Computational and Structural Biotechnology Journal 20 (2022) 5952–5961

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Recognition motifs for importin 4 [(L)PPRS(G/P)P] and importin 5 [KP(K/Y)LV] binding, identified by bio-informatic simulation and experimental *in vitro* validation

Athanasios A. Panagiotopoulos [a,1], Konstantina Kalyvianaki [a,1], Paraskevi K. Tsodoulou [b], Maria N. Darivianaki [b], Dimitris Dellis [c], George Notas [a], Vangelis Daskalakis [d], Panayiotis A. Theodoropoulos [e], Christos A. Panagiotidis [b], Elias Castanas [a,*], Marilena Kampa [a,*]

[a] *Laboratory of Experimental Endocrinology, School of Medicine, University of Crete, 71013, Greece*
[b] *Laboratory of Pharmacology, School of Pharmacy, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece*
[c] *National Infrastructures for Research and Technology, Athens 11523, Greece*
[d] *Department of Chemical Engineering, Cyprus University of Technology, Limassol, Cyprus*
[e] *Laboratory of Biochemistry, School of Medicine, University of Crete, 71013, Greece*

ARTICLE INFO

Article history:
Received 17 June 2022
Received in revised form 10 October 2022
Accepted 11 October 2022
Available online 26 October 2022

Keywords:
Importin 4
Importin 5
Nuclear localization signal (NLS)
Karyopherins

ABSTRACT

Nuclear translocation of large proteins is mediated through karyopherins, carrier proteins recognizing specific motifs of cargo proteins, known as nuclear localization signals (NLS). However, only few NLS signals have been reported until now. In the present work, NLS signals for Importins 4 and 5 were identified through an unsupervised *in silico* approach, followed by experimental *in vitro* validation. The sequences LPPRS(G/P)P and KP(K/Y)LV were identified and are proposed as recognition motifs for Importins 4 and 5 binding, respectively. They are involved in the trafficking of important proteins into the nucleus. These sequences were validated in the breast cancer cell line T47D, which expresses both Importins 4 and 5. Elucidating the complex relationships of the nuclear transporters and their cargo proteins is very important in better understanding the mechanism of nuclear transport of proteins and laying the foundation for the development of novel therapeutics, targeting specific importins.

1. Introduction

Protein shuttling among cellular compartments has evolved in eucaryotic cells. An elegant system is responsible for the cytoplasmic-nuclear transport, involving specialized transporters named collectively karyopherins. This family of specialized molecules comprises at least 20 different proteins, which form three distinct classes: exportins, responsible for nucleo-cytoplasmic protein translocation through the nuclear pore; importins, involved in the cytoplasmic-nuclear trafficking; and adaptor proteins, necessary in many cases for the formation of the importin-cargo protein complex (see [1] for a review).

An additional control over importin-mediated nuclear transport is provided by the small Ras related GTPase Ran, controlling the formation and the stability of importin–cargo complexes [2,3]. The direction of nuclear transport is determined by the GTP- versus GDP-bound forms of Ran in the nucleus and the cytoplasm. Ran-GDP binds to importin-cargo complexes and regulates their cytoplasmic-nuclear transport. Once in the nucleus, a GDP-GTP exchange takes place, and RanGTP causes cargo release [4,5].

Cargo proteins contain specific sequence motifs named nuclear localization signal (NLS), responsible and necessary for the identification and the binding of importins. Until recently, few NLS motifs were recognized (see references [6,7] for reviews) for Importin α (IPOα) [8,9] and the M9 NLS (recognized by importin β2, also known as transportin) [10–12], with an increasing number of proteins expressing this sequence (the monopartite classical Importin α NLS sequences are KRRR and KRKXK [13–20]). However, progress in structural and analytical biology led to the identification of a number of protein complexes with other impor-

Abbreviations: IPOα, Importin α; IPOβ, Importin β; IPO4, Importin 4; IPO5, Importin 5; IPO7, Importin 7.

* Corresponding authors.
E-mail addresses: castanas@uoc.gr (E. Castanas), kampam@uoc.gr (M. Kampa).

[1] Equally contributing authors.

https://doi.org/10.1016/j.csbj.2022.10.015

tins, but without the formal identification of other NLS motifs [21–23].

Recently, using a bio-informatics approach, based on bibliographic and simulation data, and experimental *in vitro* validation, we presented the sequence EKRKI(E/R)(K/L/R/S/T) as a recognition motif for binding with importin 7 [24]. Here, using a similar approach, we report that the sequences (L)PPRS(G/P)P and KP(K/Y)LV are recognition motifs for Importins 4 and 5 binding, respectively, involved in the trafficking of important proteins into the nucleus (in the NLS sequences the X indicates any residue, slash (/) indicates alternative residues, and an amino acid in brackets, e.g. (L), denotes an optional residue). Our discovery might have an immediate translational importance, for the development of specific pharmaceuticals targeting cytoplasmic-nuclear trafficking of proteins.

2. Material and methods

2.1. In silico methods

2.1.1. Identification of Importin 4 & 5 NLS sequences

The bio-informatics methods used for the identification of NLS motifs on cargo proteins have been presented *in extenso* in a previous publication of our group [24]. In summary, the following steps were followed:

1. Protein sequences were retrieved from the NCBI protein database (https://www.ncbi.nlm.nih.gov/protein/) in FASTA format and entered to the Swiss Model Biospace (https://swissmodel.expasy.org/interactive) [25]. PDB codes for the proteins were retrieved from the Protein databank (https://www.rcsb.org/) [26]. Only the predicted model(s) with a 100 % homology was retained. In cases, where the only receptor crystal model included small molecules, co-crystalized with the receptor, the small molecules were manually removed from the PDB files, with a text editor. For proteins that crystalized structures were not available, the best model from the Swiss Model Biospace (https://swissmodel.expasy.org/interactive) [25] was retained, based on the sequence coverage homology (at least 70 %) with an already published proteins. Subsequently, protein files (in PDB format) were uploaded to the Galaxy Refine server (https://galaxy.seoklab.org/) [27–29] and fully flexibly refined (Routine REFINE). Galaxy Routine Refine performs repeated structure disturbance on side chains, secondary structure elements, and loops, pursued by molecular dynamics simulations [30].
2. Protein 3D conformation was compared with the recently released AlphaFold database [31]. Whenever possible, a comparison of the flexible retained model with existing crystals and AlphaFold reported structures was performed with the Chimera program, V 1.14 (https://www.cgl.ucsf.edu/chimera/) [32].
3. Ran-GDP complexes, as well as Importins 4 or 5 complexes with Ran-GDP were performed in the Galaxy server (routines LigDock and GalaxyHeteromer respectively) and the retained DeltaG (ΔG) values were refined with the Routine REFINE at the Galaxy server. LigDock predicts 3D structures of protein–ligand complexes [33,34] and GalaxyHeteromer predicts 3D structures of protein–protein complexes by template-based and ab initio docking ([35]). From the returned results, the interacting (binding) interface was identified and the corresponding amino acids were retrieved, both for importins and cargo proteins. The 3D structure of the interacting amino acids for both Importin 4 and 5 were modeled in the GalaxyWeb server, routine TMB.
4. The interacting cargo protein sequences were also modeled in GalaxyWeb (routine TMB) and (rigid) binding of the two sequences was performed using the Hex 8.0.8 program (https://hex.loria.fr/) [36,37], in PDB format. ΔG (change in Gibbs free energy) values were retrieved and reported. The binding of the retained cargo sequences was repeated, after elimination of one amino acid from the N- or the C-terminus of each peptide (and remodeling the remaining peptide in GalaxyWeb).
5. After the last step of the above procedure, a graph was constructed, with the obtained ΔG values at each round. From this graph, we have retained the amino-acids whose elimination provokes a significant change (increase) in the returned ΔG value, as necessary for peptide-importin binding (see Results and Supplemental Material for concrete examples).
6. Finally, we aligned all retained peptide sequences, with the online tools of Jalview (https://www.jalview.org) [38], and retrieved the consensus sequence for importins 4 and 5 binding.

2.1.2. Molecular dynamics

2.1.2.1. System setup. The initial coordinates were obtained from the predicted structures of the binding domains of Importins 4 (NLS4, residues 403–616) and 5 (NLS5, residues 304–603), along with their docked conformations with the putative importin 4 and 5 recognition sequences (NLS) (L)PPRS(G/P)P and KP(K/Y)LV, identified herein. For NLS4, His-49, 51, 52 and 202 are protonated at the Nε site, while the rest of His at the N_δ site. Asp-177 was treated as protonated, while the rest of Asp, Glu residues are deprotonated. For NLS5, His-109 is protonated at the Nε site, while the rest of His at the N_δ site. Asp-41 is treated as protonated, while the rest of Asp residues are deprotonated. Glu-13, 25 and 168 are treated as protonated, while the rest of Glu residues are deprotonated. These protonations retain the original hydrogen bonding network and are in accordance with the propka method (PDB2PQR) [39] predictions at a physiological pH value of 7.3. The Amber ff14sb force field [40] has been employed for the protein and peptides. The systems are hydrated by around 26,000 Tip3p water molecules [41]. A concentration of KCl at ~150 mM was added, with a ~38 mM K^+ surplus to neutralize the system. Thus, eight (8) different systems were built of around 75,500 atoms in a cubic unit cell of a 9.15 nm^3 volume. These refer to: NLS4, NLS4-PPRSGP, NLS4-PPRSPP, NLS4-LPPRSGP, NLS4-LPPRSPP and NLS5, NLS5-KPKLV, NLS5-KPYLV.

