

Evolution of a Tumorigenic Property Conferred by Glycophosphatidyl-Inositol Membrane Anchors of Carcinoembryonic Antigen Gene Family Members during the Primate Radiation

Fakhraddin Naghibalhossaini,^{*†} Anne D. Yoder,^{‡§||} Martin Tobi,[¶]
and Clifford P. Stanners^{*}

[†]Departments of Cell and Molecular Biology and Zoology, Northwestern University Medical School, Chicago, IL 60611; [§]Department of Zoology, Field Museum of Natural History, Chicago, IL 60605; [¶]Wayne State University, School of Medicine, Detroit VAMC, Detroit, MI 48201; and ^{*}McGill Cancer Centre and Department of Biochemistry, McGill University, Montréal, Québec, Canada H3G 1Y6

Submitted October 2, 2006; Revised December 9, 2006; Accepted January 26, 2007
Monitoring Editor: Mark Ginsberg

GPI membrane anchors of cell surface glycoproteins have been shown to confer functional properties that are different from their transmembrane (TM)-anchored counterparts. For the human carcinoembryonic antigen (CEA) family, a subfamily of the immunoglobulin superfamily, conversion of the mode of membrane linkage from TM to GPI confers radical changes in function: from tumor suppression or neutrality toward inhibition of differentiation and anoikis and distortion of tissue architecture, thereby contributing to tumorigenesis. We show here that GPI anchorage in the CEA family evolved twice independently in primates, very likely from more primitive TM anchors, by different packages of mutations. Both mutational packages, one package found in many primates, including humans, and a second, novel package found only in the Cebidae radiation of New World monkeys, give rise to efficiently processed GPI-linked proteins. Both types of GPI anchors mediate inhibition of cell differentiation. The estimated rate of nonsynonymous mutations (K_a) in the anchor-determining domain for conversion from TM to GPI anchorage in the CEA family that were fixed during evolution in these primates is 7 times higher than the average K_a in primates, indicating positive selection. These results suggest therefore that the functional changes mediated by CEA GPI anchors, including the inhibition of differentiation and anoikis, could be adaptive and advantageous.

INTRODUCTION

The human carcinoembryonic antigen (CEA) family consists of 34 highly similar gene-like sequences, closely clustered on the long arm of chromosome 19, and it includes members made up of amino-terminal V-like Ig domains followed by a variable number of I-like Ig domains and linked to the external cell surface by both GPI and transmembrane (TM) anchors (Zimmermann, personal communication; Hammarström *et al.*, 1998; Zebhauser *et al.*, 2005). GPI-anchored family members CEA and CEACAM6 (formerly NCA) demonstrate radically different functions from TM-linked member CEACAM1 (formerly BGP) in that, when

expressed before differentiation, they inhibit the differentiation of many cell types, including myogenic differentiation of rat L6 myoblasts (Eidelman *et al.*, 1993; Sreaton *et al.*, 1997), myogenic differentiation of mouse C2C12 cells (Sreaton *et al.*, 1997), neurogenic differentiation of retinoic acid-treated mouse P19 EC cells (Malette and Stanners, unpublished data), and adipogenic differentiation of C3H10T1/2 cells (DeMarte and Stanners, unpublished data), whereas CEACAM1 has no such effect (Rojas *et al.*, 1996). They also disrupt cell polarity and tissue architecture of colonic epithelial cells assessed by both in vitro and in vivo assays (Ilantzis *et al.*, 2002). In the latter assays, the ability of CEACAM6-transfected human SW-1222 colonocytes to form colonic crypts with normal architecture in “mini-colons” produced under the kidney capsule of the nude mouse was abrogated (Ilantzis *et al.*, 2002). Anoikis, the quality control mechanism that maintains tissue architecture by destroying unanchored cells by apoptosis, is also inhibited by CEA/CEACAM6 but not by CEACAM1 expression (Ordonez *et al.*, 2000; Soeth *et al.*, 2001; Duxbury *et al.*, 2004). These effects of CEA and CEACAM6 lead to an increase in tumorigenicity of transfected rat L6 myoblasts (Sreaton *et al.*, 1997) and transfected human colonocytes (Ilantzis *et al.*, 2002) when tested for tumor formation in nude mice. In addition, transgenic mice possessing human CEA and CEACAM6 genes were shown to express CEA and CEACAM6 with the same tissue specificity as humans (Chan and Stanners, 2004); these mice were much more sensitive than wild-type littermates to the carcinogen

This article was published online ahead of print in *MBC in Press* (<http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E06-10-0884>) on February 7, 2007.

Present addresses: [†] Department of Biochemistry, Shiraz University of Medical Sciences, 71345 Zand St., Shiraz, Iran; ^{||} Departments of Biology and Biological Anthropology and Anatomy, Duke University, Box 90338, Durham, NC 27708.

Address correspondence to: Clifford P. Stanners (cliff.stanners@mcgill.ca).

Abbreviations used: CC1, CEACAM1; CC1-CMO, CEACAM1-4L external domain/*Callicebus molloch* carboxy-terminal domain chimeric construct; CEA, carcinoembryonic antigen; CMO, *Callicebus molloch*; PI-PLC, phosphatidylinositol-specific phospholipase C; TM, transmembrane.

azoymethane for the production of colonic tumors (Chan *et al.*, 2006). When the colonic expression of CEA/CEACAM6 was increased in these mice (by increasing transgene copy number) to levels seen in typical human colorectal carcinomas, the mice showed colons with extensive extreme hyperplasia and dysplasia (Chan, Camacho-Leal, and Stanners, unpublished data). Mouse and rat CEACAM1, in contrast, have been demonstrated to act in some situations as tumor suppressors (Kunath *et al.*, 1995; Luo *et al.*, 1997). Clinically, ~70% of all human cancers show deregulated overexpression of CEA and/or CEACAM6, whereas CEACAM1 is usually down-regulated (Chevinsky, 1991; Neumaier *et al.*, 1993); also, the differentiation status of human colorectal carcinomas in the majority of such tumors that present in the clinic is negatively correlated with the cell surface expression levels of CEA and CEACAM6 (Ilantzis *et al.*, 1997).

The structural feature that determines these radically different effects on the cell phenotype is the CEA-derived GPI anchor, because CEACAM1 can acquire the myogenic differentiation-inhibition properties of CEA by exchanging its TM anchor for the GPI anchor of CEA and vice versa (Screaton *et al.*, 2000). Also, the carboxy-terminal domain that is cleaved during GPI processing apparently determines the biological specificity of the GPI anchor, because the NCAM-125-determined GPI anchor cannot confer these properties (Screaton *et al.*, 2000).

