Mechanism and regulation of dUTPase nucleocytoplasmic transport with an outlook on cell cycle dependent nuclear proteome reconstruction

Ph.D. thesis

Gergely Róna

Biologist, MSc

Eötvös Loránd University, Faculty of Science

Doctorate School in Biology, Structural Biochemistry Ph.D. Program

Head of the Doctorate School in Biology:	**Anna Erdei**, Ph.D., D.Sc., professor
Head of the Structural Biochemistry Ph.D. Program:	**László Nyitray**, Ph.D., D.Sc., professor
Supervisor:	**Beáta G. Vértessy**, Ph.D., D.Sc., professor, scientific advisor

Prepared at the:

Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences

Budapest, 2015

My individual role in the preparation of my Ph.D thesis

The data obtained in my Ph.D thesis is the result of the combined effort of several scientists. I truly believe that science could only be done in a teamwork, and therefore I have written the text in 'first person plural'. This is the only way I could express the gratitude to all the people who have helped me in any way during my career.

Most of the thesis is based on my first or co-first author papers (listed in section 9. Publication list), in which the experiments were designed, carried out, evaluated and written by me. The experiments carried out by our collaborators are indicated in the text, for which I am truthfully very grateful since without their contribution this thesis would not be as it is today. In the thesis I have also listed all of my colleagues who have contributed to this work.

Names of collaborators:	**Contribution to the thesis:**
Dr. Éva Klement and Dr. Katalin F. Medzihradszky	determination of calpain cleavage sites by mass spectrometry
Dr. Zsuzsnna Környei, Máté Neubrandt and Dr. Emília Madarász	video-microscopy experiments
Dr. Mary Marfori and Dr. Bostjan Kobe	crystallization of importin-NLS complexes and solving the crystal structure
Dr. Jonathan Ellis, Dr. Ahmed Mehdi and Dr. Mikael Bodén	*in silico* screenings, Gene Ontology analysis and SNP search
Dr. András Horváth and Imre Zagyva	dissection of fruit flies for immunohistochemistry

Figures and results found in the thesis are based on the following peer-reviewed publications:

Section	**Results found in the following publication**
4.1	- Bozoky Z*, Rona G*, et al., 2011, PLoS One 6(5): p. e19546.*equal contribution
4.2	- Rona, G., et al., 2013, Acta Crystallogr D Biol Crystallogr, 2013. 69(Pt 12): p. 2495-505. - Rona, G., et al., 2014, Cell Cycle, 2014. 13(22): p. 3551-64.
4.3	- Rona, G., et al., 2014, FEBS J, 2014. 281(24): p. 5463-78

Contents

Acknowledgement

I would like to express my deepest appreciation to my supervisor, **Prof. Beáta G. Vértessy**, for the continuous support of my study and research, for her motivation, enthusiasm, and immense knowledge.

I wish to thank **Prof. László Buday** director of the Institute of Enzymology for allowing me to do my Ph.D. research at the Institute and for all his support during my time spent here. I thank him for the motivating discussions and also for all the cell lines provided by him, along with the custom made antibody he designed for my work.

I owe so much to all my collaborators. Without their contribution the present thesis could not have been written. I thank **Dr. Zsuzsnna Környei, Máté Neubrandt** and **Dr. Emília Madarász** for their help with video microcopy. I would also like to thank **Dr. Mary Marfori** and **Dr. Bostjan Kobe** for their work in X-ray crystallization, their guidance have been of great value in this study. I thank **Dr. Jonathan Ellis**, **Dr. Ahmed Mehdi** and **Dr. Mikael Bodén** for their work regarding the bioinformatics. I would also like to offer my regards to **Dr. Éva Klement** and **Dr. Katalin F. Medzihradszky** for their expertise is mass spectrometry. I thank **Imre Zagyva** for dissecting fruit flies for my work and helping me with the never ending struggle with the printers. I am grateful for **Prof. Yvonne Jones** for the cell lines, **Prof. Salvatore Caradonna** and **Prof. J. S. Elce** for the cDNA clones.

I am indebted to **Dr. Judit Tóth** for lending me her expertise and the illuminating views on a number of issues related to the project, along with **Ildikó Scheer** who assisted my in several experiments with great care. I thank **Dr. András Horváth** for the everyday discussions, for all the ideas he gave me when something really challenging came up during my work. I also wish to thank **Máté Borsos** and **Hajnalka Pálinkás**, former undergraduate students, who made great contributions and always followed the ongoing work with extreme enthusiasm which was very motivating. I owe a lot to **all my colleagues** for their aspiring guidance, invaluably constructive criticism and friendly advice throughout the years which had a great effect on forming my thesis, and also in creating a supportive environment and a lively atmosphere in the Lab.

I also would like to offer my regards **to all of those working in the Institute of Enzymology** who supported me in any respect. The people here created a great community that provides everyday professional and emotional support, not to mention all the fun programs organized which made my Ph.D years a great experience.

I thank **Dr. Michele Pagano** for giving me the opportunity to stay in his lab for half a year, at the New York University School of Medicine, where I have gained a lot of great experience.

Abbreviations

293T-HEK	human embryonic kidney 293T cells
3’-end	three prime end of DNA
5’ RACE	5’ rapid amplification of cDNA ends
5’-end	five prime end of DNA
5FdUMP	5-Fluoro-2′-deoxyuridine 5′-monophosphate
5FdUR	5-Fluoro-2′-deoxyuridine
5FU	5-Fluorouracil
ARM	Armadillo repeats
BER	base excision repair
BME	2-mercaptoethanol
bp	base pair
BSA	bovine serum albumin
cDNA	complementary DNA
cNLS	classical nuclear localization signal
Cdk1	cyclin-dependent protein kinase 1
COS7	African green monkey kidney fibroblast-like cell line
C-terminus	carboxyl-terminal end of a protein
D. melanogaster	*Drosophila melanogaster*
D. virilis	*Drosophila virilis*
Da / kDa / MDa	dalton ($1.660538921\times10^{-27}$ kg) / kilo dalton / mega dalton
dA	deoxyadenosine
DAPI	4', 6-diamidino-2-phenylindole
dC	deoxycytidine
DCDT	dCMP deaminase
DCTPP1	dCTP pyrophosphatase 1
dG	deoxyguanosine
DHF	dihydrofolate
DHFR	dihydrofolate reductase
dNDP	deoxyribonucleoside-diphosphates
dNTP	deoxynucleoside triphosphate
DsRed	*Discosoma sp.* red fluorescent protein
dT	thymidine

DTE	dithioerythritol
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
dU	deoxyuridine
dUMP	deoxyuridine monophosphate
dUPNPP	α,β-imido-dUTP
dUTP	deoxyuridine triphosphate
dUTPase	deoxyuridine triphosphate nucleotidohydrolases
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
EGTA	ethyleneglycoltetraacetic acid
FBS	foetal bovine serum
GO	Gene Ontology
HeLa	human cervical adenocarcinoma cell line
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hyper-P	hyperphosphorlyation mimicking mutation
hypo-P	hypophosphorlyation mimicking mutation
IBB	importin-β binding domain
IPTG	isopropyl-β-D-1-thiogalactopyranoside
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionization – Time Of Flight
MCF-7	mammary gland adenocarcinoma cell line
MLS	mitochondrial leader sequence / mictochondrial localization signal
MTHF	5,10-methylene tetrahydrofolate
MTX	methotrexate
NCS	newborn calf serum
NDP	ribonucleoside-diphosphate
NDPK	nucleoside-diphosphate kinase
NE	nuclear envelope
NES	nuclear export signal
NIH-3T3	Swiss mouse embryo derived fibroblast cell line
Ni-NTA	Ni-nitrilotriacetic acid
NK	deoxynucleoside kinase
NLS	nuclear localization signal

NPC	nuclear pore complex
N-terminus	amino-terminal end of a protein
Nup	nucleoporins
OD	optical density
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline with Tween-20
PCR	polymerase chain reaction
PIPES	piperazine-N, N′-bis (2-ethanesulfonic acid)
PMSF	phenylmethylsulfonyl fluoride
PPi	pyrophosphate
RNaseH	ribonuclease H
RNR	ribonucleotide reductase
RTX	raltitrexed (Tomudex)
S. cerevisie	*Saccharomyces cerevisiae*
S. pombe	*Schizosaccharomyces pombe*
S2 cells	Schneider S2 cells derived from a primary culture of late stage (20–24 hours old) *Drosophila melanogaster* embryos
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SHMT	serine hydroxymethyltransferase
SNP	single nucleotide polymorphism
Table A'X'	Table found in appendix (from 1 to 7)
TBS	TRIS Buffered Saline
TBS-T	TRIS Buffered Saline with Tween-20
THF	tetrahydrofolate
TK	thymidine kinase
TMPK	thymidylate kinase
TRIS	2-Amino-2-(hydroxymethyl)-1, 3-propanediol
TRITC	tetramethylrhodamine
TS	thymidylate synthase
TYMK	dTMP kinase
TYMS	thymidylate synthase
WR-Dv-1	*Drosophila virilis* cell line, derived from first instar larval stage
WT	wild type

1. Introduction

My thesis focuses on the nucleocytoplasmic transport processes of the human dUTPase and how these transport processes could be regulated. As a relevant beginning, I would like to provide an overview on nucleocytoplasmic trafficking in general, along with the introduction of how dUTPase function helps cells in maintaining their genomic integrity.

1.1 Nucleocytoplasmic transport processes

The greater cellular complexity of eukaryotes, as compared to prokaryotes, enabled them to evolve into multicellular organisms. This higher order complexity is partly due to the fact that eukaryotes have multiple inner membrane structures that allows them to separate and regulate their physiological processes in a spatial and temporal manner. The nucleus, containing the genetic material along with the nucleoplasm is sequestered by a double membrane, called the nuclear envelope (NE), from the cytoplasm. However, the nucleus is not completely isolated from its surroundings, since there are several processes that require the bidirectional transport of molecules between the two compartments. Eukaryotes have thus successfully uncoupled transcription and translation from each other allowing such regulation strategies to evolve that could not be utilized by prokaryotes [1].

1.1.1 Transport through the nuclear pore complexes

Transport processes between the nucleus and the cytoplasm are mediated through gates in the NE, called the nuclear pore complexes (NPCs). These structures grant selective transport between the nucleus and the cytoplasm while maintaining the distinctive composition of each. NPCs spanning the NE are composed of multiple copies of around 30 types of proteins, called the nucleoporins (Nups), exhibiting an octagonal rotational symmetry which results in an enormous complex of about 120 MDa [2, 3]. Macromolecules larger than ~ 40 kDa require active transport mechanisms to cross the barrier with the help of carrier proteins in a signal and energy dependent way. However, smaller molecules (under 40 kDa or 5 nm in diameter) can freely diffuse through the NPCs [3-6]. Nevertheless many small proteins, which need to be present in the nucleus at high amounts, are imported by active transport mechanisms (like histones [7, 8]) despite the fact that they could freely diffuse into the nucleus. The transport properties of the NPCs are truly astonishing since they can accommodate cargos up to 39 nm in diameter (~ 25 MDa) which enables ribosome subunits and viral capsids to cross the pore without disassembly [9]. The flexibility of NPCs thus allow native complexes to enter or exit

the nucleus without energy consuming unfolding and refolding events. This is an important distinctive feature of the nucleocytoplasmic transport processes compared to import mechanism into the mitochondria or the endoplasmic reticulum where proteins enter through pores with the help of translocases in an unfolded state [10-12]. This also enables multi component complexes to enter the nucleus, with a so called "piggy back" mechanism where only one of the components have a dedicated targeting signal and licenses the whole complex for nuclear entry [13]. NPCs can also mediate 300-1000 transport events per second with a 100 MDa per second mass flow rate [14, 15]. This means an enormous capacity for nucleocytoplasmic shuttling especially when taking into account that the average NPC number is around 2000 per nucleus, based on a few investigated human cell lines [16]. Thus far researchers did not find indication of saturating the import processes by increased cargo concentrations *in vivo*. Therefore rather than by the capacity of NPCs, the overall transport rate seems to be dependent of the cargo:carrier complex formation rate in the cytoplasm and their dissociation in the nucleus. Models describing import dynamics show that import is limited primarily by the concentrations of the transporter molecules and nuclear RanGTP (described in section 1.1.3) [14, 15, 17-19]. Considering the effects of the severe competition among the different signal containing cargos and their carriers is also important. Carrier proteins also have to sample a large number of potential partners within the cell, making non-specific competition with cytoplasmic proteins an important limiting factor in nucleocytoplasmic trafficking [15]. How cargo:carrier protein complexes cross the NPCs is far from being fully understood. There are several proposed mechanisms in the literature. Most of these models agree on that the specific interactions among the carrier proteins and a set of NPC proteins, called the FG-nucleoporins, are of key importance for proper translocation. These FG-nucleoporins contain 20-30 tandem phenylalanine-glycine (FG) repeats, hypothesized to form a molecular brush [20], or a hydrophobic gel [21], in the middle of the NPCs acting as a barrier for none nuclear macromolecules. In these models the specific carrier:Nup interactions are responsible for the highly selective transport process, by overcoming the limitations of the barrier [3].

1.1.2 The classical nuclear import pathway

Facilitated transport is carried out by specific soluble carrier proteins called karyopherins with those responsible for import and export processes called importins and exportins respectively. The most well established superfamily of the carriers is called the β-karyopherins [22]. The structural hallmark of the protein family (counting around about 20 members in humans) are the HEAT repeat folds formed by antiparallel helices (Figure 1) [23]. Either they bind the cargos

directly, and facilitate their transport across the NPCs, or they bind cargos through an adaptor protein, called importin-α (seven of which could be found in humans) [24]. This latter mode of transport is considered to be the classical nuclear import pathway [1]. Though there are several examples for direct importin-β:cargo binding (like cyclin B [25], parathyroid hormone-related protein [26, 27], etc.), most of the known cargos utilize the importin-β:importin-α heterodimer driven classical pathway.

The main determinant of nuclear localization is the presence or absence of the nuclear localization signal (NLS). This linear motif is directly recognized by karyopherins which targets the signal harboring cargo to the nuclear compartment. Classical NLSs contain one (monopartite) or two (bipartite) stretches of basic amino acids. A well-studied monopartite NLS is found on the SV40 large T antigen (^{126}PKKKRRV^{132}) and the prototype for bipartite NLS is found on the *Xenopus* nucleoplasmin (^{155}KRPAATKKAGQAKKKK^{170}) [28, 29].

Importin-α is a banana shaped protein consisting of two functional domains: an N-terminal importin-β binding domain (IBB), and a C-terminal domain containing two NLS binding sites (Figure 1). The IBB domain is responsible for binding importin-β and it also contains residues that resemble a functional NLS which is capable of binding to the NLS binding site in the protein. This autoinhibitory role is important in facilitating the cargo release in the nucleus, after importin-β dissociates from the complex. NLS containing cargos thus also compete with the IBB domain of importin-α which means that the cargos might preferentially bind to importin-α:importin-β heterodimers, rather than binding importin-α alone [30]. The C-terminal domain is composed of ten Armadillo (ARM) repeats, each of them built up by three α-helices. These repeating ARM motifs define the concave surface of the protein containing the two NLS-binding regions, the major and minor site. Bipartite NLSs span both binding sites, while monopartite NLSs usually bind preferentially to the major site (Figure 1). The major site is located closer to the N-terminal end while the minor site is found at the C-terminal part of importin-α. Individual amino acids of the NLS bind to specific pockets in the NLS-binding sites [31]. Conservative tryptophan, asparagine and negatively charged amino acids make up these binding pockets forming salt-bridges and hydrogen bonding interactions with the NLS [32-35]. Consecutive residues from the N-terminal lysine of the bipartite NLS, referred to as P1'- P2', are binding to the minor site; while the C-terminal basic cluster, corresponding to positions P2 - P5, binds to the major site. Therefore the main chain of the peptide runs antiparallel to the direction of the importin-α superhelix [32, 34, 35].

Several types of NLS sequences have been identified in the literature so far, and a strict sequence consensus could not be made but key requirements of NLSs were identified by

structural and thermodynamic studies. Results argue for a loose consensus sequence for the classical monopartite NLS to be: **K**(K/R)X(K/R), and the consensus sequence of the bipartite NLS has been defined as: (K/R)(K/R)$X_{10–12}$**K**$(K/R)_{2-3}$ [36, 37]. Alanine scanning of the cNLSs shed light on the contribution of each position to importin-α binding. Based on energy profiles the P2 positioned lysine (highlighted with bold in the consensus) has the biggest contribution to the binding, while positions P3 and P5 contribute energetically only 2/3rd compared to the P2 position with P4 position being the least important [38]. There have been numerous efforts to categorize the different NLSs based on their amino acid compositions and binding profiles (Table 1). Experimentally validated NLS sequences that do not match the above mentioned consensus were also taken into account [35, 39]. The binding affinity of a cNLS for importin-α correlates with the steady state nuclear accumulation and import rate of the corresponding cNLS cargo [40].

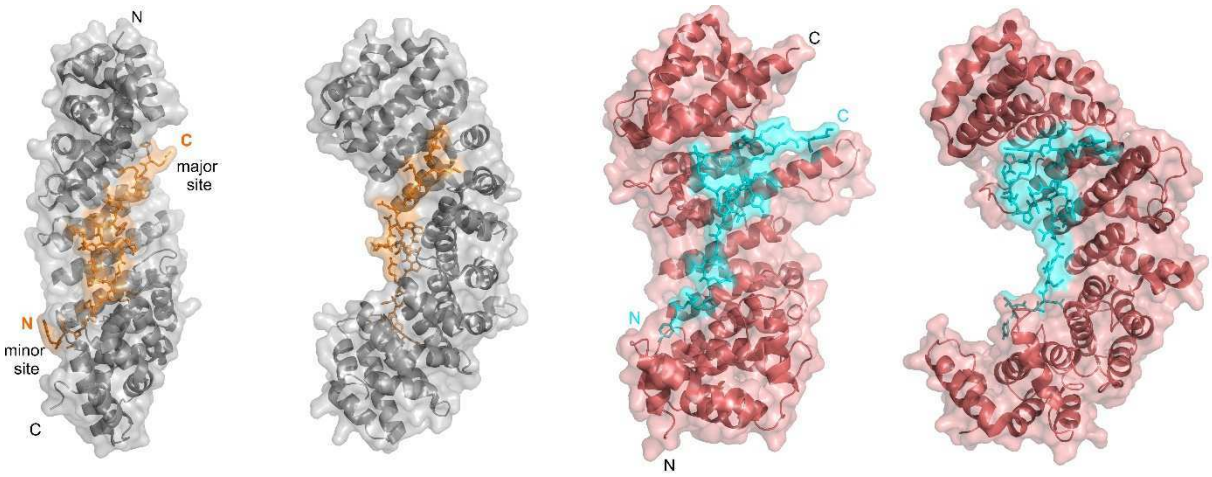


Figure 1. Structures of NLS bound importins

Mouse importin-α lacking its IBB domain (in grey color on the left) and the human N-terminal fragment (1-11 HEAT repeats) of importin-β (in salmon color on the right) are shown in ribbon representation, with the transparent surface superimposed. Importin-α was crystallized in complex with the bipartite Bimax1 NLS (orange), which has optimal binding interactions in both the minor and the major NLS binding grooves, and the linker has also several interactions with importin-α (PDB ID: 3UKW [34]). Importin-β was crystallized with the non-classical NLS^{67-94} of PTHrP (cyan) (PDB ID: 1M5N [26]). N- and C-terminal ends of the NLSs are indicated with orange and cyan color labels. N- and C-terminals of the importins are indicated in black labels.

NLS class		Consensus
Monopartite	class 1	KR(K/R)R, K(K/R)RK
	class 2	(P/R)XXKR(ˆDE)(K/R)
	class 3	KRX(W/F/Y)XXAF
	class 4	(R/P)XXKR(K/R)(ˆDE)
	class 5	LGKR(K/R)(W/F/Y)
Bipartite		$KRX_{10–12}K(KR)(KR)$
		$KRX_{10–12}K(KR)X(K/R)$

Table 1. Types of NLSs specific for importin-α binding

The table summarizes the different NLS classes that bind to distinct binding pockets of importin-α based on a screening of a random peptide library. The two types of classical NLSs (class 1 and 2) bind to the major NLS binding site of importin-α. The archetype example of class 1 is the NLS of the SV40 large T-antigen, while for class 2 the NLS of c-Myc. The two forms of noncanonical NLS (class 3 and 4) bind to the minor NLS binding site of importin-α. (ˆDE) means any amino acid except aspartic acid or glutamic acid. Class 5 type NLSs are plant specific. The optimal sequence determinant of the classical bipartite NLS is also included in the table. $X_{10–12}$ means any 10–12 amino acids. The table was adopted from the work of Kosugi and colleagues [39].

1.1.3 Directionality of nucleocytoplasmic trafficking

Ran, a member of the small Ras GTPase family, provides the energy need of the transport though the hydrolysis of GTP. Like other members of the protein family, Ran has a GTP and a GDP bound state. The nuclear compartment contains 1000 fold more RanGTP compared to the cytoplasm. This great difference in the distribution of the GTP/GDP bound Ran drives the directionality of nucleocytoplasmic trafficking [41]. Based on single molecule experiments, movement through the NPCs is bidirectional and interactions of karyopherins with NPC components do not determine directionality [42]. The direction of the transport however could be reversed by reversing the RanGTP gradient [43]. The asymmetrical distribution of the two Ran forms are maintained by two regulatory proteins. Chromatin bound RanGEF (guanine nucleotide exchange factor) catalyzes the GDP-GTP exchange, while the strictly cytoplasmic RanGAP (GTPase-activating protein), partially bound to Nups on the cytoplasmic surface of NPCs, promotes the GTP-hydrolysis on Ran. Sequestering these two regulatory factors into the nuclear and cytoplasmic compartments results in the steep gradient in RanGTP/RanGDP levels. The direction of transport is defined by the fact that the importin-β:importin-α:cargo trimeric complex disassembles once entering the nucleus because of RanGTP binding. Working just in the opposite way, exportins bind cargos in the nucleus in complex with RanGTP and as soon

as entering the cytoplasm the hydrolysis to GDP triggers their dissociation [36, 37]. In the nucleus RanGTP binds importin-β which then releases the importin-α:cargo complex. The IBB domain of importin-α, containing the autoinhibitory region, along with other protein factors helps cargo:importin-α dissociation [30]. Components have to be recycled for further import cycles, thus importin-β:RanGTP complex and importin-α bound to the CAS:RanGTP complex is exported to the cytoplasm, where they disassemble, and could participate in a new import cycle (Figure 2) [44].

