THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1989 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 264, No. 12, Issue of April 25, pp. 6906–6912, 1989
Printed in U.S.A.

Identification of a Carcinoembryonic Antigen Gene Family in the Rat

ANALYSIS OF THE N-TERMINAL DOMAINS REVEALS IMMUNOGLOBULIN-LIKE, HYPERVARIABLE REGIONS*

(Received for publication, August 10, 1988)

Vitam Kodelja, Kurt Lucas, Sabine Barnert, Sabine von Kleist, John A. Thompson, and Wolfgang Zimmermann‡

From the Institut für Immunbiologie, Universität Freiburg, Stefan-Meier-Strasse 8, D-7800 Freiburg, Federal Republic of Germany

The existence of a carcinoembryonic antigen (CEA)-like gene family in rat has been demonstrated through isolation and sequencing of the N-terminal domain exons of presumably five discrete genes (rnCGM1-5). This finding will allow for the first time the study of functional and clinical aspects of the tumor marker CEA and related antigens in an animal model. Sequence comparison with the corresponding regions of members of the human CEA gene family revealed a relatively low similarity at the amino acid level, which indicates rapid divergence of the CEA gene family during evolution and explains the lack of cross-reactivity of rat CEA-like antigens with antibodies directed against human CEA. The N-terminal domains of the rat CEA-like proteins show structural similarity to immunoglobulin variable domains, including the presence of hypervariable regions, which points to a possible receptor function of the CEA family members. Although so far only one of the five rat CEA-like genes could be shown to be transcriptionally active, multiple mRNA species derived from other members of the rat CEA-like gene family have been found to be differentially expressed in rat placenta and liver.

CEA,[1] one of the most widely used human tumor markers, belongs to a highly conserved protein family (Shively and Beatty, 1985), the members of which are encoded by approximately ten genes (Thompson *et al.*, 1987). The complete primary structures of three members of the human CEA protein family as deduced from cDNA sequences have recently been reported: CEA (Oikawa *et al.*, 1987a; Beauchemin *et al.*, 1987; Zimmermann *et al.*, 1987), a nonspecific cross-reacting antigen (NCA) (Tawaragi *et al.*, 1988; Neumaier *et al.*, 1988), and a pregnancy-specific β_1-glycoprotein (PSβG) (Watanabe and Chou, 1988a, 1988b). Amino acid sequence analyses have revealed that the CEA gene family shows homology to and can be placed within the immunoglobulin supergene family (Paxton *et al.*, 1987, Oikawa *et al.*, 1987b; Williams, 1987). CEA, NCA, and PSβG can be subdivided into a number of domains: a 34-amino acid leader sequence, a 108–110-amino acid N-terminal domain, a highly conserved 178–180-amino acid repeating unit, of which three copies can be found in CEA, one and a half in PSβG and only one in NCA and a 26-amino acid hydrophobic carboxyl region in CEA, which is two amino acids shorter in NCA and is degenerate in PSβG. The corresponding domains show a high degree of sequence conservation between the three proteins, which obviously suggests a common ancestry (Thompson and Zimmermann, 1988).

The evolution of the CEA family presents a mystery, because analyses with polyclonal antisera, which recognize human CEA, NCA, and a number of other members of this family, have been unable to unequivocally identify their counterparts in nonprimate mammals (Wahren *et al.*, 1983). So far, CEA-related molecules could only be detected in higher primates (Haagensen *et al.*, 1982; Jantscheff *et al.*, 1986). The inability to recognize CEA-like antigens below the higher primates indicates either that such molecules do not exist in these species or that they have diverged rapidly during evolution. The extremely high degree of sequence conservation between CEA and an NCA (Thompson and Zimmermann, 1988), as well as the strong homology to another human gene family member (Watanabe and Chou, 1988b), would favor the former speculation. However, a number of reports indicate the existence of oncodevelopmentally regulated glycoproteins in rodents, with very similar biochemical properties to CEA, (Abeyounis and Milgrom, 1976; Howell *et al.*, 1979; Stevens *et al.*, 1975, 1976; Martin *et al.*, 1975a, 1975b; van Hove *et al.*, 1978). In order to clarify this problem, we have made an attempt to isolate CEA-like gene fragments from a rat genomic library. The existence of CEA-like molecules in the rat would allow many studies to be made in an animal system, *e.g.* studies on the oncodevelopmental regulation of CEA gene expression, which cannot be carried out in humans.

MATERIALS AND METHODS

Tissues—Rat tissues were obtained from the strain BD II (Druckrey, 1971). The anesthetized rats were killed by cervical dislocation, and various tissues were immediately removed, washed in cold phosphate-buffered saline, frozen in liquid nitrogen, and stored at −140 °C.

Isolation of pNCA1 and pCEA5—The NCA cDNA clone pNCA1 and the CEA cDNA clone pCEA5 were isolated from a cDNA library, which had been prepared from human colon tumor mRNA (Zimmermann *et al.*, 1987). The library was screened with the 2.7-kb *Eco*RI DNA fragment from clone λ39.2, which contains the exon coding for the N-terminal domain of an NCA (Thompson *et al.*, 1987).

Screening of the Rat Gene Library—For the isolation of members of the CEA gene family, a genomic library derived from rat liver DNA was used. The library had been constructed by partial digestion of

* This work was supported by a grant from the Dr. Mildred Scheel-Stiftung für Krebsforschung. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) M22226–M22230.

‡ To whom correspondence and reprint requests should be addressed.

[1] The abbreviations used are: CEA, carcinoembryonic antigen; NCA, nonspecific cross-reacting antigen; PSβG, pregnancy-specific β_1-glycoprotein; kb, kilobase pair(s).

the genomic DNA with *Sau*3A and cloning of the DNA fragments into the *Bam*HI of the λ-phage vector EMBL3 (Shinomiya *et al.*, 1984, Frischauf *et al.*, 1983). The recombinant phages were plated onto Q359 bacteria (Karn *et al.*, 1983), transferred to nitrocellulose filters (Schleicher and Schüll, Federal Republic of Germany) in replicas and hybridized with CEA and NCA cDNA fragments labeled by random hexanucleotide priming (Feinberg and Vogelstein, 1983) at 37 °C in the presence of 40% formamide, 5 × Denhardt's solution (1 × = 0.02% each of Ficoll (Pharmacia, Federal Republic of Germany), polyvinylpyrrolidone, bovine serum albumin), 5 × SSPE (1 × SSPE = 0.18 M NaCl, 10 mM sodium phosphate, pH 7.4, 1 mM EDTA), 0.1% sodium dodecyl sulfate, 100 μg of heat-denatured calf thymus DNA/ml. After hybridization overnight, the filters were washed twice for 30 min each in 2 × SSPE, 0.1% sodium dodecyl sulfate at room temperature and at 60 °C, respectively. Positive plaques were isolated and plaque-purified twice (Maniatis *et al.*, 1982).

