Journal of Cell Science 110, 2573-2578 (1997)
Printed in Great Britain © The Company of Biologists Limited 1997
JCS3640

The centrosomal protein centrosomin A and the nuclear protein centrosomin B derive from one gene by post-transcriptional processes involving RNA editing

Christian Petzelt[1,*], Gaby Joswig[2], Antonaeta Mincheva[3], Peter Lichter[3], Hermann Stammer[2] and Dieter Werner[2]

[1]Experimental Anaesthesiology, Clinic for Anaesthesiology and Intensive Care, Charité, Humboldt-University, D-10117 Berlin, Germany
[2]Division Biochemistry of the Cell, German Cancer Research Center, PO Box 101949, D-69009, Heidelberg, Germany
[3]Division Organization of Complex Genomes, German Cancer Research Center, PO Box 101949, D-69009, Heidelberg, Germany

*Author for correspondence (e-mail: petzelt@ukrv.de)

SUMMARY

The identification of a gene encoding concomitantly a nuclear protein and an intrinsic centrosomal protein further emphasizes the close and presumably developmental relationship between the cell nucleus and the centrosome. Screening of a murine RNA-based cDNA library with an antiserum to a centrosomal protein and rescreening with the insert of an initial clone released two complete cDNAs (1.2 kbp and 2.2 kbp) coding for proteins with notable characteristics. The amino-terminal sections of centrosomin A (276 amino acid residues, molecular mass 34.5 kDa) and of centrosomin B (447 amino acid residues, molecular mass 54.8 kDa) are identical over 272 amino acid residues. The carboxy-terminal section of the larger protein comprises additional 175 amino acid residues including nuclear location signals. The mRNAs encoding centrosomin A and B derive from a single gene. Chromosomal mapping (FISH) shows only one significantly hybridizing gene locus on chromosome 19D. PCR with genomic DNA as template and primer pairs complementary to the sequence which is identical in centrosomin A and B cDNAs results in amplification of only one DNA fragment. Moreover, one exon of the genomic sequence and the centrosomin B-encoding cDNA sequence include a G which is deleted in the centrosomin A-encoding cDNA. Accordingly, the two mRNAs are the products of either alternative splicing or alternative polyadenylation in combination with RNA editing. The recombinantly expressed chimeric protein consisting of centrosomin A and the green fluorescent protein from *Aequorea victoria* accumulates in centrosomes while the corresponding fusion protein with the centrosomin B sequence is transported into nuclei.

Key words: Centrosome, Centrosomal protein, Nuclear protein, cDNA cloning, Chromosome 19D, Alternative splicing, Alternative polyadenylation, RNA editing, Green fluorescent protein (GFP), Fusion protein

INTRODUCTION

Most of the proteins found in centrosomes and in pericentrosomal regions are multifunctional because they appear additionally either at the same time or during distinct cell cycle phases in other cellular compartments and/or cellular structures (Kalt and Schliwa, 1993). An example is the CCD41 protein which is located in G_1 phase exclusively in the centrosome while it accumulates later in the cell cycle also in perinuclear vesicles (Rothbarth et al., 1993). Other proteins are only transiently associated with centrosomes, e.g. during distinct cell cycle phases. The NuMA/centrophilin protein family represents a typical example for transiently centrosome-associated proteins. These proteins reside during interphase in the cell nucleus, and only in metaphase do they become associated with the spindle poles (Zeng et al., 1994; Tang et al., 1994). In contrast, proteins which are exclusively and permanently associated with centrosomes are rare. The best known example is γ-tubulin whose intrinsic involvement in the centriole structure is well established (Horio et al., 1991; Stearns et al., 1991; Joshi et al., 1992).

Another antigen which is exclusively and permanently associated with centrosomes was detected in Ptk cells by means of a monoclonal antibody (Joswig and Petzelt, 1990). This antigen was found to be involved in a centrosomal sub-structure undergoing characteristic shape-changes during cell cycle progression. Typical ring and spider-like structures could be resolved by immuno cytochemical techniques (Joswig and Petzelt, 1990). Later, a non-species-specific antiserum (GP1) was prepared which was considered to immunoreact with this antigen because it could detect centrosomal sub-structures like those detectable by the monoclonal antibody. However, the clear identification of this intrinsic centrosomal antigen at the sequence level met difficulties because the expression cloning approach resulted in two cDNAs encoding different but closely related proteins. By applying various recombinant techniques including the

expression of constructs consisting of the centrosomin-encoding sequences and the sequence encoding the green fluorescent protein from *Aequorea victoria* it was possible to identify the intrinsic centrosomal protein (centrosomin A). Moreover it is shown that the second cDNA encodes a nuclear protein (centrosomin B) and that the two proteins are encoded by one gene.

