An Integral Membrane Protein of the Pore Membrane Domain of the Nuclear Envelope Contains a Nucleoporin-like Region

Einar Hallberg, Richard W. Wozniak, and Günter Blobel

Laboratory of Cell Biology, Howard Hughes Medical Institute, The Rockefeller University, New York 10021

Abstract. We have identified an integral membrane protein of 145 kD (estimated by SDS-PAGE) of rat liver nuclear envelopes that binds to WGA. We obtained peptide sequence from purified p145 and cloned and sequenced several cDNA clones and one genomic clone. The relative molecular mass of p145 calculated from its complete, cDNA deduced primary structure is 120.7 kD. Antibodies raised against a synthetic peptide represented in p145 reacted monospecifically with p145. In indirect immunofluorescence these antibodies gave punctate staining of the nuclear envelope. Immunogold EM showed specific decoration of the nuclear pores. Thus p145 is an integral membrane protein located specifically in the "pore membrane" domain of the nuclear envelope. To indicate this specific location, and based on its calculated relative molecular mass, the protein is termed POM 121 (pore membrane protein of 121 kD). The 1,199-residue-long primary structure shows a hydrophobic region (residues 29–72) that is likely to form one (or two adjacent) transmembrane segment(s). The bulk of the protein (residues 73–1199) is predicted to be exposed not on the cisternal side but on the pore side of the pore membrane. It contains 36 consensus sites for various kinases. However, its most striking feature is a repetitive pentapeptide motif XFXFG that has also been shown to occur in several nucleoporins. This nucleoporin-like domain of POM 121 is proposed to function in anchoring components of the nuclear pore complex to the pore membrane.

THE nuclear envelope (NE)[1] consists of three morphologically and biochemically distinct domains. The outer nuclear membrane with its attached ribosomes is continuous with the RER. One of the principal functions of this membrane system is to serve as the port of entry for all proteins, soluble and membrane integrated, destined for the membranes and compartments of the exocytotic and endocytotic pathway. The inner nuclear membrane is attached to the nuclear lamina and/or chromatin components and has been proposed to serve in the three-dimensional organization of chromatin (1). At numerous circumscribed points, the outer and inner nuclear membranes are connected with each other forming circular nuclear pores of ~100 nm diameter. These connecting bits of membrane appear to be biochemically and functionally distinct from both the outer and inner nuclear membrane and therefore can be regarded as a distinct third domain of the nuclear envelope referred to as the "pore membrane." The large nuclear pore complexes (NPCs) (estimated mass of 1.25×10^8 Daltons) (27) occupy the nuclear pores. There are as many pore membrane domains in a single nuclear envelope as there are NPCs.

As the pore membrane domain connects the outer and inner nuclear membrane it must allow lateral diffusion of integral membrane proteins from their site of integration in the outer membrane/RER to reach their final destination in the inner membrane (30). At the same time the pore membrane is expected to contain resident integral membrane proteins that perform the specific functions of this membrane domain. One of these functions is to anchor the NPC. Another one is likely to effect circumscribed fusion of the outer and inner nuclear membrane to generate new pores. Although it is not known what happens to the pore membrane during mitotic disassembly of the nuclear envelope, circumscribed fission and fusion events may occur if the pore membrane were to be disassembled as distinct vesicles, separated from the other two NE domains.

So far it has not been possible to isolate the pore membrane domain as a separate entity. It is therefore not known how many distinct resident integral membrane proteins it contains. Only one of these proteins, gp210, has so far been identified and molecularly characterized (11, 12, 43). This protein contains a single transmembrane segment. Most of its mass is exposed on the cisternal side of the pore membrane. Only 58 amino acid residues, including its COOH terminus, are exposed on the pore side of the pore membrane. The function of gp210 is unknown. Its small COOH-terminal domain would be topologically poised to interact with the NPC.

E. Hallberg's present address is Department of Biochemistry, Arrhenius-laboratory, Stockholm University, S-106 91, Stockholm, Sweden.

1. *Abbreviations used in this paper*: BRL, Buffalo rat liver; GlcNAc, N-acetylglucosamine; NE, nuclear envelope; NPC, nuclear pore complex; nup, nucleoporin; POM, pore membrane protein; TEA, triethanolamine.

In this paper we report the identification and characterization of another resident integral membrane protein of the pore membrane. This protein binds WGA. From its mobility in SDS-PAGE, this protein was estimated to be 145 kD and is therefore referred to as p145. Through cDNA cloning and sequencing we were able to deduce its entire primary structure. Its calculated relative molecular mass is 120.7 kD. Using monospecific antibodies against a synthetic peptide representing a portion of p145 we were able to localize this protein to the nuclear pores. Because of its calculated relative molecular mass and its localization in the pore membrane we suggest the alternative term POM 121 (pore membrane protein of 121 kD). POM 121 is likely to contain one or two transmembrane segments. In contrast to gp210 most of its mass (1,127 out of 1,199 residues) is most likely exposed on the pore side of the pore membrane. This domain contains XFXFG repeats that were also found in several of the nucleoporins (a collective term for all NPC proteins) suggesting that POM 121 may function in anchoring these proteins to the pore membrane.

Materials and Methods

Preparation of Rat Liver Nuclear Envelopes and Microsomes

Rat liver nuclei were isolated from 150–200 g Sprague Dawley rats after 24 h of starvation as described by Blobel and Potter (2). All solutions were buffered with 20 mM triethanolamine (TEA)-HCl, pH 7.5, and contained 0.5 mM PMSF and 1 mM DTT. After homogenization and before ultra centrifugation DTT was added to the homogenate (final concentration, 5 mM).

NEs were isolated as described by Dwyer and Blobel (6) with the following modifications. All solutions contained 0.1 mM PMSF and 1 mM DTT. The DNase concentration was increased to 2 μg/ml, and 250 ng/ml RNase was present at both nuclease digestion steps.

For the isolation of microsomes a rat liver homogenate was centrifuged for 10 min at 800 *g*. The resulting postnuclear supernatant was then centrifuged for 15 min at 12,000 *g* and the resulting postmitochondrial supernatant was centrifuged at 105,000 *g* for 60 min yielding a pellet of "microsomes."

Isolation of p145

In a first step, NEs were extracted by urea as follows: 2,500 U of NE (1 U of NE is the amount derived from 1 A_{260} U of isolated rat liver nuclei ($\sim 3 \times 10^6$) and is equivalent to ~10 μg of protein) were suspended in 30 ml of 7 M urea in 20 mM TEA, pH 7.5, 0.1 mM $MgCl_2$ and 10% sucrose and incubated for 10 min at room temperature. The suspension was then underlaid with 5 ml of 15% sucrose in the extraction solution and centrifuged for 1 h at 100,000 *g* to yield a supernatant and a pellet fraction, representing urea extracted NEs.

