J. Biochem. 119, 319-328 (1996)

Isolation and Characterization of Retinoic Acid-Inducible cDNA Clones in F9 Cells: A Novel cDNA Family Encodes Cell Surface Proteins Sharing Partial Homology with MHC Class I Molecules[1]

Zhihua Zou,* Midori Nomura,* Yoshihiro Takihara,* Teruo Yasunaga,† and Kazunori Shimada*,[2]

*Department of Medical Genetics, Division of Molecular Biomedicine, Research Institute for Microbial Diseases, and †Genome Information Research Center, Osaka University, 3-1 Yamadaoka, Suita, Osaka 565

Received for publication, September 27, 1995

Rae-1 cDNA is one of the retinoic acid (RA)-inducible cDNA clones in mouse embryonal carcinoma F9 cells. Rae-1 mRNAs were detected in mouse early embryos, but not in various tissues of adult mice. RAE-1 protein apparently consists of 253 amino acids and is likely to be a glycoprotein consisting of a leader sequence, an extracellular domain, a serine, threonine, proline-rich domain, and a transmembrane domain. Interestingly, it has a weak, but significant homology with major histocompatibility complex (MHC) class I molecules and was immunocytochemically identified as a cell surface protein. By determining partial nucleotide sequences of 17 Rae-1 cDNAs isolated from the RA-induced F9 cells, at least three different kinds of Rae-1 cDNAs were identified and were named Rae-1α, Rae-1β, and Rae-1γ cDNAs, respectively. As the overall nucleotide sequence homology among these three cDNAs was about 98%, they constitute a novel gene family which is likely to be involved in early mammalian embryogenesis.

Key words: F9 cell, membrane glycoprotein, MHC class I molecule, RA-inducible gene, retinoic acid.

Mouse embryonal carcinoma F9 cells closely resemble pluripotent embryonic stem cells in terms of morphology, biochemical characteristics, and growth properties (*1, 2*). F9 cells differentiate into primitive endoderm-like cells in response to retinoic acid (RA) (*3, 4*), and also into visceral endoderm-like cells, provided the cells aggregate in the presence of a low dose of RA (*2*). This system provides a pertinent model for analyses of early mammalian development (*1*). To elucidate molecular regulatory mechanisms involved in early mammalian development, we isolated a series of cDNA clones corresponding to those genes, the expression of which increases during RA-induced F9 cell differentiation (*5*).

We have now characterized one of the previously isolated cDNA clones, named Rae-1 (=retinoic acid early inducible cDNA clone-1) (*5*) and demonstrated using immunocytochemical techniques that it codes for a novel cell surface protein. The deduced RAE-1 protein consists of 253 amino acids and shows a weak, but significant homology with major histocompatibility complex (MHC) class I molecules (*6-8*). Rae-1 mRNAs were detected in early mouse embryos, but not in various tissues of adult mice. Interestingly, there were at least three different, but highly homologous Rae-1 cDNAs, indicating the presence of a novel gene family which is likely to be involved in early mammalian development.

MATERIALS AND METHODS

Cells—Mouse embryonal carcinoma F9 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 15% fetal calf serum (FCS), and were induced to differentiate by treatment with 10^{-6} M RA (Sigma) (*2*).

RNA Preparation and RNA Analysis—Various tissues were isolated from adult BALB/c mice. Total cellular RNAs were extracted from F9 cells and from isolated tissues by the acid guanidinium thiocyanate-phenol-chloroform method (*9*). Poly(A)$^+$ RNA was prepared on an oligo(dT)-cellulose column. For RNA blot analysis, 10 μg of total cellular RNA was denatured with glyoxal, fractionated by 1% agarose gel electrophoresis, and transferred to a nylon membrane (*10*). RNA blots were hybridized with ^{32}P-labeled Rae-1 cDNA probe (*5*), prepared by a multiprime DNA labeling system purchased from Amersham (*11*).

[1] This work was supported by a Grant-in-Aid for Scientific Research (04454170) and a Grant-in-Aid for Special Project Research (03267101) from the Ministry of Education, Science, Sports and Culture of Japan. A part of this work was also supported by grants from the Japan Medical Association and the Uehara Memorial Foundation. The nucleotide sequence data reported in this paper will appear in the GSDB, DDBJ, EMBL, and NCBI nucleotide sequence databases with the accession numbers D64160, D64161, and D64162.

[2] To whom correspondence should be addressed.

Abbreviations: DMEM, Dulbecco's modified Eagle's medium; FCS, fetal calf serum; FITC, fluorescein isothiocyanate; MHC, major histocompatibility complex; nt, nucleotide(s); PBS, phosphate-buffered saline; PCR, polymerase chain reaction; RT-PCR, reverse transcription-PCR; RA, retinoic acid; Rae-1, retinoic acid early inducible cDNA clone-1; X-gal, 5-bromo-4-chloro-3-indolyl-β-D(-)-galactopyranoside.

Cloning of a Series of Rae-1 cDNAs—Isolation of the Rae-1 cDNA clone was as described (*5*). As this clone was not carrying the full-length cDNA, a cDNA library was constructed from F9 cells treated with RA for 48 h and was screened to isolate a series of longer Rae-1 cDNAs by plaque hybridization, using a ^{32}P-labeled Rae-1 cDNA insert as a probe (*5, 10*).

Reverse Transcription-PCR (RT-PCR)—To examine the structures of mRNAs corresponding to Rae-1 cDNAs, the following two primers were prepared: primer E7, 5′-GGAGATCAGCTAATGATG-3′, which corresponds to the sense strand of Rae-1 cDNA, from nucleotide (nt) number +872 to +889 (see Fig. 3A); primer E8, 5′-ATGAGTCCCACAGAGATA-3′, which corresponds to the antisense strand of Rae-1 cDNA, from nt number 1101 to 1118 (see Fig. 3A). The following two kinds of poly(A)$^+$ RNAs were prepared for this experiment: one, from F9 cells treated with RA for 72 h and the other, from the head region of 11-day mouse embryos. The first strands of cDNAs were synthesized on these poly(A)$^+$ RNAs using primer E8. For the PCR reaction, 80 μl of the reaction mixture which contains Taq polymerase and primers E7 and E8 was added to the first strand cDNA synthesis reaction mixture, using a thermal cycler as described (*5*). These primers and the same PCR procedure were also used to analyze genomic DNA from undifferentiated F9 cells.

Subcloning of PCR Products—PCR products were subcloned by the TA cloning method, using a TA Cloning Kit (Invitrogen). Briefly, after PCR, 1 μl of the PCR products was ligated to the pCR™II vector at 14.5°C overnight, 1 μl of the ligation reaction mixture was used to transform 50 μl of competent cells, and 50 μl of transformation medium was spread on an LB agar plate containing ampicillin and X-gal.

Isolation of the Mouse rae-1 Gene—A λFIXII/mouse genomic DNA library constructed on partial *Sau*3AI digest of the 129/SVJ mouse liver DNA (1×10^9 pfu/ml) was purchased from Stratagene. Phage clones carrying the mouse *rae-1* gene were screened following the procedures described in Ref. *10*, using Rae-107 cDNA as a probe. Cloned genomic DNAs were extracted from purified phage clones and were used for restriction enzyme mapping. The exon-containing region was identified by Southern blot hybridization, and was subcloned into pBluescript II KS (Stratagene).

Genomic DNA Preparation and Southern Blot Analysis—Mouse genomic DNAs were prepared from undifferentiated F9 cells following the procedures described in Ref. *10*. For Southern blot analysis, 10 μg of mouse genomic DNA was digested with restriction enzymes (*Bam*HI, *Eco*RI, *Hin*dIII), separated by 0.7% agarose gel electrophoresis and transferred to a nylon membrane. The blots were hybridized with the Rae-107 cDNA probe or with a genomic DNA probe carrying a part of the *rae-1* gene (see Fig. 6B, probe H-B). Hybridization was carried out as described (*10*) and filters were washed three times for 5 min with an excess of 2×SSC/0.1% SDS at room temperature, and then washed twice for 15 min with 0.1×SSC/0.1% SDS at 65°C.