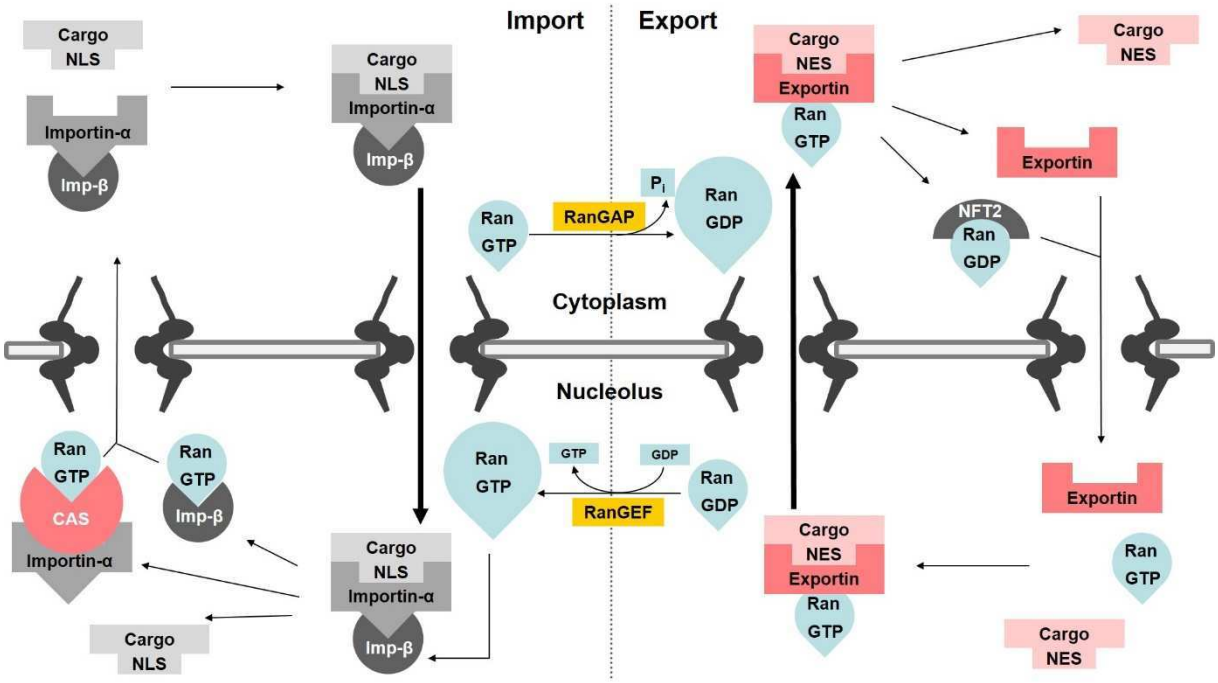


Figure 2. Overview of the classical nuclear import and export mechanism

NLS-bearing cargos are bound by the adaptor protein importin-α which is already bound to importin-β through its IBB domain. After entering the nucleus through the NPCs, binding of RanGTP to importin-β dissociates it from the importin-α:cargo complex. The NLS-containing cargo is then displaced from importin-α partly due to the competition of the NLS-like segment found in the now unbound IBB domain. Importin-α is recycled by its nuclear export factor, CAS, in complex with RanGTP. Importins are released in the cytoplasm for another import cycle after the hydrolysis of RanGTP which is stimulated by RanGAP. The steep RanGTP/RanGDP gradient is achieved by the nuclear RanGEF and the cytoplasmic RanGAP. RanGTP/GDP also cycles among the two compartments: cytoplasmic RanGDP is imported into the nucleus by NTF2 while RanGTP is exported to the cytoplasm bound to transport factors, such as importin-β, CAS or other exportins. Exportins mediate transport to the cytoplasm in complex with RanGTP, recognizing NES-bearing cargos. Cargos are released in the cytoplasm after RanGTP hydrolysis into RanGDP. Key components of the nuclear import pathway are represented in the different shades of grey, while the key components of the export machinery are in shades of red. Components responsible for the directionality of the transport processes are in blue and yellow.

1.1.4 Regulation of nuclear transport

One of the key features of limiting movement in and out of the nuclear space is that these transport processes could be regulated. Nucleocytoplasmic trafficking is regulated on several levels with new mechanisms published every day. A summary of these regulatory mechanisms will be introduced in this section, following the classification according to the work of David A. Jans [45, 46]. Posttranslational modification induced changes are the best described ways of modulating transport processes, where phosphorylation plays a center role (summarized in Figure 3). There are also growing number of examples for the role of methylation and acetylation. The widespread nature of phosphorylation driven regulation links nucleocytoplasmic transport to common signaling pathways like cell cycle regulation, immune response, DNA damage response, etc.

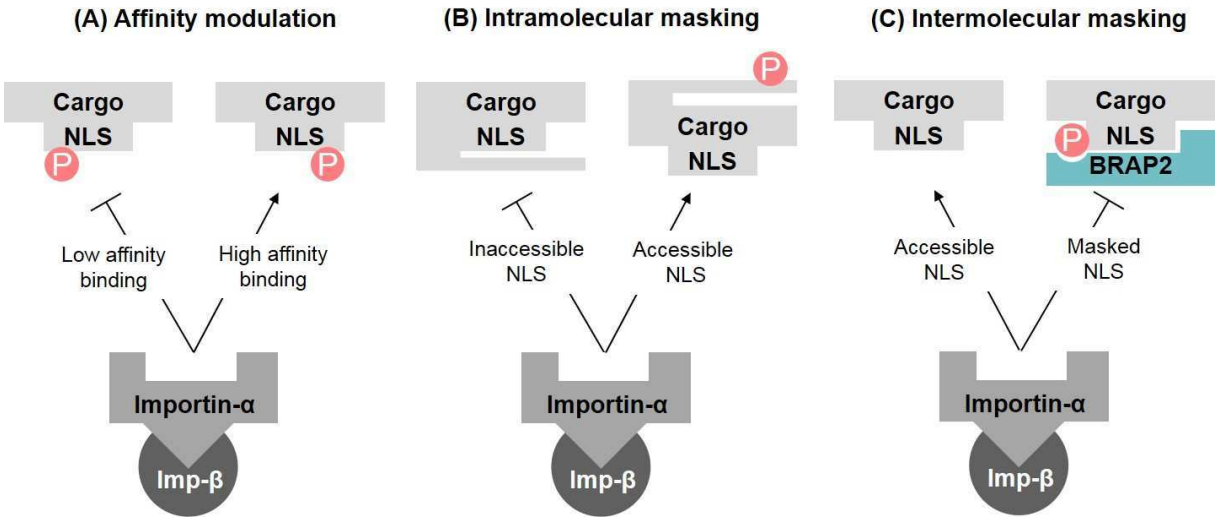


Figure 3. Overview of some possible modes of nuclear transport regulation

Figure summarizes some of the possible mechanisms of NLS regulation depending on phosphorylation based on the work of David A. Jans [45, 46]. **(A)** The affinity between cargos and importins could be modulated by phosphorylation which could either inhibit or enhance their binding. **(B)** In case of intramolecular masking, phosphorylation induces structural changes in the cargo which results in an NLS that is accessible for importins to bind. **(C)** During intermolecular masking a heterologous protein (usually overlapping the NLS) prevents the interaction of the cargos and their carriers (like the phospho-NLS binding BRAP2).

1.1.4.1 Modulating importin:cargo binding affinity

Binding characteristics and import dynamics are mainly depending on NLS:importin interactions. An obvious way to modulate this is to regulate the affinity of these complexes through the posttranslational modulation of the components. One well reported example for a

negative charge masking the NLS was described in yeast for the transcription factor Pho4. Phosphorylation near its NLS disrupts interaction with its dedicated carrier importin [47]. There are several other documented examples in the literature [45, 46, 48-50]. A few of these hits, focusing on the inhibitory phosphorylation of Cdk1, could be found in Table 2 (and references therein), where the introduced negative charge interferes directly with importin binding. Despite the several examples, the exact structural basis of how phosphorylation in the vicinity of the NLSs disrupts importin binding is not yet known to date. There are also several findings where phosphorylation enhances nuclear accumulation, because of increased NLS:importin affinity. A well-established example is CK2 mediated phosphorylation of the SV40 large T-antigen in position S111/112, which enhances affinity to importin-α two fold leading to drastic increase in nuclear import rate [51]. Nevertheless the structural reason behind the enhanced affinity due to phosphorylation remains unclear [52]. Negative charges in the linker region of bipartite cNLSs were shown to have a positive effect on importin-α binding. Generation of peptide inhibitors against the classical nuclear transport pathway resulted in bipartite NLSs which have several glutamic acid or aspartic acid residues in their linker regions. These help in maximizing the possible interactions among the cargo:carrier complex [34, 53]. A systematic position specific screening of the effects of phosphorylation in regions upstream and downstream of the NLS positive core (P2-P5) would still be feasible.

1.1.4.2 Intramolecular NLS masking

Regulation based on intramolecular NLS masking also results in the inhibition of cargo:carrier complex formation. But in this case the effect is due to phosphorylation induced structural changes in the cargo making the NLS inaccessible for importins to bind. The NLS of NF-κB is inaccessible for importin-α to bind in its p105 precursor form. After activation during immune response its C-terminal is subjected to different posttranslational modifications which results in an importin recognizable NLS [54]. In case of the *Xenopus* b-Myb protein, the C-terminal domain simultaneously inhibits DNA binding and NLS function. During embryo development b-Myb is subjected to several modifications resulting in accessible NLSs that leads to the nuclear accumulation of the protein [55].

1.1.4.3 Intermolecular NLS masking

Intermolecular masking could occur if the binding of a heterologous protein (usually overlapping the NLS) prevents the interaction of the cargos and their carriers. A recently described example is the BRCA1-binding protein 2 (BRAP2), an E3 ubiquitin ligase, that binds

phospho-NLSs. The protein was shown to reduce the nuclear accumulation of several viral and endogenous proteins depending on phosphorylation state. Though it could not completely sequester its targets in the cytoplasm, but it definitely fine tunes their localization pattern [56, 57]. DNA or RNA could also be responsible for intermolecular masking. The DNA binding region and the importin-β recognized NLS of the human sex-determining factor SRY are overlapping. Thus DNA binding inhibits importin binding, and also once bound to importin, SRY cannot bind DNA. This mechanism might also facilitate the release of SRY:importin complex, once entering the nucleus [58].

1.1.4.4 Other factors influencing nucleocytoplasmic trafficking

Among several possible other ways to regulate nuclear translocation, cytoplasmic anchoring and the contribution of the microtubular network is definitely worth mentioning.

Not only is the microtubular system essential for some viruses to reach the nucleus, but several non-viral proteins also use them for efficient nuclear translocation. Nocodazole treatment of cells significantly reduce the nuclear accumulation of p53, Rb or the parathyroid hormone related protein [59, 60]. Thus not all NLS containing cargos diffuse freely through the cytoplasm to enter the nucleus, but are transported actively to sites close to the NPCs, enhancing nuclear entry rates.

It is also important to consider the selective binding properties of the different types of importins and their difference is tissue specific expression levels [24, 61]. This also adds another level of possible regulation, since the nuclear entry of a subset of cargos could be influenced by the modulation of the expression of distinct importin-α isoforms.

Several steroid receptors, like the glucocorticoid or estrogen receptors are modulated through cytoplasmic retention. Without their ligands they are sequestered in the cytoplasm by Hsp90 through their ligand binding domains (LBD). Upon ligand binding they release Hsp90 and are imported into the nucleus licensed by their NLSs [62].

1.2 Synthesis of deoxynucleoside triphosphates, the building blocks of DNA

During the course of evolution DNA emerged as the preferred macromolecule for encoding the blueprint of life on Earth. DNA encodes, stores, and transmits the genetic information orchestrating different cellular functions. DNA encoded information is translated into RNA and protein level which then could carry out its structural, regulatory and enzymatic functions that are required for development, differentiation and normal cellular processes. DNA is synthesized by DNA polymerases during replication from its building blocks, the deoxynucleoside

triphosphates (dNTPs), within the cells. For correct replication to occur cells have to maintain a proper set of available dNTPs, which they also have to synthesize. A key step in synthesizing these precursors is the conversion of ribonucleoside-diphosphates (NDPs) into deoxyribonucleoside-diphosphates (dNDPs) by ribonucleotide reductase (RNR). Cellular dNTP levels allosterically regulate the reactions catalyzed by the enzyme in order to maintain proper dNTP levels where dTTP levels play a key role [63]. Nucleoside-diphosphate kinases (NDPKs) then convert dNDPs into dNTPs, which are readily incorporated into DNA by polymerases during replication or DNA repair. As seen in Figure 4, dATP, dCTP, dGTP and dUTP is synthesized this way, but dTTP does not have a ribonucleotide precursor [64]. Figure 4 aims to summarize the main synthesis pathways of mammals but there are significant differences among other organisms.

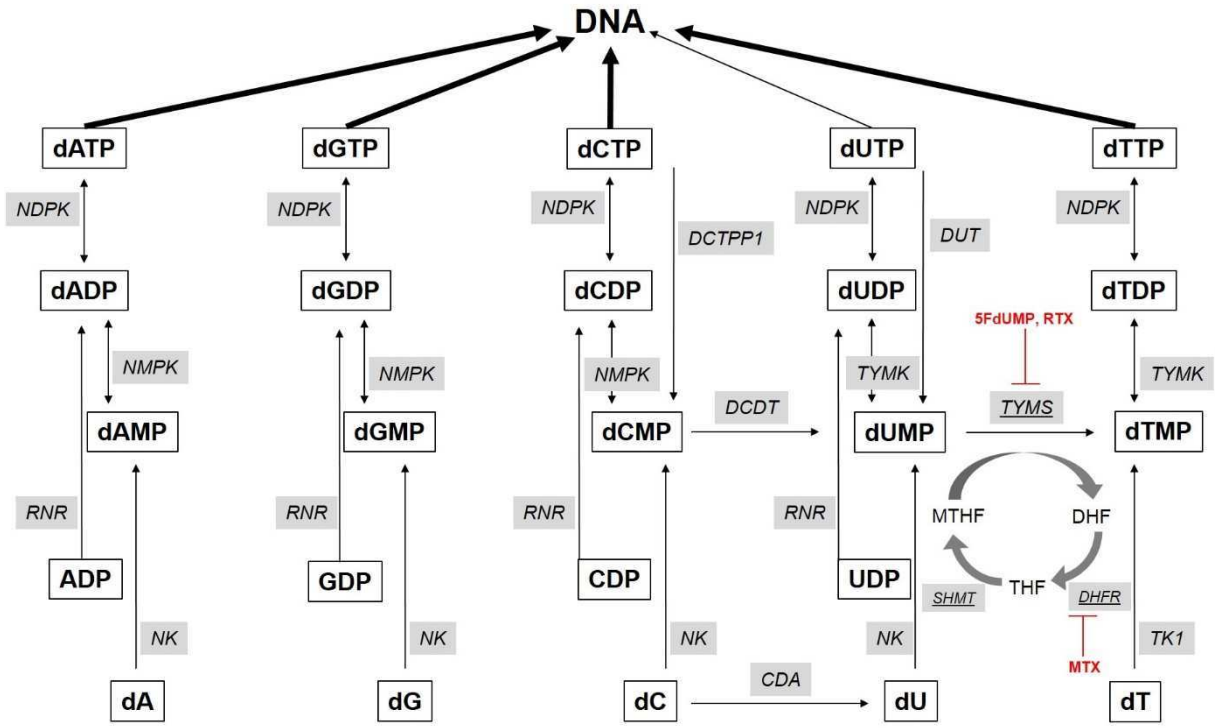


Figure 4. Overview of the synthesis of deoxynucleoside triphosphates (dNTPs) in eukaryotes

The figure summarizes the key steps in dNTP synthesis in mammals focusing on the de novo thymidylate biosynthesis (directly involved enzymes underlined). Inhibitors of the pathway is written in red. 5FdUMP the metabolite of 5FU and 5FdUR, along with raltitrexed (RTX) inhibit TYMS while methotrexate (MTX) inhibits DHFR. Abbreviations are the following: CDA: cytidine deaminase, DCDT: dCMP deaminase, DCTPP1: dCTP pyrophosphatase 1, DUT: dUTPase, DHF: dihydrofolate, DHFR: dihydrofolate reductase, MTHF: 5,10-methylene tetrahydrofolate, NDPK: nucleoside-diphosphate kinase, NK: nucleoside kinase, NMPK: nucleoside monophosphate kinase, RNR: ribonucleotide reductase, SHMT: serine hydroxymethyltransferase, THF: tetrahydrofolate, TK1: thymidine kinase 1, TYMK: dTMP kinase, TYMS: thymidylate synthase.

1.2.1 De novo thymidylate biosynthesis

The proper amount of dTTP is the result of the interplay of several enzymes linked together in the de novo thymidylate biosynthesis pathway (Figure 4). The key intermediate molecule of the pathway is dUMP which is be converted into dTMP by thymidylate synthase (TYMS) further processed into dTDP and dTTP by dTMP kinase (TYMK) and NDPKs respectively. The methyl group transfer of TYMS is irreversible, and is the rate limiting step of the pathway. The methyl group is donated from the cofactor 5,10-methylene tetrahydrofolate (MTHF) while being converted into dihydrofolate (DHF). DHF is regenerated into MTHF by the enzymes dihydrofolate reductase (DHFR) and serine hydroxymethyltransferase (SHMT) through the intermediate tetrahydrofolate (THF) [65, 66].

There are two pathways in eukaryotes that generate dUMP: one through the action of dUTPase (DUT) by hydrolyzing dUTP into dUMP and pyrophosphate plus a proton; and the other route involves dCMP deaminase (DCDT) which produces dUMP from dCMP (see synthesis possibilities in Figure 4). dCMP could either be synthesized from dC, dCDP or dCTP by deoxynucleoside kinases (NKs), through the reversible action of NMPKs and by the action of dCTP pyrophosphatase 1 (DCTPP1) respectively. These reactions are essential for the cells since these pathways feed the substrate, dUMP, for TYMS [66]. Eukaryotes also have a "backup mechanism", called the salvage pathway, which can eventually generate dTTP from dT (if available for the cells) through the action of thymidine kinase (TK). Mammals possess four types of deoxynucleoside kinases (NKs) with overlapping catalytic activity in order to be able to utilize all types of deoxyribonucleosides [67].

1.2.1.1 Localization of enzmyes involved in thymidylate biosynthesis

Despite the fact that dNTPs should be able to freely move among the nuclear and cytoplasmic compartments there are several results that point in the direction that intracellular dNTPs are not evenly distributed between the nucleus, cytoplasm and the mitochondria [65, 66] and also dynamically change during cell cycle and DNA damage [64]. It is possible that the compartmentalization of dNTPs could arise from equilibrium concentrations established by the activity and interactions of different enzymes that are exclusively localized into one of the compartments [66]. Multienzyme channeling of dNTP precursors along with their uneven distribution might be essential for the fidelity and function of the replication machinery [64, 67]. Many of the enzymes needed for proper dNTP synthesis could be found in the cytoplasm but the *de novo* thymidylate biosynthesis pathways are localized to sites of DNA synthesis,

namely to the nucleus and the mitochondria [68]. According to literature the nuclear localization of the *de novo* thymidylate biosynthesis pathway is required for the maintenance of genomic integrity [69], and the nuclear transport of the enzymes involved in the pathway (thymidylate synthase (TYMS), dihydrofolate reductase (DHFR), and serine hydroxymethyltransferase (SHMT1 and SHMT2α)) are sumoylation dependent [70, 71]. These proteins have a cell cycle dependent distribution pattern since their nuclear translocation takes place at the beginning of S-phase, and they remain in the nucleus until M-phase, while being cytoplasmic during G1-phase. This localization timing enables *de novo* thymidylate synthesis during DNA replication and repair [72]. Although the substrates and the products may freely diffuse between the nucleus and the cytoplasm it was suggested that *de novo* thymidylate biosynthesis does not occur in the cytoplasm at rates sufficient to prevent uracil incorporation into DNA [69]. Interestingly several of the enzymes, which are part of the salvage pathways, were also shown to be found in the nuclear compartment. Humans possess two thymidine kinases (TKs) which are responsible for dTMP production from thymidine. TK1 is found in the cytoplasm while TK2 is mitochondrial. However, upon DNA damage, TK1 partly re-localizes to the nucleus [73]. Similarly, the cytoplasmic ribonucleotide reductase was also shown to relocate to the nucleus upon genotoxic stress [74]. In summary, the combined nuclear presence of TK1, RNR, TYMS, DHFR and SHMT1 might provide an effective way to support proper dTTP supply for DNA repair and synthesis. Accompanying these results it was also shown that cytidine deaminase (CDA) could also enter the nucleus through the interaction of its bipartite NLS and importin-α [75] further supporting nuclear dTTP synthesis.

1.2.1.2 Role of dUTPases in nucleotide metabolism

As seen from Figure 4, dUTPase has a dual function within the cells. By eliminating dUTP, the enzyme maintains low dUTP/dTTP ratios and also supplies dUMP as substrate for TYMS to support proper levels of dTTP. The low intracellular levels of dUTP is important for genomic stability since this ensures that polymerases do not incorporate dUTP instead of dTTP in actively dividing cells and thus prevent genomic uracil accumulation [76]. Altogether dUTPase, an essential component of the *de novo* thymidylate biosynthesis pathway, participates in the regulation of the metabolism of two nucleotides (dUTP and dTTP), and consequently keeps dUTP/dTTP ratio at low levels [77]. Imbalance in the dTTP pool might also effect the amount of other dNTPs through the allosteric regulation of several enzymes involved in the dNTP synthesis pathways. Polymerases need a well-defined dNTP ratio to replicate DNA with high fidelity which is also effected by the absolute levels of dNTPs besides their ratios [64, 78]. In

summary the proper function of all enzymes involved in dTTP synthesis, including dUTPase, is required to have optimal dTTP levels in order to maintain the genetic integrity of cells [64, 79].