DNA Sequencing and Sequence Analysis—For sequencing, exon-containing subfragments of the recombinant phage DNAs were identified. After digestion with various restriction endonucleases, the resulting DNA fragments were electrophoretically separated, blotted onto a nylon membrane (GeneScreen *Plus*, New England Nuclear, Federal Republic of Germany), and hybridized with the same CEA and NCA cDNA probes used for the isolation of the genomic clones under the conditions described above. Suitable hybridizing genomic DNA and cDNA fragments were subcloned into Bluescript (Stratagene, La Jolla, CA) or M13 vectors. Sequencing was performed on single- or double-stranded templates according to Sanger (Sanger *et al.*, 1977), using universal or internal oligonucleotide primers. The oligonucleotides were synthesized by the phosphoramidite method on an Applied Biosystems 308A DNA synthesizer (Applied Biosystems, Weiterstadt, Federal Republic of Germany). The oligonucleotides still carrying the trityl group at the 5′ end were purified by high pressure liquid chromatography on a C1 Ultropac TSK TMS-250 column (LKB, Freiburg, Federal Republic of Germany) using a gradient of acetonitrile (10–25%) in 0.1 M triethylammonium acetate, pH 7, for elution. After deprotection and removal of the trityl group according to the manufacturer's protocol, the oligonucleotides could be used directly for sequencing.

For comparison of the nucleotide and amino acid sequences, the computer program "Align[2]" was used. For the determination of the similarity between amino acid sequences, pairs of amino acids with logarithms of odd matrix scores ≥ 9 (Feng *et al.*, 1985) were taken as being conservatively exchanged. Secondary structure calculations were performed with the computer program "Novotny" (PCgene, Genofit, Switzerland).

RNA Blot Hybridization—Isolation and analysis of RNA by Northern blot hybridization was performed essentially as described before (Zimmermann *et al.*, 1988). However, enrichment for poly(A) RNA was achieved by only one round of chromatography on oligo(dT)-cellulose.

RESULTS

Isolation of Genomic Clones Encoding Members of the Rat CEA Gene Family—To demonstrate the existence of a CEA-like gene family in the rat, we used a mixture of cDNA fragments, coding for regions of the human CEA and an NCA, as probes for Southern blot analyses. The location of the CEA and NCA cDNAs with respect to their mRNAs are shown in Fig. 1*A*. The identity and location of the cDNA inserts of pCEA5 and pNCA1 had been established by sequencing (for the sequencing strategy see Fig. 1*A*). The nucleotide sequences of both clones are identical with recently published sequences of full-length CEA and NCA cDNA clones (Oikawa *et al.*, 1987a; Beauchemin *et al.*, 1987; Tawaragi *et al.*, 1988; Neumaier *et al.*, 1988) except for one nucleotide difference (T instead of C) in position 519 of CEA cDNA (Beauchemin *et al.*, 1987) which does not, however, lead to an amino acid change. Hybridization of restriction endonuclease digests of rat genomic DNA with the above-mentioned cDNA fragments under nonstringent conditions yielded very weak signals.

The same probes were now used under identical hybridization conditions to screen 5 × 10^5 recombinant phages of a rat liver genomic library (Shinomiya *et al.*, 1984). Thirteen positive clones were obtained. The DNA of six phage recombinants was further characterized by restriction endonuclease analyses. Comparison of the size patterns of the DNA fragments revealed that two of these clones were identical and two others appeared to overlap. Thus, these genomic clones are apparently derived from four different chromosomal loci of the rat.

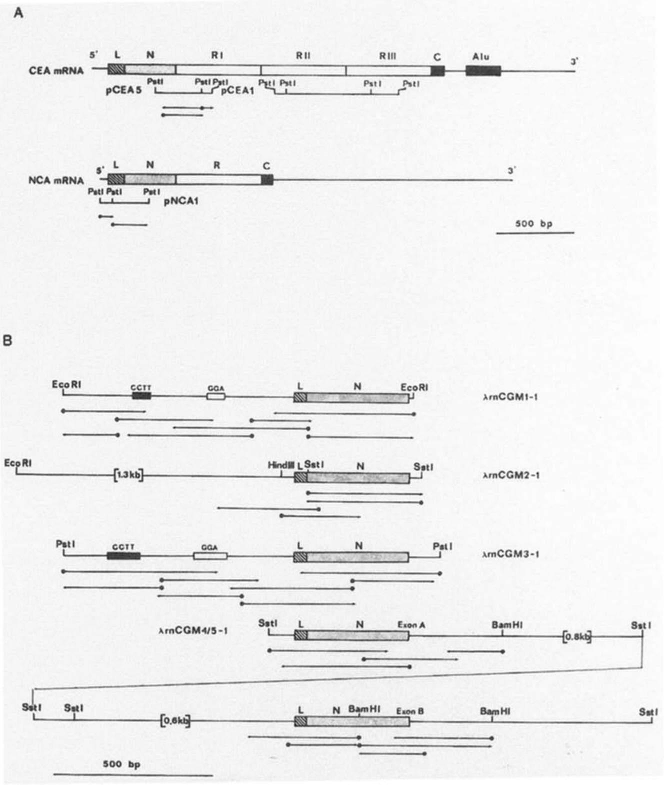

FIG. 1. *A*, structure of CEA and NCA mRNA and location of cDNA probes. A graphic representation of CEA (Beauchemin *et al.*, 1987) and NCA mRNA (Neumaier *et al.*, 1988) is shown. Homologous coding regions are depicted by the same *shading*. The coding regions are composed of a leader peptide (*L*), an N-terminal region (*N*), repeated domains (*R, RI, RII, RIII*), and a hydrophobic C-terminal region (*C*). The *box* in the 3′-untranslated region of CEA mRNA represents the truncated Alu sequence (Zimmermann *et al.*, 1987). Below, the length and location of CEA cDNA clone pCEA1 (Zimmermann *et al.*, 1987) and pCEA5 as well as the NCA cDNA clone pNCA1 are indicated. *B*, structure of the subfragments of four genomic clones containing the N-terminal domain exons of rat CEA-like genes. Leader (*L*) and N-terminal domain (*N*) encoded by the N-terminal domain exon are depicted by *hatched* and *stippled boxes*, respectively. *Solid* and *open boxes* symbolize simple repeated sequences with d(CCTT) and d(GGA) units, respectively. For graphical purposes, some fragments were shortened. The length of the regions left out is shown in *brackets*. The physical linkage of the two *Sst*I fragments of λrnCGM4/5-1 which contains the N-terminal domain exons of presumably two genes (rnCGM4 = exon A and rnCGM5 = exon B) is indicated by a *dotted line*. In *A* and *B*, only the restriction endonuclease sites relevant for subcloning and sequencing are shown. The sequencing strategy is indicated by *arrows*.