2.1.2.2. Molecular dynamics. The all-atom models, as defined previously, were used for the all-atom Molecular Dynamics Simulations. Based on published protocols [42,43], all models were relaxed and equilibrated with gradual removable of constraints on the protein backbone-heavy atoms. In a series of constant volume nVT, and constant pressure nPT ensembles, the temperature was increased from 100 K to 310 K, prior to the production runs [42,43]. For the classical Molecular Dynamics (MD) simulations, Newton's equations of motion were integrated, with a time step of 2.0 fs, for a total of 100 ns. The leap-frog integrator in GROMACS 2021 was used [44]. The production runs had been performed in the constant pressure nPT ensemble, with isotropic couplings (compressibility at $4.5x10^{-5}$). Van-der-Waals interactions were smoothly switched to zero, between 1.0 and 1.2 nm, with the Verlet cut-off scheme. Electrostatic interactions were truncated at 1.2 nm (short-range) and long-range contributions were computed within the PME approximation [45,46]. All hydrogen – heavy atom bond lengths were constrained, employing the LINCS algorithm [47]. The v-rescale thermostat (310 K, temperature coupling constant 0.5) [48] and the *Parrinello-Rahman* barostat (1 atm, pressure coupling constant 2.0) for one trajectory of 0.2 μs per model (total of 1.6 μs) [49,50] were employed.

2.1.2.3. Analysis. The last 100 ns of the production MD trajectories were used for the analysis. The NLS4-, or the NLS5-peptide binding free energies were calculated within a MMPBSA-based scheme (Molecular Mechanics Poisson-Boltzmann Surface Area) from the MD trajectories, along with the Root Mean Square Fluctuations (RMSF) per ligand residue over the same trajectories [51]. These parameters are ideal to characterize the strength of binding between protein-peptides.

2.2. In vitro methods

2.2.1. Cell culture

T47D breast cancer cells, expressing the IPO4 (Importin 4) gene, and the IPO5 (Importin 5) gene (https://maayanlab.cloud/archs4/gene/) were purchased from DSMZ (Braunschweig, Germany), and cultured in RPMI-1640 (Gibco™, Thermo Fisher Scientific) supplemented with 10 % Fetal Bovine Serum (FBS) (Qualified, Gibco™, Thermo Fischer Scientific), at 37 °C and 5 % CO_2.

2.2.2. Preparation of GFP-NLS4 and GFP-NLS5 plasmids

The plasmids encoding the putative Importin 4 and Importin 5 recognition sequences, i.e., NLS4 and NLS5, fused to the EGFP (Enhanced green fluorescent protein) at the C-terminus (EGFP-NLS4 and EGFP-NLS5, respectively) were prepared as follows: Pairs of oligonucleotides encoding the heptapeptide LPPRSGP (NLS4) or the pentapeptide KPKLV (NLS5) were synthesized and annealed *in vitro*. The 5′ end of each oligonucleotide was designed to create a single-stranded end allowing the directional cloning of the annealed oligonucleotides in vectors digested with the *Xho*I and *Bam*HI restriction endonucleases. Specifically, the sequences of the two NLS4-encoding oligonucleotides, i.e., FC-NLS4 and NLS4-RC, were: FC-NLS4: 5′- tcgaGCTTTGCCACCTAGAAGCGGACCAG- 3′, and NLS4-RC: 5′ – gatcCTGGTCCGCTTCTAGGTGGCAAAGC – 3′, while those of the NLS5-encoding oligonucleotides, i.e., NLS5-F and rNLS5, were: NLS5-F: 5′- tcgaGCTAAGCCTAAACTGGTGG – 3′ and rNLS5: 5′ - gatcCCACCAGTTTAGGCTTAGC – 3′. Note that the extraneous restriction enzyme overhang sequences are in small letters. The annealed oligonucleotides carrying *Xho*I and *Bam*HI overhangs were cloned into *Xho*I/*Bam*HI - digested pEGFP-C1 vector (Clontech, TaKaRa Bio Inc. USA) to yield plasmids pEGFP-C1-NLS4 and pEGFP-C1-NLS5, i.e., plasmids expressing the putative NLS4 or NLS5 oligopeptides fused to the carboxy-terminus of EGFP. All plasmids were verified by sequence analysis.

2.2.3. Cell transfection for GFP-NLS and IPOs silencing

Cells were seeded at an initial density of 35×10^3 cells/chamber in an 8-chamber slide with 250 μl medium and incubated for 24 h. The specific plasmids pEGFP-C1-NLS4 and pEGFP-C1-NLS5 that respectively express EGFP fusions with specific NLS4, or NLS5 sequences, or plasmid pEGFP-C1 expressing EGFP alone (control) were co-transfected using Attractene Transfection Reagent (QIAGEN, Hilden, Germany), with specific siRNAs (0.14 μg siRNA, 0.10 μg plasmid and 0.60 μl Attractene Transfection Reagent/ 10^4 cells) for IPO4 (AM16708, ID: 109561), IPO5 (AM16708, ID: 106742), or scrambled siRNAs (AM16708, ID: 149158) (Thermo Fischer Scientific, Waltham, MA USA). After 24 h, fresh medium was added and 24 h later the cells were collected and analyzed with real time PCR or fixed with 4 % paraformaldehyde.

2.2.4. RNA isolation and real time PCR

T47D cells were collected and total gene expression of importin 4 and importin 5 was measured by real-time quantitative PCR (real-time qPCR) in order to evaluate transfection efficiency. Total cell mRNA was isolated using the RNA isolation Kit (Nucleospin, Macherey-Nagel, DE), cDNA was synthesized using the PrimeScript™ RT Kit (TaKaRa Bio Inc, USA) and real time PCR was performed using the KAPA SYBR FAST qPCR Master Mix (Kapa Biosystems, Inc. Wilmington, MA, USA) as previously described [52]. The following primer pairs (synthesized by Eurofins Genomics, Ebersberg, Germany) were used (5′->3′): IPO4, forward ACGGAACAGCTCCAGATCGT, reverse AGCAAAAGCCCCATCTCTCTC, IPO5, forward CTGCTGAAGAGGCTAGACAAATG, reverse TCTGCCGCAATATCACAAACTT and Cyclophilin A, forward ATGGTCAACCCCACCGTGT, reverse TTCTGCTGTCTTTGGAACTTTGTC. In all cases transfection efficiency was around 50 %.

2.2.5. Quantification of nuclear translocation of GFP-NLS5 and GFP-NLS4

Fixed cells were mounted with Vectrashield® (Vector Laboratories, Newark, USA) and observed using inverted confocal scanning microscope (Leica SP8) with a 63x objective lens and optical zoom 2x, with oil immersion, while counterstained with DAPI (blue) to delineate the nuclear space. Image J software (https://imagej.nih.gov/) was used to quantify the fluorescence intensity ratio of GFP in the nucleus and the cytoplasm. The area (nucleus or cytoplasm) in the cell of interest was selected using the polygon selection tool and measurements of different variables were taken. To calculate the corrected total cell fluorescence (CTCF) the following formula was used:

CTCF = Integrated Density – (Area of selected cell X Mean fluorescence of background readings).

For the mean background readings ten measurements from ten different regions next to the cells were taken. The ratio of the fluorescence intensity of the nucleus to cytoplasmic region quantifies the nuclear translocation of GFP. Thirty or more cells per condition were analyzed. GraphPad Prism 8.0.1 (GraphPad Software Inc. San Diego CA) was used for parametric statistical analysis. Data were displayed as mean ± SEM. p values < 0.05 were considered statistically significant.

2.2.6. Detection of protein–protein interactions by proximity ligation assay

In order to detect *in situ* the direct interaction (distance < 40 nm) of GFP-NLS with importin 4 or 5, cells were transfected with the pEGFP-C1-NLS4 or pEGFP-C1-NLS5 plasmids with or without specific si RNA for importin 4 or 5. Cells were fixed with paraformaldehyde, and specific antibodies for GFP (GF28R, Invitrogen, mouse monoclonal antibody) and importins 4 (AA 136–289, rabbit from Antibodies-Online GmbH, Germany) or importin 5 (abx225252, rabbit from Abbexa ltd, USA) were used, for the identification of the two proteins. Subsequently, Duolink® PLA Probes (Duolink® In Situ PLA® Probe anti-Mouse PLUS and Duolink® In Situ PLA® Probe anti-Rabbit MINUS), were used, followed by Duolink® Detection Reagents for Brightfield and Wash Buffer for Brightfield (Wash Buffer A) (all pursued from Sigma-Aldrich, Sweden), following the standard manufacture's Duolink® PLA Brightfield Protocol. Briefly, cells were co-incubated with the primary antibodies (anti-Importin 4 or 5 and anti-GFP), for 1.5 h, at 37 °C in a humidity chamber. Transfected cells without the primary antibodies were used as negative controls. Cells were then stained with secondary antibodies (PLA probes -one PLUS and one MINUS) which bind to the constant regions of the primary antibodies and contain a unique DNA strand, for 1 h, at 37 °C, in a humidity chamber. The DNA probes that hybridize when the proteins interact and make circular DNA were amplified (DNA amplification 100 min in a humidity-saturated chamber at 37 °C) and visualized by HRP-labeled complementary oligonucleotide probes (incubation with HRP-probes, 1 h, at room temperature followed by two 2-min washes and incubation with HRP substrate for 15 min, at room temperature). Finally, cells were counterstained with Duolink® Detection Reagents for Brightfield Nuclear Stain and were mounted with Mounting Medium (Inova Diagnostics,

Inc, San Diego) and observed on an optical microscope (Olympus BX41) using a 100 × objective lens with oil immersion. Brown dots indicate the interacting proteins.

