A New Family of Yeast Nuclear Pore Complex Proteins

Susan R. Wente, Michael P. Rout, and Günter Blobel

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

Abstract. We have identified a novel family of yeast nuclear pore complex proteins. Three individual members of this family, NUP49, NUP100, and NUP116, have been isolated and then characterized by a combination of molecular genetics and immunolocalization. Employing immunoelectron and immunofluorescence microscopy on yeast cells, we found that the binding of a polyspecific monoclonal antibody recognizing this family was predominantly at the nuclear pore complexes. Furthermore, the tagging of NUP49 with a unique epitope enabled the immunolocalization of this protein to the nuclear pore complex by both fluorescence and electron microscopy. DNA sequence analysis has shown that the amino-terminal regions of NUP49, NUP100, and NUP116 share repeated "GLFG" motifs separated from each other by glutamine, asparagine, serine and threonine rich spacers. All three proteins lack a repetitive domain found in the two precisely described yeast nuclear pore complex proteins. Only NUP49 is essential for cell viability. NUP116-deficient cells grow very slowly and are temperature sensitive, whereas the lack of NUP100 has no detectable phenotype. NUP100 and NUP116 are homologous over their entire lengths. Interestingly, NUP100 and NUP116 are both flanked by a histidine tRNA gene and a transposon element suggesting that they may have arisen by gene duplication. We propose that subfamilies of pore complex proteins can be defined by their characteristic combinations of different modular domains.

Nuclear pore complexes are large molecular assemblies of an estimated molecular mass of $1.25 \times 10^8$ D (Reichel et al., 1990) that mediate bidirectional nucleocytoplasmic transport (Feldherr et al., Dworetzky and Feldherr, 1988; reviewed in Dingwall and Laskey, 1986). High resolution electron microscopy studies of amphibian oocyte nuclear envelopes have resulted in a three-dimensional reconstruction of the pore complex to < 10 nm resolution. Each pore complex is a cylindrical structure with a superficial octagonal symmetry and arranged such that its axis is perpendicular to the plane of the nuclear envelope. It is composed of distinct substructures referred to as spokes, rings, and a "plug" or "transporter" (Unwin and Milligan, 1982; Akey, 1989, 1990). Fibris extend from both the nuclear and cytoplasmic surfaces (Richardson et al., 1988; Ris, 1990). The pore complexes of all eukaryotic organisms are believed to be architecturally and functionally similar (Maul, 1977; Silver, 1991).

A number of monoclonal antibodies generated independently by several laboratories (Davis and Blobel, 1986, 1987; Snow et al., 1987; Park et al., 1987) have identified a group of nuclear pore complex proteins, collectively referred to as nucleoporins (Davis and Blobel, 1986). Many of these monoclonal antibodies bind in a polyspecific manner to multiple nucleoporins. In addition, they often possess extensive interspecies cross-reactivity by recognizing nucleoporins of mammals, *Xenopus*, and yeast (Davis and Blobel, 1986, 1987; Snow et al., 1987; Park et al., 1987; Featherstone et al., 1988; Aris and Blobel, 1989; Davis and Fink, 1990; Carmon-Fonseca et al., 1991).

Nucleocytoplasmic transport can be inhibited by the presence of either anti-nucleoporin monoclonal antibodies (Featherstone et al., 1988; Dabauvalle et al., 1988) or the lectin wheat germ agglutinin (WGA) (Finlay et al., 1987; Yoneda et al., 1987; Dabauvalle et al., 1988a); some vertebrate nucleoporins are post-translationally modified by the addition of O-linked N-acetylglucosamine and therefore bind WGA (Holt and Hart, 1986; Davis and Blobel, 1986, 1987; Hanover et al., 1987; Holt et al., 1987). O-linked glycosylation of the yeast nucleoporins has not yet been confirmed (Davis and Fink, 1990; Carmon-Fonseca et al., 1991). Several biochemical studies have implicated the glycosylated vertebrate nucleoporins in the mediated import of proteins (Akey and Goldfarb, 1989; Finlay and Forbes, 1990; Finlay et al., 1991; Sterne-Marr et al., 1992). However, with the use of such polyspecific reagents, attributing any discrete function to an individual nucleoporin is not possible.

Three of the nucleoporins, namely p62 of vertebrates (Starr et al., 1990; Cordes et al., 1991; Carmon-Fonseca et al., 1991) and NUP1 (Davis and Fink, 1990) and NSP1 (Nehrbass et al., 1990) of the yeast *Saccharomyces cerevisiae*, have been characterized at the molecular level. Comparison of the primary structure of the two yeast nucleoporins revealed that the central regions of both contain a unique repetitive sequence of nine amino acid residues, flanked by highly charged spacers of various lengths (Davis.
and Fink, 1990; Nehrbass et al., 1990). The two yeast nucleoporins identified so far can account for only a fraction of the total mass of the nuclear pore complex. Clearly, more components need to be characterized for an understanding of the nuclear pore complex mediated functions.

In this paper we report the identification and characterization of three additional yeast nucleoporins, NUP49, NUP100, and NUP116. These proteins share unique structural features which define them as a new family of nucleoporins.

Materials and Methods

Strains and Plasmids

Tables I and II describe the yeast strains and plasmids, respectively, that are referred to in this study. The yeast strains were grown in either YEP (1% yeast extract, 2% bactopeptone) with 2% glucose or synthetic minimal media (SD: Sherman et al., 1986) supplemented with appropriate amino acids and glucose. Yeast transformations were by the lithium acetate method (Ito et al., 1983) and general genetic manipulations of yeast cells were conducted as described by Sherman et al. (1986). Bacterial cells were cultured in Luria broth and transformed, or infected with phage, by standard methods (Sambrook et al., 1989). Bacterial strains Y1090 and Y1089 (Young and Davis, 1983) were used for λ propagation; DH5α (Hanahan, 1983; Bethesda Research Laboratories, Bethesda, MD) for pBS-KS+ based plasmids (Stratagene Inc., La Jolla, CA); JM109 (Yanisch-Perron et al., 1985) for pMAL-cRI based plasmids (Maina et al., 1988); and DH5α (Hanahan, 1983; Bethesda Research Laboratories, Bethesda, MD) for pBS-KS+ based plasmids (Stratagene Inc., La Jolla, CA); JM109 (Yanisch-Perron et al., 1985) for pMAL-cRI based plasmids (Maina et al., 1988); and DH5α (Hanahan, 1983; Bethesda Research Laboratories, Bethesda, MD) for pBS-L5 (S. R. Wente and O. M. Rosen, unpublished results). Two overlapping isolates of NUP49 were obtained from screening the AGL1 library. To obtain full length clones of the NUP49 and NUP100 loci a YEp3 based yeast genomic library (Nasmyth and Tatchell, 1980) was screened with appropriate 32P-labeled fragments from the respective AGL1 clones.

The DNA sequence of the NUP49, NUP100, and NUP116 genes was determined by the dideoxychain termination method (Sanger et al., 1977) on

