# Minimal mutations are required to effect a radical change in function in CEA family members of the Ig superfamily

**Fakhraddin Naghibalhossaini\* and Clifford P. Stanners‡**

Department of Biochemistry and McGill Cancer Centre, McGill University, Montreal, Quebec, Canada H3G 1Y6
\*Present address: Department of Biochemistry, Shiraz University of Medical Sciences, PO Box 1167, Zand Street 71345, Shiraz, Iran
‡Author for correspondence (e-mail: cliff.stanners@mcgill.ca)

## Summary

GPI anchorage in the CEA family results in the acquisition of radically different functions relative to TM anchorage, including inhibition of differentiation and anoikis, disruption of tissue architecture and promotion of tumorigenicity. CEA GPI anchors, as determined by the carboxy-terminal exon of CEA, demonstrate *biological specificity* in their ability to confer these functional changes. CEA family GPI anchorage appears to have evolved twice independently during the primate radiation, in a manner suggestive of evolution from more primitive TM-anchored CEACAM1. We show here that very few mutations in the TM exon of present-day human CEACAM1 are required to give efficient GPI anchorage and the biological specificity of CEA GPI anchors, i.e., to give the differentiation-blocking function of GPI-anchored CEA. Such a change in anchorage could therefore represent a relatively facile means for producing radical change in molecular function of Ig superfamily members during evolution.

Key words: CEACAM1, CEA, Evolution, GPI, Ig superfamily

## Introduction

The human carcinoembryonic antigen (CEA) gene family of cell surface glycoproteins, a subset of the immunoglobulin gene superfamily (Paxton et al., 1987), consists of 29 closely related gene and gene-like sequences clustered on the long arm of chromosome 19 (Hammarström et al., 1998). Their extracellular domains comprise an amino-terminal V-like Ig domain, followed by 0 to 6 I-like Ig domains. Their membrane anchorage is of two types: glycophosphatidyl inositol (GPI) (CEA, CEACAM6, CEACAM7, CEACAM8) and transmembrane (TM) (CEACAM1, CEACAM3, CEACAM4), thus dividing the family into two subgroups (Hammarström et al., 1998). Members of both subgroups function in vitro, at least, as intercellular homophilic and heterophilic adhesion molecules (for review, see Stanners and Fuks, 1998). The two subgroups can show opposite functional properties, however, especially regarding their roles in cancer. GPI-linked members CEA and CEACAM6 are up-regulated in about 50% of all human cancers (Chevinsky, 1991; Ilantzis et al., 1997; Jothy et al., 1993), whereas TM-linked CEACAM1 is usually down-regulated (Neumaier et al., 1993; Nollau et al., 1997). In fact, CEA has been suggested to function as a type of oncogene (Screaton et al., 1997) whereas rodent and human CEACAM1 have properties suggesting that they function as tumor suppressors (Hsieh et al., 1995; Kunath et al., 1995).

Consistent with the above, CEA and CEACAM6 (but not CEACAM1) have been shown to block differentiation of many cell types (Eidelman et al., 1993); (L. DeMarte, B. Malette and C. P. S., unpublished), to inhibit anoikis (the tissue architecture quality control mechanism that destroys, by apoptosis, cells that become detached from their basement membranes) (Ordoñez et al., 2000), and to disrupt epithelial cell polarity and tissue architecture (Ilantzis et al., 2002), thus increasing tumorigenicity (Ilantzis et al., 2002; Screaton et al., 1997). In the case of clinical cancer, purified tumor colonocytes from patients with colorectal carcinomas of intermediate differentiation status were shown to express cell surface levels of CEA and CEACAM6 that were inversely correlated with the degree of differentiation, as assessed by the attending pathologist (Ilantzis et al., 1997).

The structural requirements for these tumorigenic effects were shown to be self-binding external domains linked to the GPI anchor determined by the carboxy terminal exon of CEA (or CEACAM6) (Screaton et al., 2000), which is removed when the GPI anchor is attached. Thus the radical differences in function between CEA and CEACAM1 could be switched by switching their membrane anchors. Importantly, the CEA anchor was also demonstrated to possess biological specificity in that, whereas NCAM-125, a GPI-anchored isoform of NCAM, actually stimulated myogenic differentiation (Dickson et al., 1990), a construct consisting of the NCAM external domain linked to the carboxy-terminal exon of CEA blocked myogenic differentiation completely (Screaton et al., 2000).

This has focussed attention on the GPI anchor of CEA as the chief determinant of cancer-related cellular function. Rodents do not possess GPI-anchored CEA family members and several lines of evidence supports the view that they arose during the primate radiation from a primordial ancestral TM-anchored CEACAM1-like member (Hammarström et al., 1998; Zimmermann, 1998). In fact, it appears that CEA family GPI

**Table 1. Methods of site-directed mutagenesis of CEACAM1-4L**

| Construct | Mutation in CC1-4L protein | Primers used for mutagenesis |
|---|---|---|
| CC1-t[†] | A445/stop | Sense: 5′TGGAATCTCCATCCG*TTGGTTCTTCAA3′ |
| | | Antisense: 5′CTTGCCCTG**CCGGTCTTCCCGAAATGCAGAAAACATGCCAGGGCTACT**A-**TATCAGAGCAACC3′ |
| CC1-tSG | L419P420/SG | Sense: 5′CTATAATGCT**TC**A**GG**ACAAGAAAATG3′ |
| | | Antisense[‡]:5′TTTACGTTCAGCATGATGGGG3′ |
| CC1-tS | L419/S | Sense: 5′CTATAATGCT**TC**ACCACAAGAAAATG3′ |
| CC1-tG | P420/G | Sense: 5′CTATAATGCTCTA**GG**ACAAGAAAATG3′ |
| CC1-tA | L419/A | Sense: 5′CTATAATGCT**GC**ACCACAAGAAAATG3′ |
| CC1-tGT | P420I430/GT | Sense: 5′CTATAATGCTCTA**GG**ACAAGAAAATGGCCTCTCACCTGGGGCCA**C**TGCT3′ |
| CC1-tAT | L419I430/AT | Sense: 5′CTATAATGCT**GC**ACCACAAGAAAATGGCCTCTCACCTGGGGCCA**C**TGCT3′ |

CEACAM1-4L constructs, the sites of their mutations and the oligos that have been used for mutagenesis. The bold, underlined bases are mismatched nucleotides used to introduce mutations in CC1 cDNA. The dash in the antisense primer used for making CC1-t represents the deleted base corresponding to CC1-4L cDNA. The asterisks indicates the unique restriction sites in CC1-4L cDNA that have been used to replace a 296 bp fragment of CC1-4L by the mutant PCR fragment to make the CC1-t construct: *Pflm1; **Bsrf1.

[†]Other constructs have a stop codon at the same position as CC1-t, in addition to the other specified amino acid substitutions.

[‡]This antisense primer was also used for all other inverse PCR based mutagenesis.

anchors arose twice independently during primate evolution, thus implying selective advantage, in spite of conferring tumorigenic properties (F.N., A. D. Yoder, M. Tobi, K. Neiswanger and C.P.S., unpublished).