MATERIALS AND METHODS

Molecular cloning and sequencing

A first and partial cDNA clone was detected by antibodies in a plate-amplified Ehrlich ascites cell RNA-based cDNA library prepared in the λgt11 vector by sreening of 200,000 plaques (20,000 pfu per 130 cm^2 plate; Joswig et al., 1991). This was followed by screening of a plate-amplified Ehrlich ascites cell RNA-based cDNA library prepared in the λgt10 vector with the radiolabelled insert of the cDNA clone detected by antibodies. Five positives were detected on plaque lifts of 340,000 pfu (20,000 pfu per 130 cm^2 plate). Five positives were plaque-purified and the inserts were released from the phage arms by *Eco*RI digestion, and recloned in the pBluescript vector (Stratagene). Two types of phages with insert lengths of 1.2 kbp and 2.2 kbp were detected which could reflect complete cDNAs. The 3′ end sequence of the complete cDNA with the 1.2 kbp insert was identical with that of the original clone detected in the λgt11 library by antibodies. The protein encoded by this nucleotide sequence was termed centrosomin A (Joswig et al., 1991). Later, two cDNAs with 2.2 kbp inserts were sequenced (Sanger et al., 1977) using the DNA sequencing kit from Pharmacia and commercially available M13 as well as custom made primers and reverse primers. The protein encoded by this longer nucleotide sequence was termed centrosomin B. Other positives found by screening of the λgt10 library were incomplete centrosomin A or B cDNA clones.

Genomic DNA from Ehrlich ascites tumor cells was used as template for PCR amplification of a section of the gene encoding centrosomin A and B. Centrosomin A and B cDNA-fitting primer pairs were selected to obtain an amplified genomic sequence involving the single base difference between centrosomin A and B cDNA, e.g. corresponding to nucleotide positions 800-1,194 of the centrosomin B cDNA sequence. PCR amplification was performed by means of the Clontech advantage PCR kit (#k1906-1), and the ends were modified during PCR by addition of *Eco*RI sites. Following digestion with *Eco*RI the single DNA fragment was recloned in the *Eco*RI site of the pBluescript vector (Stratagene) and sequenced as described above.

Localization of the centrosomin B-encoding sequence to mouse chromosome 19D

The plasmid containing the 2.2 kbp centrosomin B-encoding sequence was labeled by nick-translation in the presence of digoxigenin-11-UTP and used as a probe for chromosomal in situ suppression hybridization as described (Lichter et al., 1990). 80 ng of the digoxigenin-labelled probe was combined with 3 μg of mouse Cot-1 DNA and 7 μg of salmon sperm DNA in 10 μl hybridization cocktail and hybridized to mouse chromosomes prepared from spleen cells of a female Balb/c animal according to established protocols (Sawyer et al., 1987). The hybridized probe was detected via rhodamine. Chromosomes were counterstained with 4′,6′-diamidino-2-phenylindole (DAPI). Images were recorded sequentially for the counterstain and rhodamine fluorescence by a CCD camera (Photometrics), aligned and electronically overlayed using the NIH image 1.41 software package.

Constructs for expression of green fluorescent protein, centrosomin A-GFP and centrosomin B-GFP

The nucleotide sequence encoding the (modified) green fluorescent protein (GFP) from *Aequorea victoria* (pGreen Lantern, Life Technologies) was PCR amplified (Clontech, advantage PCR kit, #k1906-1) and the ends were modified during PCR by addition of restriction sites: *Sal*I (5′ end) and *Hin*dIII (3′ end). The product was cloned into the *Sal*I/*Hin*dIII sites of the pBluescript KS + vector (Stratagene) to obtain the plasmid pBlue-GL. The coding sequences for centrosomin A and B were PCR amplified with cDNA-fitting primer pairs and restriction sites were added: *Kpn*I (5′ end) and *Sal*I (3′ end). Cloning into the pBlue-GL (*Kpn*I/*Sal*I) resulted in inframe fusions (pBlue-cent A-GL and pBlue cent B-GL). The GL sequence and the fused sequences cent A-GL and cent B-GL were released from the vector (*Kpn*I/*Sal*I) and cloned into the *Kpn*I/*Sal*I sites of the eukaryotic expression vector pcDNA3 (Invitrogen).

Microinjection of plasmids

Purified plasmids (Qiagen) were dissolved in 150 mM KCl, 10 mM phosphate buffer, pH 7.2 (0.5 μg/μl) and microinjected into Ptk cells using a Transjector (Eppendorf). Care was taken to inject exclusively into the cytoplasm. Living cells were inspected at the times indicated in Fig. 6 using the FITC-Filterset on an AXIOVERT 405 microscope (Zeiss). Usually 15-25% of the cells injected showed GFP-based fluorescence. Photographs were taken using the Fluar ×100 Objective on a Kodak Tmax ASA 400 film.