Table I. Yeast Strain Genotype and Construction

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>W303</td>
<td>Mata/Mata ade2-1/ade2-1 ura3-1/ura3-1 his3-11,15/trp1-1/trp1-1 leu2-3,112/leu2-3,112 can1-100/can1-100</td>
<td>Integrative transformation of W303 with EcoRI fragments of pSW55</td>
</tr>
<tr>
<td>W303a</td>
<td>Mata ade2-1 ura3-1 his3-11,15/trp1-1/leu2-3,112 can1-100</td>
<td>Integrative transformation of W303 with EcoRI/XhoI fragments of pSW56</td>
</tr>
<tr>
<td>SWY1</td>
<td>Mata ade2-1/ade2-1 ura3-1/ura3-1 his3-11,15/trp1-1/trp1-1 leu2-3,112/leu2-3,112 can1-100/can1-100 nup49-1::URA3/+</td>
<td>Segregant from tetrad of sporulated SWY2</td>
</tr>
<tr>
<td>SWY2</td>
<td>Mata ade2-1/ade2-1 ura3-1/ura3-1 his3-11,15/trp1-1/trp1-1 leu2-3,112/leu2-3,112 can1-100/can1-100 nup100-1::URA3/+</td>
<td>Transformation of SWY1 with pSW62, sporulation, resulting viable segregant</td>
</tr>
<tr>
<td>SWY3</td>
<td>Mata ade2-1 ura3-1 his3-11,15/trp1-1/leu2-3,112 can1-100 nup100-1::URA3/LEU</td>
<td>Transformation of SWY1 with pSW63, sporulation, resulting viable segregant</td>
</tr>
<tr>
<td>SWY11</td>
<td>Mata ade2-1 ura3-1 his3-11,15/trp1-1/leu2-3,112 can1-100 nup49-1::URA3/pSW62(LEU)</td>
<td>Transformation of SWY1 with pSW79</td>
</tr>
<tr>
<td>SWY12</td>
<td>Mata ade2-1 ura3-1 his3-11,15/trp1-1/leu2-3,112 can1-100 nup100-1::URA3/pSW63(LEU)</td>
<td>Transformation of SWY3 with pSW79</td>
</tr>
<tr>
<td>SWY19</td>
<td>Mata ade2-1 ura3-1 his3-11,15/trp1-1/leu2-3,112 can1-100 nup100-1::URA3/pSW97(LEU)</td>
<td>Integrative transformation of W303 with Xbal/SalI fragments of pSW99</td>
</tr>
<tr>
<td>SWY26</td>
<td>Mata ade2-1/ade2-1 ura3-1/ura3-1 his3-11,15/trp1-1/trp1-1 leu2-3,112/leu2-3,112 can1-100/can1-100 nup116-5::His3/+</td>
<td>Segregant from tetrad of SWY26</td>
</tr>
<tr>
<td>SWY29</td>
<td>Mata ade2-1 ura3-1 his3-11,15/trp1-1/leu2-3,112 can1-100 nup116-5::His3</td>
<td>Integrative transformation of W303 with EcoRI fragments of pSW54</td>
</tr>
<tr>
<td>SWY30</td>
<td>Mata ade2-1/ade2-1 ura3-1/ura3-1 his3-11,15/trp1-1/trp1-1 leu2-3,112/leu2-3,112 can1-100/can1-100 nup116-6::URA3/+</td>
<td>Segregant from tetrad of SWY30</td>
</tr>
<tr>
<td>SWY31</td>
<td>Mata ade2-1 ura3-1 his3-11,15/trp1-1/leu2-3,112 can1-100 nup116-6::URA3</td>
<td>Transformation of SWY30 with pSW75, sporulation, resulting segregant</td>
</tr>
<tr>
<td>SWY54</td>
<td>Mata ade2-1 ura3-1 his3-11,15/trp1-1/leu2-3,112 can1-100 nup116-6::URA3/pSW75(LEU)</td>
<td>Transformation of SWY30 with pSW76, sporulation, resulting segregant</td>
</tr>
<tr>
<td>SWY55</td>
<td>Mata ade2-1 ura3-1 his3-11,15/trp1-1/leu2-3,112 can1-100 nup116-6::URA3/pSW76(LEU)</td>
<td>Cross of SWY3 and SWY29</td>
</tr>
</tbody>
</table>

Isolation and Sequencing of NUP49, NUP100, and NUP116 Loci

A yeast genomic AGL1 expression library from Clontech Laboratories, Inc., was screened as described by Young and Davis (1991) with the following modifications. From Western blotting experiments, the signal to noise ratio for screening was optimized by varying the antibody titers, incubation buffers, and blocking conditions. Nitriloacetic filters from phage lifts were washed in 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.05% Tween 20 (TBST) with first 20 and then 2% (wt/vol) nonfat dry milk (Carnation Co., Los Angeles, CA) (30 min at room temperature). The blocked filters were incubated with a 1:5 dilution of anti-nucleoporin monoclonal antibody tissue culture supernatants in TBST/2% milk for 14 h at 4°C. Visualization of the MAb92 cross-reactive plaques was via a commercial alkaline phosphatase conjugated anti-mouse IgG (Promega Biotech, Madison, WI) and the manufacturer's substrate solutions. 25 plaque purified positive isolates were obtained from screening 3 x 106 recombinant phage, which after restriction analysis yielded 10 distinct clones. The 10 clones were partially characterized by their differential antibody cross-reactivity with the individual antibodies used to screen the library, with previously described monoclonal antibodies (MAb414, MAb350, and MAb306 [Davis and Blobel, 1986, 1987; Davis and Fink, 1990]), and with monospecific polyclonal antibodies against NSP1 (kindly provided by U. Nehrbass and E. Hurt (EMBL, Heidelberg, Germany)). One of the positives was a lacZ fusion with the gene encoding NSP1, a previously characterized yeast nucleoporin (Nehrbass et al., 1990), verifying that our screening process was capable of isolating clones encoding nucleoporins. λ DNA from plaque purified positive isolates was isolated by use of LambdaSorb (Promega Biotech) and the inserts were subcloned into either pBS-KS+ or pBS-L5 (S. R. Wente and O. M. Rosen, unpublished results). Two overlapping isolates of NUP49 were obtained from screening the AGL1 library. To obtain full length clones of the NUP49 and NUP100 loci a YEp3 based yeast genomic library (Nasmyth and Tatchell, 1980) was screened with appropriate 32P-labeled fragments from the respective AGL1 clones.

The DNA sequence of the NUP49, NUP100, and NUP116 genes was determined by the dideoxychain termination method (Sanger et al., 1977) on
The NcoI/SalI and PstI/HindII linkers were obtained via subcloning into pBS-L5 or pLC20H respectively. Notation "bp" = base pair, die of the open reading frame and coincident insertion of the selectable NUP49 was plasmid. Unique oligonucleotide primers were then utilized to sequence protein sequences of NUP116, NUPIOO, and NUP49 to the sequences in the two different markers, either the URA3 gene for a 2103 bp BamHI/HindIII fragment of NUP49 and the numeral following NsiI* designates the nucleotide position of this particular restriction site in the sequences of Figure 4.

deleted copy (SWY1). Transformation of the heterozygous diploid strain yielded a strain which when induced to sporulate resulted in the recovery of a viable Ura+/Leu + haploid strain (SYll) with the expected Mendelian frequency (25 %). A truncated NUPIOO gene (nup49-3::Aft), then BamHI/NcoI fragment into same of pSW65, replaced the indicated sequence in NUP49. Early log phase yeast cells were prepared for immunofluorescence experi-
ments by a modification of the filtration method of Kilmartin and Adams (1984). Mature yeast cells from auxotrophic diploids (nup49-1::URA3-) were selected. Southern analysis identified the diploid strains with the correct replacement of the individual gene by integration with the respective URA3 or HIS3-disrupted copies (SWY2, SWY30, and SWY26). A heterozygous diploid strain with both the nup100-1::URA3 and the nup16-5::HIS3 alleles (SWY56) was generated by a cross of SWY3 and SWY29. Southern analysis was also performed on the viable haploid strains (SWY3, SWY31, SWY29).

**Immunoelectron Microscopy**

Spheroplasts from early log phase yeast haploid cells (strain SWY12) were incubated in 0.6 M sorbitol for 15 min at 4°C and then fixed in suspension with 0.075% glutaraldehyde, 4% formaldehyde, 0.1 M cacodylate buffer, pH 7.2, for 1 h at 4°C. After centrifugation for 1 min at 1000 g, the pellets were dehydrated for 30 min in 50% ethanol followed by 60 min in 70% ethanol. The hydrophilic resin LR White was added and allowed to infiltrate overnight at 4°C. The resin was polymerized by baking at 55°C for 24 h. Thin sections were collected on nickel grids coated with formvar, stabilized with carbon. The sections were blocked with 1% BSA/0.2% fish gelatin in PBS, and then incubated for 2 h at room temperature with the appropriate tissue culture supernatant (MAb192 or monoclonal 12CA5 antibody [Berkeley Antibody Co., Richmond, CA]) diluted 1:2 in blocking buffer. After extensive washes with PBS, the grids were incubated for 1 h in a suspension of 10 nm colloidal gold coated with goat anti-mouse antibody (Amersham Corp., Arlington Heights, IL) diluted 1:20 in TBS. After the final washes, the grids were contrasted by staining in 2% aqueous uranyl acetate for 10 min. Samples were visualized with a JEOL 100CX electron microscope (JEOL USA, Inc., Peabody, MA) at 80 kV, and photographs were recorded with Kodak electron microscope film (Eastman Kodak Co., Rochester, NY).

