Chapter 3

On the role of the intrinsically disordered linker of Heh2 in nuclear transport: can a linker of 149 prolines support import of membrane proteins through the nuclear pore complex?

Petra Popken\textsuperscript{1,2,3}, Irina L. Rempel\textsuperscript{1}, Ali Ghavami\textsuperscript{2}, Anouk H.G. Wolters\textsuperscript{4}, Anne C. Meinema\textsuperscript{3,5}, Bert Poolman\textsuperscript{2,3}, Ben N.G. Giepmans\textsuperscript{4}, Patrick R. Onck\textsuperscript{2}, Liesbeth M. Veenhoff\textsuperscript{1}

\textsuperscript{1} European Institute for the Biology of Ageing (ERIBA), University of Groningen, University Medical Center Groningen, Groningen, Netherlands. \textsuperscript{2} Zernike Institute for Advanced Materials, University of Groningen, Groningen, The Netherlands. \textsuperscript{3} Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands. \textsuperscript{4} Dept. Cell Biology, University of Groningen, University Medical Center Groningen, Groningen, Netherlands. \textsuperscript{5} Department of Biology, Institute of Biochemistry, ETH Zürich, Zürich, Switzerland.

Submitted to: Journal of Biological Chemistry
Abstract
The nuclear pore complex (NPC) embedded in the nuclear envelope (NE) forms the sole gateway to the nuclear interior and the inner nuclear membrane (INM). The import of *Saccharomyces cerevisiae* INM proteins Src1/Heh1 and Heh2 is mediated by transport factors and depends on FG-repeat containing nucleoporins and the Ran-gradient (88, 120). The sorting signal consists of a nuclear localization signal (NLS) separated from the transmembrane domain by an intrinsically disordered (ID) linker of approximately 200 amino acids. What the exact role of the linker is and which properties of the linker are essential is still unclear. We altered the Stokes radius and flexibility of the linker by systematically increasing the proline content of the linker region up to 149 consecutive prolines and measured the effect on import. Combining in vivo fluorescence microscopy, immunoEM and a computational model predicting biophysical properties of the linkers, we find a striking correlation between Stokes radius and import efficiency consistent with a role for the linker as a spacer. In addition, we find that reduced flexibility hampers import, showing linker flexibility is required to enter and cross the NPC. Furthermore, efflux is severely hindered by a less flexible linker region. Thus, the flexible and extended linker plays a crucial role in membrane protein transport through the NPC.

Introduction
The nuclear pore complex (NPC) embedded in the nuclear envelope (NE) forms the sole gateway to the nucleus. This is best understood for soluble proteins (5) and less so for integral membrane proteins of the nuclear envelope. Several mechanisms have been proposed for the sorting of membrane proteins to the inner membrane of the nuclear envelope (10, 20, 82, 99, 202). The biogenesis and sorting of membrane proteins starts in the cytosol with cotranslational or posttranslational membrane insertion into the endoplasmic reticulum (ER) membrane. The membrane proteins then potentially roam the entire NE-ER network, as the membranes of the peripheral ER and the outer and inner nuclear membrane (ONM and INM) of the NE form a continuous system. To reach the INM, membrane proteins pass the NPCs where the INM and ONM are fused. The route through the *Saccharomyces cerevisiae* NPC is spacious enough for passage of membrane proteins with extralumenal domain sizes of 90 kDa and thus the majority of monomeric proteins may enter the INM through the NPC even in the absence of specific sorting signals (144).
Many proteins indeed reach the INM by diffusion through the NPC and their accumulation at the INM is through retention mechanisms (67). However, in S. cerevisiae there are two membrane proteins, Src1/Heh1 and Heh2, whose localization in the nucleus is depending on an active transport mechanism involving transport factors, FG-nups and the RanGDP/RanGTP gradient (88, 120). The sorting signal in these proteins consists of a very potent nuclear localization signal (NLS) (88, 105, 118, 120), and additionally a long, intrinsically disordered (ID) linker (120). Already in the ER, the NLS ensures recruitment of the transport factor Kap95 via the adapter protein Kap60 (118). The model proposed is that the complex then shuttles through the NPC by binding to the FG-nups in the NPC; a transport mode for which energy input is provided from the gradient Ran proteins across the NE (88, 120). The function of the long linkers separating the NLS from the transmembrane domains is less well understood.

The ID linkers in Heh1 and Heh2 are approximately 200 amino acids long (120). Shortening the linker decreases the sorting of the proteins to the INM, while randomization of the amino acid sequence, which keeps the linkers ID nature, maintained full functionality (120). Thus, the length but not the exact sequence of the linker is important for efficient import to the INM. We proposed that the linker is needed to effectively recruit Kap60 and Kap95 away from the crowded environment of the membrane, maybe as early as during protein biogenesis (118-120). For this, the linker is effectively a spacer and we expect that also less flexible structures may work. In addition, or alternatively, the ID linker may help to reach out to bind to the FG-nups inside the NPC (118-120), which involves dodging into the NPC for which flexibility may be required.