Biocomputing

The computer software package HUSAR (Heidelberg unix sequence analysis resources) supplied by S. Suhai, German Cancer Research Center, Heidelberg, Germany, was used for structure predictions, database searches, inspection of nucleotide and amino acid sequences. Especially the fast search program FASTA (Pearson and Lipman, 1988) was applied to search for similar sequences in nucleotide and in protein data bases. Putative post-translational modification sites and nuclear targeting sites were detected by means of the comparison tables of A. Bairoch (Medical Biochemistry Department, University of Geneva, Switzerland). For the prediction of secondary structures we used the programmes of Ptitsyn and Finkelstein (1983); Garnier et al. (1978); Mrazek and Kypr (1988); Lewin and Garnier (1988). Coiled-coil predictions were performed according to the method of Lupas et al. (1991).

RESULTS

cDNA clones based on the GP1 serum

Using the GP1 antiserum an initial and partial bona fide cDNA was detected in a murine RNA-based λgt11 cDNA library (Joswig et al., 1991). Further screening of a cDNA library prepared in the λgt10 vector with the radiolabelled insert of the initial cDNA released two types of cDNAs with 1.2 kbp and 2.2 kbp inserts. The 1.2 kbp sequence was found to represent a complete cDNA (accession no. X17373). The protein encoded by this shorter cDNA has been termed centrosomin A (Joswig et al., 1991). Complete sequencing of the 2.2 kbp inserts revealed that this larger cDNA does not reflect a cDNA with larger non-translated regions as suggested before (Joswig et al., 1991). It was found to represent a cDNA coding for another protein with a calculated molecular mass of 56 kDa, termed centrosomin B. The nucleotide sequence of the 2.2 kbp cDNA is available from databases under the accession no. X84651.

Sequence relations between centrosomin A and B

The nucleotide sequences of the 1.2 kbp cDNA (accession no. X17373) and that of the 2.2 kbp cDNA (accession no. X84651) are identical up to the nucleotide position 1,156. At this position the longer sequence comprises a G with the consequence that the reading frame remains open. In contrast, this base is missing in the centrosomin A-encoding cDNA sequence which results in the termination of the reading frame after four additional amino acid residues (Fig. 1).

The alignment of the amino acid sequences derived from the

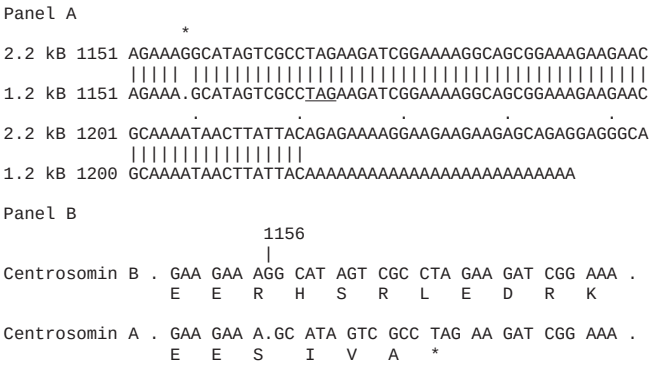

Fig. 1. Differences in the nucleotide sequences of cDNAs encoding centrosomin A and B. (A) The nucleotide sequences of the 1.2 kbp cDNA encoding centrosomin A and the 2.2 kbp cDNA encoding centrosomin B are identical up to the nucleotide position 1,156* where one base (G) is missing in the 1.2 kbp cDNA. After this position the sequences are again identical up to position 1,217 where the poly(A) tail starts in the 1.2 kbp cDNA while the sequence of the 2.2 kbp cDNA continues. (B) Due to the missing G at position 1,156* of the centrosomin A-encoding cDNA the reading frame remains only open for another four amino acid residues (SIVA). In contrast, in the 2.2 kbp cDNA encoding centrosomin B the reading frame remains open for another 175 amino acid residues.

two cDNAs shows that the amino-terminal sections of centrosomin A and B are identical over 272 amino acid residues. The carboxy-terminal section of centrosomin B comprises an additional 175 amino acid residues including three overlapping nuclear location signals. In contrast, the additional carboxy terminus of centrosomin A is composed of only a short peptide sequence consisting of four amino acid residues (Fig. 2).

The two protein sequences have in common one N-glycosylation site and multiple phosphorylation sites, e.g. motifs specific for phosphorylation by PKC, CK2 and by cAMP-dependent kinase. Centrosomin B has two additional sites for putative protein phosphorylation by tyrosine kinase. The large number of possible sites for putative protein phosphorylation suggests that at least some of these sites are either permanently or conditionally phosphorylated in vivo.

