Structure, expression and chromosomal localization of human p80-coilin gene

Edward K.L.Chan*, Saeko Takano, Luis E.C.Andrade[+], John C.Hamel and A.Gregory Matera[1]
W.M.Keck Autoimmune Disease Center and DNA Core Laboratory for Structural Analysis, Department of Molecular and Experimental Medicine, The Scripps Research Institute, 10666 N. Torrey Pines Road, La Jolla, CA 92037 and [1]Department of Genetics and Center for Human Genetics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-4955, USA

ABSTRACT

Coiled bodies (CBs) are non-capsular nuclear bodies with a diameter of $0.3-1\mu m$ and appear to be composed of coiled fibrils. Human autoantibodies to CBs recognize an 80-kD nuclear protein highly enriched in CBs, and this protein has been named p80-coilin. CBs are known to assemble and disassemble during the cell cycle, with the highest number of CBs occurring at mid to late G_1 where p80-coilin is assembled into several small nuclear body-like structures. In S and G_2 phases, CBs become larger and their number decreases and often they are undetectable during mitosis. Using a human autoantibody as a probe for expression cloning, we initially isolated a partial cDNA encoding p80-coilin. In this report, the 5′ end of the complete cDNA for p80-coilin was obtained using the 5′-RACE (rapid amplification of cDNA ends) methodology. The size of the reconstructed full-length cDNA corresponds to the 2.7-kb mRNA detected in Northern blot analysis. The complete p80-coilin protein consists of 576 amino acids with a predicted molecular mass of 62,608. A putative p80-coilin pseudogene was also detected during the rescreening of p80-coilin cDNA. To confirm the validity of the cDNA sequence, three overlapping genomic DNA clones representing the human p80-coilin gene were selected for further analysis. The complete gene for p80-coilin contains 7 exons spanning ~25kb. Sequence analysis of exons 1 and 2 in genomic DNA clones confirmed the accuracy of the 5′ cDNA sequence derived from the 5′-RACE procedure. Furthermore, the human p80-coilin gene was localized to chromosome 17q22−23 by fluorescence *in situ* hybridization.

INTRODUCTION

An increasing number of nuclear bodies or subnuclear domains have been recently reported (for review, see (1−3)). These structures include the large number of so-called 'poly(A) transcript' domains (4) that are probably composed of interchromatin granule clusters and perichromatin fibrils (3), the fewer but larger round bodies recently proposed as 'oncogenic domains' (5−7), and the large, phase-contrast visible structures described as PIKA [(8), polymorphic interphase karyosomal association] domain. Approximately 5 years ago, we 'rediscovered' a nuclear organelle known as the coiled body (CB), using human autoantibodies from patients with various rheumatic conditions (9−11). The nuclear CB was initially described as an accessory body to the nucleolus in light microscopy by the Spanish cytologist Ramon y Cajal using the silver staining method. In the late sixties, the CB was characterized by electron microscopy as a morphological structure distinct from other nuclear bodies (12,13). CBs are non-capsular structures with a diameter of 0.3−1 micron and are loosely packed with threads that have the appearance of coils. Human autoimmune antibodies to CBs recognize an 80-kD nuclear protein called p80-coilin. Affinity purified antibodies to the 80-kD protein show highly specific immunostaining (9,10), suggesting that p80-coilin is an integral component of the coiled body.

The dynamics of p80-coilin and CBs in the cell cycle and during the transition from quiescence to proliferation was reported previously (14). Most CBs disassemble prior to or during mitosis. After cell division, the reassembly of CBs occurs during G_1 phase and is preceded by the reformation of nucleoli (14). The highest number of CBs is seen at mid to late G_1, where p80-coilin is assembled into several small CBs. In S and G_2 phases, CBs become larger in size and fewer in number. It was postulated that small CBs in G_1 either fuse or reassemble into the larger CBs seen at later stages (14). Although CBs are not

---

*To whom correspondence should be addressed

[+]Present address: Division of Rheumatology, Escola Paulista de Medicina, Rua Botucatu 740, Sao Paulo, SP 04023, Brazil

observed during mitosis, the overall level of p80-coilin is about the same as in interphase, and suggests that p80-coilin is present in a soluble pool in mitosis and early G_1 and in an insoluble CB-associated compartment in later stages of interphase. We have also observed an increase in the number of CBs and the total amount of p80-coilin in proliferating versus quiescent cells (14). Spector *et al.* (15) have reported that when a REF-52 cell line is transformed by adenovirus, the number of cells with detectable CBs increases from 24% to 99% in the transformed line.

Using the human antibody as a probe for a cDNA expression library, we previously obtained a partial cDNA encoding p80-coilin (9). Subsequent attempts to obtain the full-length cDNA from several cDNA libraries using DNA hybridization uncovered a putative p80-coilin pseudogene, but failed to generate a full-length clone. In this study, the *bona fide* 5' end cDNA for p80-coilin was constructed using the 5'-RACE procedure. The sequence of the cDNA was confirmed using that of genomic clones spanning the gene, whose chromosomal localization was also determined using fluorescence *in situ* hybridization (FISH).

MATERIALS AND METHODS

Antibodies

Human anti-p80-coilin sera Op and Hu were described in our earlier study (9). Rabbit serum R288 was produced to the C-terminal 14-kD fragment of recombinant human p80-coilin as described (14). A 15-mer synthetic peptide (CGIESPSNTSSTEPA) corresponding to the C-terminus of human p80-coilin was injected into a New Zealand White rabbit #508 that eventually produced anti-p80-coilin antibodies known as serum R508. Prior to immunization, the synthetic peptide was conjugated to keyhole limpet haemocyanin (KLH) via the N-terminal Cys residue using *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester (MBS, Pierce Chemical Co., Rockford, IL) as coupling reagent (16). One mg of the peptide−KLH conjugate was mixed with Complete Freund's adjuvant and given via subcutaneous injections. Two booster injections (0.5 mg) were mixed in incomplete Freund's adjuvant and were also administered by subcutaneous injections at monthly intervals. Sera from rabbit R508 were collected before immunization and 10 days after the last booster injection and their reactivities to p80-coilin were analyzed by immunofluorescence and immunoprecipitation.

Indirect immunofluorescence

HEp-2 cell slides (Bion) were used for immunofluorescence microscopy. Human and rabbit anti-p80-coilin sera were diluted at 1:100−200 in PBS and incubated for 45−60 minutes. Secondary detecting reagents were fluorescein-labeled goat anti-human immunoglobulin or rhodamine-labeled goat anti-rabbit IgG antibodies (Caltag, South San Francisco, CA). Secondary antibodies were also incubated for 45−60 minutes and after each antibody incubation step, coverslips were washed twice in PBS for 10 min. In double labeling experiments human and rabbit sera were mixed and incubated with the cell substrate simultaneously.

cDNA cloning

In order to obtain the 5' sequence of p80-coilin mRNA, 5'-RACE was performed as per the instructions of the manufacturer (BRL). Total HEp-2 cell RNA and the oligonucleotide primer 5'-CGGAGACTTGGGATTCTTAGCC-3' was used in the

reverse transcription which is the first step in the 5'-RACE protocol. PCR products obtained after 30 cycles of amplification were subcloned into pCRII vector using blue/white selection (TA cloning kit, Invitrogen). White colonies were screened by PCR for the presence of cDNA insert and selected clones were analyzed by restriction mapping and sequencing. Nucleotide sequence was determined using dye primer cycle sequencing and Model 373A DNA sequencer from Applied Biosystems Inc. (ABI, Foster City, CA). Oligonucleotide primers were synthesized with a Model 394 DNA synthesizer (ABI). DNA sequences were determined in both strands and compiled using the alignment program SeqEd (ABI). DNA and protein sequences were analyzed by the Genetics Computer Group Sequence Analysis Software Package version 7.2 for UNIX computers (17). Alignment of protein sequences was achieved with the GAP program that employed the algorithm of Needleman and Wunsch (18).