1.2.1.3 Targeting *de novo* thymidylate biosynthesis with chemotherapy

Lack of proper amount of dNTPs shuts down replication which could eventually lead to cell death. Inhibitors of dNTP biosynthesis pathways have a pronounced effect on actively replicating cells and are therefore commonly used as part of and anti-cancer treatment therapies. Since TYMS is a key component of dTTP synthesis it is targeted by many of the well-known chemotherapeutic agents (5FU, 5FdUR, methotrexate and raltitrexed) inhibiting *de novo* thymidylate biosynthesis. 5FdUMP (the metabolite of 5FU and 5FdUR) and raltitrexed both target TYMS while the antifolate, methotrexate, inhibits DHFR which eventually also abolishes the activity of TYMS (Figure 4) [80-82]. The exact mechanisms how these drugs induce cell death are not yet clear but they drastically deplete dTTP levels and elevate the dUTP/dTTP ratio leading to high frequency dUTP incorporation by polymerases. Ironically it is the repair mechanisms that are partly responsible for the resulting genotoxicity. The repetitive futile attempts of the base excision DNA repair pathway (BER) to eliminate genomic uracil results in multiple DNA strand breaks and eventually chromosome fragmentation [83]. Significance of dUTPase in cancer treatment is demonstrated by the fact that expression level of dUTPase is in correlation with response to TYMS targeted chemotherapy [84]. Silencing dUTPase significantly sensitized cells to TYMS inhibition and lead to dUTP pool expansion demonstrating that the enzyme is a potential target of future cancer treatment [85-88].

1.2.2 Characteristics of dUTPases

dUTPase is an ubiquitous enzyme found in almost all living organism and even several types viruses (*Herpesviridae*, *Poxviridae* és *Retroviridae*) encode the protein [77, 89]. As detailed in section 1.2.1.2 dUTPases catalyse the following reaction:

$$dUTP + H_2O \rightarrow dUMP + PPi + nH^+, \text{ where n depends on the pH of the solution}$$

The enzyme is proved to be essential in several organism: *Escherichia coli* [90], *Mycobacterium tuberculosis* [91, 92], *Mycobacterium smegmatis* [93], *Saccharomyces cerevisiae* [94, 95], *Caenorhabditis elegans* [96, 97], *Trypanosoma brucei* [98], *Arabidopsis thaliana* [99] and is required during the development of *Drosophila melanogaster* [100]. In the listed organisms the effect of dUTPase depletion (based on knock-out or RNA silencing) has been characterized, which in most cases lead to elevated dUTP levels along with DNA fragmentation. These effects

could be rescued in many cases by simultaneously disrupting the BER repair pathway. In vertebrates dUTPase depletion was only achieved thus far by RNA silencing, where cells (HeLa, HT29, SW620, MCF-7) only showed modest phenotypes, but were more sensitive to TYMS inhibition [85-87]. The amount of dUTPase still present in these experiments after silencing along with thymidine in the cell culture medium might be responsible for the viability of the cells, but this will need to be addressed in the future.

Most dUTPases are homotrimeric proteins having three active sites for dUTP hydrolysis built up by the five evolutionary well-conserved protein family specific motifs (Figure 5) [101]. All three sequentially identical subunits contribute to the three active sites which are located at the subunit interfaces. The structure of the human nuclear isoform could be seen on Figure 5B.

1.2.2.1 Human dUTPases

By the usage of alternative promoters along with alternative splicing of the 5' exons, vertebrates encode a nuclear (nDUT) and a mitochondrial dUTPase (mDUT) isoform [102]. The two isoforms only differ in their N-terminals, which encode the mitochondrial leader sequence (MLS) and the NLS, but both contain the conserved core of dUTPases with the five typical motifs (Figure 5). Interestingly the longer mDUT isoform also contains the NLS but is functionally overwritten by the MLS signal since it has a strictly mitochondrial localization. The nDUT also has a very short five amino acid long specific N-terminal segment [103].

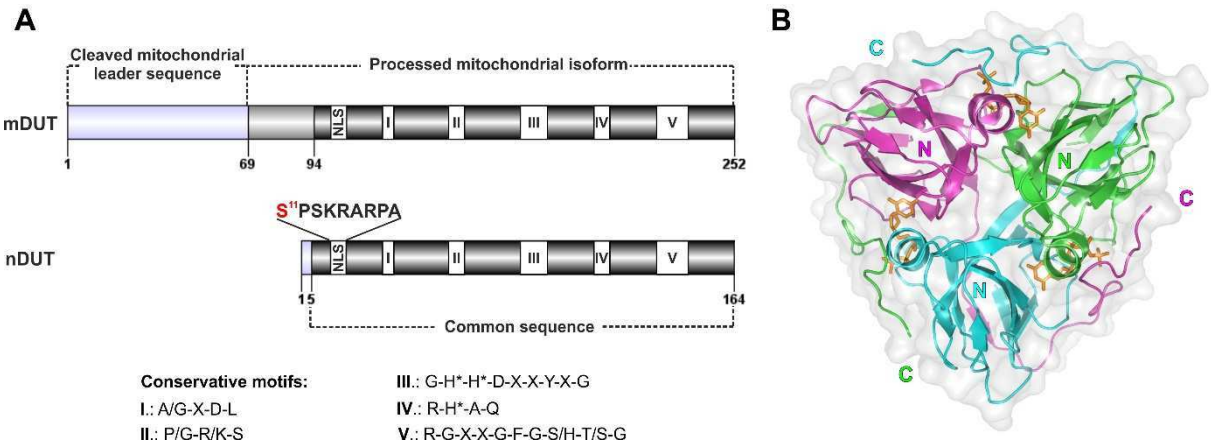


Figure 5. Structural overview of the human dUTPases

(A) Two dUTPase isoform exists in humans arising form alternative promoter and 5' exon usage. One localizes to the mitochondria (mDUT) while the other one is nuclear (nDUT). After mitochondrial translocation the mitochondrial leader sequence (MLS) is cleaved off. The nuclear isoform has a five amino acid long specific N-terminal segment, along with a consensus Cdk1 kinase site (S11 – highlighted in red) that is phosphorylated in actively replicating cells. Though mDUT also contains this position it is not phosphorylated. The dUTPase family specific conserved motifs are shown in white boxes and the consensus sequences are also given. X stands for any amino acid and H* represents hydrophobic amino acids. **(B)** Shows the crystal structure of the human nuclear dUTPase isoform bound to the α,β-imido-dUTP (orange) (PDB ID: 3EHW). Subunits are colored cyan, green, magenta and N- and C-terminal ends of each subunits are also shown. All three subunits contribute to the formation of the active site.

The nuclear isoform was reported to be phosphorylated adjacent to its NLS signal. The NLS of nDUT was experimentally defined to be the following sequence: ^{12}PSKRARPA^{19}. This sequence is essential for nuclear targeting, however total nuclear accumulation was only observed if longer segments of the N-terminal is fused to EGFP [104]. Similar targeting properties were described for the NLS segment of the nuclear isoform of uracil-DNA glycosylase (UNG2), one of the enzymes responsible for the elimination of DNA incorporated uracil residues [105]. The NLS of dUTPase corresponds to a class 2 type of cNLS and thus is predicted to utilize the classical nuclear import pathway relying on importin-α and importin-β. The nuclear dUTPase isoform was also reported to be phosphorylated on its S11 residue in actively dividing cells. Interestingly this position is also present in the mitochondrial isoform, which was not found to be phosphorylated *in vivo*. The amino acid context of this site resembles the consensus phosphorylation site of the cyclin-dependent protein kinase 1 (Cdk1): [S/T*]-P-X-[K/R], where S/T* is the phosphorylated serine or threonine residue [106]. Recombinant

nDUT could also be phosphorylated on the S11 position by Cdk1 *in vitro*. However no effects could be associated with this phosphorylation up until now. A non-phosphorylatable alanine substitution (S11A) did not have any effects on protein activity, oligomerization or localization [104, 107] but it would be also feasible to check the effects of a hyperphosphorlyation mimicking mutation. The expression of nDUT has been associated with cell cycle, linked to S-phase, while mDUT is expressed constitutively regardless of cellular division state [103, 108]. Therefore in non-dividing cells (obtained by serum starvation) the nDUT expression was barely observable, while the expression profile of mDUT did not change [103]. These results are consistent with the hypothesis that dUTPase function is mainly necessary during replication in S-phase where proper amount of the protein has to be present to maintain optimal dUTP/dTTP levels.

1.2.2.2 *Drosophila* dUTPases

Drosophila melanogaster also contains two dUTPase isoforms, which only differ in their N-terminal segments and are generated through the use of alternative splicing (Figure 25). In contrast with human dUTPase isoforms both *Drosophila* dUTPases are expressed in a cell-cycle dependent way [109]. The isoform, having a longer N-terminal segment (Nuc), contains an NLS with the following sequence: ^{10}PAAKKMKID^{18} which also resembles a class 2 type of cNLS that could utilize the classical nuclear import pathway. This isoform localizes to the nuclear compartment, while the other isoform lacking any known targeting signal on its N-terminal (shorter isoform (Cyto)) is cytoplasmic [110, 111]. Both dUTPases have a homotrimeric structure with the conserved dUTPase core containing the five typical dUTPase specific motifs. They also have a longer species specific C-terminal end which is not found in other organisms. The function of this elongated C-terminal end is not yet known [112].

While almost all *Drosophila* species encode homotrimeric dUTPases, *Drosophila virilis* seems to be an exception. Based on genetic annotations the gene encoding dUTPase is present in a repetition of three copies in one open reading frame which resulting in a single polypeptide chain. Therefore this protein achieves similar structure to the homotrimeric variants but as a monomer. Interestingly the dUTPase of *Caenorhabditis elegans* seems to have a similar structural arrangement [101]. The three domains are not sequentially identical having small differences among them, resulting in a structure we call a “pseudo-heterotrimer”. It is important to notice that the *D. virilis* dUTPase only has one NLS, located on the N-terminal of the protein, while homotrimeric nuclear dUTPases harbor three cognate NLSs (Figure 25). Experimental

evidence is still needed to confirm this pseudo-heterotrimeric dUTPase on both mRNA and protein level.

1.3 The calpain system

Calpains are calcium ion activated intracellular cysteine proteases, found in all eukaryotic organism [113]. Unlike many other proteases calpains digest their substrates in a limited fashion usually only modifying their characteristics, and functions. Calpains cleavage occurs at well-defined positions however the sequential environment differs among the substrates. Therefore calpains are considered to be regulatory proteases involved in different signaling pathways effecting many cellular processes. The calpain system plays a role in many distinct cellular functions such as the regulation of the cell cycle, differentiation, cell adhesion along with motility and apoptosis [113]. Humans have two ubiquitous calpains (with 14 members altogether) which have been investigated in detail: the μ- and the m-calpain. These heterodimer proteins are composed of a large ~80 kDa subunit responsible for the catalytic functions and a ~30 kDa regulatory small subunit [114]. *Drosophila melanogaster* has two typical ubiquitous active calpains called calpain A and B with a 94 kDa and 104 kDa molecular weight respectively and also has two atypical forms [115, 116]. Currently there are over a hundred known *in vitro* calpain substrates which were only partly validated as *in vivo* substrates. A variety of proteins could be found among these substrates ranging from cytoskeletal and membrane proteins, kinases, phosphatases and transcription factors [117]. Recently *Drosophila* dUTPases were shown to be *in vivo* substrates of calpains linking dTTP pool maintenance to calpain signaling pathways. Though the physiological role of this interaction has not yet been addressed it would also be feasible to see if it is present in humans.

2. Aims

My colleagues, Dr. Gábor Merényi, Dr. Villő Muha and Dr. Agéla Békési, have made significant contributions in describing the nuclear transport characteristics of *Drosophila melanogaster* dUTPases [109-111]. Vertebrates also have two dUTPase isoforms, but they are strictly localized to compartments where DNA synthesis occurs, namely the nucleus and the mitochondria (as detailed in section 1.2.2.1). We wished to broaden our knowledge on how the nuclear transport of dUTPase is regulated.

2.1 *Drosophila* dUTPase proved to be a calpain substrate [118]. Based on this we wished to check if the human dUTPase is also a calpain substrate or not. If so, this might have on effect of dUTPase distribution. Calpain driven partial proteolysis might limit dUTPase presence leading to altered dNTP homeostasis.

2.2 The nuclear isoform of the human dUTPase was reported to be phosphorylated in actively dividing cells, putatively by Cdk1, in the vicinity of its NLS (on the S11 residue) having no known effects [104, 107]. We wished to check the effects of this phosphorylation on intracellular localization with a glutamic acid substitution (S11E) combined with an isosteric control, a S11Q mutation. Based on NLS composition, dUTPase is expected to utilize the classical nuclear import pathway relying on importin-α. We wished to confirm this interaction through a variety of biophysical methods and systematically check the effects of phosphorylation on this putative interaction. A systematical overview of Cdk1 regulated nuclear transport processes was previously done in yeast, describing several novel proteins having a cell-cycle dependent localization pattern driven by Cdk1. Combining *in silico* screening methods along with high-throughput experimental validation was wished to generate similar findings for the human proteome.

2.3 One of the conserved features of dUTPases is that they are homotrimers. This means that the human nuclear isoform has three NLSs. We were also interested how NLS copy number influences the transport processes of the enzyme. The unique feature of *Drosophila virilis* dUTPases provided us a good model system for this. *D. virilis* encodes a dUTPase that results in a pseudo-heterotrimer (detailed in section 1.2.2.2), where the covalently linked domains are not identical and thus only have one NLS. We wished to compare the nuclear transport properties of this unique, endogenous form, with an engineered construct that harbors three NLSs. These results might shed light on how NLS copy number regulate nuclear transport processes overall.

3. Materials and methods

3.1 Methods used during characterizing the calpain-catalyzed proteolysis of the human dUTPase

3.1.1 Preparation of recombinant proteins

The large subunit of the human m-calpain (encoded in a pET-24b vector) and the small subunit (encoded in a pACpET vector) was a kind gift of Professor J. S. Elce (Queen's University, Kingston, Ontario, Canada). Protein expression was achieved from co-transformed BL21 (DE3) pLysS (Novagen, Merck Millipore, Billerica, MA, USA) cells grown in NYZM medium. The small subunit could be co-purified with the His-tagged large subunit in a complex using Ni-NTA affinity resin (Qiagen, Hilden, Germany).

The human nuclear form of dUTPase (DUT-N) was expressed in BL21 (DE3) pLysS bacteria strain form a pET-15b vector grown in Luria broth medium. In case of both protein expressions transformed cells were induced at A_{600nm}=0.6 with 0.5 mM IPTG for 4 hours at 30 ºC. Cells were harvested and lysed in lysis buffer (50 mM TRIS·HCl, pH=8.0, 300 mM NaCl, 0.5 mM EDTA, 0.1% Triton X-100, 10 mM 2-mercaptoethanol, 1 mM PMSF; 5 mM benzamidin and cOmplete EDTA free protease inhibitor cocktail tablet (Roche, Mannheim, Germany) was only added to the dUTPase lysate) assisted with sonication and cell debris was pelleted by centrifugation at 20.000 x *g* for 30 minutes. Supernatant was applied onto a Ni-NTA column and washed with lysis buffer containing 25 mM imidazole. After washing the column with lysis buffer, dUTPase was finally eluted with the following: 50 mM HEPES, pH=7.5, 30 mM KCl, 500 mM imidazole, 10 mM 2-merchaptoethanol and was dialyzed against the following buffer: 50 mM HEPES, pH=7.5, 100 mM NaCl, 5 mM $MgCl_2$, 10 mM 2-merchaptoethanol. M-calpain was eluted with the following: 50 mM TRIS·HCl pH=7.5, 300 mM NaCl, 5 mM benzamidin; 0,5 mM PMSF, 250 mM imidazole, 10 mM 2-merchaptoethanol and was dialyzed against: 50 mM TRIS·HCl pH=7.5, 150 mM NaCl, 5 mM benzamidin; 0,5 mM PMSF, 1 mM EDTA, 10 mM 2-merchaptoethanol. Protein preparations were checked on SDS-PAGE and showed >95% purity. Protein concentrations were determined according to [119] and correspond to monomers throughout the study, unless otherwise mentioned.

3.1.2 In vitro digestion assays

Digestion assays were performed in 50 mM TRIS·HCl, pH 7.5 buffer also containing 150 mM NaCl; 1 mM EDTA; 1 mM dithiothreitol; 30 μM human dUTPase and 2 μM rat m-calpain. Reaction was started by the addition of m-calpain in the presence or absence of 2 mM $CaCl_2$. For digestion in the presence of substrate analogue (α,β-imido-dUTP) the same reaction buffer was completed with 150 μM α,β-imido-dUTP and 10 mM $MgCl_2$. Reaction was stopped with 3 mM EGTA after 40 minutes, digestion products were either directly submitted to mass spectrometry or separated on SDS-PAGE gels and then analyzed by mass spectrometry.

3.1.3 dUTPase activity measurement

Calpain digested dUTPase was prepared by the following: 30 μM dUTPase and 1.75 μM m-calpain was incubated for 40 minutes at room temperature in digestion buffer (50 mM TRIS·HCl pH=7.5; 150 mM NaCl; 1 mM EDTA; 5 mM benzamidin; 0.5 mM PMSF; 1 mM DTE; 20 mM $CaCl_2$). Reaction was stopped by addition 30 mM EGTA and the mixture was applied onto a Ni-NTA affinity resin. Flowthrough contained the cleaved form of dUTPase while uncleaved fraction along with m-calpain remained on the column.

Enzymatic activity of both intact dUTPase and calpain-cleaved dUTPase was determined in steady-state pH indicator-based assay. Protons are released during the enzymatic activity of dUTPase, which is proportional to the amount of hydrolyzed dUTP:

$$dUTP + H_2O \rightarrow dUMP + PPi + nH^+, \text{ where n depends on the pH of the solution}$$

The pH change was detected based on the absorbance change of phenol red indicator at 559 nm. Reaction buffer contained 1 mM HEPES pH 7.50; 150 mM KCl; 40 μM Phenol Red indicator; 1 mM $MgCl_2$; 40 μM dUTP and 150 nM dUTPase (from the digestion mixtures produced either in the absence or presence of calcium). The reaction was followed in 1 ml reaction volume thermostated cuvette (25^{o}C) on a Jasco FP 777 spectrofluorimeter.

3.1.4 Cleavage site identification by mass spectrometry

Mass spectrometry analysis was carried out by Dr. Éva Klement and Dr. Katalin F. Medzihradszky. The calpain proteolysis cleavage products were desalted on C4 ZipTip (Merck Millipore, Billerica, MA, USA) and analyzed directly by MS on a Bruker Reflex III MALDI-TOF mass spectrometer in linear mode. Sinapinic acid was used as the matrix. External calibration was performed on the Bruker protein calibration standard I (#206355). Masses were determined by averaging seven independent measurements. Alternatively, cleavage products

were separated by SDS-PAGE, then gel bands were subjected to tryptic digestion followed by MS and MS/MS analysis as described in [118].

3.1.5 Cell culture and Western blotting

HeLa cells were purchased from ATCC (Manassas, VA, USA) and were cultured in DMEM/F12 HAM (Sigma, St. Louis, MO, USA) supplemented with 50 µg/ml Penicillin-Streptomycin (Gibco, Life Technologies, Carlsbad, CA, USA) and 10% FBS (Gibco) in a humidified 37 °C incubator with 5% CO_2 atmosphere. Cells were activated with 5 µM ionomycin (dissolved in DMSO) and the medium was complemented with $CaCl_2$ to achieve a final concentration of 2 mM. Control cells were mock treated with DMSO alone. Cells were collected at indicated time points, washed twice with PBS (10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH=7.4; 138 mM NaCl; 2.7 mM KCl) and resuspended in lysis buffer (50 mM TRIS·HCl pH 7.4; 140 mM NaCl; 0,4% NP-40; 2mM DTT; 1 mM EDTA, 1mM PMSF; 5mM benzamidin, 1x complete^{TM} EDTA free protease inhibitors (Roche)). Cell lysis was assisted with sonication. Insoluble fraction was removed by centrifugation (20,000 x *g* x 15 min at 4 °C). Protein concentration was measured with BioRad Protein Assay (BioRad, Hercules, California, USA) to ensure equivalent total protein load per lane. Proteins were resolved under denaturing and reducing conditions on a 15% polyacrylamide gel and transferred to PDVF membrane (Immobilon-P, Merck Millipore, Billerica, MA, USA). Membranes were blocked with 5% nonfat dried milk in TBS-T (50 mM TRIS·HCl, pH=7.4; 140 mM NaCl; 2,7 mM KCl; 0,05% Tween-20), incubated with primary antibody against dUTPase (1:5000) and actin (1:300, Sigma) for 2 hours at room temperature in blocking buffer. After washing the membranes, horseradish peroxidase coupled secondary antibody was applied (Amersham Pharmacia Biotech, GE Healthcare, Buckinghamshire, UK). Immunoreactive bands were visualized by enhanced chemiluminescence reagent (GE Healthcare, Buckinghamshire, UK) and recorded on X-ray films (GE Healthcare).

3.1.6 Calpain activity measurement in HeLa cell extract

To test the active calpain forms in HeLa extract, calpain activity was measured. Cells were collected and washed twice with 50 mM TRIS·HCl, pH 7.50; 150 mM NaCl. To lyse the cells, they were resuspended in 50 mM TRIS·HCl, pH 7.50 buffer containing 150 mM NaCl, 1 mM EDTA, 1 mM PMSF, 5 mM benzamidin and sonicated five times (16 µ; 10 sec). Insoluble fraction was removed by centrifugation (20.000 g × 15 min at 4 °C). Before the activity measurements, 2 mM 2-mercaptoethanol was added and the extract was incubated for 20 min

on ice. Enzyme activity was measured with a modified FRET substrate (DABCYL-TPLKSPPPSPR-EDANS) as in [117], either in the absence or presence of 2 mM $CaCl_2$. Fluorescence intensity was monitored at 640 nm with 320 nm excitation on a Jasco FP 777 spectrofluorimeter.

3.2 Methods used during studying the effect of phosphorylation adjacent to nuclear localization signals on nuclear proteome re-constitution: a case study on dUTPase

3.2.1 Cell culture

293T cells were kindly provided by Professor Yvonne Jones, while NIH3T3, MCF7 and COS7 cells were a kind gift from Professor László Buday. Cells were cultured in a humidified 37 °C incubator with 5% CO_2 atmosphere in DMEM/F12 HAM (Sigma) supplemented with Penicillin–Streptomycin solution (50 µg/ml; Gibco) and 10% FBS (Gibco) except for NIH3T3 cells where 10% NBS (Gibco) was applied.