3. Results

3.1. In silico characterization of Importins 4 & 5 NLS

For the detection of ***Importin 4 NLS*** we have used the already published interactions of RPS3A (40S ribosomal protein S3a), HGS (Hepatocyte growth factor-regulated tyrosine kinase substrate), HTT (huntingtin), TCP11L1 (*T*-complex protein 11-like protein 1) and CEBPD (CCAAT/enhancer-binding protein delta) proteins with this importin (Table 1) [53–55]. Whenever possible, the crystal structures of the proteins were retrieved from the PDB database and used. A comparison of the identified 3D conformation of the proteins used here for binding with importin 4 with the recently reported Alpha Fold structures (Supplemental Table 1) [31] accounts for the correct conformation of the 3D prediction we have used here.

Importin 4 interacted with the cargo proteins at amino acids 500–553 (Fig. 1A), while its binding with the small GTPase Ran occurred at amino acids 618–650. Importin binding to the aforementioned proteins (identified with the HEX 8.0.8 program) was performed and the interacting amino acids retrieved and also reported in Table 1. A concrete example of Importin 4 binding to Huntingtin is presented in Fig. 1A, while all interactions are shown in Supplemental Fig. 1. Using the recursive procedure of removing one amino acid at the time, from the N- or the C-terminal of the identified importin interacting peptide (see Material and Methods and Ref. [24] for details) we have identified the minimal sequence responsible for the binding of each protein to Importin 4 (Fig. 1B and Supplemental Fig. 2). Aligning the minimal sequences (Fig. 1C), whose deletion results in a substantial decrease of the binding affinity, as shown in Fig. 1B, resulted in the identification of a peptide sequence (L)PPRS(G/P)P. Its 3D conformation is shown in Fig. 1D. Interestingly, molecular dynamic analysis (Fig. 1E & F), revealed that the presence of glycine in position 6 presents a lower affinity as compared to the presence of proline at this position and that the presence of leucine at position 1 is dispensable, although its omission leads to a slightly less strong binding to Importin 4. However, it seems that the presence of Leucine at position 1, interacts and stabilizes the conformation of the NLS-related aminoacids (especially Proline, Arginine and Serine, at positions 3–5 of the NLS sequence), as found by *in silico* mutagenesis/alanine replacement (Supplemental Fig. 3). Finally, the sequence PPRSPP interaction seems the strongest among the NLS4 proposed sequences, because it exerts the lowest RMSF values (ordered), compared to the longer PPRS(G/P)P ligands (disordered) bound to NLS4 (Fig. 1E), and the lowest binding free energy (Fig. 1F).

The same methodology was followed for the identification of Importin 5 NLS. We have used proteins ACD (Adrenocortical dysplasia protein homolog), GBRAP (Gamma-aminobutyric acid receptor-associated protein), GBRL1 (Gamma-aminobutyric acid receptor-associated protein-like 1), GBRL2 (Gamma-aminobutyric acid receptor-associated protein-like 2), MLP3B (Microtubule-associated proteins 1A/1B light chain 3B), MLP3C (Microtubule-associated proteins 1A/1B light chain 3C), RPL7 (60S ribosomal protein L7) and HTT (Huntingtin), previously reported to interact with this importin (Table 1) [54,56–60]. A comparison with the reported structures from the Alpha Fold database (Supplemental Table 2) [31] reveals the accuracy of the starting structures that our approach is based on.

Importin 5 binds to Ran-GDP at amino acids 636–667, while it binds to the cargo proteins with amino acids 404–454 (Fig. 2A). The binding region of the retained cargo proteins occur at the amino acid regions identified and presented in Table 1. Performing one amino acid sequential deletions from the N- or the C- terminal of each identified regions (protein MLP3C is presented in Fig. 2A and B, while all retained proteins interaction with Importin 5 and the sequential deletion of amino acids of the identified regions are shown in Supplemental Figs. 3 and 4), we have retained the minimal sequences necessary for the interaction with Importin 5. Aligning of these sequences (Fig. 2C) revealed a pentapeptide (KP(K/Y)LV) as a putative importin 5 recognition site. Its 3D structure is shown in Fig. 2D, while molecular dynamics simulation binding on Importin 5 revealed that the presence of tyrosine at position 3 results in the lowest RMSF values (ordered), compared to the presence of lysine at this position (Fig. 2E), and the lowest binding free energy (Fig. 2F). The binding free energy for the KPKLV peptide exerts the largest standard deviation, indicating the highest disorder, or considerable instability of the bound state. Finally, all five amino acids are indispensable for Importin 5 binding, as deletion of even one decreases substantially the binding to importin (not presented).

It is to note that, during molecular simulation studies, no entropic factors were considered for the MMPBSA calculations. However, the relative differences reflect true differences, as the similarity of the peptides should reflect a constant error in all resulting energies. In addition, the presence of the solvent (water) and ions, can strongly affect the NLS4/5-peptide binding free energies calculated along the MD trajectories, so differences are expected between the docking and dynamic (MD) results.

3.2. In vitro validation of LPPRS(G/P)P and KP(K/Y)LV as Importin 4 and 5 recognition sites

As *in silico* results provide an initial prediction, here we have validated our results *in vitro*. For this, we constructed plasmids expressing the enhanced green fluorescent protein (EGFP) with NLS4 or NLS5 fused to the C-terminal part of the protein. We transfected T47D cells, expressing both Importin 4 and 5 and observed EGFP cytoplasmic to nuclear translocation by confocal microscopy. Our data are presented in Figs. 3 and 4, for Importin 4 and 5 respectively. Confocal images clearly show that EGFP localization in the nucleus is greatly enhanced when the EGFP protein has the NLS sequence for importin 4 or 5, while in cells knocked-out for importin 4 or 5 with specific siRNAs, the localization of EGFP fluorescence was predominantly cytoplasmic (Fig. 3A and 4A for representative confocal images and Fig. 3B and 4B for the quantitation of nuclear and cytoplasmic staining).

The direct interaction of EGFP-NLS proteins with their respective importins was also verified by a ligation proximity assay. As shown in Fig. 3C and 4C, brown staining (representing the physical interaction of the two proteins, i.e. Importin 4 or 5 and GFP protein) was observed only in the cells that were transfected with the plasmid containing the specific NLS sequence attached to EGFP (EGFP-NLS) and not in the control-GFP cells. Moreover, the two proteins were mainly localized in the nucleus in the presence of importins.

4. Discussion

Karyopherins, including importins and exportins, are important protein systems for the active transport of cargo proteins through the Nuclear Signal Sequence (NLS) across the nuclear pore complex. Karyopherins acquire protein transport functionality following the binding of a Ran-protein [61]. Importins interact with cargos that include transcription and splicing factors and other significant proteins with a known or presumed nuclear action [62–

Table 1
Proteins interacting with Importins 4 and 5. Table presents the protein short name, their PDB code, the corresponding references or links to the PDB page, the identified amino acid sequences interacting with importins 4 or 5 (*in silico* calculations, see text for details) and the derived Gibbs free energy changes (ΔG) of the interaction.

Importin 4				
Cargo Protein	**PDB Code**	**References**	**Interacting Amino Acids**	**ΔG (kcal/mol)**
RPS3A	6ZXG	[90]	^{46}KTLVTRTQGTKIASDGLKGR^{65}	−1831.76
HGS	4AVX	https://doi.org/10.2210/pdb4AVX/pdb	^{651}QAGPTASPAYSSYQPTPT^{668}	−407.18
HTT	6X9O	[91]	^{199}PQKCRPYLVNLLP^{211}	−2083.02
TCP11L1	4WJ3	https://doi.org/10.2210/pdb4WJ3/pdb	^{186}MMGTLCAP^{193} ^{381}DMHLPSFHLKDVLTT^{395}	−1386.74
CEBPD	1GU4	https://doi.org/10.2210/pdb1GU4/pdb	^{155}PTPPTSPEPPRSSPRQTPAPGPAREK^{180}	−3231.26
Importin 5				
ACD	5UN7	https://doi.org/10.2210/pdb5UN7/pdb	^{223}PSSMLCISENDQLILSSLGPCQRTQGP^{249}	−544.33
GBRAP	7AA8	https://doi.org/10.2210/pdb7AA8/pdb	^{46}KKKYLVPSDLTVGQFYFLIRKRI^{68}	−966.75
GBRL1	6HOI	https://doi.org/10.2210/pdb6HOI/pdb	^{46}KRKYLVPSDLTVGQFYFLIRKRI^{68}	−551.66
GBRL2	4CO7	https://doi.org/10.2210/pdb4CO7/pdb	^{46}KRKYLVPSDITVAQFMWIIRKRI^{68}	−665.54
MLP3B	5V4K	https://doi.org/10.2210/pdb5V4K/pdb	^{7}FKQRRTFEQRVEDVRLIREQHP^{28}	−549.20
MLP3C	3WAM	https://doi.org/10.2210/pdb3WAM/pdb	^{6}KIPSVRPFKQRKSLAIRQ^{23}	−781.69
RPL7	6ZMI	https://doi.org/10.2210/pdb6ZMI/pdb	^{214}SSPRGGMKKK^{223}	−1062.89
HTT	6X9O	[91]	^{1157}DDVAPGPAIKAALPSLTNPPSLSP^{1180}	−598.39