In a second step, the urea extracted NEs (15,000 U) were solubilized in 4 ml of 4% SDS, 15 mM Tris-Cl, pH 7.5, 0.15 M NaCl, 25 mM DTT, and 0.5 mM PMSF by heating at 55°C for 60 min with occasional sonication in a batch sonicator. The solubilized material was diluted to 80 ml with the appropriate reagents to give a final concentration of 0.2% SDS, 1% Triton X-100, 15 mM Tris-Cl, pH 7.5. 0.15 M NaCl, 1.25 mM DTT, and 0.1 mM PMSF. This material was incubated with 2 ml of wheat germ lectin sepharose 6 MB (Pharmacia, Uppsala, Sweden) for 4 h at 4°C and the slurry was then transferred to a column. The column was washed with 25 column volumes of wash buffer (0.2% SDS, 1% Triton X-100, 15 mM Tris-Cl, pH 7.5, 0.15 M NaCl, 1.25 mM DTT, and 0.1 mM PMSF). The wash and flow through fractions were combined and stored for analysis. The WGA column was eluted with 5 ml of 0.5 M *N*-acetylglucosamine (GlcNAc) in wash buffer and thereafter with 5 ml of 1% SDS in H_2O. The 0.5 M GlcNAc eluate from the WGA–sepharose column was concentrated by precipitation with 10% (wt/vol) TCA followed by two washes with 90% ethanol. The precipitate was solubilized in 100 μl of sample buffer (475 mM Tris-Cl, pH 8.8, 4.0% SDS, 0.1 M DTT, and 15% glycerol) and the material was subjected to SDS-PAGE using an 8% acrylamide gel. After electrophoretic separation, the proteins were transferred to nitrocellulose (40) and detected with Ponceau red. A strip of the nitrocellulose filter containing the 145-kD protein was cut out and digested with endoproteinase Lys C (Sigma Immunochemicals, St. Louis, MO) as described (8). The resulting peptide fragments were separated by reversed phase HPLC. Selected peptides were subjected to automated Edman degradation (see Fig. 5).

Isolation of mRNA, RNA Blot Analysis and Synthesis of cDNA

Total cell RNA was isolated from rat hepatoma NlSl cells (American Type Culture Collection, Rockville, MD) grown in suspension culture in DME supplemented with 10% FCS, 0.1 mM MEM-nonessential amino acids and 5 mM L-glutamine (GIBCO-BRL, Gaithersburg, MD). RNA was prepared from 2 g of pelleted cells (~800 × 10^6 cells) harvested in mid log phase by the procedure of Chirgwin and coworkers (3) with modifications (9). Poly A^+ containing RNA was isolated by oligo-dT cellulose (Boehringer Mannheim Biochemicals, Indianapolis, IN) chromatography (28).

The poly A^+ RNA was electrophoresed in a denaturing agarose gel (20), transferred to nitrocellulose (38) and probed with a random primer labeled probe corresponding to nucleotides 723–2,186 (see Fig. 5 and below).

cDNA was synthesized from 5 μg poly A^+ selected RNA using random hexamer primers and Moloney Murine Leukemia Virus reverse transcriptase (GIBCO BRL) (28).

Isolation of cDNA and Genomic Clones

On the basis of the amino acid sequence of one proteolytic fragment of p145 (see Fig. 5), partially degenerate sense (amino acids 489–494) and antisense (amino acids 506–511) primers were synthesized. SalI and XbaI restriction sites plus two extra nucleotides were inserted at the 5′-ends of the sense and antisense oligonucleotides, respectively. The oligonucleotides were purified on 15% acrylamide gels and used as primers for PCR. 4 ng of cDNA and 4 μg each of the sense and antisense primers were used in a 100 μl PCR reaction containing 10 mM Tris-Cl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, 0.2 mM each of dATP, dCTP, dGTP, and dTTP, and 2.5 U of Taq polymerase (Perkin-Elmer Corp., Norwalk, CT). Reaction conditions were 25 consecutive cycles of denaturation (95°C for 1.5 min), annealing (42°C for 2.5 min) and polymerization (72°C for 3 min). One major amplification product of 84 bp was formed. The amplification product was purified on a 4% low melting agarose gel (NuSieve GTG: FMC Bio-Products, Rockland, ME) and subcloned into pBluescript SK II (Stratagene, La Jolla, CA) and sequenced. The nucleotide sequence contained an open reading frame encoding an amino acid sequence matching the data derived from the proteolytic fragment of p145.

On the basis of the DNA sequence of this amplification product, a 34-mer antisense oligonucleotide was synthesized. The oligonucleotide was labeled at the 5′-end using ^{32}P-γ-ATP (DuPont Co., Boston, MA) and T4 polynucleotide kinase (New England Biolabs Inc., Beverly, MA) and used to screen a λZAP cDNA library derived from NlSl cell mRNA (43). Six positive plaques were isolated and their inserts (<1.5 kb) sequenced.

One of the clones isolated from the λZAP library, corresponding to nucleotides 723–2186 (see Fig. 5), was random primer labeled with ^{32}P-α-dCTP (7). This probe was used for screening of an unamplified λgtl0 cDNA library derived from Buffalo rat liver cell mRNA using an oligo dT-primer (35). 1,000,000 pfus were screened. Eight unique cDNA clones in the size range 4.8–5.5 kb were isolated, subcloned into pBluescript SK II and sequenced.

A λ DASH II rat genomic library (Stratagene) was screened using a random primer labeled (7) 719-bp restriction fragment (nucleotides −5 to 714) from the 5′ end of the λphage clone cl1. Three clones were isolated and restriction fragments hybridizing to the probe were identified by DNA blotting (32) and subcloned into pBluescript SK II. The genomic DNA fragment of the clone g301 contained sequence starting upstream of the 5′ end of the cDNA clones and ending in the Not I site at nucleotide 161 (see Fig. 5).

Hybridization Conditions

Replica lifts from the plated libraries and treatment of the nitrocellulose filters were carried out according to standard methods (28). Nitrocellulose filters, as well as the Southern and Northern blots, were prehybridized at 65°C for 4 h in hybridization solution (50% formamide, 5× Denhardt's solution, 0.2% SDS, 50 mM sodium phosphate, pH 7.7, 900 mM NaCl, 5 mM EDTA and 0.1 mg/ml denatured herring sperm DNA). Hybridization with random prime labeled probes was performed at 42°C for 24 h using ~1 ×

10^6 dpm/ml of hybridization solution. Filters were washed once at room temperature and then five times at 42°C in a solution of 0.2 × SSC and 0.1% SDS. Hybridization with the end-labeled oligonucleotide probe was performed at 37°C for 24 h using ~1 × 10^6 dpm/ml of hybridization solution containing 35% formamide. Filters were washed five times at 42°C in a solution of 2 × SSC and 0.1% SDS and positive plaques detected by autoradiography.