This study addresses the question of whether a molecular conversion from TM to GPI anchorage in present-day human CEACAM1 is easily attained and, if so, whether the conversion leads to a GPI anchor that possesses the tumorigenic specificity of CEA.

## Materials and Methods

### Mutagenesis

Human CEACAM1-4L cDNA (Beauchemin et al., 1999), cloned in pBluescript SK plasmid, was subjected to PCR-based mutagenesis using a 62-mer mutant antisense and a 27-mer sense primer (Table 1) to introduce a deletion and a substitution in order to create a stop codon at the same position as in the GPI-linked CEA family members. The 296 bp fragment between the unique Pflm1 (nucleotide 1146) and Bsrf1 (nucleotide 1442) sites in wild-type CEACAM1-4L cDNA was replaced with the mutant PCR fragment, generating CC1-t, giving a truncated CC1 protein. Other mutants were made by inverse PCR-based mutagenesis, as described previously (Clackson et al., 1992) with minor modifications, using CC1-t-pBluescript SK as template to introduce the appropriate mutations in the TM domain of human CEACAM1-4L cDNA. The oligonucleotides used for mutagenesis (see Table 1) were used with Pfu DNA polymerase (Stratagene) instead of Taq polymerase for higher fidelity of polymerization. The coding sequences of all mutants were verified by dideoxy nucleotide sequencing (Pharmacia Biotech T7 sequencing™ kit).

### Cell culture and transfection

Wild-type and mutant cDNAs were inserted into the P91023B expression vector (courtesy of R. Kaufman, Genetics Institute, Boston, MA) for expression in LR-73, a CHO-derived cell line (Pollard and Stanners, 1979), or rat L6 myoblasts (Yaffe, 1968). The cells were grown as monolayer cultures in α-MEM (LR-73) or DMEM (L6) containing 10% fetal bovine serum (FBS) as growth medium (GM; Gibco-BRL, Gaithersburg, MD) supplemented with 100 U/ml penicillin and 100 μg/ml streptomycin (Gibco-BRL) and were incubated at 37°C in a humidified atmosphere with 5% $CO_2$. Cells were co-transfected by calcium phosphate co-precipitation (Benchimol et al., 1989) with 10 μg of P91023B containing cDNAs and 1 μg pSV2neo plasmid per dish. In order to avoid complications in interpretation due to clonal differences unrelated to the effects of the expressed plasmid, pooled colonies of stable transfectants sorted by FACS for relatively high levels of expression were used. To this end, Geneticin (G-418 sulfate, Gibco-BRL)-resistant colonies were pooled and enriched for stable transfectants by culturing without G-418 selection for over 1 month then sorted for high cell surface expression of CEA family external domains by FACS, using polyclonal rabbit antihuman CEA antibody.

### FACS analysis

Transfected cells were analyzed for cell surface expression of proteins by cytofluorometric analysis (FACScan®, Becton Dickinson, Bedford, MA) using polyclonal rabbit anti-CEA antibody as primary antibody, as previously described (Zhou et al., 1993a). This antibody cross reacts extensively with CEACAM1 and also other CEA family members (Zhou et al., 1993b).

### Phosphatidylinositol-specific phospholipase C (PI-PLC) treatment of intact cells

Adherent transfected cells were removed from culture dishes with Hank's balanced salt solution lacking $Ca^{2+}/Mg^{2+}$ but containing 0.5 mM EDTA (Hank's/EDTA) for 3 minutes at 37°C, and rendered as single cell suspensions. Duplicate aliquots of $5\times10^5$ cells were washed in cold PBS and resuspended in 150 μl of serum-free culture medium diluted 1:1 with PBS + 0.2% BSA; one aliquot was treated in suspension at 37°C, in 95% air + 5% $CO_2$, with 0.2 Units PI-PLC from *Bacillus cereus* (Boehringer Mannheim) and the other without PI-PLC treatment, as a control. After 1 hour of incubation, the cells were centrifuged and resuspended in 2.5 ml cold PBS and subjected to FACS analysis as described above.

### Differential extraction of GPI-linked and TM-linked proteins by Triton X-100

Monolayer cultures of cells were removed from plastic surfaces with Hanks/EDTA and $2.5\times10^6$ LR-73 transfectant cells or $1.25\times10^6$ L6 transfectant cells were resuspended in 250 μl of ice cold lysis buffer containing 20 mM Tris-HCl pH 6.5, 150 mM NaCl, 5 mM EDTA, and 1% Triton X-100 (TX-100), supplemented with a cocktail of protease inhibitors (1 mM phenylmethyl sulfonyl fluoride, 10 μg/ml leupeptin and 10 μg/ml aprotinin). After mixing, they were incubated for 10 minutes on ice and centrifuged at 15,000 *g* at 4°C for 20

## a

```
CEA       CATCTGGAACTTCTCCTGGTCTCTCAGCTGGGGCCACTGTCGGCATCATGATTGGAGTGCTGGTTGGGGTTGCTCTGATA-TAGCAGCCCTGGTGTAGTTTCTTCATTTCAGGAAGACTG
CEACAM6   ...G.G......TC..........T......C...........C..............CCA....G........-...........T....CAT...............
CEACAM8   T.CAA....G.........C.........A.A.........A.................A....CCA...G........-...T...T.............G......A........
CEACAM7   T.CAA.C..G...A....AC...........A..G....A.................A....C....A.G........-...........T.....
CEACAM1   .TCTACC.CAAGAAAA...C......C.........T..CT.....TG.........AG...CCCT.............GC..T.......CA.GT.....G......G.......C.GCAG
CC1-tAT   .TG.ACC.CAAGAAAA...C......C...........CT.....TG.........AG...CCCT.............-...T........CA.GT.....G......G.......C.GCAG
```

## b

```
CEA           TVSASGTSPGLSAGATVGIMIGVLVGVALI*
CEACAM6       MITV..SA.V...V.....T....AR....*
CEACAM8       DALVQ.S......R...S......AR....*
CEACAM7       YE.VQAS..D....TA.S......A.M...*
CEACAM1       NYN.LPQEN...P..IA..V...VAL....AVAL

CC1-tAT       NYN.APQEN...P..TA..V...VAL....*
              415
```

**Fig. 1.** Sequence comparison of the carboxy-terminal domain of CEA gene family members: (a) Nucleotides. Underlined bases are stop codons, and dashes show a naturally occurring base deletion in comparison with the CEACAM1 gene. (b) Amino acids. Underlined amino acids are the known GPI anchor attachment sites and stars show the position of stop codons. Bold letters indicate the position of amino acid mutations in various constructed, mutant CC1-4L proteins. The TM domain of mutant GPI-linked CC1-tAT protein and the positions of its mutations (bold letters) are also shown in b for comparison.

minutes; 250 μl 2× SDS PAGE sample buffer (Laemmli, 1970) was added to the supernatant and 500 μl 1× SDS sample buffer to the pellet. Lysates were boiled for 10 minutes and 20 μl of each supernatant and sediment was analyzed by SDS-PAGE for the levels of CEACAM1 proteins by immunoblotting, as described below.