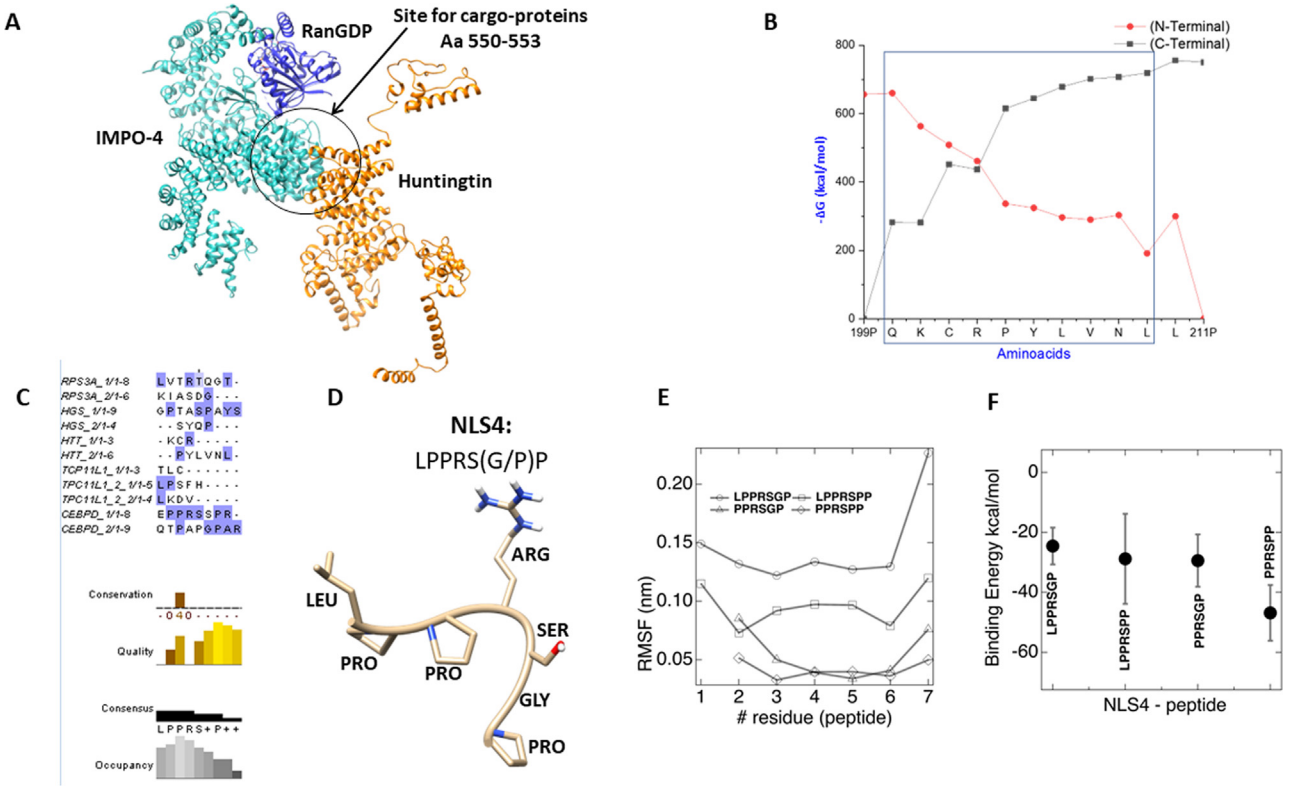

Fig. 1. In silico identification of Importin 4-NLS motif. **(A)** Docking of huntingtin (orange) on the heteroprotein complex RanGDP (blue)-IPO4 (cyan). See Results for further details. Protein interactions were calculated by the HEX 8.0.8 program [36,37], after structure optimizations in the GalaxyWEB server [27–29]. The image was made with the UCSF Chimera program [32]. **(B)** Modification of the association of the huntingtin extracted peptide sequence interacting with Importin 4 (shown in Table 1), calculated with HEX 8.0.8 [36,37] and reported as ΔG (kcal/mol) values. Black curves show the calculated ΔG of C-terminally-truncated sequences while red curves present ΔG values on *N*-terminally truncated sequences. Black box shows the retained peptide sequences, used for the prediction of the NLS sequence. See text for further details. **(C)** Alignment of minimal sequences of amino acids (presented in Fig. 1B and Table 1), with the online tool Jalview [38]. The retrieved consensus sequence is shown at the bottom. **(D)** 3D representation of the minimal consensus sequence, representing the NLS recognition motif for importin 4. **(E)** Interaction of the minimal Importin 4-NLS sequence (LPPRS(G/P)P) with Importin 4 (performed with the HEX 8.0.8 program). As shown, the presence of glycine at position 6 exhibits a higher affinity as compared to the presence of proline at this position. Finally, the presence of leucine at position 1 is dispensable, although its omission leads to a less strong but substantial binding to Importin 4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

64]. It is worth noting that the action of importins is in many cases associated with processes such as chemoresistance and oncogenesis [65–71], and therefore, their pharmacological manipulation might represent a valid approach for the advancement of novel targeted therapies. Nuclear import is a highly selective process that requires adequate receptors to recognize specific import signals.

Nuclear import signals, with a few exceptions, are typically short amino acid sequences and are found in DNA or RNA binding sites of corresponding proteins [65–68]. However, with the exception of Importin α [8,9] and the M9 (transportin) NLS [10–12], little progress has been made in the identification of other importin recognition signals. Recently, we have reported the sequence

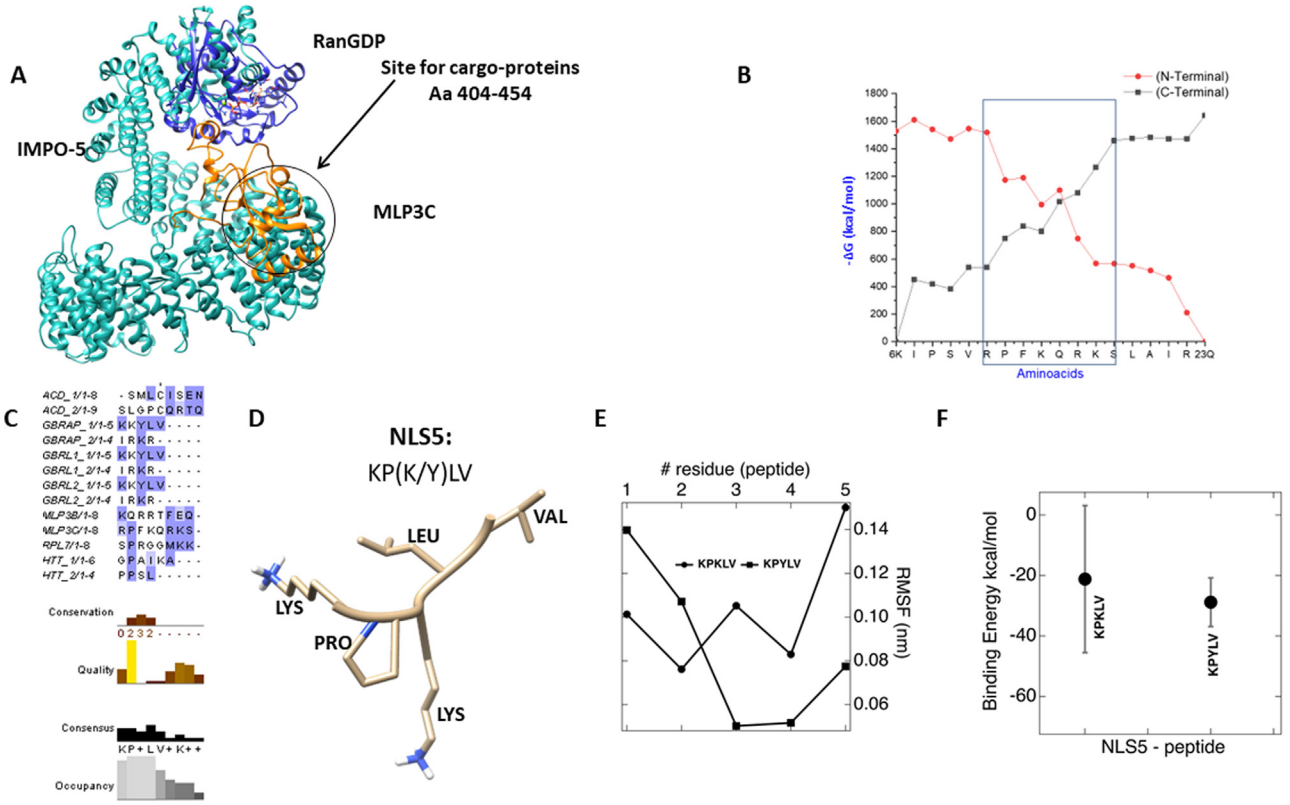

Fig. 2. In silico identification of Importin 5-NLS motif. **(A)** Docking of the microtubule-associated proteins 1A/1B light chain 3C (MLP3C, orange) on the heteroprotein complex RanGDP (blue)-IMPO5 (cyan). See Results for further details. Protein interactions were calculated by the HEX 8.0.8 program [36,37], after structure optimizations in the GalaxyWEB server [27–29]. The image was made with the UCSF Chimera program [32]. **(B)** Modification of the association of the MLP3C extracted peptide sequence interacting with Importin 5 (shown in Table 1), calculated with HEX 8.0.8 [36,37] and reported as ΔG (in kcal/mol) values. Black curves show the calculated ΔG of C-terminally-truncated sequences while red curves present ΔG values on *N*-terminally truncated sequences. Black boxes show the retained peptide sequences, used for the prediction of the NLS sequence. See text for further details. **(C)** Alignment of minimal sequences of amino acids (presented in Fig. 1B and Table 1), with the online tool Jalview [38]. The retrieved consensus sequence is shown at the bottom. **(D)** 3D representation of the minimal consensus sequence, representing the NLS recognition motif for Importin 5. **(E)** Interaction of the minimal Importin 5-NLS sequence (KP(K/Y)LV) with Importin 5 (performed with the HEX program). As shown, a higher affinity is observed when lysine is present at position 3, as compared to tyrosine. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

EKRKI(E/R)(K/L/R/S/T) as a recognition motif for binding to importin 7 [24], a result recently confirmed by another group [72], while, here, we advance the sequences (L)PPRS(G/P)P and KP(K/Y)LV as recognition sites for importin 4 and 5 binding, respectively. Remarkably, the proposed NLSs, although they contain amino acids which are known as classical NLS amino acids or non-classical NLS amino acids [73,74], here we observe that other amino acids such as tyrosine or proline can be found at an NLS motif and their combinations may contribute to a greater variety of NLS sequences for the different importins.