Sequence Analysis of the N-terminal Domain Exons and Flanking Introns of Rat CEA-like Genes—We decided to analyze the exons encoding the N-terminal domain in order to count the number of different CEA-like genes in the rat, because at present, analysis of human CEA, NCA, and PSβG cDNA implies that in contrast to the repeat domain, only one N-terminal domain exon exists per gene. A total of five subfragments were found to hybridize with a probe encoding part of the N-terminal domain of a human NCA, one each in clones of λrnCGM1-1, λrnCGM2-1, and λrnCGM3-1 and two

[2] M. Trippel and R. Friedrich, unpublished data.

rnCGM 1

```
   1 GAATTCACTCCTCAGCTCTCACAGCATAGATGGACATACAGACTCCTGAAGGCTCTTCTCTTCCCTCCACACTGGTGTGTGTCACGTACCTGTAGTGTGCACACTGGGACATGTACCTTCCCAAACCCTCACGAACAATACAGAAATATT

 151 AAATTACACTTGAATATAATTATTTTTATGTGCTATAAACATGGAAATTATGTAGACAAACCCAGAGATATCTTTTCTTCCTTCCTTCCTTCCTTCTTCCTTCCTTCCTTCCTCTTTTTCCATACTAGTTTCTGAGATTTTTTGAG

 301 GAACTGAACCTTCCAAAAAGACCATACCAATCCCTGTCCTCAAAAAGCCTTTTTTATTCTAATGGACTGGAAATCATTGTATCCAGAGGAGAAAGTCAATGATTTAGTGGAACCATAAATAGAACAGAAAACATTCAGGAAGTGAGGATT

 451 GTATGGAGGAGGAAAAAGAGGAGGAGGAGGAGGAAGAGGAGGAGGAGGAGGAGGAGGACCGAGAGCCGGTTCTCCACTCACCAGACACTTTATGGAAAGAGTGATATGGGGACACCTGAGTAGAGGATTCCACAGAGAGGAAATGACACC

 601 CTTTGAGGTTCTGAGGGCATGGAGGTCATGCTGCTCACCTCCATTAAGGGTGCATCCTACCTACAGGCTGAGGGATGCTCACACCTGCTCAGGATTGTCAACTTTTCTCTCTTCCCTTCTAGCCTCCCTCTTAACCTGCTGGCTCCTGCC
                                                                                                                                 SerLeuLeuThrCysTrpLeuLeuPr

 751 CACCACTGCCCAAGTCTCCATTGAATCCTTACCACCCCAGGTGGTTGAAGGAGAAAATGTTCTTCTACGTGTTGACAATTTGCCAGAGAATCTCATAGCCTTTGTCTGGTACAAAGGGCTGACAAACATGAGCCTCGGAGTTGCACTGTA
     oThrThrAlaGlnValSerIleGluSerLeuProProGlnValValGluGlyGluAsnValLeuLeuArgValAspAsnLeuProGluAsnLeuIleAlaPheValTrpTyrLysGlyLeuThrAsnMetSerLeuGlyValAlaLeuTy

 901 TTCACTAACCTATAACGTAACTGTGACGGGACCTGTGCACAGTGGTAGAGAGACATTGTACAGCAATGGGTCCCTGTGGATCCAAAATGTCACCCAGAAGGACACAGGATTCTACACCCTACGAACCATAAGTAATCATGGAGAAATTGT
     rSerLeuThrTyrAsnValThrValThrGlyProValHisSerGlyArgGluThrLeuTyrSerAsnGlySerLeuTrpIleGlnAsnValThrGlnLysAspThrGlyPheTyrThrLeuArgThrIleSerAsnHisGlyGluIleVa

1051 ATCAAATACATCCCTGCACCTTCATGTGTACTGTAAGTAATTCTTTGTGAATTC
     lSerAsnThrSerLeuHisLeuHisValTyr
```

rnCGM 2

```
   1 AATGTAATTCTTGTTGGAGAGTGAGTGGGGAGCCATGCAGACACGGGAGGAGAGAGACCCGTACAAAAGGTCACTCCAGCTTCGGGGGACTGGGAACATAGATGATGAAGTTTCCCTGCACCAATGAGAGCGACGCCCTCACCCCACACC

 151 TCGGCAGAAGATGAACACACCTACCTGTTCCGGACTTGGGCCTCCTCTCAGCGATCACTAAGCTTCTGACACTGATGGAGTTTTTTCCTTCTCCCTAGCTTCCTTCTTAACCTGCTGGAATGCACCCGCCGCTGCCGAGCTCACTATTGA
                                                                                                         SerPheLeuThrCysTrpAsnAlaProAlaAlaAlaGluLeuThrIleGl

 301 ATTAGTGCCACCCATGGTTGCTGAAGGCGGAAACTCCGTTTTGTTTGTGCATGAAATGCCATTGAATGTCCAGGCGTTTTACTGGTACAAACAGAGAGATCCGACGAAGAGCTATGAAGTCGCGCGGTACTTAACACCCACCAACGAAAG
     uLeuValProProMetValAlaGluGlyGlyAsnSerValLeuPheValHisGluMetProLeuAsnValGlnAlaPheTyrTrpTyrLysGlnArgAspProThrLysSerTyrGluValAlaArgTyrLeuThrProThrAsnGluSe

 451 TTCGAAGATGCCTCAGCACAGCGGCCGGAAAACCGTATTCTACAGTGGATCCCTGCTGATCAGAAACGTCACCCAGGCCGACAGTGGAGTCTACACCTTACTAACATTTAACACAGAAATGCAAAGCGAATTAACACATGTGCATCTGGA
     rSerLysMetProGlnHisSerGlyArgLysThrValPheTyrSerGlySerLeuLeuIleArgAsnValThrGlnAlaAspSerGlyValTyrThrLeuLeuThrPheAsnThrGluMetGlnSerGluLeuThrHisValHisLeuGl

 601 AGTACGCGGTAGGTGGTTGCGGGATCTCTGGGTGCTAGGGGTCGGGGTGAGCTC
     uValArg
```

rnCGM 3

```
   1 CTGCAGTGTGCACAGCAAGACATTGTGCTTTCCCGAACCCACACGAACACACTGAATTATTAAATCACACTTGAATATATTGATTTCCCTTTGCTCTGAGCCTGGGCACTATGTAGATAAGTCCATGGAAATATTAATCTTTCCTTCCTT

 151 CCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTTCTTCCTTCTTCTAGTTCTTTTCACGTTTTCCCTTTTCTTTTTCTCTCCAATTTGTTTCTAATCTATTTTCAGGAACTGAACCTTCCAAAAAGATGATTCCAGTCCCTGT

 301 CCTCACAAAGCCCTTTTCTTGTGGACTGGAAGTCAGAGTATCCAGAGAAAGGCAATGGTTTAATGGAACCTCAAACAGAACAGAAAACAATTCTGAGAGTGAGCATTGCATGAGGAAGAGGAGGAACGGGAAGAGGAGGAGGAAGAGGAG

 451 GAGGAAGAGGAGGAGGAAGAGGAATGGGAAGAGGAGGAGGAAGAGGAGGAAGAGGAGGAAGAGGTCAGACAGCTGCTTCACCTCTCACCAGACACTCTATGGGAAGAATGATATGGGGACACCTGAGTAGAGGATTCCTGGAGAGGAAAT

 601 GACAGCTTTTGAGTCTTTGAGGGCATGGAGGTCATGCTGCTCACCTCCATTAAGGGTGCATCCTACCTACAGGCTGAGGGATGCTCACACCTGCTCAGGATCGGTGACTTTTTTCTCTTCCCTTCTAGCCTCACTTTTAACCTGCTGGCT
                                                                                                                                       SerLeuLeuThrCysTrpLe

 751 CCTGCCCACCACTGCCCACGTCACCCTCAAGTCCTCACCGCCCCAGGTGGTTGAAGGAGAAAACGTTCTTCTAAGTGCTGACAATCTGCCAGAGAACATTATAGCTTTCGCCTGGTACAAAGGGGAGACCGACATGAACCGTGGAATTGC
     uLeuProThrThrAlaHisValThrLeuLysSerSerProProGlnValValGluGlyGluAsnValLeuLeuSerAlaAspAsnLeuProGluAsnIleIleAlaPheAlaTrpTyrLysGlyGluThrAspMetAsnArgGlyIleAl

 901 ACTGTATTCACTGAGGTATACTGTAAGTTTGACGGGGCCTGTGCACAGTGGTCGAGAGACATTGTACAGCGACGGGTCCCTGTGGATCAAAAATGTCACCCAGGAGGACACAGGATTTTATACCTTTCGAATCATAAATAATCATGGAAA
     aLeuTyrSerLeuArgTyrThrValSerLeuThrGlyProValHisSerGlyArgGluThrLeuTyrSerAspGlySerLeuTrpIleLysAsnValThrGlnGluAspThrGlyPheTyrThrPheArgIleIleAsnAsnHisGlyLy

1051 AATTCAATCAAATACAACCCTGTTCCTTCACGTGAAATGTAAGTAACTCTTTGTGAACTGTGGGTTTTGGGTGGTGTCCTTCCACTAGACACATAGAAGTATCAGGCCAGGGCTGTGTCTCCCTTCCCCCTGCAG
     sIleGlnSerAsnThrThrLeuPheLeuHisValLys
```

