

Genomic organization, splice variants and expression of *CGM1*, a CD66-related member of the carcinoembryonic antigen gene family

Gerhard NAGEL¹, Fritz GRUNERT¹, Taco W. KUIJPERS², Suzanne M. WATT³, John THOMPSON¹ and Wolfgang ZIMMERMANN¹

¹ Institute of Immunobiology, University of Freiburg, Germany

² Central Laboratory of the Netherlands Red Cross Blood Transfusion Service, Amsterdam, The Netherlands

³ Medical Oncology Unit, Imperial Cancer Research Fund, St. Bartholomew's Hospital, London, England

(Received December 7, 1992/February 2, 1993) – EJB 92 1747/1

The tumor marker carcinoembryonic antigen (CEA) belongs to a family of proteins which are composed of one immunoglobulin variable domain and a varying number of immunoglobulin constant-like domains. Most of the membrane-bound members, which are anchored either by a glycosylphosphatidylinositol moiety or a transmembrane domain, have been shown to convey cell adhesion *in vitro*. Here we describe two splice variants of *CGM1*, a transmembrane member of the CEA family without immunoglobulin constant-like domains. *CGM1a* and *CGM1c* contain cytoplasmic domains of 71 and 31 amino acids, respectively. The cytoplasmic region of *CGM1a* is encoded by four exons (Cyt1–Cyt4). Differential splicing of the Cyt1 exon (53 bp) leads to the formation of *CGM1c*. The presence or absence of potential protein kinase phosphorylation sites in the cytoplasmic domains and a sequence consensus motif involved in signal transduction in multichain immune recognition receptors indicates that this splice event is of functional importance. *CGM1a* mRNA, the predominant *CGM1* transcript, was found in the granulocytic lineage, but not in monocytes, lymphocytes nor in a number of tumors derived from all three germ layers. Weak staining using monoclonal antibodies Tu2 and 73 in fluorescence-activated cell scan analyses indicate low concentrations of *CGM1* protein on the surface of granulocytes. The *CGM1* protein is also recognized by CD66 antibodies. Therefore, the granulocyte-specific CD66 epitope is present on at least four CEA family members: *CGM1*, CEA, NCA-50/90 and NCA-160.

The carcinoembryonic antigen (CEA) gene family encodes the widely used tumor marker CEA (Shively and Beatty, 1985), as well as a number of closely related membrane-bound glycoproteins, some of which convey homophilic [CEA, nonspecific crossreacting antigen (NCA), biliary glycoprotein (BGP)] and/or heterophilic cell adhesion (CEA, NCA, BGP, *CGM6*) *in vitro* (Benchimol et al., 1989; Oikawa et al., 1989, 1991; Rojas et al., 1990; Zhou et al., 1990). As members of the immunoglobulin superfamily (Paxton et al., 1987), they consist of one N-terminal, immunoglobulin variable-like (IgV), and a varying number (0, 2, 3, 6) of immunoglobulin constant-like (IgC) domains (reviewed by Thompson et al., 1991). They are anchored to the cell membrane either via a glycosylphosphatidylinositol moiety (CEA,

NCA, *CGM6*) or have a hydrophobic transmembrane region (BGP). For BGP as well as for its rat (ecto ATPase) and mouse homologues (mm*CGM1/2*), isoforms have been described which differ in the length of their cytoplasmic domains (Barnett et al., 1989; Lin and Guidotti, 1989; McCuaig et al., 1992; Culic et al., 1992). The intracellular domains of human and rat BGP isoforms are phosphorylated (Afar et al., 1992; Culic et al., 1992). Furthermore, the intracellular domain of the rat homologue has been shown to be a substrate of the insulin receptor tyrosine kinase (Margolis et al., 1990). Although the functional significance of this observation is unclear, it suggests that the transmembrane members of the CEA family might be involved in signal transduction.

CEA-related antigens reveal high sequence similarities as deduced from cDNA and genomic cloning data, but they are differentially expressed (reviewed in Thompson et al., 1991). For example, CEA is found in different adenocarcinomas and some normal epithelial tissues. A coexpression of NCA-50/90 is found in the same adenocarcinomas, but this antigen is also found in granulocytes, which never express CEA. NCA-95, on the other hand, seems to be restricted to cells of the myeloid lineage and has not been found in other tissues tested thus far (Berling et al., 1990). Recently, it was demonstrated that the antibodies assigned to the granulocyte-specific clusters of differentiation (CD), CD66 and CD67 (Knapp et al., 1989), are directed against epitopes expressed by members of the CEA family. CD66 antibodies were re-

Correspondence to W. Zimmermann, Institut für Immunbiologie, Universität Freiburg, Stefan-Meier-Strasse 8, W-7800 Freiburg, Germany

Fax: +49 761 203 4016.

Abbreviations. CEA, carcinoembryonic antigen; NCA, nonspecific crossreacting antigen; BGP, biliary glycoprotein; *CGM*, CEA gene family member; PCR, polymerase chain reaction; CD, cluster of differentiation; IgC, immunoglobulin constant, IgV, immunoglobulin variable; mAb, monoclonal antibody; CHO, Chinese hamster ovary; CML, chronic myeloid leukemia; FACScan, fluorescence-activated cell scan; DOTMA, *N*-[1-(2,3-dioleloxy)-propyl]-*N,N,N*-trimethylammonium chloride.

Note. The novel nucleotide sequence data published here have been deposited with the GenBank (Los Alamos) sequence data bank and are available under the accession numbers L00692 and L00693.

ported to react with the granulocytic glycoproteins NCA-50/90 and NCA-160, whereas CD67 antibodies are reactive with NCA-95 (Watt et al. 1991; Skubitz et al., 1992; van der Schoot et al., unpublished results). Several other CEA-related antigens have been described in leukocytes, many of which have not yet been assigned to specific genes (Kuroki et al., 1990a,b, 1991, 1992).

Recently, a new CEA-related protein with transmembrane anchorage (CGM1a) was identified by cDNA cloning (Kuroki et al., 1991). The mature protein is predicted to consist of only one IgV-like domain, a transmembrane region and a cytoplasmic tail. Although of similar length, only the first third of the CGM1a cytoplasmic domain shows similarity to the intracellular tail of BGP_a (old name: TM-1 CEA) at the amino acid sequence level. It is not known whether other CGM1 isoforms with different cytoplasmic domains like BGP_a and BGP_c (TM-3 CEA) exist (Barnett et al., 1989). Furthermore, little is known about the expression pattern of CGM1 and its immunological relationship with other members of the CEA family.

As a basis for functional analyses of the intracytoplasmic domain of CEA-related proteins, we have looked for CGM1 splice variants. Here, we describe two CGM1 mRNAs created by complex differential splicing which encode isoforms with different cytoplasmic domains. Using the polymerase chain reaction (PCR) technique and monoclonal antibodies (mAbs), their expression patterns have been investigated in different leukocyte subpopulations and in tumors.

