MAN1, an Inner Nuclear Membrane Protein That Shares the LEM Domain with Lamina-associated Polypeptide 2 and Emerin*

(Received for publication, August 24, 1999, and in revised form, October 26, 1999)

Feng Lin‡, Deborah L. Blake§, Isabelle Callebaut¶, Ilona S. Skerjanc**, Lars Holmer‡, Michael W. McBurney§§, Micheline Paulin-Levasseur¶¶, and Howard J. Worman††

From the ‡Departments of Medicine and of Anatomy and Cell Biology, College of Physicians and Surgeons, Columbia University, New York, New York 10032, the §Department of Biology, University of Ottawa, Ottawa, Ontario KIN 6N5, Canada, ¶Systèmes Moléculaires et Biologie Structurale, LMCP, CNRS UMR 7590, Université Paris 6 et 7, 4 Place Jussieu, 75252 Paris Cedex 05, France, **Department of Biochemistry, University of Western Ontario, London, Ontario N6A 5C1, Canada, and ¶¶Ottawa Regional Cancer Centre and Department of Medicine, University of Ottawa, Ottawa, Ontario KIN 6N5, Canada

The “MAN antigens” are polypeptides recognized by autoantibodies from a patient with a collagen vascular disease and localized to the nuclear envelope. We now show that one of the human MAN antigens termed MAN1 is a 82.3-kDa protein with an amino-terminal domain followed by two hydrophobic segments and a carboxy-terminal tail. The MAN1 gene contains seven protein-coding exons and is assigned to human chromosome 12q14. Its mRNA is approximately 5.5 kilobases and is detected in several different cell types that were examined. Cell extraction experiments show that MAN1 is an integral membrane protein. When expressed in transfected cells, MAN1 is exclusively targeted to the nuclear envelope, consistent with an inner nuclear membrane localization. Protein sequence analysis reveals that MAN1 shares a conserved globular domain of approximately 40 amino acids, which we term the LEM module, with inner nuclear membrane proteins lamina-associated polypeptide 2 and emerin. The LEM module is also present in two proteins of Caenorhabditis elegans. These results show that MAN1 is an integral protein of the inner nuclear membrane that shares the LEM module with other proteins of this subcellular localization.

Only a few proteins other than the nuclear lamins, some with various isoforms, have been localized to the inner nuclear membrane during interphase. The first to be identified was avian lamin B receptor (LBR)† (1, 2). LBR was subsequently characterized in mammals (3–6), and an immunochromosomally cross-reactive protein has been identified in sea urchins (7). LBR has a nucleoplasmic, amino-terminal domain of approximately 200 amino acids that binds to B-type lamins and chromatin proteins and confers inner nuclear membrane retention (2, 4, 8–12). The amino-terminal domain of LBR is followed by a hydrophobic domain with eight putative transmembrane segments that is similar in sequence to sterol reductases, including two human proteins of the endoplasmic reticulum, one of which is a 7-dehydrocholesterol reductase (13).

Two other integral proteins of the inner nuclear membrane have been termed lamina-associated polypeptides (LAPs). Three related isoforms of rat LAP1 (LAP1A, LAP1B, and LAP1C) were identified by reaction with a single monoclonal antibody and shown to be integral membrane proteins associated with the nuclear lamina (14). LAP1C has a nucleoplasmic amino-terminal domain followed by one transmembrane segment, and the other LAP1 isoforms are probably of similar overall structure, arising from the same gene by alternative RNA splicing (15). LAP2 was also first identified by reaction with a monoclonal antibody and shown to be an integral membrane protein that binds to nuclear lamins and chromatin (16). Several isoforms of LAP2, which have also been called thymopoietins because they were thought to possibly be thymocyte growth factors, are generated by alternative RNA splicing (17–21). Some of the LAP2 isoforms are integral proteins of the inner nuclear membrane, with nucleoplasmic amino-terminal domains and single transmembrane segments, while others are nonmembrane nucleolar proteins.

In 1994, positional cloning of the gene responsible for X-linked Emery-Dreifuss muscular dystrophy led to the identification of Emerin that has limited sequence similarity to LAP2 (22). After its discovery, Emerin was shown to be localized to the inner nuclear membrane and lacking from this location in most patients with Emery-Dreifuss muscular dystrophy (23, 24). Emerin contains a nucleoplasmic, amino-terminal domain followed by a single transmembrane segment and a short tail in the perinuclear space/endoplasmic reticulum lumen. The nucleoplasmic domain of Emerin mediates its inner nuclear membrane targeting (25).

