Volume 315, number 2, 193–196 FEBS 11965 January 1993

Identification and differential expression of yeast *SEC23*-related gene (*Msec23*) in mouse tissues*

Renu Wadhwa[a], Sunil C. Kaul[b], Yasuhiko Komatsu[b], Yoji Ikawa[a], Akinori Sarai[c] and Yoshikazu Sugimoto[a]

[a]*Laboratory of Molecular Oncology and* [c]*Gene Bank, Tsukuba Life Science Center, The Institute of Physical and Chemical Research, Koyadai, Tsukuba, Ibaraki 305, Japan* and [b]*Fermentation Research Institute, Agency of Industrial Science and Technology, Tsukuba, Ibaraki 305, Japan*

Received 26 October 1992; revised version received 17 November 1992

We have isolated a yeast *SEC23*-related clone (*Msec23*) from mouse fibroblast cDNA library. It has an open reading frame of 1721 bp (64% homologous to *SEC23*–2.3 kb) which can potentially encode a 64.7 kDa protein (61% homologous to 85.4 kDa product of *SEC23*). The deduced Msec23 protein (Msec23p) sequence contains three successive Ig-like domains at the N-terminus followed by amphipathic α-helical regions, suggesting the potential of Msec23p to interact with other protein components. Further, *Msec23* is differentially expressed in mouse tissues with its high level in brain and fibroblasts.

Yeast *SEC23*-related gene; cDNA sequence; Differential tissue expression

1. INTRODUCTION

Molecular understanding of various biological phenomena demands the isolation and characterization of relevant cDNA clones. These are targeted by screening of cDNA libraries with specific antibodies or DNA probes, subtractive or differential hybridization. Using one of these strategies to isolate genes responsible for mortality and immortality in mouse fibroblasts we got hold of a novel clone which showed 64% homology to a member of yeast secretory pathway, *SEC23*. The latter has been demonstrated to play a role in intracellular transport from endoplasmic reticulum (ER) to Golgi both in vivo and in vitro [1]. Intracellular transport of proteins from cytoplasm to various organelles is the essential requirement for cellular structural and functional integrity. The molecules mediating such mechanism hence are the subject of great interest. In yeast more than 27 genes have been defined to have roles in secretory pathway by mutational complementation approach [2–5]. Mammalian homologue of Sec23p has been identified immunologically as a 84–85 kDa protein [6]. It is restricted to ER transitional cytoplasm as observed by immuno-electron microscopy in liver and pancreas of Sprague–Dawley rats. However, the information on its DNA structure is lacking. We accidentally isolated a yeast SEC23-related clone from mouse fibroblast cDNA library. We report here the nucleotide sequence of the novel cDNA clone (*Msec23*) along with its tissue specific expression in mouse.

Correspondence address: Y. Sugimoto, Laboratory of Molecular Oncology, Tsukuba Life Science Center, The Institute of Physical and Chemical Research, Koyadai, Tsukuba, Ibaraki 305, Japan. Fax: (81) (298) 36 9020.

*The nucleotide sequence reported in this paper has been submitted to the DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession number D12713.

2. MATERIALS AND METHODS

2.1. *Construction and screening of cDNA library*

Poly(A)$^+$ RNA prepared from CD1 ICR mouse embryonic fibroblasts (MEF) by using Fast track mRNA isolation kit (Invitrogen) was used to construct cDNA library in the λZAPII vector (Stratagene) at *Eco*RI site. The clone dealt with in the present report was isolated while screening of the MEF cDNA library with p66 cDNA probe isolated from the library with anti-p66 antibody (see Wadhwa et al. [7] for p66 protein).

2.2. *DNA sequencing*

cDNA clone (1.9 kb) in pBluescript SK (Stratagene) was sequenced by the dideoxy chain termination method using T7 and T3 primers (Sequenase 2.0 kit, US Biochemicals). The clones were deleted sequentially from 3′ end by exonuclease III using deletion kit (Takara). The complete nucleotide sequence was derived by ligating the sequence of deletion mutants and was confirmed by reverse direction cloning and sequencing. The DNA sequence was analyzed on a μVax computer.

2.3. *Northern analysis*

Total cell and tissue RNA were extracted and separated on denaturing agarose gel [8]. The ultraviolet-cross-linked membranes were hybridized with ^{32}P-labeled *Msec23* cDNA probe (1.9 kb).

