

# Molecular Characterization of NDP52, A Novel Protein of the Nuclear Domain 10, Which Is Redistributed upon Virus Infection and Interferon Treatment

Frank Koriath,\* Christian Gieffers,\* Gerd G. Maul,‡ and Jürgen Frey\*

\*University of Bielefeld, Faculty for Chemistry, Department of Biochemistry, 33615 Bielefeld, Germany; and ‡The Wistar Institute of Anatomy and Biology, Philadelphia, Pennsylvania 19104-4268

**Abstract.** The nuclear domain (ND)10 also described as POD or Kr bodies is involved in the development of acute promyelocytic leukemia and virus–host interactions. Immunofluorescence analysis using a variety of human autoimmune sera and monoclonal antibodies showed a typical dot like nuclear staining for ND10, suggesting that this structure consists of several proteins. Two of the ND10 proteins, Sp100 and PML are genetically characterized and show homology with several transcription factors.

Here we describe NDP52, an additional novel protein of the ND10. We raised a new mAb C8A2, that specifically recognizes NDP52. Immunofluorescence analysis using this mAb showed a typical nuclear dot staining as it was described for ND10. Isolation and se-

quencing of the corresponding cDNA revealed that NDP52 has a predicted molecular mass of 52 kD. The deduced amino acid sequence exhibits an extended central coiled coil domain containing a leucine zipper motif. The COOH terminus of NDP52 shows homology with LIM domains, that have recently been described to mediate protein interactions, which let NDP52 appear as a suitable candidate for mediating interactions between ND10 proteins. In vivo, NDP52 is transcribed in all human tissues analyzed. Furthermore, we show that NDP52 colocalizes with the ND10 protein PML and can be redistributed upon viral infection and interferon treatment. These data suggest that ND10 proteins play an important role in the viral life cycle.

**T**HE internal structure and the role of the different organelles of the nucleus are less well characterized than those of the cytoplasm. Immunohistochemical studies and ultrastructural analyses have identified a number of nuclear domains, the components and functions of which are only partially understood. These structures include perichromatin fibrils, interchromatin granules and nuclear bodies (Fakan et al., 1984; Puvion et al., 1984; Visa et al., 1993). Some of the nuclear domains appear in a speckled or dot-like distribution within the nucleus. The components of the splicing regions, i.e., SC35 (Fu and Maniatis, 1990), hnRNPs (Fakan et al., 1984), and snRNPs (Spector et al., 1983), are concentrated in 20–50 nuclear speckles. The distribution of these speckles varies among different cell types and can be altered by several factors affecting transcription (Antonioni et al., 1993) and by virus infection (Bridge et al., 1993).

ND10 are nuclear domains, originally characterized using mAbs and human autoantisera from patients with primary

biliary cirrhosis. At the immunohistochemical level ND10 appear as nuclear dots, where 10 indicates the approximate number of dots observed per nucleus (Ascoli and Maul, 1991). ND10 did not colocalize with kinetochores, centromeres, sites of mRNA processing, and chromosomes. Resistance of ND10 antigens to nuclease digestion and salt extraction suggested that ND10 are associated with the nuclear matrix (Ascoli and Maul, 1991). Infection of ND10 positive cells with herpes simplex virus type-1 (HSV-1)<sup>1</sup> abrogates immunofluorescent staining of ND10 (Maul et al., 1993). Using HSV-1 deletion mutants the loss of staining was found to be induced by the immediate early gene 1 product (ICP0) of HSV-1 (Maul et al., 1993; Everett and Maul, 1994; Maul and Everett, 1994).

Immunofluorescence analyses using a variety of human autoantisera and antibodies have shown that ND10 are composed of several colocalized proteins. One of the ND10 antigens, recognized by a mAb 138, was identified as a 55-kD protein, occurring as three charge isomers with

Address all correspondence to Dr. Jürgen Frey, University of Bielefeld, Faculty for Chemistry, Dept. of Biochemistry, Universitätsstraße 25, 33615 Bielefeld, Germany. Tel.: 49 521-1062079. FAX: 49 521-1066146.

Drs. Koriath and Gieffers contributed equally.

1. *Abbreviations used in this paper:* 5'RACE, rapid amplification of cDNA 5' ends; HSV-1, Herpes simplex virus type I; ND10, nuclear domain 10; NDP52, nuclear dot protein 52; IFN, interferon; snRNPs, small nuclear ribonucleoprotein particles.

a pI ranging from 7.4–7.7 (Ascoli and Maul, 1991). Two other ND10 proteins have been characterized by cDNA sequence analyses. One of these, Sp100, is an autoantigen in primary biliary cirrhosis (Szosteki et al., 1990). The transcription and expression of Sp100 were shown to be enhanced after virus infection or mitogen and interferon (IFN) treatment (Guldner et al., 1992). The deduced amino acid sequence of the human Sp100 revealed some homology to several transcriptional regulatory proteins, although the exact biological function of Sp100 is still unknown. Recent investigations have shown that ND10 contain PML (Dyck et al., 1994), the second ND10 protein that has been genetically analyzed. It is involved in acute promyelocytic leukemia (APL), a hematopoietic malignancy in which myeloid precursors are blocked at a proliferative state. This disease is associated with a specific t(15; 17) translocation that generates a fusion gene of PML and the retinoic acid receptor- $\alpha$  (de The et al., 1991; Goddard et al., 1991; Kakizuka et al., 1991; Kastner et al., 1992). Compared to the ND10-like distribution of the wild-type PML, the resulting fusion protein causes an unusual microparticulate pattern of PML and the fusion protein. Retinoic acid treatment of these APL cells restores the pattern of PML to normal appearing ND10 (Dyck et al., 1994; Koken et al., 1994; Weis et al., 1994). In vivo the administration of all trans retinoic acid has led to complete remission in APL patients (Castaigne et al., 1990). PML shows homology with several transcription factors (Kakizuka et al., 1991).

Most ND10 correspond to nuclear bodies (Koken et al., 1994). These large structures of unknown function respond to environmental stimuli like interferon and viral infection (Maul et al., 1993). Their function in the cell can only be elucidated by the identification of further associated proteins and the characterization of their behavior under the known modifying conditions.

In this study we describe a novel protein, NDP52, which is recognized by mAb C8A2, and shows the typical ND10 immunofluorescence staining pattern. NDP52 colocalizes with the ND10 proteins PML and Sp100. A full-length cDNA encoding NDP52 was isolated by screening a  $\lambda$ -ZAP cDNA expression library with mAb C8A2, together with extension of the 5' end using the 5'RACE system. NDP52 is expressed in a variety of cell lines and human tissues, and its expression and distribution are modified by different external triggers.

## Materials and Methods

### Antibodies and Cell Culture

The monoclonal antibody C8A2 was produced by immunizing mice with crude nuclei fractions prepared as described by Martelli et al. (1992). Fusion of spleen cells with the FO Myeloma and cloning of hybridoma were performed as described (Fazeka de St. Groth and Scheidegger, 1980). Hybridoma supernatants were screened for reactive mAbs by indirect immunofluorescence analysis of permeabilized MG63 cells. PML was visualized using a specific rabbit antiserum (Dyck et al., 1994).

HEp-2 carcinoma cells (CCL 23; American Type Tissue Collection, Rockville, MD) and the diploid human fibroblast cell line, MRC-5 (CCL 171) were maintained in MEM. MG63 cells (CRL 1427) were maintained in Iscove's DMEM and COS-7 cells (CRL 1651) in DMEM. Further cell lines grown for total RNA preparation were Saos-2 (HTB 85) maintained in Iscove's DMEM, SK-ES-1 (HTB 86), maintained in McCoy's 5a and

HeLa (CCL 2) maintained in Eagle's MEM. All media were supplemented with 10% FCS and cells were grown at 37°C in a humidified atmosphere containing 5% CO<sub>2</sub>. For immunofluorescence analysis, cells were grown on round coverslips in 24-well plates (Corning Glass Inc., Corning, NY). For IFN stimulation, 1,000 U/ml of the respective interferon was added to the culture medium for 18 h (IFN- $\beta$  was kindly provided by Dr. Renschler Biotechnology GmbH, Laupheim, Germany; INF- $\gamma$  was kindly provided by the DRK, Springe, Germany).

### Immunofluorescence Microscopy

HEp-2 and MRC-5 cells were fixed at room temperature for 5 min with freshly prepared 1% paraformaldehyde in PBS, washed with PBS, and permeabilized for 20 min in PBS+0.2% (vol/vol) Triton X-100 (Sigma Immunochemicals, St. Louis, MO) on ice. Nuclear antigen localization was determined after incubation of permeabilized cells with rabbit antiserum, or mAb diluted in PBS for 1 h at room temperature. For double labeling, the weaker of the two primary antibodies was complexed with a secondary biotinylated antibody and labeled with avidin-fluorescein or avidin-Texas red. The stronger reacting antibody was detected with FITC or Texas red labeled secondary antibodies (Vector Labs., Inc., Burlingame, CA). Cells were stained for DNA with 0.5 mg/ml of bisBenzimide (Hoechst 33258; Sigma) in PBS and mounted with Fluoromount G (Fisher Scientific, Pittsburgh, PA). Fluorescence images were analyzed using a Leica confocal laser scanning microscope and recorded with a Focus Imagecorder Plus graphics recorder (Leica, Inc., Deerfield, IL).

Immunofluorescence on MG63 and COS-7 cells were performed at room temperature by fixing cells with 4% (wt/vol) paraformaldehyde in PBS for 10 min. Cells were subsequently permeabilized with 0.2% (vol/vol) Triton X-100 in PBS for 15 min and incubated with 1 mg/ml mAb C8A2 in PBS for 1 h. Bound mAb was detected using a goat anti-mouse antibody, conjugated with FITC (Sigma, Deisenhofen, Germany). Finally, cells were embedded in Moviol 4-88 (Hoechst AG, Frankfurt a. M., Germany). Specimens were analyzed using a Zeiss Axiophot fluorescence microscope equipped for epifluorescence (Carl Zeiss, Inc., Thornwood, NY).

### Immunoprecipitation

For immunoprecipitation of NDP52 cells were lysed in 10 mM Tris containing 1% (wt/vol) NP40, 150 mM NaCl, pH 7.5 and centrifuged at 15,000 g for 10 min. The supernatant was incubated with mAb C8A2 coupled to protein G-Sepharose for 1 h at 4°C. After extensive washing with lysis buffer bound antigen was eluted in SDS sample buffer and analyzed by SDS-PAGE.

### Gel Electrophoresis and Immunoblot Analyses

Proteins were separated by one-dimensional SDS-PAGE (Laemmli, 1970). For immunoblot analysis proteins were subsequently transferred to nitrocellulose filters (Towbin et al., 1979). Filters were blocked by incubation with 3% (wt/vol) skimmed milk in PBS overnight and probed with 1 mg/ml mAb C8A2 for 1 h at room temperature. After washing with PBS+(wt/vol) 0.05% Tween 20 filters were incubated with goat anti-mouse IgG +IgM coupled to either horse radish peroxidase or alkaline phosphatase. Detection of bound antibody was performed after intensive washing with either the ECL detection system (Amersham Buchler GmbH, Braunschweig, Germany) or nitro blue tetrazolium chloride and 5-bromo-4-chloro-3-indolyl phosphate as substrates for the alkaline phosphatase.