MATERIALS AND METHODS

Monoclonal antibodies

The mouse mAb N1 was raised against purified NCA-50. Its specificity has been described before (Schwarz et al., 1989; Berling et al., 1990). mAbs CLBgran 10, YTH71.3.2 (CD66) and B13.9 (CD67) were characterized at the last Workshop on Leukocyte Differentiation Antibodies (Knapp et al., 1989; reagents M38, M128, M37 respectively). mAb Tu2 was purchased from Boehringer Mannheim. mAb 73 was a generous gift from F. Buchegger (Lausanne).

Isolation and characterization of cDNA clones

The anchored polymerase chain reaction (PCR) was used to isolate full-length cDNA clones. Total RNA (1 µg) from leukocytes of a chronic myeloid leukemia patient was reverse-transcribed with avian myeloblastosis virus reverse transcriptase (Promega) using an oligo(dT)-containing oligonucleotide [5'-GGAATTCGTCGACAAGCTT(T)₁₂G/C/A-3'] as a primer. This primer can hybridize at the ends of all polyadenylated mRNAs due to the redundancy at its 3' end (Leibrock et al., 1989) and it contains *EcoRI*, *SalI* and *HindIII* recognition sites for subcloning. CEA-related cDNAs were amplified by 30 cycles (denaturation: 1 min, 93°C; annealing: 30 s, 50°C; extension: 3 min, 72°C) in a Thermocycler 60 (Bio-Med) after addition of a CGM1-specific oligonucleotide deduced from the 5'-untranslated region of the known CGM1 gene sequence (Thompson et al., 1989) and *Taq* DNA polymerase (Promega) according to Sambrook et al. (1989). Selection of the gene-specific primer CGM1-5' (5'-GGAATTCCTGGAGCCCAGGCTCTTT-3') was aided by the program 'Primer' (Lucas et al., 1991). It differs in at least three positions compared to the known sequences of the other members of the CEA gene family. To the 5' end of

this primer, nucleotides were added to generate an *EcoRI* restriction site. The longest PCR product with a length of about 1100 bp was purified by agarose gel electrophoresis and subcloned into Bluescript (Stratagene). Sequencing was performed on single- or double-stranded templates with Sequenase (United States Biochemicals) or T7 DNA polymerase (Pharmacia) using internal or universal oligonucleotide primers. Protein kinase phosphorylation sites were identified using the sequence analyses package PC/Gene (Intelli-Genetics Inc.).

Determination of exon structure

The borders of the exons coding for the cytoplasmic domain were determined by sequencing. For this purpose a 2.4-kb genomic fragment was amplified by PCR from cosmid DNA of clone 8486, which contains the CGM1 gene (Tynan et al., 1992), using oligonucleotides from the transmembrane domain (MW1, position 568–584) and the 3'-untranslated region (MW2, position 973–989), subcloned into Bluescript (pBSCGM1-Cyt) and sequenced with oligonucleotides derived from the cDNA sequence as primers. The distances between the exons were inferred from the size of the amplified genomic DNA fragments. Supercoiled cosmid 8486 and plasmid pBSCGM1-Cyt DNAs were used as template with 5'-oligonucleotides TEX11 (A domain, position 629–645), CGM1-Cyt1 (cytoplasmic domain, position 597–613), CGM1-Cyt3 (cytoplasmic domain, position 712–728) and 3'-oligonucleotides CGM1-TM (transmembrane region, position 520–536), CGM1-Cyt1c (cytoplasmic domain, position 617–633), CGM1-Cyt2 (cytoplasmic domain, position 668–680) in the following combinations: TEX11/CGM1-TM, TEX11/CGM1-Cyt1c, MW1/CGM1-Cyt1c, MW1/CGM1-Cyt2, MW1/MW2, CGM1-Cyt1/CGM1-Cyt2, CGM1-Cyt1/MW2, CGM1-Cyt3/MW2. The position numbers refer to the CGM1b (TEX11; Kuroki et al., 1991) and the CGM1a sequence (see below), respectively. After subcloning, some of these PCR products were also used for the determination of exon/intron borders by sequencing. Since the distance between exons N and A could not be determined by PCR, the presence of two *KpnI* restriction endonuclease sites in the TM exon and in the 3'-flanking intron, respectively, was exploited to map the position of this exon relative to the flanking *EcoRI* sites. For this purpose, the 14.4-kb *EcoRI* DNA fragment from cosmid 17104, which hybridizes with a CEA exon A/B probe (Tynan et al., 1992), the corresponding fragment from cosmid 8486, which is truncated at its 3'-end, as well as pBSCGM1-Cyt were digested with *KpnI*. The positions of the *KpnI* sites were inferred from the size of the resulting fragments.

RNA analyses

Total RNA was isolated from fresh cells or tissues pulverized in liquid nitrogen by the guanidine thiocyanate method (Fiddes and Goodman, 1979) or by the acid phenol method of Chomczynski and Sacchi (1987). Reverse transcription and PCR for the detection of β -actin mRNA were performed as described by Lucas et al. (1991). For the reverse transcription reaction, 50 pmol 3'-oligonucleotide (5'-AGCCATGCCAATCTCATCTTGT-3') derived from the human β -actin gene was used to prime first-strand DNA synthesis from 0.5 µg total RNA. After denaturation at 93°C, the corresponding 5'-oligonucleotide (5'-ACGGCTGCTTCAGCTCCTC-3') was added for PCR. Reverse transcription

for the detection of CGM1 mRNA was carried out using a first-strand cDNA synthesis kit from Pharmacia. CGM1-specific oligonucleotides for PCR were chosen from the 5'-untranslated (CGM1-5') and 3'-untranslated region (MW2, see above). PCR conditions were as described above, except annealing was for 15 s at 64°C and 54°C for the detection of β -actin mRNA and CGM1 mRNAs, respectively. Before analysis by gel electrophoresis, samples were treated with RNase A (20 μ g/ml) at 37°C for 15 min.

Cells and tissues

HeLa cells were cultured in RPMI 1640, Chinese hamster ovary (CHO) cells in minimum essential medium α , both supplemented with 10% fetal calf serum, penicillin and streptomycin. Stable transfectants were cultured in the presence of 0.5 mg/ml G418 (Gibco-BRL). Human primary tumors, except gynecological tumors, were propagated in nude mice and were frozen and stored in liquid nitrogen immediately after resection. Leukocytes from chronic myeloid leukemia (CML) patients were isolated from peripheral blood by centrifugation through a Percoll (Pharmacia) step gradient according to the manufacturer's protocol and used directly for RNA isolation. In some cases Ficoll Paque (Pharmacia) instead of Percoll was used for the isolation of peripheral blood leukocytes. Granulocytes were purified as described before (Roos and de Boer, 1986). The purity of the granulocytes was greater than 98%, with >95% of the cells being neutrophils. Lymphocytes and monocytes were purified from the mononuclear cell fraction by countercurrent elutriation (Roos and de Boer, 1986). Cell purity was greater than 90%.