Recently, a multispanning membrane protein without a large hydrophilic domain called nurim was shown to be localized to the nuclear envelope (26). Nurim is present exclusively

* This work was supported by National Institutes of Health (NIH) Grant RO1-CA66974 to H. J. W. The confocal microscopy facility was established by NIH Grant 1S10 RR10506 and supported by NIH Grant 5 P30 CA13696 as part of the Herbert Irving Cancer Center at Columbia University. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF112299, AF112300, and AF180135–AF180142.

† Supported in part by the Centre National de la Recherche Scientifique (CNRS) program “Physique et Chimie du Vivant.”

‡‡ Supported by a Medical Research Council of Canada scholarship (development grant).

§§ Supported by a grant from the Medical Research Council of Canada.

¶¶ Supported in part as an Irma T. Hirschl Scholar and by a Greater New York Hepatology Primary Biliary Cirrhosis Seed Grant from the American Liver Foundation. To whom correspondence should be addressed: Dept. of Medicine, College of Physicians and Surgeons, Columbia University, 630 W. 168th St., 10th Floor, Rm. 508, New York, NY 10032. Tel.: 212-305-8156; Fax: 212-305-6443; E-mail: hjw14@columbia.edu.

† The abbreviations used are: LBR, lamin B receptor; LAP, lamina-associated polypeptide; GST, glutathione S-transferase; HCA, hydrophobic cluster analysis.

4840 This paper is available on line at http://www.jbc.org
in the nuclear envelope and highly immobilized in this location, strongly suggesting that it too is an inner nuclear membrane protein. It appears that multiple regions of unrin, including its transmembrane domains, contain determinants for its targeting (26).

Otefin is a protein that has been localized to the inner nuclear membrane in Drosophila (27). The deduced primary structure of otefin derived from cDNA cloning and sequencing demonstrates that it is a mostly hydrophilic protein with a hydrophobic stretch of amino acids near its carboxyl terminus (28). Although sequence analysis suggests that otefin may be an integral membrane protein, biochemical evidence indicates that it is a peripheral one and that its hydrophobic tail plays a role in its membrane association (29).

Additional proteins of the inner nuclear membrane probably exist but remain to be characterized. Some patients with rheumatic and autoimmune liver diseases have autoantibodies that recognize nuclear envelope antigens, and these antibodies can potentially be useful tools in the identification of novel nuclear membrane proteins (30, 31). Recently, autoantibodies were described in a patient with an ill-defined collagen vascular disease and shown to potentially be useful tools in the identification of novel nuclear membrane antigens that are localized exclusively to the nuclear envelope and co-fractionate biochemically with nuclear lamins (32). The three MAN antigens have apparent molecular masses of approximately 80, 58, and 40 kDa. The MAN antigen of approximately 38 kDa appears to be LAP2-β, but the other two proteins have not been characterized at the molecular level (33). We now report on the molecular characterization of one of the human MAN antigens, which we call MAN1.

### EXPERIMENTAL PROCEDURES

**P19 Cell Expression Library Screening**—An expression library was constructed of cDNA from mouse P19 cells as described (34). P19 cells were maintained in Eagle’s minimal essential medium (Life Technologies, Inc.) supplemented with 10% fetal bovine serum (Life Technologies). RNA was isolated by LiCl extraction 6 days after treatment of cells with dimethyl sulfoxide, and poly(A)-RNA was selected by standard methods (35). cDNA was synthesized using an oligo(dT) Nol1 primer-adapter and the Superscript Lambda System Kit (Life Technologies). SalI adapters were ligated to the ends, and the cDNA was subsequently subjected to restriction endonuclease digestion with NolI and size-fractionated by column chromatography. The size-fractionated cDNA was ligated into bacteriophage λ gt22, which was digested with NolI and SalI, and packaged using the Life Technologies, Inc. Lambda Packaging System (Life Technologies). Standard methods (35) were used for screening the expression library with human serum (1.29:10,000 dilution) that contained the MAN autoantibodies (36). This led to isolation of 17 cross-hybridizing cDNA clones, the longest of which was a HeLa cell cDNA of approximately 4 kilobases (9 kb, top) as a probe (second line). This led to isolation of 17 cross-hybridizing cDNA clones, the longest of which was a HeLa cell cDNA of approximately 4 kilobases (9 kb, top) as a probe (second line). A 1.6-kilobase 5′ fragment of the HeLa cell cDNA was then used to screen approximately 1 × 10^9 recombinant clones of a human leukocyte cDNA library, and 20 cross-hybridizing clones were isolated, the largest being approximately 4.7 kilobases (third line). This cDNA contained an open reading frame encoding a polypeptide of 754 amino acids with a calculated molecular mass of approximately 82,000 Da, similar to the molecular mass of the largest of the MAN antigens, which we called MAN1. We then screened a human genomic library to obtain genomic clones and sequenced a 1.8-kilobase fragment of clone GM-6 (see Fig. 3) that cross-hybridized with a 0.15-kilobase 5′ fragment of the leukocyte cDNA clone. Regions between the arrows were sequenced on both strands. The assigned translation initiation codon of MAN1 (ATG) and the location of the polyadenylate tail (+ poly A) in the cDNA clones are indicated.