To address if flexibility of the ID linker is required, we introduced polyproline stretches to reduce the disorder. Peptides with multiple consecutive prolines form polyproline helices, of which the left-handed polyproline II helix (PPII) with residues in the trans conformation is energetically favourable (126). In disordered proteins residues with PPII conformations are common (2, 201), but here we use stretches of prolines that form longer, less flexible PPII helices. A coarse-grained computational model (57) was used to gain insight in the structure and dynamics of linkers with polyproline stretches. We relate these structure predictions to in vivo measurements of the import to the INM of reporters containing linkers with different proline content, and
discuss these findings in the light of the potential transport path through the NPC. We conclude that both linker flexibility and radius are important for nuclear transport of membrane proteins.

**Results**

*Rationale*

We set out to answer how the biophysical properties of the linker regions, in particular its flexibility, impact active nuclear import of membrane proteins, and designed reporters with linkers with increasing proline content. By replacing increasing numbers of amino acids in the linker region with stretches of proline residues, the charge content is reduced, causing a more collapsed structure because of reduced repulsive energy, and flexible ID regions are exchanged by more structured helical regions. We determine the effect on import by monitoring the subcellular localization of a GFP-fusion of a transmembrane (TM) domain (10 transmembrane spanning helices derived from Sec61) with a sorting signal composed of the linker region (L) and a nuclear localization signal (NLS). The advantage of using these reporter proteins with different linkers above linker mutants of native Heh1 and Heh2 is that the native proteins are retained at the INM through interactions of their nuclear domains with other content of the nucleus, potentially masking effects of the different linkers on the import rates (118, 199).

*Expression and NE localization of membrane proteins with polyproline linkers*

Polyproline linkers such as those designed in this study are unnatural and hence a first question was if those could be expressed. We designed polyproline linkers with increasing proline content by replacing ten regions of an ID linker sequence (L) with regularly spaced stretches of 5 prolines (LP50) and increased the proline content even further by extending the polyproline stretches (LP70, LP100) until an entire 149 residue long region was replaced by prolines (LP149). This resulted in reporter proteins with domain structure GFP-h2NLS-linker-Sec61 (G-NLS-L-S), where the linker contained 57, 75, 100 and 149 prolines, the last having a continuous stretch of 149 proline residues (Figure 3.1A, for sequences see Table 3.2). All reporters expressed, although at reduced levels with increasing proline content (Figure 3.1B).

Next we asked, if the variable linker sequences interfered with subcellular
Flexible linker needed for inner nuclear membrane targeting

Indeed all variants accumulated at the NE and little protein resided at the peripheral ER. However, the efficiency of NE targeting, which was quantified by the ratio of fluorescence intensity in the NE over intensity in the peripheral ER, was affected by the amount of prolines in the linker. This ratio is an indirect readout of the import (and efflux) rate when assuming that the steady state localization of these reporters is not determined by retention but by diffusive movements within the NE/ER network, including the transport to the inner membrane. At steady state, the NE/ER ratio is largely independent of expression level (120) enabling comparison between the different strains. The NE/ER ratio decreased with increasing proline content, with an exception for the reporter with the all proline linker LP149, which showed an intermediate accumulation (Figure 3.1C, NE/ER = 24.4 ± 3.6 compared to 42.1 ± 6.4 for

Figure 3.1 Reporters with polyproline linkers express and accumulate at INM. A. Graphical representation of reporters and polyproline linkers used. Red bars are proline residues, blue bars are charged residues, others are black. B. In gel fluorescence of lysates from cells expressing indicated reporter proteins, showing full length expression (indicated by the arrowheads). C. Deconvolved images of cells expressing reporter proteins with indicated linker and quantification of NE/ER ratio. Scale bar 5 µm. Average ratios of the fluorescence at the NE over that at the peripheral ER (NE/ER) and SEM over n=30 cells is listed.
the native Heh2-linker L). In the cells expressing the reporter with the LP100 linker a small soluble fraction can be observed in the nucleus. This may be a degradation product of the reporter, containing at least the NLS and GFP, but it does not interfere with the analysis as the ratio of fluorescence intensity in different membrane compartments is used.

Thus, having many proline residues in the linker, which increases the amount of structure between the NLS and the TM-domain, does not prevent accumulation of the reporter proteins at the NE, while it does affect the efficiency of transport as judged by the NE/ER ratio.

