Analysis of the Localization and Topology of Nurim, a Polytopic Protein Tightly Associated with the Inner Nuclear Membrane*

Helmut Hofemeister and Peter O’Hare‡
From the Marie Curie Research Institute, The Chart, Oxted, Surrey RH8 0TL, United Kingdom

Nurim is an inner nuclear membrane (INM) protein that was first isolated in a visual screen for nuclear envelope-localizing proteins. Nurim lacks an N-terminal domain characteristic of other INM proteins examined to date and may represent a class of proteins that localize to the INM by a distinct mechanism. To further characterize this protein, we constructed nurim-green fluorescent protein fusions and analyzed aspects of localization, biochemistry, and membrane topology. Results from immunoprobing and protease protection assays together with other analyses indicate that nurim (total length of 262 residues) is a six transmembrane-spanning protein and contains a hairpin turn in its C-terminal transmembrane domain, resulting in the N and C termini residing on the same side of the membrane. A loop region between the fourth and fifth transmembrane domains is exposed toward the nucleoplasm and contains a region accessible for site-specific endoprotease cleavage. In biochemical fractionation, nurim remained extremely tightly bound to nuclear fractions and was released in significant quantities only in the presence of 4 M urea. Under conditions in which nuclear lamins were completely extracted, a significant population of nurim remained resistant to solubilization. This tight binding requires the C-terminal region of the protein. DNase treatment only marginally influenced its retention characteristics in nuclei. Results from consideration of sequence alignments and identification of specific topological features of nurim indicate that it may possess enzymic function. These results are discussed with reference to the retention mechanism and possible nuclear function of nurim.

The nuclear envelope consists of two distinct membrane compartments. The outer nuclear membrane is continuous with the cytoplasmic endoplasmic reticulum, whereas the inner nuclear membrane (INM)1 is underpinned by the nuclear lamina and chromatin and is characterized by the selective recruitment of a number of integral membrane proteins, resulting in a unique composition (1, 2). The outer and inner nuclear membranes are joined together at the nuclear pore complex, the center of which forms a channel controlling selective transport of molecules between the cytoplasm and the nucleoplasm (3, 4). The mechanism by which transmembrane proteins are selectively recruited to the INM has been subject to considerable investigation. So far, no signal sequence, equivalent to the basic nuclear localization sequence of soluble nuclear proteins, has been found for INM proteins. It is thought that nuclear transmembrane proteins localize to the INM by a diffusion and retention mechanism (5–7). After synthesis at the endoplasmic reticulum and insertion into endoplasmic reticulum membranes, proteins diffuse through the membrane and around the junction of the nuclear pore complex. At the INM, these proteins appear to form selective interactions with other resident nuclear components, resulting in enrichment in the INM. A major component of the nucleus underpinning its structure and stability is the nuclear lamina (8–10), and interactions with lamina components (e.g. lamina A/C or B) are thought to represent at least one mechanism for INM protein localization (11, 12). Although a growing number of integral nuclear membrane proteins have been identified recently (13), only a few have been analyzed in any detail. The lamin B receptor (LBR) was one of the first characterized INM proteins (14–16). It encompasses a predicted eight-transmembrane C-terminal domain together with a hydrophilic 200-residue N-terminal domain, which anchors the LBR in the INM and can transfer INM localization to other membrane proteins (17). The N-terminal domain has been reported to bind lamin B, DNA, and the heterochromatin protein HP1 (18–20). Although the precise relative contributions of each of these activities are unclear, it is thought that these interactions are involved in targeting and retention of the LBR to the INM and perhaps vice versa, in the case of HP1, in targeting it to the nuclear membrane (19). Other known INM proteins include the lamin-associated proteins LAP1 and LAP2, emerin, and MAN1 (21–25), which each contain a large hydrophilic N-terminal domain. In LAP2, the N-terminal domain encompasses subregions involved in lamin B binding and in binding to chromatin (2, 6, 26). Emerin has also been reported to interact with both lamins A and B (27, 28). Interestingly, the N-terminal domains of MAN1, emerin, and LAP2β encompasses a well conserved region of 40 residues termed the LEM domain, which has been shown to bind BAF, a chromatin-associated protein (29, 30). Nevertheless, the precise contributions of each of these binding interactions to selective association with the INM remains to be established; and for example, in LBR, selective INM recruitment was also reported for the C-terminal transmembrane-spanning region when the N-terminal domain was deleted (17).

In this study, we present a further characterization of a recently identified INM protein, nurim (31). Unlike the other INM proteins analyzed so far, nurim lacks a large hydrophilic N-terminal domain and contains just four or five residues up-
stream of its first transmembrane segment. Nurim was nevertheless observed to be tightly associated with the INM; and based on the lack of any similarity in sequence or organization to the other INM proteins, nurim was classified as a new kind of INM protein. Here, we examine INM association and the topological model of nurim in the membrane using green fluorescent protein (GFP) fusion derivatives of the protein in localization studies, biochemical extraction studies, and endoprotease cleavage analysis. Consistent with earlier results, nurim is bound extraordinarily tightly to the INM, requiring of INM protein. Here, we examine INM association and the orientation of the termini in membrane fractions, together with consideration of the conservation of a particular C-terminal motif, indicate a revision of the organization of nurim topology in the membrane to a six-transmembrane domain (TMD) model. In addition, we present the results from analysis of sequence conservation in nurim that highlight similarities to a class of enzymes, isoprenylcysteine carboxymethyltransferases (ICMTs), indicating that nurim may possess an enzyme activity for the INM.