**Library Screening by Hybridization**—The cDNA insert of the mouse λ gt22 clone was 32P-labeled using either the Multiprime Labeling Kit (Amersham Pharmacia Biotech, Quebec, Canada) or Random Priming System Kit (New England Biolabs, Beverly, MA) for hybridization screening of the same library and an undifferentiated P19 cell library in λ Uni-Zap XR (Stratagene, La Jolla, CA). The 32P-labeled mouse cDNA was also used to screen a human HeLa cell (S3) cDNA library (Stratagene). To isolate overlapping cDNA clones, various cDNA fragments were similarly 32P-labeled and used to screen a human leukocyte 5′-stretch plus cDNA library (CLONTECH, Palo Alto, CA) and a human whole blood λ DASH II genomic library (Stratagene). All hybridization screening was performed using standard methods (35) at 42 °C in 6× SSPE, 5× Denhardt’s solution, 0.1% SDS, 50% formamide, and 0.1 mg/ml salmon sperm DNA with probe-specific activities between 4 × 10^8 and 1 × 10^9 Bq/ml. Nitrocellulose filters (Schleicher and Schuell) were washed at 68 °C in 1× SSC and 0.1% SDS and exposed to XAR-5 film (Eastman Kodak, Inc.) with an intensifying screen at ~80 °C for 12–16 h. Hybridizing clones were isolated from libraries by repeated screenings, and cDNA inserts or fragments were isolated and subcloned for sequencing using standard methods (35). Restriction mapping of genomic clones in bacteriophage λ was performed according to Rackwitz et al. (36).

**DNA Sequencing**—DNA sequencing was performed on a 373A Auto-
tions (37) were performed on a RoboCycler Gradient 96 thermocycler (Stratagene) with forward primer of sequence 5'-CTTTTCTCTTCAGCTCCGCCGT-3' and reverse primer of sequence 5'-CTTTTCTCTTCAGCTCCGCCGT-3'. Reaction parameters were initial denaturation at 95 °C for 5 min and then 35 cycles with denaturation at 95 °C for 45 s, annealing at 55 °C for 45 s, and extension at 72 °C for 1 min followed by a final extension for 7 min at 72 °C. Reaction products were analyzed by electrophoresis on 2% agarose gels. Chromosomal assignment was performed through the Whitehead Institute/MIT World Wide Web server at with LOD score linkage set at a value of 15.

**Northern Blotting—**Complementary DNA probes were 32P-labeled using the Random Priming System Kit, and Northern blotting was performed as described (35) using human tissue and cancer cell line blots (CLONTECH).

**MAN1 Fusion Protein Expression—**A plasmid to express a MAN1 glutathione S-transferase (GST) fusion protein was constructed using the pGEX-4T-1 vector (Amersham Pharmacia Biotech). To amplify cDNA encoding MAN1 from amino acid 1 to 360, polymerase chain reaction was performed using the Gene Amp Kit (Hoffmann-La Roche) and a PCR System 2400 thermocycler (Perkin-Elmer) with MAN1 cDNA as template. Oligonucleotide primers for polymerase chain reaction were designed to contain BamHI and EcoRI restriction endonuclease sites for cloning into pGEX-4T-1. Amplified polymerase chain reaction products were purified using the PCR Purification Kit (Promega, Madison, WI) and ligated into pGEX-4T-1 that had been digested with the same restriction endonucleases. Ligation reaction mixtures were used to transform competent *Escherichia coli* strain DH5α (Life Technologies), and the recombinant plasmids were confirmed by DNA sequencing. GST fusion proteins were produced and purified by glutathione-Sepharose (Amersham Pharmacia Biotech) chromatography as described elsewhere (38).

**Electrophoresis and Immunoblotting—**SDS-polyacrylamide gel electrophoresis was performed under reducing conditions according to Laemmli (39). Methods used for immunoblotting were those that have been described previously (3, 32). Antibodies were affinity-purified from serum against fusion proteins according to the method of Smith and Fisher (40).