Stable transfectants

CGM1a cDNA and CEA cDNA (Pèlegri et al., 1992) were excised from Bluescript (BS KS M13+) by double digestion with *Eco*RI and *Hind*III, which cut in the flanking polylinker region. The inserts were cloned into the expression vector pBEHpUC9 (Artelt et al., 1988) downstream of the SV40 early promoter. The construction of the NCA-50/90 expression vector has been described before (Kolbinger et al., 1989). Transfection of HeLa cells was performed with minor modifications as described by Felgner et al. (1987). 10^6 HeLa cells on a 10-cm plastic Petri dish were transfected using *N*-[1-(2, 3-dioleoyloxy)-propyl]-*N,N,N*-trimethylammonium chloride (DOTMA; Boehringer Mannheim) as transfection reagent according to the manufacturer's protocol. 1 μ g pSV2neo plasmid DNA (Southern and Berg, 1982), which confers resistance toward the neomycin derivative G418, and 20 μ g of the CGM1a expression construct were incubated together with 30 μ l DOTMA in 500 μ l phosphate-buffered saline (NaCl/P_i; 137 mM NaCl, 2.7 mM KCl, 8.1 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.4) for 10 min at room temperature. HeLa cells were incubated with the transfection mixture in 8 ml RPMI, 10% fetal calf serum for 20 h at 37°C, 10% CO₂. Control cultures were transfected with 1 μ g pSV2neo DNA alone. Selection for cells expressing the bacterial *neo* gene was carried out in the presence of 0.5–1.0 mg/ml G418. After about 4 weeks, colonies were isolated using cloning cylinders and analyzed for surface expression of CGM1 by flow cytometry. The production of CGM6- and BGP_a-expressing transfectants has been described previously (Berling et al., 1990; Kuroki et al., 1991). BGP_a and CGM7 transfectants were a kind gift from M. Kuroki and Y. Matsumoto, Fukuoka University.

Flow cytometry analysis

Flow cytometry (Leiserson, 1985) of cloned G418-resistant cells was performed with a fluorescence activated cell scan (FACScan) flow cytometer (Becton Dickinson) using the LysisII software for data processing. Cells were removed from plastic dishes with NaCl/P_i containing 0.5 mM EDTA. All subsequent steps were performed at 4°C with NaCl/P_i containing 1.1 mM CaCl₂ and 0.4 mM MgSO₄ and 3% fetal calf serum (NaCl/P_i/fetal calf serum) 0.5×10^6 cells in 35 μ l NaCl/P_i/fetal calf serum were incubated with an excess of mouse or rat monoclonal antibodies (0.25–1.25 μ g) for 20 min. A monoclonal antibody of IgG1 isotype (control IgG1) directed against a non-biological hapten (Immunotech) served as a negative control. After washing twice with 200 μ l ice-cold NaCl/P_i/fetal calf serum, 25 μ l biotinylated anti-mouse polyvalent Ig (Sigma) or anti-rat IgG (Amersham International; both diluted 1:50 with NaCl/P_i/fetal calf serum) were added to the cells for another 20 min followed by a wash. Then the cells were washed as above and incubated with 25 μ l phycoerythrin-linked streptavidin (1:80, Southern Biotechnology Associates) for 20 min, washed and resuspended in 300 μ l NaCl/P_i/fetal calf serum for FACScan analysis. Alternatively, anti-mouse IgG directly coupled to phycoerythrin (1:50, Southern Biotechnology Associates) was used as second step antibody.

RESULTS

Cloning of CGM1 cDNAs

Several CEA-related antigens not yet assigned to their corresponding genes have been found in granulocytes (Kuroki et al., 1990a). As granulocytic precursor cells are known to contain larger amounts of CEA-related mRNAs in comparison to mature granulocytes (Berling et al., 1990), RNA from leukocytes of a CML patient, which are enriched for myeloid precursors, was used as template for generation of full-length CGM1 cDNAs by PCR. Two cDNA species were isolated, sequenced and shown by comparison with the published genomic sequence (Thompson et al., 1989) to code for CGM1. The sequence of the two cDNAs is shown in Fig. 1A together with the deduced domain structure (Fig. 1B). The longer cDNA (CGM1a) consists of a 54-bp 5'-noncoding region, a 756-bp open reading frame, and 317-bp 3'-untranslated region. It codes for a putative preprotein of 252 amino acids with an *M_r* of 27077. Comparison of the deduced amino acid sequence with other members of the CEA family (Thompson et al., 1991) revealed that the encoded protein consists of a 34-amino-acid leader peptide (L), a 108-amino-acid IgV-like domain (N), a putative transmembrane domain of 39 amino acids and a 71-amino-acid cytoplasmic domain, but lacks immunoglobulin constant like (IgC) domains found in most other members of the CEA family. Two potential sites for *N*-glycosylation are present in the N domain at conserved positions in comparison with NCA-50/90 (Neumaier et al., 1988). The sequence of the shorter cDNA (CGM1c; 1022 bp) shows two deletions of 53 and 52 bp, respectively. The encoded protein (*M_r* 22780) is nearly identical to that of CGM1a with respect to the IgV-like (N) and transmembrane (TM) domain, but differs in the cytoplasmic domain. The cytoplasmic tail of CGM1c consists of only 31 amino acids and has a different amino acid sequence. This sequence deviation is caused by the 53-bp deletion at the border between the putative transmembrane and cytoplasmic domains, which

Table 1. Nucleotide sequence of exon/intron borders of the CGM1 gene. Capital letters represent coding regions, intron sequences are shown in lower-case letters. Splice donor and acceptor motifs are underlined. The borders of the leader (L) and the N-terminal domain (N) exons have been published elsewhere (Thompson et al., 1989). Abbreviations for the exons are as in legend to Fig. 1; nd, not determined; 3'UT, 3'-untranslated region.