The isoelectric point of the immunoprecipitated NDP52 was analyzed by two-dimensional near equilibrium pH gradient electrophoresis (NEPHGE) (O'Farrell et al., 1977) and isoelectric focussing (IEF), using Servalyt Precotes with a pH range of 3–10 (Serva Feinbiochemica GmbH, Heidelberg, Germany) according to the instructions of the manufacturer.

### Isolation and Preliminary Analysis of cDNA Clones

A directional, size fractionated cDNA library of human osteosarcoma MG63 cells was constructed in  $\lambda$ -ZAP using the ZAP-cDNA Gigapack II Gold cloning Kit (Stratagene GmbH, Heidelberg, Germany). Approximately 10<sup>6</sup> phages were used for immunoscreening with mAb C8A2 according to the instructions of the picoBlue Immunoscreening Kit (Stratagene GmbH, Heidelberg, Germany). Two phages, clearly reacting with mAb C8A2, were isolated and plaque purified. After transformation of the positive  $\lambda$ -ZAP phages into Bluescript plasmids by in vivo excision re-

striction analysis revealed that both plasmids contained identical 2.3-kb inserts (clone pFK52).

### Cloning of the 5' End of the NDP52 cDNA by Polymerase Chain Reaction

Amplification of the 5' coding region of NDP52 was performed using the 5'RACE system (Life Technology GmbH, Eggenstein, Germany). MG63 total RNA was obtained by lysis of the cells in 4 M guanidinium thiocyanate, 10 mM Tris-HCl, pH 7.5, after CsCl gradient centrifugation (Chirwin et al., 1979). A 20-mer primer (GSP1: GATGAAATGCTGGGTGA-AGG) was used to prime reverse transcription of MG63 total RNA using SuperScript Reverse Transcriptase (Life Technology GmbH). The generated first strand cDNA was tailed at its 3' end with terminal deoxynucleotidyl-transferase in the presence of dCTP. PCR was performed with Taq polymerase in a 50  $\mu$ l reaction volume, using a nested primer (GSP2: TATAATGACATGTGACGTCC) containing an AatII restriction site, plus the anchor primer (Life technology), complementary to the synthetic Oligo-dC cDNA tail. The 5' end of the anchor primer contains the XbaI, SpeI, and ClaI restriction sites. The RACE product was extracted with phenol-chloroform, ethanol precipitated and resuspended in 10 mM Tris-HCl, 0.1 mM EDTA, pH 7.6. The amplified DNA was subsequently digested with the restriction enzymes AatII and SpeI, and cloned into the corresponding restriction sites of the Bluescript plasmid containing the library insert of clone pFK52 to generate the construct pFK52c.

### Sequence Analysis

The pFK52c insert was cloned into the pUC 18 vector in both orientations. A set of deletion clones was generated with the double-stranded Nested deletion kit (Pharmacia Biotech Europe GmbH, Freiburg, Germany). The different clones were sequenced by the dideoxy method (Sanger et al. 1977) using the DIG Taq Sequencing Kit (Boehringer Mannheim GmbH, Mannheim, Germany). Overlapping sequences were combined to generate the complete sequence of both strands. The part containing the RACE product, corresponding to the 5' end was sequenced from four different clones to exclude PCR errors.

Sequence similarity searches of the EMBL-GENBANK and SwissProt databases were performed with the FASTA (Pearson, 1990) and BLITZ (Smith and Waterman, 1981) programs. Prediction of the secondary structure was performed using the algorithm of Kyte and Doolittle (1982), Chou and Fasman (1978), and the coil algorithm of Lupas et al. (1991), implemented in the GENEPRO (Riverside Scientific Enterprises, Bainbridge Island, WA) and the PC GENE (Intelligenetics, Geel, Belgium) software packages.

### In Vitro Transcription and Translation

The construct pFK52c was used in a coupled transcription and translation assay using the TnT coupled Reticulocyte Lysate system (Promega Corp., Madison, WI). The transcript was generated with T7 RNA polymerase from the circular plasmid template and translated in the presence of [<sup>35</sup>S]methionine in a combined reaction. The translation product was analyzed by SDS-PAGE and subsequent fluorography using Amplify (Amersham Buchler GmbH, Braunschweig, Germany).

### Recombinant Expression of NDP52

A fragment of pFK52c (bp 53–1334) comprising the complete coding region for NDP52 was generated by digestion with NcoI, which site contains the predicted start codon and BglII, that cuts in the 3' untranslated region of the NDP52 cDNA. This fragment was cloned into the corresponding restriction sites of the *Escherichia coli* expression vector pQE60 (Stüber et al., 1990) to generate the construct pQEK52. This construct was introduced into the *E. coli* strain M15prep4. NDP52 expression was induced by the addition of 2 mM IPTG to the exponentially growing bacteria culture. After 3 h cells were pelleted and lysed in SDS sample buffer. Total cell lysate was analyzed in an immunoblot assay with mAb C8A2.

### Northern Blot Analysis

The transcription of NDP52 in human tissues was analyzed by hybridizing a commercially provided Multiple Tissue Northern Blot (Clontech Laboratories, Inc., Palo Alto, CA) with a <sup>32</sup>P-labeled PstI-fragment of the NDP52 cDNA (bp 1–1024). Each lane contains approximately 2  $\mu$ g of pu-

rified poly(A)<sup>+</sup> RNA from one of eight different human tissues standardized to contain the same amount of  $\beta$ -actin mRNA. After stringent washing with 0.2 $\times$  SSC/0.1% SDS for 1 h at 65°C, the blot was exposed to autoradiography for 1.5 h.

The transcription of NDP52 was further analyzed in the cell lines MG63, Saos-2, SK-ES-1, HeLa, COS-7 and human primary skin fibroblasts. Total RNA was prepared as described above. 15  $\mu$ g total RNA from each cell line were separated by agarose gel electrophoresis under denaturing conditions and blotted onto a nylon filter according to Sambrook et al. (1989). Hybridization was performed as described above.

### HSV-1 and Adenovirus 5 Infection

Wild-type HSV-1 strain 17 syn<sup>+</sup> was used in this study. The mutant strain FXE was derived from this strain and contains a deletion in the RING finger region of the immediate early transcription factor ICP0, which renders it ineffective for transactivation (Everett, 1989; Maul and Everett, 1994). All viruses were propagated and titrated in BHK C13 cells grown in modified Eagle's medium supplemented with 10% newborn calf serum. HEp-2 cells were infected with 10 plaque-forming units (pfu). Wild-type HSV-1 and HSV-1 FXE-infected cells were fixed after 5 h p.i. Parallel tests with mAb against ICP4 indicated that wild-type HSV-1 infected 95% and the FXE mutant 85% of the cells. Human fibroblasts (MRC-5) were infected with 20 pfu adenovirus 5 (Ad5) in medium without serum for 1 h after which serum was added to a final concentration of 10%. Cells were fixed at various times after infection and tested by immunofluorescence labeling with various antibodies.

## Results

### Isolation of a mAb Recognizing a ND10 Antigen

mAbs that recognize components of ND10 were raised in mice immunized with crude nuclei prepared from human osteosarcoma cells (MG63). Hybridomas were generated as described (Fazeka de St. Groth and Scheidegger, 1980) and screened by indirect immunofluorescence analysis for nuclear staining on MG63 cells. One mAb, C8A2, showed a typical dot-like nuclear staining. The number and size of these dots varied in MG63 cells, but  $\sim$ 10 dots per nucleus were observed. A similar staining pattern was also observed in the WI38 human fibroblasts and HOS TE85 human osteosarcoma cells whereas neither HeLa cells nor primary human skin fibroblasts showed any dot-like staining with mAb C8A2 (data not shown). The characteristic immunofluorescence staining pattern with the mAb C8A2 was very similar to that of ND10, indicating that the C8A2 antigen could be part of this structure. Cell extractions with NP-40, followed by DNaseI digestion and washing with 2.0 M NaCl (described in Ascoli and Maul, 1991) did not effect the immunofluorescence staining pattern, indi-

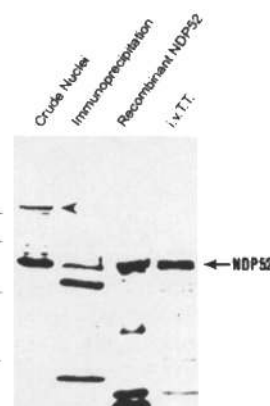
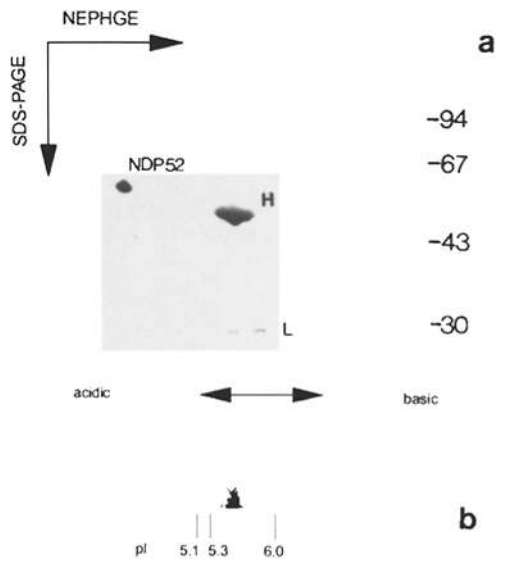


Figure 1. Analysis of NDP52 in SDS-PAGE. Lanes 1–3 show immunoblot detections of NDP52 with mAb C8A2 in crude nuclei fraction of MG63 cells (lane 1), immunoprecipitation with mAb C8A2 from MG63 cells (lane 2) and in total *E. coli* cell lysates expressing recombinant NDP52 (lane 3). Lane 4 shows a fluorography from an SDS-PAGE of the [<sup>35</sup>S] labeled in vitro translation (*i.v.T.T.*) product of the NDP52 cDNA (pFK52c).