DNA Sequence Analysis—pBluescript plasmids containing the cloned or subcloned cDNA inserts were sequenced by the dideoxy chain termination method (*12*), adapted for denatured plasmid templates, using as primers T3 and T7 (*13*). To sequence a part of each of the Rae-1 cDNAs, we used as a primer an octadecamer sequence, 5′-CATTTACAAGTCACCATG-3′ which corresponds to the nt sequence from codons 121 to 126 of Rae-1 cDNA (see Fig. 3A). PCR products subcloned in the TA cloning vector were sequenced on both strands by use of the T7 Sequencing Kit (Pharmacia) with the T7 primer, T3 primer, and M13 reverse primer.

The FASTA and BLAST programs at the Human Genome Center, Tokyo, were used to search for homologous sequences in nucleic acid and protein databases (*14, 15*).

DNA Constructions—The c-Myc epitope (*16*) was introduced within the RAE-1 protein, using the following procedures. Two complementary oligodeoxyribonucleotides, one with the sequence 5′-AATTCGAGCAGAAGCTGATCTCCGAGGAGGACCTCG-3′ and the other with 5′-AATTCGAGGTCCTCCTCGGAGATCAGCTTGTCCTCG-3′, were chemically synthesized and were annealed to form a double-stranded DNA fragment containing a reading frame for the human c-Myc epitope, EQKLISEEDL, flanked by *Eco*RI sites (*16*). This fragment was cloned into one of the two *Eco*RI sites in Rae-107 cDNA, which was isolated and characterized in this work. The tagged Rae-107 cDNA was subcloned into the mammalian expression vector pHβAPr-1-neo, and was fused to the human β-actin promoter and enhancer (*17*).

DNA Transfection and Selection—Twenty-four hours prior to DNA transfection, F9 cells were plated in DMEM supplemented with 15% FCS at the density of around 5×10^5 cells per 90-mm-diameter tissue culture dish. The medium was changed 2 h before the transfection. Cells were incubated for 16 to 20 h with transfecting DNAs, then precipitated with calcium phosphate (20 μg of the transfecting DNAs per 90-mm-diameter dish) (*18*), followed by the addition of fresh medium containing 15% FCS. After 24 h, each culture was divided into ten 90-mm-diameter dishes, and the selection for G418-resistant clones was initiated by adding G418 (GIBCO) to the medium at the final concentration of 600 μg/ml. After 12 days, G418-resistant clones were picked up and expanded for the following analysis.

Immunocytochemical Examination (19)—Cells were rinsed in phosphate-buffered saline (PBS) (137 mM NaCl/2.7 mM KCl/4.3 mM $Na_2HPO_4\cdot7H_2O$/1.4 mM KH_2PO_4, pH 7.4) and fixed in a fresh solution of 2% (w/v) formaldehyde in PBS at room temperature for 10 min (*20*). After extensive washing in PBS, cells were incubated with the primary antibody, *i.e.*, a monoclonal antibody to the human c-Myc, (MAb)9E10 (*21*) (Oncogene Science), in PBS for 1 h, at 37°C in a humidified atmosphere. Cells were again washed extensively in PBS and then incubated in a solution containing 200-fold diluted fluorescein isothiocyanate (FITC)-conjugated goat anti-mouse IgG (TAGO) for 1 h at room temperature (*22*). Finally, the cells were washed in PBS, covered with 50 mM sodium carbonate (pH 9.5) and 90% glycerol, and examined using an Olympus microscope, model BX50, equipped with a reflective light fluorescence system BX-FLA.

RESULTS

Expression of Rae-1 mRNAs—The levels of Rae-1

mRNAs in F9 cells were examined by Northern blot analysis. Rae-1 mRNAs were not detected in undifferentiated F9 cells, but were detected in F9 cells about 12 h after RA treatment, and their levels increased for up to 72 h (Fig. 1A) (*5*). As previously reported, they were not detected in various tissues of adult mice (*5*). A significant expression of Rae-1 mRNAs was detected in 11- and 14-day embryos, and at low levels in placentas and extra-embryonic tissues, but not in an 18-day embryo (Fig. 1B). Interestingly, Rae-1 mRNAs were clearly detected in the head of an 11-day embryo, but not in that of an 18-day embryo (Fig. 1B). These findings suggest that, in the course of early mouse development, the expression of Rae-1 mRNAs is controlled not only in a developmental stage-specific manner, but also in a tissue-specific manner.

Structures of Rae-1 cDNAs and Deduced RAE-1 Protein—Although Rae-1 mRNAs were detected as a 1.8 kb band in the RA-treated F9 cells, the size of the previously isolated Rae-1 cDNA was about 1.2 kb (*5*). To obtain cDNAs covering the full-length Rae-1 mRNAs, we rescreened a cDNA library constructed from F9 cells treated with RA for 48 h, using Rae-1 cDNA as a probe, and isolated 16 new clones (Fig. 2).

As Rae-109 cDNA contained the longest cDNA insert, we first constructed a physical map and found what seemed to be a block of deletion (Fig. 2). Therefore, we determined the entire nt sequence of the 2nd longest clone, Rae-107 cDNA (Fig. 2). Rae-107 cDNA contains the entire coding region; however, as the 5′- and 3′-noncoding regions of the Rae-107 cDNA are shorter than those of the Rae-109 cDNA, we compiled the sequences of these regions by determining the corresponding regions of the Rae-109 cDNA. The compiled 5′-end region contained two putative initiation ATG codons (Fig. 3A). Since the sequences around the 2nd ATG codon fit Kozak's rule (A/G N N

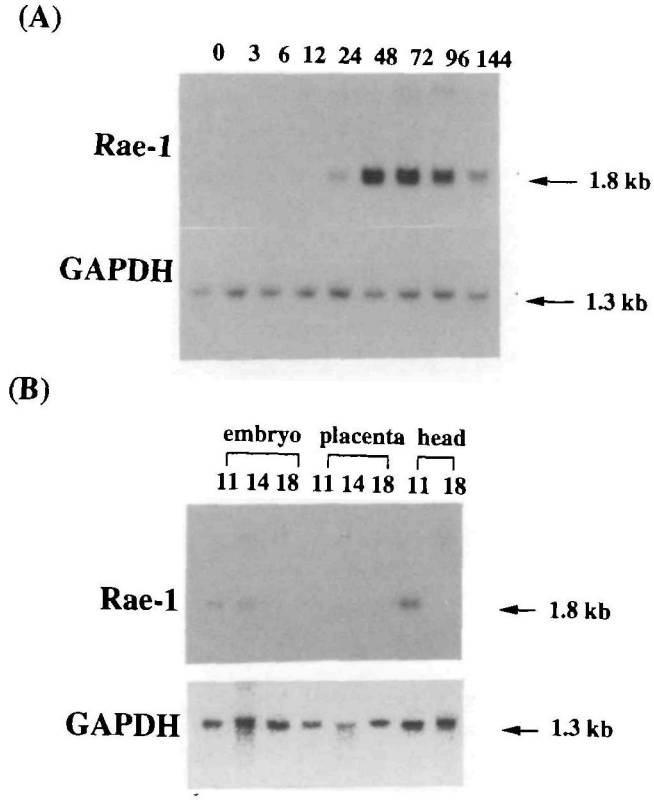

Fig. 1. **Northern blot analysis of Rae-1 mRNAs.** (A) Time course of the Rae-1 mRNAs. Total cellular RNAs were extracted from F9 cells after treatment with RA (10^{-6} M) for 3 to 144 h. Ten micrograms each of these RNAs was denatured with glyoxal and fractionated on a 1% agarose gel, transferred onto nitrocellulose filters, and hybridized with ^{32}P-labeled Rae-1 cDNA probe. The intensity of GAPDH mRNA is shown in the lower part of the figure as a control for the amount of RNA loaded in each lane. The time of RA treatment is shown above each lane. Position and size of mRNAs hybridizable with Rae-1 cDNA are indicated on the right. (B) Expression of Rae-1 mRNAs in embryos and placentas. Ten microgram samples of total cellular RNAs prepared from 11-, 14-, and 18-day embryos and placentas, and the heads of 11- and 18-day embryos were analyzed by RNA blotting, using as a probe a ^{32}P-labeled Rae-1 cDNA. The intensity of GAPDH mRNA is shown in the lower part of the figure as a control for the amount of RNAs loaded in each lane. Other symbols and procedures are as described in (A).