Exon	Intron		splice acceptor	Exon
	splice donor			
L	...CTGCTCACAG	<u>gt</u> gagtgaggattcctgggagtgg...	...aatattgactgatgcctttctccctcctag	CCTCACTTCT...
M	...CATGTATACC	<u>gt</u> gagatttccacatgacctctgg...	nd	AAGAAAATGC...
TM	...AACTGGAAG	<u>gt</u> accacagctttttcccatccttc...	...acctcccaaatgacctgaccttttctag	AACCAGCATC...
Cyt1	...CTTGCCTCG	<u>gt</u> gagtgctcctctcagccagggtg...	...cctgtgctcactcctctctctgtttccag	GCCGTGGTCC...
Cyt2	...TGCCCTCTCG	<u>gt</u> aagcctgtccccctccagccctc...	...cctccacagaccctcctcctgtcccaag	ATGTCCCTC...
Cyt3	...CATCTATGAG	<u>gt</u> gagtggtggccacggatgtctg...	...agggtccagccctctctctgttttaacag	GAATTGCTAA...

ons of the CGM1 gene were determined by PCR and by restriction endonuclease mapping. The data imply that the CGM1 gene spans a region of about 15 kb and is contained within three *EcoRI* DNA fragments of 2.7, 4.9 and 14.4 kb which have been identified and mapped recently (Thompson et al., 1989; Tynan et al., 1992). The size of the larger *EcoRI* DNA fragment was determined by hybridization of *CGM1*-containing cosmid DNAs with an oligonucleotide from the 3'-untranslated region (data not shown).

CD66 but not CD67 antibodies identify epitopes expressed by CGM1

The deduced amino acid sequence of the N-domain of CGM1 is very similar to that of NCA-50/90 (90% similarity; Neumaier et al., 1988). Since it has recently been shown that NCA-50/90 on granulocytes is recognized by CD66 antibodies (see introduction) and NCA-95, another NCA species reacts with CD67 antibodies, we decided to test whether CD66 and CD67 epitopes are also present on the CGM1 protein. To achieve this, we transfected HeLa cells with a CGM1a cDNA-containing expression vector together with a G418 resistance-conferring plasmid. G418-resistant cell clones were analyzed for CGM1 expression by flow cytometry using polyclonal anti-CEA and anti-NCA antisera. One HeLa cell clone was found by Western blot analyses to express a CGM1 protein of 30–35 kDa (data not shown). This CGM1-expressing HeLa cell clone, together with control HeLa transfectants which express CEA (HeLa-CEA), NCA-50/90 (HeLa-NCA), CGM6 (HeLa-CGM6) and the *neo* gene alone (HeLa-Neo) as well as a BGP α - and a CGM7-producing CHO cell line, were used for flow cytometric analysis with antibodies assigned to CD66 (CLBgran10) and CD67 (B13.9). The BGP α transfectant cell line was included since the BGP gene in granulocytes probably codes for the CD66 antigen NCA-160 (see Discussion). The CD66, but not the CD67, antibody reacted with the CGM1-expressing HeLa cell clone (Fig. 3). The former crossreacted with CEA, NCA and BGP α transfectants to a varying extent, whereas the CD67 antibody reacted only with the CGM6-expressing cell line. A second CD66 antibody (YTH71.3.2) exhibited the same antigen recognition pattern as CLBgran10 (data not shown).

Analysis of the expression pattern of CGM1

Since the two CGM1 cDNA clones were isolated from a mixture of white blood cells from a CML patient, we analysed various types of leukocytes for the presence of CGM1 mRNA by the PCR technique. Oligonucleotides derived from the 5'- and 3'-untranslated region of CGM1 mRNA were used as primers. These oligonucleotides allow the specific detection of both splice variants CGM1a and CGM1c as 994-bp and 889-bp DNA fragments, respectively. The intactness of the RNA samples was confirmed by analyzing for the presence of β -actin mRNA. This reaction was carried out both with and without reverse transcription prior to PCR to detect any genomic DNA contamination with processed β -actin pseudogenes (Ng et al., 1985). The β -actin cDNA fragments, which were obtained after reverse transcription, are shown (Fig. 4B). No or only weak signals were found when reverse transcriptase was omitted (data not shown). We found CGM1a mRNA to be present in normal granulocytes and leukocytes of CML patients (Fig. 4A, lanes 6–10). CGM1c mRNA could be detected in total leukocyte RNA from one

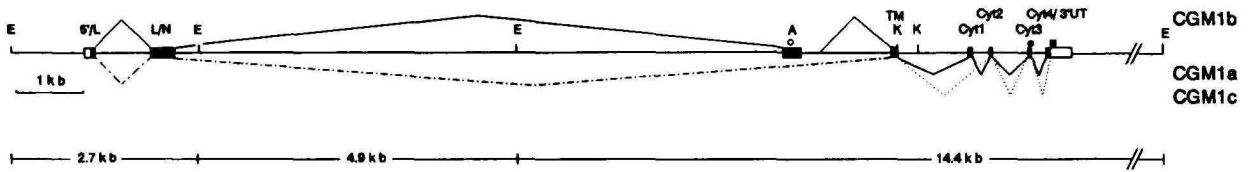


Fig. 2. Exon structure of the CGM1 gene. Exons corresponding to the 5'- and 3'-untranslated regions (UT) and to coding regions are symbolized by open and filled-in boxes, respectively (for abbreviations see legend to Fig. 1A). The transcriptional start site has not been determined experimentally. Therefore, the size of the first exon is only approximate. E, K, *EcoRI* and *KpnI* restriction endonuclease sites. Only the *KpnI* sites within the 14.4-kb *EcoRI* DNA fragment have been determined. Splice events leading to CGM1a, CGM1b (Kuroki et al., 1991) and CGM1c are indicated by lines (continuous lines, CGM1b- and CGM1a-specific splice events; dotted lines, splice events observed for CGM1c; dashed lines, splice events common to CGM1a and CGM1c). Stop codons are marked by an open (CGM1b) or closed circle (CGM1c) or a filled-in square (CGM1a).