**Figure 2.** pI analysis of NDP52 by NEPHGE and IEF. Immunoprecipitated NDP52 was separated by NEPHGE in the first dimension and by SDS-PAGE in the second dimension (a). Separated proteins were transferred to nitrocellulose and detected by mAb C8A2. Heavy and light chain of mAb C8A2 are indicated by H and L. For one-dimensional IEF, immunoprecipitated NDP52 was separated by isoelectric focussing and transferred onto nitrocellulose sheets. NDP52 was detected by a rabbit antiserum raised against a recombinant COOH-terminal fragment of NDP52 (b). The pIs of marker proteins are indicated below the blot strip. NDP52 has an pI of 5.5.

cating a stable association of the C8A2 antigen with the nuclear matrix (data not shown). Immunoblot analyses of the C8A2 antigen using crude nuclear fractions specifically identified a protein with an apparent molecular mass of 55 kD as detected using the ECL system (Fig. 1, lane 1), and a weaker signal at ~110 kD. Immunoprecipitation of the antigen with mAb C8A2 from whole cell lysates and subsequent immunoblotting also revealed a 55-kD protein (Fig. 1, lane 2). NEPHGE analysis with an immunoprecipitate of C8A2 (Fig. 2 a) showed that the antigen recognized by C8A2 was an acidic protein, with a pI of 5 to 6 with no detectable charge isomers. One-dimensional IEF with the immunoprecipitate of mAb C8A2 allowed a more precise determination of the pI at 5.5 (Fig. 2 b). Thus the antigen recognized by C8A2 is clearly distinct from the previously described NDP55, which has a pI of 7.4–7.7 with three charge isomers (Ascoli and Maul, 1991).

Because mAb C8A2 recognized the denatured antigen in immunoblot analysis, we used this mAb to screen a directional MG63 cDNA library constructed in  $\lambda$ -ZAP.

**Isolation of the cDNA Encoding the C8A2 Antigen**

Approximately  $6 \times 10^5$  recombinant phages of this library were screened for expression. Two positive phages were plaque-purified and subsequently transformed into Bluescript plasmids by in vivo excision. Restriction analysis revealed that both phages contained identical cDNA inserts of ~2.3 kb. One cDNA insert was truncated by ExoIII deletion cloning from both ends. The subclones generated were sequenced by cycle sequencing using Digoxygenin-

|   |      |
|---|------|
| ACTCTGCCCTGTTGCTGTCGCGCCGCTGCTGGTGTGCTGCTCCCTGGACCCCTACC    | 54   |
| ATGGAGGAGACCATCAAAGATCCCCCACATCAGCTGTCTTGTGGATCACTGTCTATTC  | 114  |
| M E E T I K D P P T S A V L L D H C H F                     | 20   |
| TCTCAGGTCACTTTAACAGTGTGGAGAAGTTCTACATCCCTGGAGGGGACGTACATGT  | 174  |
| S Q V I F N S V E K F Y I P G G D V T C                     | 40   |
| CATTATACCTTCACCCAGCATTTTCATCCCTCGTGAAGGATTGGATTGGCATCTTTAGA | 234  |
| H Y T F T Q H F I P R R K D W I G I F R                     | 60   |
| GTGGGGTGAAGACAACCCGTGAGTATTACACCTTCATGTGGTACTTTGCCATTGAC    | 294  |
| V G W K T T R E Y Y T F M W V T L P I D                     | 80   |
| CTAAACAACAATCAGCTAAACAGCAGGAAGTCCAATTCAAAGCTTACTACCTGCCAAG  | 354  |
| L N N K S A K Q Q E V Q F K A Y Y L P K                     | 100  |
| GATGATGAGTATTACAGTTCGCTATGTGGATGAGGATGGTGGTCCGGGGAGCAAGT    | 414  |
| D D E Y Y Q F C Y V D E D G V V R G A S                     | 120  |
| ATTCCTTCCAATCCGTCAGAAAATGAGGAAGACATCCCTGGTGTGTACCACTCAGGGA  | 474  |
| I P F Q F R P E N E E D I L V V T T Q G                     | 140  |
| GAGGTGAAGAGATTGAGCAGCACAACAGGAGCTTGAAGAAAACAGGAGTGAAG       | 534  |
| E V E E I E Q H N K E L C K E N Q E L K                     | 160  |
| GACAGCTGTATCAGCTCCAGAAGCAGAATCAGACATGAGGCTGAGCTCAAAGAAG     | 594  |
| D S C I S L Q K Q N S D M Q A E L Q K K                     | 180  |
| CAGGAGGAGTAGAAACCTACAGAGCATCAATAAGAAGTGAAGTGAAGTGAAGAA      | 654  |
| Q E E L E T L Q S I N K K L E L K V E K                     | 200  |
| CAGAAGGACTATTGGGAGCAGAGCTGCTTCAACTGAAAGAAAACCAAGAGATGTC     | 714  |
| Q K D Y W E T E L L Q L K E Q N Q K M S                     | 220  |
| TCAGAAAATGAGAAGATGGGAATCAGAGTGGATCAGCTTCAGGCCAGCTGTCACTCAA  | 774  |
| S E N E K M G I R V D Q L Q A L S T Q                       | 240  |
| GAGAAAAGAAATGGAGAAGCTTGTTCAGGAGATCAAGATAAGCAGAGCAGTATAGCAG  | 834  |
| E K E M E K L V Q G D Q D K T E Q L E Q                     | 260  |
| CTGAAAAGGAAAATGACCACCTTCTCAGTTAACTGAACAGGAAGGAGGACAGAA      | 894  |
| L K K E N D H L F L S L T E Q R K D Q K                     | 280  |
| AAGCTCGAGCAGCAGTGGAGCAAAATGAAGCAGAATGAACTACTGCAATGAAGAAACA  | 954  |
| K L E Q T V E Q M K Q N E T T A M K K Q                     | 300  |
| CAGGAATTAATGGATGAAACTTGCACCTGCAAAAAGCTGAGTGAGAGCAAAATTATA   | 1014 |
| Q E L M D E N F D L S K R L S E N E I I                     | 320  |
| TGTAATGCTCTGCAGAGACAGAAAGAGAGATTGGAAGGAGAAATGATCTTTGAAGAGG  | 1074 |
| C N A L Q R Q K E R L E G E N D L L K R                     | 340  |
| GAGAACAGCAGATTGCTCAGTTACATGGTCTGGATTTAATCTTTGCCGTATCAAGTA   | 1134 |
| E N S R L L S Y M G L D F N S L P Y Q V                     | 360  |
| CCTACTTCAGATGAAGGAGCGCAAGCAAAATCCAGGACTTGCCATGGAAAACCCATAT  | 1194 |
| P T S D E G G A R Q N P G L A Y G N P Y                     | 380  |
| TCTGGTATCCAAGAAAGTCTTCCCCAGCCGCTCCATCAAGAAATGCCATATCGC      | 1254 |
| S G I G A T E S S S P S P L S I K K C P I C                 | 400  |
| AAAGCAGATGATTTGTGATCACACCTTGGAGCAACAGCAGATGCAGCCCTTTGTTTC   | 1314 |
| K A D D I C D H T L E Q Q Q G P L C F                       | 420  |
| AATTGCTCAATTTGTGACAAGATCTTCCAGCTACAGAGAAGCAGATCTTTGAAGACC   | 1374 |
| N C P I C D K I F P A T E K Q I F E D H                     | 440  |
| GTGTTCTGCCACTCTC  | 1392 |
| V F C H S L   | 446  |
| TGAGTATCCAACCTCTTGGATGTATACAGAGATTTATAGAATAGAACCTATAGCTTCT  | 1452 |
| ACCATGAGTTATAGTCAAGATCCCTGCCTAACCTGAAATATTAGGGATTTACTCAGC   | 1512 |
| CCTGCTGCCGCTAACAGTGGAGTTATGTCAGTGTCAAGGTCACTGTTAAGGGCTTCT   | 1572 |
| GCTGCCATCTTGTGGTGTCTACCTTTAAGTCCGCAATCTAGCTGTATCATCCTCTC    | 1632 |
| ACCTGTCACTTCTGAGGGTCTCAGTACAAGGGCCCTGGATGGAGCAACCTGGGTAT    | 1692 |
| TCACAACAGGCTGACTGATACATAGTATTAGTTTCCAAGTGTCCCACTGCCATTC     | 1752 |
| AAAGTCAGCCCTTGAATGATTTGTTCTCAGTCTAACCCCTGGGGCAGAGATTTGGCCG  | 1812 |
| AGTTGAGAATTCCTCTCTCATCCTTGGTGTGCTTCTCCAATGATTTGTATGATA      | 1872 |
| CTAGCCAAAATGCGGTGGCAAGAGCTCAGAAAATCCAATTTGGATACCAAGGTTTCTC  | 1932 |
| ATGTTAATTTCTCAGCCCCAAAGAGCATCTTACTCTGACCTTACAGCAGGAAGTATT   | 1992 |
| GTTTCAGTCACAGAAAGCTTTTCTGGTACCTCTGGTTAGCAGCTTCTACTCTGATAT   | 2052 |
| TTCTATGTACATAGCTTTTATTGTTGTAATCCTTCTTAAATGGTTAAATAGGATTTT   | 2112 |
| AGCACTATGGTTTGCAGTTTCTGAGTAGGTGAGTTTGAATATGGGTAATCAGAA      | 2172 |
| AATGAGCACTTGTAACTCTTAACTAAATAAATAAATACTCTTCTTATTCAGGGA      | 2232 |
| CTTAGGTAATTTAAATAAACCCTTCAATTTATGGCTTCTGTTTGAAGCTCATGGAAA   | 2292 |
| ATTGTGATCAAAAGGGCTATGGGAAGGGCAGCCCGCAATGATTTCTCTCACCTGTG    | 2352 |
| TTAAGATTAATAAAAAAGAGTGTCTGGCAGTAAAAAAA                      | 2399 |

**Figure 3.** cDNA sequence and deduced amino acid sequence of NDP52. The part of the sequence isolated by the 5'RACE system is indicated by an arrow above the sequence. The predicted leucine zipper is marked by a dotted line. Conserved cysteine residues of the LIM-like domain are indicated by asterisks below the sequence. Potential polyadenylation sites are underlined. These sequence data are available from EMBL/Genbank/DBJ under accession number U22897.

labeled primers. Overlapping sequences were aligned to generate the complete cDNA sequence from both strands. The resulting cDNA sequence showed a continuous open reading frame (ORF) starting from the 5' end. The first predicted methionine in this sequence was not likely to be the translation initiation site *in vivo* since the predicted size of the encoded protein was significantly smaller (44 vs. 55 kD) than that indicated by immunoblotting. Therefore, the 5'RACE system was applied to generate a full-length cDNA as described in Materials and Methods. Two primers, GSP1 and -2, were chosen lying in the analyzed 5' region of the cDNA, with the nested GSP2 primer containing a natural, unique AatII restriction site. In agarose gel electrophoresis the 5'RACE-PCR product appeared as a single DNA band of ~220 bp. After digestion with AatII and SpeI (latter restriction site introduced by the anchor primer) the PCR product was ligated into the identically treated Bluescript plasmid carrying the C8A2 antigen cDNA library insert. To exclude PCR errors within the amplified product, 4 cDNA clones were analyzed by dideoxy sequencing. The sequence from the 3' end of the PCR product, comprising 110 bp, was identical with the already known 5' end of the previously isolated C8A2 antigen cDNA (Fig. 3, arrow above the sequence), confirming the specific amplification of the target cDNA. The 5' end of the PCR product contained an additional 90-bp sequence, which was continuously in-frame with the ORF.