3. RESULTS AND DISCUSSION

It would rather be fair to state that the present report is the fortuitous off-shoot of our main project which

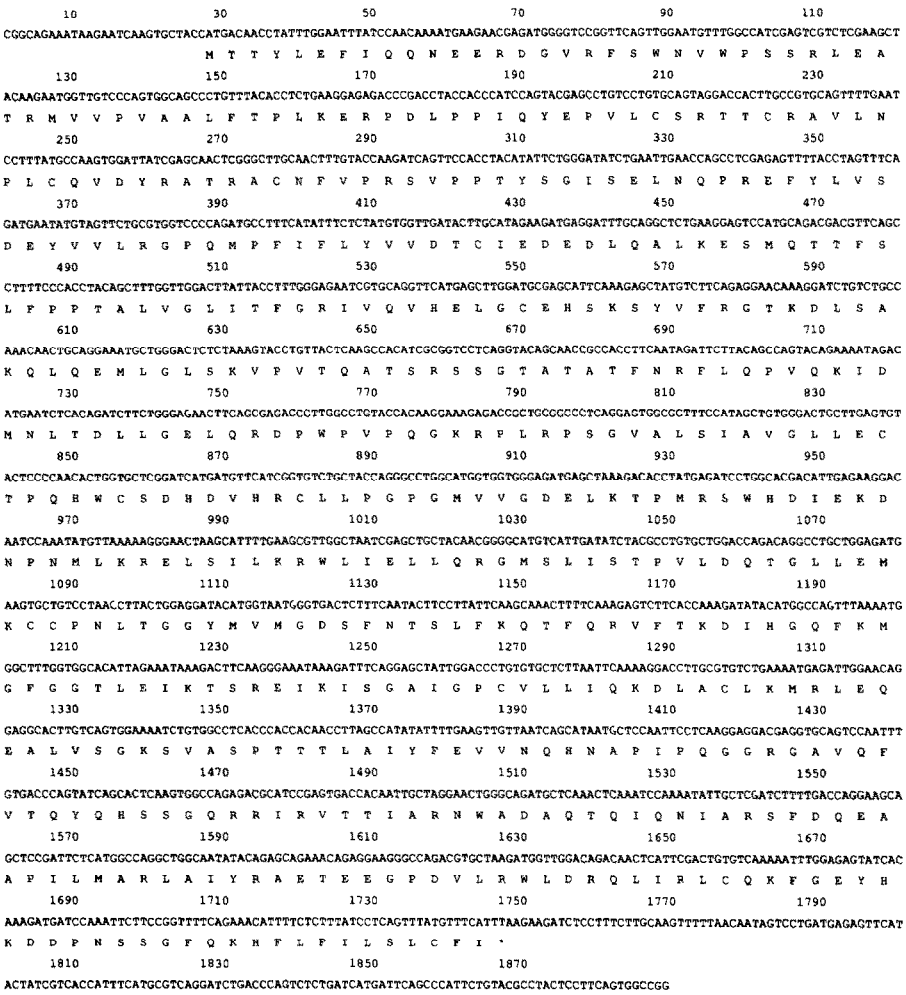

Fig. 1. *Msec23* sequence. Complete nucleotide sequence of *Msec23* isolated from MEF cDNA library along with the deduced amino acid sequence.

aims to isolate genes involved in mortal and immortal phenotypes in mouse fibroblasts. We have identified and characterized a cytosolic 66-kDa protein which is associated with the mortal phenotype of mouse fibroblasts [7]. During the screening of MEF cDNA library with 1.6 kb *Eco*RI fragment of *p66* cDNA which was aimed to obtain the 5′ upstream sequence of *p66* cDNA clone, clone 18 was isolated. However, upon partial sequencing it showed minor homology to *p66* cDNA but was found to be 64% homologous to yeast *SEC23* cDNA sequence which functions in intracellular transport from ER to Golgi [1]. Thus we have called it *Msec23*. Complete nucleotide sequence along with the predicted amino acid sequence is given in Fig. 1. It has an open reading frame of 1721 bp (28–1749). That ATG at 28 bp position acts as the initiation codon is supported by the presence of 5′ upstream in-frame stop codon TAA at 10 bp position as well as by the Kozak rule [9]. The open reading frame has the potential to encode a leucine-rich protein (p*I* 7.9) consisting of 572 amino acids corresponding to the molecular weight of 64.7 kDa, in contrast to yeast SEC23 protein (YSEC23p) with 768 amino acids with the molecular weight of 84 kDa [1] and immunologically detected 84-85 kDa mammalian Sec23p [6].