**EXPERIMENTAL PROCEDURES**

**Plasmids**—To create the nurim-GFP constructs, nurim was amplified from plasmid pVL154 (31), kindly provided by Tom Rapoport. The sense (5'-GGACGCGACTTCCCCTGACGCTTCTGAT) and antisense (5'-TATAAGGGATCATCTGCTCCTCCCATCATC) primers used to amplify nurim (amino acids 2–262) contained restriction sites for subsequent cloning. The PCR fragment was digested with HindIII and KpnI and inserted between the HindIII and KpnI sites of the pEGFP-C3 vector (Clontech) to yield plasmid pGFP-nurim. For the C-terminal GFP fusion construct, the sense (5'-GGCGCGAGGCTTCTCCCCATGAGGATCCT) and antisense (5'-TACCTCTGCTCCCTCCCATGCTTGGG) primers were used, and the amplified fragment was digested with HindIII and AgeI and inserted into the similarly digested pEGFP-N1 vector to yield plasmid pNurim-GFP. The truncated variant containing residues 1–187 (pNurim.C1-GFP) was derived from pNurim-GFP using the endogenous restriction sites SmaI and AgeI. After digestion and blunt end formation using T4 polymerase, the plasmid was religated to yield the correct in-frame deletion mutant. The sequences of all constructs were confirmed by sequencing.

**Cells and Transfections**—HeLa cells were grown on GlutaMAX I supplemented with 10% fetal calf serum, 100 units/ml penicillin, and 100 μg/ml streptomycin and incubated at 37 °C in an atmosphere containing 5% CO2. For transfections, 2×10^6 cells were plated in 35-mm dishes and transfected with 0.3 g plasmid DNA/5 ml of PBS, the cells were incubated on ice for 5 min in buffer containing 10 mM HEPES (pH 7.9), 80 mM KCl, 16 mM NaCl, 1.5 mM MgCl2, 1 mM dithiothreitol (DTT), 30% glycerol, 0.5% Triton X-100, and Complete® protease inhibitor mixture (Roche Applied Science). Cells were then imaged for direct GFP fluorescence or indirect immunofluorescence as described above. Images were routinely acquired using a Zeiss LSM 410 confocal microscope with a Plan-Apochromat 63× oil immersion objective lens (numerical aperture of 1.4) and zoom factors ranging from 1 to 8 of the LSM 410 acquisition software.

**Cell Homogenization and Extraction**—To isolate nuclei, cells from 60-mm dishes were washed with PBS and harvested in 1 ml of PBS supplemented with 5 mM EDTA and 1 mM DTT. Detached cells were then transferred to an Eppendorf tube, pelleted at 800 × g for 3 min, and resuspended in 1 ml of hypotonic homogenization buffer (10 mM HEPES, 10 mM KCl, 1.5 mM MgCl2, and 1 mM DTT) and centrifuged at 800 × g for 5 min at 4 °C. The nuclear pellets were further purified by resuspension and centrifugation three times in cushion of 1 ml extraction buffer. The 1.9 ml extract supernatant was collected and resuspended at 20,000 × g for 5 min at 4 °C. Nuclei were then resuspended in 1 volume of extraction buffer and used directly or frozen at −70 °C. For further extraction with progressively more stringent conditions (see Fig. 5), purified nuclei were resuspended in 250 μl of extraction buffer containing 0.5% Triton X-100 and Complete® protease inhibitor mixture. The samples were centrifuged at 800 × g for 5 min at 4 °C to separate cytosolic and nuclear pellets. A final treatment of extraction buffer containing 0.5% Triton X-100 and Complete® protease inhibitor mixture with 0.25 μg/ml EndoLys-C (Roche Applied Science) and 20 μg/ml of RNase A (QIAGEN Inc.) for 1 h at 37 °C followed by centrifugation at 20,000 × g for 5 min at 4 °C. Equal aliquots of DNase I-treated and untreated nuclei were further distributed into 250 μl of extraction buffer containing 0.5% Triton X-100 supplemented with 250 μM (NH4)2SO4, 2 mM urea, 4 M urea, or 20 mM DTT. The samples were then incubated for 30 min on ice and centrifuged at 21,000 × g for 5 min at 4 °C. Soluble supernatants were transferred into fresh tubes and precipitated using chloroform/methanol. Precipitated material from the different supernatant and pellet fractions was resuspended in 30 μl of SDS sample buffer, sonicated, and boiled for 15 min as a final step prior to separation by SDS-PAGE and immunoblotting as described below.

For preparation of the cytosolic membrane fractions used in the immunoprobe assays, cells (2×10^6 cells plated in 35-mm dishes) were homogenized in 400 μl of hypotonic homogenization buffer by passage through a 21-gauge needle 30 times. Homogenates were incubated for 18 h at 4 °C in 1× extraction buffer with or without 1% Triton X-100. The pellets were resuspended in 30 μl of SDS sample buffer and analyzed by SDS-PAGE and immunoblotting.

**Proteolytic Assay**—For assays of nurim cleavage with the site-specific protease endoprotease Lys-C (EndoLys-C), nuclear samples were incubated in 1× extraction buffer with or without 1% Triton X-100. Samples were incubated for 18 h at 5 °C with 0.25 mg/ml EndoLys-C (Roche Applied Science). The reactions were stopped by addition of SDS sample buffer supplemented with Complete® protease inhibitor and, for the analysis of EndoLys-C cleavage products, on 10–20% gradient gels (Inovitrogen). Following electrophoresis, proteins were transferred onto Immobilon-P membranes (Millipore Corp.), which were blocked by incubation in methanol according to the manufacturer’s protocol and then dried for 15 min at room temperature prior to incubation with primary antibody. For immunodetection, the membranes were incubated overnight with the following primary antibodies: anti-GFP polyclonal anti-
body, 1:10,000 (RDI); anti-lamin A/C monoclonal antibody, 1:500 (Novo-
ecstra Laboratories Ltd.); and anti-lamin B1 monoclonal antibody,
1:500. After washing, the membranes were incubated with horseradish
peroxidase-conjugated anti-mouse or anti-rabbit IgG secondary anti-
bodies (Bio-Rad), and proteins were detected by enhanced chemilumi-
nescence using the ECL West Pico reagent (Pierce). Membranes were
exposed to Fuji Super RX film for 1–45 min at room temperature.