The relatively small size of the PCR product, as well as its reproducibility, makes it unlikely that a truncation of the full-length product occurred due to incomplete first-stand synthesis. Therefore, the combined sequence of the 5'RACE PCR product and the cDNA library insert was considered to be the full-length cDNA of NDP52. The combined cDNA construct in pSK Bluescript, designated pFK52C, was subsequently used for *in vitro* transcription and translation of the encoded protein. Analyses of the [<sup>35</sup>S]methionine-labeled translation product by SDS-PAGE and subsequent fluorography (Fig. 1, lane 4) showed that the translation product was identical in size to that of the protein precipitated from MG63 cells by mAb C8A2, suggesting that the isolated cDNA encodes the full-length protein.

#### Recombinant Expression of the C8A2 Antigen in *E. coli*

To confirm that the isolated cDNA encodes the protein recognized by mAb C8A2, the complete predicted ORF of the cDNA was cloned into the *E. coli* expression vector pQE60 (Stüber et al., 1990), generating the construct pQEK52C. At the 5' end of the cDNA, a natural NcoI site containing the presumptive start ATG was used for cloning. Total cell lysates of induced *E. coli* cultures carrying the pQEK52C expression plasmid were subsequently analyzed by immunoblot analysis. The expressed protein was

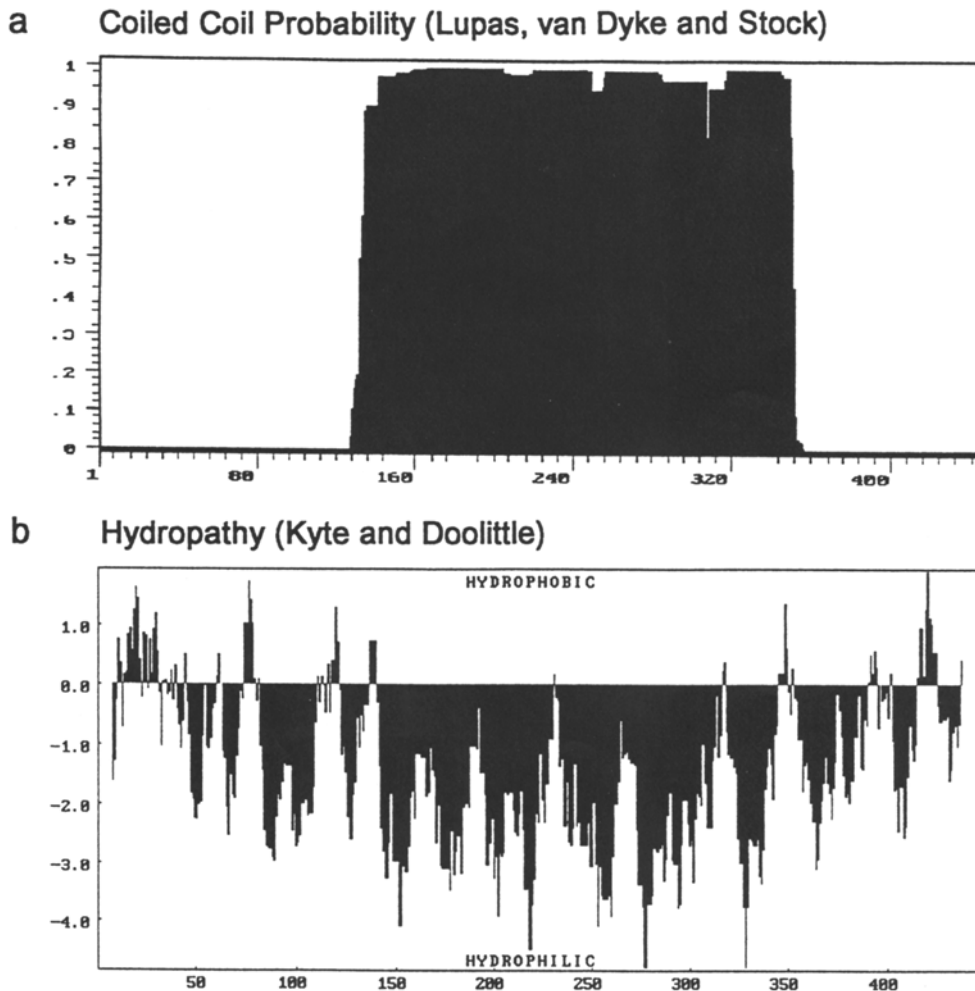


Figure 4. Secondary structure predictions of NDP52. The different values for secondary structure properties are indicated on the y axis plotted against the amino acid sequence on the x axis. The probability for coiled coil regions is shown in a. A hydropathy plot is shown in b.

specifically detected using mAb C8A2 (Fig. 1, lane 3) and showed the same apparent  $M_r$  as the protein precipitated from MG63 cells by mAb C8A2. These data clearly demonstrate that mAb C8A2 recognizes the protein encoded by the isolated cDNA. In addition a rabbit antiserum was raised against a purified recombinant protein encoded by the isolated cDNA, consisting of the last 164 COOH-terminal amino acids, which reacted identically as mAb C8A2 in immunofluorescence and Western blot analyses (data not shown).

### Sequence Analysis of the NDP52 cDNA

The cDNA is 2399 bp in length (Fig. 3), and contains an ORF from bp 55 to 1392 that encodes a protein with a predicted molecular mass of 52 kD. The slightly higher apparent mass of the native and recombinant protein in SDS-PAGE (Fig. 1) might be explained by stretches of acidic residues within the primary sequence, as described for several other proteins, e.g., Sp100 (Szosteki et al., 1990) and aspartyl  $\beta$ -hydroxylase (Korioth et al., 1994). Based on the predicted molecular mass and the localization, we named the encoded protein, nuclear dot protein 52 (NDP52). The presumptive start codon lies within the 5' region generated by the RACE system. Although the ORF continues to the 5' end of the cDNA, this ATG is likely to be the initiation codon since: (a) the recombinant protein expressed product in *E. coli*, using this ATG as the start codon and the protein immunoprecipitated from MG63 cells migrated to identical positions after SDS-PAGE (Fig. 1) and (b) the sequence (ACCATGG) surrounding the presumptive start ATG perfectly matches to the optimal sequence for eukaryotic initiation of translation (Kozak, 1986). Besides the polyadenylation signal 25 bp upstream of the poly-A tail, two more are found at bp 2108 and 2247 (Fig. 3). Homology searches in the current EMBL and GenBank databases using the Fast A algorithm showed that two expressed sequence tags (accession numbers T23881 and T07433) describe parts of our cDNA. There is no further strong homology with other sequences in the databases. However a slight homology of 51% over 700 bp (bp 500–1200) was observed with the cDNA for involucrin (AC: M25314), a marker of keratinocyte differentiation (Rice and Green, 1979; Eckert and Green, 1986). Homology of 57% over 250 bp (bp 480–730) was also observed with the cDNA of the myosin heavy chain within the region encoding the coiled coil tail of myosin.

### NDP52 Is Predicted to Have a Central Coiled Coil Domain

Hydropathy analysis of the predicted amino acid sequence of NDP52 shows a central hydrophilic region of  $\sim$ 200

amino acids (Fig. 4 b) with a high potential to form an  $\alpha$ -helical structure. This region (amino acids 134–350) has a high probability of forming a coiled coil structure as predicted by the coil algorithm of Lupas et al. (1991) (Fig. 4 a). The probability distribution predicts a sharply defined region for this coiled coil structure within NDP52. The absence of proline residues further supports this model. The amino acid sequence of this region shows a 20–30% homology with the coiled coil regions of different myosin and kinesin heavy chains. Another feature of this central region is a leucine zipper, ranging from amino acid 324–345, within the predicted coiled coil region. This motif, which is found in several DNA binding proteins, is thought to promote dimerization by forcing a coiled-coil like structure that contains the hydrophobic leucine residues internally (Busch and Sasone-Corsi, 1990). The distribution pattern of the leucine residues at position 1 of the heptamers, forming the coiled coil domain, extends with interruptions throughout the complete coiled coil region. Together, our data support the model of NDP52 homodimerization by an extended central coiled coil domain.

The small clusters of hydrophobic amino acids found in the NH<sub>2</sub>- and COOH-terminal regions flanking the central coiled coil domain (Fig. 4 b) have a higher probability of forming  $\beta$ -sheets and -turns, suggesting a nonhelical, globular structure for the NDP52 termini. Potential phosphorylation sites are found for cAMP- and cGMP-dependent protein kinases, PKC, casein kinase II and tyrosine kinase, although the biological significance of these sites is still unclear. NDP52, which was characterized as a nuclear protein, contains no bipartite or single cluster nuclear localization signal within the predicted amino acid sequence (Dingwall and Laskey, 1991).

Homology searches with the NH<sub>2</sub>- and COOH-terminal regions of NDP52 in the SWISS PROT database using the BLITZ algorithm revealed no significant homology for the NH<sub>2</sub> terminus. However, the last 50 amino acids of the COOH terminus, which contain 7 cysteine residues, show homology to oocyte and gastrula zinc finger proteins, and especially to proteins containing a LIM domain (Sadler et al., 1992). This domain, which was originally identified in a subset of transcription factors, is characterized by one histidine and seven cysteine residues conserved within a stretch of  $\sim$ 60 amino acids. Analysis of 20 LIM domain sequences has defined the consensus sequence: CX<sub>2</sub>CX<sub>16-23</sub>HX<sub>2</sub>CX<sub>2</sub>C<sub>2</sub>CX<sub>16-21</sub>CX<sub>2-3</sub>(C,H,D) (Sadler et al., 1992). Alignment of the NDP52 COOH terminus with LIM domains of homologous proteins (Fig. 5) shows that six of the seven cysteine residues of the NDP52 COOH terminus (asterisks above the sequence) perfectly match with those of the LIM domain. In addition to the conserved cysteine and histidine residues, the NDP52 COOH terminus shows

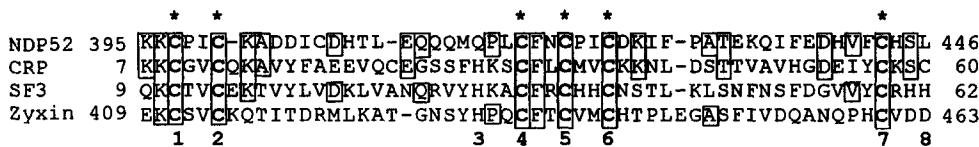


Figure 5. Alignment of the NDP52 COOH terminus with LIM domains. The residues of the LIM domain consensus sequence are indicated by numbers below the

sequences. Conserved residues of the NDP52 COOH terminus are indicated by asterisks above the sequence. LIM domains from the following proteins are shown: cysteine-rich protein (CRP) (Sadler et al., 1994), Pollen specific protein SF3 (SF3) (Baltz et al., 1992), and Zyxin (Sadler et al., 1992). All residues matching with NDP52 are shown in boxes.

