Supplemental Materials

Molecular Biology of the Cell

Kralt et al.

Supplemental Material to:

Conservation of inner nuclear membrane targeting sequences in mammalian Pom121 and yeast Heh2 membrane proteins

Annemarie Kralt*, Noorjahan B. Jagalur†, Vincent van den Boom‡, Ravi K. Lokareddy§, Anton Steen*, Gino Cingolani§, Maarten Fornerod† and Liesbeth M. Veenhoff*

Contents:

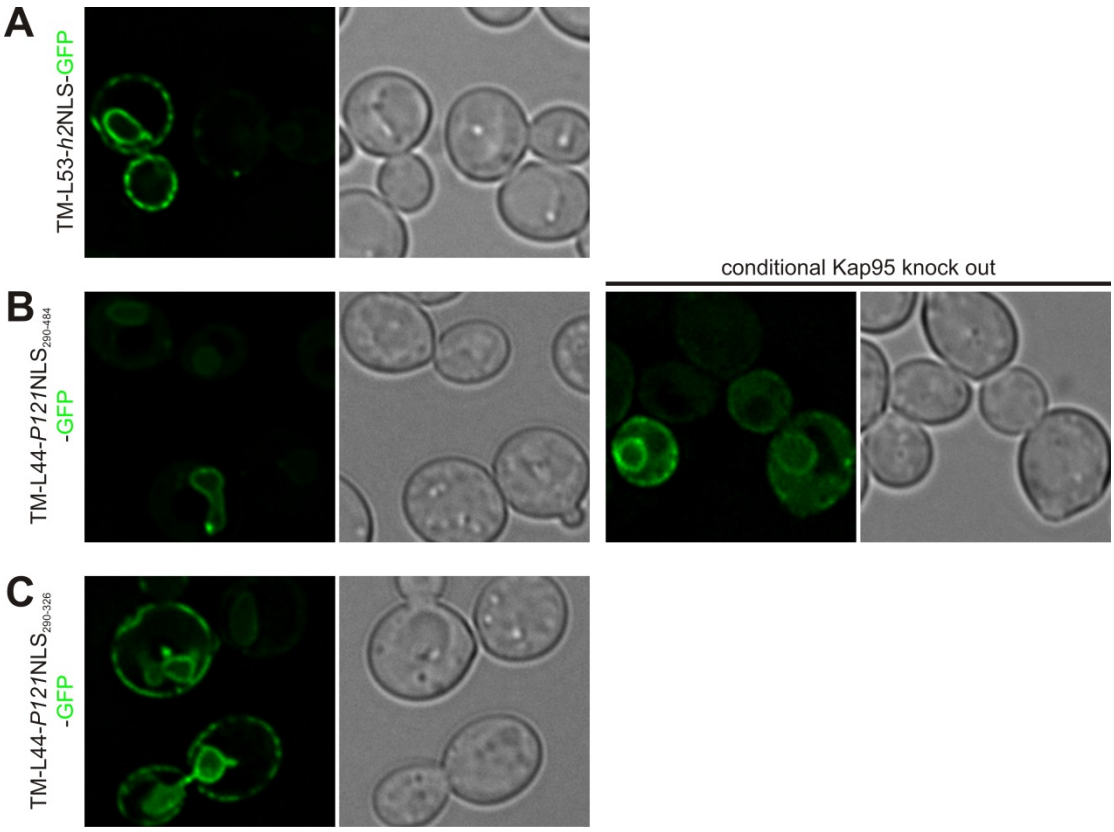

Figure S1 An ID linker region between the transmembrane segment and the NLS is required for INM import of proteins with TM-L-NLS topology (A) TM-L53-*h2*NLS-GFP, with an ID linker region that was shortened from 180 to 53 residues, did not accumulate at the INM, but was dispersed over the NE-ER network. (B) TM-L44-*Pom121*NLS$_{290-484}$-GFP with reduced ID linker region and the full *Pom121*NLS still localized to the INM in a Kap95-dependent manner. (C) TM-L44-*Pom121*NLS$_{290-326}$-GFP, which encodes the shortened ID linker and only the IBB-like region of *Pom121*NLS, localized to the NE-ER network and did not accumulate at the INM. The accumulation in (B) may be explained as the ~70 residues of the IBB-like and the spacing region to the third stretch of basic residues of the *Pom121*NLS region could function as part of the ID linker region, and the second C-terminal importin binding site in the *Pom121*NLS region mediates INM localization of this reporter.