3.2.2 Plasmid constructs and cloning

dUTPase nuclear isoform was amplified from an expression vector, previously described in [120] and cloned into the XhoI/KpnI restriction sites in-frame with DsRed-Monomer (with oligos dutN1F and dutN1R) in the pDsRed-M-N1 vector (Clontech, Mountain View, CA, USA), yielding the DsR-DUT termed fusion conscruct. DsR-DUT was further cloned into the NdeI/XhoI sites of the vector pET-20b (Novagen) for recombinant protein expression (with oligos dutpETF and dutpETR). The NLS reporter construct was created by fusing β-galactosidase with DsRed-Monomer (termed pGal-DsRed). β-galactosidase was amplified lacking its start codon from vector pCAUG (with oligos galN1F and galN1R), and was cloned it into the KpnI/EcoRI sites of the vector pDsRed-M-N1, thus generating the vector termed pGal-DsRed. Single stranded oligonucleotide pairs, listed in Table A7, encoding different NLS peptides were cloned into the NheI/EcoRI sites of the pGal-DsRed vector after annealing. In addition, the vector pHM830 (Addgene, Cambridge, MA, USA) (AflII/XbaI sites) was also used to generate constructs for the NLSs that showed a strong tendency for aggregation when used in context of the previously described pGal-DsRed construct [121]. Human tumor protein p53 cDNA (NM_000546.2) was purchased from OriGene (Rockville, MD, USA). P53 was fused to DsRed-Monomer, by cloning it into the XhoI/BamHI sites of a modified pEGFP-C1 vector (Clontech) (with primers p53_F and p53_R), where EGFP was replaced by DsRed-Monomer (within the NheI/XhoI sites of the vector). Ubiquitin-activating enzyme E1 (UBA1)

cDNA (NM_003334.2) was purchased from OriGene and the fusion construct was made by cloning it into the KpnI/BamH1 sites of the pDsRed-Monomer-N1 vector (Clontech) (with primers UBA1_F and UBA1_R). Human Uracil-DNA glycosylase 2 (UNG2) cDNA was a generous gift of Professor Salvatore Caradonna and was cloned into the XhoI/KpnI sites of the pDsRed-Monomer-N1 vector (with primers UNG2_F and UNG2_R). Site-directed mutagenesis was performed by the QuickChange method (Stratagene, Santa Clara, CA, USA). Primers used for cloning and mutagenesis were synthesized by Eurofins MWG GmbH (Ebersberg, Germany) and are summarized in Table A7. All constructs were verified by sequencing at Eurofins MWG GmbH.

3.2.3 Fluorescence imaging and analysis of DsRed-tagged constructs

For DNA transfections Lipofectamine^{TM} LTX (Life Technologies, Carlsbad, CA, USA) was used according to the manufacturer's instruction. Briefly, subconfluent cultures of 293T cells grown in 35 mm Petri dishes were incubated with 1-2 µg DNA along with 10 µl LTX reagent in serum-free medium, for 16 hours. Protein transfection was performed according to the manufacturer's protocol using Pro-DeliverIN^{TM} reagent (OZ Biosciences, Marseille, France). In brief, 8-10 µg protein and 15 µl transfection reagent was used to deliver DsRed-tagged proteins into the cells for 14-18 hours. Image analysis to quantify relative subcellular localization was performed from single-cell measurements using ImageJ 1.46j (NIH, Bethesda, USA), where the mean nuclear (F_n) and cytoplasmic (F_c) fluorescence ratio ($F_{n/c}$) was measured within each cell. Statistical analysis of the relative subcellular localization changes was carried out by InStat 3.05 software (GraphPad Software, San Diego California, USA) using the non-parametric Mann-Whitney test. Differences were considered statistically significant at $p<0.05$. Images were either acquired with a confocal microscope (Zeiss LSM 710) or a widefield microscope (Leica DM IL LED Fluo) equipped with a Leica DFC345 FX monochrome camera throughout the study.

3.2.4 Live-cell microscopy and evaluation

Live cell microscopy was carried out by Dr. Zsuzsanna Környei and Máté Neubrandt. Time-lapse recordings were performed on a Zeiss 200M inverted microscope equipped with an AxioCam Mnr camera and controlled by the AxioVision 4.8 software. Cells were cultured in Ibidi dishes and kept at 37°C in a humidified 5% CO_2 atmosphere within custom-made microscope stage incubator (CellMovie Bt., Dept. of Biological Physics, Eötvös Lóránd University, Budapest, Hungary). Phase contrast and fluorescent images were acquired every 5

minutes for at least 24 hours using a 10X magnification objective. After transfection, the cells were washed three times with serum-containing medium. Time-lapse imaging started one hour after changing the medium. Addition of serum resulted in the flattening of the cells and mitogenic serum factors boosted cell proliferation.

Plasmid transfection experiments. Kinetic analysis was carried out in Gepasi 3.30 software (Virginia Bioinformatics Institute, USA) with the help of Dr. Judit Tóth. The kinetic treatment of the imaging data addresses the gross kinetics of nuclear dUTPase accumulation and does not aim at carrying out a detailed analysis of the underlying processes. The quantification of fluorescence in single cells from each frame was performed using ImageJ 1.46j where the mean nuclear (F_n) and cytoplasmic (F_c) fluorescence were measured. Data points represent mean values extracted from 16 cells in triplicates. The time axis was defined relative to the visual observation of cytokinesis i.e. t = 0 at cytokinesis termination. The observed fluorescence intensity increase in the nucleus could be analyzed; as the total fluorescence of the cytoplasmic and nuclear compartments (F_{n+c}) was constant during the time period of the analysis. Single exponential kinetics fitted well to the rising phase of the nuclear accumulation curves in both the WT and the S11Q mutant cell lines. The considerable lag in nuclear fluorescence accumulation in the WT cells was not included in the kinetic analysis due to the lack of information on building a comprehensive kinetic model for the whole trafficking process.

Protein transfection experiments. These image sequences were not subjected to densitometry analyses due to lower intensity of the intracellular fluorescent signal as well as to the higher background (Figure 14). The time elapsed between the onset of cytokinesis and the appearance of fluorescent signal within the nucleus was determined by careful visual observation (Figure 14C). Considerable nuclear accumulation of fluorescent proteins was declared when the fluorescent intensity within the nucleus exceeded that within the cytoplasm. Parallel phase contrast images were used to determine the onset of cell cleavage.

3.2.5 Immunofluorescence and immunoblot analysis

Phosphorylation of the constructs after cellular delivery and its cell cycle dependency were investigated with immunoblot and immunofluorescence analysis respectively. Cells were grown on glass coverslips, fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X-100 in PBS for 10 minutes and blocked with 5% goat serum, 1% BSA in PBS for 2 hours. Samples were incubated with primary antibodies overnight at 4°C in blocking buffer. After extensive washing steps, cells were incubated with secondary antibodies (1:1000) coupled to either Rhodamine Red™-X or Alexa 488 (Molecular Probes, Life Technologies, Carlsbad, CA,

USA), stained with 1 µg/ml DAPI and finally embedded in FluorSave™ Reagent (Calbiochem, Merck Millipore, Billerica, MA, USA). Antibodies to detect the following proteins were used in immunofluorescence: anti-hDUT (1:2000) [86, 122], histone-H3-phospho-S10 (1:2000, Abcam, Cambridge, UK). A polyclonal antibody to the phosphorylated dUTPase NLS (EETPAI/pSer/PSKRAC) (anti-P-hDUT) was custom-produced and affinity-purified by GenScript Corporation (Piscataway, NJ, USA) and was used in 1:100 dilution. Immunoblot analysis was carried out as described in section 3.1.5. Lysis buffer was also supplemented with PhosSTOP™ phophatase inhibitor cocktail tablet (Roche) according to the recommendation of the manufacturer. Antibodies to detect the following proteins were used during western blotting: anti-hDUT (1:5000) [86], anti-P-hDUT (1:200).

3.2.6 Recombinant protein production

dUTPase constructs (including the DsRed-tagged forms) were expressed and purified as detailed in section 3.1.1. N-terminally truncated, His-tagged mouse importin-α2 (NP_034785) lacking 69 N-terminal residues (residues 70-529; importin-α ΔIBB) encoded in a pET-30a vector was obtained from Dr. Bostjan Kobe and purified on Ni-NTA column. Importin-α ΔIBB was expressed in Rosetta BL21 (DE3) pLysS bacteria strain grown in Luria broth medium. Cells were induced at A_{600nm}=0.5 with 0.6 mM IPTG for 4 hours at 25 ºC. Cells were harvested and lysed in lysis buffer (20 mM HEPES, pH=7.0; 500 mM NaCl; 0.5 mM EDTA; 0.1% Triton X-100; 10 mM 2-mercaptoethanol; 1 mM PMSF; 5 mM benzamidin; 1xcOmplete EDTA free protease inhibitor cocktail tablet (Roche); 1 mg/ml lysozyme; 5 mM imidazole) for 30 minutes on ice. Lysis was assisted by passing through the cell suspension of a 25G hypodermic needle 40x times. Insoluble fraction was pelleted by centrifugation at 20.000 x *g* for 30 minutes. Supernatant was applied onto a Ni-NTA column. Washing the column was done with following buffer: 20 mM HEPES, pH=7.0; 1 M NaCl; 20 mM imidazole; 5 mM 2-mercaptoethanol. Importin-α ΔIBB was eluted with the following: 20 mM HEPES, pH=7.0; 500 mM NaCl; 10 mM 2-merchaptoethanol; 125 mM imidazole and was finally dialyzed against the following buffer: 20 mM TRIS·HCl, pH=7.8; 125 mM NaCl; 10 mM 2-merchaptoethanol. Further purification of dUTPase constructs and importin-α was achieved by gel filtration on a Superdex 200HR column (GE Healthcare), in 20 mM TRIS·HCl, pH=7.8; 125 mM NaCl, 1 mM $MgCl_2$; 2 mM DTT. Protein concentrations were determined by UV spectrophotometry using molar absorption coefficients determined by ProtParam (ExPASy) based on the amino acid sequence of the constructs.

3.2.7 Analytical gel filtration and native-PAGE analysis

Analytical gel filtration was conducted on Superdex 200HR column (GE Healthcare). Samples were applied in a total volume of 500 µl, at a concentration of 25 µM for importin-α and 8.3 µM dUTPase (1:1 stoichiometry), latter concentration corresponding to monomers. Gel filtration of the complexes was carried out after co-incubation at room temperature for 15 minutes. Fractionation was started at identical elution volumes for all samples, where 0.5 ml fractions were collected, and analyzed on SDS-PAGE. Native-PAGE analysis was performed according to the instructions of the manufacturer of Mini-PROTEAN® Tetra Cell system (BioRad), using a discontinuous buffer systems; consisting of a 4% stacking gel (pH=7.8) and a 10% resolving gel (pH=9.0).

3.2.8 Isothermal titration microcalorimetry (ITC)

ITC experiments were carried out at 20ºC on a Microcal ITC_{200} instrument (Malvern Instruments Ltd, Malvern, UK) [123]. Proteins were dialyzed into 20 mM HEPES (pH=7.5), 100 mM NaCl, 5 mM $MgCl_2$, 10 mM 2-mercaptoethanol and were used at 45 µM (importin-α, in the cell) and 460 µM (dUTPase, in the syringe) concentration. Protein concentrations correspond to monomers (subunits of dUTPase). Aliquots of 2 µl were used for at least 15 injection steps (except for the first step of 0.5 µl that was not considered in the analysis). As a control, dUTPase was also injected into the buffer to allow for considering mixing and dilution heat effects, which was subtracted. Data analysis was performed using Origin 7.5 software (Microcal Software, Inc.). The binding isotherms were fitted with an independent binding sites model 'One Set of Sites'.

3.2.9 Circular dichroism measurements

Circular dichroism (CD) spectra were recorded on a JASCO 720 spectropolarimeter using a 1 mm path length thermostated cuvette. Thermal unfolding measurements were done in phosphate buffer (100 mM Na_2HPO_4, 18 mM KH_2PO_4, 2 mM $MgCl_2$, pH=7.5) using 5 µM importin-α or 5 µM dUTPase or the mixture of these two proteins. Temperature was increased by 1°C/minute between 15°C-70°C and the CD signal at 210 nm was monitored. The fraction of the folded protein was calculated by the method previously reported [124] and plotted against temperature to give unfolding curves. Briefly, data were fitted to the equation $f_D = (\theta - (\theta_N + s_N*T))/((\theta_D + s_D*T) - (\theta_N + s_N*T))$, where f_D is denaturated protein fraction, θ is an optical property measured at given temperature (T), θ_D and θ_N are the intersections of fitted linears of totally unfolded or native protein's measured values respectively, and s_D and s_N are the slopes

of fitted linears of totally unfolded or native protein measured values respectively. T_m was determined as the inflection point of the $f_D(T)$ function.

3.2.10 Peptide Synthesis

Peptide synthesis was carried out by Dr. Fruzsina Babos and Dr. Anna Magyar. Peptides corresponding to the WT dUTPase NLS (AISPSKRARPAEV) and the S11E dUTPase NLS (AIEPSKRARPAEV) were synthesized by solid phase peptide synthesis with Fmoc/tBu strategy (SYRO, MultiSyntech automated synthesizer) on Rink-Amide MBHA resin with double-coupling protocol using 1,3-diisopropylcarbodiimide (DIC) and 1-hydroxybenzotriazole (HOBt) as coupling reagents. The crude products were purified by semipreparative RP-HPLC and the purified compounds were characterized by analytical HPLC (Knauer system using a Phenomenex Jupiter C18 column) and mass spectrometry (Bruker Daltonics Esquire 3000+ ion trap mass spectrometer).

3.2.11 Crystallization and X-ray crystallography

Protein crystallization and data collection was carried out by Dr. Mary Marfori and Dr. Bostjan Kobe. Crystals of mouse importin-α ΔIBB:dUTPase NLS peptide complexes were obtained in conditions similar to the previously determined importin-α ΔIBB:NLS complexes [32, 33]. Peptides AISPSKRARPAEV and AIEPSKRARPAEV were used as wild type and S11E (mutation underlined) mutant dUTPase NLS sequences. Importin-α ΔIBB was concentrated to 15-20 mg/ml using Centricon-30 (Millipore) and stored at −20 °C. The crystals were obtained using co-crystallization, by combining 1 µl of protein solution, 0.5 µl of peptide solution (peptide/protein molar ratio of 3.5), and 1 µl of reservoir solution on a coverslip and suspending over 0.5 ml of reservoir solution. Single crystals were obtained with a reservoir solution containing 0.65–0.70 M sodium citrate (pH 6.0) and 10 mM dithiothreitol after 15–20 days. Single crystals were cryo-protected in reservoir solution supplemented with 25% glycerol, and data were collected at the Australian Synchrotron MX1 beamline using *Blu-Ice* software [125]. Data were processed using XDS [126] and SCALA [127, 128], and molecular replacement was performed by PHASER [129] using an importin-α ΔIBB structure (PDB ID 3UKW; NLS coordinates removed [34]. Clear difference density was observed in the NLS binding site of importin-α ΔIBB, and the dUTPase NLS side-chains could be unambiguously assigned. The dUTPase peptide backbone was manually built in COOT [130]. Iterative cycles of refinement (PHENIX [131]) and model-building (COOT) gave final models with good geometry (Table

A2). Data were deposited to the PDB with the IDs 4FDR and 4FDS for the complexes of importin-α ΔIBB with wild-type and mutant dUTPase NLS segments, respectively.

3.2.12 Bioinformatics

Bioinformatics was done by Dr. Jonathan Ellis, Dr. Ahmed Mehdi and Dr. Mikael Bodén. Predikin [132] and NucImport [133] software was used to analyze the human proteome. Protein sequences were obtained from UniProt [134] for the complete human proteome including all known isoforms, as defined by UniProt complete proteome sets (representing a total of 71,809 sequences).

To predict the location of nuclear localization signals we used NucImport [133]. NucImport indicates the probability of nuclear import, type of classical NLS (as categorized by [39]) and its exact location in any query protein sequence. Apart from sequence properties, the prediction is based on known (human) protein interactions that are retrieved from BioGRID [135]. We refer to the predicted proteome set as those proteins that were assigned a type-1 classical NLS with a probability of 0.95 or greater.

Cdk1 phosphorylation sites were predicted for all potential sites, i.e., all Ser and Thr residues, using Predikin [132]. As we used the Cdk1 matrix to score all potential phosphorylation sites in the human proteome, we have the complete distribution of scores associated with Cdk1. Converting these to a cumulative density allowed us to (empirically) determine *p*-values associated with each Predikin score (the *p*-value is the probability of achieving a score at least as high as the one observed).

We looked for enrichment of phosphorylation at the P0 and P-1 sites relative to the (predicted) NLS in each protein in the nuclear proteome by counting, for each NLS site, all potential phosphorylation sites (i.e., all Ser and Thr sites) that do not occur at NLS site of interest and recorded whether they are above or below a threshold (Predikin *p*-value = 0.1). From these counts, the ratio of phosphorylation sites/potential sites can be calculated for the background, P0 and P-1 positions. We assessed whether observations at P0 and P-1 differed from the background by performing a χ^2 analysis.

Gene Ontology (GO) term enrichment analysis was performed for identified proteins using Fisher's exact test. Specifically, we used all proteins predicted to have a type-1 classical NLS and a predicted phosphorylation site at either P0 or P-1 as a foreground, and all “reviewed” human proteins in UniProtKB as background. We used the Gene Ontology official release of human annotations (as of February 2012). For each biological process GO term, we counted the number of proteins in the foreground set and the background set with this term. A one-tailed

Fisher's exact test establishes the *p*-value of the term: the probability of finding this protein count or more extreme (greater proportion in the foreground). The *p*-value was corrected for multiple testing (shown as *E*-value). A term is thus assigned a small *E*-value only if proteins annotated with that term occur in the foreground set with a higher prevalence than can be statistically explained by chance (i.e. proteins picked randomly from the background set).

3.3 NLS copy number variation governs efficiency of nuclear import: case study on dUTPases

3.3.1 Cell culture

Schneider S2 cells (derived from *Drosophila melanogaster*) were purchased from Gibco. WR-Dv-1 cell line (derived from *Drosophila virilis*) was purchased from DGRC (Bloomington, IN, USA). Schneider S2 and WR-Dv-1 cells were cultured in Schneider Insect Medium (Sigma) supplemented with 10% FBS (Gibco) and 50 µg/ml Penicillin-Streptomycin (Gibco) and kept in a 26 °C incubator. *Drosophila* strains were obtained from Drosophila Species Stock Center (La Jolla, CA, USA).

3.3.2 Plasmid constructs and cloning

For bacterial expression the *D. virilis* ABC dUTPase (pseudo-heterotrimer) gene was amplified from *D. virilis* genomic DNA, isolated by MasterPure™ DNA Purification Kit (Epicentre, Madison, WI, USA), with primers DvirDF and DvirDR3 (Table A7). PCR fragment was cloned into the NheI/BglII sites of the vector pET-15b (Novagen). *D. virilis* homotrimeric AAA dUTPase was constructed by amplifying the 'A' subunit of the ABC pseudo-heterotrimer from genomic DNA with primers DvirDF and DvirDRs1 and cloned into the NheI/BglII sites of the vector pET-15b (Novagen).

For mammalian cell line expression both dUTPases were tagged by EGFP by cloning them into the KpnI/SmaI restriction sites of pEGFP-N1 (Clontech) vector after amplification from pET-15b vectors with A_F, A_R, ABC_R primer sets accordingly (see Table A7). For insect cell line expression the EGFP-tagged *D. virilis* dUTPases (AAA and ABC) were further cloned into the KpnI/NotI sites of the vector pIZ/V5-His (Life Technologies). NLS deleted constructs (ΔNLS-AAA and ΔNLS-ABC) were created by the QuickChange mutagenesis method (Stratagene) using NLSdelF and NLSdelR primers on AU1 or EGFP tagged and untagged constructs (both for AAA and ABC dUTPases) cloned into the insect expression vector (pIZ) or the bacterial expression vector (pET-15b). AU1 tagged dUTPase constructs were generated

by annealing a pair of single stranded oligonucleotides encoding the tag (A_Au1_F and A_Au1_R) yielding compatible ends to ligate it into the BamHI and KpnI sites of the pIZ-A-EGFP vector, replacing EGFP with AU1 tag on the C-terminal of the dUTPase. For the AU1 tagging of ABC dUTPase, the ORF was amplified via PCR (ABC_Au1_F, ABC_Au1_R) where the reverse primer encoded the AU1 tag and the product cloned into the KpnI and EcoRI sites of an empty pIZ vector. *D. melanogaster* GFP tagged nuclear and cytoplasmic dUTPase isoforms were described previously in [110]. Primers used in this study were synthesized by Eurofins MWG GmbH and are summarized in Table A7. All constructs were verified by sequencing at Eurofins MWG GmbH.

3.3.3 Plasmid transfection and immunocytochemistry

The subcellular localization of dUTPase constructs were investigated using different mammalian and insect cell lines. DNA transfections were performed according to the manufacturer's recommendations using FuGENE HD reagent (Roche). In case of S2 cells, coverslips were coated with concanavalin-A (Sigma) to flatten cells and aid visual inspection. Immunocytochemistry was carried out as described in section 3.2.5. Endogenous dUTPase was detected using the polyclonal antibody previously described in [109] at 1:10000 dilution and a monoclonal antibody specific for lamin Dm0 (ADL67.10 from DSHB, Iowa City, Iowa, USA) was used at 1:400 dilution to stain the nuclear membrane. To aid visual inspection of the localization pattern, cells were counterstained with 1 µg/ml DAPI (4',6-diamidino-2-phenylindole, Sigma) and 0.5 µg/ml phalloidin-TRITC (Sigma). Same image acquisition settings were applied in each case when comparing different constructs in each cell type. Image analysis to quantify relative subcellular localization was performed from single-cell measurements as described in section 3.2.3.

3.3.4 Immunohistochemistry and immunoblot analysis

Localization of endogenous dUTPase was visualized in *D. virilis* and *D. melanogaster* tissues (data is only shown for larval wing discs but was similar in all inspected tissues). Dissected Drosophila organs prepared by Imre Zagyva and Dr. András Horváth. Ovary, testis, larval imaginal wing discs, gut and salivary gland samples were collected and were immediately fixed in 50% n-Heptane and 50% PEM-formaldehyde (100 mM PIPES, 1 mM $MgCl_2$, 1 mM EGTA, 2.5% Tween-20, 4% PFA, pH=6.9) for 30 minutes, with vigorous shacking at room temperature then washed with inactivating buffer (50 mM TRIS, 150 mM NaCl, 0.5% Tween-20, pH=7.4). Blocking was performed in 5% goat serum, 1.5% BSA, 0.1% Tween-20, 1% Triton-X 100,

0.001% NaN_3, in PBS, pH=7.4 for 4 hours at RT. Tissues were incubated in primary antibodies diluted in blocking buffer (anti-dUTPase,1:10000; anti-lamin Dm0, 1:400), at 4 °C for 16 h. Samples were further washed with blocking buffer for 8 hours at RT. Secondary antibodies were applied in 1:1000 dilution coupled to either Alexa 488 or Alexa 633 (Molecular Probes) in blocking buffer for 2 h, RT. Tissues were stained with DAPI and 0.5 µg/ml phalloidin-TRITC and embedded in FluorSave™ (Calbiochem) reagent. Images were acquired with a Zeiss LSCM 710.