rnCGM 4

```
   1 GAGCTCTGGGAAGGCAGAAGTGTGATTTTTTAAAAAACCAACAGATTTCACCTGCTCAATATCGATGGTTGCTCTGTCTTCCCTTTTAGCCTCCCTTCTAACCTGTTGGCTCCTGACTACTGCCCAGGTCAACATTGAATCGGTGCCATT
                                                                                                 SerLeuLeuThrCysTrpLeuLeuThrThrAlaGlnValAsnIleGluSerValProPh

 151 CAATGTGGTTGAAGGGGAAAACGTCCTTCTTCTTGTCCACAATCTGCCAGAGAATCTCATAGCCTTTGCCTGGTATAGAGGGCTGAGGAAAATTGGAGTATACATACTGAACACTGAAGTAAGTGTGACGGGGCCAATGTACAGCGGTAG
     eAsnValValGluGlyGluAsnValLeuLeuLeuValHisAsnLeuProGluAsnLeuIleAlaPheAlaTrpTyrArgGlyLeuArgLysIleGlyValTyrIleLeuAsnThrGluValSerValThrGlyProMetTyrSerGlyAr

 301 AGAGACAGTGTACAGCAATGGTTCCCTGTGTATCCGCAATGTCACCCAGAAGGACACAGGATTCTACACTCTACGAACAGTCAACACACGTGGAGAAACTGTATCAACAACATCCTTGTACCTCTATGTGTACAGTAAGTGATACTTTGT
     gGluThrValTyrSerAsnGlySerLeuCysIleArgAsnValThrGlnLysAspThrGlyPheTyrThrLeuArgThrValAsnThrArgGlyGluThrValSerThrThrSerLeuTyrLeuTyrValTyr

 451 GAACTCTGGGTGTTGTGTGGGGTTCATTCCGTAGACACACACAGAAGAGGCAGGCCTACCTACCCTTTGCATTGTGTCTCCTTATTGAGGTGTGAACATTTAACTCAGGCTAAGGAGAGTAATGCCAATTGAATAGAATCCTTCTTTTGA

 601 CTTTACCTTGTAGTCAGCTGGATGTGTGGTTAACTCAGTGAAGGACATCAGCCCTTGTCTAGACTTCTGGGGTTCTTAGCAGTAATGTGTCCTTGGGAAAGACCTTGAGGGAAGGAGATTGGGTTTGAATGAGATAGCCATAGGATCC
```

rnCGM 5

```
   1 AGCGTAGGCAGGAGACTCCACACCTCAGCTGACCACTGGACACAGCTGCTCGGACTCAGGCACCATCTTAGCCAAATACTAAAGTCCTGATGTTGACGGATCTCTCTTCCCTTCTAGCCTCTCTTTTCATCTGTGGGCGTCCTTTTAACC
                                                                                                                            SerLeuPheIleCysGlyArgProPheAsnP

 151 CTGCCAAGCTCACTATTGAATCAGTGCCGCCCAGTGTTGCTGAAGGGGGAAGCGTTCTTCTCCTCGTTCACAATCTCCAGGACGAGCTTCGAGGGTTTTTCTGGTACAAAGGGGCGTCTATGTCTAGCAACCATGAGATAGCCCGATACA
     roAlaLysLeuThrIleGluSerValProProSerValAlaGluGlyGlySerValLeuLeuLeuValHisAsnLeuGlnAspGluLeuArgGlyPhePheTrpTyrLysGlyAlaSerMetSerSerAsnHisGluIleAlaArgTyrA

 301 GAACAGCAAAGAATTCAAGTGTGCCAGGCCCTGCCCACAGTGGTAGAGAGACGGTGTACAGCAATGGATCCCTCCTGCTCCAGAATGTCACCCGGAATGACACTGGGTTCTACACCCTACGCACTCTGAAAAGACATCAGAAAATGGAAT
     rgThrAlaLysAsnSerSerValProGlyProAlaHisSerGlyArgGluThrValTyrSerAsnGlySerLeuLeuLeuGlnAsnValThrArgAsnAspThrGlyPheTyrThrLeuArgThrLeuLysArgHisGlnLysMetGluL

 451 TGGCACACGTGCAACTTCAGGTGGACAGTAAGTGATTTTCCGTGATCGTTCAGTGCTGGGTGGGTCTTTGACACACAGGACTGTCACCCCTGGCATGTGGCTACCTCCTCTCTGCCTTTTTATCCCCATGTTGTGGTTAACCACTATGTG
     euAlaHisValGlnLeuGlnValAsp

 601 CAGGACACATGTGATGGAAAGAAATGCCCATGGGTCAGACTTATCATCTGACTCTCCCCTGTATCAAGGACAGTAACTCAACCCTAGGTGCTAGACTCTGCCCAGTCATCTGGGGCATCTTGCCATGCAACGTGAGGAAACCATGGATCC
```

FIG. 2. **Nucleotide and predicted amino acid sequences of the N-terminal domain exons and flanking intron regions of rat CEA-like genes.** The sequences of the N-terminal domain exons of presumably five separate CEA-like rat genes rnCGM1, rnCGM2, rnCGM3, rnCGM4, and rnCGM5 are shown. Exon/intron borders are indicated by *arrows*. The simple repeated sequences found upstream of the N-terminal domain exons of clones λrnCGM1-1 and λrnCGM3-1 are *underlined*. The 25-base pair direct repeated within the $[d(GGA)]_n$ sequence of λrnCGM3-1 is indicated by *dotted lines*.

in clone λrnCGM4/5-1 (Fig. 1*B*). These fragments were subcloned and sequenced according to the strategy shown in Fig. 1*B*. The nucleotide and the deduced amino acid sequences of the N-terminal domain exons are presented in Fig. 2. In each genomic fragment, one open reading frame flanked by canonical splice acceptor and donor sequences (Mount, 1982) could be identified (Fig. 2). The overall similarity among these homologous rat exons lies between 56 and 84% at the nucleotide level and between 39 and 76% at the amino acid level (Fig. 3). This rather low degree of conservation of the amino acid sequence among the different members of the rat CEA gene family increases strongly, if conservative exchanges are allowed (Fig. 4).