DNA Sequencing, DNA, and Protein Sequence Analysis

Sequencing of double stranded DNA was performed according to the dideoxy method (29) using 17-mer oligonucleotide primers (34, 37). Hydropathy analysis (17, 19), secondary structure predictions (10) were performed on a Macintosh personal computer using software obtained from DNASTAR, Inc. (Madison, WI). DNA and protein homology searches were performed in GenBank By FASTA (25).

Production of Anti-Peptide Antibody

A 13-mer synthetic peptide (residues 485–496, see Fig. 5) was synthesized with an additional cysteine at the NH_2-terminus. The peptide was coupled to keyhole limpet hemocyanin (Calbiochem Corp., LaJolla, CA) using m-maleimidobenzoyl-*N*-hydroxysuccinimide ester (Pierce, Rockford, IL) and injected into rabbits as described (15). The antiserum was affinity purified on a column consisting of the synthetic peptide (2 mg) coupled to sulfolink (1 ml) according to the manufacturers manual (Pierce). Antiserum was diluted fourfold in PBS and cycled though the column overnight at 4°C. The column was then washed with 50 ml of PBS and eluted with 0.1 M glycine-HCL, pH 2.8. Fractions containing antibody were pooled and the pH adjusted to 7.4.

Western Blot Analysis

Proteins were separated on 8% SDS-PAGE and then electrotransferred to nitrocellulose (40). For probing with WGA the nitrocellulose sheets were blocked in TBS-T (20 mM Tris-Cl, pH 7.5, 150 mM NaCl, 0.1% Tween 20) supplemented with 0.1% gelatin and then incubated with biotinylated WGA (Vector Laboratories, Inc., Burlingame, CA) at a 1:500-fold dilution in TBS-T for 3 h at room temperature. After washing four times in TBS-T the sheets were incubated with streptavidine coupled to HRP (Vector Laboratories) for 30 min at room temperature. After three washes in TBS-T and two washes in TBS the filters were developed for 5–20 min at room temperature in a mixture consisting of 40 ml TBS, 8 ml methanol containing 3 mg/ml 4-chloro-1-napthol and 20 μl H_2O_2.

Blots to be probed with antibodies against a synthetic peptide of p145 (see above) or SSRα (21) were blocked in TBS-T supplemented with 2% dry milk. The sheets were washed and incubated with affinity purified anti-peptide antibodies in blocking buffer for 1 h at room temperature. The filters were washed four times and incubated with HRP coupled to donkey anti-rabbit IgG (Amersham, UK) at a 1:5,000-fold dilution in blocking buffer for 30 min. Detection of immunoreactivity was performed as described in "ECL" detection system manual (Amersham, UK).

Preparation of BRL Cell Lysates

A 100 mm tissue culture dish containing a confluent monolayer of Buffalo rat liver (BRL) cells was washed three times with PBS. 1 ml of SDS-PAGE sample buffer was added and the plates were scraped to recover a whole cell lysate fraction. This material was then sonicated and heated at 95°C for 10 min in preparation for electrophoresis.

Immunofluorescence

Immunofluorescence was carried out on a subconfluent monolayer of BRL cells grown on coverslips. The cells were washed twice in PBS (20 mM sodium phosphate, 0.9% sodium chloride, pH 7.5) at room temperature, fixed in 3.7% formaldehyde in PBS for 20 min on ice and permeabilized with 0.5% Triton X-100 in PBS for 2 min on ice. The fixed and permeabilized cells were blocked in wash buffer (PBS, 0.1% Tween 20, 2% dry milk) for 20 min at room temperature. Cells were probed with affinity purified anti-peptide (p145) antibodies for 1 h at room temperature followed by 4 × 2-min washes in wash buffer. After a 40 min incubation with FITC-labeled donkey anti-rabbit IgG (1:100 dilution in wash buffer) the coverslips were washed 4 × 2 min in PBS and then mounted in a solution of 1 mg/ml *p*-phenylene diamine in 90% glycerol, pH 8.0. The samples were examined by a Zeiss Axiophot microscope (Carl Zeiss, Inc., Thornwood, NY) and the images were recorded on Kodak T-MAX 400 ASA film (Eastman Kodak Co., Rochester, NY).

Immunoelectron Microscopy

BRL cells were pelleted and fixed in PBS containing 2% paraformaldehyde and 0.05% glutaraldehyde. The material was then infused with 2.3 M sucrose in PBS for 30 min at room temperature and then frozen in liquid nitrogen. Ultrathin frozen sections were prepared as described (39). The sections were incubated with affinity purified rabbit anti-peptide (p145) antibodies for 2 h at room temperature, followed by goat anti-rabbit IgG bound to 10-nm gold particles. The grids were washed and stained as described (13).

Results

The Nuclear Envelope Contains an Integral Membrane Protein (p145) That Reacts with WGA

The proteins of isolated NEs and microsomes (ER) from rat liver were separated by SDS-PAGE, transferred to nitrocellulose and probed with WGA (Fig. 1 *A*) or antibodies against SSRα, a marker for an integral membrane protein of the ER (41) (Fig. 1 *B*). As expected, SSRα was found both in the ER and NE fractions (Fig. 1 *B*) and, as it is an integral membrane protein, it was not extracted from NEs by 7.0 M urea (Fig. 1 *B*, compare lanes *sup* and *pellet*). Probing with WGA yielded no reactive peptides in the ER fraction (Fig. 1 *A*) and showed several WGA-reactive polypeptides in the NE fraction (Fig. 1 *A*). Only one of these polypeptides (indicated by an arrow in Fig. 1 *A*) was not extracted by 7.0 M urea (compare lanes *sup* and *pellet*), suggesting that it is an integral membrane protein. Because of its 145,000 M_r, estimated from its mobility on SDS-PAGE, this integral membrane protein is referred to as p145. As p145 can be enzymatically labeled with UDP-galactose (data not shown) the most likely cause for its strong interaction with WGA is that it is

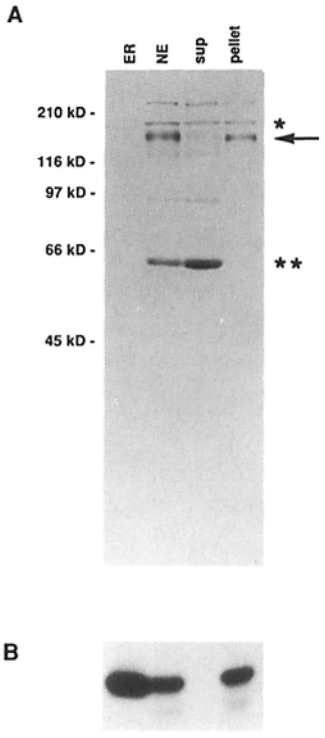

Figure 1. p145 is an integral membrane protein that binds to WGA and that is unique to the NE. A microsomal fraction (*ER*) (20 μg), a purified nuclear envelope fraction (*NE*) (20 μg), and equivalent amounts of a 7.0 M urea extract of the NE (*sup*) and extracted NE (*pellet*) were subjected to SDS-PAGE. The separated proteins were transferred to nitrocellulose and either probed with WGA (*A*) or antibodies to the α subunit of the signal sequence receptor (*B*). Note that p145 is absent from the ER. Relative molecular mass standards are indicated on the left; arrow points to p145; single star indicates p180 (35) and double star indicates p62 (33); both are nucleoporins.