Other related sequences

At the time of submission no extended similarities with other nucleotide or amino acid sequences were found in data bases which could point to a structural or functional relationship between centrosomins and other proteins. Only blocks consisting of up to five amino acid residues were found to be comprised in the centrosomin sequences and in the sequences of the nuclear nuclear lamina proteins lamin C and B (Joswig et al., 1991). Recent data base releases contain a partial sequence of a human cDNA (EST clone BL-29-9, HS0028) which shows 84.8% identity with the centrosomin B-encoding cDNA sequence suggesting that this EST clone represents the human analogue of the murine centrosomin B cDNA. An unpublished sequence (MM14172, 5,119 bp) may reflect another posttranscriptional variant because it overlaps significantly with the nucleotide sequences encoding centrosomin A and B. The data base description of this sequence points to a ‘162 kDa component of a multi-protein complex phosphorylated by Src’. However, this transcript seems to be either rare or cell type specific because no cDNA was detected during our screening experiments that could represent this transcript. Moreover, this

Fig. 2. Alignment of the amino acid sequences encoded by the 1.2 kbp cDNA (centrosomin A) and the 2.2 kbp cDNA (centrosomin B). The numbers indicate the heptad frames. Asterisks indicate discontinuities in these frames. The position of the nuclear targeting signals in the centrosomin B sequence is underscored.

large transcript could not be detected on northern blots of RNA from the Ehrlich ascites tumor probed with the radiolabelled centrosomin B-encoding sequence (not shown). It should be noted that the *Drosophila melanogaster cnn* gene which has been termed centrosomin (Li and Kaufman, 1996) is neither related to the centrosomin A sequence published previously (Joswig et al., 1991) nor to the centrosomin B sequence published here.

Secondary structure predictions for centrosomin A and B

Different secondary structure prediction programs listed in the Materials and Methods propose α-helical regions for centrosomin A and B. The α-helical regions contain heptads in which positions two and five are frequently occupied by non-polar amino acid residues. These heptad frames are indicated in Fig. 2 by numbers. Accordingly, it can be suggested that centrosomin A and B are able to form coiled-coil structures. However, it is obvious that the α-helical domains are interrupted by sequences which do not tend to adopt the α-helical conformation. Proteins with such interrupted coiled-coil domains, missing significant long-range periodicity in the distribution of either acidic or basic residues, are considered to form loose and semi-regular meshworks rather than closely packed filamentous structures (Parry, 1994).

Centrosomin A and B mRNAs are transcripts of one gene

Chromosomal mapping by fluorescence in situ hybridization

(FISH) shows one locus which hybridizes with the complete centrosomin B-encoding cDNA probe (Fig. 3). Microscopic evaluation revealed specific fluorescent signals on mouse chromosome 19 in the distal part of band D. Of 28 metaphase cells examined, 15 showed signals on one of the two chromosome 19 homologs whereas 10 revealed signals on both homologs. Since no fluorescent signal was found consistently on another chromosomal region, localization of the gene for centrosomin A/B is given as 19D. Chromosome assignment was also confirmed by dual color FISH with a simultaneously hybridized cosmid probe localized on 19D (D19cw42; Boyle et al., 1992).

The in situ hybridization results are supported by PCR amplification of a single genomic DNA fragment by primer pairs fitting to regions where the coding sequences for centrosomin A and for centrosomin B are identical. Application of different primer pairs released only one single fragment. For example, one of the PCR products generated by primer pairs including nucleotide positions 800-1,194 of the centrosomin B cDNA is shown in Fig. 4. The size and the nucleotide sequence of the genomic fragment indicates that the centrosomin B-encoding sequence between positions 800-1,194 is interrupted on the gene level by three introns (not shown).

Since the two transcripts derive from one gene it is obvious that centrosomin A and B mRNAs are the products of either alternative splicing of the same pre-mRNA or alternative polyadenylation. However, the difference in a single base at position 1,156 between the genomic sequence and the centrosomin B-encoding sequence on one side (Fig. 5) and the centrosomin A encoding sequence on the other side (Fig. 1) cannot be explained by alternative splicing and alternative polyadenylation. Since the genomic sequence contains this base it is obvious that the formation of centrosomin A mRNA involves a template G deletion. Template G deletion is a documented RNA editing process (Jaques et al., 1994) and RNA editing is a posttranscriptional process that has been identified in an increasing number of eukaryotic organisms (Adler and Hajduk, 1994). Accordingly, a combination of alternative splicing and RNA editing resulting in template G deletion could best explain the generation of centrosomin A mRNA.