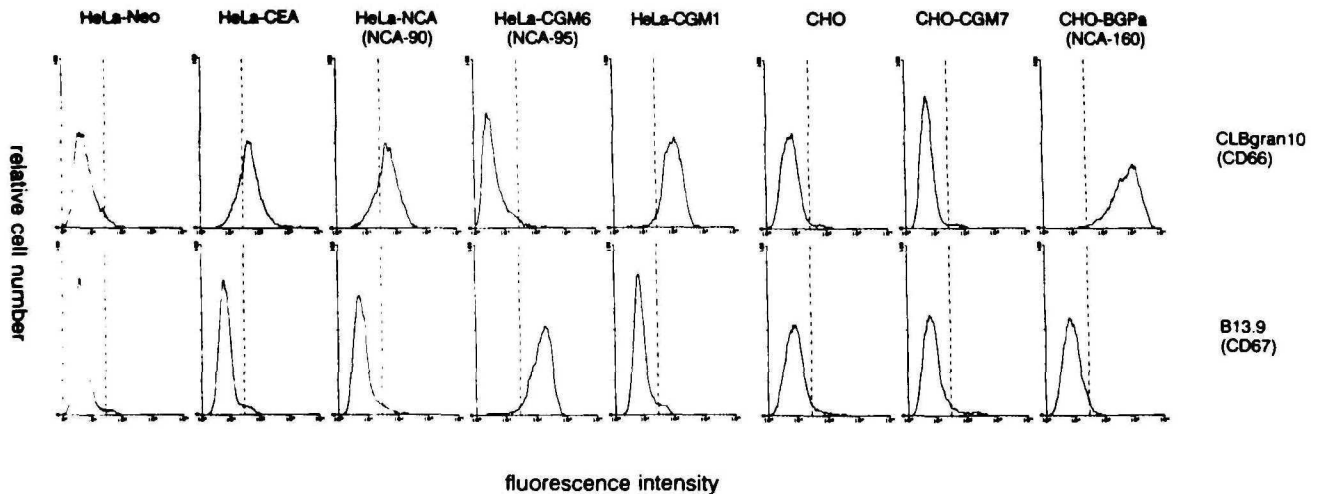


Fig. 3. Determination of the reactivity of CD66 and CD67 mAbs with transfectants expressing various members of the CEA family by flow cytometry. Stable HeLa and CHO cell transfectants expressing members of the CEA gene family (deviating antigen names are indicated in brackets) were analysed with mAb CLBgran10 (CD66) or with B13.9 (CD67). HeLa cells transfected with pSV2neo (HeLa-Neo) alone and the parental CHO cell line served as negative controls. More than 95% of the negative control cells show a relative fluorescence of less than 30 which is indicated by the broken vertical line. The CGM7 transfectant was proved to express the CGM7 antigen (P. Jantschkeff, unpublished results; data not shown).

CML patient only (Fig. 4A, lane 7). The identity of the amplification products obtained was confirmed by stringent hybridisation with an oligonucleotide derived from the transmembrane domain of CGM1 (data not shown). Using total RNA from monocytes and lymphocytes as templates, no signals were obtained (Fig. 4A, lanes 4, 5).

In order to analyze the expression pattern of CGM1 further, we tested RNAs isolated from a number of tumors and metastases derived from all three germ layers as well as a few normal tissues. In contrast to the results obtained with RNA from CML leukocytes or normal granulocytes, no CGM1 mRNAs could be detected in the following tissues: adenocarcinomas (one each) of the pancreas, lung, kidney, stomach and colon, three mucinous carcinomas of the ovary, one small-cell and one large-cell carcinoma of the lung, one squamous carcinoma of the gall bladder and one of the lung, one hepatocellular carcinoma, one osteosarcoma, one sarcoma of the ovary, one teratoma, one melanoma and one testis carcinoma as well as normal spleen, placenta and colonic mucosa.

These data indicate that the CGM1 gene is specifically expressed in the granulocytic lineage. Therefore, we tested whether the CGM1 protein can be detected on the surface of mature granulocytes. Since several NCA-related antigens are expressed on granulocytes, CGM1-specific antibodies are

needed. For this purpose we analyzed a large panel of mAbs (87) prepared against CEA, CEA-crossreacting antigens or various CEA/NCA-expressing cell lines for the presence of CGM1-specific antibodies. The panel of antibodies was available from the groups who participated in a workshop for the characterization of antibodies directed against CEA-related antigens (Nap et al., 1992). Two of the antibodies, which recognized CGM1 on transfectants (mAb Tu2 and mAb 73) were only crossreactive with CEA, but not with NCA, CGM6, CGM7 nor with BGPa transfectants. However, crossreaction with the product of *CGM2*, the only other expressible gene of the CEA subgroup, cannot be ruled out at present (Thompson et al., 1991). Since CEA is not present on granulocytes, these mAbs were used to analyze surface expression of the *CGM1* product by flow cytometry. The results shown in Fig. 5 suggest that the CGM1 product is present on the surface of mature granulocytes although in low amounts. This is inferred from the observation that granulocytes show only a low fluorescence signal when analyzed with Tu2 and mAb 73 compared to the signal obtained with mAb N1 that reacts with both CGM1 and NCA-90, but not with CEA, BGPa and the products of *CGM6* and *CGM7* (Fig. 5; Berling et al., 1990, and unpublished results). This is not due to a higher affinity of N1, since mAbs N1, Tu2 and 73 recognize the CGM1 transfectant equally well (Fig. 5).

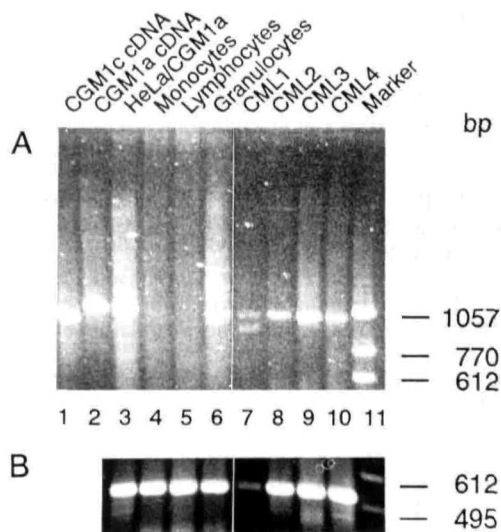


Fig. 4. Analysis of CGM1 expression in hematopoietic cells by PCR. (A) For the specific amplification, oligonucleotides from the 5'- and 3'-untranslated region of CGM1 mRNA were used. As positive controls 6 ng supercoiled pCGM1a (lane 2) and pCGM1c cDNA (lane 1) plasmids and total RNA from the HeLa-CGM1a transfectant (lane 3) were tested. Total RNA from monocytes, lymphocytes and granulocytes (lanes 4–6) of healthy donors and from leukocytes of four different CML patients (lanes 7–10) was analyzed. 20% of the amplification products (except for CGM1 cDNAs and the CML1 sample where 2% and 70%, respectively, were used) were separated by electrophoresis on a 1.8% agarose gel, stained with ethidium bromide and visualized under ultraviolet light. The integrity of the RNA samples was examined by amplification of a β -actin mRNA fragment after reverse transcription (B). *HincII*-digested bacteriophage FX174 DNA was used as size marker (lane 11).

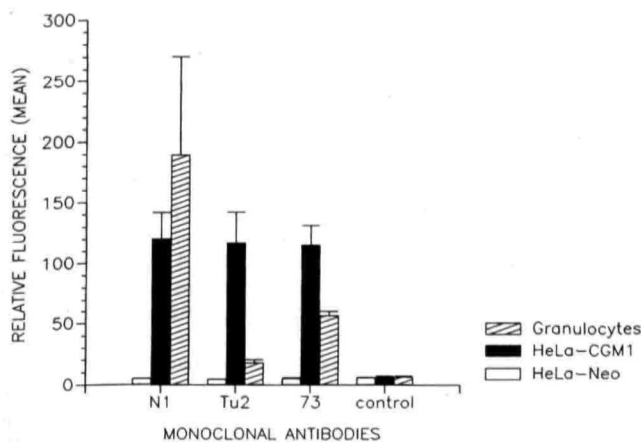


Fig. 5. Expression of CGM1 on the surface of granulocytes. The expression of CGM1 on granulocytes was determined by FACScan analyses (hatched columns). HeLa transfectants expressing CGM1 (HeLa-CGM1; filled-in columns) or transfected with pSV2neo alone (HeLa-Neo; open columns) served as positive and negative controls, respectively. As a further negative control, cells were incubated with an irrelevant isotype matched mAb (control). The mean fluorescence of duplicates from two independent experiments are shown. Bars indicate the deviations from the mean values.