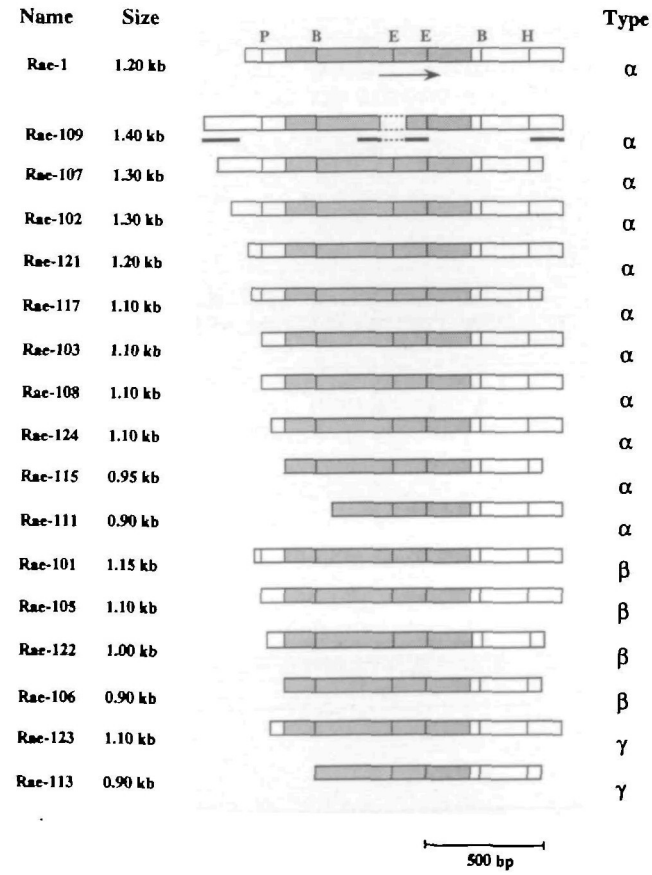

Fig. 2. **Schematic presentation of the isolated Rae-1 cDNAs.** Names, sizes, and cleavage maps of 17 Rae-1 cDNA clones are shown. The restriction sites: B, *Bam*HI; E, *Eco*RI; H, *Hind*III; P, *Pst*I. The putative coding regions are shaded. Rae-109 cDNA contained a deletion in the putative coding region, indicated by dotted lines (see text). Both strands of the entire Rae-107, Rae-101, and Rae-123 cDNAs, and the entire single strands of Rae-105 and Rae-113 cDNAs were determined. Both strands of the underlined parts of the Rae-109 cDNA were also determined. All the other cDNAs were partially sequenced using as a primer an octadecamer oligodeoxyribonucleotide (see "MATERIALS AND METHODS" and Fig. 3A), and the length and direction of DNA regions sequenced by this method are indicated by an arrow below the Rae-1 cDNA.

(A)

```
1     GCGCCGCTTGCCCGGGCCGCTGTAGTCAGTTACCTGCTCCGCGGGGCGAACACGTGCGGGGGCAGACAGACACTCC
77    AGCTTCGCAGGAGGCCGGCGCCCCTGCACCGCAGGTGCAGCTCACAGCTGGACCCACAGACCAAATGACACATAAGCAAC
157   TACTTGAAAACAGCTGGGAGCACTCTAGCAAGAAGCTCAAAGGACAGAGAAGTGTGTGCAGAGGTCAGAAACCTGCCTGG
237   AATCCCCTGGAGCAGCAGTCACAGAAGGTTGTGAGCTGCTCGATTTCAGAGGCTGCAGTTCAAGACACCAACTTTTTTCT
317   TTCTTCACCTTTCCTGACACTATTGCCAATAGTCCCTGAGGGCGCAGCTGTTGCCACAGTCACATCTGTCAGAAGAAACC

      Leader sequence
      M  A  K  A  A  V  T  K  R  H  H  F  M  I  Q  K  L  L  I  L                    20
397   ATG GCC AAG GCA GCA GTG ACC AAG CGC CAT CAT TTT ATG ATT CAG AAG CTG TTA ATT CTA

                        | Extracellular domain                  *  @
      L  S  Y  G  |  Y  T  N  G  L  D  D  A  H  S  L  R  C  N  L  T                    40
457   CTG AGC TAT GGA | TAC ACC AAC GGG CTG GAT GAT GCA CAC TCT CTT AGG TGC AAC TTG ACC

                     #                                          *
      I  K  D  P  T  P  A  D  P  L  W  Y  E  A  K  C  L  V  D  E                       60
517   ATC AAG GAT CCT ACC CCA GCA GAT CCT CTC TGG TAT GAA GCG AAG TGC TTA GTG GAT GAA

                              @                    #
      I  L  I  L  H  L  S  N  I  N  K  T  M  T  S  G  D  P  G  E                       80
577   ATA CTT ATC CTC CAT TTA AGT AAC ATA AAC AAG ACC ATG ACT TCA GGT GAC CCA GGG GAA

            @        ⌀        ⌀  *     ⌀              ⌀           *  ⌀
      T  A  N  A  T  E  V  G  E  C  L  T  Q  P  L  K  D  L  C  Q                       100
637   ACA GCA AAT GCC ACT GAA GTG GGG GAA TGT TTG ACA CAA CCT CTG AAA GAT TTG TGC CAG

            ⌀
      K  L  R  N  K  V  S  N  T  K  V  D  T  H  K  T  N  G  Y  P                       120
697   AAG TTG AGG AAC AAG GTG TCT AAC ACC AAA GTG GAC ACT CAC AAG ACC AAT GGT TAC CCA
                                                                  ΔΔ→

                                                          #
      H  L  Q  V  T  M  I  Y  L  Q  S  Q  G  Q  I  P  S  A  T  W                       140
757   CAT TTA CAA GTC ACC ATG ATT TAT CTG CAA AGC CAG GGC CAA ATT CCT AGT GCC ACC TGG

            @                                          @
      E  F  N  I  S  D  S  Y  F  F  T  F  Y  T  E  N  M  S  W  R                       160
817   GAA TTC AAC ATC AGT GAC AGT TAC TTC TTC ACC TTC TAC ACA GAG AAT ATG AGC TGG AGA
                                                   ←ΔΔ

      S  A  N  D  E  S  G  V  I  M  N  K  W  K  D  D  G  E  F  V                       180
877   TCA GCT AAT GAT GAA TCA GGG GTT ATC ATG AAT AAA TGG AAA GAT GAT GGG GAA TTT GTG

                              *
      K  R  L  K  F  L  I  P  E  C  R  Q  E  V  D  E  F  L  K  Q                       200
937   AAA CGA TTG AAA TTC TTG ATA CCC GAA TGC AGA CAG GAA GTT GAT GAA TTC TTA AAG CAG

                              | STP-rich domain
                                 #              #                    #
      P  K  E  K  P  R  |  S  T  S  R  S  P  S  I  T  Q  L  T  S  T                    220
997   CCC AAG GAA AAG CCA AGA | TCA ACC TCA AGG TCC CCC AGT ATC ACC CAG CTT ACA TCA ACT

      #  ^  ^  ^  ^  ^  #                    | Transmembrane domain
      S  P  L  P  P  P  S  H  S  T  S  |  K  K  G  F  I  S  V  G  L                    240
1057  TCC CCG CTT CCA CCT CCC AGC CAC TCT ACT TCT | AAG AAA GGA TTT ATC TCT GTG GGA CTC

      I  F  I  S  L  L  F  A  F  A  F  A  M                                           253
1117  ATC TTC ATA TCT TTA TTA TTT GCA TTT GCA TTT GCA ATG TGA AGAGGAAATATTATACCCAAGGA
1182  GAGGCCAGAGGAGGGCATCGGATCCTGTGGAACTGGAGTTACAGTTGGTTTCCAATACTGGGAACTGAATCCAGGACCT
1261  CTCCAAGAACAGCAAGTGCTCTTTGCTACTGAGCCTTATCTCCCTCCTCTAATACATTTCTTAATAAGAGAATATAGTT
1340  TGAATATTATTAAACCCAGGGCAAGTCCCACAGTAAGCTTCACAATAAAGTTGATTGTTTCAGAATCAAAAAGGAAAAC
1419  ACATTTTAAAATGACTGTCATGGCCAGTTATCACATCGTTTGTAGCAATGTATGATGGAGGACGTCCAGGTCCGTCTAA
1498  AAAAATGCCAGGCAATACAGCAACATCTAATAAAACTGGTAAATTTAAA
```