### Endoglycosidase H (Endo H) treatment

For glycosidase treatment, LR(CC1-t) transfectant cells were lysed by TX-100 containing lysis buffer, as described above, and both supernatant and sediment were dissolved in 1× Endo H denaturing buffer provided by the enzyme supplier (New England Biolabs, Beverly, MA), then denatured by boiling at 100°C for 10 minutes. Twenty μg of total protein was treated with 1250 Units Endo H at 37°C for 1 hour then 15 μg total protein of each supernatant and sediment were resolved by SDS-PAGE and immunoblotting, as described below.

### SDS-PAGE and immunoblotting

Aliquots of supernatant and sediment of TX-100 extracts or Endo H-treated samples, prepared as described above, were analyzed by electrophoresis on SDS-polyacrylamide (7.5%) gels and transferred electrophoretically to 0.45 μm nitrocellulose membranes (Protran™, XYMOTECH Biosystem INC.). Membranes were incubated at 4°C overnight with rabbit polyclonal antihuman CEA antibody, diluted 1:3000. After washing for 30 minutes with buffer containing 0.1% Tween 20, membranes were incubated in 1:4000 diluted HRP-conjugated anti-rabbit antibody, and washed and visualized by ECL reagent according to the manufacturer's instructions (Amersham Life Science, Pittsburgh, PA).

### Adhesion assay

Adhesion assays were performed on LR-73 transfectants by visual assessment of the percentage of cells remaining as single cells as a function of time in suspension using a hemocytometer, as previously described (Benchimol et al., 1989).

### Myogenic differentiation assay

Cultures of transfectants were seeded at $10^4$ cells/cm$^2$ in 60 mm dishes in GM and incubated at 37°C for 3 days, after which the medium was changed to DMEM + 2% horse serum (differentiation serum) and the incubation continued for a further 7 days. The cultures were stained with Hematoxylin and the extent of differentiation assessed by the percentage of nuclei in single myotube cells with three or more nuclei relative to the total number of nuclei. Three independent experiments testing for the ability of various experimental and control transfectants to differentiate were carried out for each of two independent sets of transfectants, with equivalent results. For anti-myosin immunofluorescent staining, cultures were induced to differentiate in multiwell chamber slides (Nunclon®; Nunc, Inc., Naperville, IL). Immunofluorescent labeling of permeabilized cells was performed as described previously (De Giovanni et al., 1993) using anti-myosin heavy chain monoclonal 47A antibody (courtesy of P. A. Merrifield, University of Western Ontario. London, ON).

## Results

### Sequence analysis

Previous structure-function studies on GPI-linked proteins have indicated that the signal required for GPI anchor addition (glypiation) resides completely in the C-terminal domain, which is cleaved in the process (Englund, 1993). These signals comprise a stop codon at the C-terminal end and, proceeding upstream, a short stretch of hydrophobic residues (8-21 amino acids), a hydrophilic/neutral spacer region (usually 5-12 residues) and, finally, the cleavage/attachment site of the protein to the GPI anchor, denoted the ω-site (Chen et al., 2001; Coyne et al., 1993). The ω and ω+2 residues are both critical for anchor addition and are usually residues with small side chains; ω−1 residues have also been observed to affect the efficiency of glypiation (Coussen et al., 2001).

Comparison of the nucleotide and amino acid sequences of the C-terminal hydrophobic exons of GPI with TM-anchored CEA family members (Fig. 1) indicates the presence of a stop codon at the same position in all GPI-linked members, whereas TM-linked CEACAM1-4L [the longest splice variant of human CEACAM1 (Beauchemin et al., 1999)] has no stop codon at this position and an extended downstream open reading frame of 82 amino acids (TM plus the cytoplasmic domain). As expected, in CEA there is a strongly hydrophobic region of about 14 amino acids immediately upstream of the stop codon followed by a 12 amino acid hydrophilic region up to the known ω site, an A residue (Fig. 1b).
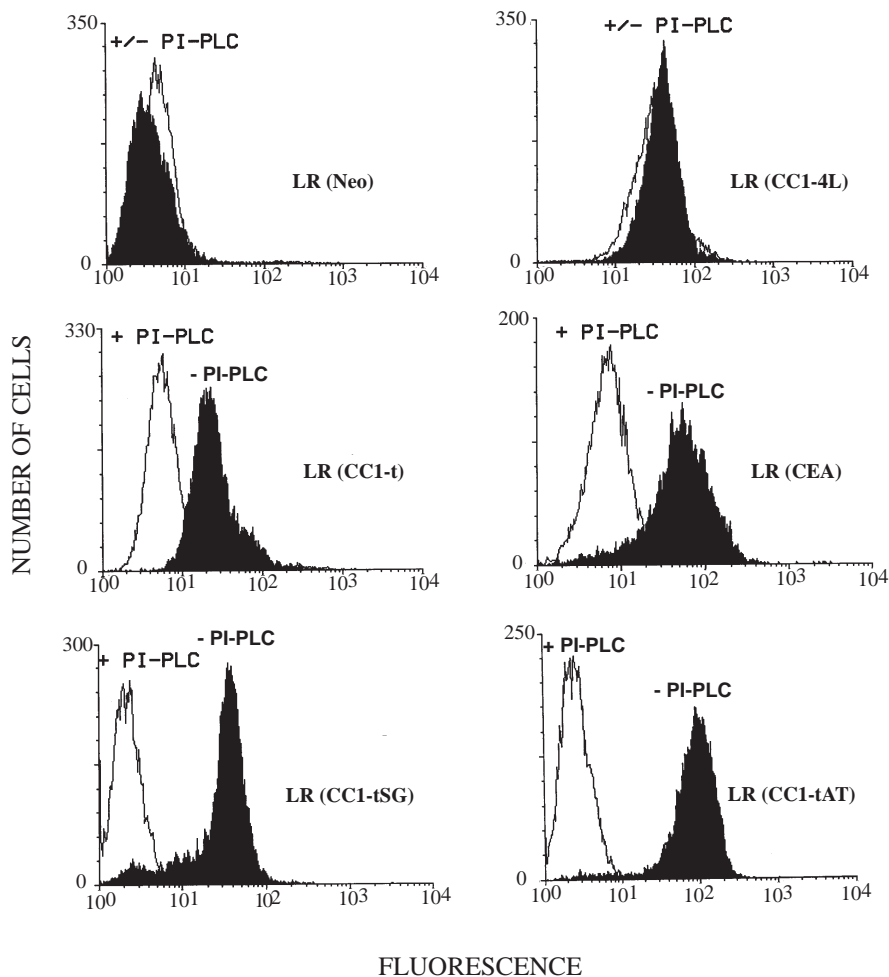
Fig. 2. FACS analysis before and after PI-PLC treatment. FACS profiles giving CEA and CC1-mutant cell surface expression on LR-73 transfectant cells with and without treatment with PI-PLC to remove GPI-anchored molecules. Profiles for CEA (positive GPI-linked control), CC1-4L (negative TM-linked control) and Neo (background control) transfectants are also shown.