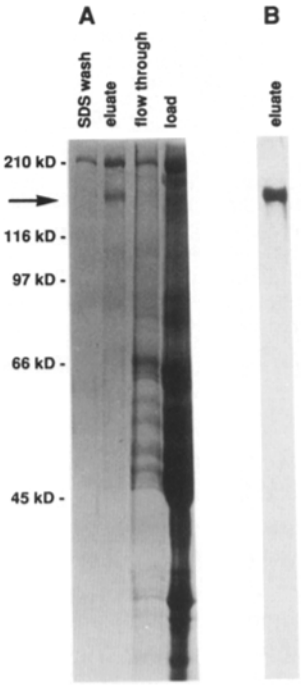

Figure 2. Purification of p145. NEs were extracted with 7.0 M urea. The extracted envelopes were solubilized with SDS–Triton X-100 and the proteins separated by WGA sepharose affinity chromatography. Proteins of equivalent aliquots (40 U) of the load, of the flow through fraction, of the 0.5 M GlcNAc eluate and of the SDS-wash were analyzed by SDS-PAGE and stained with Coomassie blue (*A*), or, in the case of the eluate fraction, were transferred to nitrocellulose and probed with WGA (*B*). Relative molecular mass standards are indicated on the left. Arrow indicates position of p145.

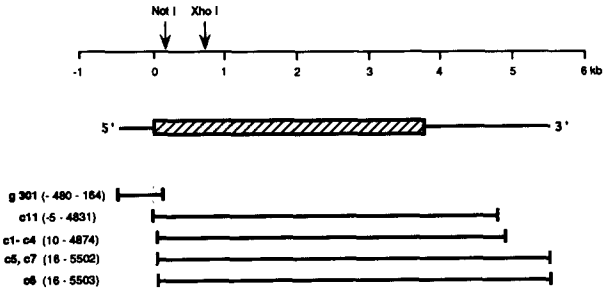

Figure 3. Schematic representation of clones of cDNA and genomic DNA. Eight individual cDNA clones (c1–c7 and c11) and one denomic clone (g301) forms a 5983 nucleotides long overlapping sequence. The 3,597-nucleotide-long open reading frame starting at position +1 is illustrated by a hatched box and the position and the length (in nucleotides) of each of the clones is indicated. The positions of the NotI (+161) and XhoI (+714) restriction sites are indicated on the scale bar.

modified by single GlcNAc residues. Since p145 is not detected in the ER fraction, it is likely that it is also absent from the outer nuclear membrane domain of the NE and is instead located either in the inner membrane or the pore membrane domain of the NE.

Purification of p145

To purify p145 for partial protein sequencing, NEs were first extracted with 7.0 M urea and then solubilized by SDS. After the addition of Triton X-100, the solubilized proteins were subjected to affinity chromatography on WGA sepharose. The proteins of the various column fractions (as well as the load) were analyzed by SDS-PAGE and stained with Coomassie blue (Fig. 2 *A*). Most of the integral membrane proteins of the NE did not bind to WGA sepharose and were found in the flow through fraction (Fig. 2 *A*). Only two proteins, one 145 kD and the other 210 kD were found in the fraction eluted with 0.5 M GlcNAc (Fig. 2 *A*). When the SDS-PAGE separated polypeptides of the eluate fraction were transferred to nitrocellulose and probed with WGA, only the 145-kD protein reacted (Fig. 2 *B*). This protein was then subjected to sequence analysis. We found that the NH_2-terminus of p145 was blocked, but we obtained sequences from several proteolytic fragments of p145 (see Fig. 5).

A 210-kD protein, which by immunoblotting was identified as gp210 (data not shown), unexpectedly coeluted with p145 (Fig. 2 *A*). The nature of this interaction is unclear since gp210 is present in the flow through fraction, the eluate, and an SDS wash of resin after elution (Fig. 2).

Molecular Cloning and Nucleotide Sequence

On the basis of peptide sequence derived from the endoproteolytic fragment of p145 (amino acids 489–511, see Fig. 5) a corresponding PCR product was synthesized and its DNA sequence was used to construct a p145 specific oligonucleotide probe. Screening of a cDNA library constructed from oligo-dT and randomly primed mRNA from the rat hepatoma cell line N1S1 (43) yielded only partial clones (<1.5 kb). We therefore screened an unamplified oligo-dT primed λgt10 cDNA library derived from Buffalo rat liver cells (35). We isolated and sequenced eight individual cDNA clones from this library in the range of 4.8–5.5 kb (Fig. 3).

The overlapping cDNA clones establish a 5,508-bp contiguous sequence (Fig. 3, and see Fig. 5). This is consistent with the size of the p145 mRNA (5.5 kb) determined by Northern blots of poly A^+ RNA from N1S1 cells (Fig. 4). The cDNA clone extending furthest in the 5′ direction, c11, contains an initiation ATG and five upstream nucleotides, CCGCG, that conform with the proposed consensus sequence for a translation initiation site (18). However, we could not be sure that this site is the actual translation initiation site. If translation were to begin at this site, the open reading frame would extend for 3,597 nucleotides and encode a protein of only 120,711 Daltons, i.e., significantly below the 145 kD estimated from its mobility on SDS–polyacrylamide gels. Our attempts to isolate other cDNA clones that would extend further 5′ of cDNA clone c11 failed. We

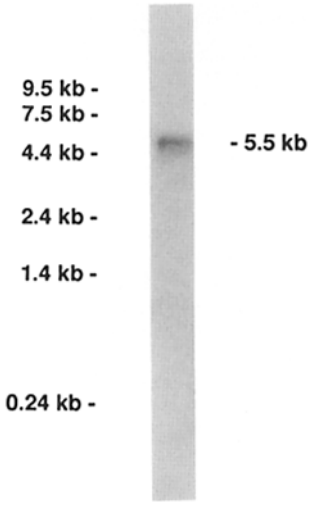

Figure 4. Northern blot analysis. 5 μg poly A+ RNA of rat hepatoma N1S1 cells was separated on an agarose gel, transferred to nitrocellulose and probed with a cDNA fragment comprising nucleotides 723–2,186 (see Materials and Methods). The size of the RNA markers are indicated on the left. The size of the mRNA hybridizing to the p145 specific cDNA probe was estimated to be 5.5 kb, as indicated.