Genomic DNA cloning

p80-coilin cDNA insert from clone pJEL1 (9) was labeled by the method of Feinberg and Vogelstein (19) and used for DNA hybridization screening of a human placenta genomic DNA library constructed with the λFIX II vector (Stratagene, La Jolla, CA, Cat # 946205). Three positive clones (λ80g24, λ80g11, λ80g20) were plaque-purified and examined by restriction enzyme analysis and Southern blotting. Restriction fragments derived from *Hind*III or *Eco*RI digestions were identified and subcloned into pBluescript SK (Stratagene). Exons were identified by direct DNA sequence determinations using either purified λFIX II phage DNA or plasmid containing *Hind*III/*Eco*RI fragments subcloned from the λFIX II phage DNA.

Northern blotting

Total cellular RNA was isolated from HEp-2 cells (20). 10μg of total cellular RNA or 2μg of poly(A)$^+$ RNA from different human tissues (Clontech, Palo Alto, CA) was fractionated by electrophoresis through 1% agarose containing 2.2M formaldehyde and transferred to nitrocellulose membrane. The RNA blot was hybridized with [^{32}P]-labeled cDNA fragment and blots were washed at 51°C in 0.1×SSC and 0.1% SDS.

Immunoprecipitation and Western blotting

Plasmid DNA was used as substrate in *in vitro* transcription using SP6 RNA polymerase and the resulting RNA was translated *in vitro* using a rabbit reticulocyte lysate (Promega, Madison, WI) in the presence of [^{35}S]-methionine (Trans ^{35}S-label, 70% methionine and 15% cysteine, ICN Biochemicals). Immuno-precipitation was carried out using Protein A-bound Sepharose CL-4B (Pharmacia, Uppsala, Sweden), essentially as described (21). Immunoprecipitates were analyzed by 10% gel SDS−PAGE followed by autoradiography. Western blotting was performed using whole cell extracts from MOLT-4 cells, nuclear extracts from HeLa cells, or *in vitro* translation products as described (21).

Fluorescence *in situ* hybridization

The complete clone λ80g24 was used as probe for fluorescence *in situ* hybridization (FISH). Approximately 2 μg of the phage DNA were nick-translated with biotin-11-dUTP and hybridized to human metaphase chromosomes, essentially as described (22). An excess of human Cot 1 DNA (BRL) was used instead of total human DNA as a repetitive sequence competitor. After
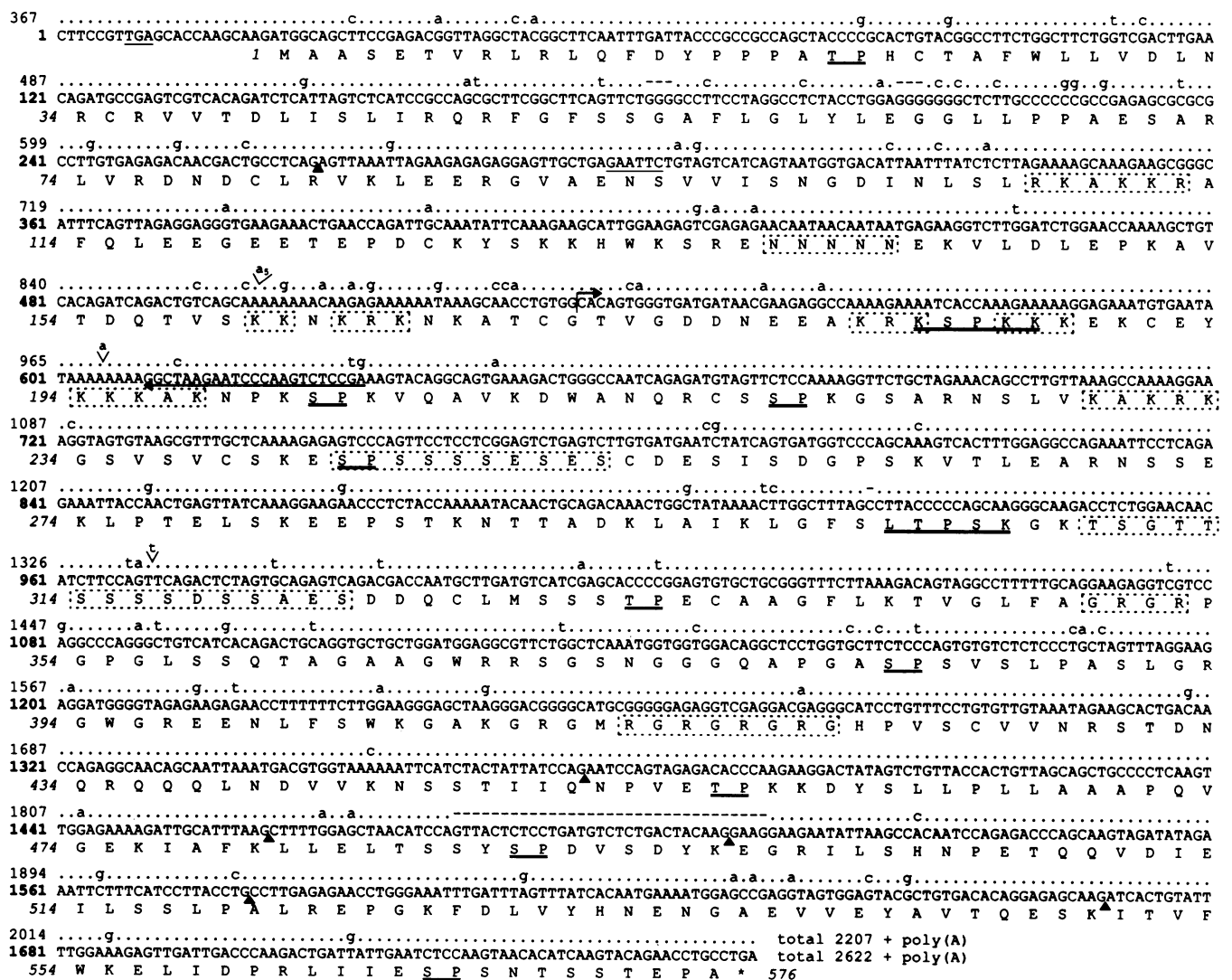
hybridization, the slides were washed in 50% formamide, 2×SSC at 37°C, followed by washes in 0.2×SSC, 55°C. The DNA counterstain DAPI was used to generate a G/Q-banding pattern on the chromosomes. Digital imaging microscopy using a cooled charge-coupled device camera was performed as described (23). Image registration was accomplished using highly plane-parallel bandpass filters which limit image displacement ± 1 pixel when changing filters (24).