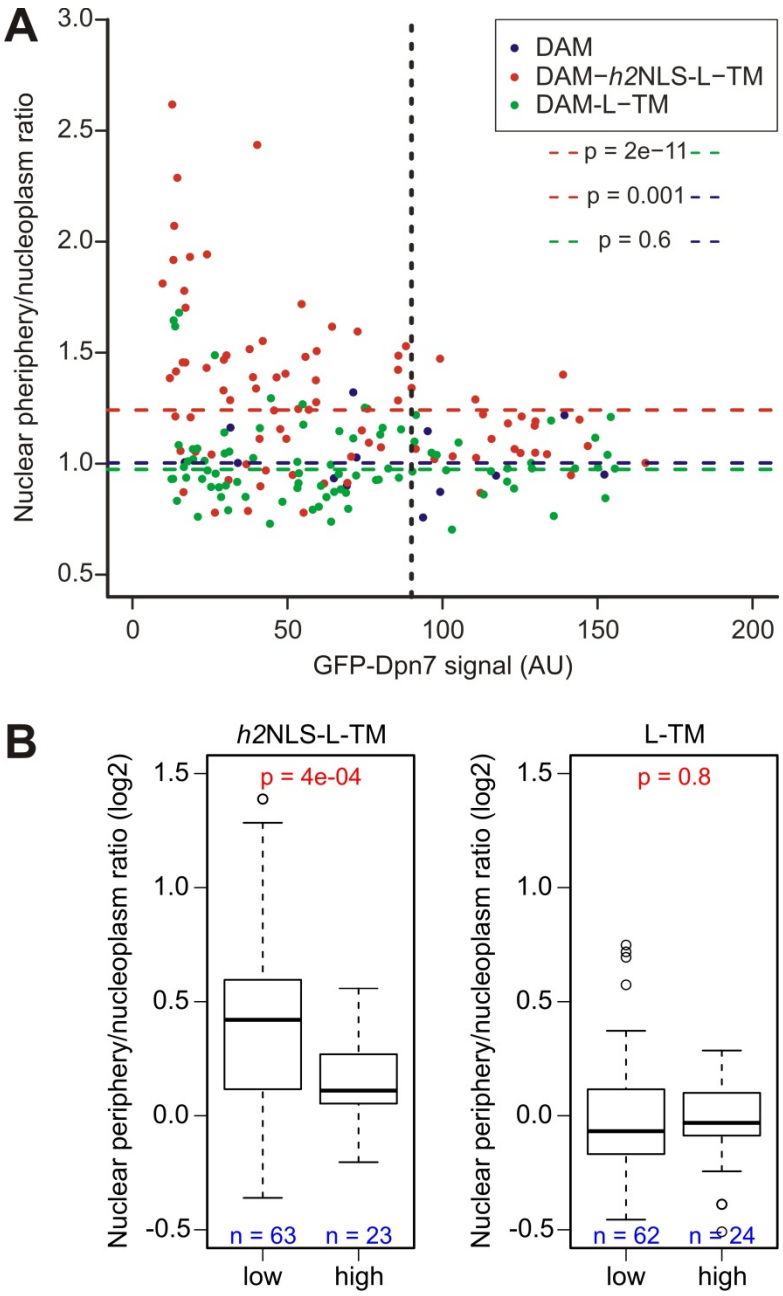

Figure S2 DAM methylation of DNA at the nuclear periphery confirms the NLS dependent INM localization of these proteins (A) The accumulation of GFP-Dpn7 at the periphery of the nucleus in cells expressing DAM (blue), DAM-*h2*NLS-L-TM (red) and DAM-L-TM (green) is represented as the ratio of average fluorescent signal at the cell periphery over the average fluorescent signal in the nucleoplasm (periphery-nucleoplasm ratio or P/N ratio). The periphery-nucleoplasm ratio was plotted against the average GFP-Dpn7 signal in the nucleus. The mean ratios are indicated by the dotted lines. An arbitrary cut-off between cells that express high and low levels of GFP-Dpn7 was set at 90 AU (black dotted line). P-values are calculated using the Wilcoxon test. (B) Boxplots comparing the P/N ratios of D-*h2*NLS-L-TM and D-L-TM expressing cells with low or high GFP-Dpn7 expression levels. No significant difference was found for cells expressing L-TM, but a significant difference was found for cells expressing *h2*NLS-L-TM. Logged data follows a normal distribution according to Shapiro-Wilk and P-values are calculated using Student t-test.