Immunoblotting of the supernatant fractions from PI-PLC-treated LR(CC1-t) cells confirmed that only the higher $M_r$, TX-100 insoluble species was released by PI-PLC and therefore GPI-linked and at the cell surface (Fig. 4a). We conclude that the lower $M_r$, TX-100-soluble band was not GPI processed and was sequestered inside the cell. Densitometric analysis of an ECL-developed western blot showed that about 10% of the total CC1-t protein in the cell is GPI processed. In parallel experiments with LR-73 transfectants producing naturally occurring GPI-linked CEA family members, CEA (Fig. 3a), CEACAM6 (Fig. 3b), or CEACAM8 (data not shown), the lower $M_r$ TX-100-soluble unprocessed band could not be detected (although some higher $M_r$, fully processed soluble protein was seen). These results show that the introduction of a CEA-like stop codon in the CC1-4L TM domain is sufficient to confer GPI cell membrane linkage, but with low efficiency.

## Sequence changes required to give efficient GPI anchorage

### Mutations giving stop codon

The nucleotide sequence differences between CEA and CEACAM1 required to produce a stop codon (a one bp deletion and a one bp substitution) were introduced at the corresponding position in CEACAM1-4L cDNA. The mutated cDNA, denoted 'CC1-t' for 'truncated', was expressed at a relatively low level on the cell surface of stable transfectants of LR-73 cells (Fig. 2). The GPI-anchorage of the cell surface CC1-t was demonstrated by sensitivity to PI-PLC: FACS profiles of PI-PLC-treated cell populations showed a reduction of cell surface levels of both CC1-t and CEA but no reduction of TM-anchored CEACAM1-4L (Fig. 2).

The GPI linkage of the CC1-t protein was also verified by a cold Triton X-100 solubilization assay, in which GPI-anchored proteins are insoluble. Immunoblot analysis of cold TX-100 extracts of CEA, CEACAM1-4L and CC1-t transfectants indicated the presence of mainly insoluble (pellet fraction) CEA, both insoluble and soluble (supernatant fraction) CC1-t protein and only soluble CEACAM1-4L protein (Fig. 3a). The soluble CC1-t protein, however, was represented as a major lower $M_r$ band. The reduced $M_r$ was presumably due to incomplete glycosylation during intracellular processing, since it was Endo-H sensitive (Fig. 4b), a test for proteins that fail to pass through the Golgi (Dunphy et al., 1985).

### Mutations giving increased GPI processing efficiency

The precise TM cleavage site ($\omega$-site) for the CC1-t protein is unknown but, on the basis of previous studies on GPI-linked proteins (Englund, 1993; Micanovic et al., 1990), the two best possible candidates are either N417 or A418. The $\omega$+2 residues would then be L419 or P420, respectively. However, both of the latter residues have been found to be weak as GPI processing signals; $\omega$+2 residues are normally A, G or S (Gerber et al., 1992; Kodukula et al., 1993). We therefore substituted the two candidate $\omega$+2 residues in CC1-t protein for the residues in CEA at these positions (S419 and G420). PI-PLC treatment of stable transfectants of this mutant, denoted LR(CC1-tSG), showed that the protein reaching the cell surface was GPI-linked (Fig. 2). The efficiency of GPI-processing of the mutant CC1-tSG protein by the Triton X-100 solubility assay gave 77% insolubility at the higher $M_r$ (Table 2), suggesting a much improved processing efficiency.

To determine which of the two above substitutions in CC1-t protein was effective for higher GPI-processing, CC1-tS and CC1-tG with single L419S, and P420G substitutions, respectively, were constructed. GPI linkage of the two mutant proteins was shown by PI-PLC treatment of stable transfectants of LR-73 cells, as before (data not shown). The cold TX-100 solubility assay indicated 80% insoluble

**Table 2. GPI-processing efficiency of CC1-4L mutants**

| Construct | Cell line | |
|---|---|---|
| | LR-73 | L6 |
| CC1-t | 8% | 10% |
| CC1-tSG | 77% | ND |
| CC1-tS | 80% | 77% |
| CC1-tG | 5% | ND |
| CC1-tA | 80% | 78% |
| CC1-tGT | 56% | ND |
| CC1-tAT | 100% | ND |

GPI-processing efficiency of CC1-4L mutants transfected into the Chinese hamster cell line LR-73 and rat L6 myoblasts, measured by scanning ECL-developed immunoblots.

ND, not determined.

CC1-tS but only 5% insoluble CC1-tG (Fig. 3b and Table 2). Since mutating L419 to S but not P420 to G could markedly improve GPI processing, we conclude that N417 and not A418 is most likely the ω site for GPI addition in CC1-t.

In spite of the quite dramatic improvement in processing activity in CC1-tS over CC1-t, the efficiency does not match the 100% efficiency of naturally occurring GPI-linked members of the CEA family. Changing the Ser residue in the ω+2 position of CC1-tS to Ala, a more favorable residue in other GPI-linked proteins (Kodukula et al., 1993; Moran and Caras, 1994), to produce CC1-tA, however, failed to further improve the efficiency (Fig. 3d and Table 2).

There is a conserved Thr in GPI-anchored CEA, CEACAM6, CEACAM8 and in virtually all primate GPI-producing C-terminal hydrophobic exons (F.N., A. D. Yoder, M. Tobi, K. Neiswanger and C.P.S., unpublished) at a position corresponding to Ile430 in both CEACAM1-4L (Fig. 1) and in all other TM-anchored CEA family members that, in CC1-t, tends to shorten the hydrophilic region required for efficient processing (Furukawa et al., 1997). To investigate the role of this residue, an inefficiently processed CEACAM1 mutant with the same residues at the ω and ω+2 sites as CEA, i.e., CC1-tG, was used. CC1-tGT, with Ile430 replaced by Thr, was therefore constructed from CC1-tG. The immunoblot of TX-100 extracts of LR(CC1-tGT) showed a marked improvement of GPI-processing relative to the CC1-tG protein (56% of CC1-tGT protein versus 5% of CC1-tG) (Fig. 3c and Table 2). These results indicate that the evolutionarily conserved Thr430 residue in fact has an important role in GPI-processing of CEA family members.

Finally, mutant CC1-tAT, with both advantageous mutations, L419A and I430T, was constructed and tested for processing efficiency. This mutant protein was well expressed on the surface of LR-73 cells and sensitive to PI-PLC (Fig. 2). The GPI-processing efficiency for CC1-tAT was essentially 100%, since there was no detectable lower $M_r$ cold TX-100-soluble band (Fig. 3d and Table 2).

The above results were confirmed in a second cell line, rat L6 myoblasts, in which CEA family proteins are produced from expressed transfected constructs at consistently lower $M_r$ (data not shown) because of lower levels of glycosylation. Stable transfectants of L6 cells expressing the mutant constructs were tested for GPI anchorage and processing activity as above. The results showed PI-PLC sensitivity of cell surface protein (data not shown) and efficiency of GPI anchor addition very similar to that seen in the CHO-derived LR-73 transfectants (Table 2).