Comparison of the amino acid sequence of Msec23p with that of YSEC23p revealed a strong homology in the entire region, with 40% identical amino acids and 61% homology (Fig. 2). These sequences have no strong homology with other sequences in the database. However, close inspection of the amino acid sequences indicates that there are three homologous domains in the N-terminal halves of the proteins. These domains consist of 90–110 amino acids (D1, 15–120; D2, 120–215; D3, 215–315 in the Msec23p, and 10–115, 115–220, 220–320 in YSEC23p). In each domain, most of the hydrophobic amino acids and turn-forming amino acids such as Pro, Gly and polar amino acids are conserved between the two sequences. This conservation of structurally important amino acids suggests that the corre-

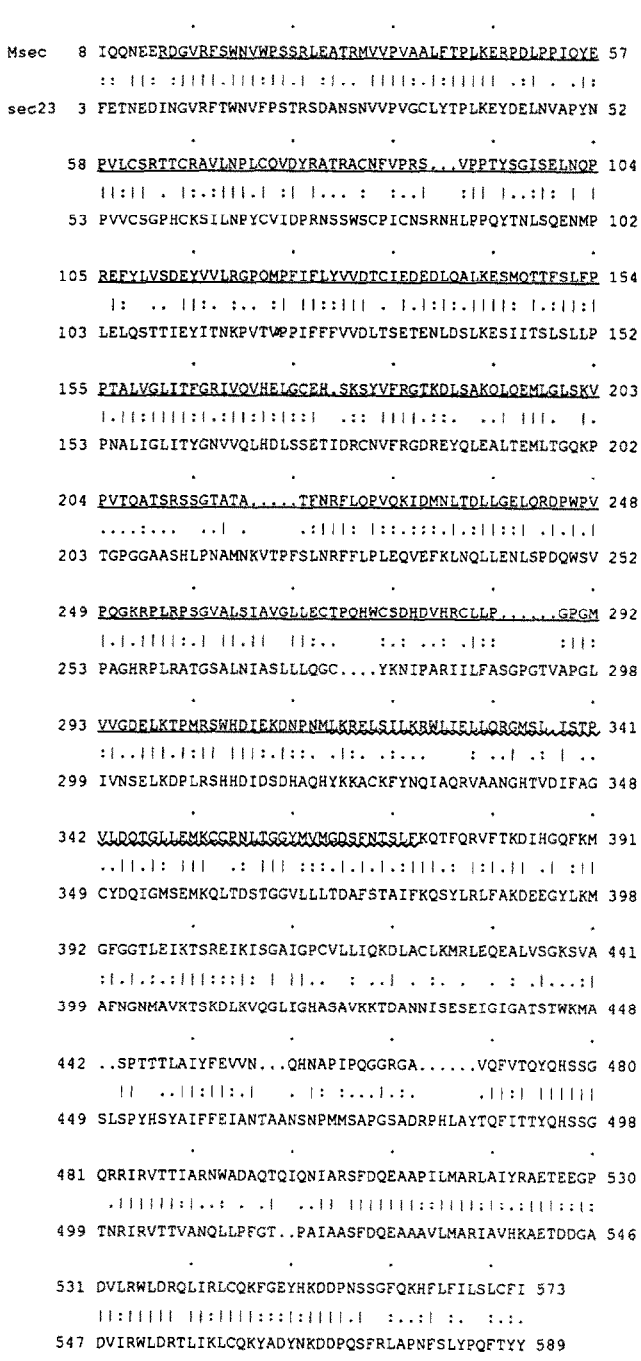

Fig. 2. Alignments of amino acid sequences of mouse Msec23p (Msec) with YSEC23p (sec23). Matches are indicated by vertical bars (identical amino acids) and dots (related amino acids). Underlining indicates three successive Ig-like domains (D1, D2 and D3) of Msec23p. Dotted line indicates helical region of Msec23.