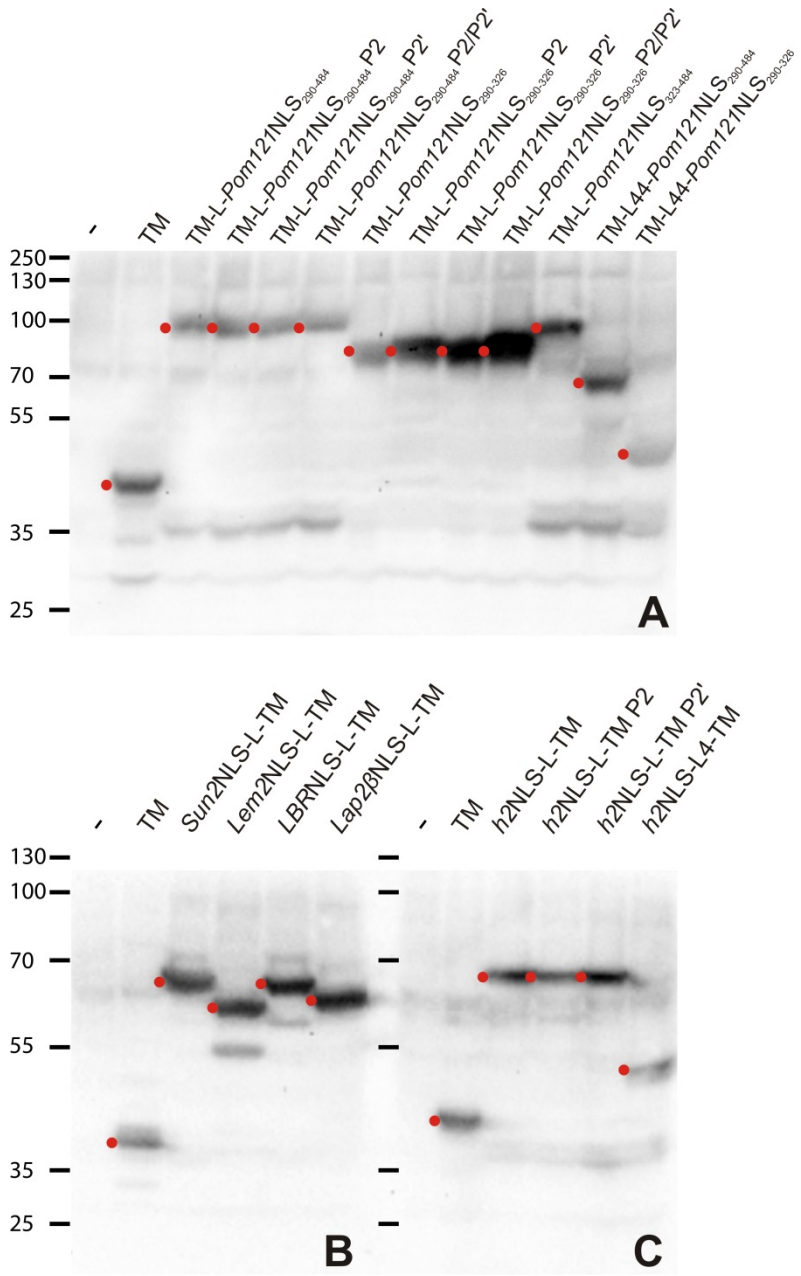

Figure S3 Visualization of expression of NLS-L-TM and TM-L-NLS reporter proteins in yeast by Western blotting. Cell lysates were prepared from equal amounts of *S.cerevisiae* KAP95AA cells expressing C-terminal (A,C) and N-terminal GFP-tagged reporter proteins and anti-GFP was used for immuno-detection.

Table S1: Data collection and refinement statistics.

	ΔIBB-importin α1:Pom121NLS
Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	78.3, 89.8, 97.9
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	15-1.80
No. reflections (tot/unique)	992,426 / 62,116
R_{sym}	5.7 (50.1)
$I / \sigma I$	34.7 (2.1)
Completeness (%)	97.8 (98.3)
Redundancy	3.6 (3.2)
Refinement	
Resolution (Å)	15-1.80
No. reflections used	60,036
R_{work} / R_{free}	18.2 / 20.9
No. atoms	
Protein	3,241
Ligand (Pom121NLS)	449
Water	607
B-factors (Å^2)	
Protein	28.6
Ligand/ion	71.8
Water	39.6
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.09

Values in parentheses are for highest-resolution shell (1.86-1.80).
The R_{free} was calculated using 5% of randomly selected reflections.

Table S2: Structural alignment of Pom121NLS with other NLSs visualized crystallographically in complex with importin α and Kap60.

NLS type	Minor Binding Site															Major Binding Site							PDB id
	P_{-1}'	P_0'	P_1'	P2'	P_3'	P_4'	P_5'									P_1	P2	P_3	P_4	P_5			
SV40T-ag	P	K	K	**K**	R	K										P	K	**K**	K	R	K	V	1EJL/1BK6*
hPLSCR1-NLS																G	**K**	I	S	K	HWTGI		1Y2A
hPLSCR4-NLS	S	I	I	**R**	K	W	N																3Q5U
Guα-NLS	G	Q	K	**R**	S	F	S																3ZIN
A89-NLS	L	G	K	**R**	K	Y	W																4B8P**
B54-NLS	L	G	K	**R**	K	R	H																2YNS**
TPX2			K	**R**	K	H										P	V	**K**	M	I	K		3KND
C-Myc			K	**R**	V	K	L									P	A	A	**K**	R	V	K	1EE4*
NP	A	V	K	**R**	P	A	A						T	K	K	A	K	**K**	K	K	L	D	1EJY/1EE5*
Kap60-IBB	E	L	R	**R**	R	R	D	T	Q	Q	V	E	L	R	K	L	A	**K**	R	R	N	F	1WA5*
h1NLS	T	R	K	**R**	K	D	P				D	S	D	D	W	N	S	**K**	E	N	K	ID	4XZR *
h2NLS	T	N	K	**R**	K	R	E	Q	I	S	T	D	N	E	A	P	K	**K**	K	R	K	KRSSKANK	4PVZ*
Pom121NLS	K	E	K	**K**	K	R	T				V	A	E	E	D	E	N	**K**	R	R	R	HDSS	4YI0*

* denotes yeast importin α (Kap60).
** denotes rice importin α. In all other cases, mammalian importin α was co-crystallized with NLSs α