We conclude that CEACAM1-4L can be converted into an efficiently processed GPI-linked protein by relatively few mutations in its TM domain: the introduction of a stop codon, a mutation at the ω+2 site and a mutation extending the length of the hydrophilic spacer region suffice.
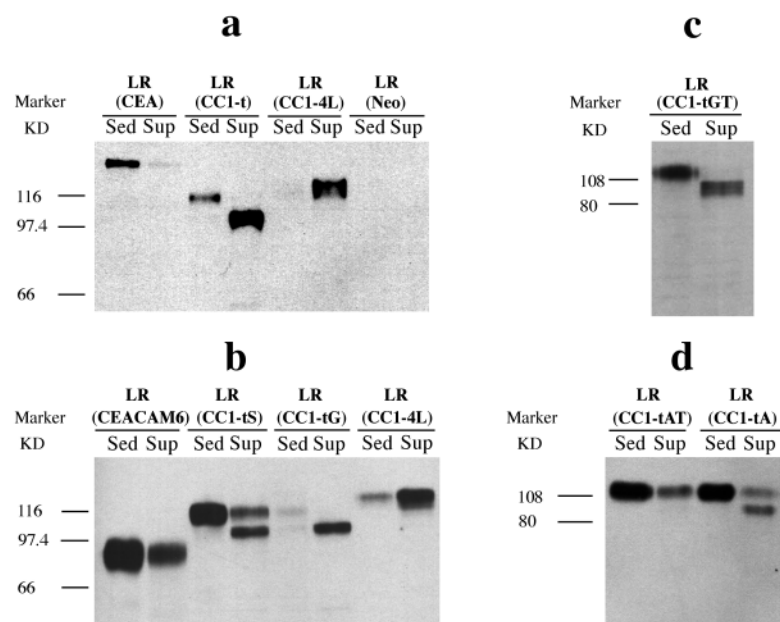


**Fig. 3.** GPI-processing evaluation of various CC1-4L mutants by cold nonionic detergent solubility assay. Immunoblot analysis of cold Triton X-100 extracts of LR-73 transfectant cells. Neo transfectants and naturally occurring CEA family members, CEA, CEACAM6 (GPI-linked), and CC1-4L (TM-linked) were used as controls. a-d represent results for four separate immunoblots.

## Minimally mutated CEACAM1 has radically changed function

### Intercellular adhesion function

In order to test whether the mode of membrane linkage of CEACAM1-4L changes its function as an intercellular adhesion molecule (Rojas et al., 1990), LR(CEACAM1-4L) and LR(CC1-tAT) transfectant cells were subjected to an aggregation assay in suspension (Benchimol et al., 1989). The results, shown in Fig. 5, indicate that CC1-tAT functions well as a homophilic intercellular adhesion molecule.

### Differentiation blocking function

GPI-anchored CEA and CEACAM6 inhibit the differentiation of many different cell types, whereas TM-anchored CEACAM1 does not (see Introduction). The molecular requirements for this tumorigenic function are self-binding external domains attached to CEA-type GPI anchors (Screaton et al., 2000). To test whether the minimal mutations that conferred efficient GPI anchorage on CEACAM1 also radically changed CEACAM1 from a differentiation neutral or stimulatory molecule
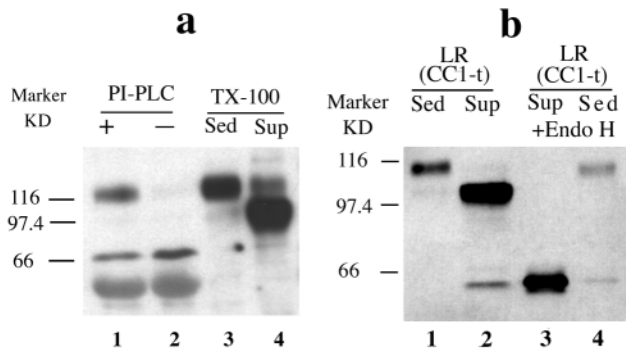
**Fig. 4.** Nature of proteins in inefficiently GPI-processed CC1 mutant, CC1-t. (a) Immunoblot of the supernatant of LR(CC1-t) transfectants with and without PI-PLC treatment. PI-PLC treatment released only one band (lane 1), which corresponds to the TX-100-insoluble, higher $M_r$ band (lane 3). Two other low $M_r$ bands, which could be seen in both PI-PLC-treated and untreated cells (lane 1 and 2) but not in TX-100 extracts (lane 3 and 4), are presumably the result of other cross-reacting components in the reaction medium (see Materials and Methods). (b) Immunoblot of both TX-100-soluble and -insoluble fractions with (lanes 3 and 4) and without (lanes 1 and 2) treatment with Endo H. Only the major lower $M_r$, TX-100-soluble band (lane 3) was sensitive to Endo H digestion.

(Rojas et al., 1996) to a molecule that blocked myogenic differentiation, CC1-tAT was expressed in rat L6 myoblasts and the transfectants tested for their ability to fuse into myotubes. Unlike L6(CC1-4L) transfectants (which were about 90% fused) and like L6(CEA) transfectants, L6(CC1-tAT) transfectants were completely unable to fuse and differentiate into myotubes (Fig. 6). The complete block in myogenic differentiation by CC1-tAT at the level of cell fusion was confirmed at the biochemical level by the complete absence of staining with anti-myosin antibody (Fig. 7).

## Discussion

In this work we have investigated the molecular genetic requirements for GPI-anchorage of CEA family members in the context of evolution. Since only TM-anchored CEACAM1-like genes have been discovered in rodents, we and others have suggested a recent evolution of GPI-anchored CEA family members from a primordial TM-anchored gene (Hammarström et al., 1998; Stanners et al., 1995; Stanners et al., 1992; Zimmermann, 1998). In view of the rather dramatic changes in function implicated by the acquisition of the GPI anchor (see Introduction), we decided to examine the ease of derivation of GPI linkage from the TM domain of present-day human CEACAM1. We show here that very few mutations are required to convert CEACAM1 into an efficiently processed GPI-linked molecule: the introduction of a stop codon and two upstream substitutions in the TM domain of CEACAM1 suffice.

Two properties of mutant CEACAM1 proteins were used to develop a quantitative assay for GPI processing efficiency. Owing to the linkage of GPI-linked proteins to saturated acyl chains, these proteins associate with sphingolipids and cholesterol in membrane rafts and become insoluble in detergent TX-100-containing solutions at low temperatures (Brown and London, 1997). Also, CEA family members are
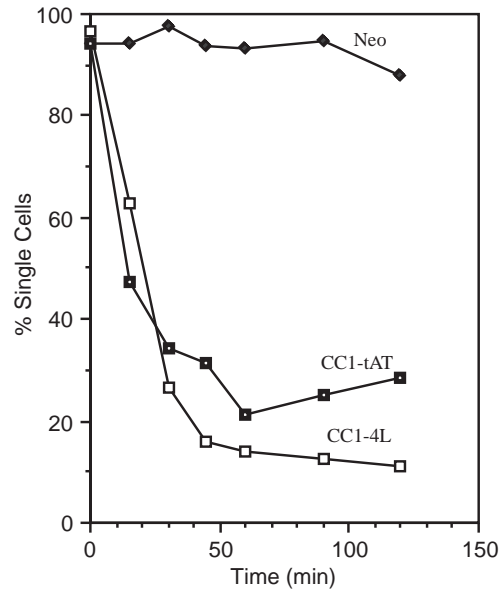


**Fig. 5.** Homotypic intercellular adhesion mediated by mutant GPI-anchored CEACAM1 protein. CC1-tAT transfectants of LR-73 cells were subjected to the homotypic adhesion assay in suspension. The percentage of single cells was measured as a function of time in suspension. Neo (vector alone) and CC1-4L transfectant cells were used as negative and positive controls, respectively. The mean expression levels of CC1-4L and CC1-tAT by FACS analysis were 280 and 98 fluorescence units, respectively.

highly glycosylated, constituting up to 50% of their molecular mass; failure to cleave the hydrophobic C terminus and attach a GPI anchor results in retention in the ER and incomplete glycosylation, giving a faster gel mobility, detectable by immunoblotting.