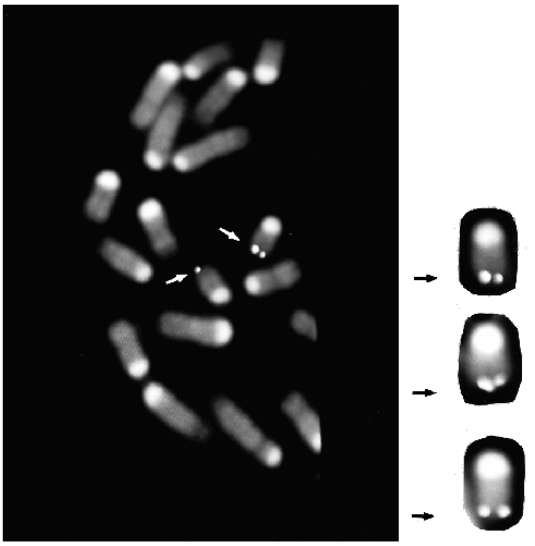

Fig. 3. Chromosomal mapping of the mouse cDNA probe encoding centrosomin B by FISH to murine metaphase chromosomes. The hybridized probe was detected using rhodamine (arrows) and chromosomes were counterstained with DAPI. The left panel shows a section of a hybridized metaphase cell with signals on both chromosome 19 homologs, while the right panel shows three chromosome 19 homologs at higher magnification to allow the assessment of the sub-chromosomal localization.

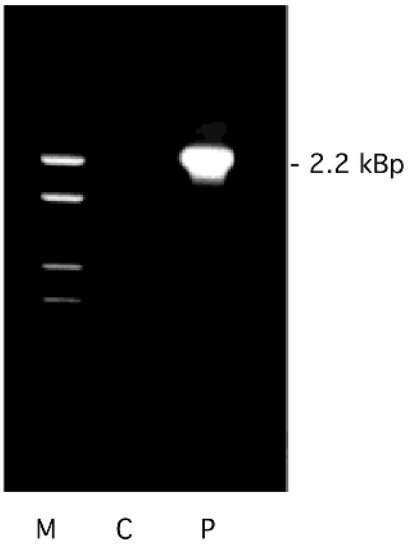

Fig. 4. Fragment of the gene encoding centrosomin A and B. Murine genomic DNA was used as template and amplified by PCR with primer pairs fitting to identical sequences of the centrosomin A and B-encoding cDNAs. The amplified DNA fragment shown (lane P) corresponds to nucleotide positions 800-1,194 of centrosomin B cDNA and with nucleotide position 800-1,193 of centrosomin A cDNA. It should be noted that only one significant DNA fragment was amplified which indicates that the transcripts encoding centrosomin A and B derive from the same gene. Lane M shows DNA markers, lane C represents a control (PCR reaction without template).

Localization of centrosomin A and B in Ptk cells

The centrosomin B sequence comprises a nuclear location signal (Fig. 2). Accordingly, centrosomin B is expectedly a nuclear

```
     Intron
     AGTGTAAAACTGTTATTCATGTTTATCTTAAAAAATGGCCTGTGGGGTAA genomic

        Intron end |
     CTGTTGAATTTGCAGGAAAAACTGAAACAGTTTGAAGAGCGCTTAGCAGA genomic
                   ||||||||||||||||||||||||||||||||||||
1101 CAGTCTGTCTACGAGGAAAAACTGAAACAGTTTGAAGAGCGCTTAGCAGA cDNA
               .         .         .         .         .
     AGAAAGGCATAGTCGCCTAGAAGATCGGAAAAGGCAGCGGAAAG...... genomic
     ||||||||||||||||||||||||||||||||||||||||||||
1151 AGAAAGGCATAGTCGCCTAGAAGATCGGAAAAGGCAGCGGAAAGAAGAAC cDNA
          *                    < Reverse primer
        1156                                  1194
```

Fig. 5. Sequence identity of the genomic sequence and the sequence of centrosomin B cDNA at nucleotide position 1,156. The PCR product shown in Fig. 4 was recloned in the pBluescript vector, sequenced and aligned with the centrosomin B-encoding sequence. The figure shows a section of this alignment which includes the exon of interest and a portion of the upstream intron. Numbering refers to the nucleotide sequence of the centrosomin B-encoding cDNA. At nucleotide position 1,156 the base G is deleted in the centrosomin A-encoding sequence (Fig. 1). The genomic nucleotide sequence shown in this figure and a longer sequence of the gene encoding centrosomin A and B are available from data bases under the accession no. Y13494.

protein. However, the exclusively nuclear localization of centrosomin A, e.g. by means of immuno histochemical techniques, meets with difficulties because the two proteins share a long stretch of sequence identity (Fig. 2). Thus, the antiserum to the recombinantly expressed centrosomin A does not distinguish between centrosomin A and B, and it immunostains centrosomes and nuclei (not shown). Moreover, antibodies to the carboxy-terminal section of centrosomin B would be non-significant because of extended sequence repeats present in this section and in centrosomin A (Fig. 2). Finally, the monoclonal antibody immunostains only the centrosomes of Ptk cells, and it does not show immuno reaction with the murine centrosomins.