**Immunochemistry and Epitope Tagging**

Early log phase yeast cells were prepared for immunofluorescence experiments by a modification of the filtration method of Kilmartin and Adams (1984). The NcoI/SalI and PstI/HindIII linkers were obtained via subcloning into pBS-L5 or pLC20H respectively. Notation "bp" = base pair, die of the open reading frame and coincident insertion of the selectable NUP49 was plasmid. Unique oligonucleotide primers were then utilized to sequence protein sequences of NUP116, NUPIOO, and NUP49 to the sequences in the two different markers, either the URA3 gene for a 2103 bp BamHI/HindIII fragment of NUP49 and the numeral following NsiI* designates the nucleotide position of this particular restriction site in the sequences of Figure 4. References for pJJ242 and pJJ215 (Jones and Prakash, 1990); pMAL-cRI (Maina et al., 1988); pRS315 (Sikorski and Hieter, 1989). Two different methods were employed to compare the protein sequences of NUP116, NUP100, and NUP49 to the sequences in the GenBank and EMBL data bases (Pearson and Lipman, 1988; Collins and Coulson, 1990). The DNA sequence similarities were found by searching the GenBank nucleic acid data base using the FASTA program (Pearson and Lipman, 1988). The NUPIOO was then made by removal of a 808 bp NsiI/XbaI fragment from the middle of the open reading frame and coincident insertion of the selectable marker URA3 (pSW55). The EcoRI fragment of pSW55 was transformed into the diploid strain W303 of S. cerevisiae and viable Ura* transformants were selected. Southern analysis was used to identify those strains in which NUP49 had been replaced by the integration of the nup49-1::URA3-disrupted copy (SWY3). Transformation of the heterozygous diploid strain SWY1 with pSW62 (an intact copy of NUP49 under control of its endogenous promoter in a centromere plasmid with the LEU2 selectable marker) yielded a strain which when induced to sporulate resulted in the recovery of a viable Ura-/Leu- haploid strain (SYll) with the expected Mendelian frequency (25%).

**Gene Disruptions**

Mutant alleles of NUP49, NUP100, and NUP116 were constructed by the integrative DNA transformation method of Rothstein (1991). A deletion of NUP49 was made by removal of a 808 bp NsiI/XbaI fragment from the middle of the open reading frame and coincident insertion of the selectable marker URA3 (pSW55). The EcoRI fragment of pSW55 was transformed into the diploid strain W303 of S. cerevisiae and viable Ura* transformants were selected. Southern analysis was used to identify those strains in which NUP49 had been replaced by the integration of the nup49-1::URA3-disrupted copy (SWY3). Transformation of the heterozygous diploid strain SWY1 with pSW62 (an intact copy of NUP49 under control of its endogenous promoter in a centromere plasmid with the LEU2 selectable marker) yielded a strain which when induced to sporulate resulted in the recovery of a viable Ura-/Leu- haploid strain (SYll) with the expected Mendelian frequency (25%).

**Table II. Plasmid Construction**

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Construction</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBS-KS+ BACKBONE</td>
<td></td>
</tr>
<tr>
<td>pSW61</td>
<td>Fragment from nucleotide 650 to 3766 of NUP116 locus in EcoRI of pBS</td>
</tr>
<tr>
<td>pSW40</td>
<td>Full length NUP49 locus (2328 bp EcoRI/NcoI fragment) in EcoRI/Sall of pBS with NcoI/Sall linker</td>
</tr>
<tr>
<td>pSW41</td>
<td>Fragment from nucleotide 715 to 4369 of NUP100 locus in EcoRI/Xhol of pBS</td>
</tr>
<tr>
<td>pSW54</td>
<td>BamHI/HindIII URA3 fragment from pJJ242 into BamHI/HindIII of pSW11</td>
</tr>
<tr>
<td>pSW55</td>
<td>NsiI/XbaI URA3 fragment from pJJ242 into NsiI/Xbal of pSW40</td>
</tr>
<tr>
<td>pSW56</td>
<td>NsiI/XbaI URA3 fragment from pJJ242 into NsiI/Xbal of pW41</td>
</tr>
<tr>
<td>pSW65</td>
<td>3370 bp NheI/NcoI fragment of NUP100 locus in YEp13 clone into XbaI/Sall of pBS with NcoI/Sall linker</td>
</tr>
<tr>
<td>pSW72</td>
<td>4928 bp BglII/Sall fragment of NUP116 locus in YEp13 clone into BamHI/Sall of pBS</td>
</tr>
<tr>
<td>pSW74</td>
<td>EcoRI/PstI TRP1 fragment of pJJ280 into NsiI/EcoRI of pSW72</td>
</tr>
<tr>
<td>pSW99</td>
<td>BamHI/PstI HIS3 fragment from pJJ215 into BamHI/PstI of pSW72</td>
</tr>
</tbody>
</table>

The NcoI/Sall and PstI/HindIII linkers were obtained via subcloning into pBS-L5 or pLC20H respectively. Notation “bp” = base pair, and the numeral following NsiI* designates the nucleotide position of this particular restriction site in the sequences of Figure 4. References for pJJ242 and pJJ215 (Jones and Prakash, 1990); pMAL-cRI (Maina et al., 1988); pRS315 (Sikorski and Hieter, 1989).
MAb192 Recognizes Yeast Nuclear Pore Complex Proteins

From a panel of monoclonal antibodies raised against either Triton X-100-treated rat liver nuclei, a rat liver nuclear pore complex/lamina fraction, or the WGA-binding proteins from Triton X-100-treated rat liver nuclei, a rat liver nuclear pore complex/lamina fraction (Davis and Blobel, 1986; Aris and Blobel, 1989; Davis and Fink, 1990; Carmo-Fonseca et al., 1991; Rout, M., and G. Blobel, manuscript in preparation). It has been demonstrated by Davis and Fink (1990) that these monoclonal antibodies, raised against rat antigens, can be successfully used to clone yeast nucleoporins from expression libraries. By a combination of immunofluorescence and Western blotting we tested a total of 29 different monoclonal antibodies from this panel for their cross-reactivity with yeast (data not shown). To isolate new nucleoporins by screening expression libraries, we used a high affinity monoclonal antibody, MAB192, that appeared to recognize previously undescribed antigens.

The indirect immunofluorecence results of MAB192 with the yeast S. cerevisiae are shown in Fig. 1. The staining with MAB192 is nuclear, distinctly punctate, and (depending on the focal plane of the particlar cell) also localized to the nuclear periphery. Several pairs of cells on the left side of this field appear to be dividing, based upon the DAPI staining and the extension of the punctate immunofluorescence staining through the bud neck (Davis and Fink, 1990). Punctate nuclear envelope staining is a characteristic and well documented property of antibodies directed against nuclear pore complex proteins in both mammalian and yeast cells (Davis and Blobel, 1986; Aris and Blobel, 1989; Davis and Fink, 1990; Nehrbass et al., 1990).

The definitive localization of MAB192 binding to the nuclear pore complexes of yeast has been accomplished by the use of postembedding immunoelectron microscopy. The binding of MAB192 to thin sections from a haploid yeast spheroplast was visualized by the presence of colloidal gold particles (10 nm) coated with a goat anti-mouse antibody. The micrographs in Fig. 2 clearly reflect the localization of MAB192 to the nuclear pore complexes in the nuclear envelope. The nuclear pore complex is visible as an electron dense patch spanning the clear area between the nuclear membranes. In all cases, gold particles localized to the nuclear envelope were coincident with the electron density typical of a pore complex. Therefore, the immunofluorescence signal with MAB192 reflects the antibody binding at pore complexes. The positions of over 400 gold particles in 50 typical cell sections was quantified (Table III). The calculation of gold particle density in the nuclear envelope versus either the nucleoplasm or the cytoplasm reflects the specificity of MAB192 for pore complex proteins. Overall, the localization of MAB192 to pore complexes in the nuclear envelope is at least 14 times higher than that in the nucleoplasm, and eightfold greater than that in the cytoplasm.