DISCUSSION

In this paper we describe the isolation of two CGM1 cDNA clones, which encode CGM1 isoforms (CGM1a and CGM1c) with different cytoplasmic domains. CGM1a is es-

entially identical to cDNA clone W264, which was recently isolated from a normal white blood cell cDNA library although seven nucleotide differences were found (Kuroki et al., 1991). Two out of four differences in the coding region result in amino acid changes. They could represent allelic variations or errors introduced during PCR amplification. A second CGM1 splice variant (CGM1b; W282) described by Kuroki et al. (1991), which in contrast to CGM1a contains part of an IgC-like domain (type A) and sequences of the downstream intron, is predicted to encode a secreted protein since the A domain exon contains a nonsense mutation (see Fig. 2). Analysis of the exon organization of the CGM1 gene reveals that the cytoplasmic region of CGM1a is encoded by four small exons. In CGM1c, the first cytoplasmic domain exon (Cyt1) is spliced out, thus creating a shift in the reading frame and leading to a different amino acid sequence. Sequences similar to the Cyt1 exon of *CGM1* are also found in human [BGP_a or TM-1 CEA and BGP_b or TM-2 CEA (Thompson et al., 1991; Barnett et al., 1989)] and rodent BGP splice variants [ectoATPase (Lin and Guidotti, 1989)]. Deletion of this exon in BGP splice variants (BGP_c, BGP_d) also results in shorter cytoplasmic domains, although with a different sequence compared to CGM1c, since the sequence following the Cyt1 region in *BGP* is not homologous to Cyt2, Cyt3 or Cyt4. Although identical with CGM1a in its overall domain organization, the *CGM7* (old name: W236) product contains an eight-amino-acid deletion in the cytoplasmic domain (Kuroki et al., 1991). From the presence of a perfect splice acceptor signal within the Cyt3 exon of CGM1 (Fig. 1A), it can be inferred that alternative splicing is probably the cause for the in-frame deletion in the *CGM7* product.

It remains to be resolved whether the different cytoplasmic domains of *CGM1*-, *CGM7*- and *BGP*-encoded proteins are of any functional importance. In this regard, however, it may be interesting that analysis of the primary structures has revealed the presence of a conserved potential protein kinase C phosphorylation site (Woodget et al., 1986) at the N-terminus of the cytoplasmic domains of all transmembrane isoforms of the CEA family so far described (Barnett et al., 1989; Lin and Guidotti, 1989; Turbide et al., 1991; Kuroki et al., 1991; McCuaig et al., 1992). Multiple protein kinase C and casein kinase II phosphorylation sites (Pinna, 1990) are also conserved at variable positions in the cytoplasmic tails. Since phosphorylation seems to influence the activity of adhesion molecules, as found for the β -subunit of CD11/CD18 integrins (Hibbs et al., 1991), the adhesive properties of the CEA subgroup members could be modulated by differential splicing within their cytoplasmic domains. Differential splicing could also affect the potential signal transduction property of CGM1, since the 'YLYL' motif, which has been shown to be associated with signal transduction (Keegan and Paul, 1992), is only present in CGM1a. Interestingly, the BGP isoforms BGP_a and BGP_b and the corresponding rat homologues (Barnett et al., 1989; Culic et al., 1992) also contain a similar motif in their cytoplasmic domains in a region without any obvious similarity to CGM1a. However, a common ancestry of the cytoplasmic domains with the 'YLYL' motif is suggested by the observation that a type-0 intron separates the codons for the first Tyr and Leu within the motif in the genes for CGM1, mouse CD3 ξ and CD3 δ (Fig. 1A; Baniyash et al., 1989; van den Elsen et al., 1986).

The investigation of the immunological relationship of the CGM1 product with other members of the CEA family using stable HeLa and CHO transfectants expressing individ-

ual members revealed that the CGM1 antigen carries the CD66 epitope. Since the extracellular part of CGM1 consists only of an IgV-like N-domain, the CD66 epitope must be located in the N domain of NCA-50/90 and NCA-160. The binding of CD66 antibodies to the BGP transfectant and to a 160-kDa protein of granulocytes strongly supports the conclusion of Drzeniek et al. (1991) that NCA-160 on granulocytes is a product of the BGP gene, although a different molecular mass has been reported for the deglycosylated forms of NCA-160 and BGP_a (Audette et al., 1987; Kuroki et al., 1990a, 1991). Thus to date, the products of four genes (*CEA*, *NCA*, *BGP* and *CGM1*) bear the epitope(s) recognized by the CD66 antibodies. Therefore, the presence of CGM1 in low but significant amounts on the surface of granulocytes has to be taken into account when the function of CEA-related antigens is studied with CD66 antibodies (Kuijpers et al., 1992).

Analysis of the expression pattern of *CGM1* revealed that CGM1 transcripts could only be detected in normal granulocytes and leukocytes of CML patients, but not in lymphocytes, monocytes nor in other tissues. The exclusive presence of mRNA in cells of the granulocytic lineage is in agreement with the fact that CGM1 belongs to CD66 which has been described as a granulocyte-specific cluster of differentiation (Knapp et al., 1989). CGM1c mRNA in addition to CGM1a could only be detected in leukocytes of one CML patient. Since the ratio of different myeloid precursor cells in peripheral blood can vary in CML patients (Raskind and Fialkow, 1987), this result suggests that CGM1 splice variants are stage-specifically expressed. However, more samples have to be analysed to clarify this.

We wish to thank S. von Kleist for her continuous support, M. Weiss and B. Fischer for excellent technical assistance, A.-M. Eades-Perner for critical reading of the manuscript and B. Malissen for bringing the exon distribution of the 'YLYL' motif to our attention. We gratefully acknowledge the BGP_a and CGM7 transfectants from M. Kuroki and Y. Matsuoka and the mAb 73 from F. Buchegger and J.-P. Mach. We also wish to thank H. Mohrenweiser for supplying the cosmid clones. This work was supported by grants from the *Dr. Mildred Scheel Stiftung für Krebsforschung* and *Deutsche Forschungsgemeinschaft*.

