid sequence shows an overall homology of 36.5% with the human butyrophilin protein. This gene is mainly expressed in small intestine, colon, testis, and leukocytes. The chromosomal location of the gene was determined on the chromosome 5q35 region by polymerase chain reaction-based analysis with both a human/rodent monochromosomal hybrid cell panel and a radiation hybrid mapping panel.

Key words Butyrophilin · rfp domain · Colon · Chromosome 5q35 · RT-PCR · cDNA library

Introduction

Butyrophilin is one of the major proteins found in the membrane layer surrounding the milk fat-secreting droplet of mammary epithelial cells (Mather and Jack 1993). While it has been studied extensively in various animals (Heid et al. 1983; Mather and Jack 1993), the function of butyrophilin remains unknown. Human butyrophilin has the general structure of a receptor consisting of two extracellular immunoglobulin (Ig) superfamily domains, a transmembrane domain and a conserved intracellular domain (Taylor et al.

1996). The first N-terminal domain is homologous to myelin/oligodendrocyte glycoprotein (Hilton et al. 1995) and the Ig variable domains of the MHC chicken B blood group system (B-G) antigens (Pham-Dinh et al. 1993; Vernet et al. 1993). The second extracellular domain is homologous to immunoglobulin constant (IgC) domains, particularly IgC1-type domains (Williams and Barelay 1988). The conserved C-terminal domain is called the rfp domain (The International FMF Consortium 1997). Here, we report the isolation of a novel cDNA, with significant homology to human butyrophilin, from a human colon mucosa cDNA library.

Results and discussion

A full length-enriched cDNA library was constructed, using the oligo-capping method (Maruyama and Sugano 1994; Suzuki et al. 1997), with mRNA isolated from the normal mucosa of colon. The sequences of the 5' end of the cDNA clones from the library were determined for 3374 clones. One clone, ColF4100, coding a novel butyrophilin-like receptor, was isolated from this library. We determined the entire sequence of this clone. The nucleotide sequence data reported here will appear in the DDBJ, EMBL, and GenBank nucleotide sequence databases with accession number AB020625.

The cDNA was 1964bp in length and had one open reading frame encoding a protein of 433 amino acids. Hydrophobicity analysis of the deduced amino acid sequence with the SOSUI system (http://www.tuat.ac.jp/~mitaku/adv_sosui/) (Hirokawa et al. 1988) suggests that ColF4100 codes for a membrane protein that exhibits at least one transmembrane domain. This domain is located at amino acid residues 166–188. We also analyzed the hydrophobicity of this amino acid sequence according to Kyte and Doolittle (1982) and the same region was suggested to be highly hydrophobic.

A homology search of the conceptual translated amino acid sequence revealed that it was homologous to several

A. Shibui · T. Tsunoda · Y. Suzuki · S. Sugano (✉)
Department of Virology, the Institute of Medical Science, University of Tokyo, Shirokane-dai, Minato-ku, Tokyo 108-8639, Japan
Tel. +81-3-5449-5286; Fax +81-3-5449-5416
e-mail: ssugano@ims.u-tokyo.ac.jp

A. Shibui · K. Sugane
Department of Parasitology, Shinshu University School of Medicine, Matsumoto, Japan

N. Seki
Genome Research Group, National Institute of Radiological Science, Chiba, Japan

ColF4100 1:MA-----FVLIIVLSFYELVSGOWVTGPGKFVQALVGEDAVFSCSLFPETS 47
 BUTY_HUMAN 1:MAVFPSSGLPRC---LLTLILQLPKLDSAPFDVIGPPEFILAVVGEDAELPCRLSPNAS 57
 RFP_HUMAN 1:MA-----SGSVAECLQOETTCPVCLQYFAEPMMLDCGHNICCAQLARCWGT 46
 RO52_HUMAN 1:MA-----SAARLTMMWEEVTCPICLDPFVEFVSIECGHSFCQECISQVGKG 46
 MOG_HUMAN 1:MASLSRPSLPSCLCSFLTLULLQVSSSYACQFRVIGPRHPIRALVGDEVELPCRISPGKN 60