Transmembrane Prediction Analysis—Prediction of transmembrane
topology within nurim was performed with several algorithms, includ-
ing hydropathy plot analysis (ProtScale program) based on the method
of Kyte and Doolittle (33), the TMHMM Version 2.0 program (34), and
the PredictProtein service PHDhtm to predict the positions and length
of individual transmembrane domains (available at ca.expasy.
org/tools/).

RESULTS

Nuclear membrane proteins that have been characterized to
date contain extended N- or C-terminal domains that are
thought to anchor the proteins to the INM by interactions with
various nuclear components, including the lamina and chroma-
tin-associated proteins (2). In this work, we wished to further
characterize nurim, a potentially different class of INM com-
ponent, about which there has been relatively little investiga-
tion since its identification (31). As seen from the protein se-
quence, nurim has almost no N-terminal residues prior to the
first hydrophobic domain (see below) and only a short C-termi-
nal domain of ~30 relatively hydrophilic amino acids. The
majority of nurim residues (>60%) are hydrophobic and are
clustered into five distinct domains (Fig. 1a), which were pre-
picted upon initial identification (31) to form five TMDs. The
predicted organization places the N terminus in the nucleo-
plasm and the C terminus on the opposite side of the INM
in the lumen of the nuclear envelope (Fig. 1b, middle). How-
ever, the actual membrane topology of nurim at the INM,
whether the N terminus is nucleoplasmic and whether the C
terminus is on the same or different sides of the INM, remains
unknown. Comparison of several different computerized struc-
ture predictions resulted in two additional possibilities for the
membrane topology of nurim. Both predictions maintained the
organization of the first four TMDs, but differed in the inter-
pretation of the last hydrophobic segment of nurim. In one
prediction (TMHMM program), the long fifth hydrophobic do-
main does not span the membrane (Fig. 1b, left), resulting in a
four-TMD model with a long C-terminal tail in the nucleo-
plasm. Alternatively, the ProtScale program indicated that the

![Fig. 1. Summary of nurim-GFP fusion constructs.](https://example.com/fig1)

(a) Summary of nurim-GFP fusion constructs. (a) Schematic illustration of nurim (total length of 262 residues). Shaded boxes indicate the five hydrophobic domains, potential TMDs. (b) Summary of the topological models from prediction programs as discussed under “Results,” indicating nurim as a polytopic membrane protein encompassing four, five, or six TMDs. (c) Schematic summary of GFP-nurim (construct 1), nurim-GFP (construct 2), and nurimΔC1-GFP (construct 3). (d) Expression of the nurim-GFP constructs in HeLa cells. Cells were transfected with 0.3 μg of the appropriate plasmids, and expression was analyzed by SDS-PAGE and Western blotting of total cell lysates with anti-GFP antibody. Lane 1, GFP-nurim; lane 2, nurim-GFP; lane 3, nurimΔC1-GFP; lane 4, GFP.
comparatively long fifth hydrophobic domain is actually a bipartite TMD, spanning the membrane twice with a hairpin turn in the middle. In this model, nurim would have a six-TMD topology, with the N and C termini on the same side of the membrane (Fig. 1b, right). Formally, in each of these models, nurim could appear in two different orientations, with the N terminus facing the luminal or nucleoplasmic side of the INM. In an attempt to further characterize nurim membrane topology and compartmentalization at the INM, we fused nurim to GFP and examined the localization by imaging and biochemical fractionation. (The examination of GFP fusion proteins appeared particularly suitable since nurim was first identified as an INM protein in a screening approach of a GFP fusion library.) We constructed three different nurim-GFP constructs (Fig. 1c). Two full-length nurim-GFP constructs (nurim-GFP, and GFP-nurim) contained GFP fused to the N- and C-terminal ends of nurim, respectively. A truncated variant of nurim was also made (nurimΔC1-GFP, residues 1–186), coding only the first four predicted TMDs, with the C-terminal end fused to GFP.

Expression and Localization of Nurim-GFP—HeLa cells were transfected with equal amounts of plasmid DNA encoding each of the nurim-GFP constructs or GFP alone, and expression levels were analyzed 24 h after transfection by SDS-PAGE and Western blotting using anti-GFP antibody (Fig. 1d). The fusion proteins migrated with a molecular mass (~53 kDa) consistent with expectation (29 kDa for nurim plus 28 kDa for GFP). GFP-nurim (Fig. 1d, lane 1) was expressed in lower amounts than nurim-GFP (lane 2) and migrated with a marginally faster electrophoretic mobility. The reason for the lower expression is difficult to attribute with certainty and could be due to general considerations of the orientation of the GFP in fusion proteins in a nonspecific manner. As anticipated, nurimΔC1-GFP migrated with a decreased molecular mass of ~47 kDa.

The localization of the nurim-GFP fusion proteins was first examined by confocal microscopy using living cells. HeLa cells were transfected as described above with each of the constructs, and the subcellular distribution was analyzed at different intervals up to 7 days after transfection (Fig. 2). Efficient expression of nurim-GFP was observed; and by 4 days post-transfection, nurim-GFP fluorescence was readily detectable in many cells, frequently in patches. In most cells, nurim-GFP was observed in a compact nuclear rim pattern, whereas in cells with relative overexpression, additional localization at the endoplasmic reticulum could be observed (Fig. 2a, panel 1). Localization at the nuclear rim and endoplasmic reticulum was confirmed in co-immunolocalization studies using calreticulin and lamin B1 as marker proteins (data not shown). Nurim-GFP appeared to be quite stable and remained detectable at least up to 7 days, by which time it appeared virtually exclusively in a nuclear rim pattern (Fig. 2a, panel 2). Similar results were obtained for the GFP-nurim construct. Deletion of the C-terminal region affected nurim localization in two ways. First, although the truncated variant nurimΔC1-GFP was observed within the nuclear envelope, it did not selectively accumulate there, mostly appearing distributed throughout the cytoplasm in a diffuse pattern (Fig. 2b, panel 3). Second, unlike the two intact constructs, expression of nurimΔC1-GFP was extinguished with time and was undetectable by 7 days (Fig. 2b, panel 4). Thus, although its expression levels were initially similar to those of its parental construct (Fig. 1d, lane 3; and Fig. 2b, panel 3; see also Fig. 5), the loss of nurimΔC1-GFP-expressing cells indicates either that the protein is much less stable or that its expression is not sustainable due to a toxic effect on the cells.