**Membrane proteins with polyproline linkers are targeted to the INM**

Having established that the reporter proteins are expressed and localize to the NE, we addressed whether the proteins are actually localized at the INM and not at the ONM. A first indication was obtained using the inducible Kap95-anchor away system. When reporters were expressed in the presence of rapamycin, Kap95-FRB traps at Pma1-FKBP at the plasma membrane, hereby disabling active import of Kap60/Kap95-dependent cargo through the NPC (64, 120). All reporters expressed under these conditions mislocalized to the peripheral ER resulting in NE/ER ratios around 1.5 (Figure 3.2A), showing that the accumulation at the NE is Kap95-dependent. This indicates that this localization is a result of nuclear import and that the proteins are localized in and targeted to the INM.

A second indication that the proteins are INM localized comes from assessing the viability of cells expressing the different INM proteins (Figure 3.2B). Previously we have seen that overexpression of reporter proteins that accumulate at the INM correlates with a reduction in cell viability, while reporters without the NLS or linker do not cause cell death (98). The reporters with polyproline linkers also show lethality upon expression, hinting at their residence at the INM (Figure 3.2B). Interestingly, there is a correlation between the viability of the strains and the NE/ER ratios, where the proteins with highest accumulation at the NE show the lowest viability (compare e.g. LP50 and LP100 in Figure 3.1C and Figure 3.2B).

Lastly, for a subset of the reporters we used quantitative immuno electron microscopy (iEM) to determine their localization in relation to the INM (Figure 3.2C). Primary anti-GFP antibodies where visualized with colloidal gold. To allow quantitative and unbiased analysis, hundreds of yeast cells where imaged using newly-developed semi-automated large-scale EM at high
Flexible linker needed for inner nuclear membrane targeting

Figure 3.2 Reporters with polyproline linkers accumulate at INM. A. Deconvolved images of cells expressing reporter proteins with indicated linker in the presence of rapamycin, which block active import. Scale bar 5 µm. Average NE/ER ratio is listed, n=30, with SEM indicated. B. Spot assay with cells expressing reporter proteins with indicated linker. Top panel: control plate, bottom panel: plate containing 0.5% Gal which induces expression. C. Two selected examples showing reporter proteins as indicated with mainly INM or mainly ONM localization. Nucleus is indicated. Arrowheads show assignment of gold particles. Green rightward: nuclear; red leftward: cytoplasmic; white upward: ambiguous; blue downward: not assigned (too far from membrane). Scale bar is 0.25 µm. D. ImmunoEM analysis to determine INM localization of h2NLS-LP149-Sec61. Percentage of gold particles counted at nuclear side (black), cytoplasmic side (grey) of the membrane, or ambiguous (lumen or on membrane, white) are shown, with number of gold particles counted listed. F-G-M2-S was included as a reference sample as this protein is excluded from the INM (144). Analysis was done blinded.
resolution (also known as nanotomy; (189)). The gold particles were assigned, in a blind-folded fashion, as nuclear (N) or cytoplasmic (C) when within 75 nm of the NE inside or outside the nucleus, respectively, or as ambiguous when they were in the lumen or at membrane (indicated by the arrowheads in Figure 3.2C). The LP149 reporters, with and without NLS, were compared with the corresponding reporters with the native linker L. The large FKBP-GFP-2xMBP-Sec61 (F-G-M2-S) was used as an control for a protein that is excluded from the INM (144). Because expression of strongly NE-accumulated reporters led to deformed nuclei, which made identification of the NE and assignment of the gold particles as INM or ONM difficult, expression levels were decreased by fusing a FKBP domain to the N-terminus of the INM reporters (F-G-NLS-L-S and F-G-NLS-LP149-S). For each reporter at least 200 gold particles were counted, and percentages of N, C or ambiguous are shown in Figure 3.2D. The distribution of the LP149 reporters are similar to their L counterparts, which shows that the different linker does not greatly affect the localization of the reporters, and that the reporter with LP149 and NLS is indeed accumulated in the INM. These results show that even with a 149 proline residues long linker, that is less flexible and has less charged residues than the native Heh2-linker, proteins can still be imported to the INM through the NPC.