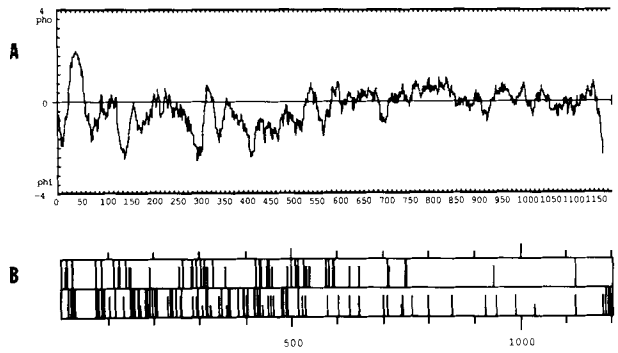

Figure 6. Distribution of hydrophobic and charged amino acids in p145. (*A*) Hydropathic values of individual amino acids residues were averaged within a 19-amino acid sliding window as described (19). Mean values were assigned to the middle amino acid residue and plotted against its position. (*B*) Positions of the acidic (*top*) and basic (*bottom*) residues are indicated, respectively, as aspartic acid (*intermediate bar*), plus glutamic acid (*full bar*) and histidine (*small bar*), plus lysine (*intermediate bar*), plus arginine (*full bar*).

Figure 5. Complete nucleotide sequence and deduced amino acid sequence of cDNA and genomic DNA clones encoding p145. Nucleotides are numbered on the right margin with the +1 coordinate assigned to the first nucleotide of the open reading frame. The deduced amino acid sequence of p145 is printed in single letter code under the first base of the nucleotide triplets. Amino acid sequences obtained by automated Edman degradation of lys C proteolytic fragments of p145 are underlined with arrows. The synthetic peptide (residues *485–496*) used for antibody production is underlined in bold. The predicted hydrophobic segment containing a potential transmembrane domain is boxed. Three polyadenylation signals (positions *4821, 4857, 5487*) in the 3′ untranslated region are underlined. An in frame stop codon (position *−66*) in the 5′ untranslated region and the termination codon of the open reading frame are underlined in bold. These sequence data are available from EMBL/GenBank/DDBJ databases under accession numbers z21513 and z21514.

therefore screened a rat genomic library and isolated a clone, g301, that contained an insert beginning 475 nucleotides upstream of the 5′ end of c11 and extending downstream to a NotI site within the cDNA clones (Fig. 3). The DNA sequence of g301 reveals an in frame TAG stop codon 66 nucleotides upstream of the first ATG of c11 strongly implicating this ATG as the initiation codon (see Fig. 5). Moreover, there were no consensus sequences for splice junction boundaries (23) between the TAG (−66) and the ATG (+1). Thus, it appears unlikely that there is an intron between these two codons that could potentially extend the open reading frame of p145 beyond the sequence contained in g301. However, there is a putative splice junction acceptor consensus sequence at position −306, suggesting that g301's nucleotide sequence 5′ to this site may be that of an intron.

cDNA Deduced Primary Structure

The 3,597-nucleotide-long open reading frame of the cloned cDNA encodes for a protein of 1,199 residues with a calculated relative molecular mass of 120,711 Daltons. The amino acid sequence of the five proteolytic fragments of p145 that were determined by Edman degradation matched the cDNA deduced amino acid sequence (Fig. 5).

The deduced primary structure of p145 is unusually rich in serines (16%), threonines (11%), and prolines (13%), which together make up 40% of the total amino acids. Serines and threonines located less than three residues from a proline have been proposed to be potential sites for GlcNAc addition (14). Most of the serines and threonines, including serine 653, are located within such a consensus motif. Peptide sequencing of one of the proteolytic fragments revealed a dehydroserine at position 653 (Fig. 5), suggesting that this amino acid was modified in the purified protein.

```
803  VFGFG    VTTAASTASTIASTSQ
824  SILFG    GAPPVTASSSAPALAS
845  IFQFG    KPLAPAASVAGTSFSQSLASSAQTAASNSSGG
882  FSGFG    GTLTTSTSAPATTSQPTLTFSNTVTPTFNIPFSASAKPALPTYPGANS
935  QPTFG    ATDGATKPAL
950  APSFG    S
956  SFTFG    NSVASAPSAAPA
973  PAAFG    GAA
980  QPAFG    GLKAS
991  ASTFG    TPAST
1001 QPAFG    STTS
1010 VFSFG    SAT
1018 TSGFG    AAAATTQTTHSGSS
1037 SSLFG    SSTPS
1047 PFTFG    GSAAPAG
1059 GGGFG    LSATPGTGSTSG
1076 TFSFG    SGQSGTTGT
1090 TTSFG    GSLSQNTLGAPSQSS
1110 PFAFS    VGSTPES
1122 KPVFG    GTS
1130 TPTFG    QSAPAPGVGTTGS
1148 SLSFG    APSTPAQGFVG
1164 VGPFG    SGAPSFSIGAGSKTPGARQRLQARRQHTRKK.
```

Figure 7. The COOH-terminal third of p145 contains 23 repeated pentapeptide motifs. The position of the first amino acid of the XFXFG pentapeptide repeats are indicated on the left. The serines and threonines in the intervening sequences are marked in bold.

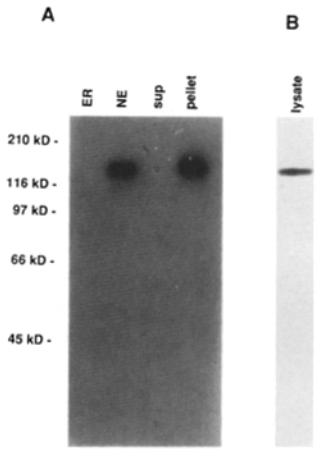

Figure 8. Characterization of antibodies against a synthetic peptide of p145. (*A*) Proteins of microsomes (*ER*), nuclear envelopes (*NE*), a urea extract of nuclear envelopes (*sup*) and the urea extracted nuclear envelopes (*pellet*) were subjected to SDS-PAGE, transferred to nitrocellulose (see also Fig. 1) and probed with affinity purified antibodies against a synthetic peptide (residues 485–496) of p145. (*B*) The total proteins of a Buffalo rat liver cell lysate (200 μg of protein) were analyzed as in *A*.

The deduced amino acid sequence contains 36 consensus sequences for various protein kinases (19 for protein kinase C, 12 for casein kinase II, two for cAMP dependent protein kinase and three for the cell cycle dependent cdc2 kinase) (22).