To further examine the association of nurim-GFP with the nuclear membrane and the effect of deletion of the C-terminal region, we analyzed localization in fixed cells with or without detergent extraction. When bound to stable structures in the nucleus, INM proteins have been shown to remain resistant to extraction by non-polar detergents (31, 35). Therefore, cells expressing the nurim-GFP constructs were extracted with 0.5% Triton X-100 (Fig. 3b) or treated with PBS as a control (Fig. 3a) prior to fixation with paraformaldehyde, and localization was examined by direct fluorescence of the GFP constructs. Typical fields are shown. For both full-length constructs (nurim-GFP
and GFP-nurim), a detergent-resistant population remained detectable at the nuclear envelope, whereas the cytoplasmic fraction was almost completely extracted (Fig. 3, compare a and b). In contrast, nurimAC1-GFP was lost from both the cytoplasmic membranes and nuclear envelope after extraction, although some punctate material that we interpreted as aggregation was sometimes observed. We note that aggregation of nurimAC1-GFP could sometimes also be observed in live cells expressing relatively high amounts of the protein (Fig. 3a, right panel, inset). The results indicate that nurim in the cytoplasmic membranes was extracted by detergent; but once localized in the nucleus, nurim became resistant to extraction in a manner involving the C-terminal region.

Resistance of Nuclear Nurim to Biochemical Extraction—Our results on the detergent resistance of nurim obtained by microscopy are consistent with the previous characterization of nurim (31). In the previous work, nurim was also found to be resistant to detergent extraction by Western blotting of the soluble and pellet fractions after Triton X-100 treatment. To further examine this tight retention of nurim in the nucleus by biochemical partitioning and whether it might depend on chromatin, isolated nuclei were alternatively subjected to a series of progressively more stringent extraction conditions with or without nuclease treatment (Fig. 4). Nuclei of nurim-GFP-expressing cells were isolated as described under “Experimental Procedures” using detergent-containing buffer (0.5% Triton X-100) in the presence or absence of DNase I (0.25 μg/μl), and the detergent-resistant nuclear fraction was then further extracted under increasingly stringent conditions by incubation in buffer containing 0.5% Triton X-100 supplemented with high ionic salt (0.25 M ammonium sulfate; Fig. 4, lane 1), chaotropic salt (2 and 4 M urea; lanes 2 and 3, respectively), or high concentrations of reducing agent (20 mM DTT; lane 4). These extracted nuclear samples were then centrifuged, and the resultant supernatants and pellets were analyzed by SDS-PAGE and immunoblotting. The fractionation of lamin A/C was examined in the same samples in parallel. The results demonstrate very tight binding of nurim to the nuclear fraction and indicate that it was resistant to nuclease treatment, including DNase I (Fig. 4, compare upper and middle panels), with or without RNase treatment (data not shown). Thus, nuclear nurim-GFP was completely resistant to extraction with Triton X-100 plus ammonium sulfate (Fig. 4, lane 1). Detectable amounts were solubilized in Triton X-100 plus 2 M urea, increasing at 4 M urea (Fig. 4, lanes 3 and 4); but even under these conditions, the majority of nurim-GFP was still resistant to extraction. Urea unfolds proteins by increasing the solubility of both polar and non-polar side chains (36); and in comparison, under this stringent extraction condition, endogenous lamin A/C was completely solubilized from the nuclear fraction (Fig. 4, lower panels, lane 3). We also tested the possibility of tight anchoring of nurim due to intermolecular disulfide bonds since nurim contains several cysteine residues that might form disulfide linkages. As shown in Fig. 4 (lane 4), incubation in detergent containing 20 mM DTT had no effect on nurim extraction, with the vast majority of the protein remaining resistant to extraction. Overall, the tight binding of nurim appears to originate mainly from strong hydrophobic interactions, but does not seem to rely on the lamina or chromatin. The resistance of nurim to extraction in the face of virtually complete lamin extraction is surprising. In any such study, it is difficult to exclude the possibility that, during the extraction procedure itself, nurim is converted into an insoluble form precisely because of the removal of interacting components. Nevertheless, these are the results of biochemical extraction procedures, and they are consistent with results from the independent approach of FRAP analysis cited above, which indicate that nurim is an extremely tightly bound, immobile nuclear component.

Examination of the Membrane Topology of Nurim—Sequence analysis of nurim clearly indicated that it is a polytopic membrane protein; but as is frequently the case, there were different possibilities for the orientation of nurim in the membrane and, in particular, whether the N and C termini are on the same or opposite sides of the membrane. In the five-TMD model, the N and C termini are on opposite sides of the membrane, whereas in the four- and six-TMD models, both ends would be on the same side of the membrane (Fig. 1b). We therefore next wished to address the topology and orientation of nurim using an immunoprobeing assay. Although nurim was clearly tightly bound to the INM once in the nucleus, in transiently expressing cells, nurim-GFP could be detected in the cytoplasm, and this was in a form that could be extracted. We anticipated that since membrane insertion proteins originate in the cytoplasm, this fraction could be used to examine the orientation of nurim in the membranes. Vesiculated cytoplasmic membrane fractions containing nurim-GFP and GFP-nurim were therefore isolated in the absence of detergent and probed for the ability to bind anti-GFP antibody. Bound antibody was then isolated by pelleting of the membrane fraction and measured by Western blotting against the heavy chain. Luminally disposed GFP should not be available for antibody binding and should not coprecipitate with the membrane fraction. To help determine the orientation within the membrane
Characterization of Nurim, an Inner Nuclear Membrane Protein