We consider that this discovery might be of importance, as these two karyopherins are implicated in major processes, such as the DNA damage-response pathway [55] and reported related with different cancers [69,75]. In addition, Importin 4 mediates the nuclear import of RPS3A and mediates the nuclear import of human cytomegalovirus UL84 protein, by recognizing a non-classical NLS [76]. UL84 is a multifunctional regulatory protein that is needed for viral DNA replication and its nuclear localization is indispensable for this activity [76].

Our approach (similar to that previously reported for the identification of an importin 7 recognition signal [24]) was validated: (1) by comparing the predicted cargo protein structures with those reported by an independent method [24]. Only minor differences were observed, validating our approach (see Supplemental Tables 1 and 2); (2) by verifying our results *in vitro*. For this, constructs of EGFP protein, containing in its C-terminal the proposed recognition sequences for Importin 4 or 5 were made, and the enhanced nuclear localization of EGFP was found; (3) by showing the EGPF-NLS for importin 4 or 5 enhanced fluorescence in the nucleus compared to the cytoplasm and providing evidence of a physical association of these proteins. These findings were not observed with EGFP lacking the proposed NLS sequences (see Figs. 3 and 4).

It is well documented that the creation of importin complexes with cargos like as RNAs, RNPs, or proteins that are destined for nuclear import, is related to importins' binding with the small GTPase Ran. The interplay between the GDP and GTP-bound form of Ran, bound to importins, determines the direction of importins' movement towards or outwards of the nucleus [2,4,5,77,78]. Importin 4 [7,78–81] and Importin 5 [7,78,81,82] have been reported to interact with Ran, which is indispensable for their action. Although is not the main target of our work, we have simulated Ran-GDP interaction with Importin 4 and 5 (Fig. 1A and 2A), in view of the 3D prediction of the structure of the two importins, bound to Ran-GDP. As expected for protein–protein interactions, the ΔG values in the binding process were very high (ΔG = -278 0.9 kcal/mol and −2670.2 kcal/mol, respectively), suggesting a rather stable interaction. Comparison of the Ran-GDP interacting region with Importin β, Importin 7 [24], and the reported here Importins 4 and 5 shows that Ran-GDP interact with the same amino acid regions (amino acids 72–81 and 132–144) with these

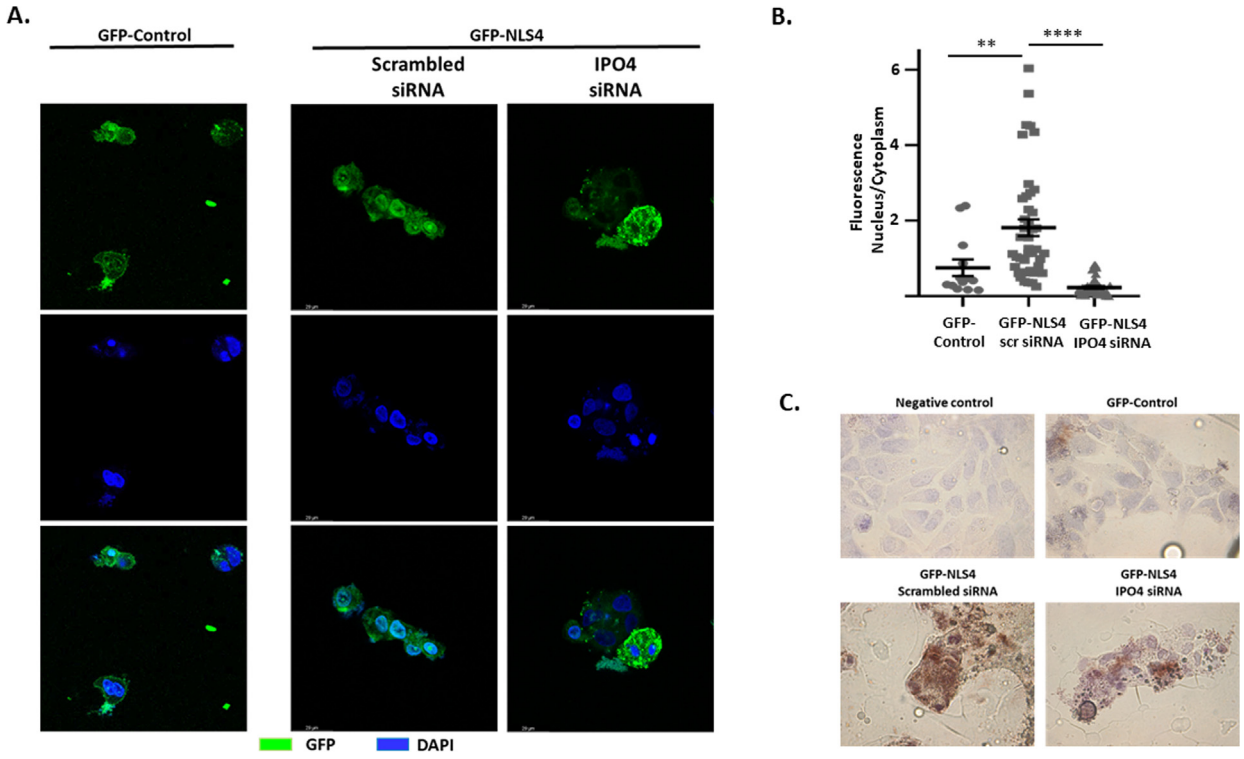

Fig. 3. **(A)** Representative confocal pictures of T47D cells transfected with plasmids expressing either the EGFP-NLS4 fusion (GFP-NLS4), which is recognized by Importin 4, or EGFP alone (GFP-Control), in the presence of either a specific siRNA for Importin 4 (IPO4) or a scrambled siRNA. Nuclei are stained with DAPI (blue). Magnification × 1260. **(B)** Intensity of fluorescence in the cytoplasm and nucleus was quantified (see Material and Methods for details) in at least 30 cells per treatment and is given as the Nucleus/Cytoplasm fluorescence ratio comparing cells with EGFP-NLS4 with cells with EGFP-Control and cells with specific IPO4 siRNA to those with the scrambled siRNA. ** denotes statistical significance $P < 0.01$ and **** $P < 0.0001$. **(C)** Representative images from the proximity ligation assay (See Material and Methods for details). T47D cells were counterstained for the nucleus (with Duolink® Detection Reagents for Brightfield Nuclear Stain) and brown staining was present whenever the EGFP protein and Importin 4 were interacting directly**.** (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Importins interacting with cargo proteins. Table presents the published interacting amino acids of selective proteins with Importins α, 4, 5 and 7 (see text for details).

Protein	Interaction with				References
	IPOα/β	IPO7	IPO4	IPO5	
Huntingtin	–	–	Residues 199–211	Residues 1157–1180	Present work and [54]
HIF1-α	–	PAS domain	Not determined	–	[92]
rPS3a	Not determined	Not determined	Not determined	Not determined	[93]
HPV18 L2	*N*-term basic stretch	–	–	*N*-term basic stretch	[94]
HPV16 L2	*N*-term basic stretch	–	–	*N*-term basic stretch	[95]
CDK5 activator p35	Not determined	Not determined	–	Residues 31–98	[85]
TAFI48	Residues 400–450	–	–	Residues 400–450	[86]
c-Jun	Residues 250–334	Residues 250–334	–	Residues 250–334	[96]
HIV-1 Rev	Residues 35–46	Residues 35–46	–	Residues 35–46	[97]
rPL23a	Residues 32–74	Residues 32–74	–	Residues 32–74	[60]
rPS7	Not determined	Not determined	–	Not determined	[60]
rPL5	Not determined	Not determined	–	Not determined	[60]
rPS3a	Not determined	Not determined	–	Not determined	[60]
H2A	–	Not determined	–	Not determined	[98–99]
H2B	Not determined	Not determined	–	Not determined	[98–99]
H3	–	Not determined	–	Not determined	[98–99]
H4	–	Not determined	–	Not determined	[98–99]

importins and with similar binding affinities. However, the stability of the proposed complex should be further tested experimentally.

