Expression of the CUB domain containing protein 1 (CDCP1) gene in colorectal tumour cells

Sara E. Perry\textsuperscript{a,1}, Philip Robinson\textsuperscript{a,1}, Alan Melcher\textsuperscript{b}, Philip Quirke\textsuperscript{b}, Hans-Jörg Bühring\textsuperscript{c}, Graham P. Cook\textsuperscript{b}, G. Eric Blair\textsuperscript{a,}\textsuperscript{*}

\textsuperscript{a} Institute of Molecular and Cellular Biology, Faculty of Biological Sciences, Garstang Building, Room 8.10a, University of Leeds, Leeds LS2 9JT, United Kingdom
\textsuperscript{b} Leeds Institute of Molecular Medicine, St. James’s University Hospital, Leeds LS9 7TF, United Kingdom
\textsuperscript{c} Department of Internal Medicine II, Division of Hematology, Immunology and Oncology, University of Tübingen, Tübingen, Germany

Received 12 December 2006; accepted 12 February 2007
Available online 20 February 2007
Edited by Veli-Pekka Lehto

Abstract Expression of CUB domain containing protein 1 (CDCP1) is upregulated in carcinoma cells. We quantitated CDCP1 gene expression in matched normal colon and tumour tissue and compared the level of expression to other genes upregulated in colorectal tumourigenesis. Furthermore, we show that the CDCP1 gene generates two transcripts which are co-expressed in normal and matched tumour tissue as well as in the majority of cell lines analysed. However, intracellular localisation studies revealed that only one of these transcripts encodes a protein that is localised to the cell surface.

Keywords: CUB domain; CUB domain containing protein 1; Epithelial cell adhesion molecule; Carcinoembryonic antigen; Colorectal cancer

1. Introduction

Tumour-associated antigens (TAAs) such as carcinoembryonic antigen (CEA) and epithelial cell adhesion molecule (Ep-CAM) are well characterised proteins used as diagnostic and prognostic markers in colorectal, breast and cervical cancers \cite{1,2} as well as therapeutic targets for immunotherapy. The CUB domain containing protein 1 (CDCP1) (also described as SIMA135 and Trask) is a recently-described tumour marker \cite{3–5}. Sherl-Mostageer et al. first described an mRNA transcript encoding this protein using representational difference analysis to identify differentially regulated genes in cancer \cite{3,8}. They described a putative transmembrane protein containing several CUB (complement protein subcomponents Clr/Cls, Urchin embryonic growth factor and Bone morphogenetic protein 1) domains. CUB domains are characterised by immunglobulin-like folds and are involved in protein:protein and protein:carbohydrate interaction \cite{6,7}. CDCP1 isoform 1 contains at least 2 putative CUB domains \cite{3–5} and has been reported to have a role in cellular adhesion linked to cell signalling via Src-family kinases and Protein Kinase C\textsuperscript{[4,8,9]}. CDCP1 was found to be upregulated in human lung and colon tumours as well as an erythroleukaemic cell line. Further studies have reported that CDCP1 was expressed on haematopoietic stem cells, mesenchymal stem cells, neuronal progenitor cells \cite{10,11}, human epidermoid carcinoma cell lines \cite{3,8}, breast and prostate carcinoma cell lines \cite{4}. Ikeda et al. have described a negative correlation of CDCP1 expression with methylation of CpG sites within its promoter region and suggest that expression of CDCP1 in cancer is attributable to such epigenetic regulation \cite{12}.

The CDCP1 gene contains nine exons and the first mRNA transcript to be described (isoform 1) is approximately 6 kb in length. The isoform 1 protein product is composed of 836 amino acids including a 29 residue signal sequence and a transmembrane domain which gives rise to a type I orientated integral membrane protein \cite{3,5}. The observed $M_r$ of CDCP1 isoform 1 protein ranges from 135 to 140 kDa \cite{3,4,8,9} which differs from its calculated $M_r$ (approximately 90 kDa) due to extensive glycosylation. Northern blot analysis suggested the presence of a shorter transcript and analysis of EST databases identified a cDNA clone encoding a shorter transcript (accession number BC021099, IMAGE clone number 4590554) \cite{13}. This suggested that a second CDCP1 transcript might be expressed in tumour cells. Here we show that the two CDCP1 transcripts are co-expressed in normal and tumour tissue.