ColF4100 48:AEAMEVRFRFRNOFHAVVHLYRDGEDWESKOMPQYRGRTFVKDSIAGGR---VSLRLKNI 104
 BUTY_HUMAN 58:AEHLELRWFRKKVSPAFLVHRDGREQEAQOMPEYRGRATLVQDGLAKGR---VALRIRGV 114
 RFP_HUMAN 47:AEITNVSCPQCRETFFQRHMRPN-R-HLANVTQLVKQLRTERPSGPGGEMGVCEKHREPLK 104
 RO52_HUMAN 47:GGS--VCPVCRQRFLLKNLRPN-R-QLANMVNNLKEISQEAREGTQGER--CAVHGERLH 100
 MOG_HUMAN 61:ATGMEVGWYRPPFSRVVHLYRNGKQDGDCAPEYRGRTELLKDAIGEGK---VTLRIRNV 117

ColF4100 105:TPSD-IGLYCCWFSSQIYDEEATWELRVAALGSLPLISIVGYVDCGIQLLCLSSFQ-PSF 162
 BUTY_HUMAN 115:RVSD-DGEYTCFFREDGSEYEAALVHLKVAALGSDPHISMQVOENCEICLECTSVGWYPEP 173
 RFP_HUMAN 105:LYCEEDQMPICVVCDSR--REHRGHSVLPLEEAVEGFKQIQNQLDHLKRVKDLKKRRRA 162
 RO52_HUMAN 101:LFCEKDGKALCWVCAQS--RKHRDHAMVPLEEAAQEQYQEKLOVALGELRRKQELAEKLEV 158
 MOG_HUMAN 118:RFSD-EGFTTCFFRDHSYQEEAAMELKVED----PFY----WVSPGVLVL-LAV----LP 163

ColF4100 163:---WRLASITLGLLCG-----ALCGVVMGMLIVFFKSKGK---IQAELG----- 200
 BUTY_HUMAN 174:QVQWRTSKGEKFPSTSES RNPE-EGIFTVAASVITRDTSTKNVSCYIQNLLLGQEKKVE 232
 RFP_HUMAN 163:QGEQARAEELISLTQMEREKIVWEFEQLYHSLKEHEYRLARLEELDLAGYN SIN-GAITQ 221
 RO52_HUMAN 159:EIAIKRADWKKTVETQKSRIHAEFVQOKNFLVEEEQROLQELEKDEREQLRILG-EKEAK 217
 MOG_HUMAN 164:---VLILQITVGLVF-----LC-----LQYRLRGK---LRAEIE----- 191

ColF4100 201:MCH-VLSLPHMVLPGPSLIHSLSLWTTLAAGWTGSTDWRRKHG-----QA-ELR----- 247
 BUTY_HUMAN 233:ISIPASSLPRLTPWIVAVAVILMVLGLLTICGSIF-FTWRLYNERPRERN-EFSSKERLL 290
 RFP_HUMAN 222:FSCNISHLSLIAQLEEKQQOPTRELLQDIDTLGRAERIRIPEPWITPP-DLQEKIHIF 280
 RO52_HUMAN 218:LAQQSQALQELISELDRRCSSAELLQEVIIVLERSSEWNLDLDITSP-ELRSVCHVP 276
 MOG_HUMAN 192:NLE-RTFDPHFLRVPCWKITLFIIVPVLGPLVALIICYNWLHRRLAGQFLELRNPFF* 247

ColF4100 248:DARK-----HAVEVTLDPETAHPKLCVS-DLITVITH 277
 BUTY_HUMAN 291:EELWKKAT-----LHAVDVTLDPDTAHPHFLYEYSKSVRL 327
 RFP_HUMAN 281:AQKCLFLTESLKQFTEKMQSDMEKIQELREAQLYSVDVTLDPDTAHPSLILSDNLQVRY 340
 RO52_HUMAN 277:GLKMLRT-----C-----AVHITLDPDTANFWLILSEDRQVRL 311

ColF4100 278:RKAPQEVPHSEKRF-TRKSVVASCGFQAGKHYWEVDVGQNVGWYVGVCRDDVDRCNNVT 336
 BUTY_HUMAN 328:EDSRQKLPEKTERFDSWPCVLGREHTSGRHYWEVEVGDRTDWAIQVCRENVMK-KGFDP 386
 RFP_HUMAN 341:SYLQQLDPDNPERFNLFPCLGSPCHAGRHYWEVEVGDKAKWTIGVCEDSVCR-KGGVT 399
 RO52_HUMAN 312:GDTOQSIPGNEERFDSYPMVLGAQEFHSKHYWEVDVTGKEAWDLGVCRDVRR-KGHFL 370