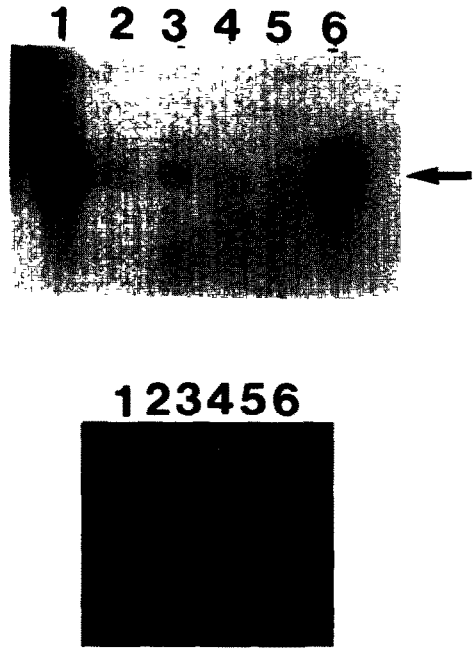

Fig. 4. Northern blot analysis of *Msec23* transcript in various mouse tissues. Lanes 1–6 are loaded with 20 μg of total RNA from brain, lung, heart, liver, kidney and fibroblasts, respectively. Lower panel shows the ethidium bromide staining of corresponding RNA lanes.

sponding domains of the two proteins have similar structure. According to the locations of hydrophobic and turn-forming amino acids, these domains are likely to form Ig-like structure with seven to nine β-strands (Fig. 3). The Ig-like structures have been found in proteins which bind to peptides and other proteins, i.e. in T cell receptor, MHC, CD4, CD8, interleukin 6 receptor and colony-stimulating factor receptor. Thus the Ig-like structure of Msec23p as well as YSEC23p accords well their role in the association with other molecules.

In addition, the C-terminal halves of Msec23p and YSEC23p have rather amphipathic α-helical characteristics in several places. In particular the mouse sequence has repeated Leu's 6–8 amino acid apart in the region between 315 and 375, next to the Ig-like domains (Fig. 2). This structure could potentially mediate its interaction with other proteins similar to the 'leucine-zipper' motif.

Msec23 is the first *SEC23*-related cDNA isolated from mammalian system though it remains to be seen whether it represents the yeast *SEC23*-corresponding gene. Msec23p (64.7 kDa) lacks the C-terminal 200 amino acids contained in YSEC23p (84 kDa) [1]. In addition, anti-YSEC23p antibody recognized 84–85 kDa protein in mouse liver and pancreas [6]. However, as can be seen in Fig. 4, brain (lane 1) and fibroblast (lane 6) expressed the high level of 3.8 kb *Msec23* transcript as compared to the others like liver where 84–85 kDa mammalian Sec23p was found to be abundant [6]. Thus *Msec23* may not be a gene encoding 84–85 kDa mammalian Sec23p. It is rather intriguing to delineate the biological significance of its differential expression and also if it can complement yeast *sec23*.