The deduced amino acid sequences of the N-terminal domains contain two to five putative *N*-glycosylation sites (consensus sequence: Asn-*X*-Ser/Thr, $X \neq$ Pro), only one of which is conserved in all five sequences (Fig. 4). The presence of multiple glycosylation sites implies that at least some of the putative CEA-like antigens in the rat are as heavily glycosylated as the human CEA, where 50–60% of its total mass is carbohydrate (Terry *et al.*, 1974).

Amino Acid Level												
	rnCGM1	rnCGM2	rnCGM3	rnCGM4	rnCGM5	CEA	NCA	PSβG	hsCGM1	hsCGM2	hsCGM3	hsCGM4
rnCGM1	▨	45	76	73	49	45	49	47	45	43	50	47
rnCGM2	60	▨	39	47	46	35	38	34	36	36	35	35
rnCGM3	84	56	▨	64	45	44	48	45	43	39	47	45
rnCGM4	80	62	72	▨	51	51	54	50	51	48	53	50
rnCGM5	65	60	61	65	▨	50	51	45	50	46	45	45
CEA	61	54	60	65	62	▨	90	57	89	66	58	58
NCA	63	55	61	65	62	94	▨	57	92	66	58	57
PSβG	65	52	62	62	58	73	71	▨	58	52	89	89
hsCGM1	64	55	61	66	63	95	94	73	▨	64	59	58
hsCGM2	58	53	58	62	59	81	81	69	82	▨	52	51
hsCGM3	64	52	60	63	56	73	72	93	73	68	▨	91
hsCGM4	64	52	60	62	57	72	70	93	72	67	93	▨
Nucleotide Level												

FIG. 3. **Comparison of the N-terminal domain exons of rat and human CEA-like genes at the nucleotide and amino acid levels.** The degrees of similarity in percent between the N-terminal domain exons of rat and human CEA-like genes were calculated after optimal alignment. The corresponding human sequences were derived from cDNA (CEA; Beauchemin *et al.*, 1987) or genomic DNA sequences (NCA, PSβG, hsCGM1-4; Thompson *et al.*, 1987 and Footnote 3).

The similarity between the 5′ flanking intron sequences of the different N-terminal domain exons is, with the exception of rnCGM1 and rnCGM3, in general very low. In the 5′-flanking regions of the N-terminal domain exon of clones λrnCGM1-1 and λrnCGM3-1, simple repeated sequences ($[d(GGA)]_n$ or $[d(AGA)]_n$ and $[d(CCTT)]_n$) of varying length are found (Figs. 1*B* and 2). These purine- and pyrimidine-rich sequences are located about 300 and 550 nucleotides, respectively, upstream from the start of the N-terminal exons.

Expression of Members of the Rat CEA Gene Family—In order to identify tissues where genes of the rat CEA gene family are transcribed, we screened, in a preliminary experiment, a number of fetal and adult rat tissues by slot blot hybridization. The 1.1-kb *Eco*RI fragment of λrnCGM1-1 containing the N-terminal domain exon (Fig. 1*B*) was used as a probe under nonstringent hybridization conditions (2 × SSPE, 60 °C). With RNA from placenta, one of the most prominent hybridization signals was obtained (data not shown). We, therefore, hybridized size-fractionated poly(A) RNA from rat placenta and, for comparison, poly(A) RNA from adult liver, with DNA fragments covering all five N-terminal domain exons. With the probe from λrnCGM1-1, three major mRNA species with lengths of 3.9, 3.2, and 2.5 kb could be detected, whereas the probes from the N-terminal exons of λrnCGM3-1 and λrnCGM4/5-1, exon A, hybridized strongly with the smallest of the three mRNA species only and to a lesser extent with a 3.0-kb species (Fig. 5, *lanes 2, 6*, and *8*). After prolonged exposure of the RNA blot hybridized with the probe from λrnCGM1-1, CEA-related mRNA species with lengths of 4.6, 3.9, 3.0, and 1.9 kb could be shown to exist in liver (Fig. 5). All observed RNA/DNA hybrids were unstable under more stringent (0.5 × SSPE, 65 °C) washing conditions (data not shown). Essentially, the same results

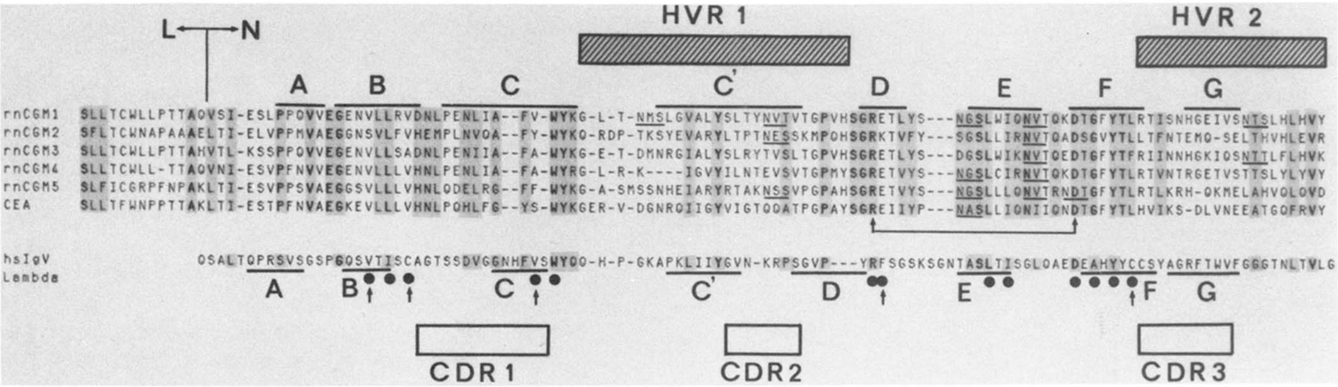
L N
HVR 1
HVR 2
A B C C′ D E F G
rnCGM1 SLLTCWLLPTTAQVSI-ESLPPQVVEGENVLLRVDNLPENLIA--FV-WYKG-L-T-NMSLGVALYSLTYNVTVTGPVHSGRETLYS---NGSLWIQNVTQKDTGFYTLRTISNHGEIVSNTSLHLHVY
rnCGM2 SFLTCWNAPAAAELTI-ELVPPMVAEGGNSVLFVHEMPLNVQA--FY-WYKQ-RDP-TKSYEVARYLTPTNESSKMPQHSGRKTVFY---SGSLLIRNVTQADSGVYTLLTFNTEMQ-SELTHVHLEVR
rnCGM3 SLLTCWLLPTTAHVTL-KSSPPQVVEGENVLLSADNLPENIIA--FA-WYKG-E-T-DMNRGIALYSLRYTVSLTGPVHSGRETLYS---DGSLWIKNVTQEDTGFYTFRIINNHGKIQSNTTLFLHVK
rnCGM4 SLLTCWLL-TTAQVNI-ESVPFNVVEGENVLLLVHNLPENLIA--FA-WYRG-L-R-K----IGVYILNTEVSVTGPMYSGRETVYS---NGSLCIRNVTQKDTGFYTLRTVNTRGETVSTTSLYLYVY
rnCGM5 SLFICGRPFNPAKLTI-ESVPPSVAEGGSVLLLVHNLQDELRG--FF-WYKG-A-SMSSNHEIARYRTAKNSSVPGPAHSGRETVYS---NGSLLLQNVTRNDTGFYTLRTLKRH-QKMELAHVQLQVD
CEA SLLTFWNPPTTAKLTI-ESTPFNVAEGKEVLLLVHNLPQHLFG--YS-WYKGER-V-DGNRQIIGYVIGTQQATPGPAYSGREIIYP---NASLLIQNIIQNDTGFYTLHVIKS-DLVNEEATGQFRVY
hsIgV
Lambda QSALTQPRSVSGSPGQSVTISCAGTSSDVGGNHFVSWYQQ-H-P-GKAPKLIIYGVN-KRPSGVP---YRFSGSKSGNTASLTISGLQAEDEAHYYCCSYAGRFTWVFGGGTNLTVLG
A B C C′ D E F G
CDR1
CDR2
CDR3