Protein profile of the membrane fractions is shown in detergent by homogenization through a narrow gauge needle. The total harvested, and membrane fractions were prepared in the absence of three nurim GFP constructs (lanes 2–4), but was not bound by control lanes 5–6. The anti-GFP antibody (heavy fragment, 50 kDa) was bound by membrane fractions containing any of the GFP fusion proteins in lanes 1–6. Each of the membrane fractions, including a control fraction lacking any GFP fusion protein (lane 1; M, mock transfected), was then incubated with anti-GFP antibody, and the membrane material was pelleted, washed, and analyzed for the presence of bound anti-GFP antibody, detected by a horseradish peroxidase-coupled secondary anti-rabbit antibody (b). The anti-GFP antibody (heavy fragment, 50 kDa) was bound by membrane fractions containing any of the three nurim GFP constructs (lanes 2–4), but was not bound by control samples or by samples containing the emerin-GFP and LBR.N1-GFP proteins (lanes 1, 5, and 6). We conclude that the N- and C-terminal ends of nurim are localized to the cytoplasmic side of the membrane.

Fraction, two additional INM proteins with known topology were included in the assay. Emerin-GFP (full-length) and an LBR-GFP variant (LBR.N1-GFP) contain an N-terminal domain and a single TMD. Both of these have been shown to be type II transmembrane proteins, with the GFP portion at their C-terminal ends disposed within the luminal side of the membrane. (11, 14, 37, 38). Homogenates were made from cells expressing nurim-GFP, GFP-nurim, nurimΔC1-GFP, emerin-GFP, or LBR.N1-GFP. To control for the total amount of the various GFP fusion proteins, samples from each of the membrane fractions were analyzed directly by SDS-PAGE and Western blotting. The results demonstrate approximately equal amounts of the fusion proteins in the starting material, with the exception of GFP-nurim, which was present in ~3-4-fold lower amounts (Fig. 5a). The membrane fractions were incubated in buffer with or without anti-GFP antibody, pelleted by centrifugation, washed to remove unbound antibody, and pelleted again. The pelleted fractions were then analyzed for the association of bound anti-GFP antibody by probing with a secondary antibody against the anti-GFP antibody (Fig. 5b). In control cells lacking any GFP fusion protein (Fig. 5b, lane 1), no anti-GFP antibody was detected, indicating that any anti-GFP antibody coprecipitation was specific and required the presence of a target GFP fusion protein. Bound anti-GFP antibody was detected in the samples for both nurim-GFP (Fig. 5b, lane 4) and GFP-nurim (lane 2). In contrast, although emerin-GFP and LBR.N1-GFP were present in similar amounts in the respective homogenates (Fig. 5a), only minor amounts of anti-GFP antibody were present in the corresponding pelleted membrane fraction (lanes 5 and 6). The results indicate, consistent with predictions, that the GFP portion of these latter two fusion proteins is disposed toward the luminal side in the membrane fraction and is not accessible to the anti-GFP antibody. Conversely, the results indicate that both the N- and C-terminal ends of nurim are positioned at the cytoplasmic side and, by inference, the nucleoplasmic side of the membrane. Anti-GFP antibody was also bound by membranes expressing the nurimΔC1-GFP variant (Fig. 5a, lane 3), indicating that it also has the GFP moiety at its C-terminal end cytoplasmically disposed. Together with the results of the prediction algorithms, these data provide strong evidence that nurimΔC1-GFP contains four TMDs. Conversely, for the intact protein, the results are not consistent with a five-TMD model. Formally, the results could be explained by nurim being limited to the four TMDs of the N-terminal region, i.e. within nurimΔC1-GFP. However, we favor the model (consistent with the results from the anti-GFP binding assays described above) in which nurim possesses six TMDs, with the last hydrophobic segment being present in a hairpin orientation and spanning the membrane twice. This proposal for nurim is supported by the additional considerations below.

Analysis of Nurim Organization by Site-specific Proteolysis—

The structure of nurim was further analyzed by proteolysis using EndoLys-C, which cleaves specifically after lysine residues. As illustrated in Fig. 6a, nurim contains five lysine residues, four of which are positioned in the interlinking loop regions between the predicted TMD (residues 86, 121, 168, and 184) and one close to the C-terminal end (residue 249). We therefore subjected extracts containing nurim-GFP, nurimΔC1-GFP, or nurim-GFP to digestion with EndoLys-C, separated the products by SDS-PAGE, and detected any cleaved products using anti-GFP antibody. Knowing the position of GFP at the N- or C-terminal end allowed us to approximate the position of any cleavage site. In establishing the parameters for this assay, we also found that native GFP itself is relatively resistant to EndoLys-C cleavage and was not cleaved to smaller products in parallel assays (data not shown).

Incubation of nurim-GFP extracts with EndoLys-C resulted in a single cleavage product migrating with a molecular mass of 35 kDa (Fig. 6b, lane 2). Based on detection by virtue of the GFP moiety and the orientation of GFP, cleavage at position 249 would not result in a large enough product, whereas cleavage at position 121 or farther N-terminal would result in a significantly larger product. Therefore, the only consistent position for this single cleavage is at position 168 or 184. (Although it is formally difficult at the resolution of this method to discriminate between these sites, we favor the main site being position 184 (see below.) In either case, the main site would be within the predicted loop between TMD4 and TMD5. The cleavage of the nurimΔC1-GFP construct yielded results consistent with this interpretation. The nurimΔC1-GFP variant contains the four N-terminal lysines, with the one at position 184 situated closely to the GFP moiety itself (fusion at position 186). Upon EndoLys-C cleavage, a single major product was again observed. This was smaller than that produced from intact nurim-GFP and migrated with a size almost identical to but marginally greater than that of native GFP (Fig. 6b, lane 5, and data not shown). This is most consistent with cleavage at position 184. Finally, when examining cleavage of GFP-nurim, we again obtained a single major product migrating now at ~40 kDa. This product is too large to represent cleavage at the more N-terminal lysine residues and is most consistent again with
cleavage at position 184. Together, these results indicate that the main or only site in nurim sensitive to EndoLys-C cleavage is in the nucleoplasmically disposed loop between TMD4 and TMD5. However, in the absence of detergent, the intact constructs nurim-GFP and GFP-nurim remained partially resistant to EndoLys-C cleavage. In contrast, nurim/H9004C1-GFP was almost completely cleaved in parallel assays. This increased sensitivity correlates with the observed lack of tight integration of nurim/H9004C1-GFP into the INM as observed above. We note also that addition of detergent resulted in virtually complete EndoLys-C cleavage of the intact nurim-GFP constructs while maintaining the qualitative pattern, resulting in a single cleavage product in each case. Although there could be several explanations for this observation, one possibility is that the presence of detergent increased EndoLys-C accessibility to the loop structure by removing an interacting component or otherwise altering the conformation of nurim around the loop (see “Discussion”). We further tested whether EndoLys-C cleavage of nurim was in anyway enhanced by DNase I treatment (i.e. possibly reflecting an interaction with chromatin), but saw no alteration in the relative sensitivity (data not shown). Although this does not rule out nurim-chromatin interaction, it is not one detected by relative sensitivity to EndoLys-C.