The CEA families of rodents and humans have been investigated in detail and show only TM-anchored members in rodents and both GPI- and TM-anchored members in humans. It has been suggested, therefore, that GPI-linked CEA family members evolved from an ancestral CEACAM1-like gene relatively recently in evolution by replication and mutation of a primordial TM exon (Stanners *et al.*, 1992, 1995; Hammarström *et al.*, 1998). We have shown recently that such mutation can be achieved relatively easily: the introduction of a stop codon alone into the TM-determining domain of human CEACAM1 converted this CEA family member into a GPI-anchored protein, albeit one that was inefficiently processed. Two further mutations upstream of the stop codon conferred the 100% processing efficiency seen in naturally occurring GPI-anchored CEA family members (Naghibalhossaini and Stanners, 2004). These considerations therefore raise an important question: can a structural feature that changes the function of CEA family members toward the perturbation of cell and tissue architecture and the inhibition of differentiation be selectively advantageous and therefore maintained in evolution? This question was addressed here by looking for key differences in CEA family membrane anchor-determining domains in species that allow tentative reconstruction of the evolution of this structural feature.

MATERIALS AND METHODS

Species Investigated

The species for which data are shown in this study are listed by their common name, scientific name, cell/tissue type investigated, and source: flying lemur, *Cynocephalus volans*, liver, Field Museum of Natural History, Chicago, IL; bat, *Tadarida brasiliensis*, Tb1lu, American Type Culture Collection, Manassas, VA; human, *Homo sapiens*, blood and chimpanzee, *Pan troglodytes*, blood, Yerkes Regional Primate Center, Atlanta, GA; rhesus macaque, *Macaca mulatta*, blood, Tulane Regional Primate Center, Covington, LA; mangabey, *Cercocebus atys*, blood, Tulane Regional Primate Center; cotton top tamarin, *Saguinus oedipus*, blood, New England Primate Center, Southborough, MA; common marmoset, *Callithrix jacchus*, blood, New England Primate Center; woolly monkey, *Lagothrix lagotherica*, blood, Los Angeles Zoo, Los Angeles, CA; spider monkey, *Ateles geoffroyi*, blood, Los Angeles Zoo; black howler monkey, *Alouatta caraya*, blood, Los Angeles Zoo; red howler monkey, *Alouatta seniculus*, blood, Los Angeles Zoo; red uakari, *Cacajao rubicundus*, blood, Los An-

geles Zoo; titi monkey, *Callicebus molloch* (CMO), blood, California Regional Primate Research Center; squirrel monkey, *Saimiri sciurus*, blood, New England Primate Center; tarsier, *Tarsisus syrichta*, liver, Duke University Primate Center, Durham, NC; and golden-crowned sifaka, *Propithecus tattersalli*, liver, Duke University Primate Center. Genomic DNA was extracted from acid citrate dextrose-anticoagulated blood, cells, or frozen tissues as described previously (Sambrook *et al.*, 1989).

Isolation and Characterization of TM Sequences

Sense, S-1 or S-2, and antisense, AS-1 or AS-2, polymerase chain reaction (PCR) primers (see Figure 1), based on the human CEA gene sequence, were used to amplify CEA-like carboxy-terminal sequences in genomic DNA of various species by using Taq polymerase (BIO/CAN Scientific Mississauga, Ontario, Canada) for 25–30 cycles at 42°C annealing and 72°C polymerization temperatures. To overcome reaction failure because of possible 3'-end mismatch of primers with template DNA and for higher fidelity of polymerization, a 5:1 mix of Taq and Pfu (Stratagene, La Jolla, CA) polymerases was used for some samples. CMO monkey cDNAs (reverse transcriptase [RT] 1–3; Figure 1) were obtained by reverse transcription-polymerase chain reaction (RT-PCR) by using Pd(N)6 random hexamer primers (Pharmacia Biotech, Piscataway, NJ) and Moloney murine leukemia virus reverse transcriptase on total RNA extracted from isolated blood leukocytes. The RT-reaction product was used for PCR amplification as described above using primer AS-2 (Figure 1) and a sense (5'-GTTTCTACTGTGTCACAATCTGCC-3') primer residing in the N-terminal IgV-like domain of human CEA family members. To obtain the upstream flanking sequence of the Ceb stop anchor-determining domain of the CMO monkey, AS-2 (Figure 1) and a sense S-2 PCR primer (Figure 1) (5'-GTACAGGTAGTCTCCT-3'), based on the sequences of human and CMO CEA family members, were used under low-stringency conditions.

Positive PCR bands from three to five pooled PCR reactions were cloned in the PCRII vector (TA cloning kit; Invitrogen, Carlsbad, CA), and multiple individual clones (Figure 1) were sequenced in both directions by using a dideoxy sequencing kit (T7 sequencing kit; Pharmacia Biotech).

Chimeric Plasmid Construct

The CC1-CMO cDNA chimera, consisting of human CEACAM1-4L (the longest splice variant of CEACAM1) extracellular domains N-A1-B1-A2 linked to the last 25 amino acids of the CMO monkey Ceb-stop GPI anchor-determining exon (see structure diagrams in Figure 5 and amino acid sequence in Figure 3) was constructed by high-fidelity Pfu polymerase PCR reactions by using appropriate primers. The chimeric protein was expected (but not shown) to use the upstream Ala residue at the site of the chimeric junction or the Ser residue two residues downstream as the ω site for cleavage and addition of the GPI anchor.

Cell Cultures and Transfection

LR-73, a Chinese hamster ovary (CHO)-derived cell line (Pollard and Stanners, 1979), and rat L6 myoblasts (Yaffe, 1968), neither of which express CEA-related proteins, were cultured in monolayer as described previously (Screaton *et al.*, 2000). Cell cultures were cotransfected with CEA or CC1-CMO cDNA inserted into the P91023B expression vector (courtesy of R. Kaufman, Genetics Institute, Boston, MA) along with pSV2Neo by calcium phosphate precipitation, as described previously (Rojas *et al.*, 1996). Pooled Geneticin (G-418, Invitrogen)-resistant colonies were sorted for high cell surface expression of CEACAM1 proteins by fluorescence-activated cell sorting (FACS) by using polyclonal rabbit antihuman CEA antibody, as described previously (Zhou *et al.*, 1993). G-418 was removed from growth media 24 h before the application of the various assays.