FIG. 4. **Comparison of the predicted amino acid sequences encoded by the rat N-terminal domain exons with the corresponding sequences of CEA (Beauchemin *et al.*, 1987) and of the variable domain of a human immunoglobulin λ chain (hsIgV λ; Köhler *et al.*, 1975).** Amino acids are identified by the single letter code. *Dashes* denote missing amino acids and were inserted to achieve optimal alignment of the compared sequences. Amino acids identical in the sequences of all CEA family members are printed in *bold letters*. Positions which contain conserved or conservatively exchanged amino acids with respect to CEA (using the system of Feng *et al.*, 1985) in all CEA family members are *shadowed gray*. Residues which are conserved in all seven sequences are similarly marked in the immunoglobulin sequence. The border between the leader (*L*) and the N-terminal domain (*N*) is indicated. The position and extent of the calculated β-strands (A–G) of rnCGM1 and hsIgV λ are indicated by over- and underlining, respectively. Residues which are highly conserved in the variable domain of the immunoglobulin superfamily are labeled by *solid circles*. The ones not conserved in the CEA family are marked with *arrows*. The location of the complementarity determining regions (*CDR*) of the variable domains of immunoglobulin light chains (Kabat *et al.*, 1987) and the hypervariable region of the N-terminal domain of the members of the CEA family (*HVR*) are shown by *open* and *hatched boxes*, respectively. Potential *N*-glycosylation sites are marked by *underlining*.

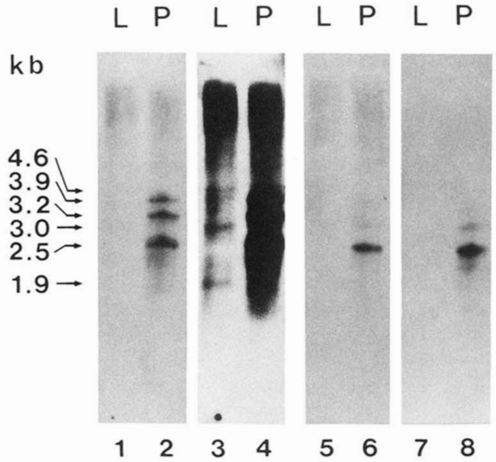

FIG. 5. **Identification of CEA-like mRNA species in rat tissues.** Two micrograms of poly(A)$^+$ RNA from rat liver (*L*) or 21 day placenta (*P*) were size-separated on a 1% agarose/methylmercury hydroxide gel by electrophoresis. The RNA was transferred to a charged nylon membrane and hybridized with ^{32}P-labeled DNA fragments containing the N-terminal domain exons. Either the complete DNA fragments as shown in Fig. 1*B* (λrnCGM1-1, *lanes 1–4*; λrnCGM3-1, *lanes 5* and *6*) or an *Sst*I/*Bam*HI subfragment from λrnCGM4/5-1 containing exon A (*lanes 7* and *8*) were used as probes. After washing in 2 × SSPE at 60 °C, the membranes were exposed to x-ray films for 6 h (*lanes 1, 2, 5,* and *6*). In order to visualize minor mRNA species or weak cross-hybridization, the membranes were reexposed for 2 days (*lanes 3, 4, 7,* and *8*). The smear in the high molecular weight range in *lanes 3* and *4* is probably due to binding of the probes to DNA present in these RNA preparations. The *numbers* on the *left* indicate the size (in kilobases) of the hybridizing mRNAs.

were obtained when probes lacking the simple repeated sequences described above were used (data not shown). The exon-containing fragments from λrnCGM2-1 (the 0.4-kb *Sst*I fragment) and λrnCGM4/5-1 (the 1.5-kb *Sst*I/*Bam*HI fragment with part of exon B) (Fig. 1*B*) did not hybridize at all under the same hybridization conditions (data not shown).

DISCUSSION

In this paper, we present data that demonstrate the existence of a CEA-like gene family in the rat of similar structure as its counterpart in humans (Fig. 4). We have sequenced five different N-terminal domain exons, which show very similar length and exact conservation of the exon/intron borders when compared with the corresponding exons of the various members of the human CEA gene family (Thompson *et al.*, 1987; Oikawa *et al.*, 1987c; Footnote 3). At present, we assume, in analogy to the human CEA gene family, that only one N-terminal domain is contained within each CEA-like antigen of the rat (Beauchemin *et al.*, 1987; Oikawa *et al.*, 1987a; Neumaier *et al.*, 1988; Tawaragi *et al.*, 1988). The five exons, therefore, presumably represent five separate genes. Until individual CEA-like antigens have been characterized in the rat, we have proposed a temporary nomenclature system for these as well as other genes (Thompson and Zimmermann, 1988). Each gene has been designated numerically according to its species (*Rattus norvegicus* = rn) as a CEA gene-family member (CGM; *e.g.* rnCGM1). Until now, the presence of CEA-like genes in rodents was not clear, because CEA-like antigens could not be detected with antibodies directed against human CEA in species below the higher primates (Wahren *et al.*, 1983). Taking into account that most epitopes are on the protein moiety of CEA (Hammarström *et al.*, 1975) and the low degree of amino acid sequence conservation between the rat and human members of the CEA gene family (Fig. 3), the lack of immunological cross-reactivity between rat CEA-like antigens and antisera against human CEA is not surprising (Johnson *et al.*, 1985). However, further experiments have to be carried out to prove that the CEA-like antigens described before (Martin *et al.*, 1975a; Stevens *et al.*, 1975, 1976; Abeyounis and Milgrom, 1976; van Hove *et al.*, 1978) indeed belong to the rat CEA family.