RESULTS
Putative pseudogene of p80-coilin
In our attempts to obtain the 5' sequence of full-length p80-coilin cDNA, we have screened multiple cDNA libraries using the 5'-most probe from the longest cDNA clone (JEL1) reported

earlier (9). Clone p80-1 was obtained from the screening of a HepG2 cell cDNA library and was thought to be an excellent candidate for the full-length cDNA of p80-coilin. However, no protein was detected when we performed *in vitro* translation using the p80-1 transcript. Sequence analysis of p80-1 revealed that it contains multiple stop codons and no methionine start site (figure 1). When the p80-1 sequence was compared to full-length p80-coilin (see below, figure 1), multiple substitutions, deletions, and insertions were detected throughout the p80-1 sequence. Interestingly, p80-1 contains a poly(A) tail and, therefore, might represent a mRNA from a putative pseudogene of p80-coilin. The possible expression of this pseudogene was not examined further, except to note that it was not detected in other cDNA library screenings.

```
367    ............................................c.........a.......c.a.........................................g.........g...................t..c.......
  1 CTTCCGTTGAGCACCAAGCAAGATGGCAGCTTCCGAGACGGTTAGGCTACGGCTTCAATTTGATTACCCGCCGCCAGCTACCCCGCACTGTACGGCCTTCTGGCTTCTGGTCGACTTGAA
     1  M  A  A  S  E  T  V  R  L  R  L  Q  F  D  Y  P  P  P  A  T  P  H  C  T  A  F  W  L  L  V  D  L  N

487    ......................g..............................at........t....---...c..........c..........a.---.c.c...c.......gg..g........t...
121 CAGATGCCGAGTCGTCACAGATCTCATTAGTCTCATCCGCCAGCGCTTCGGCTTCAGTTCTGGGGCCTTCCTAGGCCTCTACCTGGAGGGGGGGCTCTTGCCCCCCGCCGAGAGCGCGCG
 34 R  C  R  V  V  T  D  L  I  S  L  I  R  Q  R  F  G  F  S  S  G  A  F  L  G  L  Y  L  E  G  G  L  L  P  P  A  E  S  A  R

599  ...g........g.......c.............g.....................................a.g..................c....c....a.............
241 CCTTGTGAGAGACAACGACTGCCTCAGAGTTAAATTAGAAGAGAGAGGAGTTGCTGAGAATTCTGTAGTCATCAGTAATGGTGACATTAATTTATCTCTTAGAAAAGCAAAGAAGCGGGC
 74 L  V  R  D  N  D  C  L  R  V  K  L  E  E  R  G  V  A  E  N  S  V  V  I  S  N  G  D  I  N  L  S  L  [R  K  A  K  K  R] A

719  ....................a...........................a....................g.a...a...........................t...............
361 ATTTCAGTTAGAGGAGGGTGAAGAAACTGAACCAGATTGCAAATATTCAAAGAAGCATTGGAAGAGTCGAGAGAACAATAACAATAATGAGAAGGTCTTGGATCTGGAACCAAAAGCTGT
114 F  Q  L  E  E  G  E  E  T  E  P  D  C  K  Y  S  K  K  H  W  K  S  R  E [N  N  N  N  N] E  K  V  L  D  L  E  P  K  A  V

840  ..............c.....c.g...a..a.g......g.....cca.......ca.........a.....a.............
481 CACAGATCGAGCTGTCAGCAAAAAAAACAAGAGAAAAAATAAAGCAACCTGTGCACAGTGGGTGATGATAACGAAGAGGCCAAAAGAAAATCACCAAAGAAAAAGGAGAAATGTGAATA
154 T  D  Q  T  V  S [K  K] N [K  R  K] N  K  A  T  C  G  T  V  G  D  D  N  E  E  A [K  R  K  S  P  K  K  K] E  K  C  E  Y

965  ....V........c...............tg..............a.................
601 TAAAAAAAAGGCTAAGAATCCCAAGTCTCCGAAAGTACAGGCAGTGAAAGACTGGGCCAATCAGAGATGTAGTTCTCCAAAAGGTTCTGCTAGAAACAGCCTTGTTAAAGCCAAAAGGAA
194 [K  K  K  A  K] N  P  K  S  P  K  V  Q  A  V  K  D  W  A  N  Q  R  C  S  S  P  K  G  S  A  R  N  S  L  V [K  A  K  R  K]

1087 .c................................................................cg.............................c...................
721 AGGTAGTGTAAGCGTTTGCTCAAAAGAGAGTCCCAGTTCCTCCTCGGAGTCTGAGTCTTGTGATGAATCTATCAGTGATGGTCCCAGCAAAGTCACTTTGGAGGCCAGAAATTCCTCAGA
234 G  S  V  S  V  C  S  K  E [S  P  S  S  S  E  S  E  S] C  D  E  S  I  S  D  G  P  S  K  V  T  L  E  A  R  N  S  S  E

1207 ..........g...........g.............................................g......tc......-..................................
841 GAAATTACCAACTGAGTTATCAAAGGAAGAACCCTCTACCAAAAATACAACTGCAGACAAACTGGCTATAAAACTTGGCTTTAGCCTTACCCCCAGCAAGGGCAAGACCTCTGGAACAAC
274 K  L  P  T  E  L  S  K  E  E  P  S  T  K  N  T  T  A  D  K  L  A  I  K  L  G  F  S  L  T  P  S  K  G  K [T  S  G  T  T]

1326 ........ta.V...........t...........t...................a.......t.......................................t....
961 ATCTTCCAGTTCAGACTCTAGTGCAGAGTCAGACGACCAATGCTTGATGTCATCGAGCACCCCGGAGTGTGCTGCGGGTTTCTTAAAGACAGTAGGCCTTTTTGCAGGAAGAGGTCGTCC
314 [S  S  S  S  D  S  S  A  E  S] D  D  Q  C  L  M  S  S  S  T  P  E  C  A  A  G  F  L  K  T  V  G  L  F  A [G  R  G  R] P

1447 g.......a.t......g........t...................t..........c..........c..c...t.................ca.c...........
1081 AGGCCCAGGGCTGTCATCACAGACTGCAGGTGCTGCTGGATGGAGGCGTTCTGGCTCAAATGGTGGTGGACAGGCTCCTGGTGCTTCTCCCAGTGTGTCTCTCCCTGCTAGTTTAGGAAG
354 G  P  G  L  S  S  Q  T  A  G  A  A  G  W  R  R  S  G  S  N  G  G  G  Q  A  P  G  A  S  P  S  V  S  L  P  A  S  L  G  R

1567 .a............g...t...........a..........g....................................a....................g..
1201 AGGATGGGGTAGAGAAGAGAACCTTTTTTCTTGGAAGGGAGCTAAGGGACGGGGCATGCGGGGGAGAGGTCGAGGACGAGGGCATCCTGTTTCCTGTGTTGTAAATAGAAGCACTGACAA
394 G  W  G  R  E  E  N  L  F  S  W  K  G  A  K  G  R  G  M [R  G  R  G  R  G] H  P  V  S  C  V  V  N  R  S  T  D  N

1687 ...............................c....................................................................................
1321 CCAGAGGCAACAGCAATTAAATGACGTGGTAAAAAATTCATCTACTATTATCCAGAATCCAGTAGAGACACCCAAGAAGGACTATAGTCTGTTACCACTGTTAGCAGCTGCCCCTCAAGT
434 Q  R  Q  Q  Q  L  N  D  V  V  K  N  S  S  T  I  I  Q  N  P  V  E  T  P  K  K  D  Y  S  L  L  P  L  L  A  A  A  P  Q  V

1807 ..a.....................a..a............---------------------------------.....................c.................
1441 TGGAGAAAAGATTGCATTTAAGCTTTTGGAGCTAACATCCAGTTACTCTCCTGATGTCTCTGACTACAAGGAAGGAAGAATATTAAGCCACAATCCAGAGACCCAGCAAGTAGATATAGA
474 G  E  K  I  A  F  K  L  L  E  L  T  S  S  Y  S  P  D  V  S  D  Y  K  E  G  R  I  L  S  H  N  P  E  T  Q  Q  V  D  I  E

1894 ....g..............c.......................g.....................a.a...a.......c...g.....................
1561 AATTCTTTCATCCTTACCTGCCTTGAGAGAACCTGGGAAATTTGATTTAGTTTATCACAATGAAAATGGAGCCGAGGTAGTGGAGTACGCTGTGACACAGGAGAGCAAGATCACTGTATT
514 I  L  S  S  L  P  A  L  R  E  P  G  K  F  D  L  V  Y  H  N  E  N  G  A  E  V  V  E  Y  A  V  T  Q  E  S  K  I  T  V  F

2014 .....g.........................g......................                       total 2207 + poly(A)
1681 TTGGAAAGAGTTGATTGACCCAAGACTGATTATTGAATCTCCAAGTAACACATCAAGTACAGAACCTGCCTGA    total 2622 + poly(A)
554 W  K  E  L  I  D  P  R  L  I  I  E  S  P  S  N  T  S  S  T  E  P  A  *     576
```