**Expression of Epitope-tagged Protein in Transfected Cells—**To construct a plasmid encoding full-length MAN1 with a FLAG epitope tag at its amino terminus, polymerase chain reaction was performed with primers containing engineered EcoRI and SstI restriction endonuclease sites with a λ cDNA clone of MAN1 as template. The amplified product was digested with EcoRI and SstI and ligated in-frame into pBPT4, a pBluescript II KS (Stratagene)-based plasmid containing a Kozak consensus sequence, ATG, and the FLAG epitope coding sequence 5' to the multiple cloning site. The resulting recombinant construct in pBPT4, which contained a SpeI site 5' to the Kozak consensus sequence and XhoI site 3' to the multiple cloning site, was digested with SpeI and XhoI to isolate a DNA that was then ligated into pSVK3 (Amersham Pharmacia Biotech) that was digested with XhoI and XhoI (an isoschizomere of SpeI). COS-7 cells, grown in Dulbecco's modified Eagle's medium (Life Technologies) supplemented with 10% fetal bovine serum and 2 mM glutamine (Life Technologies), were transfected with the recombinant pSVK3 construct, which encoded MAN1 with a FLAG epitope tag at its amino terminus under the control of an SV40 early promoter, as described previously (8, 9).

**Extraction of Transfected Cells—**COS-7 cells were transfected as described above and harvested by low speed centrifugation. For NaOH extraction, cells were resuspended in a solution of 0.1 M NaOH, 10 mM dithiothreitol, and 0.5 mM phenylmethylsulfonyl fluoride. The suspension was then centrifuged at 100,000 × g at 4 °C for 60 min in a TL-100 centrifuge (Beckman Coulter, Fullerton, CA) to obtain supernatant and pellet fractions. For urea extraction, cells were resuspended in a solution of 20 mM Tris-HCl (pH 8.8), 25 mM NaCl, 8 M urea, 1 mM dithiothreitol, and 0.5 mM phenylmethylsulfonyl fluoride. After a 30-min incubation at room temperature, the solution was centrifuged at 100,000 × g at 18 °C for 60 min. Proteins in the supernatant fractions were precipitated with trichloroacetic acid and solubilized in 40 μl of SDS sample buffer. Pellets were solubilized directly in 40 μl of SDS sample buffer. For analysis, 10 μl of each sample were loaded on

![Fig. 3. Structural organization and chromosomal assignment of the human MAN1 gene.](image)

**FIG. 3.** Structural organization and chromosomal assignment of the human MAN1 gene. A, λ clones GM-6 and GM-9 were isolated from a human genomic DNA library and used to determine the structural organization of the MAN1 gene. Positions of restriction endonuclease sites for *XbaI* (X) and *HindIII* (H) are indicated. GM-6 and GM-9 are vertically aligned with a diagram of the MAN1 gene shown below. Exons are numbered 1–10. Exons 1–7 contained protein coding regions, which are black. The 5' part of exon 1, 3' part of exon 7, and exons 8–10 encoded untranslated sequences (white). Nucleotide sequences for MAN1 have been deposited in the GenBank™ data base under accession numbers AF180135–AF180142. Genomic clones were not isolated for all of intron 2, and a gap is shown in the diagram for the missing region. B, partial framework map of human chromosome 12q14 containing the MAN1 gene. The analysis showed the closest linkage of the MAN1 gene to marker WI-15930 and placed it 2.84 centiRays from marker AFMA066X99. Distances at the left are in centiRays.

![Fig. 4. Expression of MAN1 mRNA in human tissues and cell lines.](image)

**FIG. 4.** Expression of MAN1 mRNA in human tissues and cell lines. A 3.7-kilobase fragment of MAN1 cDNA was 32P-labeled and used to probe Northern blots of poly(A)+-enriched RNA from human tissues (left panel) and cancer cell lines (right panel). Roughly equal amounts of RNA were loaded in each lane as ascertained by probing of the same blots with a 32P-labeled β-actin cDNA probe (data not shown). For tissues, RNA was from heart (lane 1), brain (lane 2), placenta (lane 3), lung (lane 4), liver (lane 5), skeletal muscle (lane 6), kidney (lane 7), and pancreas (lane 8). For cell lines, RNA was from HL-60 promyelocytic leukemia (lane 1), HeLa S3 (lane 2), K-562 chronic myelogenous leukemia (lane 3), MOLT-4 lymphoblastic leukemia (lane 4), Raji Burkitt lymphoma (lane 5), SW480 colorectal adenocarcinoma (lane 6), A549 lung carcinoma (lane 7), and G361 melanoma (lane 8) cells. Migrations of molecular mass standards in kilobases (kb) are shown at the left.
polycrylamide slab gels for electrophoresis under reducing conditions (39). Proteins in slab gels were transferred to nitrocellulose for immuno blotting with either mouse monoclonal anti-FLAG M2 antibodies (Sigma) or rabbit antibodies against LBR (7).

