One single mRNA encodes the centrosomal protein CCD41 and the endothelial cell protein C receptor (EPCR)

Karsten Rothbarth, A. Reza H. Dabaghian, Hermann Stammer, Dieter Werner*

Division Biochemistry of the Cell (B0300), German Cancer Research Center, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

Received 14 May 1999; received in revised form 24 June 1999

Abstract The cDNA encoding the centrosomal protein CCD41 is identical with the cDNA for the endothelial cell protein C receptor. This finding is not due to an artefact, e.g. caused by selection of false positive clones. The segment of the CCD41 cDNA encoding the protein originally termed CCD41 and deletion mutants of it were fused with the nucleotide sequence encoding the enhanced green fluorescent protein (EGFP). Transfection and expression of the full length construct produces a fusion protein mainly located in cell membranes reflecting the receptor-type protein. Deletion mutants, e.g. those where the signal sequence is deleted, result in fusion proteins which are exclusively incorporated into a small perinuclear structure which is the site of the centrosome. This result suggests that post-translational modification, namely deletion of the signal sequence, is decisive for the centrosomal location of the resulting centrosomal protein while the unprocessed protein is incorporated into cell membranes.

Key words: Centrosomal protein; CCD41; Endothelial cell protein C receptor; Enhanced green fluorescent protein

1. Introduction

Eukaryotic genomes comprise encrypted messages that are displayed by post-transcriptional and post-translational mechanisms, e.g. several mRNAs encoding proteins with different functions may derive from one gene by alternative splicing, alternative polyadenylation and RNA editing. Moreover, post-translational processes increase the number of functional proteins encoded by only one gene. An example for the latter case is described in this paper.

Protein C plays an important role in blood coagulation. Its binding to the endothelial cell surface is greatly enhanced by the endothelial cell protein receptor (EPCR) which facilitates protein C activation by the thrombin-thrombomodulin complex [1–6]. Molecular cloning and sequencing of the cDNAs to the human [7], bovine [8] and murine [8] EPCRs show high interspecies similarities and reveal close identity with a murine G2/M phase-prevalent transcript identified before by differential hybridization with cell cycle phase-specific probes [9]. This finding was enigmatic because the protein derived from the G2/M phase-prevalent transcript has been shown to encode a centrosomal protein in proliferating cells [10]. The functional difference between the two proteins claimed to be encoded by the same mRNA and other discrepancies gave rise to a critical discussion [1,7,8]. However, post-translational mechanisms resulting in functionally divergent proteins were not considered. By using green fluorescent protein technology, it is possible to solve this enigma. Evidence is presented indicating that the receptor-type protein consists of the complete polypeptide sequence encoded by the full length mRNA while signal sequence deletion results in the centrosomal protein.

2. Materials and methods

2.1. Preparation of expression constructs

The enhanced green fluorescent protein (EGFP)-encoding sequence (Clontech, cat. no. 6077-1) was PCR-amplified and *Sal*I (5′ end)/-*Hin*dIII (3′ end) sites were added during amplification. Following digestion with *Sal*I/*Hin*dIII, the amplification product was ligated with the pBluescript KS+ vector digested with the same enzymes resulting in the plasmid pBlue-(*Kpn*I, *Sal*I)-EGFP-(*Not*I).

The cDNA sections encoding the polypeptides specified in Fig. 1 (no. 1–3) were PCR-amplified by appropriate primers. Restriction sites *Kpn*I (5′ end)/*Sal*I (3′ end) were added during amplification. Following digestion with *Kpn*I/*Sal*I, the fragments were cloned in the corresponding sites of the pBluescript vector containing the EGFP-encoding sequence.

The fused sequences were excised by *Kpn*I/*Not*I and recloned in the *Kpn*I/*Not*I sites of the pcDNA3 expression vector (Invitrogen).