## Structural Elements of NDP52



Figure 6. Schematic drawing of the NDP52 structure. NDP52 has a central region with a high probability of forming a coiled coil (CC) structure. At the COOH terminus of this region a leucine zipper motif (LZ) is found. The COOH terminus of NDP52 contains a LIM like domain (LIM-L) with six conserved cysteine residues of the LIM domain consensus sequence.

other amino acid matches that are conserved in some LIM proteins, including lysine residues flanking the CX2C site and a phenylalanine within the second CX2C site. To date the biological significance of these residues is unknown.

In summary, analyses of the predicted amino acid sequence of NDP52 gives evidence for three structural regions of this protein (Fig. 6). A central coiled coil domain with an integrated leucine zipper probably causes homodimerization of NDP52. The flanking terminal regions appear to be globular, with a LIM-like domain at the utmost COOH terminus.

### NDP52 Is Transcribed in Several Human Tissues and Cell Lines

The transcription of NDP52 in human tissues was analyzed by Northern blotting with a radiolabeled PstI-fragment of the NDP52 cDNA (bp 1–1024). In all human tissues analyzed, transcription of NDP52 was detected, although at variable levels (Fig. 7). Brain showed the lowest transcription level (only visible after longer exposure), whereas skeletal muscle showed the highest level. The size of the major transcript (2.5 kb) fits reasonably well with the size of the NDP52 cDNA lacking a poly(A)<sup>+</sup> tail (2.4 kb). The relative broad signal size of this major transcript can be explained by two additional polyadenylation sites found in the last 200 bp of the 3' untranslated region (Fig. 3) that might also be used for termination of transcription. Furthermore, transcripts larger than 2.5 kb were detected in all tissues analyzed. Because our immunoblot data provided no evidence for a larger isoform of NDP52 (Fig. 1), these larger transcripts are likely to be due to polyadenylation signals downstream of those contained in the isolated

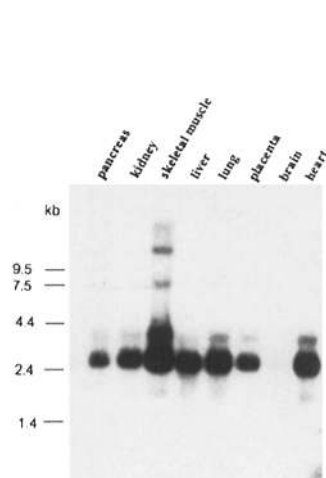


Figure 7. Northern blot analysis of NDP52 transcription in eight human tissues. A Northern blot, containing 2.5  $\mu$ g poly(A)<sup>+</sup> RNA per lane of the respective tissue, separated by electrophoresis under denaturing conditions was hybridized with a <sup>32</sup>P-labeled PstI fragment of the NDP52 cDNA. After stringent washing the blot was exposed to autoradiography. NDP52 is transcribed in all human tissues analyzed, showing multiple transcript lengths. The major transcript has a length of 2.5 kb.

cDNA.

Further analysis of NDP52 transcription in different cell lines by Northern blotting revealed transcription in the human cell lines MG63, SK-ES-1 and HeLa as well as in COS-7 cells (Fig. 8), but not in primary human skin fibroblasts even after longer exposure. The surprising absence of any NDP52 transcript, larger than the major 2.5-kb species in all cell lines presumably reflects a difference between transformed cell lines and primary tissue cells in transcription termination or processing.

### NDP52 Is a Component of the Nuclear Domain ND10

Previously, we showed that PML, Sp100 and NDP55 colocalize to precisely circumscribed nuclear domains (Dyck et al., 1994; Maul and Everett, 1994). Since NDP52 also appears at highest concentrations in nuclear domains similar to ND10 and in similar numbers, we tested for colocalization with these proteins. As shown in Fig. 9 a, PML and NDP52 colocalize precisely at these domains, except that NDP52 is also diffusely present throughout the nucleus at a low level and, in some cells, in domains reminiscent of splicing factors (Spector 1993). A human serum (no. 1785) containing autoantibodies against antigens that colocalize with the splicing factor SC35 was used in double labeling with the mouse mAb C8A2 (Fig 9 b); although other ND10-associated proteins were clearly not present in regions, containing splicing components, these regions appeared yellow, indicating colocalization of NDP52 in some cells. Thus, unlike all other ND10-associated proteins that are not specifically enriched in regions containing splicing components, NDP52 is present in more than one nuclear domain.

### HSV-1 and Adenovirus Cause Different Redistributions of NDP52

HSV-1 removes ND10 from the nucleus, and the immediate early gene product type-1 (ICP0) is responsible for this redistribution of ND10-associated proteins (Maul et al., 1993). Analysis to determine the effect of HSV-1 infection on NDP52 distribution indicated the apparent absence of both PML and NDP52 labeling of ND10 at 5 h after infection, although NDP52 remained faintly in areas shown to be domains, containing splicing components (Fig. 9 c). Since not all cells are infected, the area chosen for microscopic documentation contained a noninfected cell, providing a convenient control for successful double labeling.

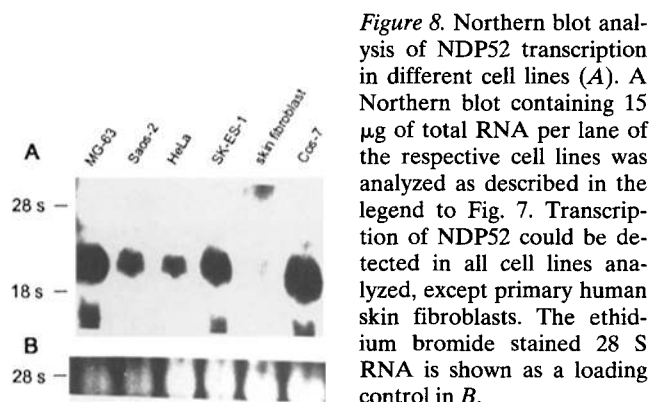


Figure 8. Northern blot analysis of NDP52 transcription in different cell lines (A). A Northern blot containing 15  $\mu$ g of total RNA per lane of the respective cell lines was analyzed as described in the legend to Fig. 7. Transcription of NDP52 could be detected in all cell lines analyzed, except primary human skin fibroblasts. The ethidium bromide stained 28 S RNA is shown as a loading control in B.

In cells infected with the HSV-1 mutant FXE, which contains a deletion at the RING finger region essential for removal of ND10 (Maul and Everett, 1994; Everett and Maul, 1994), ND10 labeling with ND10-associated proteins remained during early stages of infection, but NDP52 was not recognizable in the immunofluorescence assay of ND10 (Fig. 9 *d*). The images obtained for NDP52 were the same as with wild-type virus. In HSV-1-infected cells, NDP52 reactivity was similar to that of the other ND10 antigens, except that its removal from ND10 does not appear to be dependent on ICP0.

Unlike HSV-1, Ad 5, another DNA virus that replicates in the nucleus, segregates ND10-associated proteins into tracks 2–3  $\mu\text{m}$  in length during the early stages of infection (5–9 h p.i.) (G. Maul, personal communication). Double labeling of NDP52 with PML indicated that both proteins segregate to these tracks (Fig. 9 *e*). However, NDP52 was removed from these tracks later during infection and accumulated in the outer rim adjacent to the viral replication domains (Fig. 9 *f*). Viral replication domains were characterized by labeling with a mAb against single-stranded DNA binding protein (Pombo et al., 1994; data not shown).

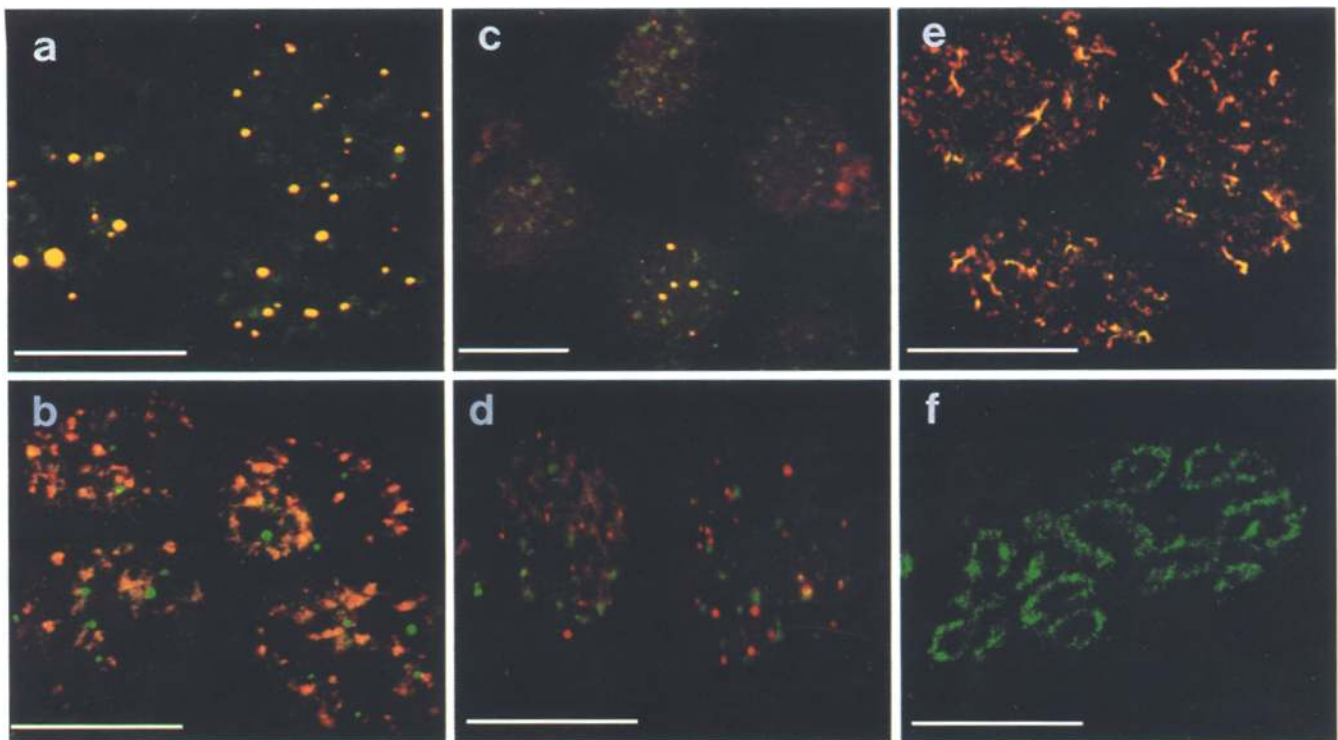
### IFN Treatment Induces NDP52

Type I and II IFN have been shown to increase the number and size of nuclear dots containing the ND10 protein Sp100

(Guldner et al., 1992). Analogous treatment of MG63 cells with 1,000 U/ml of either IFN- $\beta$  or - $\gamma$  for 18 h resulted in a similar increase in number and size of NDP52-specific dots (Fig. 10, *a* and *b*). Furthermore, some cells also showed a clear cytoplasmic localization of NDP52 after IFN treatment (Fig. 10, *c* and *d*). To investigate the effects of IFN on NDP52 transcription, a Northern blot containing identical amounts of total RNA from IFN- $\gamma$ -stimulated and unstimulated MG63 cells was hybridized as described above. As shown in Fig. 11, IFN- $\gamma$  treatment led to an enhanced transcription of NDP52, suggesting that the IFN- $\gamma$ -induced increase in NDP52 immunofluorescence staining reflects enhanced expression of the corresponding protein.