Confocal Immunofluorescence Microscopy of Transfected Cells—Transfected COS-7 cells were cultured, fixed, and prepared for immuno fluorescence microscopy as described previously (8, 9). Fluorescein isothiocyanate-conjugated anti-mouse immunoglobulin secondary antibodies were obtained from Jackson ImmunoResearch Laboratories (West Grove, PA) and used to detect mouse monoclonal anti-FLAG antibodies. Microscopy was performed on a Zeiss Axiosvert 100TV inverted microscope (Carl Zeiss). Images were processed using Photoshop software (Adobe Systems, San Jose, CA) on a Macintosh G3 computer (Apple Computer, Cupertino, CA).

Protein Sequence Analysis—The nucleotide and deduced amino acid sequences of MAN1 were compared with protein sequences in the GenBank™ data base (32). However, there are domains within MAN1 common to other proteins (see below). There are over 40 human EST sequences identical to portions of MAN1 and numerous expressed sequence tags with similar sequences from other organisms. From nucleotide 829 to 1077 (A of translation initiation codon) of the cDNA did not contain an in-frame stop codon and was presumably a partial cDNA for a novel protein. Repeated screenings of mouse P19 cDNA libraries with the 32P-labeled partial mouse cDNA did not lead to the isolation of additional cDNA clones with overlapping sequences. We therefore screened human DNA libraries to obtain overlapping clones that contained the complete coding sequence of human MAN1. The complete protein coding sequence was contained in one cDNA clone, and additional 5′ sequence was determined from one genomic DNA clone (Fig. 1).

The deduced amino acid sequence of human MAN1 is shown in Fig. 2A. From the initiation ATG to the stop codon, the protein contains 754 amino acids. The calculated molecular mass of MAN1 is 82,312 Da, and the theoretical pi is 6.18. The initiation ATG was assigned based on the fact that the deduced molecular mass of the encoded protein is virtually the same as the apparent molecular mass of the largest MAN antigen determined by SDS-polyacrylamide gel electrophoresis. The next ATG in the 3′ direction is 264 nucleotides away in the cDNA sequence, which would result in a protein approximately 10 kDa smaller. This was an in-frame stop codon 135 nucleotides 5′ to the assigned initiation ATG.

Over its entire length, MAN1 shows no significant sequence similarity to other proteins of known function in the translated GenBank™ data base. However, there are domains within MAN1 common to other proteins (see below). There are over 40 human EST sequences identical to portions of MAN1 and numerous expressed sequence tags with similar sequences from other organisms. From nucleotide 829 to 1077 (A of translation initiation codon designated as nucleotide 1), the nucleotide sequence of MAN1 is 100% identical to the sequence of a previously reported CpG island (46). CpG islands are short stretches of DNA containing a high density of nonmethylated

RESULTS

cDNA Cloning and Primary Structure of Human MAN1—The MAN antigens are three proteins with molecular masses of approximately 80, 58, and 40 kDa that are localized to the nuclear envelope and recognized by autoantibodies in serum (MAN antiserum) from a patient with a collagen vascular disease (32). Immunelectron and immunofluorescence microscopy have demonstrated that the MAN antigens are exclusively localized to the nuclear envelope, most likely the inner nuclear membrane (32). We used unfractionated MAN antiserum to screen approximately 500,000 recombinant clones of a mouse P19 cell expression library, which led to the isolation of only one cDNA clone (deposited in the GenBank™ data base under accession number AF112300) of 997 nucleotides with an open reading frame that encoded a novel polypeptide of 331 amino acids. The isolated cDNA did not contain an in-frame stop codon and was presumably a partial cDNA for a novel protein. Antibodies affinity-purified from the MAN antiserum against a β-galactosidase fusion protein containing the polypeptide expressed from this cDNA recognized proteins in HeLa and P19 cell extracts that corresponded to the proteins recognized by unpurified antiserum and exclusively labeled the nuclear envelope when examined by immunofluorescence microscopy (data not shown). The isolated cDNA clone therefore very likely encoded a portion of one of the mouse MAN antigens or a highly related protein. Repeated screenings of mouse P19 cDNA libraries with the 32P-labeled partial mouse cDNA did not lead to the isolation of additional cDNA clones with overlapping sequences. We therefore screened human DNA libraries to obtain overlapping clones that contained the complete coding sequence of human MAN1. The complete protein coding sequence was contained in one cDNA clone, and additional 5′ sequence was determined from one genomic DNA clone (Fig. 1).

The deduced amino acid sequence of human MAN1 is shown in Fig. 2A. From the initiation ATG to the stop codon, the protein contains 754 amino acids. The calculated molecular mass of MAN1 is 82,312 Da, and the theoretical pi is 6.18. The initiation ATG was assigned based on the fact that the deduced molecular mass of the encoded protein is virtually the same as the apparent molecular mass of the largest MAN antigen determined by SDS-polyacrylamide gel electrophoresis. The next ATG in the 3′ direction is 264 nucleotides away in the cDNA sequence, which would result in a protein approximately 10 kDa smaller. This was an in-frame stop codon 135 nucleotides 5′ to the assigned initiation ATG.