2.2. Cells, transfection, microscopy, microphotography

Qia-tip (Qiagen)-purified plasmids were used for transfections. Suspension cells (Ehrlich ascites tumor, EAT) were transiently transfected with 5–20 μg of plasmid DNA by electroporation using the Bio-Rad Gene Pulser II (cell density 10^7 per ml, electrode distance $D = 4$ mm, 366 V/950 μF). Transfected cells and immunostained cells were inspected by means of a Zeiss Axiovert S100 TV microscope equipped with objectives for phase contrast and fluorescence mode and a digital electronic camera (Hamamatsu C4742-95). The equipment was controlled by a Power MacIntosh computer (Cell Imaging software Openlab, Improvision, Warwick, UK).

2.3. Immunofluorescence

EAT cells were transiently transfected by electroporation with the plasmid containing the EGFP-tagged CCD41 mutant with a deleted signal sequence (specified in Fig. 1, no. 2). After 24 h, the cells were collected, washed with PBS and fixed in 100% methanol. The specimens were rehydrated before application of the antibody (anti-α-tubulin, Amersham Life Science, rhodamine red X, Dianova) in steps (70, 30 and 0% ethanol in PBS).

3. Results

3.1. CCD41 sequence and CCD41 deletion mutants

The originally published version of the CCD41-encoding nucleotide sequence [10] contains several sequencing errors due to base compression. Resequencing of the CCD41 cDNA by an automated sequencing device could eliminate these errors and the updated version is now available under the accession number X43768. A prominent feature of the CCD41 polypeptide is an initial hydrophobic signal peptide

*Corresponding author. Fax: (49) (6221) 423457.
E-mail: werner@dkfz-heidelberg.de

PII: S0014-5793(99)01074-1

```
no 1
|--->
MLTKFLLLLLLLLPGCALCNSDGSQSLH

no 2                no 3
|--->           (M)|--->
MLQISYFQDHHHVRHQGNASLGKLLTHT

LEGPSQNVTILQLQPWQDPESWERTESG

LQIYLTQFESLVKLVYRERKENVFFPLT

VSCSLGCELPEEEEEGSEPHVFFDVAVN

GSAFVSFRPKTAVWVSGSQEPSKAANFT

LKQLNAYNRTRYELQEFLQDTCVEFLEN

HITTQNMKGSQTGRSYTSLVLGILMGCF

IIAGVAVGIFMCTSGRGLLII->EGFP
```

Fig. 1. Amino acid sequence derived from the updated version of the CCD41 cDNA sequence and the sections of the sequences fused with EGFP. The fusion constructs described in Section 2 were designed to express the following proteins in fusion with EGFP. No. 1: the complete CCD41 protein including the putative signal sequence. No. 2: a polypeptide starting at the second methionine residue of the sequence (amino acid position 29) in which the signal sequence is deleted. No. 3: a polypeptide with another 18 amino acid deletion. For expression of this latter polypeptide, a primer-coded methionine was inserted which is not present in the CCD41 peptide sequence. The putative hydrophobic signal sequence is underscored.

(Fig. 1). Such signal sequences contain specific information for performing distinct functions in targeting and membrane insertion, e.g. they allow for variable membrane translocation such that the protein either remains in the cytoplasm, is inserted into the membrane or is translocated into the lumen [11]. Accordingly, the CCD41 cDNA-encoded EPCR membrane protein and the centrosomal protein could arise by alternative translocation, signal sequence cleavage or missing signal sequence cleavage. In order to test this hypothesis, expression constructs were prepared by fusing the complete coding sequence of the CCD41 cDNA and deletion mutants with the EGFP-coding sequence. The polypeptides expectedly expressed by these constructs are specified in Fig. 1 (no. 1–3).