RNA/DNA hybridization studies showed that at least some genes of this family are transcriptionally active in rat tissues. Six distinct mRNA species have been found in rat placenta and liver. The expression seems to be controlled in a tissue-specific manner and shows quantitative and qualitative differences. In placenta, the mRNA species with lengths of 3.9, 3.2, and 2.5 kb are very prominent. The latter two mRNAs are not present in liver, where, however, two additional mRNAs of 4.6 and 1.9 kb are found. All CEA-like mRNAs seen in liver are expressed at a comparatively low level. At present, these data cannot be interpreted extensively, because under stringent hybridization conditions, no transcripts could be visualized for any of the genes characterized in this paper in liver or placenta. Recently, however, we have isolated a clone from a rat placental cDNA library, and partial sequence analysis reveals identity to rnCGM1, proving this gene to be transcribed in placenta.[4] As possible explanations for this apparent contradiction, we assume that this gene is either expressed at a level too low to be visualized in DNA/RNA hybridization analyses or that it is differentially expressed during placental development. Further experiments have to be performed to prove that the other genes are active or represent pseudogenes.

When the nucleotide and derived amino acid sequences of the N-terminal domain exons of the rat CEA-like genes are compared with the corresponding exons of the human genes, a generally low sequence conservation is observed (see Fig. 3). The highest similarity at the amino acid level (54%) is found between NCA and rnCGM4 (Fig. 3). In contrast, most of the human exons and some of the rat exons show a higher degree of sequence conservation within a species (Fig. 3). For this reason, it is not possible to assign individual genes to their counterpart in the other species. In a further study, Southern analyses of human and rat genomic DNA probed with various exon-containing fragments indicate that the sequences of all members of the rat CEA-like gene family show strong divergence to their counterparts in man.[5] This may create problems in finding analogous genes in the two species by sequence comparison alone.

The N-terminal domain exons identified so far in the rat, reveal a maximal similarity of 76% at the amino acid level (rnCGM1/rnCGM3), which is much lower than the similarity among closely related CEA-like genes in man (up to 92%; Fig. 3). These data and the lack of strongly hybridizing genomic DNA fragments in the rat with rat exon-containing probes[5] argue against the existence of subgroups with closely related members in this species as has been found in the human CEA gene family.[3]

Simple repeated DNA sequences, which might serve as hot spots of recombination (Cohen *et al.*, 1982) or trigger gene conversion (Slightom *et al.*, 1980), are found upstream of the N-terminal domain exons of rnCGM1 and rnCGM3 (Fig. 2).

[3] Thompson, J. A., Mauch, E. M., Chen, F. S., Hinoda, Y., Schrewe, H., Ortlieb, B., Barnert, S., Von Kleist, S., Shively, J. E., and Zimmerman, W. (1989) *Biochem. Biophys. Res. Commun.,* in press.

[4] S. Rebstock, J. A. Thompson, and W. Zimmerman, unpublished data.

[5] Rudert, F., Zimmerman, W., and Thompson, J. A. (1989) *J. Mol. Evol.*, in press.

They consist of homopurine/homopyrimidine stretches with d(CCTT) and d(GGA) units, which are repeated to a different extent in the two genes. This type of sequence is quite abundant in the mammalian genome and is found in association with various genes (Müller *et al.*, 1987; Kelly and Trowsdale, 1985; van den Heuvel *et al.*, 1985; Delaey *et al.*, 1987; Cohen *et al.*, 1982). Deviations from the basic motif of the simple sequences in the rat CEA-like genes are mainly caused by transitions, so that the homogeneity of these homopurine/homopyrimidine stretches is not disrupted. This might be important for the still hypothetical function of such sequences, which are known to adopt an open conformation under superhelical stress, as assayed by their S1 nuclease hypersensitivity (Delaey *et al.*, 1987). The open conformation may allow easy access for factors involved in control of transcription or recombination processes. The simple DNA sequences, present in the rat CEA-like genes might, therefore, have been involved in the conservation of certain parts of these genes by gene conversion or formation of multiple CEA-like genes by unequal crossing over. The latter possible function is supported by the observed length heterogeneity of the simple sequences (Fig. 2) caused by imperfect alignment of the originally identical sequences and the presence of a 25-base pair direct repeat in the $(GGA)_n$ sequence in rnCGM3 (Fig. 2). Alternatively, slippage of DNA polymerase during replication of the simple repeated sequences could also account for this heterogeneity or formation of the direct repeat (Efstratiadis *et al.*, 1980).

Alignment of the deduced amino acid sequences of the five N-terminal domains and partial leader sequences of the rat and a comparison with the corresponding human CEA sequences, reveals strong conservation of certain amino acids. Among them are the first (serine) and the last (alanine) amino acids of the leader fragment, which are absolutely conserved in all rat (Fig. 4) and all human CEA-like antigen leader sequences analyzed so far (Thompson *et al.*, 1987; Beauchemin *et al.*, 1987; Watanabe and Chou, 1988a). These residues, therefore, could be an essential part of a putative signal peptidase recognition site (Perlman and Halvorson, 1983). In the N-terminal domain, two regions, one in the N-terminal and one in the C-terminal half, can be identified, where most of the conserved or conservatively exchanged amino acids are clustered. The latter region is flanked by two segments with low sequence similarity (Fig. 4). Recently, Williams (1987) suggested that the CEA repeat halves (Thompson and Zimmerman, 1988) reveal a close structural similarity to the constant domains of the immunoglobulins, whereas the N-terminal domain of CEA is more related to the variable domains. The latter suggestion is strengthened by the observation that 8 out of 13 amino acids, which are highly conserved in the variable domains of nearly all members of the immunoglobulin superfamily (Williams, 1987), are present in the N-terminal domain of all CEA-like antigens (Fig. 4). These critical amino acids probably play a key role in formation of the characteristic immunoglobulin fold, which in the case of the variable domain is composed of eight to nine β-strands (Williams, 1987). β-Strands very similar in number and length can also be predicted for the N-terminal domains of all CEA-like antigens of the rat, including the additional C′ strand, characteristic for the variable domain of the immunoglobulins (only shown for rnCGM1; Fig. 4). This C′ strand and two additional regions in the variable domain of immunoglobulin light chains form the antigen combining site or complementarity determining regions (Kabat *et al.*, 1987; Fig. 4). In the N-terminal domain of the CEA-like molecules, two hypervariable regions (HVR1 and HVR2) are found to completely overlap two of the three complementarity determining regions of immunoglobulin light chain variable domains (Fig. 4). In the immunoglobulins, the β-strands form two β-sheets, which are held together by a disulfide bond and a salt bridge between an arginine and aspartic acid. The latter two residues are also absolutely conserved in all members of the CEA gene family so far analyzed in rat. The ionic bond between arginine and aspartic acid is probably sufficient to stabilize the three-dimensional structure of the variable domain in immunoglobulins in the absence of the disulfide bridge (Williams, 1987). Taking these results together, the N-terminal domain of the CEA-like antigens is probably similarly folded to the variable domains of the immunoglobulins. This feature of the various members of the CEA family may indicate, in analogy to the immunoglobulin superfamily, a recognition or a receptor function for the CEA-like antigens, each conveying a separate function.