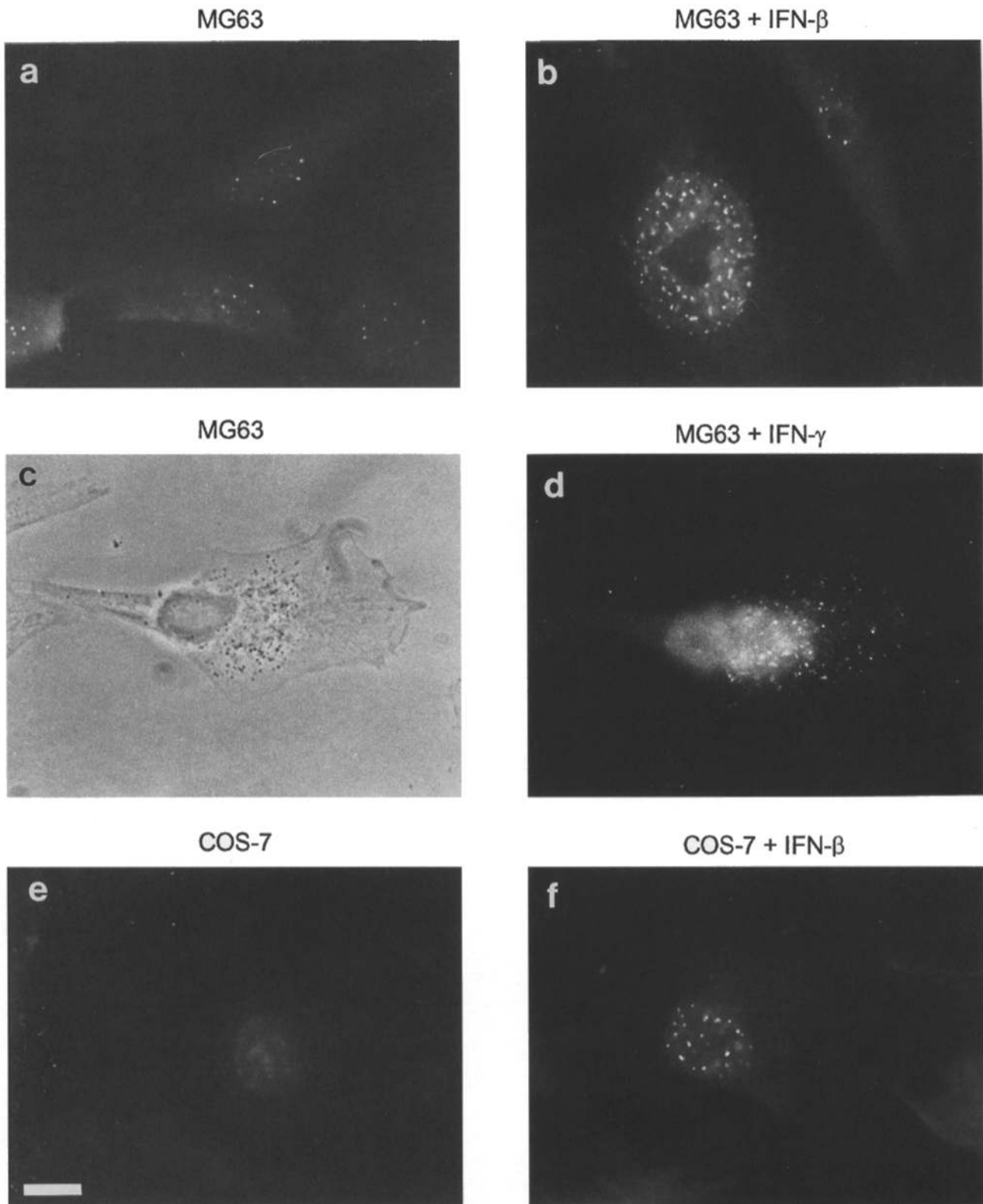
### IFN Treatment Reorganizes NDP52 into ND10 in COS Cells

To analyze the conservation of the NDP52 epitope recognized by the mAb C8A2, we used this antibody for immunoprecipitation of NDP52 from COS-7 cells, which were previously shown to transcribe the corresponding gene (Fig. 8). mAb C8A2 immunoprecipitated NDP52 from COS-7 cells in amounts comparable to those from MG63 cells (not shown). However, in immunofluorescence analyses COS-7 cells did not show the typical nuclear dot staining, but instead a diffuse staining of the complete nucleus (Fig. 10 *e*). In light of the relatively close evolutionary



**Figure 9.** NDP52 colocalized with PML and is redistributed upon virus infection. All micrographs were obtained as single optical sections by laser confocal microscopy. Those double labeled were scanned simultaneously for both labels using pseudocolors so as to reflect the fluorescent dye, i.e., green for FITC and red for Texas red. (*a*) HEP-2 cells double labeled for PML and NDP52 resulting in yellow due to perfect overlapping stain. (*b*) MRC-5 fibroblasts labeled with human serum 1785 recognizing splicing factors (*red*) and C8A2 (*green*). (*c*) HEP-2 cells 2 h after HSV-1 infection double labeled with anti PML antibodies (*red*) and C8A2 (*green*). ND10 are absent except in one uninfected cell. (*d*) HEP-2 cells 5 h after infection with the HSV-1 mutant FXE double labeled with anti PML antibodies (*red*) and C8A2 (*green*). (*e*) MRC-5 cells infected with Ad5 for 8 h and double labeled with anti PML antibodies (*red*) and mAb C8A2 (*green*) resulting in an overlapping image of fibrous tracks. (*f*) MRC-5 cells infected with Ad5 for 16 h and labeled with mAb C8A2 showing the outer rim of Ad5 replication domains. Bar, 10  $\mu\text{m}$ .





**Figure 10.** Interferon treatment of MG63 and COS-7 cells. (a) Immunofluorescence analysis of MG63 cells shows a typical dot like staining in the nucleus. (b) Treatment with IFN- $\beta$  and - $\gamma$  leads to an increase in number and size of NDP52-specific dots. INF- $\gamma$  treatment of MG63 cells partially leads to a redistribution of NDP52 to the cytoplasm (d) as shown by comparison with phase contrast of the same cell (c). (e) COS-7 cells show a diffuse fluorescence staining of the nucleus with mAb C8A2. (f) Treatment with IFN- $\beta$  however, leads in  $\sim$ 10% of the cells to the appearance of NDP52-specific nuclear dots. Bar, 10  $\mu$ m.

proximity of human and primates, we analyzed the effect of the generally species-specific human IFN- $\beta$  on NDP52 distribution in COS-7 cells. Immunofluorescence analysis revealed a general increase in nuclear staining and also NDP52-specific dots of varying intensity in  $\sim 10\%$  of the cells (Fig. 10 f). Identical results were gained with HeLa cells, which showed without stimulation a diffuse nuclear distribution of NDP52, and NDP52-specific dots after treatment with IFN- $\beta$  (data not shown). Thus NDP52 is not only distributed in ND10, but in some cell lines also in a diffuse nuclear localization from which it can be reorganized to ND10 by treatment with IFN- $\beta$ .

## Discussion

The nuclear domain ND10, also described as POD or Kr bodies, has been characterized using a variety of human autoimmune sera and monoclonal antibodies (Ascoli and Maul, 1991; Stuurman et al., 1992; Xie et al., 1993; Dyck et al., 1994; Koken et al., 1994; Weis et al., 1994). The fact that several different antibodies recognize this structure suggests that ND10 is a multiprotein complex. Thus far the cDNA sequences for two ND10 proteins, Sp100 and PML, have been characterized (Szostecki et al., 1990; Kazikuka et al., 1991; de The et al., 1991). A direct interaction of PML and Sp100 could not be demonstrated (Koken et al., 1994), suggesting, that other structures or proteins are required for their association in colocalization experiments.

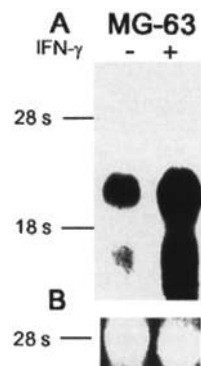
Here we describe the isolation and characterization of a cDNA encoding the novel ND10-associated protein, NDP52. The mAb, C8A2, which specifically recognizes NDP52, was generated in mice immunized with crude nuclei fractions from the human osteosarcoma cell line MG63. Immunofluorescence analysis of MG63 cells revealed a typical nuclear dot staining with mAb C8A2 that was very similar to the ND10 staining described by Ascoli and Maul (1991). Subsequent colocalization experiments with ND10-specific antibodies confirmed that mAb C8A2 recognizes a component of this structure. Immunoblot analyses of crude nuclei fractions used for immunization and for immunoprecipitation revealed that mAb C8A2 specifically recognizes a protein with an apparent molecular mass of 55 kD. NEPGHE analyses and IEF of the immunoprecipitated antigen of C8A2 characterized it as an acidic protein with a pI of 5.5 without charge isoforms. The combined use of expression screening and the 5' RACE system enabled the isolation of a cDNA encoding

the antigen recognized by mAb C8A2. Based on its localization in immunofluorescence analyses and the predicted molecular mass we designated this antigen nuclear dot protein (NDP) 52. The isolated cDNA was characterized by recombinant expression in *E. coli* and specific detection of the recombinant protein by mAb C8A2 in immunoblot analysis of total cell extracts. A rabbit antiserum raised against the recombinant COOH terminus of the encoded protein yielded identical results in immunofluorescence and immunoblot analyses to those obtained with mAb C8A2 (data not shown).

The protein encoded by the NDP52 cDNA consists of 446 amino acids with a predicted mass of 52 kD. Secondary structure analysis of the amino acid sequence predicts an extended coiled domain (Lupas et al., 1991) in the central region of the NDP52 protein consisting of heptamers with hydrophobic residues at positions 1 and 4. This region also shows homology to parts of the myosin and kinesin heavy chains, which are known to have a coiled coil structure in vivo. Thus, the observed homology is probably due to the structural requirements for a coiled coil domain. A leucine zipper motif is found at the COOH-terminal end of the predicted coiled coil region. Leucine zippers are known to promote homo- and heterodimerization in several transcription factors, probably by forming a coiled coil-like structure (Struhl, 1989). However, the clusters of basic amino acids that act as DNA binding domains in transcription factors and that are usually located adjacent to the leucine zipper are not found in the NDP52 sequence. Consistent with these predictions is the presence of an additional protein band of  $\sim 110$  kD detected with mAb C8A2 in immunoblot analysis of nuclei fractions. Based on the considerations above, this band likely represents unseparated homodimers. Further investigations will determine whether such homo- or heterodimerization occurs in vivo. Coiled coil domains described for several proteins form homodimers that usually aggregate to filaments (Cohen and Parry, 1990). The colocalization of several proteins within ND10 probably also requires aggregation of these proteins by still unknown interactions. Interestingly, coiled coil regions are also predicted for the two genetically characterized ND10-associated proteins, PML and Sp100 (Koken et al., 1994), suggesting that these domains may lead to an aggregation of these and as yet uncharacterized proteins to form the complex ND10 structure.