A hydropathy plot of the deduced amino acid sequence reveals one single region (amino acids 29–72), which is hydrophobic enough to contain a transmembrane segment (Fig. 6 *A*). The hydrophobic segment, which is flanked by charged residues, is 44 amino acids long. Part of this sequence could form a single transmembrane α-helix. Alternatively, it may form two transmembrane α-helices adjacent to each other (see below, Fig. 11). Except for the hydrophobic segment, the NH_2-terminal half of the deduced amino acid sequence is much richer in charged residues as compared to the COOH-terminal half (Fig. 6 *B*).

The COOH-terminal third of the deduced amino acid sequence of p145 contains a repetitive pentapeptide motif of XFXFG separated by sequences rich in serines and threonines (Fig. 7). Similar repeats and serine and threonine-rich spacers occur in mammalian nucleoporins p62 (33) and NUP153 (35). XFXFG is also part of a repeated nonapeptide in the middle domains of the yeast nucleoporins NSP1 (16) and NUP1 (5). Because of this repeat the COOH-terminal third of p145 shows similarity (20% to NUP153 in an 508 residue long overlap; 31% to p62 in an 140 residue long overlap; 20% to NSP1 in an 230 residue long overlap; 18% to NUP1 in an 397 residue long overlap) to the corresponding regions of these proteins (see Fig. 5 in ref. 35). These data suggest that the COOH-terminal third of p145 has characteristics in common with nucleoporins.

Subcellular Distribution and Immunolocalization

Antibodies were raised against a synthetic peptide within a region of the predicted amino acid sequence of p145 that did not show homology to nucleoporins. After affinity purification these antibodies recognized a single protein of 145 kD on a nitrocellulose blot of SDS-PAGE separated proteins of a buffalo rat liver cell lysate (Fig. 8 *B*). When rat liver subcellular fractions were analyzed the antigen was exclusively found in the nuclear envelope fraction, and not in the ER fraction, and remained in the membrane pellet after extraction of rat liver nuclear envelopes with 7 M urea (Fig. 8 *A*). These data are in good agreement with those of Fig. 1.

The affinity purified anti-peptide antibodies were used in indirect immunofluorescence microscopy of fixed and permeabilized Buffalo rat liver cells. As shown in Fig. 9 *A*, the antibodies gave rise to a punctate staining in a focal plane tangential to the upper surface of the cell nucleus. When a focal plane through the equator of the nucleus was chosen a punctate nuclear rim staining pattern was seen (Fig. 9 *B*). Double immunofluorescence with mAb 414, which recognizes proteins of the NPC (4), yielded exactly superimposable staining (not shown) suggesting that p145 is located at or close to the NPC.

The affinity purified anti-peptide antibodies were also used to localize p145 in cryosections of BRL cells. Fig. 10 *A* shows immunogold labeling of the nuclear pores. At a slightly higher magnification immunogold decorates a single nuclear pore (Fig. 10 *B*). The immunolocalization data demonstrate that p145 is located at or close to the nuclear pore. As p145 is an integral membrane protein of the pore membrane domain of the NE and has a calculated relative molecular mass of 120.7 kD, it is termed POM 121.

The precise topology of POM 121 in the pore membrane

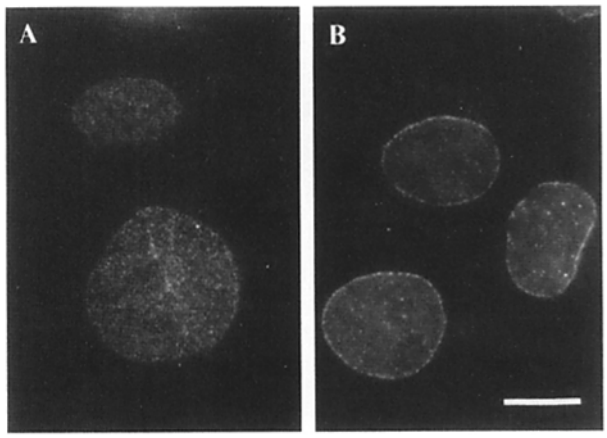

Figure 9. Localization of p145 by indirect immunofluorescence. Coverslips containing fixed and permeabilized Buffalo rat liver tissue culture cells were probed with affinity purified anti-peptide antibodies (see Fig. 8). FITC-labeled donkey anti-rabbit IgG was used as secondary antibody. A tangential view shows punctate staining of the nuclear surface (*A*). An equatorial view shows punctate rim staining (*B*). Bar, 10 μm.