The GPI-processing efficiency was therefore assessed by the presence of cold detergent-soluble bands of lower than normal molecular mass. The justification for this assumption derives from both precedent and experiment. Thus, uncleaved GPI anchoring signals have been reported to function as intracellular retention signals for GPI-unprocessed proteins (Moran and Caras, 1992), resulting in retention within the ER and rapid degradation (Oda et al., 1996). In the case of CEACAM1 mutant constructs that were inefficiently GPI-processed, the high $M_r$ TX-100-insoluble band was shown to be the mature, properly localized GPI-linked protein, as expected. In the TX-100-soluble fraction, two bands were obtained, one of normal size, the other of lower $M_r$. A normal size soluble band was seen for all GPI-anchored CEA family members and may reflect a proportion of molecules that are membrane localized but not situated in membrane rafts. The lower $M_r$ soluble band was shown to be selectively sensitive to endoglycosidase H (Endo H), an enzyme that cleaves the high-mannose oligosaccharides added in the ER that have not been processed by β-$N$-acetylglucosamine transferase I and α-mannosidase II in the medial Golgi cisternae (Dunphy et al., 1985). Since no unprocessed CEACAM1 mutant protein was found in the extracellular medium, indicative of an extracellular pool of improperly processed protein, the level of the high $M_r$ TX-100-insoluble band relative to the total level, as seen on the immunoblots, was used as a measure of the GPI-processing
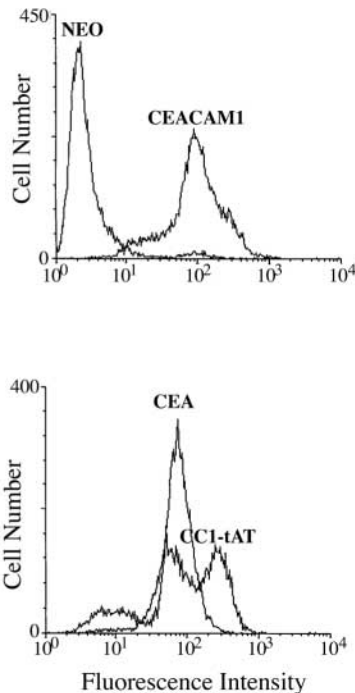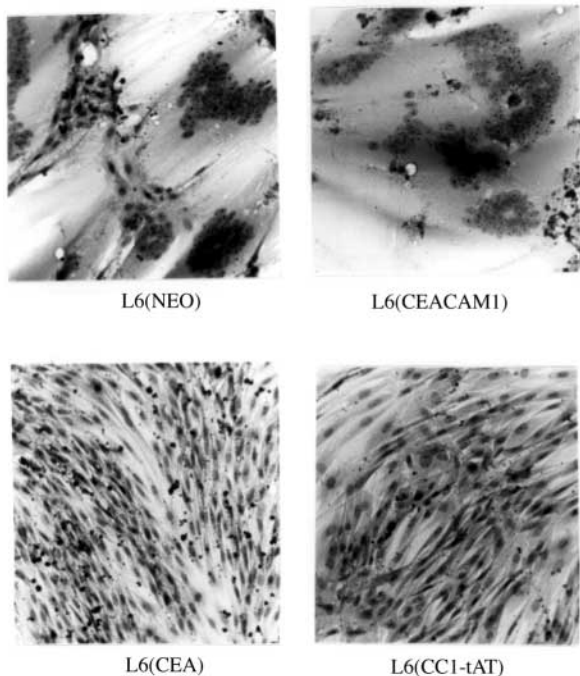
**Fig. 6.** Effect of mutant GPI-anchored CEACAM1 protein on morphological myogenic differentiation. Photomicrographs of Hematoxylin-stained cultures of various L6 transfectants incubated in differentiation medium for 7 days. Magnification, 400×. FACS profiles show the relative cell surface expression level of CEA proteins in the cell cultures tested.

efficiency. The rate of degradation of the intracellular lower $M_r$ incompletely glycosylated soluble material was not assessed, however, so that our measures of the percentage of GPI-processed protein based on steady state levels are likely to represent overestimates of the true processing efficiencies.

The introduction of a stop codon in CEACAM1 at the CEA stop codon position alone could confer GPI-linkage, although inefficiently (~10% processing efficiency). This indicates that a latent GPI-anchoring signal must exist in the CEACAM1-4L TM domain that becomes functional upon removal of the cytoplasmic tail, as has been seen in other systems (Bell et al., 1994). Two other substitutions, one, L419A, 'improving' the ω+2 site, the other, I430T, extending the length of the hydrophilic spacer immediately downstream of the ω+2 site, were all that was required to increase the processing efficiency of this mutant construct to 100%, i.e., that of all naturally occurring GPI-anchored CEA family members. The dramatic improvement in GPI processing efficiency observed for the I430T mutation, a single amino acid substitution in the hydrophilic spacer region, could not have been predicted from current knowledge. It is of interest, therefore, that this very substitution was observed as part of *both* packages of mutations giving GPI anchorage that arose independently during the primate radiation (F.N., A. D. Yoder, M. Tobi, K. Neiswanger and C.P.S., unpublished).