Figure 1. Nucleotide and amino acid sequence of human p80-coilin. Full-length mRNA (pC15/pJEL1 combined, 2622nt; upper case) and the mRNA for a putative pseudogene for p80-coilin (2,224nt, lower case) shown here were submitted to GenBank/EMBL under accession numbers U06632 and L06522 respectively. For the full-length mRNA, an in-frame stop codon TGA in the 5' noncoding sequence is underlined. The short arrow (nucleotide 535) shows the 5' end of the longest cDNA clone pJEL1 previously obtained from cDNA library screening (9). The long arrow (nucleotides 632−610) represents the oligonucleotide primer used for the 5'-RACE method generating clone pC15 (nucleotides 1 to 632). The arrowheads indicate the exon−exon junctions. The mRNA for the putative pseudogene has multiple nucleotide substitutions, deletions (e.g. at nucleotides 547, 570, 1847), and insertions (at nucleotide 860, 966, 1337) resulting in numerous stop codons in all three reading frames. An internal *EcoRI* site was detected in both nucleotide sequences (underlined). Regions of Ser/Thr-rich, Arg/Lys-rich, Gly/Arg repeats, and an Asn repeat are boxed. In addition, putative phosphorylation sites for cdc2 kinase are double underlined.

5' sequence of p80-coilin

An alternative strategy using the 5'-RACE methodology was employed to obtain the 5' sequence of p80-coilin mRNA. The longest insert (632bp) was observed in clone pC15 and sequence analysis showed that this insert overlapped with the 5' region of the cDNA sequence identified in our earlier study (figure 1). The 5' sequence contained an open reading frame with an ATG at nucleotide 23 in good agreement with the Kozak consensus sequence (25). There is an in-frame TGA stop codon 12 bp upstream of the ATG start site. *In vitro* translation of the pC15 transcript yielded a labeled polypeptide of ~20kD which was not recognized by human or rabbit anti-p80-coilin sera in immunoprecipitation (data not shown). Drs Zheng'an Wu and Joe Gall, Carnegie Institute, Baltimore, have independently obtained the 5' sequence by 5'-RACE and exchanged sequence data with our laboratory. A full-length cDNA (clone p80.6), constructed by PCR using mRNA from HeLa cells and primers from the 5' and 3' sequences of p80-coilin, was subcloned into pCRII vector and kindly provided to us by Drs Wu and Gall. The complete nucleotide sequence of clone p80.6 was determined for both strands in our laboratory and two nucleotide substitutions were detected in positions 1633 (C→T) and 1771 (T→C) compared to the sequence reported earlier (9). Both nucleotide substitutions do not affect protein translation.

Sequence analysis of p80-coilin

The complete p80-coilin gene encodes a protein of 576 amino acids with a molecular mass of 62.6kD calculated using the GCG program PEPTIDESORT. The slower gel migration observed for this protein may be associated with several charged regions primarily located in the central domain (exon 2, see below). This domain includes several Arg/Lys-rich regions that may also encode potential nuclear localization signals (RKAKKR and KAKRK, see figure 1), and several Gly/Arg repeats (figure 1). Other regions of interest are the two Ser/Thr-rich regions, that are potential phosphorylation sites for casein kinase II, and several putative phosphorylation sites for cdc2 kinase (figure 1).

Northern blot analysis and expression of p80-coilin

Figure 2A shows that both the 3' probe pJEL2 (9) and the 5' probe pC15 hybridized to a band migrating at 2.7 kb. This result is in good agreement with the cDNA sequence of 2622 nucleotides not including the poly(A) tail (figure 1). When mRNA samples from different human tissues were hybridized with the p80-coilin probe, all samples showed a single band at 2.7 kb (figure 2B). Compared to the 2.2kb bands detected by the control probe for the non-snRNP splicesosomal protein SC35 (figure 2C), the expression of p80-coilin in brain and skeletal muscle appeared to be higher than in placenta, heart, and lung.

When the full-length p80-coilin cDNA was used as a template for *in vitro* transcription and translation, a band of approximately 80 kD was the major labeled product (Fig. 3). In the immunoprecipitation assay, the *in vitro* translation product was recognized by a human anti-p80-coilin serum and two rabbit anti-p80-coilinsera (R288 and R508), whereas no reactivity was observed for normal human serum and preimmune sera from the same rabbits (Fig. 3). When the p80-coilin from MOLT-4 cells, as detected by immunoblotting, and the *in vitro* translation product were compared by side-by-side analysis in the same polyacrylamide gel, it was apparent that the *in vitro* translation product migrated slightly faster than the cellular protein (Fig 4). This slight difference might be explained by post-translational modification(s) of p80-coilin in MOLT-4 cells (26).
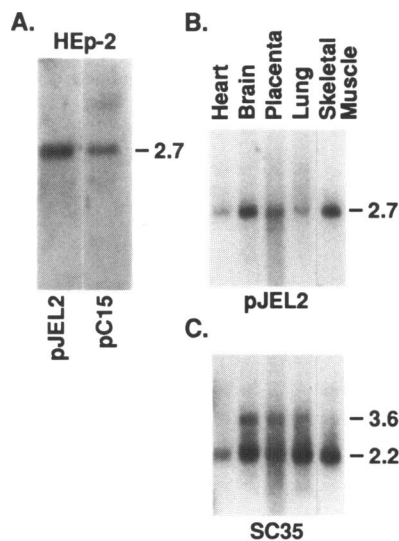
Figure 2. Northern blot analysis of human p80-coilin. **A.** 10μg of HEp-2 total cellular RNA was hybridized with [³²P]-labeled cDNA fragments from pJEL2 which contained the 3' half of p80-coilin cDNA (9), or from clone pC15 which was the 5' region of p80-coilin obtained by the 5'-RACE method. Both 5' and 3' cDNA probe hybridized to a single 2.7 kb mRNA. **B** and **C.** RNA blots containing poly (A)⁺ mRNA (2 μg/lane) from normal human tissues were hybridized with pJEL2 (B), stripped and reprobed with a cDNA of SC-35 as control (C). SC35 is a non-snRNP splicesosomal protein which shows relatively uniform mRNA expression with respect to the 2.2kb message (43).
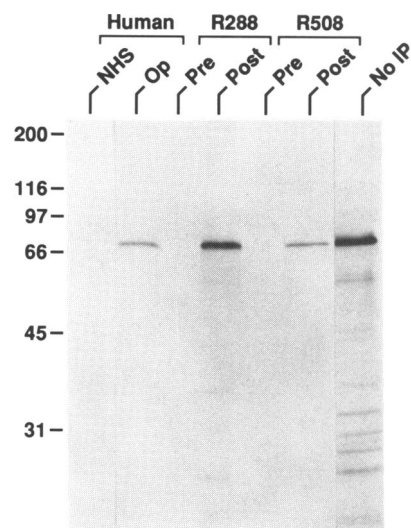
Figure 3. Immunoprecipitation of *in vitro* translation product of full-length cDNA for p80-coilin. R288: serum from a rabbit immunized with C-terminal 14-kD recombinant fragment of p80-coilin (9). R508: serum from a rabbit immunized with a 13-mer peptide corresponding to the predicted C-terminal sequence of p80-coilin. Human anti-p80-coilin serum Op, normal human serum (NHS), and sera from rabbits before (Pre) and after (Post) immunization were tested in immunoprecipitation as described in Materials and Methods using p80-coilin *in vitro* translation product (No IP).