Consequently, an alternative technology had to be applied for the differential localization of the two proteins. Fusion proteins between a functional protein and GFP are expected to indicate the intracellular transport and the localization of the functional protein with high probability (Chalfie et al., 1994). Using this system it was possible to differentiate between the location of centrosomin A and B at the cellular level (Fig. 6). Transfection with the expression vector containing only the GFP-encoding sequence resulted in expressed GFP protein uniformly distributed in all cellular compartments. Following transfection of the centrosomin A-GFP fusion construct the protein becomes initially located in the centrosome. Only after prolonged expression does the centrosome become obviously 'overloaded' and some of the fusion protein is deposited in cytoplasmic vesicles. In contrast, the expressed centrosomin B-GFP fusion protein is initially transported into nuclei. Accordingly, centrosomin A is an exclusive centrosomal protein while centrosomin B is a nuclear protein.

DISCUSSION

Centrosome duplication is an absolute prerequisite for cell division (Maniotis and Schliwa, 1991). However, the molecular basis for the generation of cell polarity and the centrosome's cross-connection to other processes regulating the cell cycle is still unknown. Findings indicating a brief period of centrosome 'maturation' during which the two organelles are not identical (Rothbarth et al., 1993) and the observation that the M-phase controlling protein kinase $p34^{cdc2}$ is transiently associated with the centrosome (Bailly et al., 1989) mark only preliminary steps in the elucidation of the molecular mechanisms of the centrosome's function.

Centrosomes cannot be formed de novo. An existing centrosome has to serve as a 'template' for the assembly of the proteins composing the daughter centrosomes (Maniotis and Schliwa, 1991). However, most of the centrosome-associated proteins identified so far are only transiently located in the centrosome