Over its entire length, MAN1 shows no significant sequence similarity to other proteins of known function in the translated GenBank™ data base. However, there are domains within MAN1 common to other proteins (see below). There are over 40 human EST sequences identical to portions of MAN1 and numerous expressed sequence tags with similar sequences from other organisms. From nucleotide 829 to 1077 (A of translation initiation codon designated as nucleotide 1), the nucleotide sequence of MAN1 is 100% identical to the sequence of a previously reported CpG island (46). CpG islands are short stretches of DNA containing a high density of nonmethylated

Fig. 6. Extraction of MAN1 and LBR from cells with NaOH and urea. Cells were transfected to express FLAG-tagged MAN1 and extracted with 8 M urea or 0.1 M NaOH. Suspensions were fractionated by centrifugation into pellets (P) and supernatants (S). Proteins in collected fractions were separated by electrophoresis on 10% SDS-polyacrylamide gels and transferred to nitrocellulose for immunoblotting with anti-FLAG antibodies (top) or anti-LBR antibodies (bottom). Migration of molecular mass standards in kDa is indicated at the left.

Fig. 5. The MAN autoantibodies recognize a GST fusion protein containing the first 360 amino acids of MAN1, and antibodies affinity-purified against this fusion protein recognize the MAN antigens in HeLa cell extracts. Left (Coomassie), a Coomassie Blue-stained polycrylamide slab gel of proteins in HeLa cell extracts (lane 1) and a GST-MAN1 fusion protein and lower molecular mass probable degradation products (lane 2). Middle (Serum), autoradiogram of immunoblot in which the proteins in HeLa cell extracts (lane 1) and GST-MAN1 fusion protein (lane 2) were probed with MAN antiserum (diluted 1:10,000). Right (Antibody), autoradiogram of immunoblot in which proteins in HeLa cell extracts (lane 1) and GST-MAN1 fusion protein (lane 2) were probed with antibodies affinity-purified from MAN antiserum against the GST-MAN1 fusion protein. These antibodies recognize proteins in the HeLa cell extracts of the same molecular mass as the autoantibodies in the MAN antiserum. Migration of molecular mass standards in kDa are indicated at the left.
and cell lines examined appeared to have MAN1 mRNA; however, the amount varied. Steady state MAN1 mRNA was noticeably low in A549 lung carcinoma cells, while, in the tissues examined, placenta appeared to have the highest amount. The fact that there was only one major band recognized by MAN1 cDNA on Northern blots suggests that no major isoforms arise by alternative RNA splicing.

The Amino-terminal Domain of MAN1 Is Recognized by MAN Antiserum—To confirm that the protein encoded by the MAN1 cDNA is actually a MAN antigen recognized by the MAN antiserum, we expressed the first 360 amino acids of MAN1 as a GST fusion protein. We affinity-purified antibodies from the serum that bound to this fusion protein and performed immunoblotting to determine if the purified antibodies recognized any of the MAN antigens in HeLa cell lysates (Fig. 5). Antibodies in the MAN antiserum recognized two major polypeptides on immunoblots of proteins from whole HeLa cell lysates of approximately 82 and 60 kDa; the MAN antigen with an apparent molecular mass of approximately 40 kDa was not strongly recognized in these whole cell lysates as in nuclear envelopes (32). The MAN antiserum also recognized the GST-MAN1 fusion protein and a probable degradation product. Antibodies that were affinity-purified from the serum against the GST-MAN1 fusion protein recognized proteins with the same molecular masses that were recognized by the unpurified MAN antiserum. Recognition of the GST-MAN1 fusion protein by the MAN antiserum confirms that the isolated cDNA actually encodes the 82-kDa MAN antigen. Recognition of both the 82- and 60-kDa MAN antigens by the affinity-purified antibodies suggests that they share a common epitope(s).