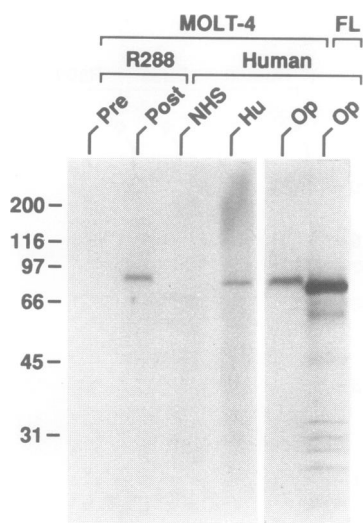## Anti-C-terminal peptide antibody showed specific binding to CBs

Fig. 5 shows the staining of CBs by rabbit anti-C-terminal peptide antiserum R508. Compared to the human anti-p80-coilin serum which gave CB and other diffuse nucleoplasmic staining, serum R508 showed highly specific staining of nuclear bodies colocalized with the CBs detected by the human serum, and thus confirming that this anti-C-terminal peptide antiserum recognized CBs. Rabbit serum 508 also showed CB staining in several human cell lines tested including MOLT-4, HEp-2, and HeLa S3. The reactivity of serum R508 suggests that this may be a highly special probe for CBs.

Gene for human p80-coilin

To obtain an independent confirmation for the 5' sequence derived by the RACE method, the analysis of genomic sequence was initiated. Among ~ 10^6 λFIXII phages screened by hybridization using both 5' and 3' cDNA probes, three overlapping clones were identified and analyzed by restriction mapping and DNA sequence determinations. Figure 6 shows the mapping data for the complete gene for p80-coilin. There are a total of 7 exons whose sequences are identical to the cDNA sequence in Fig. 1. The exon—exon junctions are indicated in Fig. 1 and the sequences for exon—intron junctions are shown in Fig. 6b. Southern blot analysis of human DNA from peripheral blood lymphocytes using either endonuclease *Hind*III or *Eco*RI showed characteristic bands consistent with the map shown in Fig. 6a (data not shown).

Chromosomal location

Fluorescence *in situ* hybridization using λ80g24 revealed exclusive labeling to chromosome 17 (Fig. 7A). Thirty-one metaphase spreads were scored, of which sixteen were chosen for imaging. The p80-coilin gene signals are within 17q22-23. An ideogram of human chromosome 17 is depicted in Fig. 7B. Interesting genes that have been mapped to this region of human chromosome 17 include a putative tumor suppressor gene NM23 (27) and B-cell neoplasm associated locus BCL-3 (28).

DISCUSSION

In this report, we have described the complete cDNA and gene sequence for a major protein of the coiled body, a nuclear body that is found in both mammalian cells and plants. The 5' region of the mRNA for human p80-coilin was obtained using the 5'-RACE method. This region most likely encoded the complete N-terminus of p80-coilin since there was a stop codon 12-nucleotides 5' to the methionine start site. The accuracy of this 5' sequence was confirmed by DNA sequencing of a genomic clone for human p80-coilin. The slight difference in migration between the *in vitro* translation product from the full-length cDNA clone and the cellular p80-coilin protein may be explained by putative post-translational modification. When HeLa cells were metabolically labeled with [^{32}P]-phosphate, p80-coilin could be detected in a standard immunoprecipitation assay showing that p80-coilin was phosphorylated during the overnight labeling procedure [Andrade and Chan, unpublished data, (26)]. It is not clear if phosphorylation alone can account for the difference observed. Carmo-Fonseca *et al.* (26) have described the hyperphosphorylation of p80-coilin during the mitotic phase of the cell cycle and, showed that phosphoserine is the only residue detected in both interphase and mitotic cells, from the analysis of phosphoamino acids (26). Although the sites of p80-coilin phosphorylation have not been mapped, the presence of the Ser/Thr-rich regions and potential sites for cdc2 kinases are consistent with the described mitotic hyperphosphorylation (26). Using a synthetic peptide corresponding to the putative C-terminus of p80-coilin for the production of antibody, it was shown that the resulting antiserum was a functional reagent for
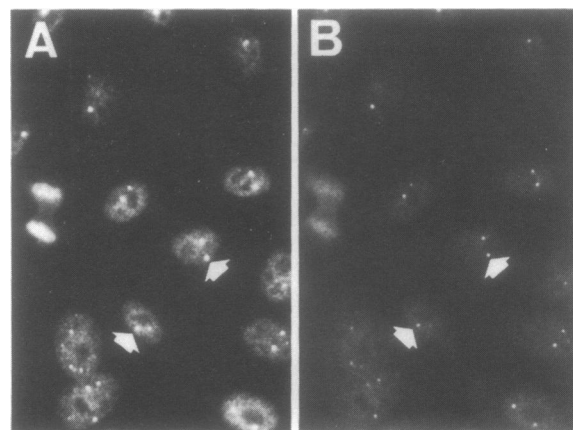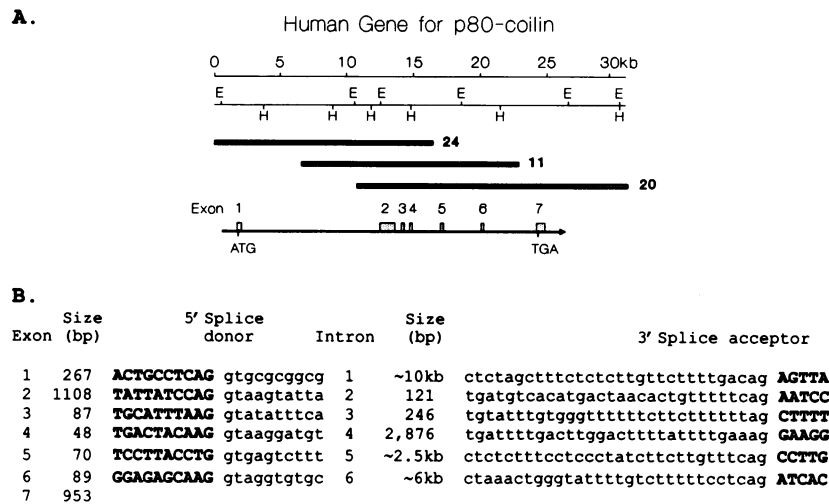
Figure 4. The comparison of p80-coilin in MOLT-4 whole cell extract and in *in vitro* translation. MOLT-4 whole cell extracts and [^{35}S]-methionine labeled *in vitro* translation products were analyzed on a 12.5 % polyacrylamide gel and transferred to nitrocellulose membrane. The identity of p80-coilin in a MOLT-4 whole cell extract was determined by immunoblotting using rabbit serum R288 and human anti-p80-coilin sera Hu and Op. Comparing the MOLT-4 p80-coilin and *in vitro* translation product from the full-length cDNA (FL), the cellular p80-coilin migrated slightly slower than the *in vitro* translation product. This difference might be explained by the presence of post-translational modification(s) of p80-coilin in MOLT-4 cells.