Table S3: Plasmids for S. cerevisiae

Plasmid Name (Reporter name)	**Encodes[3]**
pAK80 (GFP-TM)[1]	GFP-Heh2$_{300-379}$
pACM023 (GFP-h2NLS-L-TM)[1]	GFP-Heh2$_{93-379}$
pAK99 (*Sun2*NLS-L-TM)[1]	GFP-*Hs*Sun2$_{18-106}$-Heh2$_{138-379}$
pAK100 (*Lem2*NLS-L-TM)[1]	GFP-*Mm*Lem2$_{128-145}$-Heh2$_{138-379}$
pAK102 (*LBR*NLS-L-TM)[1]	GFP-*Hs*LBR$_{55-113}$-Heh2$_{138-379}$
pAK103 (*Lap2β*NLS-L-TM)[1]	GFP-*Hs*Lap2β$_{248-265}$ -Heh2$_{138-379}$
pAK01 (TM-GFP)[2]	TM-GFP (encodes synthetic TM)
pAK54 (TM-L-*Pom121*NLS$_{290-484}$)[2]	TM-Heh2$_{145-315}$-Pom121$_{290-484}$-GFP
pAK55 (TM-L-*h2*NLS)[2]	TM-Heh2$_{145-315,93-145}$-GFP
pAK95 (TM-L-*Pom121*NLS$_{290-326}$)[2]	TM-Heh2$_{145-315}$-Pom121$_{290-326}$-GFP
pAK96 (TM-L-*Pom121*NLS$_{290-326}$_P2)[2]	TM-Heh2$_{145-345}$-Pom121$_{290-326}$ (K313→A) -GFP
pAK97 (TM-L-*Pom121*NLS$_{290-326}$_P2')[2]	TM-Heh2$_{145-345}$-Pom121$_{290-326}$ (K295→A) -GFP
pAK105 (TM-L-*Pom121*NLS$_{290-484}$_P2/P2')[2]	TM-Heh2$_{145-345}$-Pom121$_{290-484}$ (K295 and 313→A) -GFP
pAK107 (TM-L- *Pom121*NLS$_{323-484}$)[2]	TM-Heh2$_{145-345}$-Pom121$_{323-484}$-GFP
pAK93 (TM-L44-*Pom121*NLS$_{290-484}$)[2]	TM-Heh2$_{145-153,305-315}$-Pom121$_{290-484}$
pAK94 (TM-L53-*h2*NLS)[2]	TM-Heh2$_{145-153,305-315,93-145}$
pAK104 (TM-L44- *Pom121*NLS$_{290-326}$)[2]	TM-Heh2$_{145-153,305-315}$-Pom121$_{290-326}$

[1]Originates from pUG34, *MET25* promoter is replaced by *GAL1* promoter, N-terminal GFP fusion, HIS3

[2]Originates from pUG23, *MET25* promoter is replaced by *GAL1* promoter, C-terminal GFP-fusion, HIS3

[3]In the Heh2 proteins residues 93-138 are encoding the NLS region and residues 138-379 encode the linker region plus the first TM region. Residues 145-315 encode a major part of the linker region.

Table S4: S. cerevisiae strains

Strains name	Genotype	Sources
HHY118 (Kap95AA)	MATα tor1-1 frp1::NAT PMA1-2xFKBP12::TRP1 KAP95-FRB::KAN	(Haruki et al., 2008)
heh2Δ	BY4742 *heh2::KAN*	Yeast KO collection
AS1	BY4742 *heh2::GFP-HEH2-NAT*	This study
AS2	BY4742 *heh2::GFP-HEH2(Δh2*NLS*)-NAT*	This study
AS5	BY4742 *heh2::GFP-HEH2(Δh2*NLS*,Pom121*$NLS_{291-320}$*)-NAT*	This study
AS4	BY4742 *heh2::GFP-NAT*	This study
BY4741	MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	Invitrogen
nup84Δ	BY4741 *nup84Δ::KAN*	Yeast KO collection