ColF4100 337:LSPNNGYWVLRITTEHLVFTFNPHFISLPPSTPPTRVGVFLDYEGCTISFNTNDQ-SLI 395
 BUTY_HUMAN 387:MTPENGFWAVELY-GNGYWALTPLRTPPLAGPPRRVGIFLDYESCDISFYNMNDG-SDI 444
 RFP_HUMAN 400:SAPONGFWAVSLWYGKEYWALTSPMTALPLRTPLORVGIFLDYDAEVSFYNVTER-CHT 458
 RO52_HUMAN 371:LSSKSCFWTIWLNKQKYEAGTYPQTPHLQVPPCQVGIFLDYEAQMVSYNITDHGSLI 430

ColF4100 396:YTLITCQFEGLLRPYIQHAMYDEEKG-TPIFICPVSWG* 433
 BUTY_HUMAN 445:YTFSNVTFSCPLRPFFCLWSS--GK--KPLTICPIADGPERVTIVIANAQDLSKEIPLSPM 500
 RFP_HUMAN 459:FTFSHATFCGFPVRPYFSLSYS-GGKSAAPLITCPMSGIDGFSGHVGNHGHSMETSP* 513
 RO52_HUMAN 431:VSFSECAFTICPLRPFFSPGFNDGGKNTAPLITLCPLN-----IGSQGSTDY* 475

BUTY_HUMAN 501:GEESAPRDADTLHSLIPTQPSQGAP* 526

proteins. The multiple alignments of amino acid sequences of the ColF4100 protein, human butyrophilin precursor (BUTY) (Taylor et al. 1996), human Zinc-finger protein RFP (RFP) (Takahashi et al. 1988), human 52-kD RO protein (RO52) (Itoh et al. 1991), and human myelin oligodendrocyte glycoprotein precursor (MOG) (Hilton et al. 1995) are shown in Fig. 1. The amino acid sequence of ColF4100 was 41.7% identical to human RO52 and 40.4% identical to human RFP. Both proteins had high homology at the C-terminal domain, which is called the rfp domain. This domain is evolutionarily conserved, but its function is poorly understood. Human MOG had 36.1% identity with the extracellular immunoglobulin (Ig) superfamily domains. Human BUTY had 36.5% identity, but it was most homologous to ColF4100 because its identity was shown on the entire amino acid sequence.

The expression level of this gene was hardly detectable by Northern blot analysis. Thus, we examined the tissue distribution of the transcript in various human adult tissues by reverse transcription-coupled polymerase chain reaction (RT-PCR). Primers used for RT-PCR were to amplify the 440bp of the 3'-untranslated region of the cDNA (Fig. 2). The *ColF4100* gene was mainly expressed in small intestine, colon, testis, and leukocytes. A low but significant band was detected in spleen. We also examined RT-PCR in heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas, but no significant band was detected (data not shown).

Chromosomal assignment of the *ColF4100* gene was done by PCR analysis of a human/rodent somatic cell hybrid panel and a radiation hybrid panel as described previously (Saito et al. 1997). The specific amplified PCR primers were designed at the 3'-untranslated region of the gene (5'-CCAGACACAGCCAA GGGAGAG-3', 5'-GTCTATTCCGTCCCCACTTCA-3'; PCR product size was 237 bp). First, a specific amplified product for humans was detected only from the hybrid containing human chromosome 5 (data not shown). Then, we performed further mapping analysis, using a PCR-based radiation hybrid panel (Genebridge 4; Research Genetics, Huntsville, AL) with the same primers as those used in the assay for the human/rodent somatic cell hybrid panel. Statistical analysis of the radiation hybrid data was performed using the RHMAPPER software package (<http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl>). The data vector for the gene was 1000000100 0010000000 1100000000 1001010100 1110010000 0001000001 0110000010 0011001100 1111010000 001 and the consequent report indicated that the gene was placed to 13.58 cR distal from WI-6737, which was cytogenetically mapped to the 5q35 region.

We cloned the *ColF4100* gene, a novel human gene

Fig. 1. Multiple alignment of ColF4100 (accession number AB020625), human butyrophilin precursor (*BUTY_HUMAN*; accession number U39576), human Zinc-finger protein RFP (*RFP_HUMAN*; accession number J03407), human 52-kD RO protein (*RO52_HUMAN*; accession number M34551), and human myelin-oligodendrocyte glycoprotein precursor (*MOG_HUMAN*; accession number X74511). Identities are indicated by black background. Asterisks denote terminal codons.