Figure 5. Immunostaining of CBs in HEp-2 cells using anti-C-terminal peptide serum R508. Cells were stained with a human serum with anti-p80-coilin antibodies (A) and the rabbit serum R508 (B). Serum R508 showed specific staining for CBs that were recognized by the human serum (arrows). In addition, the human serum also gave nucleoplasmic and mitotic chromosomal staining that are unrelated to p80-coilin.

A.

Human Gene for p80-coilin

B.

Exon	Size (bp)	5' Splice donor	Intron	Size (bp)	3' Splice acceptor
1	267	**ACTGCCTCAG** gtgcgcggcg	1	~10kb	ctctagctttctctcttgttcttttgacag **AGTTA**
2	1108	**TATTATCCAG** gtaagtatta	2	121	tgatgtcacatgactaacactgtttttcag **AATCC**
3	87	**TGCATTTAAG** gtatatttca	3	246	tgtatttgtgggttttttcttctttttttag **CTTTT**
4	48	**TGACTACAAG** gtaaggatgt	4	2,876	tgattttgacttggactttttatttttgaaag **GAAGG**
5	70	**TCCTTACCTG** gtgagtcttt	5	~2.5kb	ctctctttcctccctatcttcttgtttcag **CCTTG**
6	89	**GGAGAGCAAG** gtaggtgtgc	6	~6kb	ctaaactgggtattttgtcttttttcctcag **ATCAC**
7	953				

Figure 6. The human gene for p80-coilin. **A.** The human p80-coilin gene spans 25kb and is composed of seven exons. Exon 1 contains the translation initiation codon while the stop codon is located in exon 7. **B.** The 5' and 3' splice sites for all the exons are shown.
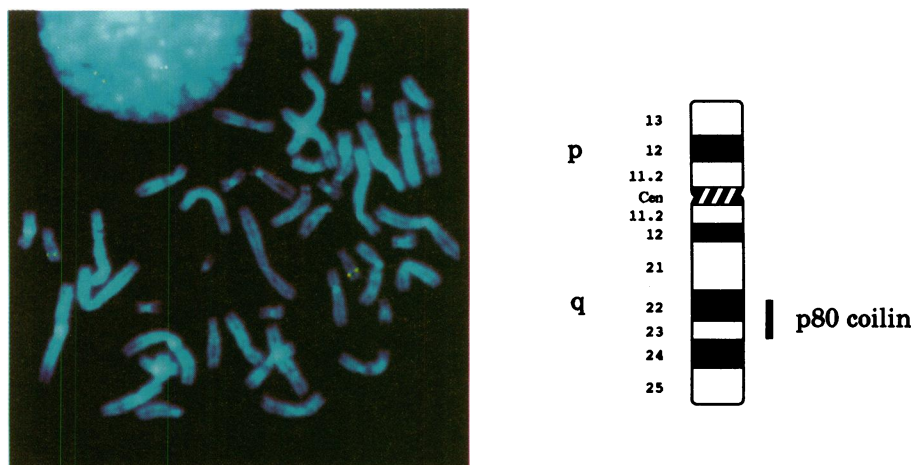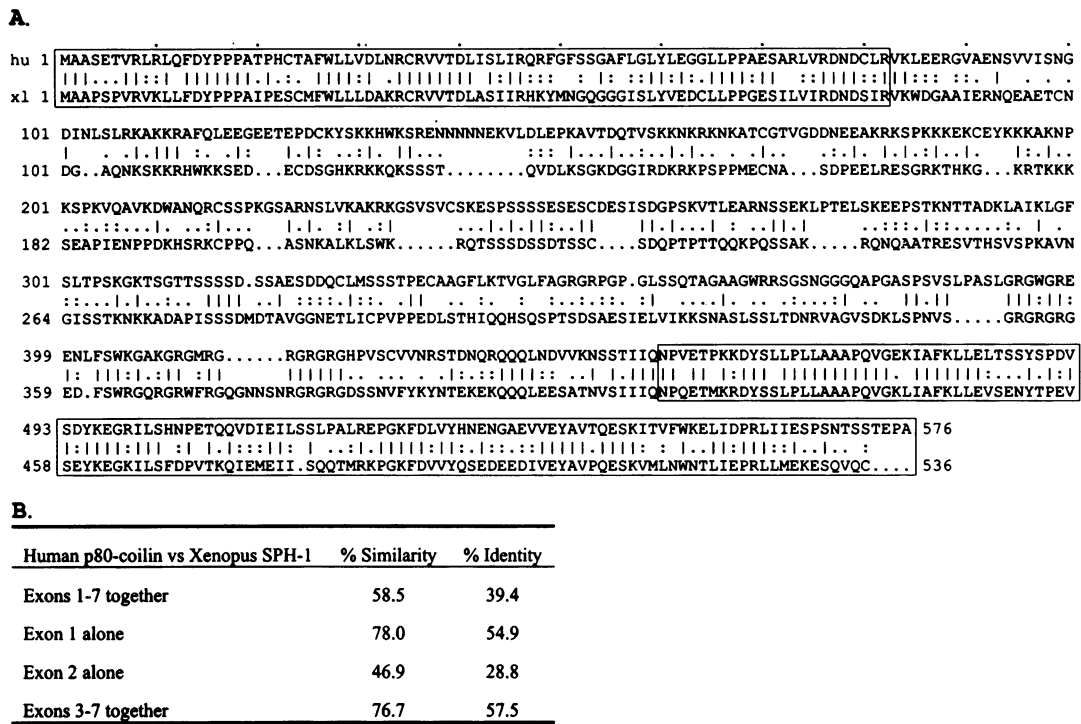

Figure 7. Chromosomal localization of the human p80-coilin gene to 17q22−23. **A.** *In situ* hybridization of biotinylated λ80g24 to a human metaphase spread. **B.** A schematic representation of a G-banded human chromosome 17. The location of the p80-coilin gene is shown by a vertical line.

the detection of CBs in many human cell types. The fact that this serum (R508) also recognized the *in vitro* translation product of the full-length p80-coilin essentially confirmed the authenticity of the cDNA clones.