GPI Anchorage Assays

Cells were removed from culture dishes with Hanks' balanced salt solution lacking Ca^{2+} / Mg^{2+} but containing 0.5 mM EDTA, and duplicate aliquots of 5×10^5 cells were either treated in suspension at 37°C for 1 h with 0.2 U of phosphatidylinositol-specific phospholipase C (PI-PLC) from *Bacillus cereus* (Roche Diagnostics, Laval, Quebec, Canada) or without treatment as a control. Residual levels of cell surface CEACAM1-containing protein were measured by FACS analysis as described previously (Zhou *et al.*, 1993).

For the cold nonionic detergent solubility assay, cells were suspended for 10 min in ice-cold lysis buffer containing 1% Triton X-100 at pH 6.5 and then centrifuged at $15,000 \times g$ at 4°C for 20 min. SDS-solubilized pellet and supernatant fractions were analyzed by polyacrylamide gel electrophoresis and immunoblotting, by using rabbit polyclonal antihuman CEA antibody (Screaton *et al.*, 1997).

Myogenic Differentiation and Fusion Assay

To promote fusion and differentiation, rat L6 myoblasts were seeded at 10^4 cells/cm² in 60-mm tissue culture plastic Petri dishes in 4 ml of growth medium (GM) (DMEM plus 10% fetal bovine serum), and cultured without changing the medium for 3 d. When cells were nearly confluent, the medium was replaced with 4 ml of differentiation medium (DMEM plus 2% horse serum), and the cells were cultured for an additional 7 d. The ability of cells

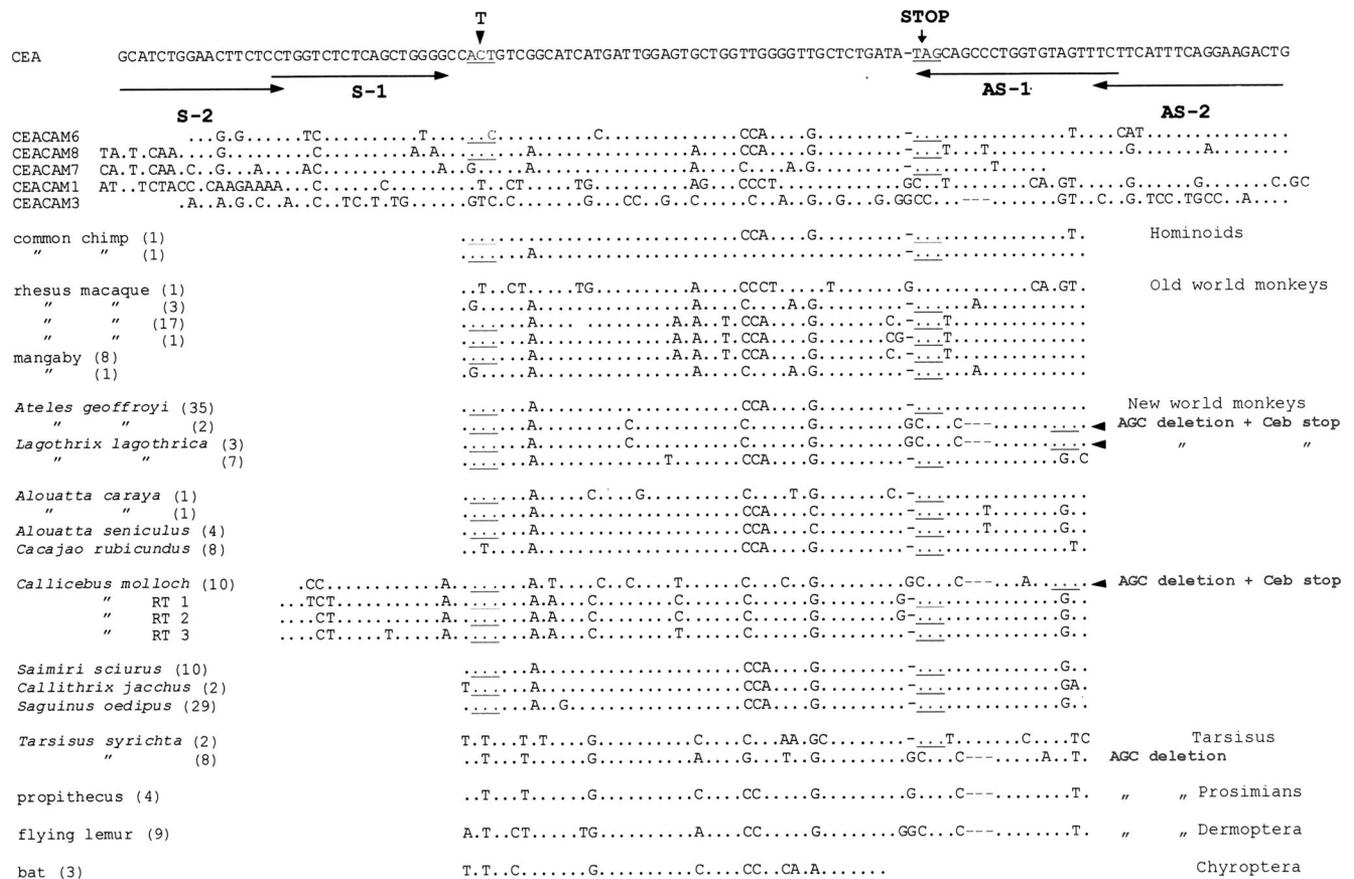


Figure 1. Nucleotide sequence alignment of anchor-determining domains of CEA family members of human and other species compared with human CEA. Identical nucleotides are indicated as dots. Dashes show naturally occurring deletions. Stop codons are underlined and marked by word "STOP." The conserved Thr codon in GPI-linked members, shown to be important in GPI processing, is underlined and marked by a bold letter T. Positions of two sets of S and AS PCR primers are shown by horizontal arrows. Numbers of sequenced independent clones of PCR products for each species are indicated in brackets. Sequences of three cloned RT-PCR-amplified bands of CMO monkey cDNA (see *Materials and Methods*) are shown as RT 1–3.

to fuse was assessed by determining the number of nuclei in cells with more than three nuclei and dividing by the total number of nuclei in all cells, as described previously (Eidelman *et al.*, 1993). For anti-myosin immunofluorescent staining, cultures were induced to differentiate in multiwell chamber slides (Nunc; Nalge Nunc, Naperville, IL). As a measure of biochemical differentiation, immunofluorescent labeling of permeabilized cells was performed as described previously (De Giovanni *et al.*, 1993) by using anti-myosin heavy chain monoclonal 47A antibody (courtesy of P. A. Merrifield, University of Western Ontario, London, Ontario, Canada).