3.2. Expression and cellular location of the CCD41-EGFP fusion proteins missing the signal sequence

The assumption of functionally different proteins encoded by the same mRNA might be due to artefacts, e.g. the selection of false positive clones during expression cloning of the EPCR protein [7] or due to a false positive antibody developed against the recombinantly expressed CCD41 antigen [10]. Any artefact of this kind would be revealed by green fluorescent protein technology because this approach provides reliable information on the intrinsic cellular distribution of any protein encoded by a known nucleotide sequence. A large bulk of literature indicates that EGFP-tagged proteins are directed to the physiological location of the functional protein. For example, a well-studied EGFP-tagged centrosomal protein (centrosomin A) is incorporated into the centrosome while a closely related EGFP-tagged protein comprising a nuclear location signal is transported into nuclei [12]. By application of this technology, it becomes evident that EGFP-tagged CCD41 proteins missing the signal sequence are first and exclusively concentrated at dense cytoplasmic spots located at the expected site of the centrosome (Fig. 2B). This result supports the validity of previous immuno cytochemical data on the centrosomal location of a protein encoded by the CCD41 cDNA [10]. The level of this protein is regulated in a strictly cell cycle-dependent manner when expressed from its own promoter [10]. In contrast, when controlled by the strong CMV promoter in transfected cells, its expression is high throughout the cell cycle. This unscheduled cell cycle-independent expression obviously conflicts with the cell's regulation system resulting in the polarity-related mitotic disorders shown in Fig. 2C and D. At 24–48 h post-transfection, the

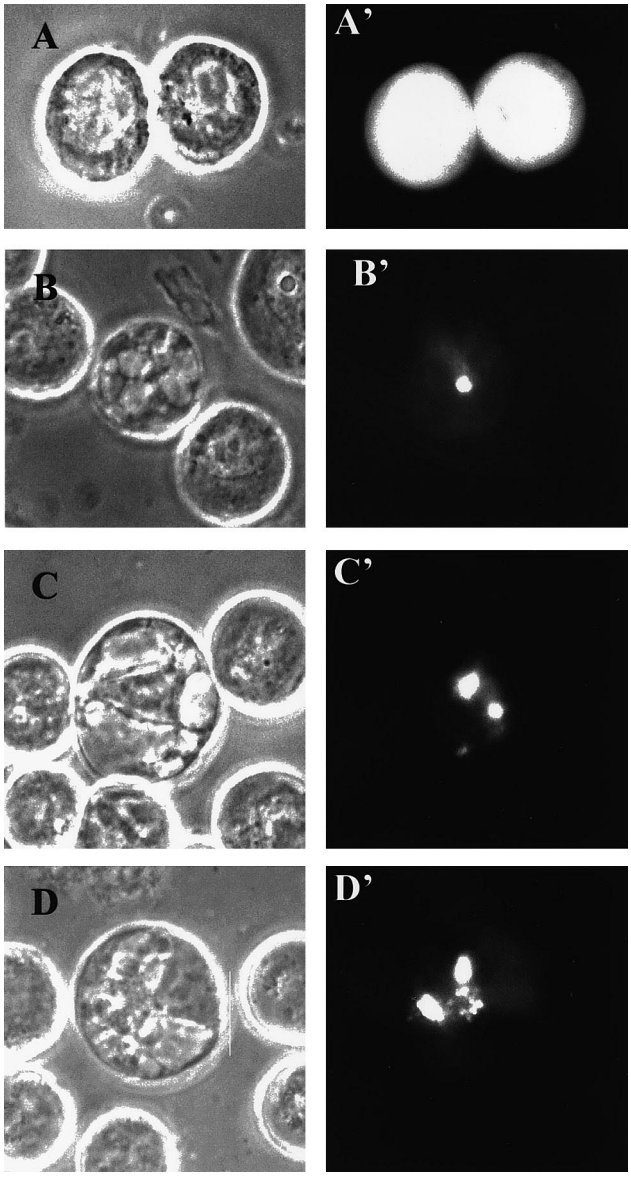

Fig. 2. Centrosomal location of EGFP-fused, signal sequence-depleted CCD41 proteins. A, A′: EAT cells transfected with the control plasmid expressing only EGFP, (A) in phase contrast, (A′) in green fluorescence mode. The non-functional EGFP is highly expressed and not enriched in any cellular compartment. B–D: Cells transfected with the fusion construct specified in Fig. 1, no. 3. B, B′: 12–24 h post-transfection, the fusion protein appears in a single cytoplasmic spot. C and D: Later, 12–48 h post-transfection, the transfected cells become enlarged and multinucleated while the fusion protein is located in several cytoplasmic spots. It should be noted that transfection with the construct specified in Fig. 1, no. 2, results in the same effects (not shown).