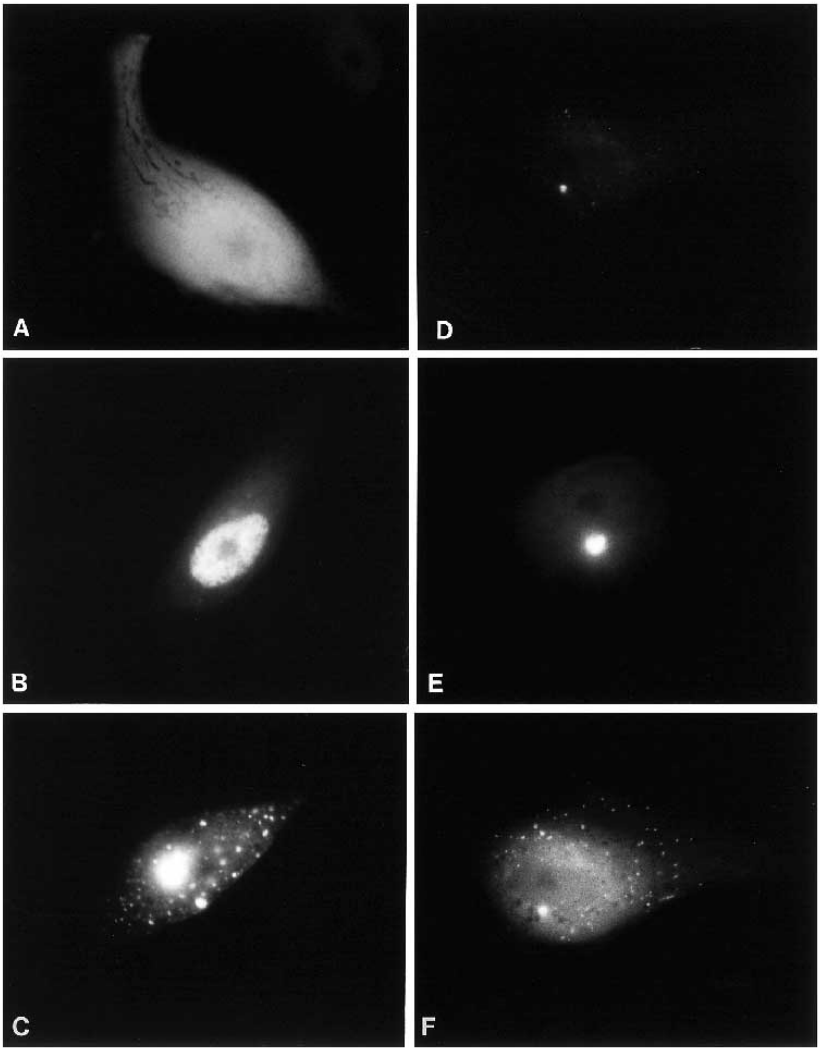

Fig. 6. Intracellular transport and localization of chimeric proteins between centrosomin A/B and GFP in living Ptk cells. (A) Following injection of the control plasmid encoding only GFP the fluorescent protein is uniformly distributed in the cell, in the nucleus as well as in the cytoplasm. Only mitochondria appear to be somewhat more weakly stained. (B and C) Following injection of the plasmid encoding the fusion protein consisting of centrosomin B and GFP only the cell nucleus is stained. Eight hours post injection the strong and exclusively nuclear fluorescence is associated with a fibrogranular network (B). Occasionally the nuclear fluorescence forms a punctuate pattern (not shown). Only about 36 hours post injection significant fluorescence is observed in the cytoplasm (C) which may reflect either overloading of the nuclear location sites or protein degradation. (D to F) Following injection of the plasmid encoding the fusion protein consisting of centrosomin A and GFP the fluorescence is first concentrated in a small dot in the nuclear periphery which is the site of the centrosome in Ptk cells (D). Eight hours post transfection the site of the centrosome becomes obviously saturated (E) and after 36 hours fluorescent protein accumulates in cytoplasmic vesicles (F) which may reflect either overloading of the centrosome or protein degradation.

and several other centrosome-associated proteins have other functions in the cell. Consequently, it is unlikely that transiently centrosome-associated proteins are involved in a basic structure which may serve as a template for centrosome duplication. On the other hand, proteins that are permanently and exclusively located in the centrosome appear to be more relevant with respect to the constitution of the basic centrosome structure. Since such proteins are rare they deserve special attention.

A candidate protein potentially involved in an intrinsic centrosomal sub-structure was first detected by a monoclonal antibody immunoreacting with an antigen permanently and exclusively present in centrosomes of PtK cells (Joswig and Petzelt, 1990). However, the molecular cloning approach released two cDNAs encoding closely related proteins. Because of the complex sequence relations between these proteins the intrinsic centrosomal protein could not be reliably identified by conventional techniques. However, by means of the new reporter molecule GFP which allows us to follow gene expression and protein localization in living cells it was now possible to identify centrosomin A as a centrosomal protein while the closely related centrosomin B has obviously its main function in the cell nucleus. The nuclear location of centrosomin B is in agreement with the nuclear targeting signal present in its amino acid sequence.

The sequence relationship between centrosomin A and B parallels that between the nuclear lamina proteins lamin A and C which are identical in their amino-terminal sections. In the latter case it has been well documented that the two mRNAs derive from the same pre-mRNA by alternative splicing (Lin and Worman, 1993). Our results indicate that centrosomin A and B are also the products of only one gene which is located on chromosome 19D. Accordingly, centrosomin A and B mRNAs could derive from the same pre-mRNA by an alternative splicing process while alternative polyadenylation cannot be ruled out. However, the posttranscriptional processes cannot explain the deletion of a single base in the centrosomin A-encoding mRNA. This base deletion is best explained by RNA editing which is a posttranscriptional process that has been identified in a number of eukaryotic organisms (Adler and Hajduk, 1994; Jacques et al., 1994).

Centrosomin A and B can now be considered to belong to the increasing group of protein families whose members reside either in the centrosome or in the cell nucleus. The involvement of closely related proteins in nuclear and centrosomal functions could reflect the origin of cell polarity. In the bacterial cell the polarity is determined by the bacterial membrane. As soon as the nucleus developed (the Dinoflagellates may serve as an example) the bacterial membrane was replaced by the nuclear membrane as the coordinator of cell polarity and the distributor of DNA. In mammalian cells, the polarizing activity became translocated into the cytoplasm where it is organized in the centrosome which is generally located in close vicinity to the outer nuclear membrane. From this history it could be suggested that the cell polarity was originally a nuclear function and that the centrosome derived from a nuclear ancestor structure. Accordingly it is not surprising that we find an increasing number of centrosomal proteins or centrosome-associated proteins which are genetically related to nuclear proteins.

This work was supported by a grant of the European Communities (Petzelt, Contract SC1*.CT91.0640). A.M. and P.L. are grateful to Prof. Dr H. zur Hausen for support.

REFERENCES

Adler, B. K. and Hajduk, S. L. (1994). Mechanisms and origins of RNA editing. *Curr. Opin. Genet. Dev.* **4**, 316-322.