RESULTS

Sequence Features

The known nucleotide sequences of the carboxy-terminal domains of human CEA family members (Figure 1) indicate the common presence at the same position of a stop codon in all GPI-anchored members (CEA, CEACAM6, CEACAM8, and CEACAM7), differing by a one-base pair deletion and a substitution relative to the sequences of TM-anchored members (CEACAM1 and CEACAM3). Interestingly, these mutations occurred in a region consisting of three trinucleotide (AGC) imperfect repeats, which might be expected to be a mutational hot spot due to slippage of replication enzymes (Freimer and Slatkin, 1996), and which, in fact, shows deletion of one of the first two repeats as well as other mutations in various family members (Figure 1). As mentioned above, the acquisition of a stop codon and a few upstream muta-

tions in TM-anchored CEACAM1 has been shown to be sufficient for conversion to efficient GPI anchorage.

Isolation and Characterization of Genomic Carboxy-Terminal Sequences

On the assumption that a stop codon in the carboxy-terminal domain of CEA family members is indicative of GPI linkage (Naghibalhossaini and Stanners, 2004), the carboxy-terminal exon nucleotide sequence of human CEA was used to design two sets of primers for PCR reactions with genomic DNA of various species (Figure 1). One of these sets (S-1 and AS-1) could amplify only sequences containing the human CEA-like stop codon, because the 3' end of the AS-1 primer is situated in the stop codon. The second set of primers (S-2 and AS-2) bracket the first set and therefore include the stop codon. Both sets were used to screen genomic DNA from various species.

PCR bands of the expected size were obtained only for primates and related species and were cloned and sequenced to confirm their validity and to compare their nucleotide and predicted amino acid sequences. We obtained CEA-related PCR products for three closely related mammalian orders: Microchiroptera (*T. brasiliensis*), Dermoptera (*Cynocephalus volans*), and various groups of primates (hominoids, Old World and New World monkeys, and *Tarsisus*) (Figure 1). All other

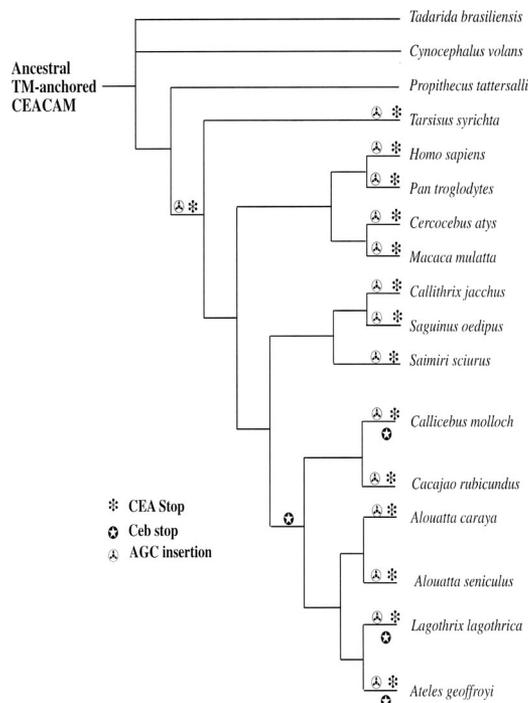


Figure 2. Evolution of key mutations in the anchor-determining exons of CEA gene family members giving GPI anchorage relative to the phylogenetic tree of various primate and related species (Porter *et al.*, 1997). Because all CEA family members from more primitive species for which nucleotide sequences are available, including mouse and rat (see Figure 3), demonstrate deletion of one AGC repeat (see Figure 1), it is assumed that the ancestral precursor CEA family member anchor-determining exon possessed this deletion. The CEA family anchor-determining exons of later species are therefore considered to have acquired an AGC insertion at this position. Because it is not possible to trace the lineage of individual CEA family members, this diagram shows only the presumed occurrence of mutations in the CEA family anchor determining exon during its evolution.

species were negative with both sets of primers with both *Taq* and *Taq/Pfu* polymerase mixes used at relatively low stringency, including tree shrew, opossum, rabbit, dog, cat, sheep, goat, cow, horse, mouse, and rat. Mouse and rat are known to possess only homologues of the human TM-anchored CEACAM1 gene and were negative because of known mismatches with the primers used. The failure to detect specific PCR products in the other species therefore does not exclude the existence of more distantly related CEACAM1-like homologues.

Molecular Evidence for a Second Independent Evolution of GPI Anchorage in the CEA Family

Comparative nucleotide sequence analysis of these CEA-like carboxy-terminal domains indicates that the package of mutations giving GPI anchorage in the present-day human CEA family, which we shall term the “CEA Stop,” arose in a common ancestor of tarsiers and anthropoids (Haplorrhini) (Figures 1 and 2). Surprisingly, a second GPI mutational package arose: three species in the Cebidae family of New World monkeys (*Callicebus molloch*, *Lagothrix lagothrica*, and *Ateles geoffroyi*) showed a CEA family member with a novel anchor-determining exon containing a different type of stop codon-generating mutational mechanism, consisting mainly of a triplet deletion followed by a new stop codon four

amino acids downstream of the CEA stop codon. We shall term this GPI mutational package the “Ceb Stop.”

The predicted amino acid sequences for these anchor-determining domains are shown in Figure 3. Both the sequences with the CEA package and the Ceb package possess, in addition to the stop codons, upstream changes (relative to the sequences of TM-anchored domains) necessary for efficient GPI processing (Englund, 1993). The latter (Figure 3, open arrowheads) represent changes toward hydrophilicity, one of which, the I-to-T change (first bolded amino acid in consensus sequences; also underlined in Figure 1 as the ACT codon), has been shown directly to markedly enhance the GPI processing efficiency of CEACAM1 mutated to include the CEA stop (from 5 to 56%) (Naghbalhossaini and Stanners, 2004). The “consensus” sequences obtained from all species for five such residues (indicated in bold in Figure 3), including the above-mentioned T residue, are strikingly similar for both the CEA-stop and Ceb-stop GPI domains, and they differ from these residues in the consensus sequence of TM domains. These observations predict that the Ceb package would also signal to give efficient GPI processing.

The Novel Mutational Ceb Package Gives Efficient GPI Processing

To verify experimentally whether the Ceb mutational package leads to the expression of a cell surface GPI-linked protein, a chimeric cDNA consisting of 25 amino acids from the carboxy-terminal exon of the CMO monkey containing the Ceb package linked to the extracellular domains of human TM-anchored CEACAM1-4L was constructed. This chimeric cDNA, denoted “CC1-CMO,” was stably transfected into the CHO-derived cell line LR-73, and it was expected to use the upstream Ala residue at the site of the chimeric junction or the Ser residue two residues downstream as the ω site for cleavage and addition of the GPI anchor (Figure 3). The cell surface localization of the expressed CC1-CMO chimeric protein was confirmed by FACS analysis of whole transfectant cells; the GPI cell surface anchorage of the protein was demonstrated by its complete sensitivity to PI-PLC digestion (Figure 4A).