Subconfluent cells in T25 flasks were transfected with 6.5 µg DNA supplemented with 30 µl transfection reagent for 48 h. Immunoblotting was carried out as described in section 3.1.5. Antibodies were used in the following dilution: dUTPase (1:50000) and GFP (Sigma, 1:1000).

3.3.5 Recombinant protein production

dUTPase constructs and importin-α were expressed and purified as described in section 3.1.1 and 3.2.6 but the expression of dUTPase constructs were carried out at 25°C.

3.3.6 Analytical gel filtration and native-PAGE analysis

Analytical gel filtration and native-PAGE analysis was conducted as described in section 3.2.7. Samples were applied at a concentration of 32.7 µM for importin-α, 10.9 µM for AAA dUTPase and 32.7 µM for ABC dUTPase (to achieve 1:1 stoichiometry of NLSs and importins) during gel filtration.

3.3.7 Isothermal titration microcalorimetry (ITC)

ITC experiments were carried out and analyzed as described in section 3.2.8 with some modifications. Proteins were dialyzed into 20 mM HEPES (pH=7.5); 100 mM NaCl; 5 mM $MgCl_2$; 2 mM Tris(2-carboxyethyl)phosphine; 5% glycerol and were used at the following concentrations (these concentrations are given in monomers for the AAA dUTPase and correspond to pseudo-heterotrimer for the ABC dUTPase): AAA dUTPase was used at 400 µM while ΔNLS-AAA dUTPase was used at 550 µM in the syringe, along with 40 µM and 55 µM importin-α in the cell respectively. ABC and ΔNLS-ABC dUTPase were used at 250 µM in the syringe, along with 25 µM importin-α in the cell. Aliquots of 1.5 µl were used for at least twenty injection steps (the first step of 0.5 µl was not considered in the analysis).

3.3.8 5'-Rapid Amplification of cDNA Ends (5' RACE)

In order to isolate potential alternative transcripts from *D. virilis* embryos and ovaries and to determine their 5' ends, 5' RACE was performed using the 5'/3' RACE Kit (Roche) according to the manufacturer's instructions. Primers used are summarized in Table A7.

Briefly, *D. virilis* embryos and ovaries were harvested and immediately stored in RNA*later*™ (Ambion, Life Technologies, Carlsbad, CA, USA), snap-frozen in liquid nitrogen, and stored at −80 °C until use. TRIzol reagent was used to isolate total RNA from the homogenized samples according to the manufacturer's recommendation (Life Technologies) and was further purified after DNase (Qiagen) treatment with RNA Clean-up Kit (Macherey-Nagel, Düren, Germany). The quality of RNA samples were analyzed by gel electrophoresis. Total RNA (1 µg) was reverse transcribed with a gene specific primer (SP1 green), in a 20 µl reaction volume. RNase Cocktail™ Mix (Ambion) and RNaseH (NEB, Ipswich, MA, USA) was added and incubated at 37 °C for 15 min. The resulting RNA free cDNA preparation was further purified using the High Pure PCR Product Purification Kit (Roche). After poly-A tailing, cDNA was amplified by RedTaq DNA polymerase (Sigma) using a gene specific primer (SP2) and the Oligo dT-anchor primer provided with the kit. A second nested PCR was performed using another gene specific primer (SP3) and the PCR anchor primer provided with the kit. Final RACE products were cloned into the SalI/EcoRI sites of pBluescript SK+ vector (Stratagene) for sequencing. Sequencing was performed by Eurofins MWG GmbH using the M13 uni(-43) primer.

3.3.9 *In silico* splicing site prediction

DNA sequences of dUTPases and surrounding genomic regions were downloaded from FlyBase (*D. melanogaster*: **CG4584** and *D. virilis*: **GJ10455**). *In silico* splice site predictions were done with the Human Splicing Finder, Version 2.4.1 (http://www.umd.be/HSF/) with conservative settings [136].

4. Results and discussion

4.1 Characterizing the calpain-catalyzed proteolysis of the human dUTPase

As compared to digestive proteases like trypsin, calpains perform a more fine-tuned role by regulating the function of numerous substrates by limited proteolysis [137]. Intracellular proteolysis events are key steps in signaling. The identification of substrates of specific proteolytic events are indispensable to decipher the cellular network and significance of these pathways. This work is particularly complex in the case of calpains that play roles in several pathways. A search for *in vivo* calpain substrates in a *Drosophila melanogaster* system, provided a list of new calpain substrates with hits belonging to several different cellular pathways [118]. Our aim was to investigate if human dUTPase may also be susceptible to calpain-catalyzed proteolysis, as it was reported in *Drosophila melanogaster*; and if so, how does it influence dUTPase function. During our study we used the rat m-calpain however we do not believe this might cause any significant difference since the two proteins share 94 % identity and 99 % similarity.

4.1.1 *In vitro* calcium-dependent proteolysis of human dUTPase by calpain

We prepared an *in vitro* digestion assay with the human nuclear dUTPase isoform and rat m-calpain. We found that dUTPase is digested by m-calpain in a calcium-dependent manner. During digestion two distinct dUTPase fragments could be visualized after SDS-PAGE analysis (Figure 6). The intensity of the large m-calpain subunit also drops in the presence of calcium because of auto-proteolysis. It has been previously shown that the substrate analogue α,β-imido-dUTP (dUPNPP) provides a partial protection against limited tryptic digestion, due to the conformational change it induces at the C-terminal end of dUTPase [120, 138]. However in case of m-calpain, the presence of dUPNNP did not have a protective effect, no alteration of the proteolytic process could be seen. This suggests that the calpain cleavage sites are not located in the vicinity of the substrate binding pocket and that the cleavage probably takes place at the flexible N-terminal end.

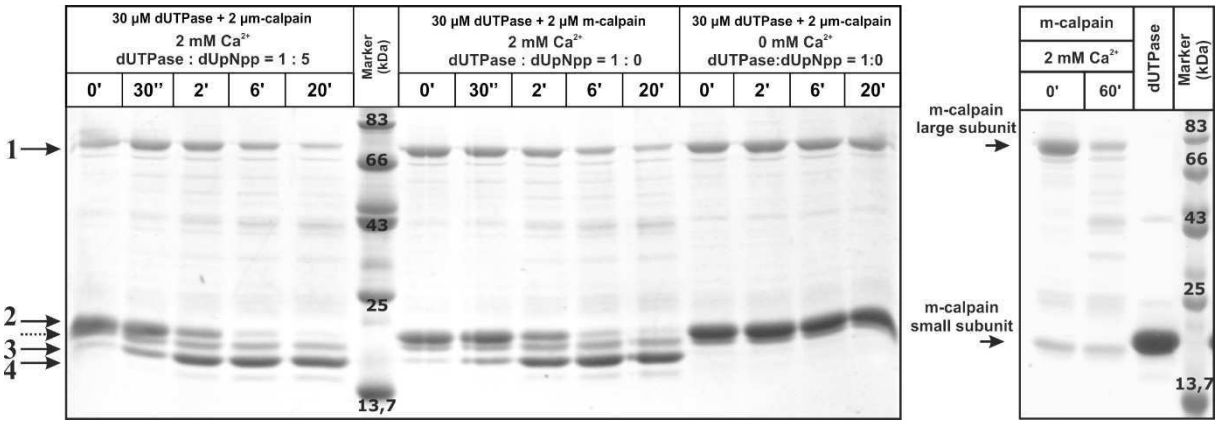


Figure 6. Calcium-dependent proteolysis of the human dUTPase by m-calpain

Left panel: SDS-PAGE analysis of the i*n vitro* digestion of the human dUTPase by m-calpain in the presence or absence of substrate analogue. Control experiment was carried out in the absence of calcium. Dotted arrow points at the small subunit of m-calpain. Lane 1: m-calpain large subunit, Lane 2: intact dUTPase, Lane 3 and 4: dUTPase fragments. Reaction was stopped after the indicated incubation times. **Right panel:** shows calpain and dUTPase samples on their own to facilitate clear-cut identification of the bands corresponding to the calpain small subunit and intact human dUTPase that run very close to each other on the gel.

This hypothesis was further addressed by measuring the catalytic activity of the m-calpain cleaved fragments and intact dUTPase forms. Trypin cleavage close to the C-terminal disrupts catalytic activity of the enzyme [120]. However after m-calpain cleavage no difference in catalytic function was observed (Figure 7), indicating that calpain digestion does not alter dUTPase activity.

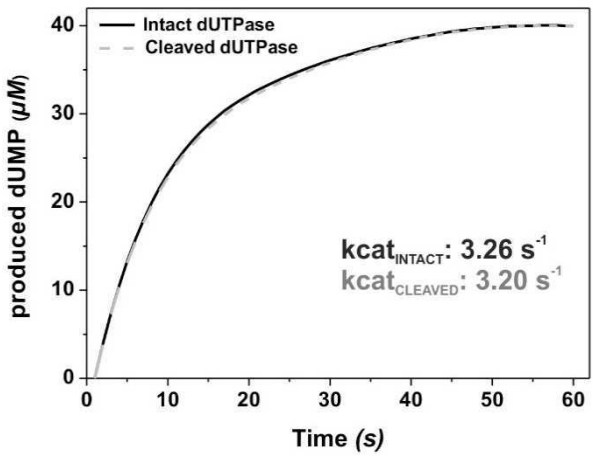


Figure 7. Effect of calpain cleavage on dUTPase activity

dUTPase activity was followed for the intact (black solid line) and cleaved (grey dashed line) proteins. Catalytic rate constants (kcat) are indicated in the graph, which were 3.26 s^{-1} for the intact and 3.20 s^{-1} for the m-calpain digested dUTPase.

4.1.2 Mass spectrometry revealing the calpain cleavage sites

To identify the specific cleavage sites, fragments were analyzed by mass spectrometry. Three cleavage sites were identified in the two fragments separable on SDS-PAGE (Figure 8 and Table A1).

Figure 8. Mapping calpain cleavage sites

The N-terminal segment of the 164 amino acid long human dUTPase is presented. The three calpain cleavage sites are indicated in red (^{4}SE^{5}; ^{7}TP^{8}; ^{31}LS^{32}). The NLS (^{12}P-A^{19}) is indicated with a box, where cleavage at the ^{31}LS^{32} position results in the loss of this segment. M^{24} marks the position from where structural information is available of the enzyme (PDB ID: 2HQU).

The first two cleavage sites are found at the very beginning of the flexible N-terminal very close to each other (^{4}SE^{5}; ^{7}TP^{8}). The third site (^{31}LS^{32}) can be located in the three-dimensional structure of the human dUTPase and cleavage at this site results in a truncated protein that lacks the NLS segment (located between residues ^{12}P-A^{19}, [104]).

Results of *in vitro* digestion experiments showed that human dUTPase is cleaved by calcium activated m-calpain within the flexible N-terminus (Figure 6 and 8). Although catalytic function is unperturbed in the N-terminally truncated form (Figure 7), cellular localization is expected to be altered, since the truncated dUTPase species does not contain the NLS segment, indispensable for nuclear import [104]. Nucleocytoplasmic trafficking of trimeric dUTPases in eukaryotic cells depends on its NLS since the 53.25 kDa molecular mass of these proteins prevents their passive diffusion into the nucleus. In *Drosophila melanogaster* the lack of the NLS signal of the nuclear isoform was shown to lead to the exclusion of dUTPase from the cell nucleus [110].

4.1.3 Degradation of the dUTPase in HeLa cells *via* calcium-activation

To test if the *in vitro* results may also be applied for intact cells, we wished to follow dUTPase levels in HeLa cells in the absence or presence of calcium activation. First we wanted to see if the HeLa cell line contains a significant calpain pool that could be activated by calcium under our experimental conditions. We performed activity measurements with the calpain specific substrate and HeLa cell extract in the absence and presence of calcium. Results show that the

presence of calcium led to calpain activation, indicating the presence of an active calpain pool (Fig. 3) while without calcium no activity could be observed.

Treating HeLa cells with ionomycin led to significant dUTPase degradation after 24 h. Some decrease in the m-calpain level could also be detected (due to auto-proteolysis), while the dUTPase and m-calpain levels in the control samples were unaffected. However we could not detect any dUTPase fragments, presumably due to further degradation processes (Figure 9) also the antibody used might not recognize the resulting fragments. To monitor equivalent amounts of total protein applied per lane, the western blots were also developed against actin.

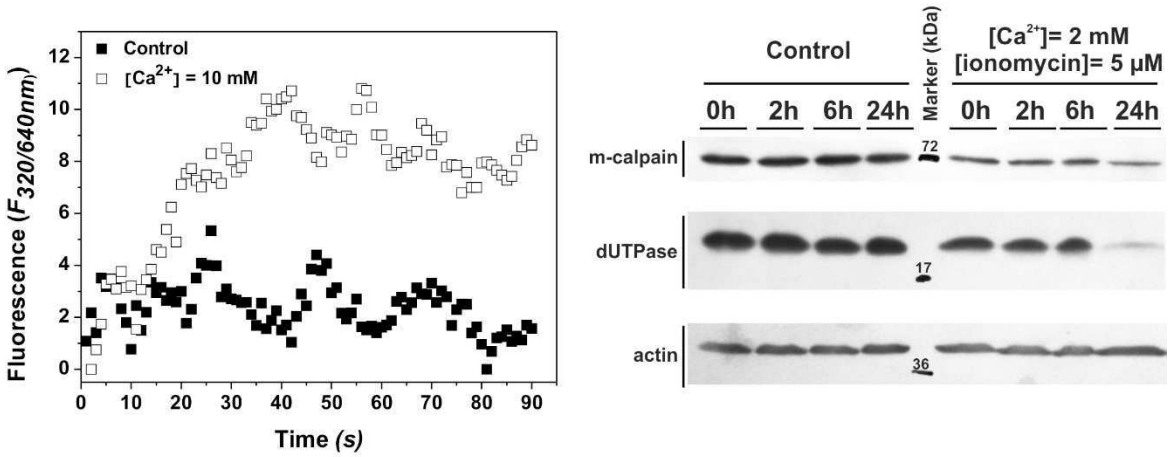


Figure 9. Degradation of cellular dUTPase pool upon calcium induction

Left panel: Calpain activity in HeLa cellular extract was recorded in the absence of calcium (control, black squares), or the presence of 2 mM calcium (empty squares). **Right panel:** dUTPase and m-calpain levels were monitored for 24 hours in cells treated (active calpain pool) and not treated (inactive calpain pool) by ionomycin. Decrease in dUTPase levels could be observed after 24h in the treated cells. Actin was used as a loading control.

We hypothesize that within the cells, calpain-cleaved dUTPase may be further degraded by other proteolytic events. Possible m-calpain mediated proteolysis results in dUTPase fragments having an E^5 or a S^{32} residue on their N-terminal, both possibly destabilizing the enzyme, based on the N-end rule [139], by making the fragments more susceptible to proteasomal degradation. Also degradation of dUTPase in HeLa cells was only observed after 24 hours of calcium-induction, which we believe might be due to the fact that the cytoplasmic m-calpains may get hold of their nuclear substrates mainly during cell division when the nuclear envelope breaks

down. It is widely observed that calpains harbor several nuclear substrates [140, 141], and even play role in the regulation of cell division [142-145].

Based on our results, we propose a hypothesis that calpain activation may control dUTPase localization and function through limiting its overall availability. Upon calcium-induction, and following mitosis-coupled release of the nuclear dUTPase pool into the cytoplasm during the cell cycle, calpain cleaves the N-terminal of dUTPase removing the NLS segment. Truncated dUTPase is unable to enter the nucleus; it remains in the cytosol and gets eventually degraded. It is still uncertain whether nuclear availability of the enzyme is generally required for eukaryotic cells to fulfill its function in genome integrity maintenance (discussed later in section 4.2.3 in detail), however in eukaryotic cells the usual presence of a nuclear localization signal at the N-terminus argue for the imminent nuclear function of dUTPase. Also lack of dUTPase results in perturbed dUTP/dTTP ratios and provokes cell death (detailed in section 1.2.1.3). This pathway of calpain action may act in parallel with other described roles of calpain in apoptosis (e.g. [146, 147]).

4.2 Effect of phosphorylation adjacent to nuclear localization signals on nuclear proteome re-constitution: a case study on dUTPase

Besides limited proteolysis, protein localization and availability could be regulated on several levels including reversible post-translational modifications such as phosphorylation. Based on earlier results the nuclear isoform of the human dUTPase was suggested to be phosphorylated by Cdk1 at the Ser11 position, which is just upstream the NLS signal [104, 107]. Possible effects were only investigated on a S11A mutant, that could not be phosphorylated (hypo-phosphorylation mimicking mutant) which showed no difference compared to the WT enzyme in respect of activity and subunit association. So we aimed to determine the effect of Cdk1 phosphorylation on cellular localization of dUTPase by utilizing different point mutations at the phosphorylation Ser11 site. Glutamic acid mimics a constitutively phosphorylated serine residue (P-mimic, hyper-P mutation) and may be optimally paired with studies of the isosteric but not charged non-phosphorylatable (hypo-P mutation) mutant using glutamine [148].

4.2.1 Effect of NLS phosphorylation on dUTPase localization

4.2.1.1 Localization pattern of WT, S11E (hyper-P) and S11Q (hypo-P) dUTPase-DsRed constructs

The WT (phosphorylatable) DsRed-labeled dUTPase (WT DsR-DUT) is mainly nuclear in asynchronous cell populations, whereas the hyper-P S11E mutant is solely cytoplasmic (Figure 10). The hypo-P S11Q mutant is also nuclear, indicating that the introduced negative charge is responsible for the altered localization pattern of the S11E mutant (Figure 10). Localization of the different dUTPase constructs were confirmed in a number of cell types (MCF-7, COS7, HeLa, NIH-3T3 and 293T cell lines) with diverse genetic backgrounds differing in both origin and modes of transformation (Figure 10), showing that the intracellular distribution of the constructs are not cell line specific.

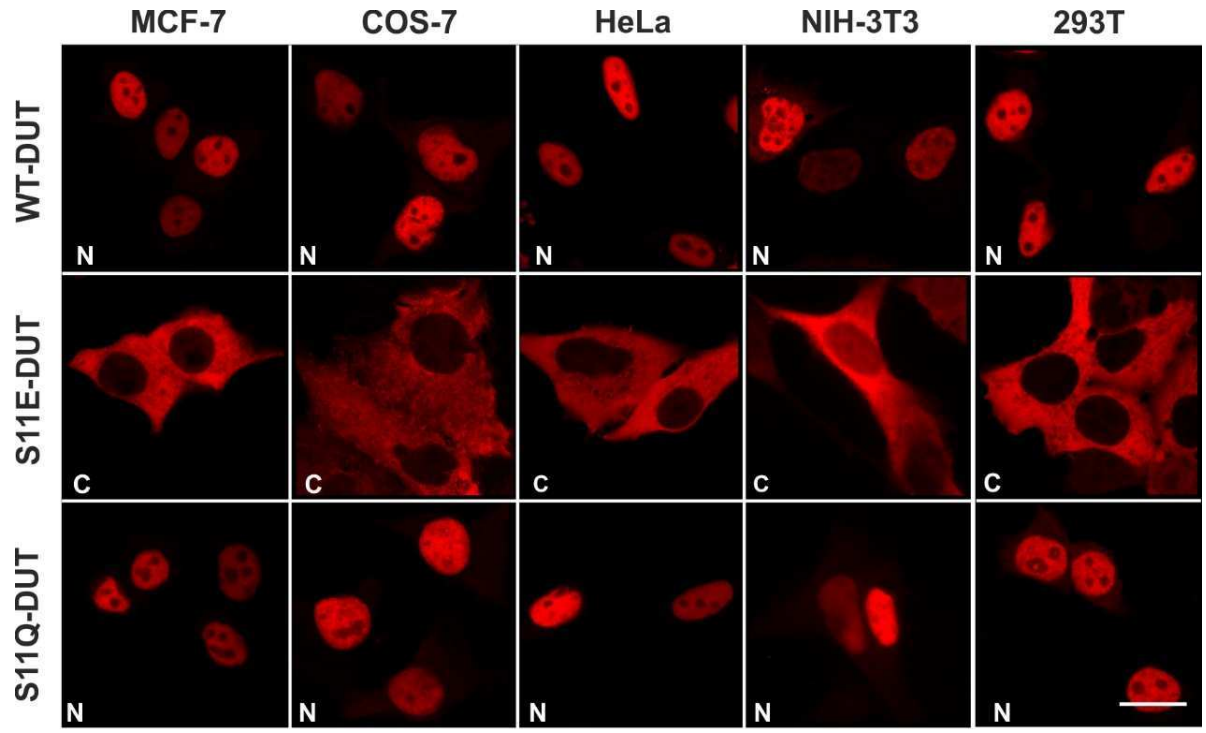


Figure 10. Localization of different phosphorylation status mimicking dUTPase mutants

The DsRed-labeled WT human dUTPase (nuclear isoform) and hypophosphorylation mimicking S11Q mutant is mainly nuclear in asynchronous cell lines, whereas the hyperphosphorylation mimicking S11E mutant is cytoplasmic. Localization pattern is the same in MCF-7, COS7, HeLa, NIH-3T3 and 293T cell lines. N: nuclear; C: cytoplasmic. Scale bar represents 20 μm.

4.2.1.2 Contribution of active nuclear export processes to the localization pattern

Leptomycin B, a well-described specific inhibitor of one of the major exportins (CRM1), was used in order to check the possible contribution of active nuclear export to the strict nuclear exclusion of the S11E mutant [149]. Since leptomycin had no effect on the localization pattern we can conclude that the exclusive cytoplasmic localization of the S11E mutant dUTPase construct is probably due to perturbation of its nuclear import and not due to accelerated nuclear export (Figure 11).