To establish the importance of CDCP1 expression in colorectal cancer relative to CEA and Ep-CAM, we have compared their expression in colorectal tumour and matched normal tissue. We further investigated the expression of CDCP1 in a panel of colorectal and cervical cancer cell lines and report for the first time the expression of CDCP1 in cervical carcinoma cell lines and the presence of the less-well characterised isoform 2 transcript. Expression of CDCP1 cell-surface protein was found to vary between colorectal carcinoma cell lines but this should not detract from its importance as a tumour marker as such variation was also observed for both Ep-CAM and CEA.

2. Materials and methods

2.1. Cell culture

A panel of colon carcinoma cell lines (SW480, HCT116, DLD-1, WiDr, HRT18, HCT8, COLO 741, COLO 320, HT-29), cervical carcinoma cell lines (CaSki, SiHa and HeLa), the lung carcinoma cell line...
A549 and Chinese hamster ovary cells (CHO) were maintained as monolayer cultures in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% (v/v) foetal calf serum, 100 U/mL penicillin and 100 μg/mL streptomycin (Sigma–Aldrich).

2.2. RNA isolation, cDNA synthesis, quantitative and non-quantitative PCR

RNA was isolated from 10⁶ cells using TRIzol (Gibco Life Technologies) and treated with DNase (Promega) following the manufacturer’s protocols. RNA isolated in this manner and RNA isolated from colon tumours and normal tissue obtained commercially (Ambion) was used to synthesise cDNA using Superscript II reverse transcriptase (Invitrogen) according to the manufacturer’s protocols. Quantitative real-time PCR was performed using the SYBR green method using approximately 0.1 μg cDNA template, 1 μM of both forward and reverse primers and SYBR green master mix (Applied Biosystems). The samples were prepared in 96 well plates and analysed using a ABI PRISM 7900 HT (Applied Biosystems) following the manufacturer’s protocols. Using Primer Express software (Applied Biosystems) the following primers were designed: Forward (5'-GAAGGTGAAGGTCGGAGTC-3') and reverse (5'-GATGATGACACAGACGTTCAGTAGATAGTA-3').

2.3. Plasmids

A pcDNA3.1 plasmid encoding C-terminal FLAG-tagged CDCP1 isoform 1 (pcDNA3.1-CDCP1iso1) was kindly provided by Stephen Solhoff, Beth Israel Deaconess Medical Center, Boston, USA [8]. CDCP1 isoform 2 was obtained from the IMAGE consortium (clone 4590554) and the coding sequence sub-cloned from the original pOTB7 vector into pcDNA3 (Clontech) (pcDNA3-CDCP1iso2) using EcoRI and XhoI restriction sites.

2.4. Transfection and immunofluorescence microscopy

Sub-confluent CHO cells grown on glass cover slips were transfected with the plasmid expression vectors pcDNA3.1-CDCP1iso1, pcDNA3-CDCP1iso2 and empty pcDNA3 using Lipofectamine (Invitrogen) following the manufacturer’s protocol. Twenty four hours post-transfection, the cells were washed twice with phosphate-buffered saline (PBS) and fixed with 4% paraformaldehyde in PBS. The cells were washed with PBS and blocked by incubation in PBS/10% (v/v) NGS and resuspended with 1:50 in PBS/10% (v/v) NGS. Analysis was performed using a FACScalibur benchtop flow cytometer and CellQuest software (Becton Dickinson).

3. Results and discussion

Since CDCP1 is known to be upregulated in tumour cells, we first set out to validate its expression as a colorectal cancer marker and potential therapeutic target by analysing its expression in colon tissue compared to that of the well-characterised tumour markers Ep-CAM and CEA. RNA samples from colorectal tumour and matched normal tissue were subjected to quantitative real time PCR using primers specific for sequences within exon 3 of CDCP1 which is common to both CDCP1 mRNA isoforms (Fig. 1). Comparison of CDCP1, Ep-CAM and CEA expression in these samples demonstrated equally significant upregulation of these genes (Fig. 2), confirming the potential of CDCP1 as a novel tumour marker.