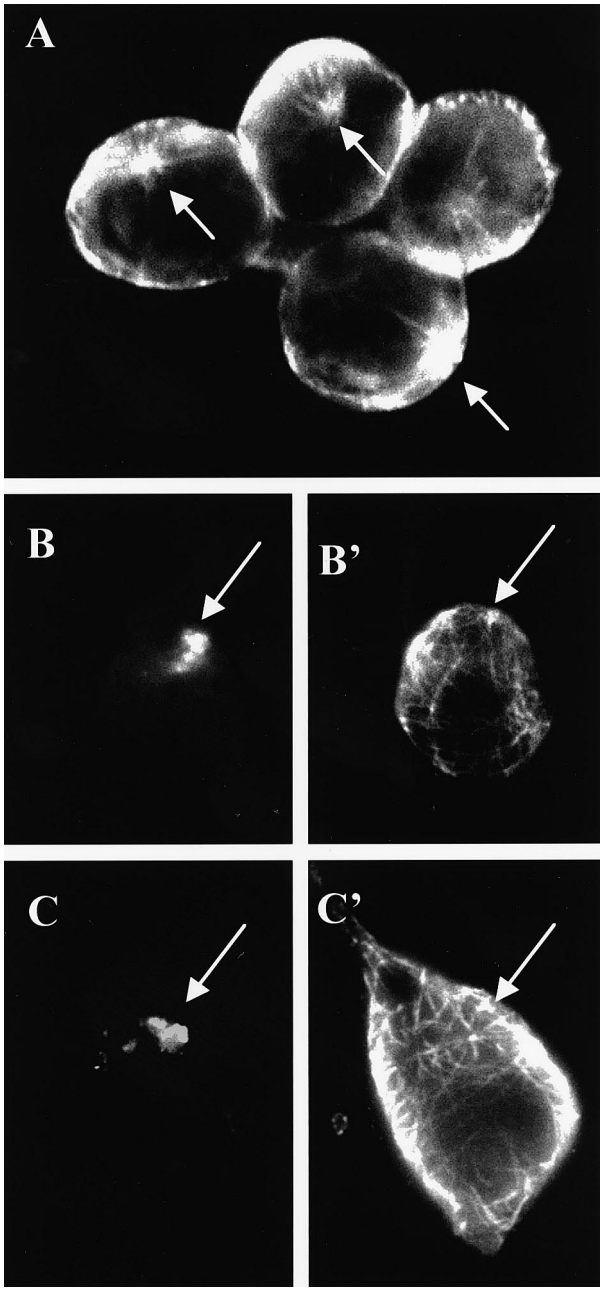

Fig. 3. Organization of the microtubular system in non-transfected cells and in cells transfected with a construct expressing EGFP-fused, signal sequence-depleted CCD41 protein. A: In non-transfected cells, the α-tubulin antibody (red fluorescence) indicates the position of the centrosome (arrows) by concentric arrays of microtubules. The location of the fusion protein (green fluorescence) in transfected cells is shown in B and C. In this case, the microtubules (red fluorescence, B′ and C′) show a more network-like appearance and are not concentrated at a single spot. It appears that the centrosome overloaded with the CCD41 centrosomal protein version has lost its microtubular-organizing capacity.

genetically manipulated cells become multinucleated and they exhibit multicentric locations of the fusion proteins. Moreover, in a fraction of cells, mitosis appears to be arrested in the late prophase (not shown). These polarity-related disorders reveal the relationship between a CCD41 cDNA-encoded protein and the centrosome which is the organelle responsible for cell polarity. This conclusion is further supported by transfection-induced disorders of the microtubular system (Fig. 3B anm C).