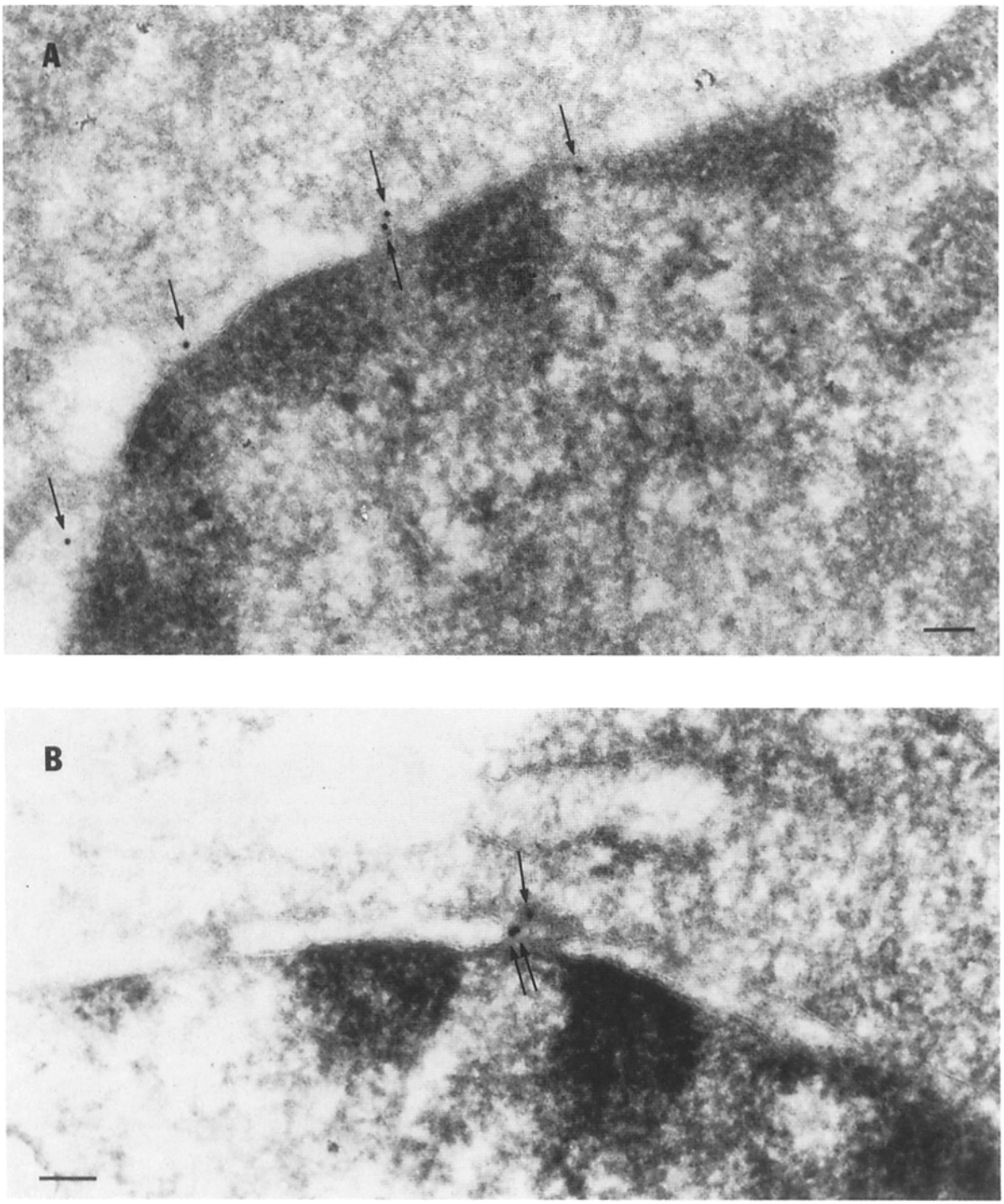

Figure 10. Localization of p145 by immunoelectron microscopy. Buffalo rat liver cells were cryosectioned as described in Materials and Methods. The ultra-thin sections were first incubated with affinity purified anti-peptide antibodies (see Fig. 8) and then with anti-rabbit IgG conjugated to 10-nm gold particles. In *A* four nuclear pores are labeled with one or two gold particles each (*arrows*). In *B* the outer and inner membrane and the connecting pore membrane are well preserved and three gold particles decorate a single nuclear pore (*arrows*). Bars, 100 nm.

domain remains to be determined. Fig. 11 shows two models, one with a single transmembrane segment and the NH_2 terminus exposed on the cisternal side, and the other one with two transmembrane segments and the NH_2 terminus exposed on the pore side of the pore membrane domain. In both models the bulk of POM 121 (1127 of 1199 amino acids) including the COOH terminus, the putative consensus sites for various kinases (see above) and GlcNAc addition, and the XFXFG repeats (see above) are exposed on the pore side of the pore membrane domain.

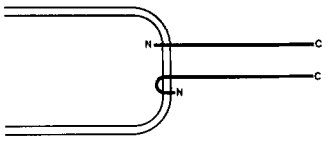

Figure 11. Proposed models for the topology of p145 in the pore membrane domain of the nuclear envelope. A short NH_2-terminal region (*N*) located on the cisternal side of the pore membrane, is followed by a single transmembrane segment and the bulk COOH-terminal regional. (*C*) exposed on the pore side of the pore membrane (*upper model*). The lower model shows two transmembrane segments with the NH_2-terminal exposed on the pore side of the pore membrane (for details see Fig. 5 and Discussion).

Discussion

Our data here shows that POM 121 is a novel integral membrane protein that resides in the pore membrane domain of the rat liver nuclear envelope. The most interesting feature of the cDNA deduced primary structure of POM 121 is the presence of a repetitive pentapeptide motif XFXFG that is also present in several mammalian (33, 35) and yeast (5, 24) nucleoporins (nups). Thus, POM 121 possesses a nup-like domain but unlike nups, it contains in addition one (or perhaps two) transmembrane segment(s). This suggests that POM 121's nup-like domain is part of the NPC and that POM 121 therefore may function as a membrane anchor for components of the NPC. Another feature that POM 121 shares with those nups that have so far been molecularly characterized is an abundance of various consensus sites for phosphorylation (26, 35). In mitosis the NPC is reversibly disassembled and detached from the pore membrane (4). It is likely that numerous phosphorylation/dephosphorylation events coordinate disassembly/reassembly and detachment/reattachment. Moreover, like several nups, POM 121 is most likely modified by single GlcNAc residues at serine/threonine and therefore strongly interacts with WGA.

We have not yet obtained data on the topology of POM 121 in the pore membrane. However the presence of a hydrophobic region between residues 29 and 72 of the 1,199-residue-long protein indicate that the membrane anchor (with one or two adjacent transmembrane segments, see Fig. 11) is at the NH_2-terminal region. The bulk of POM 121 (residues 73–1199) containing the repetitive XFXFG motifs as well as all consensus sequences for phosphorylation is likely to face the pore side rather than the cisternal side of the pore membrane. This suggests a highly asymmetric distribution of the mass of POM 121 on the two sides of the pore membrane. The other pore membrane protein that has so far been molecularly characterized, gp210, also exhibits a highly asymmetric distribution of its mass. In this case, however, most of the mass is located on the cisternal side of the pore membrane domain. Targeting studies with gp210 showed that its transmembrane segment (but not its pore side-exposed COOH-terminal tail) is the dominant topogenic signal for sorting gp210 to the pore membrane domain (42). Specific interactions with the transmembrane segment of another pore membrane protein has been suggested (42). Perhaps POM 121's transmembrane segment interacts with that of gp210 and the dimer would then be retained in the pore membrane domain via interaction of POM 121's large nup-like domain with the NPC.

It is likely that POM 121 is identical to an integral membrane protein of 145,000 M_r (estimated from its mobility on SDS-PAGE) that was identified by Snow et al. (31). They demonstrated that on immunoblots this and other proteins are recognized by polyspecific monoclonal antibodies. By immunoelectron microscopy these polyspecific antibodies labeled nuclear pores. However, it remains unclear whether one, several, or all of these cross-reactive proteins contributed to this decoration of the NPCs. Therefore it was not possible to conclude that the p145 protein identified by Snow and collaborators is a pore membrane protein. Our immunolocalization data here with monospecific antibodies allowed unequivocal localization of POM121 to the pore membrane domain of the nuclear envelope. The deduced amino acid sequence shows that POM 121 shares repetitive XFXFG pentapeptide motifs with several nups (see above) and thus might be part of the epitope recognized by the polyspecific mAbs of Snow et al. (31). While this suggests that POM 121 is identical to the p145 identified by Snow et al., this remains to be experimentally established.

We thank Helen Shio and Eleana Sphicas for preparation of the immunoelectron microscopy specimens and members of the Rockefeller University/Howard Hughes Medical Institute Biopolymer facility for protein sequencing. The λgt10 cDNA was a kind gift of Jun Sukegawa. The anti-SSRα was kindly provided by Giovanni Migliaccio and Chrisopher Nicchitta. We also thank Hiroshi Murakami and Jun Sukegawa for valuable advice.