Fig. 2. Reverse transcription-coupled polymerase chain reaction RT-PCR analysis of ColF4100 in multiple human tissues. Primers used for RT-PCR amplify the 440bp of the 3' untranslated region of the cDNA. The primers were (5'-TCATATGTCCAGTGTCTGG-3') and (5'-CCTGCAAGCATGATTACATTCG-3'). The cDNA templates for RT-PCR were from the human tissues of poly(A)⁺ RNAs and were purchased from Clontech (Palo Alto, CA, USA). PCR was carried out in a final volume of 50 µl containing 1 × GC buffer I (Takara, Kyoto, Japan), 0.2mM each primer, 400mM each dNTP, 5 µl (~1 ng) of template cDNA, and 2.5 units of LA-Taq DNA polymerase (Takara, Kyoto, Japan). For the positive control, 50pg of the ColF4100 clone from a full-length enriched cDNA library was used for the template. Temperature and time scheduled were 35 cycles of 94°C for 1 min., 55°C for 1 min, and 72°C for 1 min. PCR products were separated on 2.0% agarose gel. The lower panel shows the result with a Human G3PDH Control Amplimer Set (Clontech). The control G3PDH cDNA was also purchased from Clontech and 5 µl (~1 ng) was used for PCR.

which shares significant homology with the human butyrophilin gene, from a human colon mucosa cDNA library. Both ColF4100 protein and human butyrophilin protein have the general structure of a receptor and have two Ig superfamily extracellular domains. Although the function of human butyrophilin remains unknown, there is a hypothesis that butyrophilin may also provide protection against infection for breast-fed infants (Taylor et al. 1996). Such a possibility for ColF4100 is to be investigated.

References

- Heid HW, Winter S, Bruder G, Keenan TW, Jarasch ED (1983) Butyrophilin, an apical plasma membrane-associated glycoprotein characteristic of lactating mammary glands of diverse species. *Biochim Biophys Acta* 728:228-238
- Hilton AA, Slavin AJ, Hilton DJ, Bernard CCA (1995) Characterization of cDNA and genomic clones encoding human myelin oligodendrocyte glycoprotein. *J Neurochem* 65:309-318
- Hirokawa T, Seah, B.-C., Mitaku S. (1988) SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics*. 14:378-379
- Itoh K, Itoh Y, Frank MB (1991) Protein heterogeneity in the human: Ro/SSA ribonucleoproteins. *J Clin Invest* 87:177-186
- Kyte J, Doolittle R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105-132
- Maruyama K, Sugano S (1994) Oligo-capping: a simple method to replace

- the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* 138:171–174
- Mather IH, Jack LJ (1993) A review of the molecular and cellular biology of butyrophilin, the major protein of bovine milk fat globule membrane. *J Dairy Sci* 76:3832–3850
- Pham-Dinh D, Mattei MG, Nussbaum JL, Roussel G, Pontarotti P, Roeckel N, Mather IH, Artzt K, Lindahl KF, Dautigny A (1993) Myelin/oligodendrocyte glycoprotein is a member of a subset of the immunoglobulin superfamily encoded within the major histocompatibility complex. *Proc Natl Acad Sci USA* 90:7990–7994
- Saito T, Seki N, Ishii H, Ohira M, Hayashi A, Kozuma S, Hori T (1997) Complementary DNA cloning and chromosomal mapping of a novel phosphatidylinositol kinase gene. *DNA Res* 4:301–305
- Suzuki Y, Yoshitomo-Nakagawa K, Maruyama K, Suyama A, Sugano S (1997) Construction and characterization of a full length and 5'-end-enriched cDNA library. *Gene* 200:149–156
- Takahashi M, Inaguma Y, Hiai H, Hirose F (1988) Developmentally regulated expression of a human "Finger"-containing gene encoded by the 5' half of the *ret* transforming gene. *Mol Cell Biol* 8:1853–1856
- Taylor MR, Peterson JA, Ceriani RL, Couto JR (1996) Cloning and sequence analysis of human butyrophilin reveals a potential receptor function. *Biochim Biophys Acta* 1306:1–4
- The International FMF Consrtium (1997) Ancient missense mutations in a new member of the RoRet gene family are likely to cause familial mediterranean fever. *Cell* 90:797–807
- Vernet C, Boretto J, Mattei MG, Takahashi M, Jack LJ, Mather IH, Rouquier S, Pontarotti P (1993) Evolutionary study of multigenic families mapping close to the human MHC class I region. *J Mol Evol* 37:600–612
- Williams AF, Barclay AN (1988) The immunoglobulin superfamily — domains for cell surface recognition. *Annu Rev Immunol* 6:381–405