The nucleotide sequence databases contain two CDCP1 transcripts; isoform 1 (accession number NM133974) and isoform 2 (accession number BC020199) suggesting alternative

---

![Fig. 1. Schematic representation of CDCP1 isoform 1 and isoform 2 transcripts.](image-url)
splice variants. Alignment of these transcripts with the \textit{CDCP1} gene allowed us to identify the origins of these isoforms (Fig. 1A). The isoform 2 transcript is 1.4 kb in length and the open reading frame continues from the exon 4 boundary into the adjacent intron giving rise to a putative secreted protein of 343 amino acids with identity to the first 341 residues of isoform 1 including one putative CUB domain of isoform 1 but lacking the transmembrane domain (Fig. 1B). Its expression in tumour cells and its protein product have yet to be described so we first used immunofluorescence microscopy to characterise the intracellular localisation of the two isoforms. CHO cells transfected with expression vectors coding for either \textit{CDCP1} isoform 1 or 2 (Fig. 3) showed that only isoform 1 encodes a protein localised to the cell surface (Fig. 3E), consistent with the presence of the transmembrane domain in this molecule. In contrast, isoform 2 product could only be detected intracellularly following permeabilisation (Fig. 3H). Cells transfected with empty expression vector showed no staining (data not shown). This demonstrates that the isoform 2 transcript encodes a protein that is not present at the surface of cells. Interestingly, we did not detect a morphological change similar to that observed by Bhatt et al.\cite{4} in MDA-468TR cells overexpressing \textit{CDCP1} isoform 1. This difference may be attributable to the cell lines used.

We next analysed expression of the two transcripts in carcinoma tissues and cell lines. We designed an isoform 1 reverse primer specific to the sequence spanning the exon 4–5 boundary and an isoform 2 reverse primer specific to a sequence within the 3’ UTR of the isoform 2 transcript. Using a common forward primer specific to a sequence within exon 3 of \textit{CDCP1}, PCR generated an isoform 1-specific 455 bp product and an isoform 2-specific 294 bp product. Reverse transcriptase (RT) PCR, using mRNA derived from CHO cells transfected with expression vectors coding for either isoform 1 or isoform 2 revealed that these primers were specific for the particular isoform (Fig. 4A) Samples where RT was omitted to control for residual genomic DNA were negative (data not shown). Using these primers, we were able to identify both iso-

Fig. 2. Expression of \textit{CDCP1}, Ep-CAM and CEA in colorectal tissue. Quantitative real-time PCR was performed on cDNA synthesised from an RNA sample obtained from human colon adenocarcinoma and matched normal tissue (Ambion). Primers specific to a region in exon 3, common to both \textit{CDCP1} isoforms, were used to compare expression of \textit{CDCP1} transcripts to those of \textit{Ep-CAM} and \textit{CEA} relative to \textit{GAPDH} control mRNA. Values represent the mean of six independent experiments and error bars ± S.E. \textit{P} < 0.05 between all tumour and matched normal samples for all targets.

Fig. 3. Immunofluorescence analysis of \textit{CDCP1} expression in transfected cells. Immunofluorescence microscopy of CHO cells transfected with expression vectors coding for \textit{CDCP1} isoform 1 (A–F) and isoform 2 (G–L), detergent-permeabilised (P) (A–C, G–I) to reveal surface and intracellular \textit{CDCP1} and fixed intact cells (NP) (D–F, J–L) to reveal cell-surface\textit{CDCP1}. Cells were treated with the anti-\textit{CDCP1} mouse monoclonal antibody CUB1 followed by Alexa 594-conjugated anti-mouse IgG antibody fragment and analysed using the ApoTome (Zeiss).
form 1 and isoform 2 transcripts in colorectal carcinoma tissue (Fig. 4B) and cell lines (Fig. 4C) as well as cervical and lung tumour cell lines (Fig. 4D). Interestingly, the presence of both isoform 1 and isoform 2 transcripts was invariant and two of the colorectal cell lines (Colo741 and Colo320) were found to be negative for CDCP1 expression. This result was further substantiated by analysis of CDCP1 surface expression by flow cytometry, a measure of CDCP1 isoform 1 protein expression (Fig. 5A and B) and in accordance with recently published data [12]. Surface expression of CDCP1 on the cells tested varied, however surface expression was detected on colorectal, cervical and lung carcinoma cell lines with notably high levels observed for HeLa cells. The colorectal cell lines Colo320 and Colo741 expressed no significant level of isoform 1 protein on their cell surface as judged by flow cytometry, consistent with the lack of detectable CDCP1 transcripts in these cells. This apparent lack of CDCP1 expression in these cell lines is most notable and may distinguish between different types of colorectal cancer.