Disentangling import and efflux efficiency

To interpret the NE/ER ratio’s or the iEM data in terms of nuclear import efficiencies it is critical to consider all parameters that determine the steady state localization of the reporters. Most relevant are synthesis rates, retention mechanisms and the rate of efflux and import through the NPC. We first compared the rates of efflux between the membrane proteins with different linker regions. The Kap95-AA system was used to study the efflux of reporters from the INM (120). When rapamycin and glucose were added to cells expressing G-NLS-L-S, a gradual decrease in NE/ER ratio was observed over time (Figure 3.3A). Rapamycin blocked active import, because it tethered Kap95 at the PM, and so the only transport remaining is the passive diffusion through the INM/ONM/ER network. This resulted in a net efflux from the INM to the ONM and ER, resulting in a decreasing accumulation at the NE. The simultaneous addition of glucose halted expression, reducing the impact of newly synthesized proteins interfering with the measurements (see Figure 3.2A: the newly synthesized proteins would remain in the ER because the
Flexible linker needed for inner nuclear membrane targeting

import is blocked, which would also decrease the NE/ER ratio). So, the gradual decreasing NE/ER ratio of the reporter with the native linker showed that it leaks out from the INM as expected. In contrast, the NE/ER ratio for the reporter with LP149 did not decrease after blocking the import and stopping synthesis (Figure 3.3A). This showed that this reporter did not leak from the INM, either because it is retained in the INM at a nuclear component, or it could not passively go through the NPC.

The lack of diffusion to the ER upon depletion of Kap95 prompted us to validate if the protein with LP149 is mobile at the INM. Fluorescence recovery after photobleaching (FRAP) measurements showed the mobility of G-NLS-LP149-S to be comparable with G-NLS-L-S (Figure 3.3B), indicating that the lack of relocalization to the ER is not due to retention but rather to an inability to efflux through the NPC. Considering that accumulation is a result of import rate over efflux rate (assuming synthesis and degradation rates to be much slower than import), a low efflux rate would result in a higher NE/ER ratio if import rates are the same. The lack of leak for the reporter with

**Figure 3.3** Mobility of reporters with L or LP149. A. Time traces of leak of reporter proteins from INM after abolishing import by adding rapamycin which depletes Kap95 and stopping synthesis by adding glucose. White: G-NLS-L-S, black: G-NLS-LP149-S. Each time point is average of at least 30 cells, error bar: SEM. B. Mobility measurements of reporters at INM using FRAP. D is diffusion coefficient averaged over n cells. SEM is indicated.

<table>
<thead>
<tr>
<th>Reporter</th>
<th>Mobile fraction</th>
<th>D (μm²/s)</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>G-NLS-L-S</td>
<td>75 ± 3 %</td>
<td>1.4 ± 0.3</td>
<td>19</td>
</tr>
<tr>
<td>G-NLS-LP149-S</td>
<td>87 ± 3 %</td>
<td>1.3 ± 0.4</td>
<td>12</td>
</tr>
</tbody>
</table>
LP149 thus shows that also the import is less efficient with this linker when compared to the reporters with similar NE/ER ratios. We conclude that the LP149-reporter is actively imported to the INM at reduced rates compared to other linkers, and its efflux through the NPC is severely hindered by the all-proline linker.

**Important features of the linker for import efficiency**

Next, we addressed which properties of the linker, stokes radius or flexibility, are responsible for the decreased import to the INM. The polyproline linkers have more structure, making them less flexible than other linkers previously tested, but also the charge content is decreased, which might cause a more collapsed structure. Here, a coarse grained molecular dynamics model was used to simulate the dynamic structure of the linkers (56, 57). Using this model the Stokes radius can be predicted, which is a measure of how extended the linker is. The Stokes radius would be a better way to compare the polyproline linkers used in this study with the randomized and truncated ID linkers that have been used before, than the linker length in amino acids. The Stokes radii were calculated for the native and truncated linkers from (120) and plotted against the NE/ER ratio of the reporter proteins (domain structure GFP-h2NLS-linker-TM, measured at confocal microscope). Indeed the calculated stokes

![Stokes Radius (Å) vs NE/ER ratio](image)

**Figure 3.4** Correlation between NE/ER ratio of membrane proteins and Stokes radius of linker. NE/ER ratio (from confocal microscope) of GFP-h2NLS-linker-TM reporters with different linkers from (120) plotted against Stokes radius predicted with the coarse-grained MD model (56, 57). Truncations of native linker (L, white circles), randomized linker 1 (LR1, dark grey triangles) and randomized linker 2 (LR2, light grey squares). Error bars are SEM for NE/ER ratio and SD for Stokes radius.
Flexible linker needed for inner nuclear membrane targeting

radius and the measured NE/ER ratio show a striking correlation (Figure 3.4).

Because the Stokes radius of linkers correlates with accumulation at the INM, the effect of the proline stretches on Stokes radius and overall conformation were investigated. The model predicts that the Stokes radius decreases with more prolines present, up to LP100 (Figure 3.5A). Snapshots show that the short proline stretches, that can form PPII helices, fold back on top of each other. Combined with the lower charge content this makes a more collapsed structure more favourable. In the simulations LP149 formed an elongated structure, a long PPII helix, that folds back upon itself. This elongated conformation has a Stokes radius that is comparable to the native linker.