Table S5: Plasmids for HEK293T cells

Plasmid Name (<u>Reporter name</u>)	**Encodes[3]**
pAK111 (2GFP)[1]	2GFP
pAK112 (2GFP-*h2*NLS)[1]	2GFP-$Heh2_{93-145}$
pAK113 (2GFP-*h2*NLS_P2)[1]	2GFP-$Heh2_{93-145}$(K126→A)
pAK114 (2GFP-*h2*NLS_P2')[1]	2GFP-$Heh2_{93-145}$(R103-A)
pAK115 (2GFP-*Pom121*$NLS_{290-484}$)[1]	2GFP-$Pom121_{290-484}$
pAK118 (2GFP-*Pom121*$NLS_{290-484}$P2/P2')[1]	2GFP-$Pom121_{290-484}$(K295 and 313→A)
pAK119 (2GFP-*Pom121*$NLS_{290-326}$)[1]	2GFP-$Pom121_{290-326}$
pAK120 (2GFP-*Pom121*$NLS_{290-326}$_P2)[1]	2GFP-$Pom121_{290-326}$(K313→A)
pAK121 (2GFP-*Pom121*$NLS_{290-326}$_P2')[1]	2GFP-$Pom121_{290-326}$(K295→A)
pAK122 (2GFP-*Pom121*$NLS_{290-326}$_P2P2')[1]	2GFP-$Pom121_{290-326}$(K295 and 313→A)
pAK123 (2GFP-*NP*NLS)[1]	2GFP-*X*/Nucleoplasmin$_{153-172}$
pAK37 (GFP-*h2*NLS-L-TM)[1]	GFP-$Heh2_{93-378}$
pAK38 (GFP-L-TM)[1]	GFP-$Heh2_{138-378}$
pAK108 (GFP-*h2*NLS-L37-TM)[1]	GFP-$Heh2_{93-145,289-378}$
pAK49 (DAM-*h2*NLS-L-TM)[2]	DAM-$Heh2_{93-378}$
pAK50 (DAM-L-TM)[2]	DAM-$Heh2_{138-378}$

[1]Originates from peGFP-c1

[2]Originates from pIND/V5-HisA (Invitrogen and (Vogel *et al.*, 2006)

Table S6: Amino acid sequences of reporter proteins in S. cerevisiae[1]