Received for publication 23 February 1993 and in revised form 7 May 1993.

References

1. Blobel, G. 1985. Gene gating: a hypothesis. *Proc. Natl. Acad. Sci. USA.* 82:8527–8529.
2. Blobel, G., and V. R. Potter. 1966. Nuclei from rat liver: isolation method that combines purity with high yield. *Science (Wash. DC).* 154:1662–1665.
3. Chirgwin, J. M., A. E. Przbyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* 18:5294–5299.
4. Davis, L. I., and G. Blobel. 1986. Identification and characterization of a nuclear pore complex protein. *Cell.* 45:699–709.
5. Davis, L. I., and G. R. Fink. 1990. The NUP1 gene encodes an essential component of the yeast nuclear pore complex. *Cell.* 61:965–978.
6. Dwyer, N., and G. Blobel. 1976. A modified procedure for the isolation of a pore complex-lamina fraction from rat liver nuclei. *J. Cell Biol.* 70:581–591.
7. Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6–13.
8. Fernandez, J., M. DeMott, D. Atherton, and S. Mische. 1992. Internal protein sequence analysis: Enzymatic digestion for less than 10 micrograms of protein bound to polyvinylidine difluoride or nitro cellulose membranes. *Anal. Biochem.* 201:255–264.
9. Freeman, G. J., C. Clayberger, R. DeKruyff, D. S. Rosenblum, and H. Cantor. 1983. Sequential expression of new gene programs in inducer T-cell clones. *Proc. Natl. Acad. Sci. USA.* 80:4094–4098.
10. Garnier, J., D. J. Osguthorpe, and B. Robson. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120:97–120.
11. Gerace, L., Y. Ottaviano, and C. Kondor-Koch. 1982. Identification of a major polypeptide of the nuclear pore complex. *J. Cell Biol.* 95:826–837.
12. Greber, U. F., A. Senior, and L. Gerace. 1990. A major glycoprotein of the nuclear pore complex is a membrane-spanning polypeptide with a large lumenal domain and a small cytoplasmic tail. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:1495–1502.
13. Griffiths, G., K. Simons, G. Warren, and K. T. Tokuyasu. 1983. Immunoelectron microscopy using thin, frozen sections: application to studies of the intracellular transport of Semliki forest virus spike glycoproteins. *Methods Enzymol.* 96:466–485.
14. Haltiwanger, R. S., W. G. Kelly, E. P. Roquemore, M. A. Blomberg, L. Dennis Dong, L. Kreppel, T. Chou, and G. W. Hart. 1992. Glycosylation of nuclear and cytoplasmic proteins is ubiquitous and dynamic. *Biochem. Soc. Trans.* 20:264–269.
15. Harlow, E., and D. Lane. 1988. Antibodies: A Laboratory Manual. Cold

Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
16. Hurt, E. C. 1988. A novel nucleoskeletal-like protein located at the nuclear periphery is required for the life cycle of *Saccharomyces cervisiae*. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:4323–4334.
17. Klein, P., M. Kanehisa, and C. DeLisi. 1985. The detection and classification of membrane-spanning proteins. *Biochim. Biophys. Acta.* 815: 468–476.
18. Kozak, M. 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* 115:887–903.
19. Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105–132.
20. Lehrach, H., D. Diamond, J. M. Wozney, and H. Boedker. 1977. RNA molecular weight determination by gel electrophoresis under denaturing conditions, a critical reexamination. *Biochemistry.* 16:4743–4751.
21. Migliaccio, G., C. V. Nicchitta, and G. Blobel. 1992. The signal sequence receptor, unlike the signal recognition particle receptor, is not essential for protein translocation. *J. Cell Biol.* 117:15–25.
22. Moreno, S., and P. Nurse. 1990. Substrates for p34 cdc2 in vivo veritas? *Cell.* 61:549–551.
23. Mount, S. M. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res.* 10:459–472.
24. Nehrbass, U., H. Kern, A. Mutvei, H. Horstmann, B. Marshallsay, and E. C. Hurt. 1990. NSP1: a yeast nuclear envelope protein localized at the nuclear pores exerts its essential function by its carboxy-terminal domain. *Cell.* 61:979–989.
25. Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA.* 85:2444–2448.
26. Radu, A., G. Blobel, and R. Wozniak. 1993. NUP155 is a novel nuclear pore complex protein that contains neither repetitive sequence motifs nor reacts with wheat germ agglutinin. *J. Cell Biol.* 121:1–9.
27. Reichelt, R., A. Holzenburg, E. L. J. Buhle, M. Jarnik, A. Engel, and U. Aebi. 1990. Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. *J. Cell Biol.* 110:883–894.
28. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.
30. Smith, S., and G. Blobel. 1993. The first membrane spanning region of the lamin B receptor is sufficient for sorting to the inner nuclear membrane. *J. Cell Biol.* 120:631–637.
31. Snow, C. M., A. Senior, and L. Gerace. 1987. Monoclonal antibodies identify a group of nuclear pore complex glycoproteins. *J. Cell Biol.* 104:1143–1156.
32. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503–517.
33. Starr, C. M., M. Donofrio, M. K. Park, and J. A. Hanover. 1990. Primary sequence and heterologous expression of nuclear pore glycoprotein p62. *J. Cell Biol.* 110:1861–1871.
34. Strauss, E. C., J. A. Kobori, G. Siu, and L. E. Hood. 1986. Specific-primer-directed DNA sequencing. *Anal. Biochem.* 154:353–360.
35. Sukegawa, J., and G. Blobel. 1993. A nuclear pore complex protein that contains zinc finger motifs, binds DNA and faces the nucleoplasm. *Cell.* 72:29–38.
36. Suzuki, M. 1989. SPKK, a new nucleic acid-binding unit of protein found in histone. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:797–804.
37. Tabor, S., and C. C. Richardson. 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA.* 84:4767–4771.
38. Thomas, P. S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. USA.* 77: 5201–5205.
39. Tokuyasu, K. T. 1973. A technique for ultracryotomy of cell suspensions and tissues. *J. Cell Biol.* 57:551–565.
40. 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.* 76:4350–4354.
41. Wiedmann, M., T. V. Kurzchalia, E. Hartmann, and T. A. Rapaport. 1987. A signal sequence receptor in the endoplasmic reticulum membrane. *Nature (Lond.).* 328:830–833.
42. Wozniak, R. W., and G. Blobel. 1992. The single transmembrane segment of gp210 is sufficient for sorting to the pore membrane domain of the nuclear envelope. *J. Cell Biol.* 119:1441–1449.
43. Wozniak, R. W., E. Bartnik, and G. Blobel. 1989. Primary structure analysis of an integral membrane glycoprotein of the nuclear pore. *J. Cell Biol.* 108:2083–2092.