It should be noted that the obvious misfunction of the centrosome in transfected cells is not due to the EGFP-tag of the vector-encoded protein because the same mitotic disorders are observed in cells transfected with expression constructs missing the EGFP portion (not shown) and EGFP expression per se does not show any mitotic disorder (Fig. 2A).

3.3. Expression and cellular location of the CCD41-EGFP fusion protein comprising the signal sequence

Following transfection with the construct comprising the complete coding section of the CCD41 cDNA in fusion with EGFP, significant fluorescence is mainly found in the plasma membrane. In this case, the site of the centrosome is only weakly labelled (Fig. 4A) which points to a delayed or reduced rate of precursor protein processing. The post-translational process resulting in the signal sequence-depleted centrosomal protein appears to be disturbed by the CMV-dependent and non-regulated expression of the full length CCD41 sequence. Since the unprocessed CCD41 protein is transported and obviously integrated in the cell membrane, it exhibits the expected characteristics of the EPCR protein. Thus, by genetic manipulation, it can be shown that the CCD41 mRNA has the potential to encode a centrosomal protein and a mem-

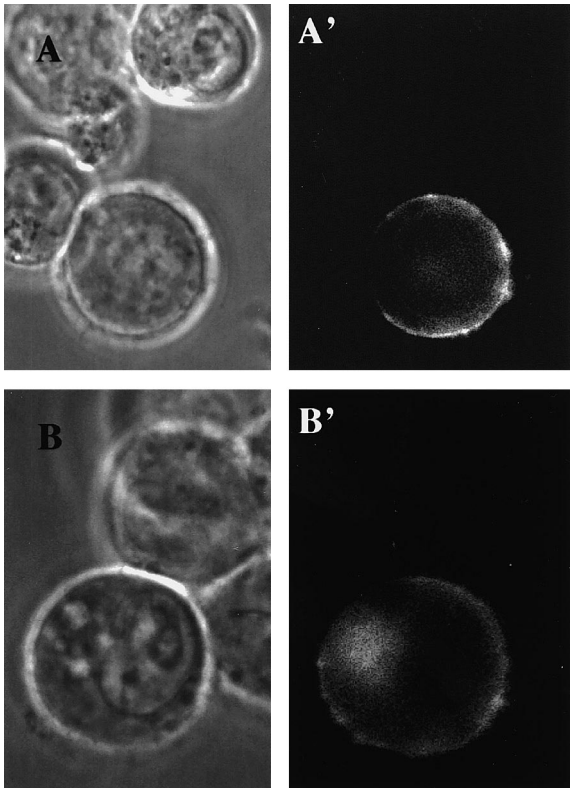

Fig. 4. Evidence indicating that the unprocessed CCD41 protein (Fig. 1, no. 1) is a membrane protein. A and B: Cells transfected with the vector construct expressing the complete coding section of the CCD41 cDNA in fusion with EGFP. The expressed green fluorescent fusion protein becomes mainly located in the cell membrane. A fluorescent cytoplasmic area which could reflect the site of the centrosome is weakly labelled in only a fraction of transfected cells (B′). A and B: Phase contrast. A′ and B′: Green fluorescent mode.

brane protein. Cell type or cell cycle-dependent post-translational mechanisms are decisive for the release of the variant products.

4. Discussion

Fluorescence in situ hybridization shows that there is only one gene locus hybridizing specifically with the CCD41 probe (2H1-3) [13]. Accordingly, the centrosomal protein and the EPCR protein must derive from the same gene either by post-transcriptional or post-translational processes.

Vector-expressed and EGFP-tagged CCD41 protein versions missing the signal sequence are directly incorporated into centrosomes and no traces of the fusion proteins can be detected in the plasma membrane of transfected EAT cells. This result is consistent with immuno cytochemical data on the initial location of the CCD41 protein expressed by its own promoter in rapidly proliferating cells [10]. Accordingly, it can be concluded that the centrosomal protein encoded by the CCD41 mRNA arises from the precursor protein by deletion of the signal sequence. The cleavage of the signal sequence must occur in such a way that the protein remains cytosolic. However, it cannot be excluded that the full length protein encoded by the CCD41 mRNA is first integrated into membranes from where the truncated version is recovered in a recycling step. Finally, the truncated protein must be targeted to the centrosome by a still unknown sequence motif. If over-expressed under vector control, it creates the centrosomal dysfunctions described above.