Fig. 3A

[ATG] G) (*23*), we propose that this is the initiation codon (Fig. 3A). There were two poly-adenylation signals in the 3′-noncoding region (Fig. 3A), corresponding to the presence of 2 kinds of Rae-1 cDNAs with different lengths of the 3′-noncoding region (Fig. 2). The hydropathy profile of the deduced RAE-107 protein revealed 3 hydrophobic regions; 2 close to the N-terminal region, and 1 at the C-terminal region (Fig. 3B). The region between the 2nd and 3rd hydrophobic regions is highly hydrophilic. The initiating methionine is followed by a 23-amino-acid-residue, predominantly hydrophobic sequence corresponding to the first hydrophobic region close to the N-terminus and having the characteristics of a leader sequence (Fig. 3, A and B). Comparisons with previously documented leader sequences suggested that a cleavage is introduced after the 24th glycine residue (Fig. 3A) (*24*). The resulting mature RAE-107 protein consists of 229 amino acids with a predicted molecular mass of about 26 kDa, and consists of the following three domains: (i) an extracellular domain, (ii) a serine, threonine, and proline-rich (STP-rich) domain, and (iii) a C-terminal transmembrane domain (Fig. 3, A and B).

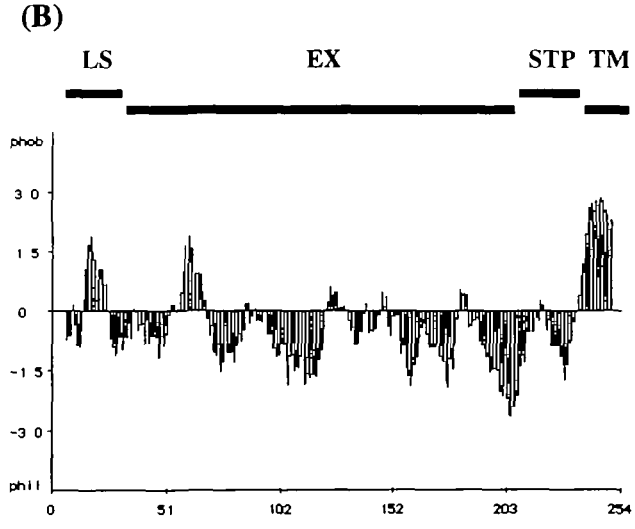

Fig. 3. **Structures of Rae-1 cDNAs and deduced RAE-1 protein.** (A) A compiled nt sequence of Rae-107 and Rae-109 cDNAs and a deduced amino acid sequence of RAE-107 protein. Nucleotide sequences of the 5′- and 3′-noncoding regions are compiled from those of Rae-107 and Rae-109 cDNAs, and those for the coding region are from Rae-107 cDNA (Fig. 2). Two putative initiation ATG codons in the 5′-noncoding region and 2 polyadenylation signal sequences in the 3′-noncoding regions are double-underlined. The predicted amino acid sequence is shown in single-letter code above the nt sequence. The numbers at the left and right sides are those of nt and amino acids, respectively. ΔΔ→ and ←ΔΔ indicate the 5′ and 3′ ends of the deletion present in Rae-109 cDNA, respectively (see Fig. 2). Five cysteine residues, 5 putative *N*-glycosylation sites and 8 possible *O*-linked glycosylation sites are denoted by •, @, and #, respectively, and a PLPPP sequence is denoted by ∧. The boundaries of domains are marked with vertical lines. Six conserved amino acid residues between mouse RAE-1 protein and human MHC class I molecules in the α-helix of α1 domain (see text and Ref. 7) are denoted by ϕ. The nt sequences of Rae-109 cDNA cover two DNA regions, one from nt numbers 1 to 46, and the other, from nt numbers 1401 to 1546, and those of Rae-107 cDNA cover a DNA region from nt numbers 47 to 1400. The octadecamer sequence used as a primer to partially sequence the other Rae-1 cDNAs is boxed, and the sequences used to classify Rae-1 cDNAs are underlined. (B) The hydropathy profile of the deduced RAE-107 protein. Four solid horizontal bars indicate the approximate locations and sizes of the following regions: LS, leader sequence; EX, extracellular domain; STP, serine, threonine, proline-rich domain; TM, transmembrane domain.

(i) The extracellular domain consists of 182 amino acids and contains 5 potential *N*-linked glycosylation sites (*25*), 3 possible *O*-linked glycosylation sites (PS/T, S/TNNP) (*26, 27*), and 5 cysteine residues (Fig. 3A).

(ii) The STP-rich domain consists of 25 amino acids and 19 of them are serine, threonine or proline residues. There are 5 possible *O*-linked glycosylation sites (Fig. 3A) (*26, 27*). One PLPPP sequence present in this region may form a hinge-like domain (*28*).

(iii) The C-terminal 14 amino acids predominantly consist of hydrophobic residues and probably form a transmembrane domain. This region may serve as an anchor to the membrane (Fig. 3A).

All these structural features suggest that RAE-107 protein is a heavily glycosylated, novel cell surface protein.

The Rae-109 cDNA contained one 111 bp long deletion within its coding region and the deleted region covered a reading frame of 37 amino acids (Figs. 2 and 3A). As the ends of the deleted region have the GT and AG sequences (Fig. 3A), the mRNA corresponding to the Rae-109 cDNA is probably generated by alternative splicing (*29*).