The GPI linkage and efficiency of GPI processing of the CC1-CMO protein was verified by a cold nonionic detergent solubilization assay. GPI-linked proteins form a complex in the cell membrane with sphingolipids and cholesterol, which is insoluble in nonionic detergents such as Triton X-100 at 4°C (Brown and London, 1997; Simons and Ikonen, 1997). TM-linked proteins, in contrast, are soluble under the same conditions. Immunoblot analysis of soluble and insoluble fractions after cold Triton X-100 extraction of LR-73 transfectant cells showed that TM-anchored CEACAM1-4L (denoted CC1-4L) was soluble and GPI-anchored CEA was mainly insoluble, as expected. Under the same conditions, CC1-CMO was mainly insoluble, indicating GPI anchorage (Figure 4B). We have shown that incomplete GPI-processing of CEACAM1 proteins with mutant anchor-determining domains could be distinguished by the presence of a less glycosylated lower-molecular weight, Triton X-100-soluble band that is localized intracellularly (Naghbalhossaini and Stanners, 2004). Because no such band could be seen in the immunoblot of the CC1-CMO chimeric protein (Figure 4B), we conclude that this chimeric protein, as with all of the naturally occurring GPI-linked human CEA family members (Naghbalhossaini and Stanners, 2004), is efficiently GPI processed.

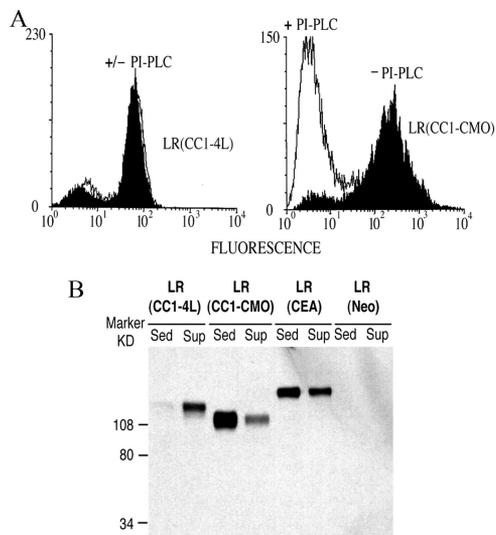


Figure 4. Evaluation of GPI-processing efficiency of the CC1-CMO protein by PI-PLC sensitivity (A) and cold nonionic detergent solubility (B). (A) CC1-CMO, CEACAM1-4L (CC1-4L), and control Neo LR-73 transfectants were treated with PI-PLC, and cell surface levels before (black profiles) and after (white profiles) treatment were assessed by FACS analysis. (B) Western blot analysis of sedimented and supernatant fractions of cold Triton X-100 extracts of CC1-CMO, CC1-4L, CEA, and control Neo LR-73 transfectants.

fectants as positive and negative controls, respectively. The results indicate that although L6 (Neo) (vector only control) and L6 (CEACAM1) transfectants started to fuse on day 3 and showed >70–80% fusion after 7 d in differentiation medium, L6 (CC1-CMO) transfectants, like CEA transfectants, were incapable of fusing and forming myotubes, even after 9 d in differentiation medium (Figure 5, top). The complete block in morphological myogenic differentiation by CC1-CMO was confirmed at the biochemical level by the complete absence of staining with anti-myosin antibody (Figure 5, bottom). These results demonstrate that the novel Ceb GPI anchor found in CMO monkeys can, like the human CEA-derived GPI anchor, confer the property of abrogation of myogenic differentiation.

DISCUSSION

Our results are most easily interpreted under the assumption that the acquisition of GPI anchorage in the CEA family during evolution most likely occurred by mutation of the TM domain of more primitive TM-anchored CEA family members. The evidence for this view includes parsimony analysis of human CEA family members (Hammarström *et al.*, 1998) and “silent site” mutation analysis of rat/mouse and human CEA family sequences (Zimmermann, 1998), both of which put TM-anchored CEA family members as the ancestral genes from which all others evolved. In addition, exhaustive searches have failed to discover GPI-anchored CEA family members in mouse and rat (Zimmermann, 1998; Kammerer *et al.*, 2004), and the PCR analysis carried out in this work failed to detect GPI anchor-determining sequences in any species except for those in the primate radiation and some close relatives; also, anchor-determining sequences of some very primitive primates, such as *propithecus* and primate-related species such as flying lemur, were found to be of the TM type (Figure 1). It is thus reasonable to assume that CEA family GPI anchorage evolved in the primate

radiation by mutation from TM type anchor-determining sequences.

The second novel package of mutations giving efficient GPI anchorage found in the Cebidae radiation of New World monkeys was presumably also derived from a primitive TM exon, but one lacking a AGC repeat, that existed before the primate radiation in a common ancestor of *Dermoptera*, *Propithecus*, and *Tarsius*. These anchorage-determining mutations are related to current well-corroborated phylogeny (Porter *et al.*, 1997) in Figure 2. The failure to detect the Ceb package in the *Alouatta* branch of the Cebidae family is probably due to the relatively few independent TM exon clones that could be obtained for these species.

Did these two independent GPI anchor-determining mutational packages arise and become fixed in the primate germline by chance? There are two major views regarding the possible mechanism of amino acid replacements in proteins in the course of evolution: either the amino acid substitutions have had selective advantage by which they improve the functional attributes of the molecule or they have been replaced by chance (random genetic drift), being selectively neutral (Kreitman and Akashi, 1995). The latter seems unlikely for the following reasons.