REFERENCES

- Afar, D. E. H., Stanners, C. P. & Bell, J. C. (1992) Tyrosine phosphorylation of biliary glycoprotein, a cell adhesion molecule related to carcinoembryonic antigen, *Biochim. Biophys. Acta* **1134**, 46–52.
- Artelt, P., Morelle, C., Ausmeier, M., Fritzek, M. & Hauser, H. (1988) Vector of efficient expression in mammalian fibroblastoid, myeloid and lymphoid cell via transfection or infection, *Gene* **68**, 213–219.
- Audette, M., Buchegger, F., Schreyer, M. & Mach, J.-P. (1987) Monoclonal antibody against carcinoembryonic antigen (CEA) identifies two new forms of crossreacting antigens of molecular mass 90 and 160 kDa in normal granulocytes, *Mol. Immunol.* **24**, 1177–1186.
- Baniyash, M., Hsu, V. W., Seldin, M. F. & Klausner, R. D. (1989) The isolation and characterization of the murine T cell antigen receptor zeta chain gene, *J. Biol. Chem.* **264**, 13252–13257.
- Barnett, T. R., Kretschmer, A., Austen, D. A., Goebel, S. J., Hart, J. T., Elting, J. J. & Kamarck, M. E. (1989) Carcinoembryonic antigens: Alternative splicing accounts for the multiple mRNAs that code for novel members of the carcinoembryonic antigen family, *J. Cell Biol.* **108**, 267–276.
- Benchimol, S., Fuks, A., Jothy, S., Beauchemin, N., Shirota, K. & Stanners, C. P. (1989) Carcinoembryonic antigen, a human tumor marker, functions as an intercellular adhesion molecule, *Cell* **57**, 327–334.
- Berling, B., Kolbinger, F., Grunert, F., Thompson, J. A., Brombacher, F., Buchegger, F., von Kleist, S. & Zimmermann, W. (1990) Cloning of a carcinoembryonic antigen family member expressed in leukocytes of chronic myeloid leukaemia patients and bone marrow, *Cancer Res.* **50**, 6534–6539.
- Chomczynski, P. & Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate/phenol/chloroform extraction, *Anal. Biochem.* **162**, 156–159.
- Culic, O., Huang, Q.-H., Flanagan, D., Hixson, D. & Lin, S.-H. (1992) Molecular cloning and expression of a new rat liver cell-CAM105 isoform: differential phosphorylation of isoforms, *Biochem. J.* **285**, 47–53.
- Drzeniek, Z., Lamerz, R., Fenger, U., Wagener, C. & Haubeck, H. D. (1991) Identification of membrane antigens in granulocytes and colonic carcinoma cells by a monoclonal antibody specific for biliary glycoprotein, a member of the carcinoembryonic antigen family, *Cancer Lett.* **56**, 173–179.
- van der Elsen, P., Georgopoulos, K., Shepley, B.-A. & Orkin, S. (1986) Exon/intron organization of the genes coding for the delta chains of human and murine T-cell receptor/T3 complex, *Proc. Natl Acad. Sci. USA* **83**, 2944–2948.
- Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M. & Danielsen, M. (1987) Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure, *Proc. Natl Acad. Sci. USA* **84**, 7413–7417.
- Fiddes, J. C. & Goodman, H. M. (1979) Isolation, cloning and sequence analysis of the cDNA for the α -subunit of human chorionic gonadotropin, *Nature* **281**, 351–356.
- Hibbs, M. L., Jakes, S., Stackner, S. A., Wallace, R. W. & Springer, T. A. (1991) The cytoplasmic domain of the integrin lymphocyte function-associated antigen 1 β subunit: sites required for binding to the intercellular adhesion molecule 1 and the phorbol-ester-stimulated phosphorylation site, *J. Exp. Med.* **174**, 1227–1238.
- Keegan, A. D. & Paul, W. E. (1992) Multichain immune recognition receptors: similarities in structure and signaling pathways, *Immunol. Today* **13**, 63–68.
- Knapp, W., Dörken, B., Rieber, P., Schmidt, R. E., Stein, H. & Kr. von dem Borne, A. E. G. (1989) CD antigens 1989, *Int. J. Cancer* **44**, 190–191.
- Kolbinger, F., Schwarz, K., Brombacher, F., von Kleist, S. & Grunert, F. (1989) Expression of an NCA cDNA in NIH/3T3 cells yields a 110-kDa glycoprotein, which is anchored into the membrane via glycosylphosphatidylinositol, *Biochem. Biophys. Res. Commun.* **161**, 1126–1134.
- Kuijpers, T. W., Hoogerwerf, M., van der Laan, L., Nagel, G. & van der Schoot, E. (1992) CD66 non-specific crossreacting antigens are involved in neutrophil adherence to cytokine-activated endothelial cells, *J. Cell Biol.* **118**, 457–466.
- Kuroki, M., Matsuo, Y., Kuroki, M. & Matsuoka, Y. (1990a) Non-specific cross-reacting antigen (NCA) expressed by human granulocytes: six species with different peptide sizes and membrane anchoring forms, *Biochem. Biophys. Res. Commun.* **166**, 701–708.
- Kuroki, M., Matsuo, Y., Ohtani, T., Minowada, J., Kuroki, M. & Matsuoka, Y. (1990b) A novel CEA-cross-reacting antigen of molecular mass 140 kDa expressed on human lymphoid cell lines, *Mol. Immunol.* **27**, 689–696.
- Kuroki, M., Arakawa, F., Matsuo, Y., Oikawa, S., Misumi, Y., Nakazato, H. & Matsuoka, Y. (1991) Molecular cloning of nonspecific cross-reacting antigens in human granulocytes, *J. Biol. Chem.* **266**, 11810–11817.
- Kuroki, M., Matsuo, Y., Kinugasa, T. & Matsuoka, Y. (1992) Three different NCA species, CGM6/CD67, NCA-95, and NCA-90, are comprised in the major 90–100-kDa band of granulocyte NCA detectable upon SDS-polyacrylamide gel electrophoresis, *Biochem. Biophys. Res. Commun.* **182**, 501–506.
- Leibrock, J., Lottspeich, F., Hohn, A., Hofer, M., Hengerer, B., Masiakowski, P., Thoenen, H. & Barde, Y. A. (1989) Molecular