Classification of Rae-1 cDNAs and Homologies among the Rae-1 Members—During the course of these studies, we noted that there may be several sequence heterogeneities in Rae-1 cDNAs, so partial nt sequences of all the other clones were determined, using as a primer an octadecamer corresponding to the nt sequence from codons 121 to 126 (Fig. 3A). Three different kinds of Rae-1 cDNAs were identified and were named Rae-1α, Rae-1β, and Rae-1γ cDNAs, respectively: 11 of the 17 clones corresponded to Rae-1α, 4 corresponded to Rae-1β, and 2, to Rae-1γ cDNAs (Fig. 2). These results suggest that Rae-1α mRNAs are the predominantly expressed Rae-1 mRNAs in the RA-treated F9 cells. We chose Rae-101 and Rae-123 cDNAs as representatives corresponding to the Rae-1β and

TABLE I. **Nucleotide substitutions in the 1st, 2nd, and 3rd positions of codons.**

Compared family members[a]	Counted region[b]	Positions of codons		
		1st	2nd	3rd
Rae-1α *vs.* Rae-1β	entire	6 (6)[c]	7 (7)	4 (2.5)[d]
Rae-1α *vs.* Rae-1γ	entire	7 (7)	4 (4)	7 (4)
Rae-1β *vs.* Rae-1γ	entire	7 (7)	7 (7)	9 (7)
Rae-1α *vs.* Rae-1β	rep.seq.	3 (3)	4 (4)	2 (1.5)[d]
Rae-1α *vs.* Rae-1γ	rep.seq.	5 (5)	2 (2)	3 (3)
Rae-1β *vs.* Rae-1γ	rep.seq.	2 (2)	4 (4)	5 (5)

[a]Rae-1α, Rae-1β, and Rae-1γ indicate Rae-107, Rae-101, and Rae-123 cDNAs, respectively. [b]The "entire" lines show numbers of substitutions present in the entire coding region, and the "rep.seq." lines show only those present in the repeatedly sequenced region (see Fig. 3A). [c]Numbers were taken from the sequence data deposited in DDBJ, and those in parentheses indicate numbers of nonsynonymous substitutions. [d]In Rae-1β, the codon 194, GTT (V), was substituted by ATG (M) (data not shown). There are two possible ways leading to this substitution: one is GTT (V)→ATT (I)→ATG (M), and the other, GTT (V)→GTG (V)→ATG (M). In the former, the 3rd position substitution is nonsynonymous, and in the latter, it is synonymous. As we could not distinguish between these two, we counted this substitution 0.5 nonsynonymous.

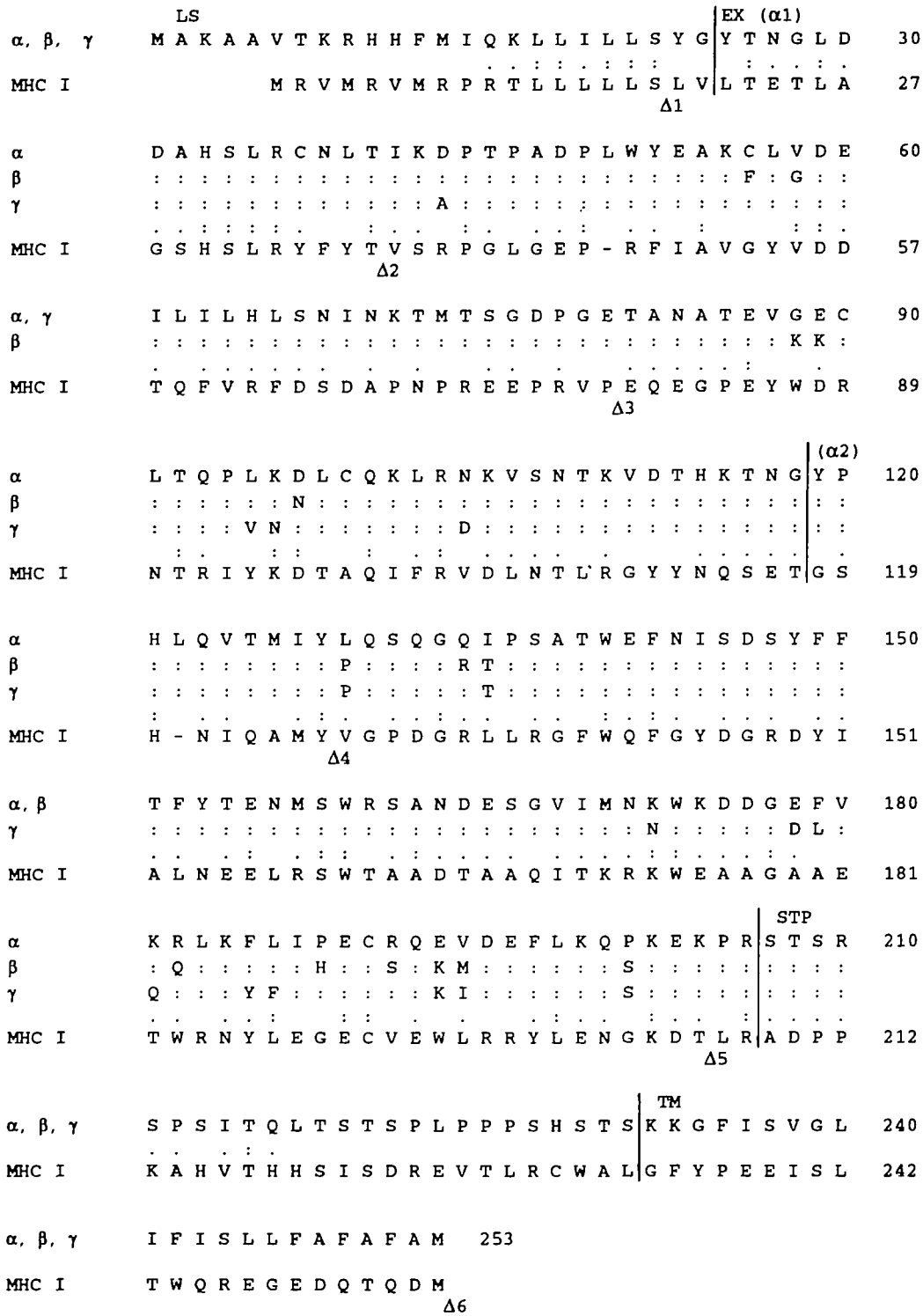

Fig. 4. **Comparison of the deduced amino acid sequences among Rae-1 family members and bovine MHC class I molecule.** The deduced amino acid sequences of Rae-107(α) cDNA are shown in lines labeled with α at the left, and those of the Rae-101(β) and Rae-123(γ) cDNAs are shown in lines labeled with β or γ, only where they differ from the Rae-107(α) sequences, respectively. Each colon represents an identical amino acid residue. Those lines labeled with α, β, γ or α, β or α, γ indicate that there is no difference among the sequences shown in these lines, respectively. The numbers at the right side are those of amino acids (see Fig. 3A). The alignment of amino acid sequence of RAE-107 protein with that of bovine MHC class I molecule (BOLA-A) is shown in lines labeled with MHC I at the left. Above the MHC I lines, a colon represents an identical amino acid residue and a dot, a conservative amino acid substitution. Insertions and deletions made in the MHC I sequence during optimization are marked with a dash and Δ, respectively: Δ1 corresponds to amino acid residues #18-19, GV; Δ 2, #38 G; Δ 3, #78-79 WM; Δ 4, #127-129 GCD; Δ 5, #206 L; Δ 6, from #256 to the C-terminal. The other symbols were as described in the legend to Fig. 3B.

Rae-1γ cDNAs, respectively (Fig. 2), and determined their entire nt sequences (data not shown).

The nt sequence homologies among these three cDNAs are about 98%, indicating that the generation of this family is a relatively recent event. The deduced amino acid sequence homology between RAE-101(β) and RAE-123(γ) is 92.1%, which is significantly lower than that between the other two combinations of the family members, *i.e.*, the homology between RAE-107(α) and RAE-101(β) is 94.5% and that between RAE-107(α) and RAE-123(γ) is 94.1%. The nt substitutions were relatively rare in the 5′- and 3′-noncoding regions (data not shown), and those present in the coding region were classified by their positions in codons (Table I). Their distributions were random and showed no preference for the 3rd positions (Table I). This is in contrast to the finding that nt diversities for a gene are larger at synonymous sites, as for the 3rd positions of codons, than at nonsynonymous sites, such as the 1st and 2nd positions of codons (*30*). In the case of Rae-1 family cDNAs, we found that even in the 3rd positions, most nt substitutions are nonsynonymous (Table I). We classified the nt substitutions present in the repeatedly sequenced region and confirmed that, although the numbers are small, almost all substitutions in the 3rd positions are nonsynonymous (Table I). Interestingly, these amino acid substitutions were distributed only within the extracellular domain, and were clustered in a region close to the C-terminal end of the domain (Fig. 4). These results suggest that positive selection operates on the extracellular domain of the RAE-1 family proteins.