Because the binding of MAB192 was primarily, although not exclusively, to the nuclear pore complexes, the distribution of MAB192 reactive proteins during nuclear subfractionation was investigated. Yeast nuclei from S. uvarum were prepared by the method of Rozijn and Tonino (1964) with the modifications described by Rout and Kilmartin (1990). The nuclei are recovered in a single fraction with a total yield of ~80% as reflected by monitoring the fractionation of NUP1 (Fig. 3 B, lane 7), a known yeast nucleolar protein (Schimmang et al., 1989; Henriquez et al., 1990). This enriched nuclear fraction clearly contains several major coenriching bands with apparent molecular sizes of 49, 54, 65, 100, and 118 kD that are recognized by MAB192 (Fig. 3 A). A 35- and a 12-kD protein are present in spheroplasts (lane J) and a postnuclear supernatant (lane 2) but do not coenrich with the nuclei. The combination of the immunolocalization and cell fractionation data indicates that MAB192 recognizes a unique group of yeast proteins, the majority of which behave as nuclear pore complex proteins.

Results

MAb192 Recognizes Yeast Nuclear Pore Complex Proteins

Because the binding of MAB192 was primarily, although not exclusively, to the nuclear pore complexes, the distribution of MAB192 reactive proteins during nuclear subfractionation was investigated. Yeast nuclei from S. uvarum were prepared by the method of Rozijn and Tonino (1964) with the modifications described by Rout and Kilmartin (1990). The nuclei are recovered in a single fraction with a total yield of ~80% as reflected by monitoring the fractionation of NUP1 (Fig. 3 B, lane 7), a known yeast nucleolar protein (Schimmang et al., 1989; Henriquez et al., 1990). This enriched nuclear fraction clearly contains several major coenriching bands with apparent molecular sizes of 49, 54, 65, 100, and 118 kD that are recognized by MAB192 (Fig. 3 A). A 35- and a 12-kD protein are present in spheroplasts (lane J) and a postnuclear supernatant (lane 2) but do not coenrich with the nuclei. The combination of the immunolocalization and cell fractionation data indicates that MAB192 recognizes a unique group of yeast proteins, the majority of which behave as nuclear pore complex proteins.

Isolation of Genes Encoding New Yeast Nucleoporins NUP49, NUP100, and NUP116

We isolated clones from a yeast genomic λgt11 expression

The Journal of Cell Biology, Volume 119, 1992 708
Figure 1. Punctate nuclear rim staining of yeast cells by MAb192 as detected by indirect immunofluorescence microscopy. Formaldehyde fixed yeast cells (diploid strain W303) were incubated with MAb192, and binding was detected with an FITC-conjugated goat anti-mouse IgG (top). The coincident DAPI staining is also shown (bottom). Bar, 5 μm.

library based upon their cross-reactivity with MAb192. Only three of the isolates that were recognized by MAb192, clones 4A, 4B, and 4C, will be presented in this paper. Whole bacterial cell extracts from lysogenic strains for the respective positive clones were prepared before and after induction with 10 mM isopropylthiogalactoside (IPTG). The lysates were analyzed by SDS-polyacrylamide gel electrophoresis and subsequent Western blots were probed with
MAb192 (Fig. 3 C). Each of the three clones produced an IPTG induced, distinct major polypeptide strongly recognized by MAb192. Only one of the clones (4C) is a lacZ fusion protein (confirmed by DNA sequencing).

The complete DNA sequence of these three positive clones was determined, as well as the sequence of overlapping clones that were isolated from either the λgt11 library or a YEpl3-based, yeast genomic plasmid library (Nasmyth and Tatchell, 1980). The resultant nucleotide and predicted amino acid sequences are presented in Fig. 4, A–C. All three

Figure 2. MAb192 recognizes proteins in the nuclear pore complex of yeast cells. Immunolabeling of thin sections from postembedded yeast haploid spheroplasts was performed with MAb192 as described in the Methods. The micrograph in A shows a thin section of an entire cell. The 10-nm gold particles labeling nuclear pore complexes are indicated by arrowheads. A magnification of a cluster of labeled nuclear pore complexes in a different thin section is shown in B. Quantification of the MAb192 labeling in 50 representative cell sections is presented in Table III. Bars, 0.2 μm.
sequences contain a single large uninterrupted open reading frame. These three distinct genes are capable of encoding proteins of 1113, 959, and 473 amino acids with predicted sizes of 116.4, 100.0, and 49.2 kD, respectively. These values have assumed the use of the first AUG codon in each open reading frame, which in all three cases lies immediately downstream of an A residue at position –3, a highly conserved element for an optimal translational start site (Hamilton et al., 1987). Further upstream, sequences resembling the TATA consensus for transcriptional initiation exist (Struhl, 1987). Finally, the appropriate sequences for efficient transcriptional termination are found downstream of the first translational termination codon in all the open reading frames (Zaret and Sherman, 1982).

The genes have been designated NUP116, NUP100, and NUP49, where NUP stands for nucleoporin or nuclear pore complex protein as described by Davis and Fink (1990), and the numbering reflects the predicted molecular mass (kD) from the amino acid sequence.

The Amino-Terminal Regions of These New Nucleoporins Contain “GLFG” Repeats

Comparison of the amino acid sequences of NUP49, NUP100, and NUP116 to the sequences in the GenBank and EMBL data bases revealed that all three are novel proteins. Analysis of their primary sequence suggests that all three have at least two distinct structural regions that span the amino- and carboxy-terminal halves of each. Fig. 5 A diagrams these potential domains and the position of acidic and basic residues throughout the sequences. The amino-terminal half of each protein possesses an unusual charge distribution, with the virtual absence of acidic residues and the presence of a few well-spaced basic residues. In contrast, the carboxy-terminal half contains a mixture of charged residues resulting in an overall predicted isoelectric point of 5.94 for NUP49, 9.38 for NUP100, and 9.32 for NUP116. As reflected by the different shading of the carboxy-terminal regions in Fig. 5 A, the carboxy-terminal half of NUP49 diverges greatly from that of NUP100 and NUP116. However, the carboxy-terminal regions of NUP100 and NUP116 are strikingly similar. Over their entire carboxy-terminal domains, they are ~35% identical and 32% conserved for an overall similarity value of 67%. This homology is especially evident in the carboxy-terminal 150 amino acids of each (Fig. 5 B) where nearly 54% of the residues are identical and 24% are conserved (totaling 78%). We failed to detect any notable similarities between the carboxy-terminal regions of NUP49, NUP100, and NUP116 and any proteins in the data bases mentioned above.

A closer examination of the primary amino acid sequence of the amino-terminal regions reveals striking and characteristic amino acid repeat motifs. The repeats have been grouped in Fig. 6 with the four consecutive glycine-leucine-phenylalnine-glycine (GLFG) residues constituting the central core, and with a GLFG motif composed of at least three of the four residues. As verified by dot matrix analyses of the sequences with themselves, this GLFG motif is repeated 13, 29, and 33 times, respectively, in the individual amino-terminal regions of NUP49, NUP100, and NUP116. The spacer sequences between the GLFG repeats are unusually rich in asparagine (N), glutamine (Q), serine (S), and threonine (T) residues, so that even though these regions are relatively uncharged they are polar in nature. We have designated the amino-terminal regions of NUP49, NUP100, and NUP116 the “GLFG” domain.

Secondary structure analysis for NUP49, NUP100, and NUP116 suggests a complete segregation of helix forming residues to their carboxy-terminal halves (Gibrat et al., 1987). In contrast, the GLFG domain is predominantly composed of stretches of amino acid residues with a preference for maintaining either an extended or coil conformation, separated by putative turns. Therefore, the differences in the primary structure composition of the amino- and carboxy-terminal halves of these new nucleoporins may reflect different structural (and functional) domains.