MAN1 Fractionates as an Integral Membrane Protein—To confirm that MAN1 is an integral membrane protein, we examined its fractionation into urea and NaOH. Upon extraction of HeLa cell nuclear envelopes with 0.1 M NaOH or 8 M urea, a portion of the 82-kDa reactive species remained in the membrane fractions; however, some was also present in the extracts (data not shown). This may have occurred because the MAN antiserum recognizes several proteins, including LAP2/thymopoietin-α, a nonmembrane protein with a predicted molecular mass of approximately 75 kDa. Therefore, we used anti-FLAG antibodies to detect MAN1 with a FLAG epitope tag in transfected cells. COS-7 cells were transfected with a plasmid encoding full-length MAN1 with a FLAG epitope tag at its amino terminus under the control of an SV40 early promoter. Virtually all of the full-length FLAG-tagged MAN1 was detected in the pellet fractions after extraction of cells with 8 M urea or 0.1 M NaOH (Fig. 6). LBR, another integral protein of the inner nuclear membrane, demonstrated the same fractionation pattern. Urea and NaOH extracts contained a polypeptide with a lower molecular mass than MAN1 that reacted with anti-FLAG antibodies, which was likely to be a nonmembrane proteolytic fragment generated during the extraction procedures. Anti-FLAG antibodies did not recognize any polypeptides in nontransfected cells (data not shown). These results confirm the prediction based on sequence analysis that MAN1 is an integral membrane protein.

Expression of Human MAN1 Shows That It Is Localized to the Nuclear Envelope—We demonstrated that FLAG-tagged MAN1 expressed from its cDNA in transfected cells was targeted to the nuclear envelope, probably the inner nuclear membrane. Detection of the expressed protein with anti-FLAG antibodies in cells examined by confocal immunofluorescence microscopy showed that it was localized exclusively to the nuclear envelope in most cells (Fig. 7). In a few cells in which

2 M. Paulin-Levasseur, unpublished observation.
MAN1 appeared to be expressed at very high levels, weak labeling of the endoplasmic reticulum was observed. A mostly exclusive nuclear envelope localization in transfected cells, along with "back-up" into the endoplasmic reticulum at high expression levels, is characteristic for integral proteins of the inner nuclear membrane (8, 9, 11). In occasional overexpressing cells, intranuclear fluorescence was observed (data not shown), which may have resulted from invaginations of the inner nuclear membrane or a nonmembrane degradation fragment of MAN1 that was transported to the nucleus.

**MAN1 Shares the LEM Module with LAP, Emerin, and Proteins of C. elegans**—We used HCA and other protein sequence analysis programs to analyze the sequence of MAN1 and compare it with those of other proteins in the publicly accessible data bases. This analysis revealed that MAN1 shares a highly conserved domain with inner nuclear membrane proteins LAP2 and emerin as well as two proteins of C. elegans—F42H11.2, W01G7.5, the LEM modules (Fig. 8B). Both MAN1 and W01G7.5 contain structurally similar globular domains of approximately 85 amino acids near their carboxyl termini. We did not identify any other proteins that contained a similar sequence.

Schematic diagrams of the LEM-containing proteins are shown in Fig. 5C. MAN1 and W01G7.5 contain two hydrophobic stretches that are putative transmembrane segments, and LAP2-β and emerin contain only one (Fig. 5C, teal). In emerin, MAN1, and W01G7.5, the LEM modules (Fig. 5C, dark blue) are at the amino termini. In LAP2-β, the LEM module starts at amino acid 108; however, a LEM-like module predicted by HCA to be similar to the LEM domain (see below) is also located at the amino terminus (Fig. 5C, light blue). C. elegans protein MAN1, and W01G7.5, the LEM modules (Fig. 8B). MAN1 and W01G7.5 contain similar sequence similarity to MAN1 (Fig. 8B). Both MAN1 and W01G7.5 contain structurally similar globular domains of approximately 85 amino acids near their carboxyl termini. We did not identify any other proteins that contained a similar sequence. Schematic diagrams of the LEM-containing proteins are shown in Fig. 5C. MAN1 and W01G7.5 contain two hydrophobic stretches that are putative transmembrane segments, and LAP2-β and emerin contain only one (Fig. 5C, teal). In emerin, MAN1, and W01G7.5, the LEM modules (Fig. 5C, dark blue) are at the amino termini. In LAP2-β, the LEM module starts at amino acid 108; however, a LEM-like module predicted by HCA to be similar to the LEM domain (see below) is also located at the amino terminus (Fig. 5C, light blue). C. elegans protein MAN1, and W01G7.5, the LEM modules (Fig. 8B). MAN1 and W01G7.5 contain similar sequence similarity to MAN1 (Fig. 8B). Both MAN1 and W01G7.5 contain structurally similar globular domains of approximately 85 amino acids near their carboxyl termini. We did not identify any other proteins that contained a similar sequence. Schematic diagrams of the LEM-containing proteins are shown in Fig. 5C. MAN1 and W01G7.5 contain two hydrophobic stretches that are putative transmembrane segments, and LAP2-β and emerin contain only one (Fig. 5C, teal). In emerin, MAN1, and W01G7.5, the LEM modules (Fig. 5C, dark blue) are at the amino termini. In LAP2-β, the LEM module starts at amino acid 108; however, a LEM-like module predicted by HCA to be similar to the LEM domain (see below) is also located at the amino terminus (Fig. 5C, light blue). C. elegans protein
F42H11.2 does not contain a predicted transmembrane segment and has a LEM domain starting at amino acid 425. *C. elegans* protein F42H11.2 contains two ankyrin domains (Fig. 8C, yellow) near its amino terminus. *C. elegans* W01G7.5 and MAN1 also contain the similar domains of approximately 85 amino acids near their carboxyl termini (Fig. 8C, orange). The polypeptide encoded by the mouse cDNA isolated by expression cloning with MAN antiseraum also contains a LEM domain at its amino terminus that is 100% identical to the one in human MAN1 (data not shown).