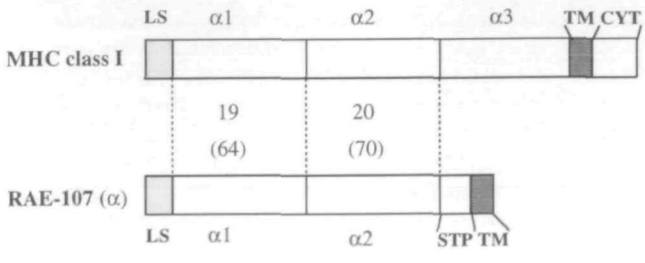

Fig 5 **Structural similarity between RAE-1 protein and MHC class I molecule.** A scheme showing the degree of deduced amino acid sequence homology between RAE-107(α) protein and bovine MHC class I molecule Symbols: α1, α2, and α3 represent extracellular domains; CYT, cytoplasmic domain. The other symbols are as described in the legends to Fig. 3B. In each domain, the numbers indicate the sequence identity in % and those in parentheses indicate the sequence similarity in %.

TABLE II. **Analyses of the types of expressed Rae-1 mRNAs and the corresponding *rae-1* genes.** The detailed procedures of the experiments were described in the "MATERIALS AND METHODS", sections "*Reverse Transcription-PCR (RT-PCR)*" and "*Subcloning of PCR Products*".

Sources of poly(A)$^+$ RNAs or genomic DNAs	Examined numbers	Types of mRNAs or genomic DNAs		
		Rae-1α	Rae-1β	Rae-1γ
Poly(A)$^+$ RNAs from F9 cells treated with RA for 72 h	50	38 (76%)	6 (12%)	6 (12%)
Poly(A)$^+$ RNAs from heads of 11-day mouse embryos	50	7 (14%)	23 (46%)	23 (46%)
Genomic DNAs from undifferentiated F9 cells	31	9 (29%)	11 (35%)	11 (35%)

Deduced RAE-1 Proteins Share Partial Homology with MHC Class I Molecules—Comparison of the deduced amino acid sequences of RAE-1 proteins with those in protein databases revealed that the RAE-1 proteins show a weak, but significant homology to MHC class I molecules (Figs. 4 and 5) (*31*). MHC class I molecules consist of L, α1, α2, α3, transmembrane, and cytoplasmic domains (Fig. 5) (*8*). On the other hand, RAE-1 proteins have no regions corresponding to the α3 and the cytoplasmic domains, but do have regions corresponding to the L, α1, α2, and transmembrane domains, and have an STP-rich domain which is absent in the MHC class I molecules (Figs. 4 and 5). Therefore, the extracellular domain of RAE-1 proteins can be divided into α1 and α2 subdomains (Figs. 4 and 5). Mouse RAE-107 protein and bovine MHC class I molecule (BOLA-A) showed identities (similarities) of 19% (64%)

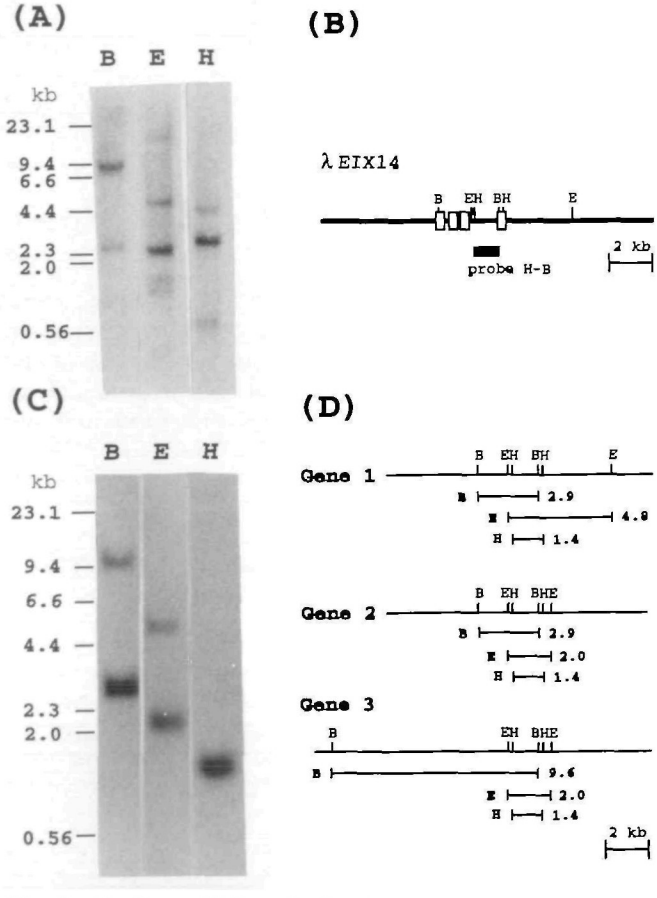

Fig. 6 **Southern blot analysis of mouse genomic DNA.** (A) Mouse genomic DNAs (10 μg each) were digested with restriction enzymes, fractionated on 0.7% agarose gels, transferred onto nitrocellulose filters, and hybridized with the 1.3 kb Rae-107 cDNA probe. λ DNA digested with *Hin*dIII was loaded as a size marker. B, *Bam*HI, E, *Eco*RI; H, *Hin*dIII. (B) Relevant restriction map of the genomic DNA cloned in λEIX14. Boxes indicate approximate sizes and locations of the sequences present in Rae-107(α) cDNA, from nt numbers 488 to 1546 (Fig. 3A). The location of probe H-B, used for the Southern blot analysis, is shown below the map. (C) Mouse genomic DNAs were digested, fractionated, Southern blotted, and hybridized with the H-B probe as described in (A). (D) Deduced restriction maps of the putative *rae-1* gene family members, gene 1, 2, and 3. The estimated sizes of *Bam*HI, *Eco*RI, and *Hin*dIII fragments are shown below the maps.

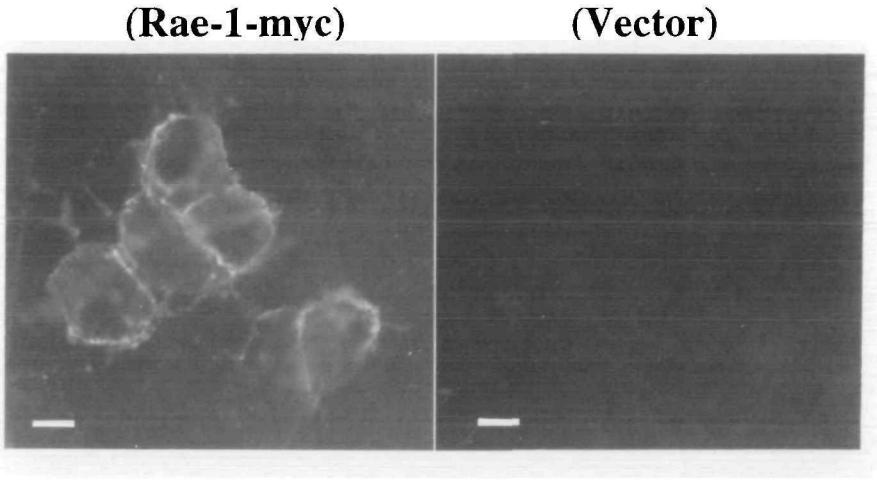

Fig 7. **Immunocytochemical examination of the F9-36 cells.** The F9-36 cells transfected with pHβAPr-1-neo/Rae-107-myc DNA are shown in (Rae-1-myc) The F9-pHβAPr-1-neo cells transfected with a vector are shown in (Vector) Experimental procedures are described in "MATERIALS AND METHODS". Bars. 20 μm.

and 20% (70%) between their α1 and α2 domains, respectively (Fig. 5).