GFP-TM (pAK80) (TM: Heh2$_{300-379}$)	MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGYGVQC FARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK LEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKD PNEKRDHMVLLEFVTAAGITHGMDELYKSRTSGSRMKANKTKRGIDIMKPFIAHL*FIWLWNGAIFLSIICP* *ILFGLWYREQRIQVGYCGHEKPLKSLAISAFPQTERVDSVLQAYRP*
GFP-NLS-L-TM (L-TM: Heh2$_{138-379}$) (At "_***NLS***_" the below indicated NLS(-like) sequences are introduced)	MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGYGVQC FARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK LEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKD PNEKRDHMVLLEFVTAAGITHGMDELYKSRTS_***NLS***_PPESPPQSKSDGKATSADLTSELETVEELHKKDS SDDKPRVKELPKPELPNLKVSNEFLAQLNKELASAATENYDHSIKSTDLSSIRIETEEPVGPSTGAETRNESE VMENINLEVQPEVKEAKEELTKISETFDNQDEEDTSRLSSKKNIRSPKGRTRHFIANKTKRGIDIMKPFIAHL *FIWLWNGAIFLSIICPILFGLWYREQRIQVGYCGHEKPLKSLAISAFPQTERVDSVLQAYRP*
*h2*NLS (pACM023) (*Saccharomyces cerevisiae*)	HGENLLGQGVKDENVETNKRKREQISTDNEAKMQIQEEKSPKKKRKKRSSKANK
*Sun2*NLS (pAK99) (*Homo sapiens*)	GSGSSSSGGSSVAGSQSTLFKDSPLRTLKRKSSNMKRLSPAPQLGPSSDAHTSYYSESLVHESWFPPRSSLE ELHGDANWGEDLRVRRRRG
*Lem2*NLS (pAK100) (*Mus musculus*)	GSGLSYPPHAGPGPLRRRAS
*LBR*NLS (pAK102) (*Homo sapiens*)	GSKPLTSFRQRKGGSTSSSPSRRRGSRSRSRSRSPGRPPKSARRSASASHQADIKEARREV
*Lap2β*NLS (pAK103) (*Rattus norvegicus*)	GSRESTRGSRRTPRKRVETS
TM-GFP (pAK01) (TM: synthetic)	*MPRYAQLVSDVRETQPFASIALSKLPREHGCYGVQIRQERYWLGFLIPCIISLFIAGNWLWIF*LHAIFPKMI DIGRKTKNASRTSGSPGLQEFDIKLIDTVDLDMSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATY GKLTLKFICTTGKLPVPWPTLVTTFGYGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK
TM-L-NLS-GFP (TM: synthetic, L: Heh2$_{145-315}$) (At "_***NLS***_" the below indicated NLS sequences are introduced)	*MPRYAQLVSDVRETQPFASIALSKLPREHGCYGVQIRQERYWLGFLIPCIISLFIAGNWLWIFL*HAIFPKMI DIGRKTKNASRTSGSKSDGKATSADLTSELETVEELHKKDSSDDKPRVKELPKPELPNLKVSNEFLAQLNKE LASAATENYDHSIKSTDLSSIRIETEEPVGPSTGAETRNESEVMENINLEVQPEVKEAKEELTKISETFDNQD EEDTSRLSSKKNIRSPKGRTRHFIANKTKRGIDIMKPFI_***NLS***_DIKLIDTVDLDMSKGEELFTGVVPILVELD GDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGYGVQCFARYPDHMKQHDFFKSA MPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQ KNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAA GITHGMDELYK
*Pom121*NLS$_{290-484}$ (pAK54/pAK105) (*Rattus norvegicus*)	ALKEK**K**KRTVAEEDQLHLDGQEN**K**RRRHDSSGSGHSAFEPLVANGVPAAFVPKPGSLKRSLASQSSDDH LNKRSRTSSVSSLTSTCTGGIPSSSRNAITSSYSSTRGVSQLWKRSGPTSSPFSSPASSRSQTPERPAKKTREE EPCHQSSSSAPLVTDKESPGEKVTDPATGKQQSLWTSPPTPGSSGQRKRKIQ
*h2*NLS (pAK55) (*Saccharomyces cerevisiae*)	AVKDENVETNKRKREQISTDNEAKMQIQEEKSPKKKRKKRSSKANKPPESPPQS
Pom121NLS$_{290-326}$ (pAK95/pAK96/pAK97) (*Rattus norvegicus*)	ALKEK**K**KRTVAEEDQLHLDGQEN**K**RRRHDSSGSGHSA
*Pom121*NLS$_{323-484}$ (pAK107) (*Rattus norvegicus*)	AGHSAFEPLVANGVPAAFVPKPGSLKRSLASQSSDDHLNKRSRTSSVSSLTSTCTGGIPSSSRNAITSSYSST RGVSQLWKRSGPTSSPFSSPASSRSQTPERPAKKTREEEPCHQSSSSAPLVTDKESPGEKVTDPATGKQQS LWTSPPTPGSSGQRKRKIQ
TM-NLS-L44/53-GFP (At "_***NLS***_" the below indicated NLS sequences are introduced)	*MPRYAQLVSDVRETQPFASIALSKLPREHGCYGVQIRQERYWLGFLIPCIISLFIAGNWLWIF*LHAIFPKMI DIGRKTKNASRTSGSKSDGKATSRGIDIMKPFI_***NLS***_DIKLIDTVDLDMSKGEELFTGVVPILVELDGDVN GHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGYGVQCFARYPDHMKQHDFFKSAMPEGY VQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKV NFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHG MDELYK
*Pom121*NLS$_{290-484}$ (pAK93) (*Rattus norvegicus*)	ALKEKKKRTVAEEDQLHLDGQENKRRRHDSSGSGHSAFEPLVANGVPAAFVPKPGSLKRSLASQSSDDHL NKRSRTSSVSSLTSTCTGGIPSSSRNAITSSYSSTRGVSQLWKRSGPTSSPFSSPASSRSQTPERPAKKTREEE PCHQSSSSAPLVTDKESPGEKVTDPATGKQQSLWTSPPTPGSSGQRKRKIQ
*h2*NLS (pAK94) (*Saccharomyces cerevisiae*)	AVKDENVETNKRKREQISTDNEAKMQIQEEKSPKKKRKKRSSKANKPPESPPQS
*Pom121*NLS$_{290-326}$ (pAK104)	ALKEKKKRTVAEEDQLHLDGQENKRRRHDSSGSGHSA