In contrast to the hydrophilic central region, the flanking NH<sub>2</sub>- and COOH-terminal regions of NDP52 show a relatively high potential to form  $\beta$ -sheets and -turns. These regions might therefore have a more globular structure compared to the extended central coiled coil region. The overall predicted structure of NDP52 i.e., a large central coiled coil domain flanked by nonhelical terminal domains appears to be quite similar to that of intermediate filament proteins.

The predicted amino acid sequence of NDP52 contains no clear nuclear location signal. Therefore, the nuclear localization can be explained by an unregulated diffusion through the nuclear pore complex (Peters, 1986) as it would be possible for a protein of 52 kD or by the existence of a nuclear location signal within the tertiary structure of NDP52. Nuclear localization of NDP52 could also



**Figure 11.** IFN- $\gamma$  treatment of MG63 cells leads to an increase in NDP52 transcription. Northern blot was performed as described in legend to Fig. 7. An autoradiography of the hybridization of identical amounts of total RNA from unstimulated (-) and stimulated (+) MG63 cells with a <sup>32</sup>P-labeled NDP52 cDNA fragment is shown in A. The ethidium bromide stained 28 S RNA is shown in B.

be explained by binding to a protein in the cytoplasm, which is targeted to the nucleus.

Sequence comparison revealed a weak homology of NDP52 with involucrin, a marker of terminal keratinocyte differentiation (Rice and Green, 1979; Eckert and Green, 1986), which comprises the central coiled coil region of NDP52 and extends over 700 bp (51% identity) including an encoded stretch of 200 amino acid (22% identity). This homology at both the gene and protein levels raises the possibility of a common evolutionary origin of these genes. Remarkably, the function of involucrin is crosslinking plasma membrane proteins via covalent  $\epsilon$ -( $\gamma$ -glutamyl)lysine bonds (Simon and Green, 1988) mediated by a transglutaminase, and several of the involved glutamine residues in involucrin are conserved in NDP52. Thus NDP52 and involucrin might share a similar function.

Homology searches with the nonhelical NH<sub>2</sub>- and COOH-terminal domains alone revealed that the cysteine rich COOH terminus of NDP52 showed homology to several members of the LIM domain protein family (Crawford et al., 1992). The LIM domain was shown to bind two Zn<sup>2+</sup> ions that are tetrahedrally coordinated at four cysteine and three cysteine plus one histidine sites (Kosa et al., 1994). Alignment of NDP52 with LIM domains of several proteins shows a perfect match between six cysteine residues of NDP52 and those of the LIM domain consensus sequence. Besides the conserved residues of the LIM domain, lysine residues flanking the CX<sub>2</sub>C sites as well as one phenylalanine residue within this site are conserved in NDP52 and some of the LIM domain proteins. Based on these homologies, the COOH terminus of NDP52 can be considered a LIM-like domain. In determining the biological significance of this LIM-like domain, it will be important to investigate whether the natural or the *E. coli* derived recombinant NDP52 actually contains bound zinc. Recent studies have confirmed that LIM domains mediate protein-protein interactions (Schmeichel and Beckerle, 1994; Wadman et al., 1994). A LIM-like domain in NDP52 that mediates similar protein-protein interactions would make this protein a likely candidate for accumulating itself and other proteins within the nuclear dots of ND10. Alternatively, NDP52 could be responsible for their specific localization within the nucleus by mediating interactions with the nuclear matrix.

Transcription of *NDP52* was detected *in vivo* in all human tissues analyzed. The variations observed in transcription levels suggest that the function of the encoded protein is not tissue specific. *NDP52* transcripts were detectable after 1.5-h autoradiography exposure of the hybridized Northern blot, indicating high *in vivo* transcription levels of this gene. Of all tissues and cells analyzed, human primary skin fibroblasts were NDP52-negative in Northern- and immunofluorescence analyses. Therefore, despite its expression in most human tissues and cell lines, NDP52 may be not essential for normal cell growth and division.

Virus replication often leads to a destruction of the nuclear organization as a consequence of the expanding space the viral replication domains occupy. However, ND10 is modified much earlier during virus infection (Maul et al., 1993; Maul and Everett, 1994), indicating, that the associated proteins may have some functions in early events

of the viral life cycle. NDP52 is eliminated from ND10 by HSV-1. In early stages of infection ICP0, one of the immediate early transactivating HSV-1 proteins, transiently colocalizes with PML and Sp100. This colocalization is also found with NDP52, except that it is removed from ND10 even if a HSV-1 mutant with a nonfunctional ICP0 is used for infection. Therefore, NDP52 is removed independently of other ND10-associated proteins and distributed throughout the nucleus with an increase of staining found in regions containing splicing factors. Although the significance of this observation is not clear it demonstrates that the colocalization of ND10-associated proteins can be abrogated by virus infection, in the course of which they may exert their different specific functions.

Ad5, a DNA virus from another family modifies ND10 in a different way. During infection with Ad5, PML is segregated to fibrous tracks together with Sp100 and NDP52 (not shown). NDP52 also transiently moves to these tracks but like Sp100 it is later recruited to the viral replication domains. The specific distribution of NDP52 in the outer rim of the developing Ad5 replication domains suggests that it localizes with the sites of viral transcription, that have recently been demonstrated by the incorporation of bromo-UTP (Pombo et al., 1994). NDP52 remains at the outer rim of the expanding replication domain of Ad5 and is finally recognized at the nuclear envelope in a distribution similar to that observed for snRNPs. Together with the findings that NDP52 distributes to some degree to areas enriched in splicing factors we may speculate that it is likely to be involved in viral transcription or posttranscriptional processing. In contrast to HSV-1, that has no late genes that require splicing, Ad5 is dependent on the host cell splicing apparatus for the processing and maturation of viral transcripts (Pombo et al., 1994). It has recently been shown, that splicing occurs at sites of transcription and not in nuclear speckles, containing splicing factors (Zhang et al., 1994). Although these domains therefore rather appear as storage or accumulation sites for excess splicing factors, the colocalization of NDP52 with these factors still supports the possibility that it might be involved in splicing processes.

Treatment of MG63 cells with IFN led to an increase in number and size of NDP52-specific dots, and to an increase of the NDP52 transcript levels. These results suggest an IFN-induced enhanced expression of NDP52, comparable to that described for the ND10 protein Sp100 (Guldner et al., 1992). In addition to their colocalization, the very similar IFN regulation of NDP52 and Sp100 argues for a shared function that might depend on a direct interaction in a multiprotein complex. However, IFN treatment had an additional effect on NDP52, inducing a cytoplasmic distribution of the protein in some cells. A similar cytoplasmic localization of ND10 proteins was recently described for PML (Everett and Maul, 1994). Relocation of PML to the cytoplasm is thought to be caused by interactions with a nuclear targeting deficient ICP0. Shutling between the nucleus and cytoplasm has been described for several proteins (reviewed in Laskey and Dingwall, 1993). In contrast to the highly regulated nuclear import of proteins, export from the nucleus was recently demonstrated to be independent of specific export signals, and is instead a consequence of lost retention (Schmidt-

Zachmann et al., 1993). This, in turn, suggests that proteins are released from the nucleus by processes, that reduce the affinity to nuclear structures. Such processes leading to the cytoplasmic relocation of ND10 proteins might comprise interactions with IFN-induced or viral proteins as well as IFN-induced modification of ND10 proteins. Interestingly, the two different in vitro inducers of relocation of ND10 proteins to the cytoplasm, i.e., IFN treatment and virus infection, reflect an identical in vivo situation; organisms respond to virus infection with production of IFNs, which induce a variety of proteins with antiviral actions (Samuel, 1991). Thus, IFN-induced enhanced expression of ND10 proteins as well as their observed relocation to the cytoplasm might be part of a common antiviral function.

Immunofluorescence analysis of COS-7 cells, that transcribe and express NDP52 revealed no dot like staining using mAb C8A2 and instead, a diffuse nuclear distribution of NDP52. IFN- $\beta$  treatment induces the relocation of NDP52 into ND10. This effect was also observed in HeLa cells (not shown). In contrast to the diffuse nuclear distribution of NDP52 in COS-7 and HeLa cells before IFN treatment (not shown), mAb138 detected the ND10 structure as nuclear dots in these cells. However, in all other cell lines investigated, the antigens recognized by both mAbs were already colocalized in ND10 without any stimulation. Therefore, redistribution of NDP52 to ND10 can either be explained, by the induction of other proteins in COS-7 and HeLa cells interacting with NDP52 to form this structure or by an IFN- $\beta$  induced modification of NDP52 increasing its affinity to ND10. Studies of IFN-induced gene expression and NDP52 modification in NDP52-positive cells without dot structures will help us to test this hypothesis.

The NDP52 protein characterized here is a novel component of ND10. In addition to its colocalization with the ND10 proteins Sp100 and PML, NDP52 also shares structural and regulatory features with them. The results for NDP52 support a model of ND10 as a multiprotein complex, whose components might function synergistically. Previously reported data on ND10 proteins and those for NDP52 suggest the involvement of ND10 proteins in the viral life cycle. The growing number of genetically characterized ND10 proteins will allow a more detailed analysis of their presumed function as well as their involvement in APL.

This work was supported by the Deutsche Forschungsgemeinschaft SFB 223, the Stifterverband für die deutsche Wissenschaft, Dr. Hilmer Stiftung, project no. T 146/02.001/91-93 and National Institutes of Health Core grant CA-10815.

Received for publication 24 January 1995 and in revised form 27 March 1995.