The change of mode of cell anchorage from TM to GPI in the minimally mutated CEACAM1, CC1-
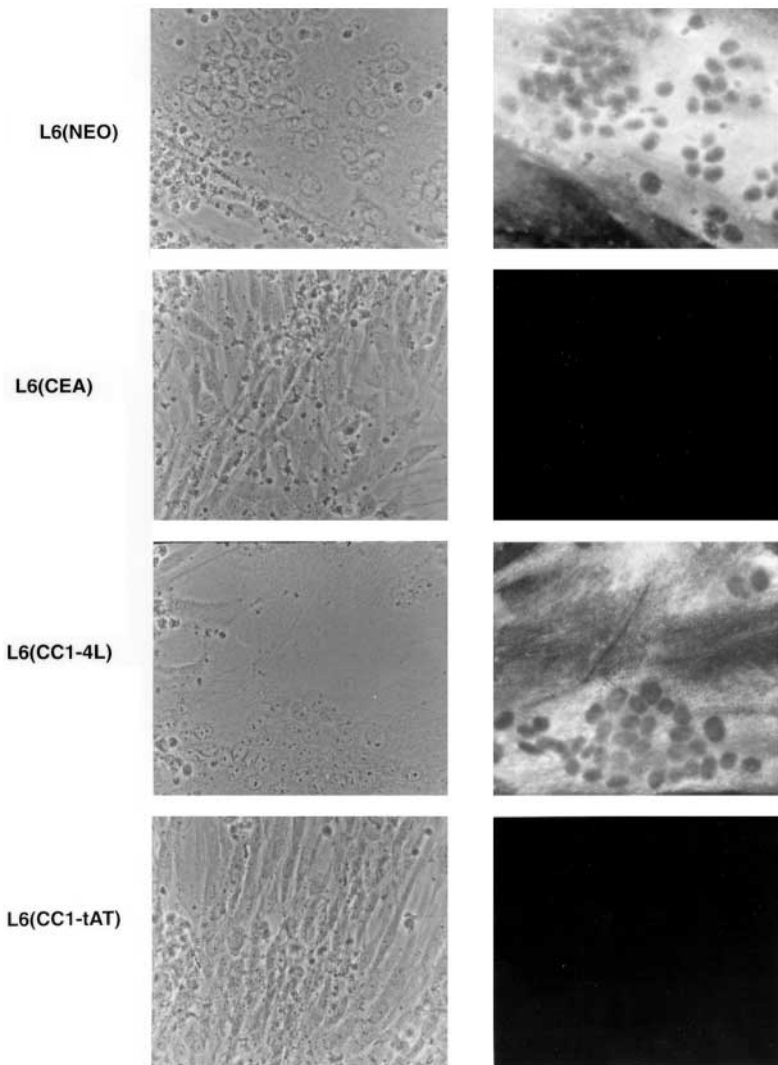


**Fig. 7.** Effect of mutant GPI-anchored CEACAM1 protein on biochemical myogenic differentiation. Left column: phase contrast light photomicrographs; right column: corresponding anti-myosin antibody plus FITC-conjugated, anti-mouse IgG-stained fluorescent photomicrographs of cultures of various transfectant cells incubated for 4 days in differentiation medium. Cytoplasmic staining of multinucleated myotubes can be seen for L6(Neo) and L6(CC1-4L), indicating myosin synthesis and myogenic differentiation. Magnification, 600×.

tAT, did not abolish the intercellular adhesive function of CEACAM1, an expected result considering the fact that both GPI-linked and TM-linked CEA family members have been found to mediate homotypic and heterotypic intercellular adhesion (Rojas et al., 1996). It did, however, radically change a sentinel function: CC1-tAT, unlike CEACAM1, completely blocked the myogenic differentiation of rat L6 myoblasts, a strong indicator of the acquisition of all of the tumorigenic properties of CEA and CEACAM6, including inhibition of anoikis and disruption of tissue architecture (see Introduction). We presume that the sequence similarities between the carboxy-terminal exons of CC1-tAT and CEA are sufficient to ensure similar biological specificity of the attached GPI anchor, which is the structure that determines the ability to block differentiation. We have speculated that the latter specificity is due to the particular nature of the sugar structures in the GPI anchor, which is directed by the sequence of the processed carboxy terminus of the pro-protein (Screaton et al., 2000).

The foregoing has led us to suggest the following speculative model for the evolution of GPI anchors in the CEA family. It has been proposed that selective constraints are relaxed after gene duplication (Lewin, 1994). A duplicated gene is relatively free to evolve to gain a new function. In this context it is plausible to think that, after duplication of a primordial TM-anchored CEA gene (or of its carboxy-terminal exon), one copy retained the original function while the other copy accumulated enough mutations in its carboxy-terminal exon to be converted to GPI anchorage. There are three imperfect AGC repeats in the stop codon region of the carboxy-terminal exon of CEA family members that, by slippage of DNA replication enzymes (Freimer and Slatkin, 1996), might represent a mutational hotspot; in fact, two different stop codons and the addition (or deletion) of one AGC repeat have been seen in various primate species in this region (F.N., A. D. Yoder, M. Tobi, K. Neiswanger and C.P.S., unpublished). We envisage the generation of such a stop codon in a primitive exon, followed by the appearance of low levels of cell surface GPI-anchored CEA family protein because of its low processing efficiency. Since GPI-linked proteins are found in very primitive organisms, such as protozoa and yeast, where in some cases most of the cell surface proteins are GPI-anchored (McConville and Ferguson, 1993), it seems assured that the machinery for producing the GPI anchors themselves was already in place at the time of this event. The inefficiently processed GPI-anchored protein isoform would then allow 'testing' for an advantageous change in function; if the latter accrued, the advantage would favor the acquisition of further mutations giving efficient processing and higher cell surface levels of expression. The efficient GPI anchor exon could then be shuffled with the numerous CEA family exons determining external domains, resulting in further diversification of function upon which selection could act. The changes conferred by GPI anchorage, that of inhibition of differentiation and the adoption of relaxed communal architecture which could delay morphogenesis during embryonic development, would require accurate regulation in order to be advantageous. Fine tuning of transcriptional control elements to restrict expression of CEA family members temporally and spatially appears to have achieved this control [(Eades-Perner et al., 1994) C. Chan and C.P.S., unpublished].

## References

**Beauchemin, N., Draber, P., Dveksler, G., Gold, P., Gray-Owen, S., Grunert, F., Hammarstrom, S., Holmes, K. V., Karlsson, A., Kuroki, M. et al.** (1999). Redefined nomenclature for members of the carcinoembryonic antigen family. *Exp. Cell Res.* **252**, 243-249.

**Bell, L. M., Solomon, K. R., Gold, J. P. and Tan, K. N.** (1994). Cytoplasmic tail deletion of T cell receptor (TCR) beta-chain results in its surface expression as glycosylphosphatidylinositol-anchored polypeptide on mature T cells in the absence of TCR-alpha. *J. Biol. Chem.* **269**, 22758-22763.

**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.

**Chen, R., Knez, J. J., Merrick, W. C. and Medof, M. E.** (2001). Comparative efficiencies of C-terminal signals of native glycophosphatidylinositol (GPI)-anchored proproteins in conferring GPI-anchoring. *J. Cell. Biochem.* **84**, 68-83.

**Chevinsky, A. H.** (1991). CEA in tumors of other than colorectal origin. *Semin. Surg. Oncol.* **7**, 162-166.

**Clackson, T., Güssow, D. and Jones, P. T.** (1992). General applications of PCR to gene cloning and manipulation. In *PCR a Practical Approach* (ed. M. J. Mcpherson, P. Quirke and G. R. Taylor), pp. 187-214. New York: IRL Press at Oxford University Press.

**Coussen, F., Ayon, A., Le Goff, A., Leroy, J., Massoulie, J. and Bon, S.** (2001). Addition of a glycophosphatidylinositol to acetylcholinesterase. Processing, degradation, and secretion. *J. Biol. Chem.* **276**, 27881-27892.

**Coyne, K. E., Crisci, A. and Lublin, D. M.** (1993). Construction of synthetic signals for glycosyl-phosphatidylinositol anchor attachment. Analysis of amino acid sequence requirements for anchoring. *J. Biol. Chem.* **268**, 6689-6693.