Importin 4 NLS, which is reported here, consists mainly of non-polar hydrophobic amino acids such as leucine and proline, positively charged basic amino acids such as arginine and neutrally charged polar amino acids such as serine. This distribution of amino acids indicates that the NLS binding site of Importin 4 recognizes hydrophobic regions of the Importin 4 (amino acids 500 to 553), partially charged. Importin 5 NLS consists mainly of polar amino acids such as lysine, proline and tyrosine and non-polar hydrophobic amino acids such as leucine and aliphatic non-polar amino acids such as valine. This distribution of amino acids indicates that the IPO5-NLS binding site of Importin 5 recognizes polar sequences partially charged (amino acids 404 to 454), as suggested from our *in silico* data.

There were previous attempts to identify NLS sequences for Importins 4 and 5 (reviewed in Table 4 of Reference [7]). The

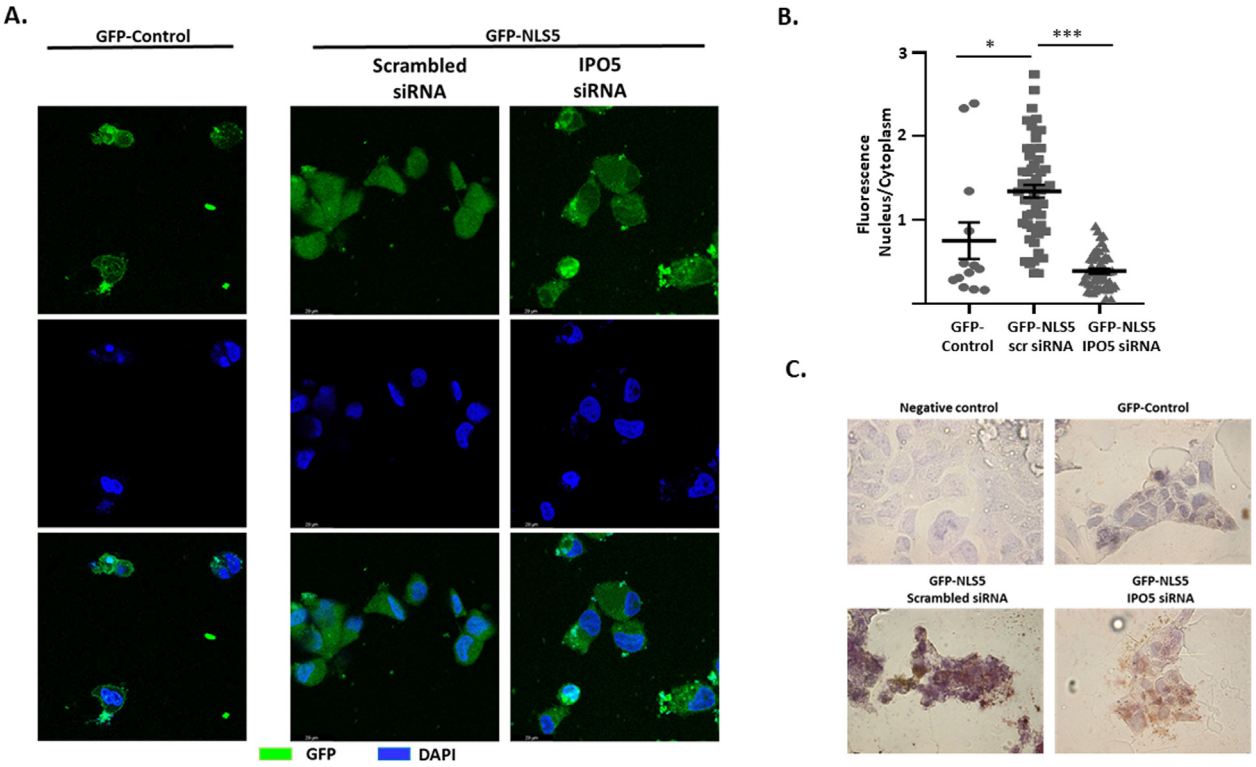

Fig. 4. **(A)** Representative confocal pictures of T47D cells transfected with plasmids expressing either the EGFP-NLS5 fusion (GFP-NLS5), which is recognized by Importin 5, or EGFP alone (GFP-Control), in the presence of either a specific siRNA for Importin 5 (IPO5) or a scrambled siRNA. Nuclei are stained with DAPI (blue). Magnification × 1260. **(B)** Intensity of fluorescence in the cytoplasm and nucleus was quantified (see Material and Methods for details) in at least 30 cells per treatment and is given as the Nucleus/Cytoplasm fluorescence ratio comparing cells with EGFP-NLS5 with cells with EGFP control and cells with specific IPO5 siRNA to those with the scrambled siRNA. * denotes statistical significance $P < 0.05$ and *** $P < 0.001$. **(C)** Representative images from the proximity ligation assay (See Material and Methods for details). T47D cells were counterstained for the nucleus (with Duolink® Detection Reagents for Brightfield Nuclear Stain) and brown staining was present whenever the EGFP protein and Importin 5 were interacting directly**.** (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

authors reported vitamin D receptor (VDR) and TP2 [79] interaction with Importin 4 (aa 4–232 of VDR and aa 87–95 of TP2, see Supplemental Figure 6 for an alignment of the proposed NLS4 sequence with $TP2_{87\text{-}95}$) and a number of proteins interacting with Importin 5. Specifically, residues 439–527 of Rag-2 [83], residues 149–243 of apolipoprotein A-I [84], residues 31–98 of CDK5 activator p35 [85], residues 400–450 of TAFI48 [86], residues 250–331 of c-Jun [87], residues 35–46 of HIV-1 Rev [88] and residues 32–74 of rPL23a [60] have been reported to bind with Importin 5. Scanning the sequences of the proposed NLS here, we found that the proposed here NLSs for Importins 4 and 5 are found in the regions that have previously been experimentally confirmed to be associated with the corresponding importins (Supplemental Table 3).

Miyauchi et al [80], in a very detailed study, showed that the vitamin D Receptor (VDR) is imported to the nucleus, through Importin 4, both in an unliganded and liganded form. The authors reported that an *N*-terminal truncated form of VDR (Δ4-232) is not transported (in an unliganded form) by Importin 4. Blast analysis of VDR revealed two occurrences of our proposed IPO4 NLS in VDR sequence (aa 32–38 and 414–420, the latter being in the ligand binding domain of the molecule), supporting the reported data. Concerning the LBD-related IPO4-NLS, binding of 1, 25 Vitamin D to its receptor caused stereochemical changes in the amino acid region 414–420 leading to effective binding to importin 4, with a calculated ΔG of −812.4 kcal/mol for liganded receptor compared with ΔG of −566.5 kcal/mol for unliganded receptor. (See Supplemental Figure 5), suggesting that only after ligand binding this NLS sequence is functional (supporting data from Miyauchi et al) [80].

From data presented here, and previous literature [1,7,13,24,89], it becomes evident that nuclear import through IPO α, 7, 4 and 5 (and perhaps other karyopherins) is a specific but redundant mechanism. In Table 2 and Supplemental Table 4, we present the currently available knowledge of single or multiple importins associations with known proteins. It becomes evident that this mechanism of multiple nuclear signals in cargo proteins leads to the control of their nuclear displacement. Below, we discuss some concrete examples: Based on *in silico* experiments, and existing literature [54], huntingtin appears to interact with both Importins 4 and 5 via amino acids ^{199}PQKCRPYLVNLLP^{211} and ^{1157}-DDVAPGPAIKAALPSLTNPPSLSP^{1180}, respectively. This can be explained as a kind of control of the action of the protein in the cell nucleus depending on the cell stimuli and the needs of the cell. c-Jun binds to transportin, importins β, 5, 7, 9, and 13 in amino acid region ^{250}PPAAPPGGRGHSHRDRIHYQADVRLEATEEIYLTPVQRPPDAAEPTSAFLPPTESRMSVSSDPDPAAYPSTAGRPHPSISEEEE^{334} [87]. Finally, protein rPL23a was reported to bind to transportin, and importins β, 5 and 7 in amino acid region ^{32}HSHKKKKIRTSPTFRRPKTLRLRRQPKYPRKSAPRRNKLDHY^{74} [60]. Comparing these sequences for the presence of IPOα, IPO4, IPO5 and IPO7 NLS sequences (Table 2 and Supplemental Table 4) the different binding regions of the different importins are confirmed, thus leading to better control of the nuclear displacement of the respective cargo proteins.

In conclusion, findings of the present work identify recognition motifs of Importin 4 (LPPRS(G/P)P) and Importin 5 (KP(K/Y)LV) on cargo proteins, important for their nuclear transfer and subsequent action. The interaction of several significant proteins that control

cell fate with these importins might represent an alternative approach in the pharmaceutical control of different protein actions and subsequently, pathophysiological states.

CRediT authorship contribution statement

Athanasios A. Panagiotopoulos: Methodology, Investigation. **Konstantina Kalyvianaki:** Methodology, Investigation. **Paraskevi K. Tsodoulou:** Investigation. **Maria N. Darivianaki:** Investigation. **Dimitris Dellis:** Investigation. **George Notas:** Validation, Writing – review & editing. **Vangelis Daskalakis:** Investigation, Validation. **Panayiotis A. Theodoropoulos:** Validation, Writing – review & editing. **Christos A. Panagiotidis:** Validation. **Elias Castanas:** Conceptualization, Methodology, Writing – review & editing. **Marilena Kampa:** Conceptualization, Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was partially supported by Greece and the European Union (European Social Fund-ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the project "Strengthening Human Resources Research Potential via Doctorate Research" (MIS-5000432), implemented by the State Scholarships Foundation (IKY)» to AAP (PhD scholarship) and a Hellenic Foundation for Research and Innovation (H.F.R.I.) Grant to MK. We further acknowledge the computational time granted from the National Infrastructures for Research and Technology S.A. (GRNET S.A.) in the National HPC facility ARIS, under project ID: pa180203 (Insdock). The authors further acknowledge the valuable assistance in microscopy of Ms Ourania Kolliniati and Dr Hara Polioudaki.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2022.10.015.