The 3’ Noncoding Sequence of Both NUP100 and NUP116 Contain a Histidine tRNA Gene and Ty Delta Element

Although there was a remarkable similarity in amino acid sequence between regions of NUP116, NUP100, and NUP49, there is no especially notable conservation throughout their

1. Abbreviations used in this paper: GLFG, glycine-leucine-phenylalanine-glycine; HA, hemagglutinin antigen.

Table III. Distribution of MAb192 and 12CA5 Antibody in Immunoelectron Microscopy Experiments

<table>
<thead>
<tr>
<th>Number of Gold Particles</th>
<th>Density (gold particles/µm²)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nucleoplasm</td>
</tr>
<tr>
<td>MAb192</td>
<td>22</td>
</tr>
<tr>
<td>12CA5</td>
<td>12</td>
</tr>
</tbody>
</table>

The number and location of 10 nm gold particles in 50 cell thin sections of postembedded yeast haploid spheroplasts, representing either the binding of MAb192 or the monoclonal antibody 12CA5 in a strain expressing epitope tagged NUP49, are shown. *Because the number of pore complexes per section was variable, the area of the entire nuclear envelope was used to calculate the density at nuclear pore complexes. Therefore these values are an underestimate. The density of gold particles was calculated by using the average area of the nucleoplasm, nuclear envelope, and cytoplasm in thin sections of five typical cells.
open reading frame DNA sequences or the 700–900 bp of upstream sequence in these clones. However, there are two regions in the 3′ noncoding sequence of NUP116 and NUP100 that possess some surprising similarities. The 72 bp of boxed DNA sequence (Fig. 4, A and B, respectively) downstream of both NUP116 and NUP100 are identical to each other and its complementary sequence is identical to genes encoding the histidine tRNA of yeast S. cerevisiae (del Rey et al., 1983). These newly identified histidine tRNA genes end 230 and 324 bp downstream of the termination codons for NUP116 and NUP100, respectively.

The DNA sequences on either side of the respective histidine tRNA genes are divergent until a region either 145 or 130 bp from the tRNA gene initiation point (underlined sequence in Fig. 4, A and B, respectively). These underlined sequences again both possess striking homologies, and in this case are strongly related to the delta (δ) element of Ty1 transposons. The δ sequence elements in the genome of budding yeast have been found as either direct repeats that flank the Ty1 transposon or as solo elements without the central portion of Ty1 (Boeke, 1989). The underlined region of the NUP116 clone, beginning at 4516 bp, is 86.3% identical to the long terminal repeat of Ty1-H3 (Boeke et al., 1988). This homology continues into the central portion of Ty1-H3 until the NUP116 clone ends. In the NUP100 clone the underlined sequence from bp 4293 to 4607 appears to be a solo δ element possessing an 81.6% identity to the Ty1-H3 δ. The NUP100 solo δ element is in an inverted orientation relative to the tRNA gene when compared with the positioning of these elements in the NUP116 clone (see diagram in Fig. 4).

Another unusual feature of the NUP116 DNA sequence is the presence of an MluI restriction enzyme site ~208 bp upstream of the translation initiation codon. MluI sites have been found within the promoters of yeast genes whose

---

**Figure 3.** MAb192 recognizes a group of yeast nuclear pore proteins. (A) Western blot analysis of the MAb192 cross-reactive proteins in yeast cell fractionation. The Western blots were probed sequentially with MAb192, rabbit anti-mouse IgG, and 125I-protein A. (Lane 1) Whole spheroplast lysate; (lanes 2 and 3) postnuclear supernatant and crude nuclear pellet, respectively. Lanes 4–8 are fractions from the sucrose gradient centrifugation of the crude nuclear pellet. (Lane 4) Sample layer of the gradient; (lane 5) the interface between the sample layer and the 2.0-M sucrose-PVP layer; (lane 6) the 2.0–2.1-M interface; (lane 7) the 2.1–2.3-M interface containing most of the nuclei; (lane 8) the remainder of the 2.3-M layer. Five discrete bands in the enriched nuclei fraction (lane 7) are identified as p118, p100, p65, p54, and p49. Molecular mass markers are indicated on the left. (B) Western blot analysis with the monoclonal antibody D77 of the yeast cell fractionation. A Western blot identical to that above was similarly processed except that the primary monoclonal D77 (Henriquez et al., 1990) was employed to recognize NOP1, a yeast nucleolar protein of apparent molecular mass 38 kD. Estimates of the total milligrams of protein for each step of the fractionation, starting from a 36-liter preparation of early log phase cells, are shown along the bottom of the gel. Protein loadings for each lane for the Western blots in A and B are in proportion to the total protein recovered for the corresponding fractions; lanes 1–3 have 1/75,000 of the total protein recovered and lanes 4–8 have 1/25,000 of the total protein recovered. (C) Agt11 clones expressing MAb192 cross-reactive proteins. Proteins of cell lysates from lysogenic E. coli strains for three positive isolates (4 A, 4 B, and 4 C), before (−) and after (+) a 90-min induction with 10 mM IPTG, were separated by electrophoresis on a 7% SDS polyacrylamide gel and transferred to nitrocellulose. The blot was probed sequentially with MAb192, rabbit anti-mouse IgG, and finally 125I-protein A. Positions of the molecular mass markers (kD) are noted. The three bands in lane 2 reflect the sensitivity of this protein to proteolysis.
mRNA levels increase at the GI/S phase boundary of the cell cycle (Andrews and Herskowitz, 1990). For example, several genes encoding essential DNA replication proteins contain one or more upstream MluI sites and are transcriptionally regulated in a cell-cycle dependent manner (Brill and Stillman, 1991; Pizzagalli, et al., 1988). In general, the number of nuclear pore complexes doubles in S phase of the yeast cell cycle (Jordan et al., 1977). Therefore, in terms of studying the assembly of new nuclear pore complexes, the MluI site in the promoter of NUP116 and the potential for transcriptional regulation at the GI/S phase transition is intriguing. This is a unique feature of the NUP116 promoter as no MluI sites are present in the upstream sequences of NUP100 and NUP49 or in the two published yeast nucleoporin sequences.

**MAb192 Binds to the GLFG Domain**

Because MAb192 cross-reacts with all three of these new proteins it was of interest to test whether the antibody was recognizing the common GLFG domain. A series of male/NUP49 fusions was constructed and expressed in the E. coli strain JM109 by induction with IPTG (Fig. 7, A and B). The proteins encoded by these plasmids are referred to here by their respective lane numbers in Fig. 7. Fusions 3, 4, and 5 are progressively truncated from the amino-terminal end of NUP49. The 6, 7, and 8 fusions encode segments of the GLFG domain. Western blots of SDS polyacrylamide gels identical to that in Fig. 7 B were probed with either MAb192, MAb350, or MAb414. In Fig. 7 C, MAb192 exclusively recognizes the GLFG containing amino-terminal domain as indicated by the loss of reactivity with fusion 5 (beginning 3’ of the last GLFG repeat). Furthermore, NUP49 contains multiple, separable epitope sites for binding MAb192 because the nonoverlapping segments from the amino-terminal domain, contained in fusions 6 and 7, both cross-react (although with varying affinity, fusion 6 being weaker and having about two GLFG repeats whereas fusion 7 has five repeats). The minimal fragment necessary for MAb192 binding may be that which is underlined in Fig. 4 C (from comparing fusion 7 and 8). The GLFG repeat alone is not sufficient for MAb192 binding as fusion 8 and a construct with a linker encoding an insertion of only the residues GGLFGN (data not shown) did not cross-react with MAb192.

This study has also demonstrated that the epitope recognized by MAb192 differs from that of MAb350 and MAb414. Fusions 3 and 6 were the only constructs that produced MAb350 (Fig. 7 D) or MAb414 (data not shown) cross-reactive proteins. Therefore, because MAb192 could recognize all the proteins encoded by fusions 3, 4, 6, and 7, MAb350 and MAb414 must bind a distinct epitope that is encoded exclusively by fusions 3 and 6. MAb350 was used to clone an epitope within the 9-amino acid repeat domain of both NUP1 and NSPI. Our results are consistent with these previous investigations if MAb350 is recognizing the single GFSFG sequence in NUP49 (see Fig. 6) that is encoded exclusively in fusions 3 and 6 and that is similar to the internal consensus of the 9-amino acid repeats (Davis and Fink, 1990).