- cloning and expression of brain-derived neurotrophic factor, *Nature* 341, 149–152.
- Leiserson, W. (1985) *Immunological Methods III* (Lefkovits, I. & Pernis, B., eds) pp. 289–315, Academic Press, New York.
- Lin, S.-H. & Guidotti, G. (1989) Cloning and expression of a cDNA coding for a rat liver plasma membrane ecto-ATPase, *J. Biol. Chem.* 264, 14408–14414.
- Lucas, K., Busch, M., Mössinger, S. & Thompson, J. A. (1991) An improved microcomputer program for finding gene-, or gene family-specific oligonucleotides suitable as primers for polymerase chain reactions or as probes, *CABIOS* 7, 525–529.
- Margolis, R. N., Schell, M. J., Taylor, S. I. & Hubbard, A. L. (1990) Hepatocyte plasma membrane ecto-ATPase (pp120/HA4) is a substrate for tyrosine kinase activity of the insulin receptor, *Biochem. Biophys. Res. Commun.* 166, 562–566.
- McCuaig, K., Turbide, C. & Beauchemin, N. (1992) mmCGM1a: a mouse carcinoembryonic antigen gene family member, generated by alternative splicing, functions as an adhesion molecule, *Cell Growth Differ.* 3, 165–174.
- Nap, M., Hammarström, M. L., Börner, O., Hammarström, S., Wagener, C., Handt, S., Schreyer, M., Mach, J. P., Buchegger, F., von Kleist, S., Grunert, F., Sequin, P., Fuks, A., Holm, R. & Lamerz, R. (1992) Specificity and affinity of monoclonal antibodies against carcinoembryonic antigen, *Cancer Res.* 52, 2329–2339.
- Neumaier, M., Zimmermann, W., Shively, L., Hinoda, Y., Riggs, A. D. & Shively, J. E. (1988) Characterization of a cDNA clone for the nonspecific cross-reacting antigen (NCA) and a comparison of NCA and carcinoembryonic antigen (CEA), *J. Biol. Chem.* 263, 3202–3207.
- Ng, S. Y., Gunning, P., Eddy, R., Ponte, P., Leavitt, J., Shows, T. & Kedes, L. (1985) Evolution of the functional human β -actin gene and its multi-pseudogene family: Conservation of noncoding regions and chromosomal dispersion of pseudogenes, *Mol. Cell. Biol.* 5, 2720–2732.
- Oikawa, S., Inuzuka, C., Kuroki, M., Matsuoka, Y., Kosaki, G. & Nakazato, H. (1989) Cell adhesion activity of non-specific cross-reacting antigen (NCA) and carcinoembryonic antigen (CEA) expressed on CHO cell surface: homophilic and heterophilic adhesion, *Biochem. Biophys. Res. Commun.* 164, 39–45.
- Oikawa, S., Inuzuka, C., Kuroki, M., Arakawa, F., Matsuoka, Y., Kosaki, G. & Nakazato, H. (1991) A specific heterotypic cell adhesion activity between members of carcinoembryonic antigen family, W272 and NCA, is mediated by N-domains, *J. Biol. Chem.* 266, 7995–8001.
- Paxton, R., Mooser, G., Pande, H., Lee, T. D. & Shively, J. E. (1987) Sequence analysis of carcinoembryonic antigen: identification of glycosylation sites and homology with the immunoglobulin supergene family, *Proc. Natl Acad. Sci. USA* 84, 920–924.
- Pèlegri, A., Terskikh, A., Hayoz, D., Chalandon, Y., Olsson, N.-O., Folli, S., Buchegger, F., Kromer, B., Schwarz, K., Martin, M., Martin, F. & Mach, J.-P. (1992) Human carcinoembryonic antigen cDNA expressed in rat carcinoma cells can function as a target antigen for tumor localization of antibodies in nude rats and as rejection antigen in syngeneic rats, *Int. J. Cancer* 52, 110–119.
- Pinna, L. A. (1990) Casein kinase 2: an 'eminence grise' in cellular regulation? *Biochim. Biophys. Acta* 1054, 267–284.
- Raskind, W. H. & Fialkow, P. J. (1987) The use of cell markers in the study of human hematopoietic neoplasia, *Adv. Cancer Res.* 49, 127–167.
- Rojas, M., Fuks, A. & Stanners, C. P. (1990) Biliary glycoprotein (BGP), a member of the immunoglobulin supergene family, functions *in vitro* as a Ca^{2+} -dependent intercellular adhesion molecule, *Cell Growth Differ.* 1, 527–533.
- Roos, D. & de Boer, M. (1986) Purification and cryopreservation of phagocytes from human blood, *Methods Enzymol.* 132, 225–243.
- Sambrook, J. E., Fritsch, J. & Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY.
- Schwarz, K., Brückel, N., Schwaibold, H., von Kleist, S. & Grunert, F. (1989) Non-specific cross-reacting antigen: characterization of specific and cross-reacting epitopes, *Mol. Immunol.* 26, 467–475.
- Shively, J. E. & Beatty, J. D. (1985) CEA-related antigens: molecular biology and clinical significance, *CRC Crit. Rev. Oncol. Hematol.* 2, 355–399.
- Skubitz, K. M., Ducker, T. P. & Goueli, S. A. (1992) CD66 monoclonal antibodies recognize a phosphotyrosine-containing protein bearing a carcinoembryonic antigen crossreacting antigen on the surface of human neutrophils, *J. Immunol.* 148, 852–860.
- Southern, P. J. & Berg, P. (1982) Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter, *J. Mol. Appl. Genet.* 1, 327–341.
- Thompson, J. A., Mauch, E.-M., Chen, F.-S., Hinoda, Y., Schrewe, H., Berling, B., Barnert, S., von Kleist, S., Shively, J. E. & Zimmermann, W. (1989) Analysis of the size of the carcinoembryonic antigen (CEA) gene family: isolation and sequencing of N-terminal domain exons, *Biochem. Biophys. Res. Commun.* 158, 996–1004.
- Thompson, J. A., Grunert, F. & Zimmermann, W. (1991) The carcinoembryonic antigen gene family: molecular biology and clinical perspectives, *J. Clin. Lab. Anal.* 5, 344–366.
- Turbide, C., Rojas, M., Stanners, C. P. & Beauchemin, N. (1991) A mouse carcinoembryonic antigen (CEA) gene family member is a calcium dependent cell adhesion molecule, *J. Biol. Chem.* 266, 309–315.
- Tynan, K., Olsen, A., Trask, B., de Jong, P., Thompson, J., Zimmermann, W., Carrano, A. & Mohrenweiser, H. (1992) Assembly and analysis of cosmid contigs in the CEA-gene family region of human chromosome 19, *Nucleic Acids Res.* 20, 1629–1636.
- Watt, S. M., Sala-Newby, G., Hoang, T., Gilmore, D. J., Grunert, F., Nagel, G., Murdoch, S. J., Tchilian, E., Lennox, E. S. & Waldmann, H. (1991) CD 66 identifies a neutrophil-specific epitope within the hematopoietic system that is expressed by members of the carcinoembryonic antigen family of adhesion molecules, *Blood* 78, 63–74.
- Woodgett, J. R., Gould, K. L. & Hunter, T. (1986) Substrate specificity of protein kinase C: use of synthetic peptides corresponding to physiological sites as probes for substrate recognition requirements, *Eur. J. Biochem.* 161, 177–184.
- Zhou, H., Fuks, A. & Stanners, C. P. (1990) Specificity of intercellular adhesion mediated by various members of the immunoglobulin supergene family, *Cell Growth Differ.* 1, 209–215.