First, conversion of a TM-linked to a GPI-linked protein requires highly ordered amino acid replacements. Only certain amino acids in specific regions of the C-terminal domain of proteins are compatible with GPI-processing (Englund, 1993; Udenfriend and Kodukula, 1995). Also, a major change in protein structure such as conversion of the mode of anchorage from TM to GPI and its functional consequences could not be considered as a neutral event. Although the complexity of the CEA gene family, with its multiple closely related genes arising presumably by a series of duplications, renders very difficult the tracing of gene lineage, it is nonetheless possible to arrive at a consensus sequence for the carboxy-terminal exon of TM-anchored members, which, even including mouse and rat sequences, is remarkably constant (Figure 3). The consensus sequences for both CEA and Ceb package GPI exons, aside from the position of the stop codon, are also remarkably similar. Both of these, however, differ markedly at certain key positions from the TM consensus sequence (Figure 3). It seems highly improbable that these changes could have occurred twice by chance. Thus, our recent work has shown that the introduction of a stop codon into the TM-anchored domain of splice variant CEACAM1-4L can give GPI linkage but with a processing efficiency of only 10%; further mutations that replace the normally hydrophobic amino acids in the TM consensus sequence downstream of the new GPI cleavage site with hydrophilic residues are required for efficient GPI membrane linkage (Udenfriend and Kodukula, 1995; Naghibalhossaini and Stanners, 2004). Mutations in these positions, such as I688T, G690S, V692T, and I693T, representing changes in four of six contiguous amino acid residues, which extend the hydrophilic spacer while retaining a downstream short hydrophobic stretch of amino acids, are necessary for efficient GPI processing (Udenfriend and Kodukula, 1995; Naghibalhossaini and Stanners, 2004). These are exactly the residues that differ between the TM and both GPI consensus sequences (Figure 3). Intuitively, the probability that different stop codons along with four substitutions all affecting efficient GPI linkage arose randomly twice during evolution without significant selection by adaptive advantage seems extremely low (Zhang and Kumar, 1997).

Second, independently of the double-independent occurrence during primate evolution of GPI anchorage, a comparison of the nucleotide sequences of a typical more primitive

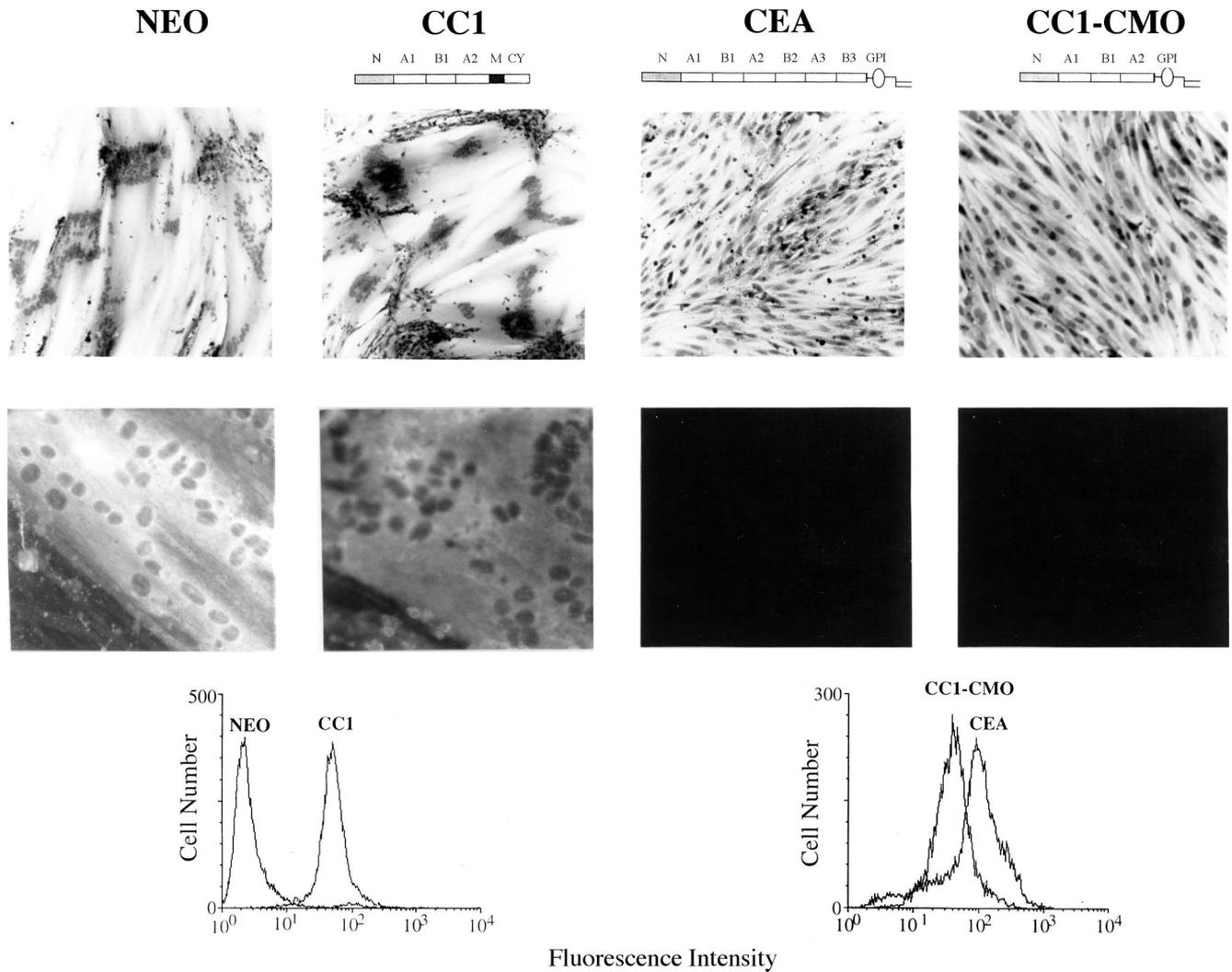


Figure 5. Effect of the CC1-CMO protein on myogenic differentiation of L6 myoblasts at the morphological and biochemical levels. Top, photomicrographs of hematoxylin-stained cultures of various rat L6 myoblast transfectants incubated in differentiation medium for 7 d. Bottom, photomicrographs obtained by fluorescent microscopy of cultures incubated for 4 d in differentiation medium and stained by anti-myosin monoclonal antibody and fluorescein isothiocyanate-conjugated, anti-mouse IgG. FACS profiles show the relative cell surface expression levels of the indicated proteins. The results were reproducible and were repeated for at least two independently isolated, pooled, total transfectant populations.

TM-anchored exon (*Tarsisus*) with that of a typical Ceb GPI-anchored exon (*Callicebus*) (both lacking one AGC repeat) yields a high rate of nonsynonymous amino acid substitutions (see *Results*). Such an accelerated rate of nonsynonymous substitution implies positive selection (Kreitman and Akashi, 1995) for conversion of the TM-anchored CEACAM into GPI-anchored proteins. The carboxy-terminal exon of the CEA family and the closely related PSG family seems to represent an exception to the other exons in that here there is evidence of sequence conservation and adaptive amino acid substitution (Streydio *et al.*, 1990). An accelerated directional amino acid replacement, which gave rise to the GPI anchors of the CEA family, suggests that the carboxy-terminal domain of this gene was subject to non-neutral mechanisms of change. It thus seems that the GPI anchor itself might be a critical element in the evolution of function in the CEA family.