[1]Green: GFP, *italics: transmembrane region*, black: ID linker region, red: NLS region with in black and black/underlined the P2 and P2' position respectively

Table S7: Amino acid sequences of reporter proteins in HEK293T cells[1]

2GFP (pAK111) Between ***brackets*** is absent in 2GFP-NLS fusions where the below indicated NLS sequences are introduced	MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQC FSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK LEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKD PNEKRDHMVLLEFVTAAGITLGMDELYKSGLRSRAMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGE GDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDG NYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGS VQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKQASNS(AVDGTAGPGSTGSR)
*h2*NLS (pAK112) (*Saccharomyces cerevisiae*)	VKDENVETNK**R**KREQISTDNEAKMQIQEEKSPK**K**KRKKRSSKANKPPESPPQS
*Pom121*NLS$_{290-484}$ (pAK113/AK114/ pAK115) (*Rattus norvegicus*)	ALKEK**K**KRTVAEEDQLHLDGQEN**K**RRRHDSSGSGHSAFEPLVANGVPAAFVPKPGSLKRSLASQSSDDH LNKRSRTSSVSSLTSTCTGGIPSSSRNAITSSYSSTRGVSQLWKRSGPTSSPFSSPASSRSQTPERPAKKTREE EPCHQSSSSAPLVTDKESPGEKVTDPATGKQQSLWTSPPTPGSSGQRKRKIQ
*Pom121*NLS$_{290-326}$ (pAK119/pAK120/pAK121/ pAK122)	ALKEK**K**KRTVAEEDQLHLDGQEN**K**RRRHDSSGSGHSA
*NP*NLS (pAK123)	LAAVKRPAATKKAGQAKKKKLDGSTGSR
*h2*NLS-L-TM *Saccharomyces cerevisiae* Heh2 $_{93-378}$ sequence C-terminally fused to GFP or DAM (pAK37/ pAK49) Underlined: absent in L-TM (pA38/ pAK50) Between brackets: absent in *h2*NLS-L37-TM (pAK108)	VKDENVETNKRKREQISTDNEAKMQIQEEKSPKKKRKKRSSKANKPPESPPQS(KSDGKATSADLTSELET VEELHKKDSSDDKPRVKELPKPELPNLKVSNEFLAQLNKELASAATENYDHSIKSTDLSSIRIETEEPVGPST GAETRNESEVMENINLEVQPEVKEAKEELTKISETFDNQDEEDTSRLSSKKNI)RSPKGRTRHFIANKTKRGI DIMKPFIAHL*FIWLWNGAIFLSIICPILFGLWYREQRIQVGYCGHEKPLKSLAISAFPQTERVDSVLQAYR*

[1]Green: GFP, *italics: transmembrane region*, black: ID linker region, red: NLS region with in black and black/underlined the P2 and P2' position, respectively