Sequence Analysis and Conservation of the C-terminal Region of Nurim—Nurim was originally identified in a screen for likely INM proteins in human cells (31). In a database search, we found nurim homologs in other mammalian species and other organisms. In mammals, nurim is quite highly conserved in, for example, the mouse (Mus musculus) nurim ortholog, sharing 94% sequence identity with human nurim. Orthologs in other phyla were found, including the fish Takifugu rubripes and the insect genomes of Drosophila melanogaster and Anoph-
eles gambiae. The D. melanogaster sequence was the most distant nurim ortholog found in eukaryotes, but was clearly related, coding for a 253-amino acid protein sharing 27% sequence identity and 46% conservation with human nurim (Fig. 7a). We found no evidence for the presence of a nurim ortholog in the completed genomes of various lower metazoan organisms, e.g. the nematode Caenorhabditis elegans and unicellular eukaryotic organisms such as Saccharomyces cerevisiae (but see below). Consideration of the alignment of the nurim ortholog proteins (Fig. 7a) revealed several features relevant to organization and topology. Thus, nurim proteins are each organized in five quite highly conserved domains localized in the area of the five hydrophobic domains of nurim, and each of the sequences was predicted to be a TMD. For illustration, the predicted TMDs are underlined in the alignment. In addition to the TMD organization and sequence conservation, a domain of significant sequence conservation is present in the largest interlinking loop, between TMD4 and TMD5. The amino acid sequence
\[\text{ELMGLKQVYXXGXPX}_5\text{RX}_{25}\text{RHP} \ldots\] is almost completely conserved in all nurim orthologs and spans from the end of TMD4 to the entry to TMD5, which is itself also well conserved. According to our model, this conserved amino acid sequence will be present in the large loop exposed to the nucleoplasmic face of the INM. We note that the loops between TMD1 and TMD2 and between TMD3 and TMD4, which, according to our results, are localized on the luminal side of the membrane, are the regions with the least homology. On the other hand, the loop between TMD2 and TMD3 is predicted to be on the nucleoplasmic side, and it encompasses significantly conserved residues.

Considering the absence of a nurim orthologs in C. elegans, S. cerevisiae, and Schizosaccharomyces pombe as indicated above, we were surprised to find a very clear ortholog of nurim in the prokaryote Mycobacterium tuberculosis. The mycobacterial protein Rv3238 (strain H37Rv) shares 22% identity and 44% conservation with human nurim (Fig. 7a). Moreover, the putative mycobacterial protein conforms precisely in secondary structure prediction to the nurim model (see below). Interestingly, the region linking TMD4 and TMD5 is also well conserved, confirming to the consensus found in the mammalian species. Clear homologs to Rv3238 in M. tuberculosis were also found in other species of mycobacteria.

Homology of Nurim to ICMTs—In further computer analysis of the regions of conservation between TMD4 and TMD5 in

---

**FIG. 7. Analysis of the sequences and organization of nurim orthologs.**

**a**, protein sequence alignment. The follow sequences are aligned: Homo sapiens nurim (N_Hum; accession number NP_009174.1, gi:2528391), an incomplete sequence of a nurim ortholog protein identified in T. rubripes (N_Fugu; accession number CAA80104153.1, gi:22419496, missing N-terminal sequence is indicated by dashes), a nurim ortholog in D. melanogaster (N_Dr; accession number AAL28322.1, gi:16765806), the predicted protein Rv3238c in M. tuberculosis strain H37Rv (accession number NP_217755.1, gi:15610374), and the ICMT Ste14p in S. cerevisiae (accession number NP_010698.1, gi:6320618). Accession numbers are from the NBI ReSeq data base. Black boxes indicate identical amino acids in all sequences; light- and dark-gray boxes indicated conserved amino acids at the same position in three or four sequences. Dashes indicate short gaps that give better overall similarity. Underline denotes putative transmembrane regions of the proteins. Sequences were aligned using ClustalX software with parameter settings as follows: gap extension penalty, 0.2; gap opening penalty, 10; and protein weight matrix, Gonnet series. The alignment was then manually modified using the data from the membrane prediction program PHDhtm. **b**, hydrophobicity plots of nurim (H. sapiens), Rv3238 (M. tuberculosis), and Ste14p (S. cerevisiae) generated according to the algorithm of Kyte and Doolittle (33). The predicted secondary structure is indicated by the horizontal black line through the hydrophobicity plot, with gray boxes indicating the potential TMDs. To emphasize the similarity between proteins, the hydrophobic domains are highlighted by vertical gray shading across all plots. The plots emphasize the overall greater similarity between nurim and Rv3238 (compare upper and middle panels) than between nurim and Ste14p (compare upper and lower panels).
Characterization of Nurim, an Inner Nuclear Membrane Protein