Analyses of the Types of Expressed Rae-1 Family Members in RA-Treated F9 Cells and Mouse Embryos—To estimate the relative levels of the Rae-1α, Rae-1β, and Rae-1γ mRNAs, poly(A)$^+$ RNAs prepared from F9 cells treated with RA for 72 h and from the heads of 11-day mouse embryos were reverse-transcribed into the first-strand cDNAs using primer E8, then amplified by PCR using primers E7 and E8, and the amplified products were cloned as described in "MATERIALS AND METHODS." Fifty clones were randomly picked up for each type of poly(A)$^+$ RNA and the inserts were sequenced. The relative copy numbers of the *rae-1* genes corresponding to each of the three types of Rae-1 mRNAs were also estimated by similar methods. All the results are summarized in Table II, and support not only the presence of three types of Rae-1 mRNAs, but also the presence of equal copy number(s) of the corresponding *rae-1* genes. In the RA-induced F9 cells, Rae-1α mRNA was predominantly detected, while in the head region of 11-day mouse embryos, Rae-1β and Rae-1γ mRNAs were the predominantly detected family members (Table II). These results suggest that the regulatory mechanisms controlling the expression of the family members are different.

Southern Blot Analysis of the rae-1 Gene—To gain insight into the structure of the *rae-1* gene, genomic DNAs isolated from undifferentiated F9 cells were digested with several restriction endonucleases and were analyzed by Southern blotting, using the Rae-107 cDNA as a probe. Although several bands, including faint ones were detected (Fig. 6A), we could not conclude from this result that the *rae-1* gene has multiple members. Accordingly, we decided to isolate *rae-1* gene(s) from a mouse genomic DNA library, using the Rae-107 cDNA as a probe. We found that one of the isolated clones, named λEIX14, carries a DNA fragment covering a part of the Rae-107(α) cDNA, from nt numbers 488 to 1546 (Figs. 3A and 6B). We prepared a 1.2 kb *Hin*dIII/*Bam*HI fragment from λEIX14 and analyzed the mouse genomic DNAs by Southern blotting, using this fragment as a probe (Fig. 6B, probe H-B). The results are shown in Fig. 6C: two bands were detected in *Bam*HI digests and the signal of the 2.9 kb band was apparently stronger than that of the 9.6 kb band; two bands were detected in *Eco*RI digests and the signal of the 2.0 kb band was significantly stronger than that of the 4.8 kb band; only one band of 1.4 kb showing a strong signal was detected in *Hin*dIII digests. Based on these results, we deduced *rae-1* gene structures which support the presence of three different genes (Fig. 6D). Apparently, gene 1 corresponds to the *rae-1* gene cloned in λEIX14. Recently, we isolated two other phage clones, named λEIX2 and λEIX16, and found that λEIX2 carries a part of gene 2, which corresponds to Rae-1β cDNA, and λEIX16 carries a part of gene 3 which corresponds to Rae-1γ cDNA (unpublished data). Isolation of the genomic DNAs covering all the Rae-1 family members is in progress in our laboratory.

Localization of RAE-1 Protein in F9 Cells—The deduced amino acid sequences of RAE-1 proteins strongly suggest that they are cell surface glycoproteins. To determine the localization of RAE-1 proteins in F9 cells, we constructed an epitope-tagged Rae-1α cDNA derivative, Rae-107-myc DNA, which contained a DNA insert encoding a part of the human c-Myc protein within the reading frame of Rae-107 cDNA (see "MATERIALS AND METHODS"). The Rae-107-myc DNA was subcloned into a mammalian expression vector, pHβAPr-1-neo DNA (*17*), and was then transfected into F9 cells. Among 30 G418-resistant transfectants, the one showing the highest levels of Rae-107-myc mRNAs was selected and named F9-36. The localization of RAE-1α-Myc protein in the F9-36 cells was determined by using an immunocytochemical method (*19*) and a monoclonal antibody to the human c-*myc* proto-oncogene product (*21*). High levels of immunoreactivity were observed on the surface of F9-36 cells, but not on the surface of F9 transfectants with the control pHβAPr-1-neo DNA (Fig. 7) (*17*). These results show that RAE-1 proteins are indeed present on the cell surface.

DISCUSSION

We obtained evidence for a novel cDNA family encoding cell surface proteins and probably involved in early mouse embryogenesis. Three different but highly conserved members were identified and were named Rae-1α, Rae-1β, and Rae-1γ cDNAs.

Characteristics of the Predicted Structure of RAE-1 Proteins—The deduced RAE-1 protein showed a weak, but significant homology to the α1 and α2 domains of bovine MHC class I molecule (Figs. 4 and 5). MHC class I

molecules are membrane-bound cell surface glycoproteins with a molecular mass of 45 kDa (*8*). There is no cysteine residue in the $\alpha 1$ domain of the MHC class I family. However, 2 of the 3 cysteine residues present in the $\alpha 2$ domain are highly conserved among the MHC class I family, and maintain the globular structure of the domain by forming an S-S bridge (*32, 33*). In the RAE-1 protein, only one cysteine residue is present in the region corresponding to the $\alpha 2$ domain of MHC class I molecule, and 4 residues are present in the region corresponding to the $\alpha 1$ domain (Fig. 3A). Accordingly, the three-dimensional structures of the extracellular domain of RAE-1 proteins are likely to be different from those of the MHC class I family.

Interestingly, 6 of the 9 invariant amino acids present in the α-helix region of the $\alpha 1$ domain of MHC class I molecules are conserved in the corresponding region of RAE-1 proteins (Fig. 3A) (*7*). All of these 6 residues are distributed on the surface of each turn of the α-helix, forming an upper lip-like structure, facing the lower lip-like structure formed by the $\alpha 2$ domain, and surrounding the antigen recognition site (*6*). These features suggest that the invariant residues are important to construct the backbone of the α-helix, and suggest that the corresponding region of the extracellular domain of RAE-1 protein has a similar α-helix structure.

Possible Functions of RAE-1 Proteins—As RAE-1 proteins showed a low but significant homology to MHC class I molecules, they may share similar functions. The MHC class I molecule associates with a non-MHC-encoded, non-membrane-bound light chain, β_2-microglobulin through its $\alpha 3$ domain, binds immunogenic peptides and presents them to cytotoxic T lymphocytes (*8*). However, it is difficult to assign similar functions, because RAE-1 proteins have no region corresponding to the $\alpha 3$ domain of MHC class I molecules, and because their extracellular structure probably differs from that of the MHC class I molecules.

MHC class I molecules have also been implicated in morphogenesis and cell-to-cell interaction in embryonic development (*34*). In determining lineages of individual cells and enabling controlled morphogenesis and organogenesis, molecular mediators function as diffusible factors recognized by specific cell surface receptors, or cell surface and extracellular matrix components regulating cell-to-cell or cell-to-matrix interactions (*35-37*). Rapid, dramatic changes of glycoconjugates have been observed during the differentiation of cultured embryonal carcinoma cells and during embryogenesis (*38*). These observations suggest that cell surface glycoconjugates play a vital role in cell-to-cell or cell-to-matrix interactions and in early embryonic development (*39-42*). The RAE-1 proteins may be good candidates for temporal as well as spatial modulators of these cell-to-cell interactions, because they are cell surface glycoproteins, their expression is apparently abundant in the early stage of mouse development and they probably predominate in the embryonal head region (Fig. 1). Even though Rae-1 family members are highly homologous, their protein products may function distinctively during early mouse development, because we found by RT-PCR and sequence analysis that in RA-induced F9 cells, Rae-1α is predominantly expressed over the other two members, and, in the early mouse embryos, Rae-1β and Rae-1γ mRNAs are predominantly expressed (Table II).