**NUP49 is an Essential Gene**

Gene disruption experiments have been conducted with all three genes, and Fig. 8 A displays the respective constructs that have been employed to make heterozygous diploid strains. Sporulation of the nup49-1::URA3 heterozygous, diploid strain (SWY1) generated only two viable, ura- spores per tetrad (Fig. 8 B). The lethal phenotype of the nup49-1::URA3 disruption was rescued by the presence of a single-copy plasmid bearing an intact NUP49 gene (pSW62 in haploid strain SWY11). However, the disruption was not complemented by a truncated NUP49 gene with two stop codons inserted after amino acid #76 (pSW64). Overexpression from NUP100 or NUP116 on 2 µm plasmids also does not complement the nup49-1::URA3 disruption. These experiments demonstrate that NUP49 is essential for viability. Microscopic examination of 120 dissected nup49-1::URA3 spores revealed that growth was arrested in the large budded stage before or after at most one cellular division.

**Single Disruptants of NUP100 and NUP116 are Viable**

The sporulation and dissection of the heterozygous nup100-1::URA3 diploid strain resulted in the recovery of four viable spores (Fig. 8 B). Replica plating of the tetrads to selective media confirmed that the Ura+ markers segregated with the expected 2:2 ratio. The nup100-1::URA3 disruption conferred no obvious growth defects nor any gross morphological differences as compared to wild type cells by electron microscopy (data not shown). When the phenotype of a gene disruption of NUP116 (nup116-6::URA3) was examined, the initial appearance of only healthy ura- colonies derived from dissected daughter spores was consistent with a lethal phenotype. However, prolonged incubation at 30°C revealed the presence of slow growing Ura+ colonies. When these nup116-6::URA3 haploid strains were grown in liquid culture at 30°C, their doubling times were ~2.5 times longer than that of wild type haploids. The growth of the nup116-6::URA3 strain is also temperature sensitive such that at 37°C the allele is lethal. Identical results have been obtained with a different NUP116 disruption construct that has replaced a fragment from the 5’ most BamHI site shown in Fig. 8 A to 10 bp 5’ from the stop codon with the HIS3 marker. The dramatic growth defect of NUP116-deficient cells was nearly completely rescued by the presence of a single-copy plasmid bearing an intact NUP116 gene (pSW75 in SWY54) but was not altered by a similar plasmid with an internal deletion in the NUP116 gene (pSW100).

As described in the previous sections, MAb192 recognizes the amino-terminal GLFG domain of NUP49. Therefore, the major MAb192 cross-reactive proteins in the cell fractionation shown in Fig. 3 A may all be proteins that contain GLFG motifs. Protein extracts from yeast strains that lack either NUP100 or NUP116 were prepared and analyzed by Western blotting with MAb192 (Fig. 8 C). The clear absence of the respective protein bands in the disrupted strains confirms that pl100 and pl18 are encoded by NUP100 and NUP116.

Because of the remarkable homology between NUP100 and NUP116, the phenotype of a diploid strain harboring both the nup100-1::URA3 and nup116-5::HIS3 disruptions was examined (SWY56). This doubly disrupted, heterozy-
Localization of Epitope Tagged Gene Products by Immunofluorescence Microscopy

The individual gene products were tagged with a unique epitope to independently confirm their localization to the nuclear pore complex in yeast. A sequence encoding a nine amino acid epitope derived from the influenza hemagglutinin antigen (HA) (Wilson et al., 1984) was inserted in frame at the NsiI site of the HA linker. The histidine tRNA gene sequence is boxed and the DNA sequence homologous to Ty element is underlined. The sequence data are available from EMBL/GenBank/DDBJ under accession number Z15035. (C) Nucleotide and predicted amino acid sequence of the locus. The predicted amino acid sequence (in single letter code) begins with the first AUG codon in the open reading frame and ends at the bold faced termination codon.

Figure 4. (A) Nucleotide and predicted amino acid sequence of the NUP116 locus. The predicted amino acid sequence (in single letter code) begins with the first AUG codon in the open reading frame and ends at the bold faced termination codon. The stars (below the base) designate the 5' and 3' ends of the hgtl4A clone. The arrow marks the insertion site (NsiI) for the HA linker. The histidine tRNA gene sequence is boxed and the DNA sequence homologous to Ty element is underlined. The sequence data are available from EMBL/GenBank/DDBJ under accession number Z15035. (B) Nucleotide and predicted amino acid sequence of the NUP100 locus. The predicted amino acid sequence (in single letter code) begins with the first AUG codon in the open reading frame and ends at the bold faced termination codon. The stars (below the base) designate the 5' and 3' ends of the hgtl4B clone. The arrow marks the insertion site (NsiI) for the HA linker. The histidine tRNA gene sequence is boxed and the DNA sequence homologous to Ty element is underlined. The sequence data are available from EMBL/GenBank/DDBJ under accession number Z15035. (C) Nucleotide and predicted amino acid sequence of the NUP49 locus. The predicted amino acid sequence (single letter code) begins with the first AUG codon in the open reading frame and ends at the bold faced termination codon. The underlined amino acid sequence may be the minimal region necessary for MAb192 binding. The stars (below the base) designate the 5' and 3' ends of the hgtl4C clone. The arrow marks the insertion site (NsiI) for the HA linker. The sequence data are available from EMBL/GenBank/DDBJ under accession number Z15040.
Figure 5. Structural features of NUP49, NUP100, and NUP116. (A) The domain structure of NUP49, NUP100, and NUP116. The sequence of each of the three proteins is represented here as containing at least two regions; an amino terminus (lightly shaded) and a carboxy terminus (unshaded). The boundaries of the domains are indicated by thick lines and the positions of acidic (A, top line) and basic (B, bottom line) residues are noted, respectively, as aspartic acid (intermediate bar), histidine (full bar) and asparagine (light bar). (B) Alignment of the NUP100 and NUP116 carboxy-terminal domains. An ALIGN analysis (Dayhoff et al., 1983) between the sequences of NUP100 and NUP116 revealed a significant homology. The center line designates the identical (capital letter) and conserved (:) residues.

Figure 6. A new family of nucleoporins characterized by GLFG repeats in their amino-terminal domains. The amino-terminal protein sequences of NUP49, NUP100, and NUP116 have been aligned such that the consensus GLFG repeat is shown in the left column. Bold-faced type highlights the G, L, F, or G residues in these repeats and any paired FG residues in the spacer sequences. The underlined and bold-faced "FXF" sequences are the only such double phenylalanines in each.

Wente et al. A New Family of Yeast Nucleoporins
C strongly suggests that NUP49 and NUP116 are nuclear pore complex proteins.

Exhaustive efforts were made to obtain a clear signal for the nup100-3::HA localization, including a wide variety of fixation and wash conditions. Although extremely faint punctate nuclear rim staining has occasionally been observed for a number of conditions, it has so far proven impossible to record this. The nup100-3::HA protein is being expressed at approximately the same levels as the nup116-3::HA protein as detected by Western blotting of crude spheroplast samples from all the above strains (data not shown). The HA epitope recognized by the 12CA5 antibody has been found to be relatively insensitive to formaldehyde fixation in other nucleoporins (Davis and Fink, 1990) and the homology between NUP116 and NUP100 would suggest that the problem does not lie in the positioning of the tag, which was placed at a similar point in the sequences of both proteins. We therefore propose that the relative level of immunofluorescence signal for these proteins reflects their accessibility to the antibody as defined by their position within the nuclear pore complex, the steric blocking by other nuclear pore complex proteins, and the degree to which the fixation procedure preserves the relative positions of these proteins. In the case of NUP100, these factors may have conspired to render it inaccessible to the antibody. Similar access problems have been observed in another densely packed structure, the mammalian midbody (Saxton et al., 1984). In the light of the strong homology to NUP116 and the coenrichment data discussed below, we believe that NUP100 is also a nuclear pore complex protein.