Nurim, we found that a similar motif (termed the RHP motif) has been identified previously as part of a consensus sequence within a group of enzymes known as ICMTs (39). Although primary sequence similarity across the length of nurim was not initially found for ICMT, the conservation of the RHP motif prompted us to evaluate a potential nurim-ICMT relationship. Yeast ICMT (Ste14p) is a polytopic transmembrane protein and has been shown to exhibit a six-TMD organization. As for nurim, Ste14p contains a loop region between TMD4 and TMD5, and the C-terminal TMD is present as a hairpin turn, placing the N and C termini on the same side of the membrane. The C-terminal consensus sequence characteristic for ICMT proteins also contains additional defining features, including a conserved EE or ED doublet after the final TMD region (39). In nurim, a similar acidic motif (DXXD) is present in an equivalent position. Finally, to promote a hairpin in the membrane, turn-inducing amino acid residues, located in the middle of the long hydrophobic domain, are also required (40). A pair of strong helix-disrupting amino acids (Asn-Pro in Ste14p and Asp-Arg in nurim) are located precisely within the middle of their respective hydrophobic domains (Fig. 7a, inverted black triangle), consistent with the organization into TMD5 and TMD6 (Fig. 7a). The characteristics of the TMDs and the bipartite organization of TMD5 and TMD6 are readily illustrated by the hydropathy plots shown in Fig. 7b. We note for the primary sequence a somewhat greater similarity between nurim and Rv3238 than between nurim and Ste14p. Whether or not nurim acts as an ICMT (see “Discussion”), these considerations altogether add strong weight to the proposal that nurim is a six-TMD protein and encompasses a helical hairpin structure at the C-terminal end.

DISCUSSION

In this study, we have investigated the membrane topology and characteristics of nurim, a protein first identified in a screen for INM proteins. Nurim was predicted to encompass five TMDs, placing the N and C termini on opposite sides of the membrane (31); but as is generically the case, alternative models are frequently possible.

Our biochemical data, together with further sequence consideration, provide strong evidence that nurim is most likely to be present as a six-TMD protein, with nucleoplasmic loops between TMD2 and TMD3 and between TMD4 and TMD5. Our model differs from the original mainly in that the last TMD is a long bipartite transmembrane domain containing a hairpin turn and, importantly, places the C terminus on the same side of the membrane as the N terminus. In an immunological probing assay of nurim-GFP fusion constructs, both ends were detected on the cytoplasmic/nucleoplasmic side of the membrane. This indicates a topology with an even number of TMDs. Further consideration of the results of the nurim deletion, the presence of a C-terminal hydrophobic domain including a pair of charged amino acid residues with turn propensities, and the conservation with Ste14p, whose hairpin organization in this region has been demonstrated (39), all point to the six-TMD model.

Binding of Nurim to the INM—Very tight binding of nurim to the INM was noted in biochemical extraction experiments and FRAP analysis in the original isolation (31). We also observed that nurim remained largely resistant to extraction. Even under conditions in which nuclear lamins were completely extracted, nurim was only partially solubilized from the nucleus. Cytoplasmically located nurim was readily extracted in detergent, indicating that localization at the INM (or modification within the nucleus) resulted in its extreme resistance to solubilization. We found no difference in the resistance of nurim to biochemical extraction after DNase and RNase treatments. Nurim binding is also unusual since, in those INM proteins analyzed to date, anchoring has been shown to involve large nucleoplasmic domains at their N-terminal ends (5, 7) that contain the various binding sites for interactions with the nuclear lamina and chromatin. Nurim therefore appears to be retained in the INM by an unusual mechanism, perhaps involving the assembly of some sort of scaffold within the membrane itself, by binding either to additional INM proteins or to itself. Another possibility is that nurim binds another unidentified structural component of the nucleus.

By proteolysis with the site-specific protease EndoLys-C, we have also shown that a lysine residue within the loop region between TMD4 and TMD5 is the single major exposed site available for cleavage and that detergent extraction increases accessibility. Deletion of the C-terminal region abolished the tight binding of nurim to the INM, and it is therefore tempting to speculate that this loop region contains a major determinant of INM recruitment for nurim. Although this is a reasonable conclusion, it is also tempered by the observations of Rolls et al. (31), who found that deletions in several positions throughout nurim affect INM recruitment (as defined by detergent resistance) and that no one short determinant appears uniquely critical. It may therefore be that it is the complete nurim structure, perhaps dictating intramolecular interactions and/or intermolecular multimerization, that is involved in orchestrating the structure required for stable integration. Nevertheless, the loop region between TMD4 and TMD5 contains certain interesting features, as discussed below.

Homology of Nurim to the ICMT Enzyme Family—Although this work concerns biochemical and compartmentalization studies, we found an interesting similarity between nurim and the enzyme family of ICMTs that warrants discussion. ICMTs are cytoplasmic polytopic enzymes involved in the processing of proteins containing a CAAX (where A is an aliphatic amino acid) motif at their C termini, which include, for example, the Ras proteins and, interestingly, the nuclear lamins (44, 45). The CAAX motif is a target for a series of modifications, including isoprenylation and methylation of the cysteine, which together allow the modified proteins to associate with membranes. ICMTs from different species (one known ICMT in humans) contain a conserved tripartite consensus sequence ([... RHPXHydrophobic amino acids]EE [...]) wherein the hydrophobic region forms a hairpin turn within the membrane as discussed above (39). This conserved tripartite motif is proposed to form the S-adenosylmethionine-binding motif (46). Intriguingly, this feature is also present in the C-terminal region of nurim proteins. Nurim also exhibits a similar size and membrane topology to those of Ste14p, the ICMT in S. cerevisiae. It is therefore plausible that nurim may be an ICMT. Although the main ICMT activity is present in the cytoplasm, nurim might function as a specific nuclear ICMT. The existence of a nuclear CAAX-modifying machinery has been suggested previously (47). It is therefore intriguing to speculate that nurim could be specifically involved in nuclear isoprenylcysteine methylation, perhaps of the lamins themselves (48, 49).