CB-like structures may be present in frogs and related species. There are three types of 'snurposomes' described in the germinal vesicles of amphibian oocytes (29,30). Type A snurposomes contain only U1-snRNP, type B snurposomes have all the U-snRNPs including the non-snRNP splicing factor SC-35, and type C snurposomes also have all the U-snRNPs but lack SC-35 and are the largest with diameters ranging from less than 1μm to greater than 20μm (29−31). Type C snurposomes, also known as spheres or sphere organelles, are thought to be related to mammalian CBs based on the immunostaining for p80-coilin and the presence of other snRNPs and snoRNPs that are known to be present in human CBs. Recently, Bauer *et al.* (32) described the *in vitro* assembly of CBs in *Xenopus* egg extracts and anti-

p80-coilin antibody R288 was shown to recognize these assembled structures. On the other hand, Tuma *et al.* (33) have reported the characterization of a Xenopus sphere organelle protein SPH-1 by cDNA cloning and its similarity to p80-coilin has already been noted. The similarity between these two proteins is enhanced by the present study (fig. 8) because both N- and C-terminal domains, corresponding to exon 1 and exons 3−7 respectively, showed relatively high degrees of sequence identity and similarity (fig. 8B). The major difference between the two proteins is found in the center regions, corresponding primarily to exon 2 of p80-coilin alone. Our present data may be useful for the characterization and cloning of p80-coilin homologs in other species.

The exact composition of mammalian CBs is not known, but many snRNAs and associated components are highly enriched in CBs (10). Using anti-Sm or snRNP antibody as probes, Fakan *et al.* (34) and Elicieri and Ryerse (35) were among the first to

A.

```
hu  1 MAASETVRLRLQFDYPPPATPHCTAFWLLVDLNRCRVVTDLISLIRQRFGFSSGAFLGLYLEGGLLPPAESARLVRDNDCLRVKLEERGVAENSVVISNG
      |||...||::| ||||||||.|.:. |||||:| .|||||||| |:||::: : |: ::||:|:.||||:|| ::||||::||||::: ::. | ...
x1  1 MAAPSPVRVKLLFDYPPPAIPESCMFWLLLDAKRCRVVTDLASIIRHKYMNGQGGGISLYVEDCLLPPGESILVIRDNDSIRVKWDGAAIERNQEAETCN

  101 DINLSLRKAKKRAFQLEEGEETEPDCKYSKKHWKSRENNNNNEKVLDLEPKAVTDQTVSKKNKRKNKATCGTVGDDNEEAKRKSPKKKEKCEYKKKAKNP
      | . .|.||| :. .|: |.|.: ..:|. ||... ::: |...|..:.:.|.|.. . .|. .:.|. |.|.:|.|. |:.|..
  101 DG..AQNKSKKRHWKKSED...ECDSGHKRKKQKSSST........QVDLKSGKDGGIRDKRKPSPPMECNA...SDPEELRESGRKTHKG...KRTKKK

  201 KSPKVQAVKDWANQRCSSPKGSARNSLVKAKRKGSVSVCSKESPSSSSESESCDESISDGPSKVTLEARNSSEKLPTELSKEEPSTKNTTADKLAIKLGF
      ..:.::...|. ..:|.... |.|. :| .:| ...|.|.||:..|| ||.|...| : ..||.| ..:::.|::......:. | .
  182 SEAPIENPPDKHSRKCPPQ...ASNKALKLSWK......RQTSSSDSSDTSSC....SDQPTPTTQQKPQSSAK.....RQNQAATRESVTHSVSPKAVN

  301 SLTPSKGKTSGTTSSSSD.SSAESDDQCLMSSSTPECAAGFLKTVGLFAGRGRPGP.GLSSQTAGAAGWRRSGSNGGGQAPGASPSVSLPASLGRGWGRE
      ::...|.|...:.. |||| ..| ::::.|::. .|| .. :. : :...::. :| ....|. .. .:...:| .. ||.|| |||:||:
  264 GISSTKNKKADAPISSSDMDTAVGGNETLICPVPPEDLSTHIQQHSQSPTSDSAESIELVIKKSNASLSSLTDNRVAGVSDKLSPNVS.....GRGRGRG

  399 ENLFSWKGAKGRGMRG......RGRGRGHPVSCVVNRSTDNQRQQQLNDVVKNSSTIQNPVETPKKDYSLLPLLAAAPQVGEKIAFKLLELTSSYSPDV
      |: |||:|.:|| :|| ||||||.. ... .|:.::||||:: ..| |.||||| || |:||| ||||||||||. |||||||:...|.|:|
  359 ED.FSWRGQRGRWFRGQGNNSNRGRGRGDSSNVFYKYNTEKEKQQQLEESATNVSIIQNPQETMKRDYSSLPLLAAAPQVGKLIAFKLLEVSENYTPEV

  493 SDYKEGRILSHNPETQQVDIEILSSLPALREPGKFDLVYHNENGAEVVEYAVTQESKITVFWKELIDPRLIIESPSNTSSTEPA 576
      |:||||:||| :| |.|:::||: | ..:|.||||||:||:.|::.::||||.|||: : |..||:|||:|..|.. :
  458 SEYKEGKILSFDPVTKQIEMEII.SQQTMRKPGKFDVVYQSEDEEDIVEYAVPQESKVMLNWNTLIEPRLLMEKESQVQC.... 536
```

B.

Human p80-coilin vs Xenopus SPH-1	% Similarity	% Identity
Exons 1-7 together	58.5	39.4
Exon 1 alone	78.0	54.9
Exon 2 alone	46.9	28.8
Exons 3-7 together	76.7	57.5

Figure 8. Comparison of amino acid sequences between human p80-coilin and *Xenopus laevis* SPH-1 proteins. The alignment of the two amino acid sequences (A) was generated using the GAP program which also produced the percent similarity and identity shown in (B). The boxes in panel A represent the regions corresponding to exon 1 and exons 3−7 for human p80-coilin, respectively.

show that snRNP proteins are found in human and rodent CBs. The presence of snRNAs in CBs has been reported in many studies using antibody to m₃G capped RNA (10) or hybridization with oligo probes (36,37). The variety of proteins and snRNPs detected in CBs suggests that there may be multiple functions for this nuclear substructure. With the complete cDNA for human p80-coilin, we can begin to examine the interaction of this major component of CBs with other proteins.

The cellular function of CBs remains unclear, although there are several recent studies that give some interesting insights into the role of CBs in different biological systems. From the initial description of CBs attached to the periphery of nucleoli in rat neurons, many investigators have observed the association between the nucleolus and CBs (38−41). Recently, Ochs *et al.* (41) showed that, in addition to the normal nuclear CBs which can be found sometimes close to or attached to the nucleolus, CBs are also found encased within the nucleolus in a few human breast cancer cell lines. The nucleolar CBs can be detected by immunostaining with antibodies to p80-coilin, Sm, and fibrillarin. Furthermore, nucleolar CBs show characteristic coiled fibrils as seen by electron microscopy (41). Although the data were derived from only a few cancer cell lines, their results suggest that the link between CBs and nucleoli may be more apparent in some aberrant conditions such as breast carcinoma. Interestingly, similar nucleolar CBs are also observed in hibernating dormouse liver nuclei in the recent report of Malatesta *et al.* (42). By examining brown adipose tissue and liver of hibernating, arousing, and euthermic dormice, Malatesta *et al.* (42) have shown that both nucleolar CBs and extranucleolar CBs are found in hibernating, but not in euthermic dormice. In arousing

dormice, only poorly structured CBs are observed, suggesting that they are in the process of dissociation. This study suggests that during hibernation CBs might be storage sites for snRNP and snoRNPs which are then released during arousal and ready for RNA processing functions. Whether CBs are simple storage sites for snRNPs and snoRNPs in other cell types remains to be determined. Our earlier report of CBs during the cell cycle in *in vitro* culture systems shows that small CBs are formed during mid G1 phase after the assembly of the nucleolus and as the cells progress to S and G2 phases, the number of CBs decreases while their size increases (14). In view of the findings of Malatesta *et al.* (42), it is possible to reinterpret the cell cycle changes in CBs as an excess of snRNPs that are stored away in smaller CBs that eventually reassembled into fewer but larger CBs at the later stages of the cell cycle where there is less of a cellular requirement for snRNPs. Since we have noted that the level of p80-coilin protein remains essentially the same during the cell cycle (14), it is interesting to speculate that phosphorylation of p80-coilin is associated with its role in the cycling of CBs. This report on the gene structure of human p80-coilin will hopefully provide a means for probing the function of p80-coilin and CBs in a variety of systems.