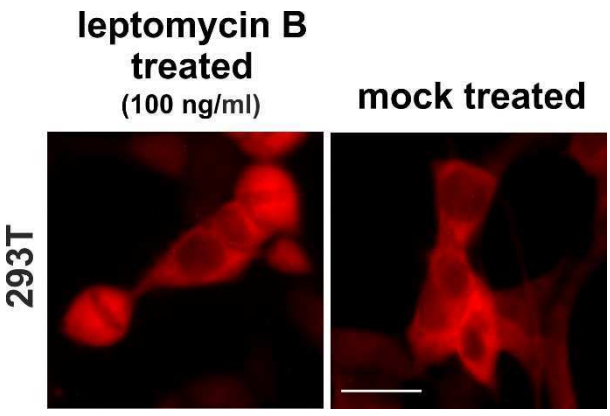


Figure 11. Nuclear export dependency of S11E dUTPase localization pattern

293T cells expressing S11E mutant DsR-DUT were subjected to leptomycin B (a selective CMR1 inhibitor) treatment. No difference was observed in intracellular distribution of the protein compared to methanol mock-treated cells. Scale bar represents 20 μm.

4.2.1.3 Timing of the S11 position phosphorylation of dUTPase

Cdk1-cyclin B is activated at the M phase of the cell cycle [150]. Therefore we investigated the putative cell-cycle dependency of the phosphorylation of the endogenous dUTPase using an M-phase marker (histone-H3-phospho-S10 antibody). All cells in an asynchronous cell population were stained with the polyclonal anti-hDUT antibody (regardless of phosphorylation status), whereas the anti-P-hDUT antibody (generated against the phosphorylated S11 NLS position of dUTPase) stains only cells in the M phase (Figure 12). Based on these results we conclude that phosphorylation of dUTPase occurs within the nucleus and is scheduled at the M phase of the cell cycle. Our results further strengthen data found in the literature proposing potential involvement of Cdk1 kinase and its partner, cyclin B, in this process.

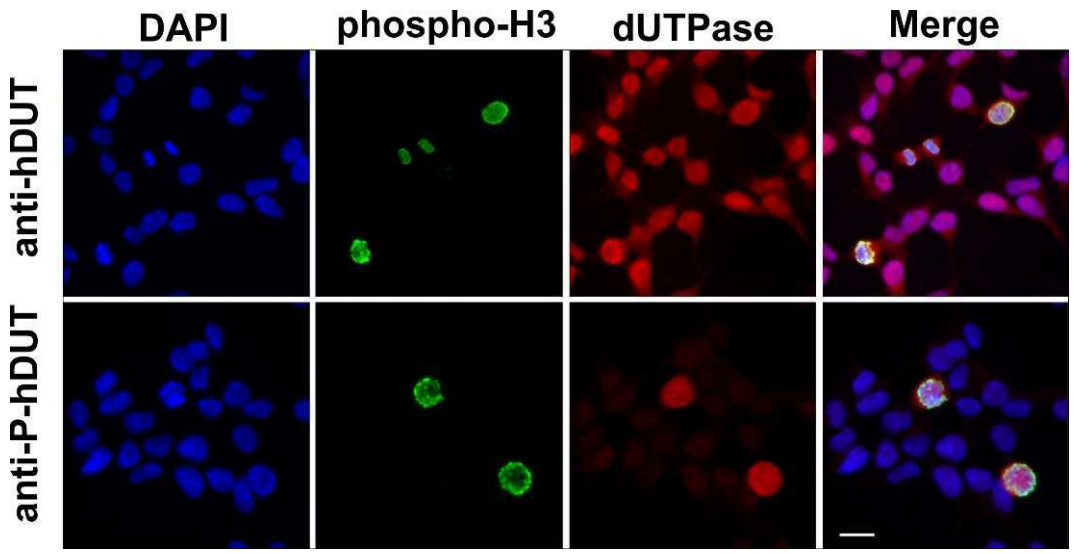


Figure 12. Cell cycle-dependent phosphorylation of dUTPase in asynchronous cells.

Asynchronous 293T cells were stained with the polyclonal antibody specific for dUTPase (anti-hDUT) or the phospho-NLS-dUTPase specific antibody (anti-P-hDUT). Immunofluorescence shows that dUTPase phosphorylation coincides with anti-histone H3 (phospho-H3) M-phase marker. Scale bar represents 20 μm.

4.2.1.4 Analysis of dUTPase distribution throughout the cell cycle with video microscopy

To follow dUTPase localization throughout the cell cycle we followed individual cells after transfection with the appropriate fluorescent constructs by video microscopy. The S11E mutant remains cytoplasmic during the entire cell cycle (data not shown), so this construct was not considered for further analysis. The WT and S11Q dUTPase pool exhibits marked cell cycle-dependent dynamic behavior (Figure 13A). When the new nuclear envelope appears, dUTPase is excluded from the nucleus and it takes a considerable time before the nuclear space is again

repopulated with WT dUTPase in the daughter cells. Nuclear re-population dynamics is markedly different for the S11Q mutant, which gets imported back to the nucleus much faster. We were able to measure the apparent rate constant for nuclear re-accumulation of the dUTPases constructs. Major difference between the WT and the S11Q mutant is that the WT protein starts re-entering the nucleus after a considerable lag (~200 minutes). However the apparent rate constants of nuclear accumulation are identical for the WT (k_{obs} = 0.0044 min^{-1} ± 7%) and for the S11Q mutant (k_{obs} = 0.0043 min^{-1} ± 8%) after the lag phase (Figure 13B). The observed apparent single exponential kinetics likely represent the result of multiple undistinguishable transport events. Importantly, the mean total fluorescence of the cells (F_{n+c}) did not change during our observations (Figure 13B, inset), indicating the steady-state of the investigated fluorescent protein pool. Western blot experiments were used to investigate whether the transfected constructs used in the video-microscopic experiments can be phosphorylated similarly to the endogenous protein (Figure 13C). As we can see in Figure 13C, 293T cells, transfected with the appropriate fusion protein-encoding plasmids, produce a WT DsR-DUT protein pool that can be phosphorylated on the S11 residue. The recognition of dUTPase by the anti-S11P-hDUT antibody is observed only if Ser11 can be phosphorylated, providing evidence for the specificity of this antibody. The endogenous dUTPase pool is also visible at a lower molecular mass position. The anti-hDUT antibody recognizes all dUTPase constructs, as well as the endogenous dUTPase pool, independently of the point mutation or phosphorylation state (Figure 13C).

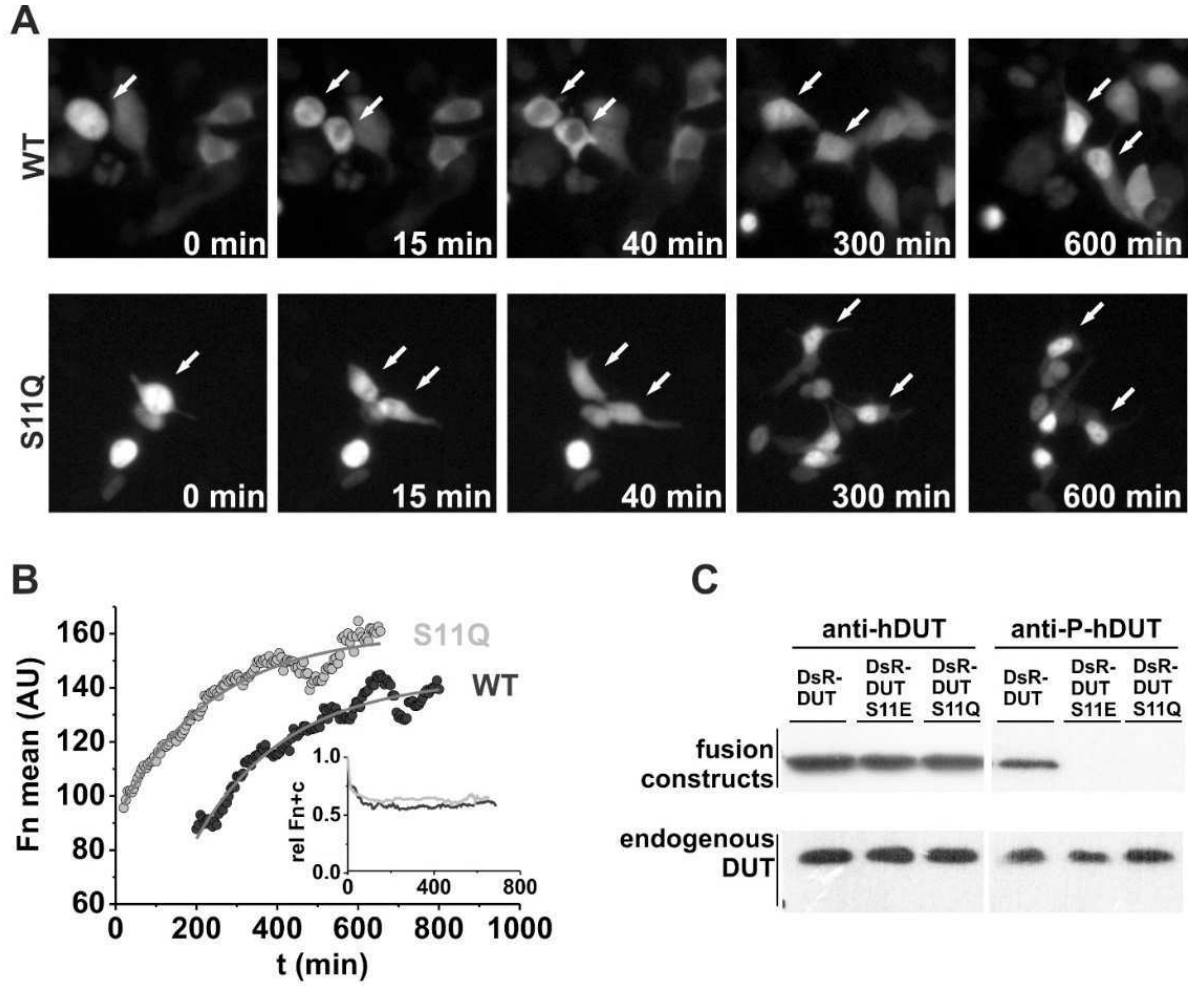


Figure 13. Phosphorylation-dependent localization pattern of dUTPase during cell cycle

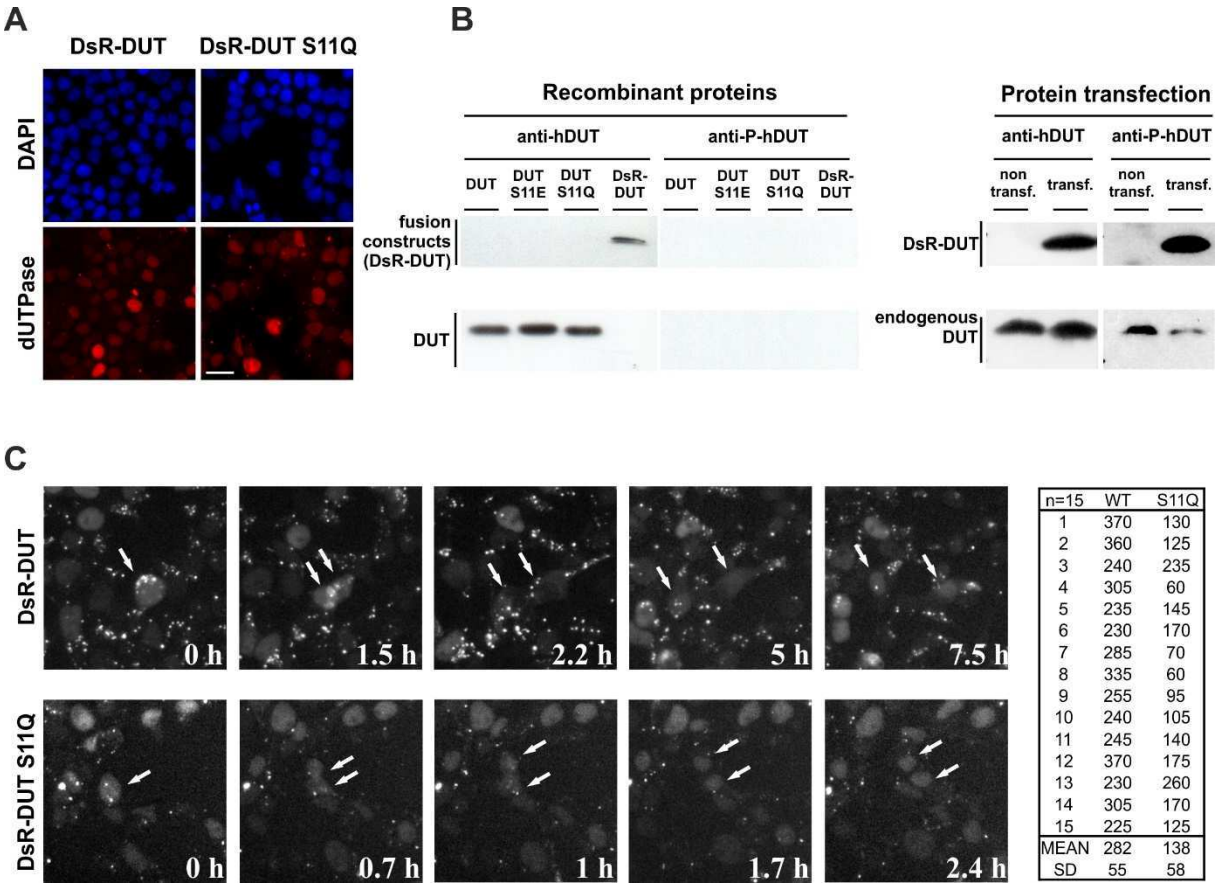
(A) Transfected 293T cells were observed for at least one full cell cycle with live-cell imaging. Still images were taken of actively dividing cells at the indicated time points. The once-nuclear dUTPase pool gets re-imported into the nucleus with different characteristics for the WT and the S11Q mutant.

(B) Kinetic analysis of protein re-import dynamics of the daughter cells indicate similar import kinetics but different lag phases for the WT protein and the S11Q mutant (k_{obs} = 0.0044 min^{-1} ± 7% and 0.0043 min^{-1} ± 8%, respectively). Inset shows that mean total fluorescence of the cells (Fn+c) did not change during our observations.

(C) Western blot only shows cognate phosphorylation event on the S11 position of the WT DsRed-tagged exogenous dUTPase (DsR-DUT).

To directly trace the nucleocytoplasmic trafficking of a given protein pool without continuous protein expression from plasmids, we repeated the video-microscopy experiments by transfecting the fluorescent proteins themselves. Cells transfected with purified WT and S11Q DsR-DUT proteins show the same dynamic events of dUTPase pool distribution as those in the

plasmid transfection experiments (Figure 14A and C). Neither forms of the recombinant proteins purified from *E. coli* are recognized by the anti-P-hDUT antibody, indicating that they are not phosphorylated at Ser11. Within the cells transfected with purified DsR-DUT protein itself, however, the cognate phosphorylation event targeting Ser11 can take place (Figure 14B).



n=15	WT	S11Q
1	370	130
2	360	125
3	240	235
4	305	60
5	235	145
6	230	170
7	285	70
8	335	60
9	255	95
10	240	105
11	245	140
12	370	175
13	230	260
14	305	170
15	225	125
MEAN	282	138
SD	55	58

Figure 14. Live cell microscopy after protein transfection.

(A) Both recombinant DsR-DUT WT and DsR-DUT S11Q localize to the nucleus after protein transfection, as in the case of plasmid transfection. Scale bar represents 20 μm.

(B) Neither forms of recombinant proteins produced in *E. coli* are recognized by the anti-P-hDUT antibody, indicating that they are not phosphorylated at Ser11. Recombinant DsR-DUT can be phosphorylated after protein transfection similarly to the endogenous protein.

(C) Still images taken from live-cell imaging of 293T cells containing WT and S11Q mutant dUTPase. Cells were followed during throughout a cell cycle after protein transfection. The differences observed between the two forms are similar to the dynamics observed with the plasmid transfection based experiments. The inserted table shows the time (minutes) elapsed between the onset of cytokinesis and the appearance of considerable fluorescent signal within the nucleus. Data were collected from fluorescent time-lapse image sequences taken in 5 min intervals. Parallel phase contrast images were used to determine the onset of daughter cell separation.

Both DNA and protein transfection-based experiments gave the same conclusions regarding the dynamic distribution pattern of WT and S11Q mutant DsR-DUT. This is potentially due to the fact that the DsRed-labeled proteins can only be detected after a considerable delay of the time following protein translation, partially because of the time required for maturation of DsRed fluorophore and because of the time required for detectable fluorophore accumulation and maturation. Furthermore, newly maturing DsRed molecules (which also went through phosphorylation in M phase) might be in steady state with a degradation process. Because of these effects, the DsR-DUT pool translated during the recording time of video-microscopy used for analysis (~12 hours) does not contribute to the fluorescent signal significantly. The observable fluorescent signal of the mature folded protein molecules thus necessarily originates from the protein pool translated during the cell cycle(s) completed prior to start of the video recording. The new *de novo* synthesized protein pool (after division) will not be readily detectable until cells enter the next cell cycle.

In summary, we described the cell cycle-dependent phosphorylation pattern of dUTPase. The strict schedule of this event is linked to the M phase, supporting the role of Cdk1 kinase. We have also shown that the phosphorylation mimicking mutation leads to exclusion from the nucleus. These results indicate that the phosphorylation of dUTPase within the nucleus has physiological implications manifested in its retarded re-import in the daughter cells, namely nuclei of the daughter cells become re-populated with dUTPase protein only after a significant delay. The non-phosphorylatable S11Q construct does not exhibit this behavior. We suggest that this mechanism may also operate for a number of similar phosphorylatable NLSs on other proteins. Assuming similar nuclear re-import characteristics as that of dUTPase, Cdk1 kinase-induced phosphorylation at these similar NLS positions would significantly alter the nuclear proteome re-establishment in the daughter cells after the M phase which will be discussed in detail in section 4.2.4. We first wanted to determine the structural basis of the impeded import of the dUTPase that results from the perturbation of importin-α:dUTPase interaction (section 4.2.2); and second, performing a large-scale study to identify proteins in the human proteome that might follow a similar pattern of nucleocytoplasmic trafficking to dUTPase (section 4.2.4).

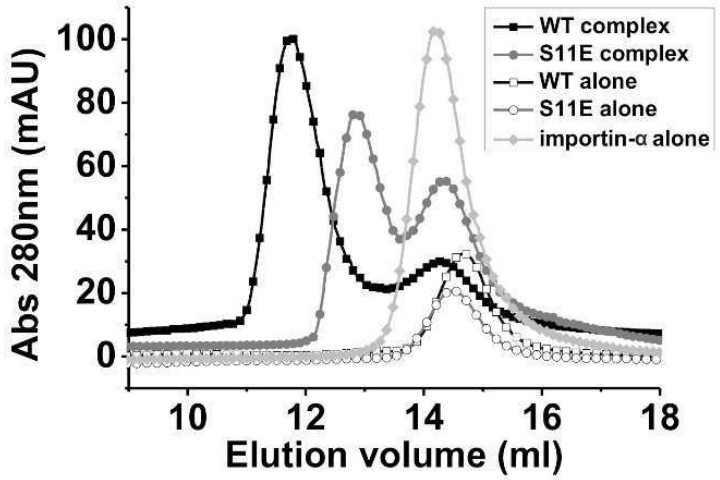
4.2.2 Molecular details of the interaction between human dUTPase and importin-α

Several independent biophysical methods were used to evaluate complex formation of WT, hypo-P S11Q and hyper-P S11E mutant dUTPases and importin-α. For this latter protein, the

construct lacking the autoinhibitory domain (importin-α ΔIBB), was used throughout in our study.

4.2.2.1 Analytical gel filtration

Analytical gel filtration showed that in the mixture of WT dUTPase and importin-α, a molecular species with higher molecular mass than either components separately is present (Figure 15). This argues for a complex formed between the two proteins. A fraction of the protein still elutes at the position corresponding to the free components, potentially representing the protein pool that did not form a complex (Figure 15). The chromatogram profile was different when the S11E mutant dUTPase was used, still having a peak eluting earlier than any of the components alone, but later than the peak associated with the WT dUTPase:importin-α complex (Figure 15). Also, the protein fraction eluting at the position corresponding to the free components is larger. This behavior might be explained by assuming that the complex formed between S11E dUTPase and importin-α is characterized with faster dynamic equilibrium, i.e. it is less stable as compared to the WT dUTPase:importin-α complex [151]. The same conclusions could be made after analyzing the fractions on SDS-PAGE (Figure 15).



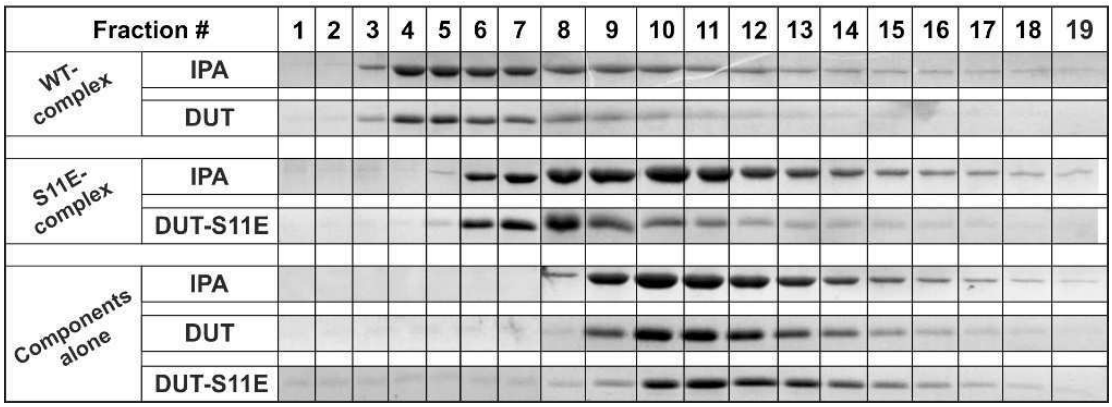


Figure 15. Analysis of dUTPase-importin-α complex formation with size exclusion chromatography
Upper panel: Chromatograms of size-exclusion chromatography showing complex formation between dUTPase and importin-α (IPA), weakened by the S11E mutation. **Lower panel:** Fractions of the gel filtration chromatography were analyzed on SDS-PAGE. The WT dUTPase:importin-α complex elutes at a smaller volume compared to the S11E dUTPase:importin-α complex or the components alone.