The major route taken by colorectal cells towards transformation is associated with chromosomal instability (CIN) resulting in an aneuploid karyotype. A similar phenomenon can be found in cervical and non-small cell carcinomas [15]. A smaller proportion of colorectal carcinomas present the microsatellite instability (MSI) phenotype. These cells exhibit genome-wide instability at the nucleotide level associated with a near-diploid karyotype [16]. This phenotype has not been associated with either cervical or non-small cell carcinoma cells [17,18]. However, CDCP1 expression was not correlated with cell lines described in the literature as exhibiting either CIN or MSI phenotype [16,19,20]. Furthermore, and in accordance with the findings of Hooper et al., CDCP1 expression did not seem to be correlated with metastatic ability as the CDCP1-negative Colo741 cells were derived from a pelvic wall metastasis [3].

Interestingly, surface expression of Ep-CAM and CEA also varied between cell lines (Fig. 5C and D). WIDR were found to have exceptionally high surface levels of CEA (Fig. 5C) which is in accordance with previous observations [21] and we observed an apparent correlation of low CDCP1 with Ep-CAM in both Colo741 and Colo320 and CEA in Colo320 (Fig. 5).

This leads us to propose that elevated expression of CDCP1 could be a result but not a direct cause of transformation which does not preclude the possibility that it is a marker of an as-yet undefined sub-set of carcinomas. Further studies investigating the expression of CDCP1 in primary colorectal carcinoma cells in relation to progression using the Duke’s staging system are currently being pursued in our laboratory.

The detection of CDCP1 isoform 2 transcripts in cell lines expressing CDCP1 is an interesting finding. This transcript encodes a signal sequence and, in accordance with our observation that it is not retained at the plasma membrane, is likely to encode a secreted soluble protein whose physiological role remains unknown. Hooper et al. describe a membrane-shed moiety of CDCP1 isoform 1 of approximately 110 kDa which would be predicted to contain most of the extracellular domain [3]. They also described strong immunohistochemical staining in the goblet cells of normal colon crypts, lumen of glands and in the glandular mucus of colon carcinoma, suggesting the production of a soluble form of CDCP1. This may be due to the membrane-shed extracellular domain of isoform 1 but it is also possible that these authors detected isoform 2. It has also been reported that CDCP1 isoform 1 undergoes proteolytic digestion by the serine protease MT-SP1 after amino acid R369 and this cleavage is linked to intracellular signaling pathways [4]. CDCP1 isoform 2 possesses the first 341 residues of isoform 1 with an additional 2 amino acids at its C-terminus. The proteolytically-released moiety of CDCP1 isoform 1 described by Bhatt et al. and isoform 2 would have high identity and both possess the first CUB domain of CDCP1 [4]. Furthermore, Bühring et al. reported that the CUB1 antibody (which we found to recognise CDCP1 isoform 2) stimulated growth of erythroid colonies from CD34+ stem/progenitor cells [10]. It is still unknown what physiological role may be played by any of these three soluble forms of CDCP1, but this warrants further investigation.

Fig. 4. Analysis of CDCP1 mRNA expression in a panel of tumour cell lines. cDNA synthesised from RNA extracted from CHO cells transfected with expression vectors (A), RNA obtained from human colon adenocarcinoma and matched normal tissue (Ambion) (B) and a panel of human tumour cell lines (C and D) was used as the template for PCRs using primer sets specific for GAPDH (G), CDCP1 isoform 1 (1) and CDCP1 isoform 2 (2). The resulting products were resolved by agarose gel electrophoresis and the expected size of the GAPDH, CDCP1 isoform 1 and isoform 2 primer set products are indicated.
In conclusion, we show that the CDCP1 gene generates two co-expressed transcripts, only one of which encodes a cell surface molecule. Expression of CDCP1 in cell lines is similar to that of the well documented tumour markers CEA and EP-CAM and like these molecules, CDCP1 might prove to be a useful therapeutic target.

Acknowledgments: We thank Yorkshire Cancer Research for support and A. Trejdosiewicz and J. Jarvis for technical assistance. H.-J. Buhring was supported by the Deutsche Forschungsgemeinschaft (SFB 685, project Z2).

References


Fig. 5. Analysis of CDCP1 isoform 1 protein product in a panel of tumour cell lines. A panel of cell lines was analysed by flow cytometry using the CUB1 mouse anti-CDCP1 antibody followed by FITC-conjugated anti-mouse IgG (A and B), FITC-conjugated anti-Ep-CAM (C) and mouse anti-CEA antibody followed by FITC-conjugated anti-mouse IgG (D). The mean fluorescence intensity (MFI) per cell of positive cells was determined using a FACScalibur benchtop flow cytometer. Values represent the average MFI of three independent experiments, error bars represent ±1 S.E.