References

[1] Cagatay T, Chook YM. Karyopherins in cancer. Curr Opin Cell Biol 2018;52:30–42.
[2] Görlich D, Mattaj IW. Nucleocytoplasmic transport. Science 1996;271:1513–9.
[3] Moore MS, Blobel G. A G protein involved in nucleocytoplasmic transport: the role of Ran. Trends Biochem Sci 1994;19:211–6.
[4] Guttler T, Gorlich D. Ran-dependent nuclear export mediators: a structural perspective. EMBO J 2011;30:3457–74.
[5] Lonhienne TG, Forwood JK, Marfori M, Robin G, Kobe B, Carroll BJ. Importin-beta is a GDP-to-GTP exchange factor of Ran: implications for the mechanism of nuclear import. J Biol Chem 2009;284:22549–58.
[6] Soniat M, Chook YM. Nuclear localization signals for four distinct karyopherin-beta nuclear import systems. Biochem J 2015;468:353–62.
[7] Chook YM, Suel KE. Nuclear import by karyopherin-betas: recognition and inhibition. Biochim Biophys Acta 1813;2011:1593–606.
[8] Kalderon D, Richardson WD, Markham AF, Smith AE. Sequence requirements for nuclear location of simian virus 40 large-T antigen. Nature 1984;311:33–8.
[9] Kalderon D, Roberts BL, Richardson WD, Smith AE. A short amino acid sequence able to specify nuclear location. Cell 1984;39:499–509.
[10] Iijima M, Suzuki M, Tanabe A, Nishimura A, Yamada M. Two motifs essential for nuclear import of the hnRNP A1 nucleocytoplasmic shuttling sequence M9 core. FEBS Lett 2006;580:1365–70.
[11] Izaurralde E, Jarmolowski A, Beisel C, Mattaj IW, Dreyfuss G, Fischer U. A role for the M9 transport signal of hnRNP A1 in mRNA nuclear export. J Cell Biol 1997;137:27–35.
[12] Bogerd HP, Benson RE, Truant R, Herold A, Phingbodhipakkiya M, Cullen BR. Definition of a consensus transportin-specific nucleocytoplasmic transport signal. J Biol Chem 1999;274:9771–7.
[13] Marfori M, Mynott A, Ellis JJ, Mehdi AM, Saunders NF, Curmi PM, et al. Molecular basis for specificity of nuclear import and prediction of nuclear localization. Biochim Biophys Acta 1813;2011:1562–77.
[14] Natalia F, Celso C. Mechanisms and signals for the nuclear import of proteins. Curr Genomics 2009;10:550–7.
[15] Miyamoto Y, Yamada K, Yoneda Y. Importin α: a key molecule in nuclear transport and non-transport functions. J Biochem 2016;160:69–75.
[16] Goldfarb DS, Corbett AH, Mason DA, Harreman MT, Adam SA. Importin α: a multipurpose nuclear-transport receptor. Trends Cell Biol 2004;14:505–14.
[17] Robbins J, Dilworth SM, Laskey RA, Dingwall C. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: Identification of a class of bipartite nuclear targeting sequence. Cell 1991;64:615–23.
[18] Fontes MRM, Teh T, Kobe B. Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-α11Edited by K. Nagai. J Mol Biol 2000;297:1183–94.
[19] Adam SA. Transport pathways of macromolecules between the nucleus and the cytoplasm. Curr Opin Cell Biol 1999;11:402–6.
[20] Terry LJ, Shows EB, Wente SR. Crossing the nuclear envelope: hierarchical regulation of nucleocytoplasmic transport. Science 2007;318:1412–6.
[21] Gajewska KA, Lescesen H, Ramialison M, Wagstaff KM, Jans DA. Nuclear transporter Importin-13 plays a key role in the oxidative stress transcriptional response. Nat Commun 2021;12:5904.
[22] Grabarczyk P, Delin M, Roginska D, Schulig L, Forkel H, Depke M, et al. Nuclear import of BCL11B is mediated by a classical nuclear localization signal and not the Kruppel-like zinc fingers. J Cell Sci 2021;134.
[23] Brandi V, Di Lella V, Marino M, Ascenzi P, Polticelli F. A comprehensive in silico analysis of huntingtin and its interactome. J Biomol Struct Dyn 2018;36:3155–71.
[24] Panagiotopoulos AA, Polioudaki C, Ntallis SG, Dellis D, Notas G, Panagiotidis CA, et al. The sequence [EKRKI(E/R)(K/L/R/S/T)] is a nuclear localization signal for importin 7 binding (NLS7). Biochim Biophys Acta Gen Subj 2021;1865:129851.
[25] Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 2006;22:195–201.
[26] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat T, Weissig H, et al. The protein data bank. Nucleic Acids Res 2000;28:235–42.
[27] Shin W, Lee G, Heo L, Lee H, Seok C. Prediction of Protein Structure and Interaction by GALAXY protein modeling programs. Bio Design 2014;2:1–11.
[28] Ko J, Park H, Heo L, Seok C. GalaxyWEB server for protein structure prediction and refinement. Nucleic Acids Res 2012;40:W294–7.
[29] Lee GR, Heo L, Seok C. Effective protein model structure refinement by loop modeling and overall relaxation. Proteins 2016;84(Suppl 1):293–301.
[30] Heo L, Park H, Seok C. GalaxyRefine: Protein structure refinement driven by side-chain repacking. Nucleic Acids Res 2013;41:W384–8.
[31] Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, et al. Improved protein structure prediction using potentials from deep learning. Nature 2020;577:706–10.
[32] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605–12.
[33] Baek M, Shin WH, Chung HW, Seok C. GalaxyDock BP2 score: a hybrid scoring function for accurate protein-ligand docking. J Comput Aided Mol Des 2017;31:653–66.
[34] Shin WH, Kim JK, Kim DS, Seok C. GalaxyDock2: protein-ligand docking using beta-complex and global optimization. J Comput Chem 2013;34:2647–56.
[35] Park T, Won J, Baek M, Seok C. GalaxyHeteromer: protein heterodimer structure prediction by template-based and ab initio docking. Nucleic Acids Res 2021;49:W237–41.
[36] Ritchie D, Grudinin S. Spherical polar Fourier assembly of protein complexes with arbitrary point group symmetry. J Appl Cryst 2016;49:158–67.
[37] Ritchie DW. Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2. Proteins 2003;52:98–106.
[38] Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2–a multiple sequence alignment editor and analysis workbench. Bioinformatics 2009;25:1189–91.
[39] Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res 2004;32:W665–7.
[40] Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. Ff14SB: improving the accuracy of protein side chain and backbone parameters from Ff99SB. J Chem Theory Comput 2015;11:3696–713.
[41] Mark P, Nilsson L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. J Phys Chem A 2001;105:9954–60.
[42] Daskalakis V, Papadatos S, Stergiannakos T. The conformational phase space of the photoprotective switch in the major light harvesting complex II. Chem Commun (Camb) 2020;56:11215–8.
[43] Petratos K, Gessmann R, Daskalakis V, Papadovasilaki M, Papanikolau Y, Tsigos I, et al. Structure and dynamics of a thermostable alcohol dehydrogenase from the antarctic psychrophile moraxella sp. TAE123, ACS. Omega 2020;5:14523–34.
[44] Berendsen HJC, Vanderspoel D, Vandrunen R. Gromacs – a message-passing parallel molecular-dynamics implementation. Comput Phys Commun 1995;91:43–56.