4.2.2.2 Native-PAGE

Native-PAGE results show that WT and S11Q mutant dUTPase forms a complex with importin-α, showing up at a new position on the gel, whereas the band of uncomplexed importin-α disappears (Figure 16). No complex formation could be detected in case of the S11E mutant dUTPase. Results indicate that the different nucleocytoplasmic localization patterns of the investigated constructs are likely attributed to modulation of the dUTPase:importin-α interaction. The WT and the S11Q mutant forms a stable complex with importin-α, whereas the S11E mutant does not.

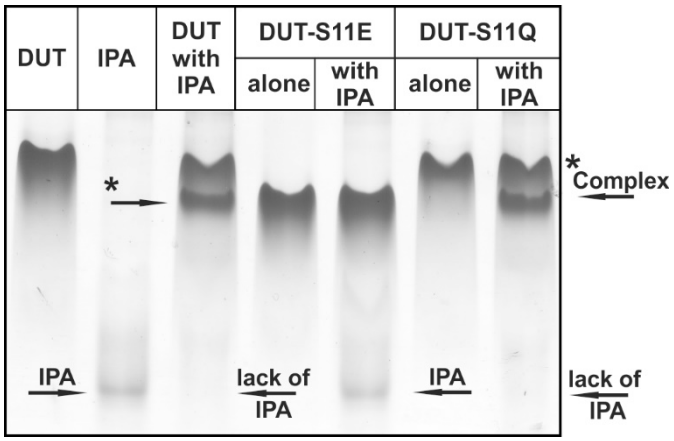
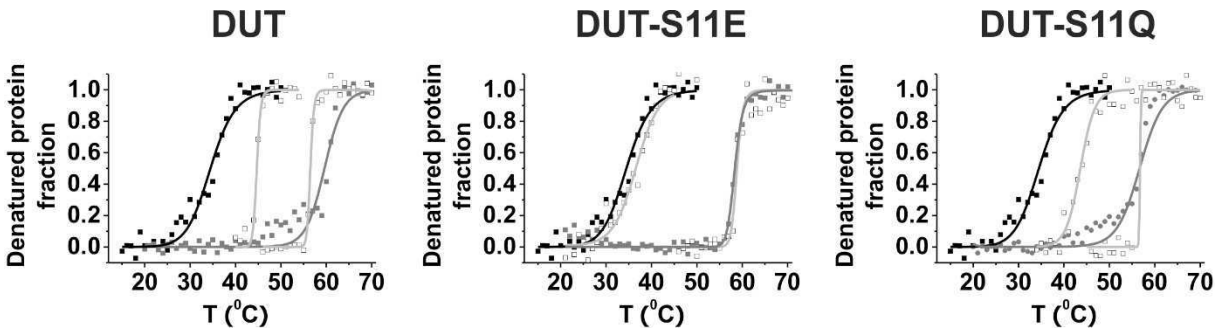


Figure 16. Analysis of dUTPase-importin-α complex formation with native-PAGE analysis
Native gel shows complex formation only between the WT / S11Q dUTPase and importin-α while no new band emerges in case of the S11E mutant. Asterisk shows putative complexes.

4.2.2.3 Thermal denaturation experiments

Complex formation was also confirmed using thermal denaturation experiments of the individual components and their mixtures. Heat-induced unfolding was followed by circular dichroism spectroscopy at 210 nm. Importin-α is stabilized against heat induced unfolding due to the interaction with dUTPase and also renders the thermal unfolding transition more cooperative (Figure 17). We observed this stabilizing effect in the mixtures of importin-α with either WT (shift from 35.7 °C to 43.6 °C) or S11Q mutant dUTPase (35.7 °C to 44.4 °C), but not with S11E mutant dUTPase (shift from 35.7 °C to 36.7 °C). These results again reinforced the previous conclusion that the phosphorylation mimicking negative charge (S11E) impedes

complex formation, whereas the isosteric control glutamine mutant (S11Q) has the same capability to bind to importin-α as the WT protein.



CD measurements *

components	T_m values detected (ºC)		
	DUT	DUT-S11E	DUT-S11Q
dUTPase	58.4 ± 1.6	58.3 ± 0.1	57.5 ± 1.3
importin-α	35.7 ± 1.5	35.7 ± 1.5	35.7 ± 1.5
dUTPase + importin-α	56.4 ± 0.6 / 43.6 ± 1.6	57.4 ± 2.0 / 36.7 ± 0.2	56.5 ± 0.3 / 44.4 ± 1.1

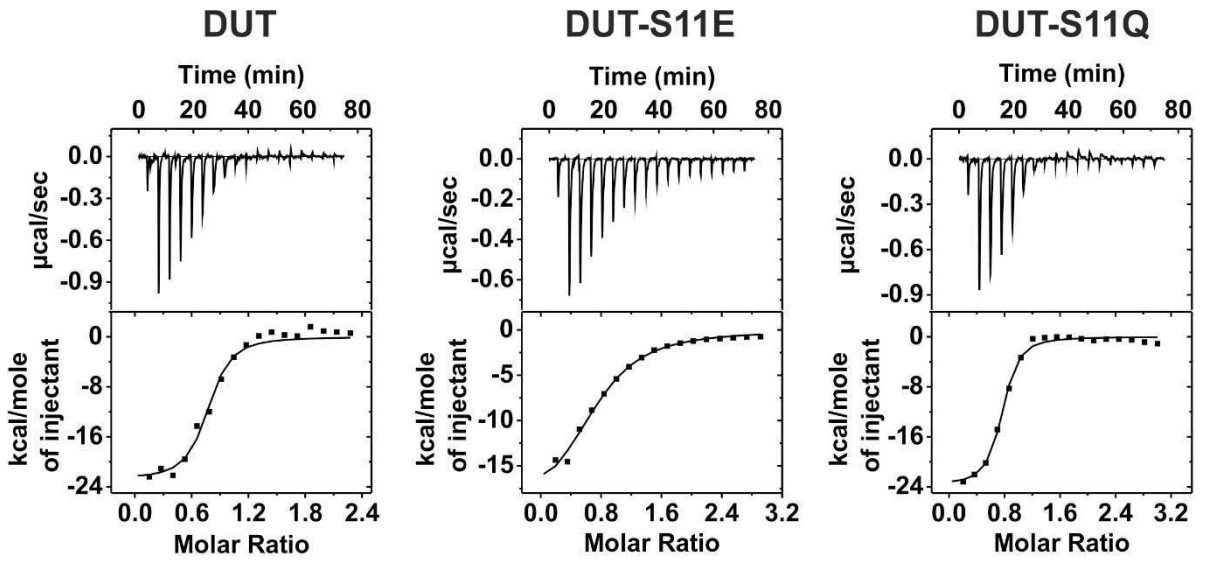
* Values extracted from independent experiments (± standard deviation).

Figure 17. Analysis of complex formation between dUTPase and importin-α using circular dichroism at 210 nm during thermal denaturation

In the denaturation scans curves are shown for importin-α (black squares), dUTPase (gray squares) and their mixtures (empty squares). In case of proteins in complex two transitions were observed, presumably reflecting partially separate unfolding events. Data obtained from the curves are presented in the table underneath.

4.2.2.4 Isothermal titration microcalorimetry

For direct quantitative analysis we performed isothermal titration microcalorimetry (ITC) assays (Figure 18). A change of one order of magnitude in the dissociation constant was observed for the S11E mutant compared to the WT and S11Q dUTPase (K_d was 0.79 μM, 0.76 μM and 9.62 μM for WT, S11Q and S11E dUTPases, respectively). The determined stoichiometry supports a model where each NLS of the trimeric dUTPase binds an importin-α (as the molar binding ratio is ~ 1.0). Thus all NLSs could be bound to importin-α at ideal conditions. It is interesting to note that this one order of magnitude difference in the K_d values translates into an "all or nothing" phonotype *in vivo*, meaning the protein is either nuclear or is completely excluded from there.



ITC measurements *

	DUT	DUT-S11E	DUT-S11Q
N (molar binding ratio)	0.74 ± 0.02	0.73 ± 0.03	0.71 ± 0.01
K_d (µM)	0.79 ± 0.26	9.62 ± 1.54	0.76 ± 0.15
ΔH (kcal mol^{-1})	-22.84 ± 0.90	-20.31 ± 1.31	-23.78 ± 0.56
ΔS (cal mol^{-1} K^{-1})	-50.0	-46.3	-53.1

* Reliability of the fittings are shown by the error values. Protein concentrations correspond to subunits.

Figure 18. Analysis of complex formation between dUTPase and importin-α using isothermal titration microcalorimetry

Top panels present baseline-corrected timeline; bottom panels shows the integrated data (black square) and the fit of the binding isotherm by an "independent binding sites" model (solid line). Values extracted are shown in the table underneath.

4.2.2.5 X-ray crystallography

In order to gain more detail of the complex formed by dUTPase and importin-α we aimed to determine the 3D structural basis of this interaction with X-ray crystallography. Crystal structures of importin-α were determined for the WT and S11E dUTPase NLS peptides (Figure 19 and 20). Both NLS peptides were clearly present in the major NLS binding site of importin-α. Crystallographic data and refinements statistics are summarized in Table A2.

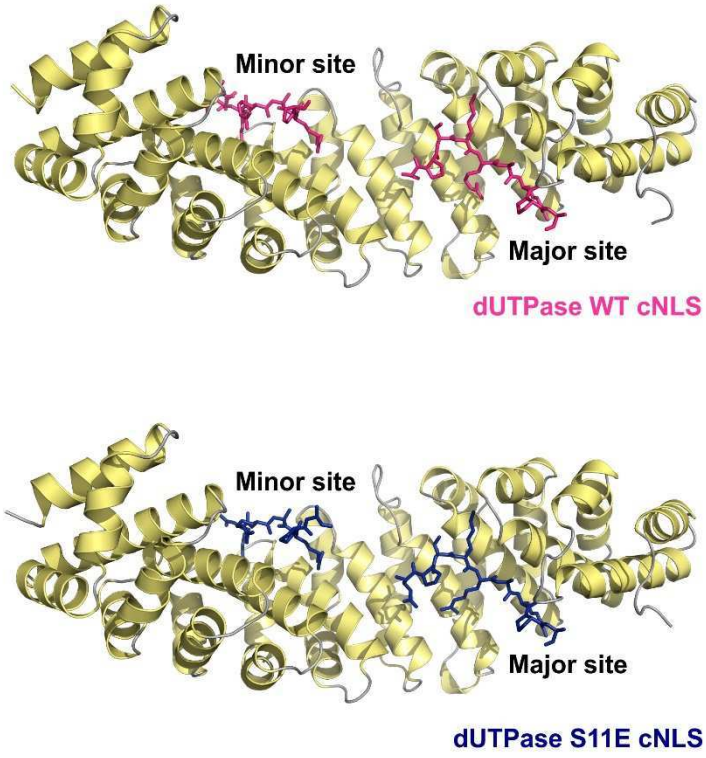
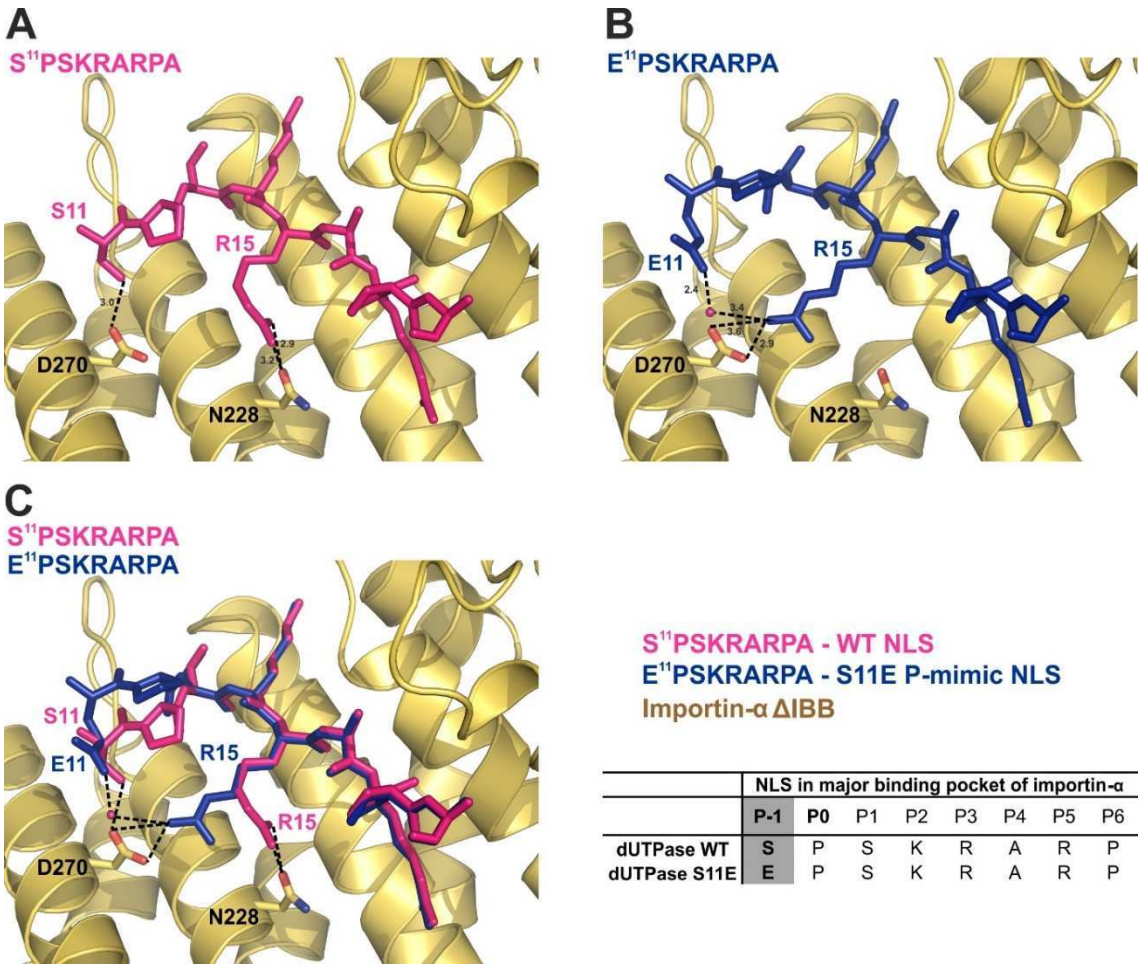


Figure 19. Structural basis of the effect on binding of phosphorylation adjacent to NLSs

Yellow cartoon shows the crystal structures of importin-α ΔIBB in complex with dUTPase WT NLS peptide (upper image, magenta stick) and dUTPase S11E phospho-mimic NLS peptide (bottom image, blue stick). cNLS stands for classical NLS.

The highly conserved Lys14 of the dUTPase peptides (P2 position) form a salt bridge with Asp192 of importin-α, analogous to all previously deposited importin-α:NLS 3D structures. The S11E phospho-mimic mutation uses a similar coordination environment within the complex (especially the C-terminal segment), not leading to major conformational changes in the structure. The introduced glutamic acid side-chain is involved with intra-NLS contact with the Arg15 position of the NLS. This leads to the alternative positioning of the Arg15 side chain in the P3 binding pocket, where it loses the contact with importin-α Asn228 side chain (Figure 20). This conformational shift is stabilized by a potential salt-bridge interaction between dUTPase NLS Arg15 (P3 position) and Asp270 of importin-α with both positions being highly conserved. Importin-α side chains Trp231, Trp273, and Arg238 (also highly conserved residues) undergo slight conformation changes in their side chain positions to accommodate the Arg15:Asp270 interaction. The main-chain of the residues N-terminal to the P2 Lys now follows a different trajectory than that of the WT peptide (Figure 20), resulting in a loss of H-bonds between Pro12 of the phospho-mimic peptide, and the importin-α surface.

The differently arranged accommodation of the S11E peptide within the NLS binding site results in the loss of several interactions between the NLS peptide and the importin-α surface leading to the weaker binding properties of the phosphorylation mimicking NLS.



	NLS in major binding pocket of importin-α							
	P-1	P0	P1	P2	P3	P4	P5	P6
dUTPase WT	S	P	S	K	R	A	R	P
dUTPase S11E	E	P	S	K	R	A	R	P

Figure 20. Close-up views of the interactions for wild-type and mutant NLS segments with importin-α
(A) Close-ups of the interactions for WT dUTPase NLS and **(B)** S11E phosphorylation mimicking NLS. The dUTPase Arg15 side-chain interaction with the Asn228 side chain of importin-α is disrupted in the case of the S11E NLS peptide forming alternative interactions with Asp270. Asn228 and Asp270 are shown in stick representation. **(C)** Superimposition of importin-α:WT dUTPase and importin-α:S11E dUTPase NLS complexes, showing different main-chain trajectories of the NLS peptides N-terminal to the P2 position.

4.2.3 Possible biological significance of dUTPase cell cycle specific localization pattern

We have seen that Cdk1 mediated phosphorylation of dUTPase within the nucleus at the G2/M phase will have a prominent effect on the localization of the protein pool in the daughter cells. dUTPase nuclear import is inhibited until the phosphate moiety is removed from the S11 position resulting in slow dUTPase nuclear re-population in the daughter cells. Nuclear accumulation reaches its maximal extent around the S phase and the protein remains nuclear until the onset of mitosis. Localization pattern of the enzymes involved in *de novo* thymidylate biosythesis proteins have a cell cycle dependent distribution pattern very similar to dUTPase

(detailed in section 1.2.1.1). dUTPase along with its preventive function by maintaining low dUTP/dTTP ratios (thus inhibiting uracil incorporation during replication) also plays a key role in *de novo* thymidylate synthesis since it ensures the substrate dUMP for TYMS [77] to synthesize dTMP. We hypothesize that dUTPase might also be necessary to accompany this enzyme complex into the nucleus for proper genomic DNA maintenance. Thus scheduled nuclear availability of *de novo* thymidylate biosynthesis enzymes, along with dUTPase, may ensure strictly regulated dNTP pool composition for DNA polymerases.

4.2.4 General implications of the effect of phosphorylation on nucleocytoplasmic protein distribution during the cell cycle

Nuclear transport processes regulated by phosphorylation events have several know examples in the literature, so we tried to look up known cases where phosphorylation in the vicinity of NLSs have an inhibitory effect on nuclear transport. The NLS sequences of the found hits were aligned as predicted to bind to the NLS-binding pockets of importin-α [34, 35] (Table 2). Phosphorylation sites were immediately N-terminal to the basic cluster of the NLS segment either located at the P0 or the P-1 positions of the NLS. In most cases the phosphorylation was attributed to the yeast Cdc28 kinase or its human orthologue, Cdk1 (Table 2 and references therein), giving the substrates cell-cycle specific localization pattern. It must be noted though that especially in the case of the closed mitosis of yeast, phosphorylation regulated active nuclear export processes also fine tune the localization pattern. These results argue for the importance of Cdc28-kinase-regulated nuclear transport for several yeast proteins involved in the regulation of cell cycle progression, DNA replication, transcriptional programs, DNA damage recognition and repair. It is tempting to hypothesize that the localization of several human proteins might also fall under the regulation of Cdk1.

Table 2. List of known proteins where phosphorylation inhibit nuclear transport

Predicted binding to the major binding site of importin-α by proteins reported in the literature to be phosphorylated in the vicinity of their NLSs (phosphorylated residues are in italics)

	NLS binding to major binding site of importin-α									Kinase	Ref.
	P-2	**P-1**	**P0**	P1	P2	P3	P4	P5	P6		
SV40 TAg	S^{123}	*T*	P	K	K	K	R	K	V	**Cdk1 / Cdc28**	[152, 153]
Swi6	G^{159}	*S*	P	L	K	K	L	K	I		[152, 154-156]
Swi5	R^{645}	*S*	P	R	K	R	G	R	P		[156-158]
Cdh1	S^{42}	*S*	P	S	R	R	S	R	P		[156, 159]
Fzr1	R^{162}	*S*	P	R	K	P	T	R	K		[160, 161]
Mcm3	K^{764}	*S*	P	K	K	R	Q	K	V		[162]
Acm1	R^{30}	*S*	P	S	K	R	R	S	Q		[156, 163]
Msa1	P^{53}	*S*	P	N	K	R	R	L	S		[156]
Psy4	E^{319}	*T*	P	R	K	R	K	P	T		[156]
Pds1	V^{70}	*S*	P	T	K	R	L	H	T		[156]
Yen1	I^{678}	*S*	P	I	K	K	S	R	T		[156]
Lmnb2	S^{411}	*S*	R	G	K	R	R	R	I	**PKC**	[164]
AhR	Y^{10}	A	*S*	R	K	R	R	K	P		[165]
v-jun	S^{246}	K	*S*	R	K	R	K	L	R		[166]

4.2.4.1 *In silico* screening using integrated bioinformatics tools

Since we mainly found evidence in yeast that phosphorylation in the vicinity of NLSs has an important role in the regulation of overall nuclear import processes, we aimed to perform a human proteome-wide bioinformatics screen to identify human proteins possessing a Cdk1 phosphorylation site at either the P0 or P-1 positions of their NLSs. A combination of two bioinformatics tools were used, one for cNLS prediction (NucImport) and one for phosphorylation site prediction (Predikin). The background frequency of phosphorylation was determined to be 0.136 by Predikin, showing how often a Ser or Thr residue is predicted to be phosphorylated not taking into account the presence of NLS. Then NucImport was applied to predict the position of classical NLSs. The frequency of predicted phosphorylation of the P0 position was 0.393 (p = 2.348e-30) and for the P-1 position is was 0.234 (p = 9.530e-05) meaning both positions are significantly enriched for phosphorylation.

We found 92 proteins with a phosphorylation site at the P0 position and 44 proteins for the P-1 position with a conservative setting of both predictors. If considering only parent proteins (not

all isoforms), this corresponds to 50 and 22, respectively (Table A3). Among the hits, there were numerous findings for which, to our knowledge, no previous experimental data have been reported on phosphorylation-dependent nuclear translocation. Based on Gene Ontology (GO) terms annotations, we found proteins involved in DNA damage recognition and repair, gene expression, epigenetics, RNA-editing and several transcription factors (Table A4). Regulated scheduling of nuclear availability has a clear significance for any of these functions. We therefore selected several hits for experimental validation.

4.2.4.2 Establishing an NLS reporter construct to evaluate NLS function

To validate our hits efficiently, we needed a sensitive model system to test the effect of phosphorylation on nuclear transport. Although several such systems are claimed to exist in the literature, we found that these previously reported constructs lack the efficiency and sensitivity required for in depth investigations. Hence we designed the DsRed-labeled β-galactosidase reporter construct, which is an inert fluorescent cargo core upon which different NLSs can be loaded. Phenotypes obtained this way were used to characterize NLS strength throughout the study. The construct is strictly cytoplasmic, unless fused to a functional NLS, such as the SV40 large T-antigen NLS (Figure 21).