In contrast, the vector-expressed and EGFP-tagged complete CCD41 protein version is almost exclusively targeted to the plasma membrane, suggesting that the protein with a non-deleted signal sequence reflects the receptor-type protein. Only weak fluorescence is found in this case at the site of the centrosome which points to a limited capacity of the cells to produce the truncated version of the protein. It should be noted that in non-manipulated cells, the expression and processing of the CCD41 precursor protein is strictly regulated in a cell cycle phase-dependent manner [10], e.g. in the G1 phase, the transcript is below the level of its detection [9] and the protein can only be found in minute amounts exclusively in the centrosome [10]. Higher transcript and protein levels are only observed late in the cell cycle, e.g. in the G2/M phase [10]. Thus, vector-dependent over-expression of the precursor protein during early cell cycle phases suggestively overspills the cell's capability to produce the centrosomal version of the protein in significant amounts. Consequently, the overflow of the precursor protein is alternatively processed and targeted to the plasma membrane.

These results show that it is possible to manipulate rapidly proliferating non-endothelial cells in such a way that they preferentially express the receptor-type protein. This procedure may be useful for investigations of the EPCR receptor in non-endothelial cells. Since endothelial cells are specialized to express the full length CCD41 protein constitutively at high levels [7,8], it is not surprising that most of it remains unprocessed and that it is found mainly in the plasma membrane. Accordingly, the CCD41 mRNA encodes the messages for two proteins with most different functions and it depends on post-translational processes which of the messages is preferentially expressed.

References

[1] Fukudome, K., Kurosawa, S., Stearns-Kurosawa, D.J., He, X., Rezaie, A.R. and Esmon, C.T. (1996) J. Biol. Chem. 271, 17491–17498.
[2] Stearns-Kurosawa, D.J., Kurosawa, S., Mollica, J.S., Ferrell, G.L. and Esmon, C.T. (1996) Proc. Natl. Acad. Sci. USA 93, 10212–10216.
[3] Laszik, Z., Mitro, A., Taylor, F.B., Ferrell, G. and Esmon, C.T. (1997) Circulation 96, 3633–3640.
[4] Regan, L.M., Mollica, J.S., Rezaie, A.R. and Esmon, C.T. (1997) J. Biol. Chem. 272, 26279–26284.
[5] Kurosawa, S., Stearns-Kurosawa, D.J., Hidari, N. and Esmon, C.T. (1997) J. Clin. Invest. 100, 411–418.
[6] Fukudome, K., Ye, X., Tsuneyoshi, N., Tokunaga, O., Sugawara, K., Mizokami, H. and Kimoto, M. (1998) J. Exp. Med. 187, 1029–1035.
[7] Fukudome, K. and Esmon, C.T. (1994) J. Biol. Chem. 269, 26486–26491.
[8] Fukudome, K. and Esmon, C.T. (1995) J. Biol. Chem. 270, 5571–5577.
[9] Lu, X., Dengler, J., Rothbarth, K. and Werner, D. (1990) Gene 86, 185–192.
[10] Rothbarth, K., Petzelt, C., Lu, X., Todorov, I.T., Joswig, G., Pepperkok, K., Ansorge, W. and Werner, D. (1993) J. Cell Sci. 104, 19–30.
[11] Martoglio, B. and Dobberstein, B. (1998) Trends Cell Biol. 8, 410–415.
[12] Petzelt, C., Joswig, G., Mincheva, A., Lichter, P., Stammer, H. and Werner, D. (1997) J. Cell Sci. 110, 2573–2578.
[13] Mincheva, A., Rothbarth, K., Werner, D. and Lichter, P. (1995) Mamm. Genome 6, 444.