Immunoelectron Microscopic Localization of Epitope Tagged NUP49 to the Nuclear Pore Complex of Yeast Cells

The expression of epitope tagged proteins was designed to facilitate the localization of an individual protein with a monospecific reagent (12CA5) versus the localization of the entire family with the polyspecific MAb192. Using the same techniques we had employed to visualize MAb192 binding with immunoelectron microscopy, the position of the epitope tagged NUP49 has been determined. Thin sections from

![Figure 7](image-url)

**Figure 7.** MAb192 recognizes multiple epitopes in the GLFG domain of NUP49. (A) A diagram compares the full length NUP49 to the fusion protein products expressed from respective pSW plasmids. The numbers along side each fusion reflect the amino acid residue boundaries of the NUP49 inserts. the malE/NUP49 plasmids were constructed as described in the Methods. The GLFG domain is lightly shaded, the NUP49 carboxy-terminal domain is unshaded, maltose binding protein is blackened, and LacZ sequences are striped. (B) Cell lysates from E. coli JM109 strains transformed with the fusion plasmids were prepared after a 2 h induction with 0.3 mM IPTG (lanes 2-8). Equivalent protein samples were separated by electrophoresis on a 7% SDS polyacrylamide gel and visualized by Coomassie blue staining. Lane I is a sample from the strain with pMAL-cRI before IPTG addition. The individual strains from whence the samples for lanes 2-8 were generated correspond to the plasmid diagrammed directly above in Fig. 7A (lane 2, pMAL-cRI; lane 3, pSW82; lane 4, pSW67; lane 5, pSW80; lane 6, pSW85; lane 7, pSW81; lane 8, pSW95). (C) A Western blot of a gel identical to that in Fig. 7B was probed sequentially with MAb192, rabbit anti-mouse IgG, and 125I-protein A. (D) A Western blot as in Fig. 7C except that the primary monoclonal antibody was MAb350.
postembedded spheroplasts of the haploid yeast strain expressing nup49-2::HA (SWY12) were successively incubated with the monoclonal 12CA5 antibody against the HA epitope and 10 nm colloidal gold coated with goat antimouse antibodies. The micrograph in Fig. 10 shows the localization of gold particles to the nuclear pore complexes. Quantification of the gold particles in 50 representative cell thin sections highlights that this is specific binding at the nuclear pore complexes (Table III). Therefore, the immunoelectron microscopy results confirm that, as with MAb192, the punctate nuclear rim staining in the epitope-tagged NUP49 immunofluorescence studies is due solely to nuclear pore complex binding.

Discussion

We have isolated and characterized three genes that encode a novel class of yeast nuclear pore complex proteins; NUP49, NUP100, and NUP116. These are nuclear pore complex proteins as concluded from a combination of immunolocalization and subcellular fractionation data. Immunoelectron microscopy with MAb192, which recognizes all three proteins, showed specific localization to the nuclear pore complex and not the nuclear envelope (Fig. 2). The localization of each individual protein has been addressed by the use of epitope tagging. Both the tagged NUP49 and NUP116 resulted in a punctate nuclear rim staining in immunofluorescence experiments. The definitive localization to the nuclear pore complexes by immunoelectron microscopy has been presented for only NUP49. Similar experiments are underway with tagged NUP116 and NUP100. However, the subcellular fractionation data substantiates the conclusion that NUP49, NUP100, and NUP116 are nuclear pore complex proteins not only by their cofractionation with yeast nuclei (as shown in Fig. 3 A), but also by their exclusive enrichment with the isolation of nuclear pore complexes (Rout, M., and G. Blobel, manuscript in preparation).

The criteria for classifying these three proteins as members of a new family are based upon comparisons with the known yeast nucleoporins. It was found that the only similarities are restricted to the GLFG domains of NUP116, NUP100, and NUP49. The amino-terminal region of NSP1 and the carboxy-terminal region of NUP1 appear to contain at most six to nine degenerate versions of the GLFG motif flanked by uncharged spacers (Fig. 11). In comparison the GLFG regions of NUP49, NUP100, and NUP116 are more extensive and more tightly conserved than those found in NUP1 and NSP1. Searches of the available protein data banks do not reveal any other sequences with such GLFG motifs.

Surprisingly, NUP49, NUP100, and NUP116 lack the 9-amino acid repeat domain found in NUP1 and NSP1. The core consensus sequence of the GLFG repeat motifs and
The 9-amino acid repeats are clearly distinct. Moreover, the spacer sequences are of an entirely different nature, highly charged for the 9-amino acid repeat domains and uncharged, Q, N, S, and T rich for the GLFG repeat regions. The aminoterminal regions of NUP49, NUP100, and NUP116 each contain a single “FXF” sequence (see Fig. 6) which is reminiscent of the central core sequence of a 9-amino acid repeat. However, the context of the surrounding sequence is not homologous to the 9-amino acid repeat domains. It is the repetitive nature of both the GLFG motif and the 9-amino acid repeat that has been the basis for the designation of these primary structure domains. Secondary structure prediction programs (Gibrat et al., 1987) suggest that the GLFG region does not form the same secondary structure element that has been predicted for the 9-amino acid repeat domains (an amphipathic β barrel, Davis and Fink, 1990).

We propose that families of yeast nucleoporins can be classified by their composition of primary structure domains. So far two different modular protein domains are distinguishable, namely the 9-amino acid repeat domain and the GLFG domain. NSP1 and NUP1 have both the 9-amino acid repeat domain and a degenerate GLFG domain, whereas the three new NUPs have only a conserved GLFG domain. Therefore, NUP49, NUP100, and NUP116 resemble each other far more than they do NUP1 or NSP1. Thus, they may constitute a class distinct from that of NUP1 and NSP1. It is possible to imagine a variety of nucleoporin subfamilies arising from other combinations of the GLFG domain, the 9-amino acid repeat domain, and any, as yet, unidentified domains.

Our epitope mapping studies have shown that MAb192 recognizes multiple sites in the GLFG domain of NUP49 and, by inference, in the other proteins it recognizes. This recognition by MAb192 may require more than a simple linear peptide sequence, and it may bind to the antigen via a characteristic secondary structure of the GLFG domain (Laver et al., 1990). This work has extended the understand-
Because MAb192 was generated from rat nuclear antigens, mammalian homologs of this yeast GLFG family can probably also be isolated. In fact, peptide sequences from several purified rat liver nuclear pore complex proteins have been found to contain GLFG motifs (Woźniak, R., and G. Blobel, unpublished results). The complete sequence of only a single mammalian nucleoporin, p62, has been reported so far. The amino-terminal region of p62 has been described as a 9-amino acid repeat domain, but the spacer sequences flanking the GFxFG repeats are uncharged and serine/threonine rich (Starr et al., 1990; Starr and Hanover, 1991; Cordes et al., 1991; Carmo-Fonseca et al., 1991). The carboxy-terminal regions of vertebrate p62 and NSP1 share significant sequence similarities (Carmo-Fonseca et al., 1991; Cordes et al., 1991; Starr and Hanover, 1991). More mammalian nucleoporins must be sequenced before the correlations to the yeast nucleoporin subfamilies can be drawn.

All three of these new genes have dramatically different phenotypic effects on the normal growth of a yeast cell. The NUP49 gene product is required for cell viability, and it has a unique carboxy-terminal region attached to the GLFG domain. The amino-terminal region of NUP1 and the carboxy-terminal region of NSP1 are also both encoded by essential genes and appear to be unique domains amongst the known yeast nucleoporins (Davis and Fink, 1990; Hurt, 1988; Nehrbass et al., 1990). The functional roles of the individual domains of the yeast nucleoporin NSP1 have been examined. The carboxy-terminal domain of NSP1 has been shown to be sufficient for localization to the nuclear pore complex and complementation of a lethal null strain. Furthermore, a temperature sensitive mutation in this carboxy-terminal domain arrested cell growth and prevented the accumulation of a nucleolar protein (Nehrbass et al., 1990; Hurt, 1990). The cellular functions of the two nonessential domains of NSP1, namely the 9-amino acid repeat region and the degenerate GLFG domain, have not yet been determined.