**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.

**Dickson, G., Peck, D., Moore, S. E., Barton, C. H. and Walsh, F. S.** (1990). Enhanced myogenesis in NCAM-transfected mouse myoblasts. *Nature* **344**, 348-351.

**Dunphy, W. G., Brands, R. and Rothman, J. E.** (1985). Attachment of terminal N-acetylglucosamine to asparagine-linked oligosaccharides occurs in central cisternae of the Golgi stack. *Cell* **40**, 463-472.

**Eades-Perner, A. M., van der Putten, H., Hirth, A., Thompson, J., Neumaier, M., von Kleist, S. and Zimmermann, W.** (1994). Mice transgenic for the human carcinoembryonic antigen gene maintain its spatiotemporal expression pattern. *Cancer Res.* **54**, 4169-4176.

**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.

**Furukawa, Y., Tsukamoto, K. and Ikezawa, H.** (1997). Mutational analysis of the C-terminal signal peptide of bovine liver 5′-nucleotidase for GPI anchoring: a study on the significance of the hydrophilic spacer region. *Biochim. Biophys. Acta* **1328**, 185-196.

**Gerber, L. D., Kodukula, K. and Udenfriend, S.** (1992). Phosphatidylinositol glycan (PI-G) anchored membrane proteins. Amino acid requirements adjacent to the site of cleavage and PI-G attachment in the COOH-terminal signal peptide. *J. Biol. Chem.* **267**, 12168-12173.

**Hammarström, S., Olsen, A., Teglund, S. and Baranov, V.** (1998). The nature and expression of the human CEA family: In *Cell Adhesion and Communication Mediated by the CEA Family: Basic and Clinical Perspectives.* vol. 5 (ed. C. P. Stanners), pp. 1-30. Amsterdam: Hardwood Academic Publishers.

**Hsieh, J. T., Luo, W., Song, W., Wang, Y., Kleinerman, D. I., Van, N. T. and Lin, S. H.** (1995). Tumor suppressive role of an androgen-regulated epithelial cell adhesion molecule (C-CAM) in prostate carcinoma cell revealed by sense and antisense approaches. *Cancer Res.* **55**, 190-197.

**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.

**Ilantzis, I., De Marte, L., Screaton, 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.

**Jothy, S., Yuan, S. Y. and Shirota, K.** (1993). Transcription of carcinoembryonic antigen in normal colon and colon carcinoma. In situ hybridization study and implication for a new in vivo functional model. *Am. J. Pathol.* **143**, 250-257.

**Kodukula, K., Gerber, L. D., Amthauer, R., Brink, L. and Udenfriend, S.** (1993). Biosynthesis of glycosylphosphatidylinositol (GPI)-anchored membrane proteins in intact cells: specific amino acid requirements adjacent to the site of cleavage and GPI attachment. *J. Cell Biol.* **120**, 657-664.

**Kunath, T., Ordoñez-Garcia, C., Turbide, C. and Beauchemin, N.** (1995). Inhibition of colonic tumor cell growth by biliary glycoprotein. *Oncogene* **11**, 2375-2382.

**Laemmli, U. K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

**Lewin, B.** (1994). Gene numbers: replication and redundancy. In *Genes V*, Vol. 1 (ed. B. Lewin), pp. 701-731. New York: Oxford University Press.

**McConville, M. J. and Ferguson, M. A.** (1993). The structure, biosynthesis and function of glycosylated phosphatidylinositols in the parasitic protozoa and higher eukaryotes. *Biochem. J.* **294**, 305-324.

**Micanovic, R., Kodukula, K., Gerber, L. D. and Udenfriend, S.** (1990). Selectivity at the cleavage/attachment site of phosphatidylinositol-glycan anchored membrane proteins is enzymatically determined. *Proc. Natl. Acad. Sci. USA* **87**, 7939-7943.

**Moran, P. and Caras, I. W.** (1992). Proteins containing an uncleaved signal for glycophosphatidylinositol membrane anchor attachment are retained in a post-ER compartment. *J. Cell Biol.* **119**, 763-772.

**Moran, P. and Caras, I. W.** (1994). Requirements for glycosylphosphatidylinositol attachment are similar but not identical in mammalian cells and parasitic protozoa. *J. Cell Biol.* **125**, 333-343.

**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.

**Nollau, P., Scheller, H., Kona-Horstmann, M., Rohde, S., Hagenmuller, F., Wagener, C. and Neumaier, M.** (1997). Expression of CD66a (human C-CAM) and other members of the carcinoembryonic antigen gene family of adhesion molecules in human colorectal adenomas. *Cancer Res.* **57**, 2354-2357.

**Oda, K., Wada, I., Takami, N., Fujiwara, T., Misumi, Y. and Ikehara, Y.** (1996). Bip/GRP78 but not calnexin associates with a precursor of glycosylphosphatidylinositol-anchored protein. *Biochem. J.* **316**, 623-630.

**Ordoñez, C., Screaton, R. A., Ilantzis, C. and Stanners, C. P.** (2000). Human carcinoembryonic antigen functions as a general inhibitor of anoikis. *Cancer Res.* **60**, 3419-3424.

**Paxton, R. J., Mooser, G., Pande, H., Lee, T. D. and 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.

**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.

**Rojas, M., Fuks, A. and Stanners, C. P.** (1990). Biliary glycoprotein, a member of the immunoglobulin supergene family, functions in vitro as a Ca2(+)-dependent intercellular adhesion molecule. *Cell Growth Differ.* **1**, 527-533.

**Rojas, M., DeMarte, L., Screaton, 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.

**Screaton, 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.

**Screaton, R. A., DeMarte, L., Draber, P. and Stanners, C. P.** (2000). The specificity for the differentiation blocking activity of carcinoembryonic antigen resides in its glycophosphatidyl-inositol anchor. *J. Cell Biol.* **150**, 613-626.

**Stanners, C. P. and Fuks, A.** (1998). Properties of adhesion mediated by the human CEA family. In *Cell Adhesion and Communication Mediated by the CEA family: Basic and Clinical Perspectives*, vol. 5 (ed. C. P. Stanners), pp. 57-71. Amsterdam: Harwood Academic Publishers.

**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.

**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.

**Yaffe, D.** (1968). Retention of differentiation potentialities during prolonged cultivation of myogenic cells. *Proc. Natl. Acad. Sci. USA* **61**, 477-483.

**Zhou, H., Fuks, A., Alcaraz, G., Bolling, T. J. and Stanners, C. P.** (1993a). Homophilic adhesion between Ig superfamily carcinoembryonic antigen molecules involves double reciprocal bonds. *J. Cell Biol.* **122**, 951-960.

**Zhou, H., Stanners, C. P. and Fuks, A.** (1993b). Specificity of anti-carcinoembryonic antigen monoclonal antibodies and their effects on CEA-mediated adhesion. *Cancer Res.* **53**, 3817-3822.

**Zimmermann, W.** (1998). The nature and expression of the rodent CEA families: evolutionary considerations. In *Cell Adhesion and Communication Mediated by the CEA Family: Basic and Clinical Perspectives,* vol. 5 (ed. C. P. Stanners), pp. 31-55. Amsterdam: Hardwood Academic Publishers.