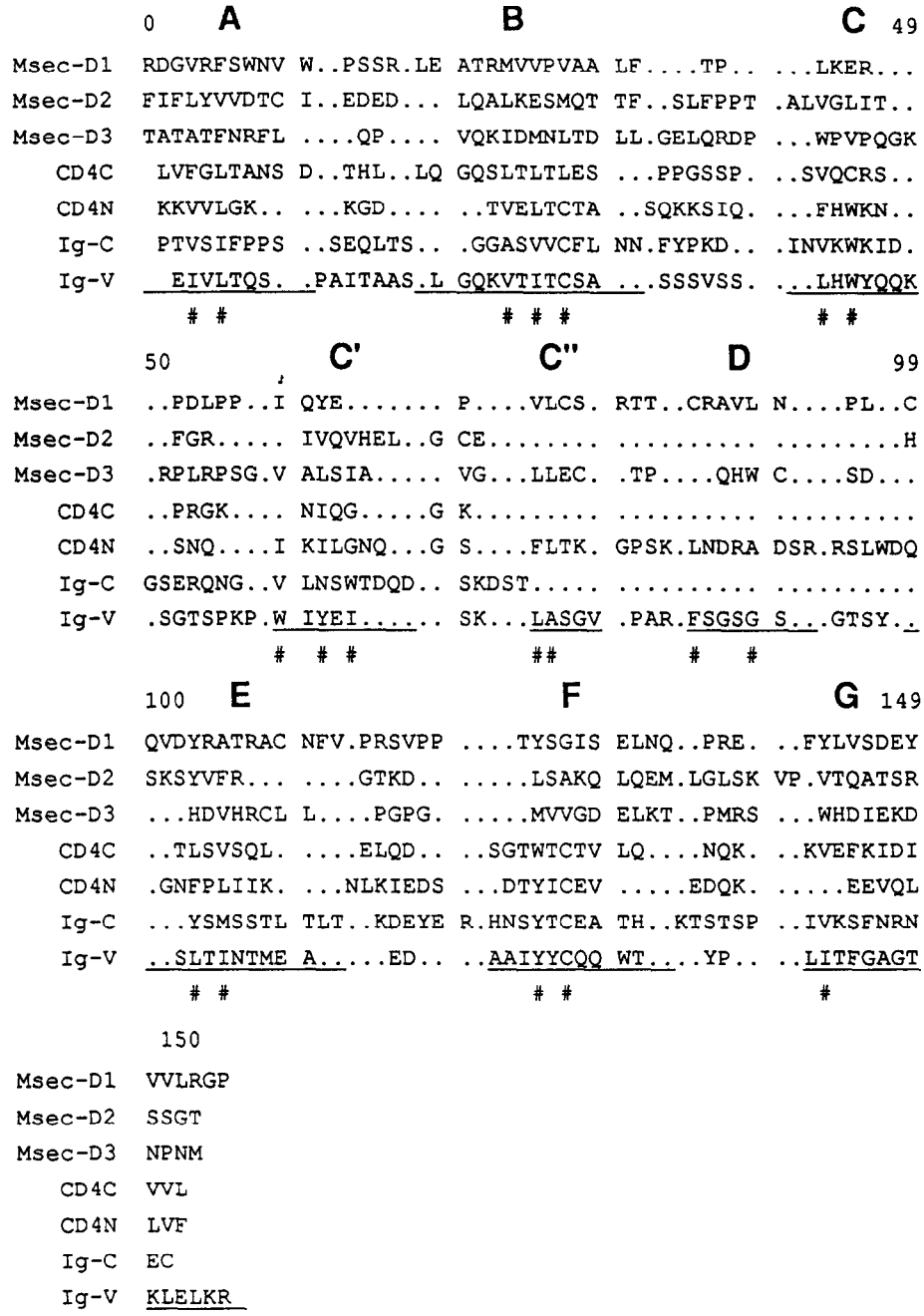

Fig. 3. Multiple sequence alignment of the Ig-like domains of Msec23p (Msec-D1 to Msec-D3), CD4 C and N domains (CD4C and CD4N), and immunoglobulin constant and variable domains (Ig-C and Ig-V). The β-strand portions are indicated by underlinings and designated by letters. # denotes the conserved hydrophobic amino acids.

REFERENCES

[1] Hicke, L. and Schekman, R. (1989) EMBO J. 8, 1677–1684.
[2] Deshaies, R. and Schekman, R. (1987) J. Cell Biol. 105, 633–645.
[3] Ferro-Novick, S., Hansen, W., Schauer, I. and Schekman, R. (1984) J. Cell Biol. 98, 44–53.
[4] Novick, P. and Schekman, R. (1979) Proc. Natl. Acad. Sci. USA 76, 1858–1862.
[5] Novick, P., Field, C. and Schekman, R. (1980) Cell 21, 205–215.
[6] Orci, L., Ravazzola, M., Meda, P., Holcomb, C., Moore, H., Hicke, L. and Schekman, R. (1991) Proc. Natl. Acad. Sci. USA 88, 8611–8615.
[7] Wadhwa, R., Kaul, S.C., Ikawa, Y. and Sugimoto, Y. (1991) Mut. Res. 256, 243–254.
[8] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) in: Molecular Cloning: a Laboratory Manual, Cold Spring Harbor, Cold Spring Harbor Laboratory, NY.
[9] Kozak, M. (1984) Nucleic Acids Res. 12, 857–872.