To test that the model is able to faithfully predict the properties of linkers with high proline content, we experimentally determined the Stokes radius of the LP50-linker. LP50 was expressed in Lactococcus lactis as a fusion protein with C-terminal TEV-cleavage site, GFP and His-tag. After purification the GFP and His-tag were cleaved off and the Stokes radius was determined by size-exclusion chromatography (Figure 3.5B). The experimental measurements at 47 Å were in good correspondence with the predicted value using the model (45 Å ± 4), showing that the model reliably predicts the Stokes radius of this polyproline linker.

Knowing that Stokes radius predicts import efficiency when similar amino acid compositions (and similar flexibility) make up the linker regions, the artificial polyproline linkers can now be used to dissect the effects of radius and flexibility. The NE/ER ratio measured for reporters with polyproline linkers was plotted against their modelled Stokes radius. The NE/ER ratio’s and stokes radii of reporters with the native Heh2-linker and 2 truncated ID linkers (120), that otherwise have the same domain topology (GFP-h2NLS-linker-Sec61) as the polyproline reporters, were included for comparison (Figure 3.5C). The plot shows a clear correlation between the hydrodynamic radius of the linker with the accumulation at the nuclear envelope for the random linkers and the polyproline linkers up to LP100, which indicates that a less collapsed structure is favourable for efficient import. However, the NE/ER ratio of the reporter with the all proline linker LP149 is lower than expected based on its Stokes radius. Considering that no efflux is observed for this protein we conclude the import rate of G-NLS-LP149-S is even lower compared to the other reporters reaching a similar NE/ER ratio. This shows that in addition to a large Stokes
radius, flexibility is important for efficient import. Altogether we conclude that a more extended structure, represented by a large Stokes radius, is required for import of these INM destined membrane proteins, but on top of that for highly efficient import flexibility is also essential.

Figure 3.5 Stokes radii of polyproline linkers and the correlation with import. A. Snapshots from modelling and Stokes radii predicted by the model of polyproline linkers. Standard deviation is indicated. B. Stokes radius determination of LP50. Top graph shows elution profile from gel filtration of LP50 after GFP was removed by TEV-cleavage. Fractions indicated were analyzed together with the loaded sample (S) on SDS-PAGE (inset). Arrowhead indicates LP50, asterisk labels uncleaved LP50-GFP. Bottom graph shows fractional volume of reference proteins (squares) and LP50 (cross) plotted against Stokes radius. C. NE/ER ratio of reporters (G-NLS-linker-S) with different linkers plotted against the Stokes radius from the computational model. White circle: L, black diamonds: proline linkers, grey triangles: LR1 truncations. Error bars are SEM for NE/ER ratio and SD for Stokes radius.
Discussion

To study how the biophysical properties of the linker impact active membrane protein import to the INM, the linker of Heh2 was replaced with linkers with increasing amounts of proline residues in a reporter protein. This yielded proteins with stretches of at least 5 consecutive prolines up to an all-proline stretch of 149 residues. Surprisingly, these reporter proteins expressed full length in *Saccharomyces cerevisiae*, although to a lower level than with the native linker. It has been shown that 3 or 4 consecutive prolines already stalls translation in yeast and that the eukaryotic translation elongation factor eIF5A is needed to relieve this ribosomal stalling (62). Also the longest stretch of consecutive prolines in the proteome of *S. cerevisiae* is 12 (in Formin (BNI1)) and the highest content of proline stretches is seen in VRP1, which has 61 prolines in 12 stretches of at least 3 prolines (116). This shows that 149 prolines in one protein is not a natural occurring sequence, but yeast is capable of handling these artificial proteins. Not only were the reporters expressed, they were also inserted in the membrane and the artificial linkers did not abolish targeting, resulting in INM localization of the reporters.

Increasing the amount of prolines in the linker reduced the flexibility of the intrinsically disordered spacer between the NLS and the transmembrane domain. Polyproline peptides have been used as molecular rulers in FRET, because they form rigid PPII helices, but it has been shown that for longer (24 residues) polyproline peptides the residues are not all in the trans conformation which results in PPII, but that some residues are in the cis conformation (63). Thus, most probably the LP149 is also not a rigid all-trans PPII helix, but some residues will be in a cis conformation, but this is still much less flexible than a ID domain. The other linkers used in this study have shorter stretches of prolines that will form PPII helices between intrinsically disordered regions, resulting in rods connected via flexible chains. The modelling of the linkers shows that instead of favouring a more extended conformation, these polyproline stretches tend to interact with and fold back on each other, making the linkers with prolines more compact than the native Heh2-linker. This is clearly seen when looking at the calculated Stokes radius, which decreases with higher proline content. Judging from the good correlation between import efficiency and Stokes radius for the LP50, LP70 and LP100 linkers, it seems that for transport to the INM the main property of the linker is a large Stokes radius. An (easily) extended linker is thus required for import to the INM. One
of the main reasons for this requirement might be the recruitment of soluble Kaps to bind the NLS. To have a linker between the membrane-embedded domain and the sorting sequence at the terminus, allows the sorting signal to move in three dimensions instead of the two-dimensional membrane (92).