Boyle, A. L., Feltquite, D. M., Dracopoli, N. C., Housman, D. E. and Ward, D. C. (1992). Rapid physical mapping of cloned DNA on banded mouse chromosomes by fluorescence in situ hybridization. *Genomics* **12**, 106-115.

Bailly, E., Dorée, M., Nurse, P. and Bornens, M. (1989). p34cdc2 is located in both nucleus and cytoplasm; part is centrosomally associated at G2/M and enters vesicles at anaphase. *EMBO J.* **8**, 3985-3995.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**, 802-805.

Garnier, J., Osguthorpe, D. J. and Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* **120**, 97-120.

Horio, T., Uzawa, S., Jung, M. K., Oakley, B. R., Tanaka, K. and Yanagida, M. (1991). The fission yeast gamma-tubulin is essential for mitosis and is localized at microtubule organizing centers. *J. Cell Sci.* **99**, 693-700.

Jacques, J. P., Hausmann, S. and Kolakofsky, D. (1994). Paramyxovirus mRNA editing leads to G deletions as well as insertions. *EMBO J.* **13**, 5496-5503.

Joswig, G. and Petzelt, C. (1990). The centrosomal cycle. Its visualization in PtK cells by a monoclonal antibody to a 32 kDa protein. *Cell Motil. Cytoskel.* **15**, 181-192.

Joswig, G., Petzelt, C. and Werner, D. (1991). Murine cDNAs coding for the centrosomal antigen centrosomin A. *J. Cell Sci.* **98**, 37-44.

Joshi, H. C., Palacios, M. J., McNamara, L. and Cleveland, D. W. (1992). Gamma-tubulin is a centrosomal protein required for cell cycle-dependent microtubule nucleation. *Nature* **356**, 80-83.

Kalt, A. and Schliwa, M. (1993). Molecular components of the centrosome. *Trends Cell Biol.* **3**, 122-128.

Lewin, J. M. and Garnier, J. (1988). Improvements in a secondary structure prediction method based on a search for local sequence homologies and its use as a model building tool. *Biochim. Biophys. Acta* **955**, 283-295.

Lichter, P., Tang C. C., Call, K., Hermanson, G., Evans, G. A., Housman, D. and Ward, D. C. (1990). High resolution mapping of human chromosome 11 by in situ hybridization with cosmid clones. *Science* **247**, 64-69.

Li, K. and Kaufman, T. C. (1996). The homeotic target gene centrosomin encodes an essential centrosomal component. *Cell* **85**, 585-596.

Lin, F. and Worman, H. J. (1993). Structural organization of the human gene encoding nuclear lamin A and nuclear lamin C. *J. Biol. Chem.* **268**, 16321-16326.

Lupas, A., van Dyke, M. and Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* **252**, 1162-1164.

Maniotis, A. and Schliwa, M. (1991). Microsurgical removal of centrosomes blocks cell reproduction and centriole generation in BSC-1 cells. *Cell* **67**, 495-504.

Mrazek, J. and Kypr, J. (1988). Computer program Jamsek combining statistical and stereochemical rules for the prediction of protein secondary structure. *Comput. Appl. Biosci.* **4**, 297-302.

Parry, A. D. (1994). NuMA/centrophilin: sequence analysis of the coiled-coil rod domain. *Biophys. J.* **67**, 1203-1206.

Pearson, W. R. and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Nat. Acad. Sci. USA* **85**, 2444-2448.

Ptitsyn, O. B. and Finkelstein, A. V. (1983). Theory of protein secondary structure and algorithm of its prediction. *Biopolymers* **22**, 15-25.

Rothbarth, K., Petzelt, C., Lu, X., Todorov, I., Joswig, G., Pepperkok, R., Ansorge, W. and Werner, D. (1993). cDNA-derived molecular characteristics and antibodies to a new centrosome-associated and G_2/M phase-prevalent protein. *J. Cell Sci.* **104**, 19-30.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci. USA* **74**, 5463-5467.

Sawyer J. R., Moore, M. M. and Hozier, J. C. (1987). High resolution G-banded chromosomes of the mouse. *Chromosoma* **95**, 350-358

Stearns, T., Evans, L. and Kirschner, M. (1991). Gamma-tubulin is a highly conserved component of the centrosome. *Cell* **65**, 825-836.

Tang, T. K., Tang, C. J., Chao, Y. J. and Wu, C. W. (1994). Nuclear mitotic apparatus protein (NuMA): spindle association, nuclear targeting and differential subcellular localization of various NuMA isoforms. *J. Cell Sci.* **107**, 1389-1402.

Zeng, C., He, D. and Brinkley, B. R. (1994). Localization of NuMA protein isoforms in the nuclear matrix of mammalian cells. *Cell Motil. Cytoskel.* **29**, 167-176.

(Received 9 July 1997 – Accepted 20 August 1997)