On the other hand, nurim appears to be a distinct protein class and actually more similar to the mycobacterial protein Rv3238 than to ICMTs. Furthermore, since the tripartite consensus motif is also present in the C-terminal region of the LBR and its structural homolog, the sterol Δ14-reductase ER24 (41), neither of which have been reported to have methyltransferase- or S-adenosylmethionine-binding activities, it may be that the conservation of this tripartite motif underpins conservation of structural organization within the protein-membrane interface rather than the enzymic function. With regard to INM localization, we plan to alter the human ICMT sequence to that of
Characterization of Nurim, an Inner Nuclear Membrane Protein

nurim (particularly within the TMD4–TMD5 region) and examine whether it can be converted from a cytoplasmic form to one tightly bound in the nucleus and vice versa.

Finally, as indicated, the greatest homology to eukaryotic nurim proteins is actually found in the prokaryote M. tuberculosis. The mycobacterial protein Rv3238 is conserved in sequence and organization and, in particular, in the loop region between TMD4 and TMD5, containing the additional defining features of the nurim orthologs that are lacking in the ICMT family. The function of this mycobacterial protein is unknown. If this protein is found to play a role in the pathogenicity of the bacteria in the host cell, the homology to nurim may be an important feature relevant to its activity.

Acknowledgments—We thank Tom Rapoport for the generous provision of plasmids and Christine Hrycyna and Susan Michaelis for helpful advice during the course of this work.

REFERENCES

1. Burke, B., and Stewart, C. L. (2002) Nat. Rev. Mol. Cell. Biol. 3, 575–585
2. Holmer, L., and Worman, H. J. (2001) Cell. Mol. Life Sci. 58, 1741–1747
3. Fahrenkrog, B., Stoffler, D., and Aebi, U. (2001) Curr. Top. Microbiol. Immunol. 259, 85–117
4. Marte, B. (2001)
5. Georgatos, S. D. (2001)
6. Furukawa, K., Glass, C., and Kondo, T. (1997) Biochem. Biophys. Res. Commun. 238, 240–246
7. Worman, H. J., and Courvalin, J. C. (2000) J. Membr. Biol. 177, 1–11
8. Weber, K. (1986) Nature 320, 402
9. Steurman, N., Heins, S., and Aebi, U. (1998) J. Struct. Biol. 122, 42–66
10. Goldman, R. D., Gruenbaum, Y., Moir, R. D., Shumaker, D. K., and Spann, T. P. (2002) Genes Dev. 16, 533–547
11. Ellenberg, J., Sigga, E. D., Moreira, J. E., Smith, C. L., Presley, J. F., Worman, H. J., and Lippincott-Schwartz, J. (1997) J. Cell Biol. 136, 1183–1206
12. Faisner, R. (2001) J. Cell Sci. 114, 3791–3792
13. Schirmer, E. C., Flores, L., Guan, T., Yates, J. R., III, and Gerace, L. (2003) Science 301, 1380–1382
14. Worman, H. J., Evans, C. D., and Blobel, G. (1990) J. Cell Biol. 111, 1535–1542
15. Worman, H. J., Yuan, J., Blobel, G., and Georgatos, S. D. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 8531–8534
16. Schulze, K., Lin, F., and Worman, H. J. (1994) J. Biol. Chem. 269, 11312–11317
17. Soullam, B., and Worman, H. J. (1995) J. Cell Biol. 130, 15–27
18. Ye, Q., and Worman, H. J. (1994) J. Biol. Chem. 269, 11306–11311
19. Ye, Q., and Worman, H. (1996) J. Biol. Chem. 271, 14653–14656
20. Simos, G., and Georgatos, S. D. (1992) EMBO J. 11, 4027–4036
21. Senior, A., and Gerace, L. (1988) J. Cell Biol. 107, 2029–2036
22. Martin, L., Crimando, C., and Gerace, L. (1995) J. Cell Biol. 126, 8822–8828
23. Harris, C. A., Andryuk, P. J., Cline, S. W., Matthew, S., Siekierka, J. J., and Goldstein, G. (1995) Genomics 28, 198–205
24. Biene, S., Macerini, E., Rivella, S., Mancini, M., Regis, S., Romeo, G., and Tonio, D. (1994) Nat. Genet. 8, 323–327
25. Lin, F., Blake, D. L., Callebaut, I., Skerjanc, I. S., Holmer, L., McBurney, M. W., Paulin-Levasseur, M., and Worman, H. J. (2000) J. Biol. Chem. 275, 4840–4847
26. Furukawa, K., and Kondo, T. (1998) Eur. J. Biochem. 251, 729–733
27. Fairley, A. E., Kendrick-Jones, J., and Ellis, J. A. (1999) J. Cell Sci. 112, 2571–2582
28. Clements, L., Manilal, S., Leve, D. R., and Morris, G. E. (2000) Biochem. Biophys. Res. Commun. 267, 709–714
29. Lee, K. K., Haraguchi, T., Lee, R. S., Koujin, T., Hirooka, Y., and Wilson, K. L. (2001) J. Cell Sci. 114, 4567–4573
30. Furukawa, K. (1999) J. Cell Sci. 112, 2485–2492
31. Rolls, M., Stein, P. A., Taylor, S. S., Ha, E., McKewen, F., and Rapoport, T. A. (1999) J. Cell Biol. 146, 29–44
32. Greaves, R., and O’Hare, P. (1989) J. Virol. 63, 1641–1650
33. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132
34. Krogh, A., Larsson, B., von Heijne, G., and Sonnhhammer, E. L. (2001) J. Mol. Biol. 305, 547–550
35. Gant, T. M., and Wilson, K. L. (1997) Proc. Natl. Acad. Sci. U. S. A. 11312–11317
Analysis of the Localization and Topology of Nurim, a Polytopic Protein Tightly Associated with the Inner Nuclear Membrane

Helmut Hofemeister and Peter O'Hare

J. Biol. Chem. 2005, 280:2512-2521.
doi: 10.1074/jbc.M410504200 originally published online November 12, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M410504200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 49 references, 23 of which can be accessed free at http://www.jbc.org/content/280/4/2512.full.html#ref-list-1