[45] Darden T, York D, Pedersen L. Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. J Chem Phys 1993;98:10089–92.
[46] Yeh IC, Berkowitz ML. Ewald summation for systems with slab geometry. J Chem Phys 1999;111:3155–62.
[47] Hess B, Bekker H, Berendsen H, Fraaije J. LINCS: a linear constraint solver for molecular simulations. J Comput Chem 1997;18:1463–72.
[48] Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys 2007;126:014101.
[49] Nose S, Klein ML. Constant pressure molecular-dynamics for molecular-systems. Mol Phys 1983;50:1055–76.
[50] Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. J Appl Phys 1981;52:7182–90.
[51] Miller 3rd BR, McGee Jr TD, Swails JM, Homeyer N, Gohlke H, Roitberg AE. MMPBSA.py: an efficient program for end-state free energy calculations. J Chem Theory Comput 2012;8:3314–21.
[52] Kalyvianaki K, Drosou I, Notas G, Castanas E, Kampa M. Enhanced OXER1 expression is indispensable for human cancer cell migration. Biochem Biophys Res Commun 2021;584:95–100.
[53] Luck K, Kim DK, Lambourne L, Spirohn K, Begg BE, Bian W, et al. A reference map of the human binary protein interactome. Nature 2020;580:402–8.
[54] Haenig C, Atias N, Taylor AK, Mazza A, Schaefer MH, Russ J, et al. Interactome mapping provides a network of neurodegenerative disease proteins and uncovers widespread protein aggregation in affected brains. Cell Rep 2020;32:108050.
[55] Wang J, Sarkar TR, Zhou M, Sharan S, Ritt DA, Veenstra TD, et al. CCAAT/enhancer binding protein delta (C/EBPdelta, CEBPD)-mediated nuclear import of FANCD2 by IPO4 augments cellular response to DNA damage. Proc Natl Acad Sci U S A 2010;107:16131–6.
[56] Chou CW, Tai LR, Kirby R, Lee IF, Lin A. Importin beta3 mediates the nuclear import of human ribosomal protein L7 through its interaction with the multifaceted basic clusters of L7. FEBS Lett 2010;584:4151–6.
[57] O.H. Lee, H. Kim, Q. He, H.J. Baek, D. Yang, L.Y. Chen, J. Liang, H.K. Chae, A. Safari, D. Liu, Z. Songyang, Genome-wide YFP fluorescence complementation screen identifies new regulators for telomere signaling in human cells, Mol Cell Proteomics, 10 (2011) M110 001628.
[58] Behrends C, Sowa ME, Gygi SP, Harper JW. Network organization of the human autophagy system. Nature 2010;466:68–76.
[59] Chao HW, Lai YT, Lu YL, Lin CL, Mai W, Huang YS. NMDAR signaling facilitates the IPO5-mediated nuclear import of CPEB3. Nucleic Acids Res 2012;40:8484–98.
[60] Jakel S, Gorlich D. Importin beta, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. EMBO J 1998;17:4491–502.
[61] Kosyna FK, Depping R. Controlling the gatekeeper: therapeutic targeting of nuclear transport. Cells 2018;7.
[62] Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U. Nuclear export of microRNA precursors. Science 2004;303:95–8.
[63] Kau TR, Way JC, Silver PA. Nuclear transport and cancer: from mechanism to intervention. Nat Rev Cancer 2004;4:106–17.
[64] Tran EJ, King MC, Corbett AH. Macromolecular transport between the nucleus and the cytoplasm: Advances in mechanism and emerging links to disease. Biochim Biophys Acta 1843;2014:2784–95.
[65] Lin KC, Lin MW, Hsu MN, Yu-Chen G, Chao YC, Tuan HY, et al. Graphene oxide sensitizes cancer cells to chemotherapeutics by inducing early autophagy events, promoting nuclear trafficking and necrosis. Theranostics 2018;8:2477–87.
[66] Conforti F, Zhang X, Rao G, De Pas T, Yonemori Y, Rodriguez JA, et al. Therapeutic effects of XPO1 inhibition in thymic epithelial tumors. Cancer Res 2017;77:5614–27.
[67] Conforti F, Wang Y, Rodriguez JA, Alberobello AT, Zhang YW, Giaccone G. Molecular pathways: anticancer activity by inhibition of nucleocytoplasmic shuttling. Clin Cancer Res 2015;21:4508–13.
[68] Saenz-Ponce N, Pillay R, de Long LM, Kashyap T, Argueta C, Landesman Y, et al. Targeting the XPO1-dependent nuclear export of E2F7 reverses anthracycline resistance in head and neck squamous cell carcinomas. Sci Transl Med 2018;10.
[69] Zhang W, Lu Y, Li X, Zhang J, Lin W, Zhang W, et al. IPO5 promotes the proliferation and tumourigenicity of colorectal cancer cells by mediating RASAL2 nuclear transportation. J Exp Clin Cancer Res 2019;38:296.
[70] Zheng M, Tang L, Huang L, Ding H, Liao WT, Zeng MS, et al. Overexpression of karyopherin-2 in epithelial ovarian cancer and correlation with poor prognosis. Obstet Gynecol 2010;116:884–91.
[71] Carden S, van der Watt P, Chi A, Ajayi-Smith A, Hadley K, Leaner VD. A tight balance of Karyopherin beta1 expression is required in cervical cancer cells. BMC Cancer 2018;18:1123.
[72] Garcia-Garcia M, Sanchez-Perales S, Jarabo P, Calvo E, Huyton T, Fu L, et al. Mechanical control of nuclear import by Importin-7 is regulated by its dominant cargo YAP. Nat Commun 2022;13:1174.
[73] Neira JL, Rizzuti B, Abian O, Araujo-Abad S, Velazquez-Campoy A, de Juan Romero C. Human enzyme PADI4 binds to the nuclear carrier importin alpha3. Cells 2022;11.
[74] Chang CW, Counago RM, Williams SJ, Boden M, Kobe B. Distinctive conformation of minor site-specific nuclear localization signals bound to importin-alpha. Traffic 2013;14:1144–54.
[75] Xu X, Zhang X, Xing H, Liu Z, Jia J, Jin C, et al. Importin-4 functions as a driving force in human primary gastric cancer. J Cell Biochem 2019;120:12638–46.
[76] Lischka P, Sorg G, Kann M, Winkler M, Stamminger T. A nonconventional nuclear localization signal within the UL84 protein of human cytomegalovirus mediates nuclear import via the importin alpha/beta pathway. J Virol 2003;77:3734–48.
[77] Dahlberg JE, Lund E, Goodwin EB. Nuclear translation: what is the evidence? RNA 2003;9:1–8.
[78] Nagai M, Yoneda Y. Small GTPase Ran and Ran-binding proteins. Biomol Concepts 2012;3:307–18.
[79] Pradeepa MM, Manjunatha S, Sathish V, Agrawal S, Rao MR. Involvement of importin-4 in the transport of transition protein 2 into the spermatid nucleus. Mol Cell Biol 2008;28:4331–41.
[80] Miyauchi Y, Michigami T, Sakaguchi N, Sekimoto T, Yoneda Y, Pike JW, et al. Importin 4 is responsible for ligand-independent nuclear translocation of vitamin D receptor. J Biol Chem 2005;280:40901–8.
[81] Strom AC, Weis K. Importin-beta-like nuclear transport receptors, Genome Biol, 2 (2001) REVIEWS3008.
[82] Deng T, Engelhardt OG, Thomas B, Akoulitchev AV, Brownlee GG, Fodor E. Role of ran binding protein 5 in nuclear import and assembly of the influenza virus RNA polymerase complex. J Virol 2006;80:11911–9.
[83] Ross AE, Vuica M, Desiderio S. Overlapping signals for protein degradation and nuclear localization define a role for intrinsic RAG-2 nuclear uptake in dividing cells. Mol Cell Biol 2003;23:5308–19.
[84] Kyung Min C, Sun-Shin C, Sung Key J. A novel function of karyopherin beta3 associated with apolipoprotein A-I secretion. Mol Cells 2008;26:291–8.
[85] Fu X, Choi YK, Qu D, Yu Y, Cheung NS, Qi RZ. Identification of nuclear import mechanisms for the neuronal Cdk5 activator. J Biol Chem 2006;281:39014–21.
[86] Dynes JL, Xu S, Bothner S, Lahti JM, Hori RT. The carboxyl-terminus directs TAF (I)48 to the nucleus and nucleolus and associates with multiple nuclear import receptors. J Biochem 2004;135:429–38.
[87] Waldmann I, Walde S, Kehlenbach RH. Nuclear import of c-Jun is mediated by multiple transport receptors. J Biol Chem 2007;282:27685–92.
[88] Freedman ND, Yamamoto KR. Importin 7 and importin alpha/importin beta are nuclear import receptors for the glucocorticoid receptor. Mol Biol Cell 2004;15:2276–86.
[89] Marfori M, Lonhienne TG, Forwood JK, Kobe B. Structural basis of high-affinity nuclear localization signal interactions with importin-alpha. Traffic 2012;13:532–48.
[90] Ameismeier M, Zemp I, van den Heuvel J, Thoms M, Berninghausen O, Kutay U, et al. Structural basis for the final steps of human 40S ribosome maturation. Nature 2020;587:683–7.
[91] Guo Q, Bin H, Cheng J, Seefelder M, Engler T, Pfeifer G, et al. The cryo-electron microscopy structure of huntingtin. Nature 2018;555:117–20.
[92] Chachami G, Paraskeva E, Mingot J-M, Braliou GG, Görlich D, Simos G. Transport of hypoxia-inducible factor HIF-1α into the nucleus involves importins 4 and 7. Biochem Biophys Res Commun 2009;390:235–40.
[93] Jäkel S, Mingot JM, Schwarzmaier P, Hartmann E, Görlich D. Importins fulfil a dual function as nuclear import receptors and cytoplasmic chaperones for exposed basic domains. EMBO J 2002;21:377–86.
[94] Klucevsek K, Daley J, Darshan MS, Bordeaux J, Moroianu J. Nuclear import strategies of high-risk HPV18 L2 minor capsid protein. Virology 2006;352:200–8.
[95] Darshan MS, Lucchi J, Harding E, Moroianu J. The l2 minor capsid protein of human papillomavirus type 16 interacts with a network of nuclear import receptors. J Virol 2004;78:12179–88.
[96] Waldmann I, Wälde S, Kehlenbach RH. Nuclear import of c-Jun is mediated by multiple transport receptors. J Biol Chem 2007;282:27685–92.
[97] Arnold M, Nath A, Hauber J, Kehlenbach RH. Multiple importins function as nuclear transport receptors for the Rev protein of human immunodeficiency virus type 1. J Biol Chem 2006;281:20883–90.
[98] Mühlhäusser P, Müller EC, Otto A, Kutay U. Multiple pathways contribute to nuclear import of core histones. EMBO Rep 2001;2:690–6.
[99] Baake M, Bäuerle M, Doenecke D, Albig W. Core histones and linker histones are imported into the nucleus by different pathways. Eur J Cell Biol 2001;80:669–77.