An important factor in this analysis is that GPI anchors derived from the expression of different carboxy-terminal exons can determine specificity of function (see Introduc-

tion). The Ceb package GPI anchor-determining domain of *C. molloch* was, in fact, found to completely block L6 myogenic differentiation when linked to CEACAM1 external domains (CC1-CMO construct), exactly as was shown for the common human CEA GPI anchor linked to the same external domains (Screaton *et al.*, 2000). It has previously been shown that the external domains of NCAM-125, when attached to the GPI-determining carboxy-terminal domain of CEA, also suffice to block myogenic differentiation (although the NCAM GPI-determining carboxy-terminal domain does *not* suffice) (Screaton *et al.*, 2000). Mutations in the external domains of CEA that obliterate self-binding delete the differentiation blocking property (TaHERI *et al.*, 2003). This tumorigenic effect seems to be, therefore, specific to CEA family-derived GPI anchor attached to nonspecific self-binding external domains. There is considerable diversity in the structural modifications found on GPI glycolipids within and between species and cell types, although the core structure is conserved (McConville and Menon, 2000), which presumably confers this specificity.

As Figure 1 shows, there are both types of GPI-encoding package of mutations in individual members of the Cebidae family. An important question would be why a new (Ceb) GPI-encoding CEACAM sequence evolved when there was already a CEA-type GPI-encoding motif present in this group of primates. It has been suggested that one of the mechanisms by which members of a pair of gene duplicates can escape mutational decay is subfunctionalization (Lynch and Force, 2000). By this mechanism, for example, a gene that is originally expressed in two tissues may diverge into two copies, each being expressed uniquely in one of the two tissues. Once such a partitioning of expression pattern has become fixed in a population, the two copies will be maintained indefinitely by natural selection. In our case, a primitive TM-anchored CEACAM could duplicate, and each copy could become tissue-specifically expressed and then mutate independently to become GPI-anchored, with each giving selective advantage in its respective tissue. The tissue specific pattern of expression of the new GPI-determining exon in New World monkeys could provide additional insight regarding its *in vivo* function.

Thus, with the likely assumption that GPI anchor acquisition conferred positive selection during evolution, the radical changes in function—from tumor suppression or neutrality toward suspension of differentiation and allowance of more freedom in tissue architecture, thereby contributing to adult tumorigenesis—must also have contributed selective advantage. Sequence convergence leading to structural and functional convergence may have occurred. Lipid-like GPI structures allow more lateral movement in the membrane than TM anchors, because they are not linked to the cell cytoskeleton through cytoplasmic domains, and they would be expected to change radically the repertoire of cell surface elements with which CEA family members could interact. These novel interactions can apparently confer a cell and tissue state maintaining a lack of architecture and inhibition of cellular specialization (Benchimol *et al.*, 1989). Unlike many oncogenes, this is not achieved by stimulating cellular proliferation. CEA is expressed at higher levels during embryonic development (Benchimol *et al.*, 1989). We speculate that such a cell and tissue state, although it exacts the price of increased predilection for cancer in adult life, is advantageous during development where delay of specialization and structure by the explicit deployment of GPI-anchored CEA family members could offer improved temporal and spatial control of morphogenesis.

ACKNOWLEDGMENTS

We thank Drs. Calvin A. Porter (Department of Biology, Xavier University, New Orleans, LA), Laurence Haney (Field Museum of Natural History, Chicago, IL), Neal Clapp (Marmoset Research Center, Oak Ridge Associated Universities, Oak Ridge, TN), and institutions listed in *Materials and Methods* for provision of specimens used in this study. We thank Luisa DeMarte for extensive help in the preparation of the manuscript. This work was supported by grants to C.P.S. from the Canadian Institutes of Health Research and the National Cancer Institute of Canada with funds from the Canadian Cancer Society of Canada.