Based on HCA plots, it is possible that an additional regular secondary structure (probably also an α-helix) exists at the beginning of the LEM domain, before the first α-helix predicted by PHD (Fig. 9). Using HCA, a globular domain of similar size to the LEM domain can also be identified at the amino terminus of LAP2-β, clearly separated from the LEM module (amino acids 108–151) by a nonglobular region of approximately 60 amino acids (Fig. 9). The existence of hydrophobic cluster similarities as well as sequence identities (20.4% with the LAP2 LEM domain) strongly suggests that this amino-terminal domain could be a divergent member of the LEM family that we call a LEM-like module.

**DISCUSSION**

Only a few proteins of the inner nuclear membrane have been characterized at the molecular level. At the present time, LBR, emerin, nurim, and various isoforms of LAP1 and LAP2 are the only integral proteins of the inner nuclear membrane whose primary structures are known. We have now completed the cDNA cloning and sequencing of MAN1, which is predicted to be an integral protein of the inner nuclear membrane that contains a 470-amino acid nucleoplasmic domain, two transmembrane segments, and a nucleoplasmic carboxyl-terminal tail domain. Biochemical extraction and expression experiments confirm that MAN1 is an integral membrane protein with this subcellular localization.

Our sequence analysis of MAN1 has identified the LEM domain that is shared with two other proteins of the inner nuclear membrane, LAP2-β (and other LAP2 isoforms) and emerin. The LEM module is located in the nucleoplasmic portions of MAN1, emerin, and LAP2-β, suggesting that it may be a protein or nucleic acid interaction domain. The LEM domain does not correspond to the lamin-binding domain of LAP2-β, which has been mapped to the stretch between amino acids 298 and 373 of this protein (49). The chromosome-binding domain of LAP2-β has been mapped to the stretch between amino acids 1 and 85 (50), which also does not contain the LEM domain but does contain the LEM-like module predicted by HCA. Although it is tempting to speculate that the LEM domain functions in binding to other proteins, this hypothesis remains to be tested experimentally. Recently, Furukawa (51) has shown that LAP2-β binds to the rat orthologue of human BAF (barrier to autointegration factor). The minimal domain of LAP2-β necessary for binding to BAF in the yeast two-hybrid assay is amino acids 68–195; deletion of amino acids 67–137 abolishes the interaction (51). This suggests that the LEM domain of LAP2-β may be at least partly involved in its binding to BAF.

A noticeable feature of the LEM module is the absence of cysteine residues, which usually help small polypeptide domains (<50 amino acids) fold into stable structures through covalent disulfide bridges. Although the MAN1 amino-terminal domain is synthesized in the cytoplasm, where disulfide...
bridges generally do not form, their absence from the LEM module is nonetheless interesting from a structural point of view. Ten of the 12 LEM family proteins, emerin, is mutated in or absent from patients with X-linked Emery-Dreifuss muscular dystrophy (55). It is intriguing to speculate that mutations in genes encoding other LEM family proteins such as MAN1, whose gene we have assigned to chromosome 12q14, may be responsible for similar types of muscular dystrophy. Along these lines, autosomal dominant scapulopelvic muscular dystrophy has been mapped to chromosome 12 (56).

Based on their overall similarity, C. elegans protein W01G7.5 is probably the orthologue of human MAN1. Both have amino-terminal LEM domains followed by primarily nonglobular stretches and then primarily globular regions as predicted by HCA. These stretches, however, cannot be aligned because of their different lengths. MAN1 and W01G7.5 also contain two hydroporphic stretches separated by primarily globular stretches. Following their second hydroporphic segments, both MAN1 and W01G7.5 have similar domains of approximately 85 amino acids near their carboxyl termini. The identification of an orthologue in C. elegans, an organism readily amenable to genetic manipulation, should be helpful in deciphering the physiological function of MAN1.

Acknowledgments—We thank Jean-Claude Courvalin for initial recognition of the LEM-like domain, Jean-Paul Mornon for invaluable guidance of the LEM-like domain, Birgit Terjung and Racine M. Inns for assistance with confocal microscopy.