In this study the LP149 linker clearly behaves distinctly. The Stokes radius calculated for this linker is rather large because of its extended conformation, even higher than the native linker. However, the accumulation at the NE is much lower than expected, especially when regarding the non-detectable efflux. Using this information it can be concluded that the role of the linker is not just extending into the cytoplasm to find and bind karyopherins that facilitate the transport of the protein over the NPC, because based on its Stokes radius the extended LP149 should be able to do this. Also spanning the distance from the membrane to the FG-nups should be possible with the all proline linker. Apparently flexibility is important for efficient import, possibly to slice through the scaffold of the NPC, where it needs to wiggle between the spokes of the NPC, as has been proposed before (120) or alternatively to align the highly curved pore membrane in case of a more peripheral passage way.

The lack of efflux observed for the reporter with the LP149 linker points to an even more important role of a flexible linker in leaking from the INM to the ER than in import. One explanation could be that leak of membrane proteins is going through the lateral channel of the NPC, which is a more spatially confined route (144). The extended and less flexible all-proline linker might be hindered in the diffusion through the lateral channel, hereby inhibiting efflux. Future high resolution imaging studies may resolve the dynamics of the NPC structure during transport of membrane proteins, and resolve the distinct routes of passive influx/efflux and active import of membrane proteins.

**Materials & Methods**

*Strains, plasmids & growth conditions*

All experiments were done in the KAP95-AA strain (W303, MATα tor1-1 fpr1::NAT PMA1-2×FKBP12::TRP1 Kap95-FRB::KanMX) (64, 120), with the exception of the iEM control sample F-G-2M-S which was expressed in strain Htb2-FRB (W303 MATα; tor1-1; fpr1::NAT; Htb2-FRB::kanMX6) (144). Plasmids (Table 3.1) were generated by standard molecular biology techniques and validated by sequencing; details are available upon request. Amino acid sequences of the linker used in this study are listed in Table 3.2.
Cells were grown at 30°C in selective drop-out medium, supplemented with 2% D-raffinose. Reporter proteins were expressed under control of the GAL1 promoter by 2h induction with 0.5% D-galactose. Rapamycin (LC laboratories, Woburn MA) was used at a final concentration of 5 µg/mL where appropriate. Incubation with 1% D-glucose was used to stop expression.

Spot assays were done using exponentially growing cells, diluted to 1x10⁶ - 1x10⁷ with drops of 4 µL on SD-His plates supplemented with 2% D-raffinose and 0.5% D-galactose or only 2% D-raffinose as a control.

**Fluorescence Microscopy**

Imaging was done on a DeltaVision Deconvolution Microscope (Applied Precision) at 30°C, using InsightSSITM Solid State Illumination of 488 nm and an Olympus UPLS Apo 100x oil objective with 1.4NA. Detection was done with a CoolSNAP HQ2 camera. SoftWoRx software was used, and image-stacks were deconvolved using standard settings. Data was analyzed with open source software Fiji (160).

FRAP was done on a LSM780 NLO confocal microscope (Carl Zeiss MicroImaging, Jena, Germany) at 30°C , using a “PlanNeofluar” 63x/1.3NA CorrDIC water or glycerine immersion objective, 34channelQuasar detector and ZEN acquisition software. Measurements were done essentially as described in (118). Data were analyzed using the ZEN2010B software package (Carl Zeiss) and fitted to diffusion equation for membrane proteins (39).

**Purification and Stokes radius determination of LP50**

*L. lactis* NZ9000 (96) was used as expression host for LP50. Liquid cultures were grown standing at 30°C in M17 medium (Oxoid, Hampshire UK) supplemented with 0.5% glucose and 5 µg/mL Chloramphenicol. Expression of LP50 was induced at OD 0.5 with 0.01% nisin supernatant for 4h. Purification and Stokes radius determination of LP50 was done as was described for h2NLS-L in (120).