REFERENCES

- Benchimol, S., Fuks, A., Jothy, S., Beauchemin, N., Shirota, K., and Stanners, C. P. (1989). Carcinoembryonic antigen, a human tumor marker, functions as an intercellular adhesion molecule. *Cell* 57, 327–334.
- Brown, D. A., and London, E. (1997). Structure of detergent-resistant membrane domains: does phase separation occur in biological membranes? *Biochem. Biophys. Res. Commun.* 240, 1–7.
- Chan, C. H., Cook, D., and Stanners, C. P. (2006). Increased colon tumor susceptibility in azoxymethane treated CEABAC transgenic mice. *Carcinogenesis* 27, 1909–1916.
- Chan, C. H., and Stanners, C. P. (2004). Novel mouse model for carcinoembryonic antigen-based therapy. *Mol. Ther.* 9, 775–785.
- Chevinsky, A. H. (1991). CEA in tumors of other than colorectal origin. *Semin. Surg. Oncol.* 7, 162–166.
- De Giovanni, C., Lollini, P. L., Dolcetti, R., Landuzzi, L., Nicoletti, G., D'Andrea, E., Scotland, K., and Nanni, P. (1993). Uncoupling of growth inhibition and differentiation in dexamethasone-treated human rhabdomyosarcoma cells. *Br. J. Cancer* 67, 674–679.
- Duxbury, M. S., Ito, H., Zinner, M. J., Ashley, S. W., and Whang, E. E. (2004). CEACAM6 gene silencing impairs anoikis resistance and *in vivo* metastatic ability of pancreatic adenocarcinoma cells. *Oncogene* 23, 465–473.
- Eidelman, F. J., Fuks, A., DeMarte, L., Taheri, M., and Stanners, C. P. (1993). Human carcinoembryonic antigen, an intercellular adhesion molecule, blocks fusion and differentiation of rat myoblasts. *J. Cell Biol.* 123, 467–475.
- Englund, P. T. (1993). The structure and biosynthesis of glycosyl phosphatidylinositol protein anchors. *Annu. Rev. Biochem.* 62, 121–138.
- Freimer, N. B., and Slatkin, M. (1996). Microsatellites: evolution and mutational processes. *Ciba Found. Symp.* 197, 51–67; discussion 67–72.
- Gillespie, J. H. (1991). Protein evolution. In: *The Causes of Molecular Evolution*, New York: Oxford University Press, 3–63.
- Hammarström, S., Olsen, A., Teglund, S., and Baranov, V. (1998). The nature and expression of the human CEA family: basic and clinical perspectives. In: *Cell Adhesion and Communication Mediated by the CEA Family*, ed. C. P. Stanners, Amsterdam, The Netherlands: Hardwood Academic Publishers, 1–30.
- Ilantzis, C., De Marte, L., Sreaton, R. A., and Stanners, C. P. (2002). Deregulated expression of the human tumor marker CEA and CEA family member CEACAM6 disrupts tissue architecture and blocks colonocyte differentiation. *Neoplasia* 4, 151–163.
- Ilantzis, C., Jothy, S., Alpert, L. C., Draber, P., and Stanners, C. P. (1997). Cell-surface levels of human carcinoembryonic antigen are inversely correlated with colonocyte differentiation in colon carcinogenesis. *Lab. Invest.* 76, 703–716.
- Kammerer, R., Popp, T., Singer, B. B., Schlender, J., and Zimmermann, W. (2004). Identification of allelic variants of the bovine immune regulatory molecule CEACAM1 implies a pathogen-driven evolution. *Gene* 339, 99–109.
- Kreitman, M., and Akashi, H. (1995). Molecular evidence for natural selection. *Annu. Rev. Ecol. Syst.* 26, 403–422.
- Kunath, T., Ordóñez-García, C., Turbide, C., and Beauchemin, N. (1995). Inhibition of colonic tumor cell growth by biliary glycoprotein. *Oncogene* 11, 2375–2382.
- Luo, W., Wood, C. G., Earley, K., Hung, M. C., and Lin, S. H. (1997). Suppression of tumorigenicity of breast cancer cells by an epithelial cell adhesion molecule (C-CAM1): the adhesion and growth suppression are mediated by different domains. *Oncogene* 14, 1697–1704.
- Lynch, M., and Force, A. (2000). The probability of duplicate gene preservation by subfunctionalization. *Genetics* 154, 459–473.
- McConville, M. J., and Menon, A. K. (2000). Recent developments in the cell biology and biochemistry of glycosylphosphatidylinositol lipids (review). *Mol. Membr. Biol.* 17, 1–16.
- Naghibalhossaini, F., and Stanners, C. P. (2004). Minimal mutations are required to effect a radical change in function in CEA family members of the Ig superfamily. *J. Cell Sci.* 117, 761–769.
- Nei, M., and Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3, 418–426.
- Neumaier, M., Paululat, S., Chan, A., Matthaes, P., and Wagener, C. (1993). Biliary glycoprotein, a potential human cell adhesion molecule, is down-regulated in colorectal carcinomas. *Proc. Natl. Acad. Sci. USA* 90, 10744–10748.
- Ohta, T. (1997). Role of random genetic drift in the evolution of interactive systems. *J. Mol. Evol.* 44, S9–S14.
- Ordóñez, C., Sreaton, R. A., Ilantzis, C., and Stanners, C. P. (2000). Human carcinoembryonic antigen functions as a general inhibitor of anoikis. *Cancer Res.* 60, 3419–3424.
- Pollard, J. W., and Stanners, C. P. (1979). Characterization of cell lines showing growth control isolated from both the wild type and a leucyl-tRNA synthetase mutant of Chinese hamster ovary cells. *J. Cell. Physiol.* 98, 571–585.
- Porter, C. A., Page, S. L., Czelusniak, J., Schneider, H., Schneither, M.P.C., Sampio, I., and Goodman, M. (1997). Phylogeny and evolution of selected primates as determined by sequences of the ϵ -globin locus and 5' flanking regions. *Int. J. Primatol.* 18, 261–295.

- Rojas, M., DeMarte, L., Sreaton, R. A., and Stanners, C. P. (1996). Radical differences in functions of closely related members of the human carcinoembryonic antigen gene family. *Cell Growth Differ.* 7, 655–662.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Analysis and cloning of eukaryotic genomic DNA*, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 9.2–9.62.
- Sreaton, R. A., DeMarte, L., Draber, P., and Stanners, C. P. (2000). The specificity for the differentiation blocking activity of carcinoembryonic antigen resides in its glycosylphosphatidylinositol anchor. *J. Cell Biol.* 150, 613–626.
- Sreaton, R. A., Penn, L. Z., and Stanners, C. P. (1997). Carcinoembryonic antigen, a human tumor marker, cooperates with Myc and Bcl-2 in cellular transformation. *J. Cell Biol.* 137, 939–952.
- Simons, K., and Ikonen, E. (1997). Functional rafts in cell membranes. *Nature* 387, 569–572.
- Soeth, E., Wirth, T., List, H. J., Kumbhani, S., Petersen, A., Neumaier, M., Czubayko, F., and Juhl, H. (2001). Controlled ribozyme targeting demonstrates an antiapoptotic effect of carcinoembryonic antigen in HT29 colon cancer cells. *Clin. Cancer Res.* 7, 2022–2030.
- Stanners, C. P., DeMarte, L., Rojas, M., Gold, P., and Fuks, A. (1995). Opposite functions for two classes of genes of the human carcinoembryonic antigen family. *Tumour Biol.* 16, 23–31.
- Stanners, C. P., Rojas, M., Zhou, H., Fuks, A., and Beauchemin, N. (1992). The CEA family: a system in transitional evolution? *Int. J. Biol. Markers* 7, 137–142.
- Streydio, C., Swillens, S., Georges, M., Szpirer, C., and Vassart, G. (1990). Structure, evolution and chromosomal localization of the human pregnancy-specific beta 1-glycoprotein gene family. *Genomics* 7, 661–662.
- Taheri, M., Saragovi, H. U., and Stanners, C. P. (2003). The adhesion and differentiation-inhibitory activities of the immunoglobulin superfamily member, carcinoembryonic antigen, can be independently blocked. *J. Biol. Chem.* 278, 14632–14639.
- Udenfriend, S., and Kodukula, K. (1995). Prediction of omega site in nascent precursor of glycosylphosphatidylinositol protein. *Methods Enzymol.* 250, 571–582.
- Yaffe, D. (1968). Retention of differentiation potentialities during prolonged cultivation of myogenic cells. *Proc. Natl. Acad. Sci. USA* 61, 477–483.
- Zebhauser, R., Kammerer, R., Eisenried, A., McLellan, A., Moore, T., and Zimmermann, W. (2005). Identification of a novel group of evolutionarily conserved members within the rapidly diverging murine *Cea* family. *Genomics* 86, 566–580.
- Zhang, J., and Kumar, S. (1997). Detection of convergent and parallel evolution at the amino acid sequence level. *Mol. Biol. Evol.* 14, 527–536.
- Zhou, H., Fuks, A., Alcaraz, G., Bolling, T. J., and Stanners, C. P. (1993). Homophilic adhesion between Ig superfamily carcinoembryonic antigen molecules involves double reciprocal bonds. *J. Cell Biol.* 122, 951–960.
- Zimmermann, W. (1998). The nature and expression of the rodent CEA families: evolutionary considerations. In: *Cell Adhesion and Communication Mediated by the CEA Family*, ed. C. P. Stanners, Amsterdam, The Netherlands: Hardwood Academic Publishers, 31–55.