In contrast with the other yeast nucleoporins, the predicted amino acid sequences of NUP100 and NUP116 are highly conserved over their entire length and therefore, individually, they lack a unique domain. This suggests they may have been derived by evolution from a common ancestral gene. The histidine tRNA genes in our NUP100 and NUP116 clones are the fourth and fifth histidine tRNA genes found to date. In 1983, del Rey et al. (1983) reported that there were at least seven histidine tRNA genes in the genome of yeast S. cerevisiae. The striking DNA conservation between the three noncoding regions of NUP100 and NUP116 may very well be a remnant of a previous recombination event. Because of the dramatically different phenotypes observed in NUP116-deficient compared to NUP100-deficient cells, it may actually be the nonhomologous regions between these two proteins that are important for their discrete functions. Further characterization of the slow growth, temperature sensitive phenotype of cells lacking NUP116 may also yield functional insights.

If domains with extensive stretches of the GLFG repeat motif or the 9-amino acid repeat motif are exclusive features of nucleoporins, then the type of domain usage in nuclear pore complex proteins could be a reflection of common structural or functional properties of a given nucleoporin subfamily. For example, the localization of proteins with

---

**Figure 11.** Yeast nucleoporin subfamilies. The proposed domain structures of NUP1, NSPI, NUP49, NUP100, and NUP116 are shown. The amino-terminal regions of NUP49, NUP100, and NUP116 are characterized by the GLFG motif (boxes on a line). The central regions of both NUP1 and NSPI are composed of a 9-amino acid repeat domain (diamonds; Davis and Fink, 1990; Nehrbass et al., 1990). The amino-terminal domain of NSPI and the carboxy-terminal domain of NUP1 resemble the GLFG domains in NUP49, NUP100, and NUP116. This limited similarity is due, coincidentally, to the lack of acidic residues and only spaced basic residues and to the presence of, at most, six to nine degenerate GLFG repeats in these regions of NSPI and NUP1. The number of individual boxes or diamonds reflects the number of repeats in the respective domain. The homologous carboxy-terminal regions of NUP100 and NUP116 are shaded whereas unique domains amongst these proteins are unshaded. The only common feature between all five of these nucleoporins is the presence of the GLFG domain. Based upon the complexity of the MAb192 cross-reactivity with isolated yeast nuclei, we believe that there are at least five different nuclear pore complex proteins that contain extensive stretches of the conserved GLFG motif and therefore belong to this new family of nucleoporins. MAb350 and MAB414, for example, recognize an epitope distinct from MAb192 and contained within the 9-amino acid repeat do-

---

**Figure 11.** Yeast nucleoporin subfamilies. The proposed domain structures of NUP1, NSPI, NUP49, NUP100, and NUP116 are shown. The amino-terminal regions of NUP49, NUP100, and NUP116 are characterized by the GLFG motif (boxes on a line). The central regions of both NUP1 and NSPI are composed of a 9-amino acid repeat domain (diamonds; Davis and Fink, 1990; Nehrbass et al., 1990). The amino-terminal domain of NSPI and the carboxy-terminal domain of NUP1 resemble the GLFG domains in NUP49, NUP100, and NUP116. This limited similarity is due, coincidentally, to the lack of acidic residues and only spaced basic residues and to the presence of, at most, six to nine degenerate GLFG repeats in these regions of NSPI and NUP1. The number of individual boxes or diamonds reflects the number of repeats in the respective domain. The homologous carboxy-terminal regions of NUP100 and NUP116 are shaded whereas unique domains amongst these proteins are unshaded. The only common feature between all five of these nucleoporins is the presence of the GLFG domain. Based upon the complexity of the MAb192 cross-reactivity with isolated yeast nuclei, we believe that there are at least five different nuclear pore complex proteins that contain extensive stretches of the conserved GLFG motif and therefore belong to this new family of nucleoporins. MAb350 and MAB414, for example, recognize an epitope distinct from MAb192 and contained within the 9-amino acid repeat domain (Davis and Fink, 1990).

Based upon the complexity of the MAb192 cross-reactivity with isolated yeast nuclei, we believe that there are at least five different nuclear pore complex proteins that contain extensive stretches of the conserved GLFG motif and therefore belong to this new family of nucleoporins. MAb350 and MAB414, for example, recognize an epitope distinct from MAb192 and contained within the 9-amino acid repeat domain (Davis and Fink, 1990).

---

**Figure 11.** Yeast nucleoporin subfamilies. The proposed domain structures of NUP1, NSPI, NUP49, NUP100, and NUP116 are shown. The amino-terminal regions of NUP49, NUP100, and NUP116 are characterized by the GLFG motif (boxes on a line). The central regions of both NUP1 and NSPI are composed of a 9-amino acid repeat domain (diamonds; Davis and Fink, 1990; Nehrbass et al., 1990). The amino-terminal domain of NSPI and the carboxy-terminal domain of NUP1 resemble the GLFG domains in NUP49, NUP100, and NUP116. This limited similarity is due, coincidentally, to the lack of acidic residues and only spaced basic residues and to the presence of, at most, six to nine degenerate GLFG repeats in these regions of NSPI and NUP1. The number of individual boxes or diamonds reflects the number of repeats in the respective domain. The homologous carboxy-terminal regions of NUP100 and NUP116 are shaded whereas unique domains amongst these proteins are unshaded. The only common feature between all five of these nucleoporins is the presence of the GLFG domain. Based upon the complexity of the MAb192 cross-reactivity with isolated yeast nuclei, we believe that there are at least five different nuclear pore complex proteins that contain extensive stretches of the conserved GLFG motif and therefore belong to this new family of nucleoporins. MAb350 and MAB414, for example, recognize an epitope distinct from MAb192 and contained within the 9-amino acid repeat domain (Davis and Fink, 1990).

Based upon the complexity of the MAb192 cross-reactivity with isolated yeast nuclei, we believe that there are at least five different nuclear pore complex proteins that contain extensive stretches of the conserved GLFG motif and therefore belong to this new family of nucleoporins. MAb350 and MAB414, for example, recognize an epitope distinct from MAb192 and contained within the 9-amino acid repeat domain (Davis and Fink, 1990).
both the 9-amino acid repeat and the GLFG repeat domains (NUP1 and NSPI) might be to different substructures of the nuclear pore complex than that of proteins with only a GLFG domain (NUP49, NUP100, and NUP116). It is also possible that the GLFG domain is designed to mediate a particular nuclear pore complex property, be it for interactions with another class of nucleoporins or with transport substrates. The current literature is filled with numerous and diverse examples of similar domains and amino acid repeat motifs that are responsible for a common property of functionally distinct proteins (Doolittle, 1989); such as the “EGF domains” in the LDL receptor (Südhoff et al., 1985), the leucine zippers and Zn-fingers of DNA binding proteins (Harrison, 1991), or the motifs shared by RNA binding proteins (Bandziulis et al., 1989). As more is learned about all the nucleoporins, it is likely that the combinations of unique and shared domains will mirror distinct and common functional features of the nucleoporin subfamilies.

At this stage we propose that the essential NUP49 has a distinct function from that of the other two essential nucleoporins, NUP1 and NSPI. Furthermore NUP49 may play a different role than NUP100 and NUP116 because of its unique carboxy-terminal domain. Our future work will focus on localizing the individual GLFG family members to the substructures of the yeast nuclear pore complex and delineating the functional role of these proteins in nucleocytoplasmic traffic.

This paper is dedicated to George E. Palade. We are especially indebted to Eleana Sphicas for providing her technical expertise in conducting the immunoelectron microscopy experiments. We are also grateful to Thomas Meier for generously providing the MAB192 cell line. We thank Chris Hardy for numerous reagents and discussions concerning yeast cloning and genetics; Richard Wozniak for providing peptide sequences; Edward Eisenstein for advice and assistance with protein domain (NUP49, NUP100, and NUP116). It is also possible that the combinations of unique and shared domains will mirror distinct and common functional features of the nucleoporin subfamilies.

Received for publication 11 June 1992 and in revised form 3 August 1992.

References