**Modeling**

Simulations are performed using a one-bead-per-amino-acid coarse-grained molecular dynamics model (56, 57). The model accounts for hydrophobic and electrostatic interactions and can differentiate between all 20 naturally existing amino acids. The force field parameters of proline are specifically calibrated against the experimentally measured end to end distance of polyproline segments (163). The proteins are set at the fully extended conformation at the beginning of the simulations and are allowed to freely move and rotate through the medium during simulation. The Stokes radii are calculated according to the
### Table 3.1 Plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Descriptive name</th>
<th>Source</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pACM023</td>
<td>GFP-h2NLS-L-TM (G-NLS-L-TM)</td>
<td>pUG34, GAL1 promoter (instead of Met25), GFP fused to h2NLS-L-TM</td>
<td>(120)</td>
</tr>
<tr>
<td>pSI6</td>
<td>GFP-h2NLS-L-Sec61 (G-NLS-L-S)</td>
<td>pACM023, where the TM is replaced for Sec61 (Residues 25-471)</td>
<td>This study</td>
</tr>
<tr>
<td>pSI7</td>
<td>GFP-h2NLS-LP50-Sec61 (G-NLS-LP50-S)</td>
<td>pSI6, where L is replaced for LP50</td>
<td>This study</td>
</tr>
<tr>
<td>pSI8</td>
<td>GFP-h2NLS-LP70-Sec61 (G-NLS-LP70-S)</td>
<td>pSI6, where L is replaced for LP70</td>
<td>This study</td>
</tr>
<tr>
<td>pSI9</td>
<td>GFP-h2NLS-LP100-Sec61 (G-NLS-LP100-S)</td>
<td>pSI6, where L is replaced for LP100</td>
<td>This study</td>
</tr>
<tr>
<td>pSI11</td>
<td>GFP-h2NLS-LP149-Sec61 (G-NLS-LP149-S)</td>
<td>pSI6, where L is replaced for LP149</td>
<td>This study</td>
</tr>
<tr>
<td>pSI20</td>
<td>GFP-ΔNLS-L-Sec61 (G-ΔNLS-L-S)</td>
<td>pSI6, where h2NLS is removed</td>
<td>This study</td>
</tr>
<tr>
<td>pSI26</td>
<td>GFP-ΔNLS-LP149-Sec61 (G-ΔNLS-LP149-S)</td>
<td>pSI20, where L is replaced by LP149</td>
<td>This study</td>
</tr>
<tr>
<td>pAS004</td>
<td>FKBP-GFP-h2NLS-L-Sec61 (F-G-NLS-L-S)</td>
<td>pSI6 with N-terminal 2×FKBP12</td>
<td>This study</td>
</tr>
<tr>
<td>pPP037</td>
<td>FKBP-GFP-h2NLS-LP149-Sec61 (F-G-NLS-LP149-S)</td>
<td>pSI11 with N-terminal 2×FKBP12</td>
<td>This study</td>
</tr>
<tr>
<td>pPP011</td>
<td>FKBP-GFP-2xMBP-Sec61 (F-G-M2-S)</td>
<td></td>
<td>(144)</td>
</tr>
<tr>
<td>pSI15</td>
<td>h2NLS-LP50-GFP</td>
<td>Vector based on pNZ for expression of h2NLS-LP50-GFP in <em>L. lactis</em></td>
<td>This study</td>
</tr>
<tr>
<td>pACM044</td>
<td>GFP-h2NLS-LR1(138)-Sec61</td>
<td>Adapted from pACM023, where the L is replaced by a randomized and truncated version of 138 residues and TM is replaced for full-length Sec61 (Residues 25-471)</td>
<td>(120)</td>
</tr>
<tr>
<td>pACM046</td>
<td>GFP-h2NLS-LR1(78)-Sec61</td>
<td>pACM044 with linker truncated to 78 residues</td>
<td>(120)</td>
</tr>
</tbody>
</table>
Flexible linker needed for inner nuclear membrane targeting

Amino acid sequence of linkers (including h2NLS) used in this study. NLS sequence is underlined. First TM of Sec61 starts after this sequence.