Further study on the Rae-1 cDNA family should shed light on molecular mechanisms of cell-to-cell interactions in early mammalian development.

We thank M. Ohara for many helpful comments on the manuscript. We also thank the National Center for Biotechnology Information, USA, for providing access to the network BLAST server, and the Human Genome Center, Institute of Medical Science, The University of Tokyo, Japan, for the network BLAST and FASTA servers.

REFERENCES

1. Martin, G.R. (1980) Teratocarcinomas and mammalian embryogenesis. *Science* **209**, 768-776
2. Rudnicki, M.A. and McBurney, M.W. (1987) Cell culture methods and induction of differentiation of embryonal carcinoma cell lines in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach* (Robertson, E.J., ed.) pp. 19-49, IRL Press, Oxford
3. Kurkinen, M., Cooper, A.R., Barlow, D.P., Jenkins, J.R., and Hogan, B.L.M. (1983) Gene expression during parietal endoderm differentiation in mouse embryos and teratocarcinoma cells in *Teratocarcinoma Stem Cell* (Silver, L.M., Martin, G.R., and Strickland, S., eds.) Vol. 10, pp. 389-401, Cold Spring Harbor Conferences on Cell Proliferation, Cold Spring Harbor, New York
4. Strickland, S. and Mahdavi, V. (1978) The induction of differentiation in teratocarcinoma stem cells by retinoic acid. *Cell* **15**, 393-403
5. Nomura, M., Takihara, Y., and Shimada, K. (1994) Isolation and characterization of retinoic acid-inducible cDNA clones in F9 cells: One of the early inducible clones encodes a novel protein sharing several highly homologous regions with a *Drosophila* polyhomeotic protein. *Differentiation* **57**, 39-50
6. Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. (1987) Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* **329**, 506-512
7. Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. (1987) The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* **329**, 512-518
8. Hood, L., Kronenberg, M., and Hunkapiller, T. (1985) T cell antigen receptors and the immunoglobulin super-gene family. *Cell* **40**, 225-229
9. Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156-159
10. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
11. Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6-13
12. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
13. Hattori, M. and Sakaki, Y. (1986) Dideoxy sequencing method using denatured plasmid templates. *Anal. Biochem.* **152**, 232-238
14. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410
15. Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**, 2444-2448
16. Klar, A., Baldassare, M., and Jessell, T.M. (1992) F-Spondin: A gene expressed at high levels in the floor plate encodes a secreted protein that promotes neural cell adhesion and neurite extension. *Cell* **69**, 95-110
17. Gunning, P., Leavitt, J., Muscat, G., Ng, S.-Y., and Kedes, L.

(1987) A human β-actin expression vector system directs high-level accumulation of antisense transcripts. *Proc. Natl. Acad. Sci. USA* **84**, 4831-4835
18. Wigler, M., Silverstein, S., Lee, L.-S., Pellicer, A., Cheng, Y.-C., and Axcel, R. (1977) Transfer of purified Herpes simplex thymidine kinase gene to cultured mouse cells. *Cell* **11**, 223-232
19. Polak, J.M. and Van Noorde, S. (1983) *Immunocytochemistry: Practical Application in Pathology and Biology*, John Wright PSG, Littleton, MA
20. McNeill, H., Ozawa, M., Kemler, R., and Nelson, W.J. (1990) Novel function of the cell adhesion molecule uvomorulin as an inducer of cell surface polarity. *Cell* **62**, 309-316
21. Evan, G., Lewis, G.K., Ramsay, G., and Bishop, J.M. (1985) Isolation of monoclonal antibodies specific for human c-*myc* proto-oncogene product. *Mol. Cell. Biol.* **5**, 3610-3616
22. Dodd, J. and Jessell, T.M. (1985) Lactoseries carbohydrates specify subsets of dorsal root ganglion neurons projecting to the superficial dorsal horn of rat spinal cord. *J. Neurosci.* **5**, 3278-3294
23. Kozak, M. (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12**, 857-872
24. von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* **14**, 4683-4690
25. Mann, K., Deutzmann, R., Aumailley, M., Timpl, R., Raimondi, L., Yamada, Y., Pan, T.-C., Conway, D., and Chu, M.-L. (1989) Amino acid sequence of mouse nidogen, a multidomain basement membrane protein with binding activity for laminin, collagen IV and cells. *EMBO J.* **8**, 65-72
26. Jentoft, N. (1990) Why are proteins *O*-glycosylated? *Trend. Biochem. Sci.* **15**, 291-294
27. Wilson, I.B.H., Gavel, Y., and von Heijne, G. (1991) Amino acid distributions around *O*-linked glycosylation sites. *Biochem. J.* **275**, 529-534
28. Viitala, J., Carlsson, S.R., Siebert, P.D., and Fukuda, M. (1988) Molecular cloning of cDNAs encoding lamp A, a human lysosomal membrane glycoprotein with apparent M_r 120,000. *Proc. Natl. Acad. Sci. USA* **85**, 3743-3747
29. Sharp, P.A. (1981) Speculations on RNA splicing. *Cell* **23**, 643-646
30. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution*, Cambridge University Press, Cambridge
31. Ennis, P., Jackson, A.P., and Parham, P. (1988) Molecular cloning of bovine class I MHC cDNA. *J. Immunol.* **141**, 642-651
32. Ferguson, W.S., Terhorst, C.T., Robb, R.J., and Strominger, J.L. (1979) Localization of the disulfide bridges of human histocompatibility antigens. *Mol. Immunol.* **16**, 23-28
33. Martinko, J.M., Halpern, R., Adlersberg, J.B., and Nathenson, S.G. (1981) Structure of murine major histocompatibility complex alloantigens: Identification of cysteine residues involved in two interchain disulfide bonds of H-2K^b. *Mol. Immunol.* **18**, 883-888
34. Ohno, S. (1977) The original function of MHC antigens as the general plasma membrane anchorage site of organogenesis-directing proteins. *Immunol. Rev.* **33**, 59-69
35. Bernfield, M. and Sanderson, R.D. (1990) Syndecan, a developmentally regulated cell surface proteoglycan that binds extracellular matrix and growth factor. *Phil. Trans. R. Soc. Lond. B* **327**, 171-186
36. Humphries, M.J. (1990) The molecular basis and specificity of integrin-ligand interaction. *J. Cell. Sci.* **97**, 585-592
37. Yamada, K.M. (1983) Cell surface interactions with extracellular materials. *Annu. Rev. Biochem.* **52**, 761-799
38. Hakomori, S., Fukuda, M., and Nudelman, E. (1982) Role of cell surface carbohydrates in differentiation: Behavior of lactosaminoglycan in glycolipids and glycoproteins in *Teratocarcinoma and Embryonic Cell Interactions* (Muramatsu, T., Gachelin, G., Moscona, A., and Ikawa, Y., eds.) pp. 179-200, Academic Press, New York
39. Fukuda, M. (1991) Lysosomal membrane glycoprotein. *J. Biol. Chem.* **266**, 21327-21330
40. Lasky, L.A., Singer, M.S., Dowbenko, D., Imai, Y., Henzel, W.J., Grimley, C., Fennie, C., Gillett, N., Watson, S.R., and Rosen, S.D. (1992) An endothelial ligand for L-selectin is a novel mucin-like molecule. *Cell* **69**, 927-938
41. Riou, J.-F., Umbhauer, M., Shi, D.L., and Boucaut, J.-C. (1992) Tenascin: A potential modulator of cell-extracellular matrix interactions during vertebrate embryogenesis. *Biol. Cell.* **75**, 1-9
42. Simmons, D.L., Satterthwaite, A.B., Tenen, D.G., and Seed, B. (1992) Molecular cloning of a cDNA encoding CD34, a sialomucin of human hepatopoietic stem cells. *J. Immunol.* **148**, 267-271