ACKNOWLEDGEMENTS

REFERENCES

1. Brasch,K. and Ochs,R.L. (1992) Exp. Cell Res. 202, 211−223.
2. Lamond,A.I. and Carmo-Fonseca,M. (1993) Trends Cell Biol. 3, 198−204.
3. Spector,D.L. (1993) Annu. Rev. Cell Biol. 9, 265−315.
4. Carter,K.C., Taneja,K.L. and Lawrence,J.B. (1991) J. Cell Biol. 115, 1191−1202.
5. Ascoli,C.A. and Maul,G.G. (1991) J. Cell Biol. 112, 785−795.
6. Weis,K., Rambaud,S., Lavau,C., Jansen,J., Carvalho,T., Carmo-Fonseca,M., Lamond,A.I. and Dejean,A. (1994) Cell, 76, 345−356.
7. Dyck,J.A., Maul,G.G., Miller, Jr., Chen,J.D., Kakizuka,A. and Evan,R.M. (1994) Cell, 76, 333−343.
8. Saunders,W.S., Cooke,C.A. and Earnshaw,W.C. (1991) J. Cell Biol. 115, 919−931.
9. Andrade,L.E.C., Chan,E.K.L., Raska,I., Peebles,C.L., Roos,G. and Tan,E.M. (1991) J. Exp. Med. 173, 1407−1419.
10. Raska,I., Andrade,L.E.C., Ochs,R., Chan,E.K.L., Chang,C.M., Roos,G. and Tan,E.M. (1991) Exp. Cell Res. 195, 27−37.
11. Raska,I., Ochs,R.L., Andrade,L.E.C., Chan,E.K.L., Burlingame,R., Peebles,C.L., Gruol,D. and Tan,E.M. (1991) J.Struct.Biol. 104, 120−127.
12. Monneron,A. and Bernhard,W. (1969) J. Ultrastruct. Res. 27, 266−288.
13. Hardin,J.H., Spicer,S.S. and Greene,W.B. (1969) Anat. Rec. 164, 403−431.
14. Andrade,L.E.C., Tan,E.M. and Chan,E.K.L. (1993) Proc. Natl. Acad. Sci. USA, 90, 1947−1951.
15. Spector,D.L., Lark,G. and Huang,S. (1992) Mol. Biol. Cell, 3, 555−569.
16. Liu,F.-T., Zinnecker,M., Hamaoka,T. and Katz,D.H. (1979) Biochemistry, 18, 690−697.
17. Devereux,J., Haeberli,P. and Smithies,O. (1984) Nucleic Acids Res. 12, 387−395.
18. Needleman,S.B. and Wunsch,C.D. (1970) J. Mol. Biol. 48, 443−453.
19. Feinberg,A.P. and Vogelstein,B. (1983) Anal. Biochem. 132, 6−13.
20. Chirgwin,J.M., Przbyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) Biochemistry, 18, 5294−5299.
21. Chan,E.K.L., Hamel,J.C., Buyon,J.P. and Tan,E.M. (1991) J. Clin. Invest. 87, 68−76.
22. Lichter,P., Tang,C.C., Call,K., Hermanson,G., Evan,G.A., Housman,D. and Ward,D.C. (1990) Science, 247, 64−69.
23. Ward,D.C., Lichter,P., Boyle,A., Baldini,A., Menninger,J. and Ballard,S.G. (1991) In Lindsten,J. and Pettersson,U. (eds.), Etiology of human disease at the DNA level. Raven Press Ltd. London, pp. 291−303.
24. Ballard,S.G. and Ward,D.C. (1993) J. Histochem. Cytochem. 41, 1755−1759.
25. Kozak,M. (1987) Nucleic Acids Res. 15, 8125−8148.
26. Carmo-Fonseca,M., Ferreira,J. and Lamond,A.I. (1993) J. Cell Biol. 120, 841−852.
27. Varesco,L., Caligo,M.A., Simi,P., Black,D.M., Nardini,V., Casarino,L., Rocchi,M., Ferrara,G., Solomon,E. and Bevilacqua,G. (1992) Genes, Chromosomes & Cancer, 4, 84−88.
28. Yano,T., Sander,C.A., Andrade,R.E., Gauwerky,C.E., Croce,C.M., Longo,D.L., Jaffe,E.S. and Raffeld,M. (1993) Blood, 82, 1813−1819.
29. Gall,J.G. (1991) Science, 252, 1499−1500.
30. Wu,Z., Murphy,C., Callan,H.G. and Gall,J.G. (1991) J. Cell Biol. 113, 465−483.
31. Wu,C.-H.H. and Gall,J.G. (1993) Proc. Natl. Acad. Sci. USA, 90, 6257−6259.
32. Bauer,D.W., Murphy,C., Wu,Z., Wu,C.-H.H. and Gall,J.G. (1994) Mol. Biol. Cell, (in press).
33. Tuma,R.S., Stolk,J.A. and Roth,M.B. (1993) J. Cell Biol. 122, 767−773.
34. Fakan,S., Leser,G. and Martin,T.E. (1984) J. Cell Biol. 98, 358−363.
35. Eliceiri,G.L. and Ryerse,J.S. (1984) J. Cell Physiol. 121, 449−451.
36. Huang,S. and Spector,D.L. (1992) Proc. Natl. Acad. Sci. USA, 89, 305−308.
37. Matera,A.G. and Ward,D.C. (1993) J. Cell Biol. 121, 715−727.
38. Lafarga,M., Hervas,J.P., Santa Cruz,M.C., Villegas,J. and Crespo,D. (1983) Anat. Embryol. (Berl.), 166, 19−30.
39. Vagner-Capodano,A.M., Mauchamp,J., Stahl,A. and Lissitzky,S. (1980) J. Ultrastruct. Res. 70, 37−51.
40. Lafarga,M., Andres,M.A., Berciano,M.T. and Maquiera,E. (1991) J. Comp. Neurol. 308, 329−339.
41. Ochs,R.L., Stein,T.W.,Jr. and Tan,E.M. (1994) J.Cell Sci. 107, 385−399.
42. Malatesta,M., Zancanaro,C., Martin,T.E., Chan,E.K.L., Amalric,F., Luhrmann,R., Vogel,P. and Fakan,S. (1994) Exp. Cell Res. 211, 415−419.
43. Fu,X.-D. and Maniatis,T. (1992) Science, 256, 535−538.