<table>
<thead>
<tr>
<th>Linker</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>VKDENVETNKRKREQIQSTDNEAKMQIQEKSPKKKRKKRRSSKANPPQSDVASEEESPPLQFDEVENKNNPPPEREDGKDELPTEQPVKTELIPPLKVLIGARKVPYNQK</td>
</tr>
<tr>
<td>LP50</td>
<td>VKDENVETNKRKREQIQSTDNEAKMQIQEKSPKKKRKKRRSSKANPPQSDVASEEESPPLQFDEVENKNNPPPEREDGKDELPTEQPVKTELIPPLKVLIGARKVPYNQK</td>
</tr>
<tr>
<td>LP70</td>
<td>VKDENVETNKRKREQIQSTDNEAKMQIQEKSPKKKRKKRRSSKANPPQSDVASEEESPPLQFDEVENKNNPPPEREDGKDELPTEQPVKTELIPPLKVLIGARKVPYNQK</td>
</tr>
<tr>
<td>LP100</td>
<td>VKDENVETNKRKREQIQSTDNEAKMQIQEKSPKKKRKKRRSSKANPPQSDVASEEESPPLQFDEVENKNNPPPEREDGKDELPTEQPVKTELIPPLKVLIGARKVPYNQK</td>
</tr>
<tr>
<td>LP149</td>
<td>VKDENVETNKRKREQIQSTDNEAKMQIQEKSPKKKRKKRRSSKANPPQSDVASEEESPPLQFDEVENKNNPPPEREDGKDELPTEQPVKTELIPPLKVLIGARKVPYNQK</td>
</tr>
<tr>
<td>LR1(138)</td>
<td>VKDENVETNKRKREQIQSTDNEAKMQIQEKSPKKKRKKRRSSKANPPQSDV&gt;VSYPKTELIDPENDALPFEPSRIEKLDENIITSVSRDKRGSPNRVLGITSAPPPRELDAEETPLPPPETNDNESLSPPNPNHEPEKDKPPPVPNKVMIPKLVIGARKVPYNQK</td>
</tr>
<tr>
<td>LR1(078)</td>
<td>VKDENVETNKRKREQIQSTDNEAKMQIQEKSPKKKRKKRRSSKANPPQSDV&gt;VSYPKTELIDPENDALPFEPSRIEKLDENIITSVSRDKRGSPNRVLGITSAPPPRELDAEETPLPPPETNDNESLSPPNPNHEPEKDKPPPVPNKVMIPKLVIGARKVPYNQK</td>
</tr>
</tbody>
</table>

method explained in (57), in which each protein is simulated for $10^7$ steps and the average Stokes radius for the generated conformations (excluding the first $10^6$ steps) is calculated using the HYDRO computer program (22). The standard deviation of the Stokes radii (Fig. 4 and S1) is calculated based on the generated conformations after the first $10^6$ steps.

Immunoelectron Microscopy
Sample preparation essentially the same as published before (189). Cells for immuno-EM (iEM) were fixed in 2% paraformaldehyde (Merck, Germany) and 0.2% glutaraldehyde (Polysciences, Germany) in PHEM buffer (60 mM
PIPEC, 25 mM HEPES, 2 mM MgCl$_2$, 10 mM EGTA, pH 6.9), washed with PHEM buffer and incubated for 1h with 1% periodic acid in PHEM buffer. After washing the cell pellet was infiltrated with 12% gelatin (Sigma, Germany) in PHEM buffer. The sample was cut in blocks and incubated in 2.3 M sucrose in PHEM buffer overnight. The blocks were mounted on aluminium pins and put in liquid nitrogen. 70nm thin sections were cut with a cryo-ultramicrotome (Leica UC7, Austria) and picked up with a 2.3M sucrose (J.T. Baker, the Netherlands) in 2% methylcellulose (Sigma, USA) solution. The sections were placed on copper grids (Veco, the Netherlands) coated with formvar (Sigma, USA) support film sputtercoated (Leica, Austria) with carbon. Grids were placed on a solid gelatin plate at RT (Sigma, Germany; 2% in 0.1 M phosphate buffer), which was placed in an 37˚C incubator (30min). Grids were washed with PBS/glycine (Sigma, Japan) and subsequently blocked with PBS/1%BSA (Sanquin, Netherlands). Next, GFP was probed (Abcam, rabbit anti GFP; 1:200; 2 hours), washed with PBS/0.1%BSA and incubated with protein A gold (purchased from G. Posthuma, UMC-Utrecht; 1:50 for 30 minutes). After washing with PBS, grids were post-fixed with 1% glutaraldehyde in PBS. After washing with MilliQ water the samples were contrasted with uranylaloxalate-acetate (4% uranylacetate in water, 3.8% oxalic acid, pH 7) for 5 minutes and methylcellulose/uranylacetate (9ml 2% methylcellulose, 1ml 4% uranylacetate) for 5 minutes on ice.

Data was recorded using the Supra55 scanning electron microscope at 29kV (SEM; Zeiss, Oberkochen), at 2.5nm pixel size using the transmission detector essentially the same as described before (171). Large-area scans were generated using the external scan generator ATLAS (Fibics, Canada) and tiles were stitched using VEviewer (Fibics) and exported both as a single TIF and a high-resolution html file, which contains the raw data and is open-access available via www.nanotomy.org.

Acknowledgement

This work was supported by the Zernike Institute for Advanced Materials, the Netherlands Organization for Scientific Research (VIDI to LMV), and the European Research Institute for the biology of Ageing. Part of the work has been performed at the UMCG Imaging and Microscopy Center (UMIC), which is sponsored by NWO-grants 175-010-2009-023, ZonMW 91111.006 and STW 12718. The computations were carried out at the Centre for High Performance Computing and Visualization of the University of Groningen.