Acknowledgments—We wish to thank M. Weiss for her excellent technical assistance, G. Scherer for the gift of the rat genomic library, and G. Igloi for the synthesis of oligonucleotides. Assistance with the computer search by F. Jehle and help with oligonucleotide purification by K. Schwarz is gratefully acknowledged. The automated DNA synthesizer used for oligonucleotide synthesis was a gift from the "Fonds der Chemischen Industrie" to H. Kössel.

REFERENCES

Abeyounis, C. J., and Milgrom, F. (1976) *J. Immunol.* **116,** 30–34

Beauchemin, N., Benchimol, S., Cournoyer, D., Fuks, A., and Stanners, C. P. (1987) *Mol. Cell. Biol.* **7,** 3221–3230

Cohen, J. B., Effron, K., Rechavi, E., Ben-Neriah, Y., Zakut, R., and Givol, D. (1982) *Nucleic Acids Res.* **10,** 3353–3370

Delaey, B., Dirckx, L., Decourt, J. L., Claessens, F., Peeters, B., and Rombauts, W. (1987) *Nucleic Acids Res.* **15,** 1627–1641

Druckrey, H. (1971) *Arzneim.-Forsch.* **21,** 1274–1278

Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C., and Proudfoot, N. J. (1980) *Cell* **21,** 653–668

Feinberg, A. P., and Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13

Feng, D. F., Johnson, M. S., and Doolittle, R. F. (1985) *J. Mol. Evol.* **21,** 112–125

Frischauf, A. M., Lehrach, H., Poustka, A., and Murray, N. (1983) *J. Mol. Biol.* **170,** 827–842

Haagensen, D. E., Jr., Metzgar, R. S., Swenson, B., Dilley, W. G., Cox, C. E., Davis, S., Murdoch, J., Zamcheck, N., and Wells, S. A., Jr. (1982) *J. Natl. Cancer Inst.* **69,** 1073–1076

Hammarström, S., Engvall, E., Johansson, B. G., Svensson, S., Sundblad, G., and Goldstein, I. J. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72,** 1528–1532

Howell, J. H., Russo, A. J., and Goldrosen, M. H. (1979) *Cancer Res.* **39,** 612–618

Jantscheff, P., Indzhiia, V., and Micheel, B. (1986) *Arch. Geschwulstforsch.* **56,** 113–116

Johnson, F. E., LaRegina, M. C., Devine, J. E., Herbold, D. R., and Palmer, D. C. (1985) *Cancer Detect. Prev.* **8,** 471–476

Kabat, E. A., Wu, T. T., Reid-Miller, M., Perry. H. M., and Gottesman, K. S. (1987) *Sequences of Proteins of Immunological Interest*, United States Department of Health and Human Services, Public Health Service, National Institutes of Health, Bethesda, MD

Karn, J., Brenner, S., and Barnett, L. (1983) *Methods Enzymol.* **101,** 3–19

Kelly, A., and Trowsdale, J. (1985) *Nucleic Acids Res.* **13,** 1607–1621

Köhler, H., Rudofsky, S., and Kluskens, L. (1975) *J. Immunol.* **114,** 415–421

Maniatis, T., Fritsch, E. F., and Sambrook, J. (eds) (1982) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Martin, F., Knobel, S., Martin, M., and Bordes, M. (1975a) *Cancer Res.* **35,** 333–336

Martin, F., Martin, M. S., Bordes, M., and Knobel, S. (1975b) *Int. J. Cancer* **15,** 144–151

Mount, S. M. (1982) *Nucleic Acids Res.* **10,** 459–472

Müller, R. M., Taguchi, H., and Shibahara, S. (1987) *J. Biol. Chem.* **262,** 6795–6802

Neumaier, M., Zimmermann, W., Shively, L., Hinoda, Y., Riggs, A. D., and Shively, J. E. (1988) *J. Biol. Chem.* **263,** 3202–3207

Oikawa, S., Nakazato, H., and Kosaki, G. (1987a) *Biochem. Biophys. Res. Commun.* **142,** 511–518

Oikawa, S., Imajo, S., Noguchi, T., Kosaki, G., and Nakazato, H. (1987b) *Biochem. Biophys. Res. Commun.* **144,** 634–642

Oikawa, S., Kosaki, G., and Nakazato, H. (1987c) *Biochem. Biophys. Res. Commun.* **146,** 464–469

Paxton, R., Mooser, G., Pande, H., Lee, T. D., and Shively, J. E. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 920–924

Perlman, D., and Halvorson, H. O. (1983) *J. Mol. Biol.* **167,** 391–409

Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467

Shinomiya, T., Scherer, G., Schmid, W., Zentgraf, H., and Schütz, G. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 1346–1350

Shively, J. E., and Beatty, J. D. (1985) *CRC Crit. Rev. Oncol./Hematol.* **2,** 355–399

Slightom, J. L., Blechl, A. E., and Smithies, O. (1980) *Cell* **21,** 627–638

Stevens, R. H., England, C. W., Osborne, J. W., Cheng, H. F., and Richerson, H. B. (1975) *J. Natl. Cancer. Inst.* **51,** 1011–1013

Stevens, R. H., Englund, C. W., Osborne, J. W., Cheng, H. F., and Hoffman, K. L. (1976) *Cancer Res.* **36,** 3260–3264

Tawaragi, Y., Oikawa, S., Matsuoka, Y., Kosaki, G., and Nakazato, H. (1988) *Biochem. Biophys. Res. Commun.* **150,** 89–96

Terry, W. D., Henkart, P. A., Coligan, J. E., and Todd, C. W. (1974) *Transplant. Rev.* **20,** 100–129

Thompson, J. A., and Zimmermann, W. (1988) *Tumor Biol.* **9,** 63–83

Thompson, J. A., Pande, H., Paxton, R. J., Shively, L., Padma, A., Simmer, R. L., Todd, C. T., Riggs, A. D., and Shively, J. E. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 2965–2969

van Hove, L., Delacourt, M., Park, B., Sobis, H., and Vandeputte, M. (1978) *Int. J. Cancer* **21,** 731–740

van den Heuvel, R., Hendriks, G., Quaks, W., and Bloemendal, H. (1985) *J. Mol. Biol.* **185,** 273–284

Wahren, B., Gadler, F., Gahrten, G., Hammarström, S., Hareland, Y., Hyden, N., Ljungdahl, E., Mahlen, A., Ruden, U., and Wiklund, M. (1983) *Ann. N. Y. Acad. Sci.* **417,** 344–358

Watanabe, S., and Chou, J. Y. (1988a) *J. Biol. Chem.* **263,** 2049–2054

Watanabe, S., and Chou, J. Y. (1988b) *Biochem. Biophys. Res. Commun.* **152,** 762–768

Williams, A. F. (1987) *Immunol. Today* **8,** 298–303

Zimmermann, W., Ortlieb, B., Friedrich, R., and von Kleist, S. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 2960–2964

Zimmermann, W., Weber, B., Ortlieb, B., Rudert, F., Schempp, W., Fiebig, H.-H., Shively, J. E., von Kleist, S., and Thompson, J. A. (1988) *Cancer Res.* **48,** 2550–2554