## References

- Antonioni, M., M. Carmo-Fonseca, J. Ferreira, and A. I. Lamond. 1993. Nuclear organization of splicing snRNPs during differentiation of MEL cells in vitro. *J. Cell Biol.* 123:1055-1068.
- Ascoli, C. A., and G. G. Maul. 1991. Identification of a novel nuclear domain. *J. Cell Biol.* 112:785-795.
- Baltz, R., C. Domon, D. T. N. Pillay, and A. Steinmetz. 1992. Characterization of a pollen specific cDNA from sunflower encoding a zinc finger protein. *Plant J.* 2:713-721.
- Bridge, E., M. Carmo-Fonseca, A. I. Lamond, and U. Pettersson. 1993. Nuclear organization of splicing snRNPs in adenovirus infected cells. *J. Virol.* 67:5792-5802.
- Busch, J. B., and P. Sasone-Corsi. 1990. Dimers, leucine zipper and DNA-binding domains. *Trends Genet.* 6:36-40.
- Castaigne, S., C. Chomienne, M. T. Daniel, P. Ballerini, R. Berger, P. Fenaut, and L. Degos. 1990. All-trans retinoic acid as a differentiation therapy for acute promyelocytic leukemia: I Clinical results. *Blood.* 76:1704-1717.
- Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* 18:5294-5299.
- Chou P. Y., and G. D. Fasman. 1978. Prediction of the secondary structure of proteins from the amino acid sequence. *Adv. Enzymol. Relat. Areas Mol. Biol.* 47:45-147.
- Cohen, C., and D. A. D. Parry. 1990.  $\alpha$ -Helical coiled coils and bundles: how to design an  $\alpha$ -helical protein. *Prot. Struct. Funct. Genet.* 7:1-15.
- Crawford, A. W., J. W. Michelsen, and M. C. Beckerle. 1992. An interaction between zyxin and  $\alpha$ -actinin. *J. Cell Biol.* 116:1381-1393.
- de The, H., C. Lavau, A. Marchio, C. Chomienne, L. Degos, and A. Dejean. 1991. The PML-RAR $\alpha$  fusion mRNA generated by the t(15;17) translocation in acute promyelocytic leukemia encodes a functionally altered RAR. *Cell.* 66:675-684.
- Dingwall, C., and R. A. Laskey. 1991. Nuclear targeting sequences - a consensus? *Trends Biochem. Sci.* 16:478-481.
- Dyck, J. A., G. G. Maul, W. H. Miller, J. D. Chen, A. Kakizuka, and R. M. Evans. 1994. A novel macromolecular structure is a target of the promyelocyte-retinoic acid receptor oncoprotein. 1994. *Cell.* 76:333-343.
- Eckert, R. L., and H. Green. 1986. Structure and evolution of the human involucrin gene. *Cell.* 46:583-589.
- Everett, R. D. 1989. Construction and characterization of herpes simplex virus type I mutants with defined lesions in immediate early gene 1. *J. Gen. Virol.* 70:1185-1202.
- Everett, R. D., and G. G. Maul. 1994. HSV-1 IE protein Vmw110 causes redistribution of PML. *EMBO (Eur. Mol. Biol. Organ.) J.* 13:5062-5069.
- Fakan, S., G. Leser, and T. E. Martin. 1984. Ultrastructural distribution of nuclear ribonucleoproteins as visualized by immunocytochemistry on thin sections. *J. Cell Biol.* 98:358-363.
- Fazekas de St. Groth, S., and D. Scheidegger. 1980. Production of monoclonal antibodies: strategy and tactics. *J. Immunol. Methods.* 35:1-24.
- Fu, X. D., and T. Maniatis. 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature (Lond.)* 343:437-441.
- Goddard, A. D., J. Borrow, P. S. Freemont, and E. Solomon. 1991. Characterization of a zinc finger gene disrupted by the t(15;17) in acute promyelocytic leukemia. *Science (Wash. DC).* 254:1371-1374.
- Guldner, H. H., C. Szostecki, T. Groetzinger, and H. Will. 1992. IFN enhance expression of Sp100, an autoantigen in primary biliary cirrhosis. *J. Immunol.* 149:4067-4073.
- Kakizuka, A., W. H. Miller, Jr., K. Umeson, R. P. Warrell, Jr., S. R. Frankel, V. V. S. Murty, E. Dmitrovsky, and R. M. Evans. 1991. Chromosomal translocation t(15;17) in human acute promyelocytic leukemia fuses RAR $\alpha$  with a novel putative transcription factor, PML. *Cell.* 66:663-674.
- Kastner, P., A. Perez, Y. Lutz, C. Rochette-Egly, M. P. Gaub, B. Durand, M. Lanotte, R. Berger, and P. Chambon. 1992. Structure, localization, and transcriptional properties of two classes of retinoic acid receptor  $\alpha$  fusion proteins in acute promyelocytic leukemia (APL): structure similarities with a new family of oncoproteins. *EMBO (Eur. Mol. Biol. Organ.) J.* 11:629-642.
- Koken, M. H. M., F. Puvion-Dutilleul, M. C. Guillemain, A. Viron, G. Linares-Cruz, N. Stuurman, L. de Jong, C. Szostecki, F. Calvo, C. Chomienne et al. 1994. The t(15;17) translocation alters a nuclear body in a retinoic acid-reversible fashion. *EMBO (Eur. Mol. Biol. Organ.) J.* 13:1073-1083.
- Kornoth, F., C. Gieffers, and J. Frey. 1994. Cloning and characterization of the human gene encoding aspartyl  $\beta$ -hydroxylase. *Gene (Amst.)* 150:395-399.
- Kosa, J. L., J. W. Michelsen, H. A. Louis, J. I. Olsen, D. R. Davis, M. C. Beckerle, and D. R. Winge. 1994. Common metal ion coordination in LIM domain proteins. *Biochemistry.* 33:468-477.
- Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell.* 44:283-292.
- Kyte J., and R. F. Doolittle. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157:105-132.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.)* 227:680-685.
- Laskey, R. A., and C. Dingwall. 1993. Nuclear shuttling: The default pathway for nuclear proteins. *Cell.* 74:585-586.
- Lupas, A., M. Van Dyke, and J. Stock. 1991. Predicting coiled coils from protein sequences. *Science (Wash. DC).* 252:1162-1164.
- Martelli, A. M., R. S. Gilmour, V. Bertagnolo, L. M. Neri, L. Manzoli, and L. Cocco. 1992. Nuclear localization and signalling activity of phosphoinositidase C in Swiss 3T3 cells. *Nature (Lond.)* 358:242-245.
- Maul, G. G., H. H. Guldner, and J. G. Spivack. 1993. Modification of discrete nuclear domains induced by herpes simplex virus type 1 immediate early gene product (ICP0). *J. Gen. Virol.* 74:2679-2690.
- Maul, G. G., and R. D. Everett. 1994. The nuclear location of PML, a cellular member of the C3HC4 zinc-binding domain protein family, is rearranged during herpes simplex virus infection by the C3HC4 viral protein ICP0. *J. Gen. Virol.* 75:1223-1233.

- O'Farrell, P. Z., H. M. Goodman, and P. H. O'Farrell. 1977. High resolution two dimensional electrophoresis of basic as well as acidic proteins. *Cell*. 36: 83-96.
- Pearson, W. R. 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* 101:202-211.
- Peters, R. 1986. Fluorescence microphotolysis to measure nucleocytoplasmic transport and intracellular mobility. *Biochim. Biophys. Acta.* 864:305-359.
- Pombo, A., J. Ferreira, E. Bridge, and M. Carmo-Fonseca. 1994. Adenovirus replication and transcription sites are spatially separated in the nucleus of infected cells. *EMBO (Eur. Mol. Biol. Organ.) J.* 13:5075-5085.
- Puvion, E. A., A. Viron, C. Assens, E. H. Leduc, and P. Jeanteur. 1984. Immunocytochemical identification of nuclear structures containing snRNPs in isolated rat liver cells. *J. Ultrastruct. Res.* 87:180-189.
- Rice, R. H., and H. Green. 1979. Presence in human epidermal cells of a soluble protein precursor of the cross-linked envelope: activation of the cross-linking by calcium ions. *Cell*. 18:681-694.
- Sadler, I., A. W. Crawford, J. W. Michelsen, and M. C. Beckerle. 1992. Zyxin and cCRP: Two interactive LIM domain proteins associated with the cytoskeleton. *J. Cell Biol.* 119:1573-1587.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*. Second Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 7.37-7.45.
- Samuel, C. E. 1991. Antiviral actions of interferon. Interferon-regulated cellular proteins and their surprisingly selective antiviral activities. *Virology*. 183: 1-11.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463-5467.
- Schmeichel, K. L., and M. C. Beckerle. 1994. The LIM domain is a modular protein-binding interface. *Cell*. 79:211-219.
- Schmidt-Zachmann, M. S., C. Dargemont, L. C. Kuhn, and E. A. Nigg. 1993. Nuclear export of proteins: The role of nuclear retention. *Cell*. 74:493-504.
- Simon, M., and H. Green. 1988. The glutamine residues reactive in transglutaminase-catalyzed cross-linking of involucrin. *J. Biol. Chem.* 263:18093-18098.
- Smith, T. F., and M. S. Waterman. 1981. Identification of common molecular subsequences. *J. Mol. Biol.* 147:195-197.
- Spector, D. L., W. H. Schrier, and H. Bush. 1983. Immunoelectron microscopic localization of snRNPs. *Biol. Cell.* 49:1-10.
- Spector, D. L. 1993. Macromolecular domains within the cell nucleus. *Annu. Rev. Cell Biol.* 9:265-315.
- Struhl, K. 1989. Helix-turn-helix, zinc finger, and leucine zipper motifs for eukaryotic transcriptional regulatory proteins. *Trends Biochem. Sci.* 14:137-140.
- Stüber, D., H. Natile, and G. Garotta. 1990. System for high-level production in *E. coli* and rapid purification of recombinant proteins. *Immunol. Methods.* 4: 121-152.
- Stuurman, N., A. de Graaf, A. Floore, A. Jossen, B. Humbel, L. de Jong, and R. van Driel. 1992. A monoclonal antibody recognizing nuclear matrix-associated nuclear bodies. *J. Cell Science.* 101:773-784.
- Szostecki, C., H. H. Guldner, H. J. Netter, and H. Will. 1990. Isolation and characterization of a cDNA encoding a human nuclear antigen predominantly recognized by autoantibodies from patients with primary biliary cirrhosis. *J. Immunol.* 145:4338-4347.
- Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. USA.* 79:4350-4354.
- Visa, N., F. Puvion-Dutilleul, J.-P. Bachellerie, and E. Puvion. 1993. Intracellular distribution of U1 and U2 snRNA visualized by high resolution in situ hybridization: revelation of a novel compartment containing U1 but not U2 snRNA in HeLa cells. *Eur. J. Cell Biol.* 60:308-321.
- Wadman, I., J. Li, R. O. Bash, A. Forster, H. Osada, T. H. Rabbitts, and R. Baer. 1994. Specific in vivo association between the bHLH and LIM proteins implicated in human T cell leukemia. *EMBO (Eur. Mol. Biol. Organ.) J.* 13: 4831-4839.
- Weis, K., S. Rambaud, C. Lavau, J. Jansen, T. Carvalho, M. Carmo-Fonseca, A. Lamond, and A. Dejean. 1994. Retinoic acid regulates aberrant nuclear localization of PML-RAR $\alpha$  in acute promyelotic leukemia cells. *Cell*. 76:345-356.
- Xie, K., E. J. Lambie, and M. Snyder. 1993. Nuclear dot antigens may specify transcriptional domains in the nucleus. *Mol. Cell. Biol.* 13:6170-6179.
- Zhang, G., K. L. Taneja, R. H. Singer and M. R. Green. 1994. Localization of pre-mRNA splicing in mammalian nuclei. *Nature (Lond.)*. 372:809-812.