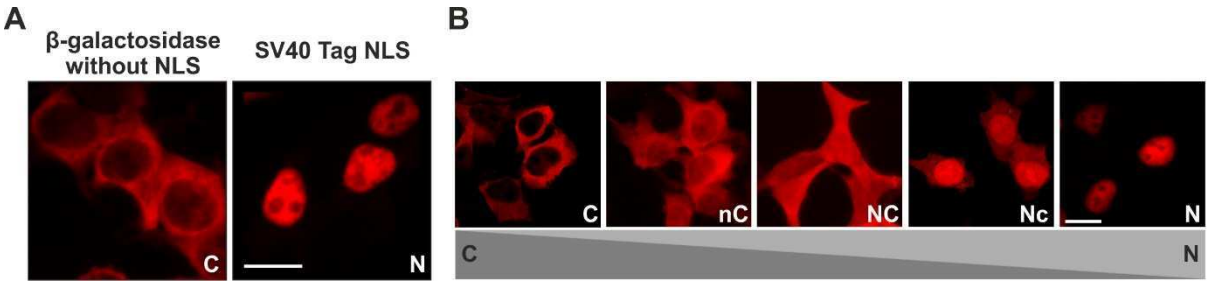


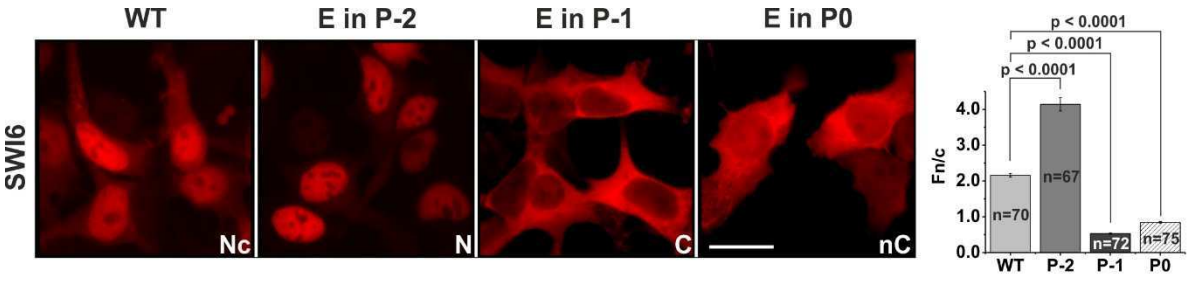
Figure 21. Representation of relative NLS activity

(A) The β-galactosidase-DsRed (pGal-DsRed) reporter construct is strictly cytoplasmic, unless fused to a functional NLS, such as the SV40 large T-antigen NLS. Scale bar represents 20 µm.

(B) The reporter construct was fused to different NLSs resulting in constructs having different degree of nuclear accumulation capabilities. Localization was categorized into five types: N (nuclear), Nc (mainly nuclear), NC (even distribution between the nucleus and cytoplasm), nC (mainly cytoplasmic), C (cytoplasmic). This nomenclature is used throughout the study. Scale bar represents 20 µm.

We tested the system on the well described NLS of Swi6 which is already known to be regulated by an inhibitory phosphorylation on its P-1 position by Cdc28 (see references in Table 2).

Mutations were introduced in a way that the phosphorylation mimicking negative charge was placed either at the P-2, P-1 or P0 positions, while the structurally important proline residue was not perturbed (Figure 22). The results argue that the exact position of the phosphorylated residue is a crucial defining factor of the localization of the NLS-containing cargo. This is in good agreement with previous work proposing that phosphorylation at upstream positions enhance nuclear import while phosphorylation at P0 or P-1 positions impedes it [53].



	NLS in major binding pocket of importin-α									Localization
	P-2	**P-1**	**P0**	P1	P2	P3	P4	P5	P6	
Swi6 wt	A	**S**	P	L	K	K	L	K	I	Nc
Swi6 E in P-1	A	**E**	P	L	K	K	L	K	I	C
Swi6 E in P-2	**E**	P	A	L	K	K	L	K	I	N
Swi6 E in P0	G	A	**E**	P	K	K	L	K	I	nC

Figure 22. Position-specific effect of phosphorylation on NLSs

(A) Images show the NLS reporter system, pGal-DsRed, fused to the WT and mutant NLSs of Swi6. Fn/c ratios (± standard error of the mean) were determined in the indicated amount of cells (n). Significant alteration in nuclear accumulation could be observed among the different constructs. Scale bar represents 20 μm.

(B) Glutamic acid at P-2 or P0 was introduced by insertion of Ala in P0 or deletion of Leu in P1. Sequences were aligned as predicted to bind to the NLS-binding pockets of importin-α.

4.2.4.3 Validation of the selected hits with the NLS reporter construct

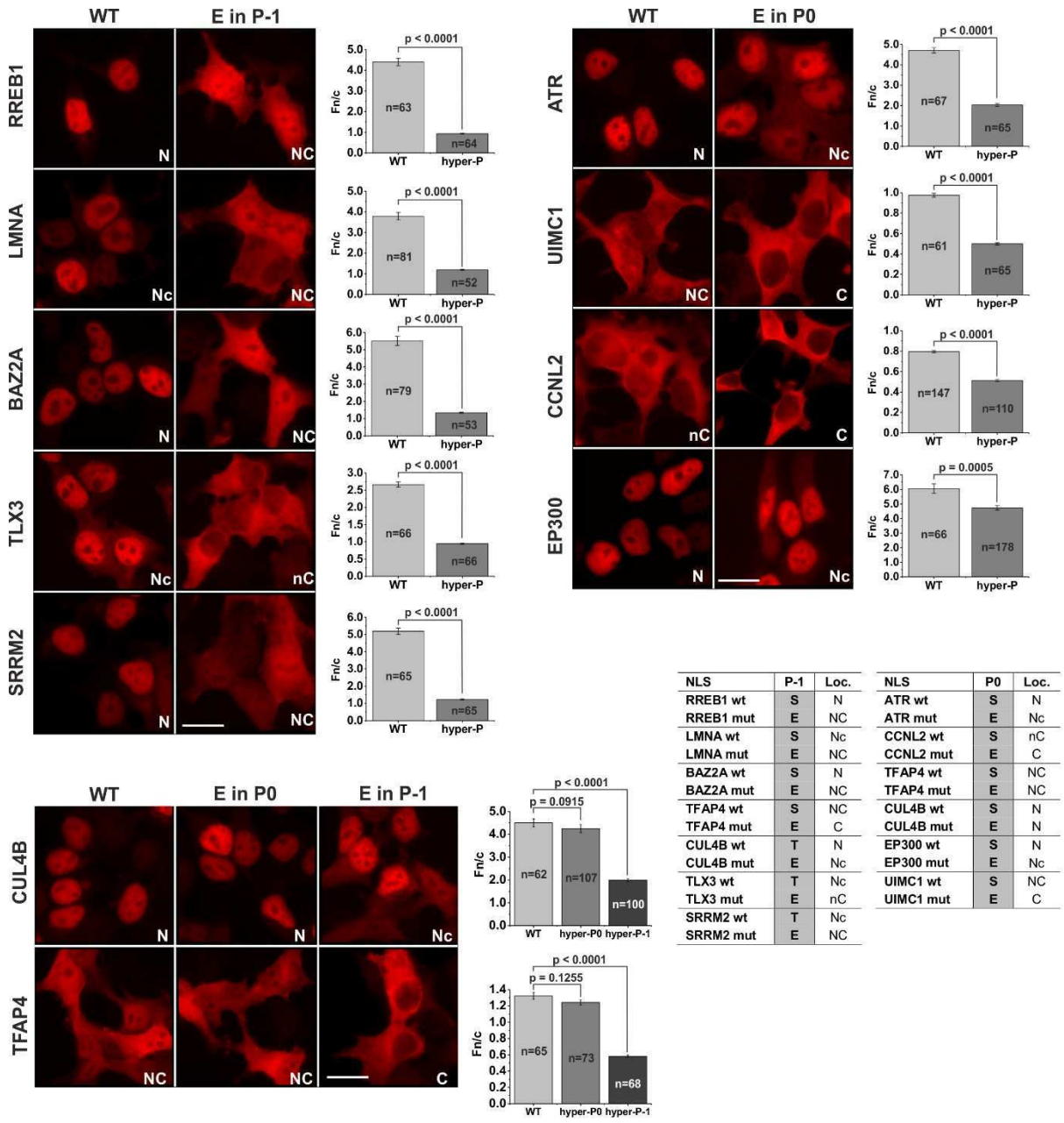
We subjected a selection of proteins involved in a variety of cellular functions found in the initial bioinformatics screening. Six proteins with predicted Cdk1 phosphorylation sites at the P0 position, and seven proteins with predicted Cdk1 phosphorylation sites at the P-1 position (Table 3) were selected.

Table 3. List of the selected proteins that were validated by the NLS reporter system

Proteins for which the phosphorylation of the particular NLS adjacent residues were experimentally confirmed, according to the Phosida database (http://www.phosida.com/) [167, 168] are indicated in italics.

Function	Protein name	Abbreviation	NLS sequence	Ref.
DNA damage recognition and repair	*Ataxia telangiectasia and Rad3-related protein*	ATR	**S**PKRRRLS	[169]
	BRCA1-A complex subunit RAP80	UIMC1	**S**VKRKRRL	[170, 171]
	Cullin-4B	CUL4B	**TS**AKKRKL	[172]
	Transcription factor AP-4	TFAP4	**S**PKRRRAE	[173]
	Histone acetyltransferase p300	EP300	**S**AKRPKLS	[174]
	Ras-responsive element-binding protein 1	RREB1	**S**PLKRRRL	[175]
Regulation of gene expression	Ras-responsive element-binding protein 1	RREB1	**S**PLKRRRL	[176]
	Histone acetyltransferase p300	EP300	**S**AKRPKLS	[177]
	Transcription factor AP-4	TFAP4	**S**PKRRRAE	[173, 178, 179]
Epigenetics	Histone acetyltransferase p300	EP300	**S**AKRPKLS	[180]
	Bromodomain adjacent to zinc finger domain protein 2A	BAZ2A	**S**PSKRRRL	[181, 182]
	Cullin-4B	CUL4B	**TS**AKKRKL	[183]
RNA editing/splicing	Ser-Arg repetitive matrix protein 2	SRRM2	**T**PAKRKRR	[184]
	Cyclin-L2	CCNL2	**S**PKRRKSD	[185]
Cell cycle regulation	Cullin-4B	CUL4B	**TS**AKKRKL	[186-188]
	Cyclin-L2	CCNL2	**S**PKRRKSD	[185, 189]
Development	T-cell leukemia homeobox protein 3	TLX3	**T**PPKRKKP	[190]
	Transcription factor AP-4	TFAP4	**S**PKRRRAE	[173]
Nuclear skeleton	*Lamin A*	LMNA	**S**VTKKRKL	[191, 192]

The efficiency of nuclear targeting differed among the different NLSs (Figure 23). Because the permeability of our reporter construct is the same (depending on size and molecular shape), the steady-state $F_{n/c}$ ratios can be used as a measure of the import rate itself. However the resulting localization data show that in each case, replacement of the appropriate Ser or Thr residues (in P-1 or P0 position), predicted to be phosphorylated by Cdk1, by Glu in the NLSs always leads to significantly weaker nuclear accumulation or even to complete nuclear exclusion, compared to the WT. These results confirmed the validity of the *in silico* analysis and strengthened the observation that the P-1 and P0 position specific phosphorylations have an inhibitory effect on nuclear translocation also in a human model system.



NLS	P-1	Loc.
RREB1 wt	S	N
RREB1 mut	E	NC
LMNA wt	S	Nc
LMNA mut	E	NC
BAZ2A wt	S	N
BAZ2A mut	E	NC
TFAP4 wt	S	NC
TFAP4 mut	E	C
CUL4B wt	T	N
CUL4B mut	E	Nc
TLX3 wt	T	Nc
TLX3 mut	E	nC
SRRM2 wt	T	Nc
SRRM2 mut	E	NC

NLS	P0	Loc.
ATR wt	S	N
ATR mut	E	Nc
CCNL2 wt	S	nC
CCNL2 mut	E	C
TFAP4 wt	S	NC
TFAP4 mut	E	NC
CUL4B wt	S	N
CUL4B mut	E	N
EP300 wt	S	N
EP300 mut	E	Nc
UIMC1 wt	S	NC
UIMC1 mut	E	C

Figure 23. Evaluation of the proteins identified by computational analysis: cellular screens for NLS activity

Validation of the localization pattern of the selected hits. NLSs were loaded into the pGal-DsRed reporter system, and localization was adressed in 293T cells. Graphs show $F_{n/c}$ ratios for the indicated constructs determined in the indicated amount of cells (n). Phosphorylation mimicking mutations at the appropriate Ser/Thr position significantly reduced nuclear accumulation. Table summarizes localization pattern of the different constructs and their mutants. Scale bar represents 20 µm.

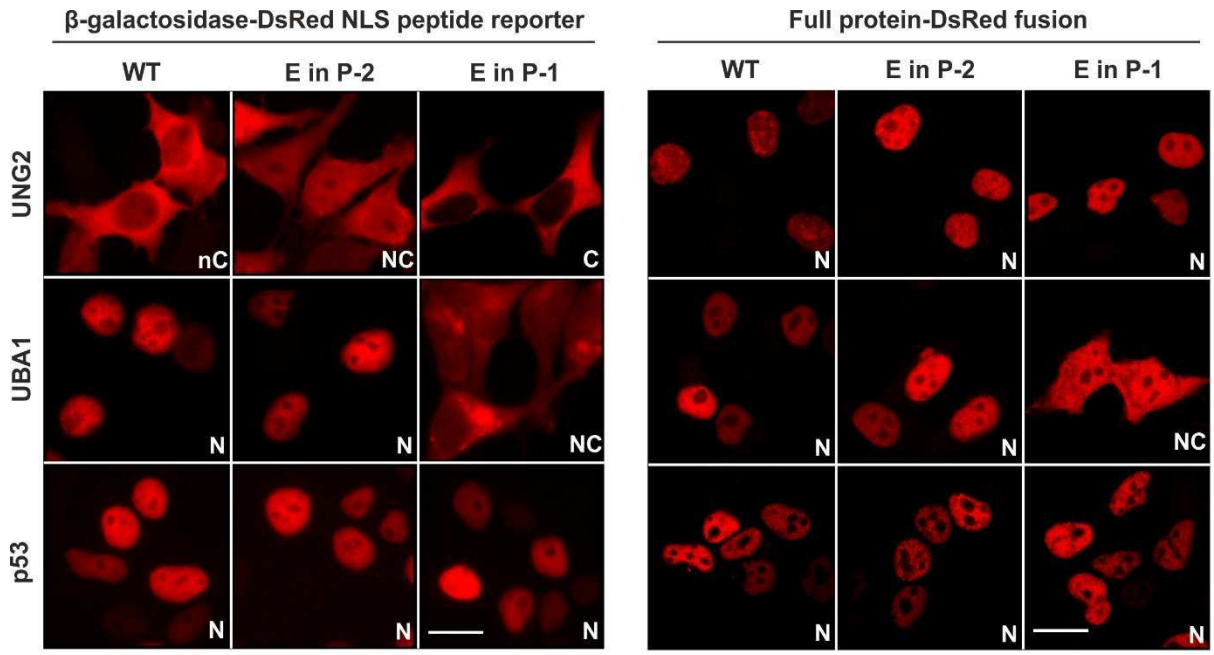
4.2.4.4 Possible biological significance of the cell cycle specific localization pattern

Our screening has identified a number of human proteins potentially sharing a similar Cdk1-driven regulatory pattern (Table 3 and Table A3) as dUTPase, which are involved in a variety of cellular functions such as DNA damage recognition and repair, transcriptional regulation, cell cycle control, epigenetics and RNA editing. Clearly, for proteins involved in such functions, the fine-tuned regulation of nuclear availability is of high significance. Regulation of cell cycle by cullin-4B along with cyclin-L2 could be an example for this, which latter is also a hub among pre-mRNA splicing, apoptosis induction and cell-cycle arrest pathways in cancer cells. Many of our hits act in an interconnected manner; for example, during DNA damage, the protein kinase ATR phosphorylates BAZ2A, UIMC1, RREB1, and SRRM2 [193]. For proteins involved in DNA repair (ATR, CUL4B, UIMC1, TFAP4, EP300), the tight connection between cell-cycle checkpoints and DNA damage recognition and repair may be the underlying reason for their scheduled absence or presence within the nucleus [194]. We found two examples where an SNP might overwrite the Cdk1 driven regulation, however, no data for either frequency or physiological relevance of these SNPs has been reported to date (Table A3).

4.2.4.5 Other known Cdk1 substrates having phospho-NLSs with no known effect

There are some examples in the literature where NLS-adjacent Cdk1-driven phosphorylation has been reported previously but its actual effects were only addressed by using hypo-P mimicking mutants. Therefore, we checked the effect of phosphorylation mimicking mutations at the previously established Cdk1 sites of UNG2 (S14) [195, 196], UBA1 (S4) [197] and p53 (S315 and S312 in human and in mouse, respectively) [198] (Figure 24). These phosphorylation sites are predicted to be located in the P-2 position. For localization studies the NLSs were checked in our pGal-DsRed NLS reporter construct while the full length ORFs were fused with DsRed-monomer. As expected based on our previous findings, a negative charge introduced at this position did not abolish the nuclear localization of either of these constructs (both with the NLS reporter and the full length proteins). However, if we moved this negative charge to the P-1 position with mutagenesis, the perturbation of nuclear import is clearly observable in case of UBA1 and UNG2 (Figure 24) with the NLS reporter constructs. With UNG2 it is clearly visible that the nuclear localization is enhanced with the S14E mutant (in P-2 position). So phosphorylation of the P-2 positions of these proteins might enhance nuclear accumulation after mitosis. The nuclear transport of UNG2 does not solely rely on its NLS (S^{14}PARKRHA) as other sequence motifs also have a role in proper localization of UNG2 [105]. This might be the reason why the full length UNG2 does not have the same localization pattern as its NLS reporter

construct. p53, which harbors a bipartite NLS sequence, might have the flexibility to compensate these negative effects by the additional binding to the minor NLS-binding site.



	minor pocket					NLS in major binding pocket of importin-α							Localization	
	P1'	P2'	linker	P-2	P-1	P0	P1	P2	P3	P4	P5	P6	reporter	Full ORF
UNG2 wt				**S**	P	A	R	K	R	H	A	P	nC	N
UNG2 E in P-2				**E**	P	A	R	K	R	H	A	P	NC	N
UNG2 E in P-1					**E**	P	R	K	R	H	A	P	C	N
UBA1 wt				**S**	P	L	S	K	K	R	R	V	N	N
UBA1 E in P-2				**E**	P	L	S	K	K	R	R	V	N	N
UBA1 E in P-1					**E**	P	S	K	K	R	R	V	NC	NC
P53 wt	K	R	…	**S**	P	Q	P	K	K	K	P	L	N	N
p53 E in P-2	K	R	…	**E**	P	Q	P	K	K	K	P	L	N	N
p53 E in P-1	K	R	…		**E**	P	P	K	K	K	P	L	N	N

Figure 24. Effect of Cdk1 phosphorylation on UNG2, UBA1 and p53 localization

UNG2 (residue S14), UBA1 (residue S4) and p53 (residue S315) phosphorylation at the P-2 positions of their NLSs were mimicked by Glu substitution in the pGal-DsRed reporter system or were mutated in full length proteins fused to DsRed-monomer. Additional mutations were carried out to move the negative charge introduced by Glu to the P-1 position. Localization was tested in 293T cells. Scale bar represents 20 μm.

These results also draw our attention to the fact that the full-length protein context may modulate the basic effects of NLS phosphorylation, and should be addressed for each of our findings individually. The cell cycle-dependent changes in the nuclear proteome have an important role in selecting the correct set of proteins to be present in the nucleus during the different stages of the cell cycle. We hypothesize that Cdk1 has an important role in regulating nuclear proteome re-establishment after cell division through inhibitory phosphorylation of P-

1 and P0 positions in NLSs or through enhancing nuclear transport by P-2 position phosphorylation. Several proteins will only appear in the nucleus of the daughter cells after a significant delay following dephosphorylation or *de novo* protein synthesis. Dephosphorylation in the cytoplasm may require some time and might be under yet another level of regulation, giving cells further plasticity in setting up their nuclear proteome composition. The exact purpose of this regulation might differ for each protein, and should be checked individually in detail.

4.3 NLS copy number variation governs efficiency of nuclear import: case study on dUTPases

Intracellular localization of the nuclear isoform of the human dUTPase seems to be regulated by several different ways. We have seen thus far that calpain activation dependent proteolysis along with cell cycle dependent phosphorylation by Cdk1 limits dUTPase presence in the nucleus. Most dUTPases are homotrimeric enzymes, having one NLS on each subunit. This arrangement might give an opportunity to fine tune the already established ways of nuclear transport regulation. Namely each NLS could be altered individually meaning each could be inactivated through phosphorylation by Cdk1 or cleaved off by calpains. It is interesting to speculate whether dUTPases with different NLS copy number might have different efficiency in nuclear targeting. Having this interesting question in mind, here we focus on the peculiar occurrence of a pseudo-heterotrimeric dUTPase in *Drosophila virilis* that contains only one NLS and investigate its localization pattern as compared to the homotrimeric dUTPase isoforms of *Drosophila melanogaster*. Although the interaction of individual NLSs with importin-α has been well characterized, the question how multiple NLSs of oligomeric cargo proteins could affect their trafficking has been less frequently addressed in adequate detail. We tried to exploit the exceptional characteristics of *Drosophila* dUTPases to evaluate the regulatory potential of NLS copy number in nuclear import.

4.3.1 *D. virilis* has a single dUTPase isoform, corresponding to a pseudo-heterotrimer termed 'ABC'

In *D. virilis*, genome annotations (Gene symbol: Dvir_GJ10455, Gene ID: 6627974) indicated the presence of a unique dUTPase where three copies of the dUTPase subunits are covalently linked in one continuous polypeptide (termed ABC pseudo-heterotrimer) with only one NLS on the N-terminal. Figure 25 shows the comparison of domain arrangement of the dUTPases from *D. melanogaster* and *D. virilis*, indicating the (potential) NLS segment (verified for *D. melanogaster* dUTPase in [110]), as well as the five characteristic conserved motifs of dUTPases (marked with roman numbers in the schematics).

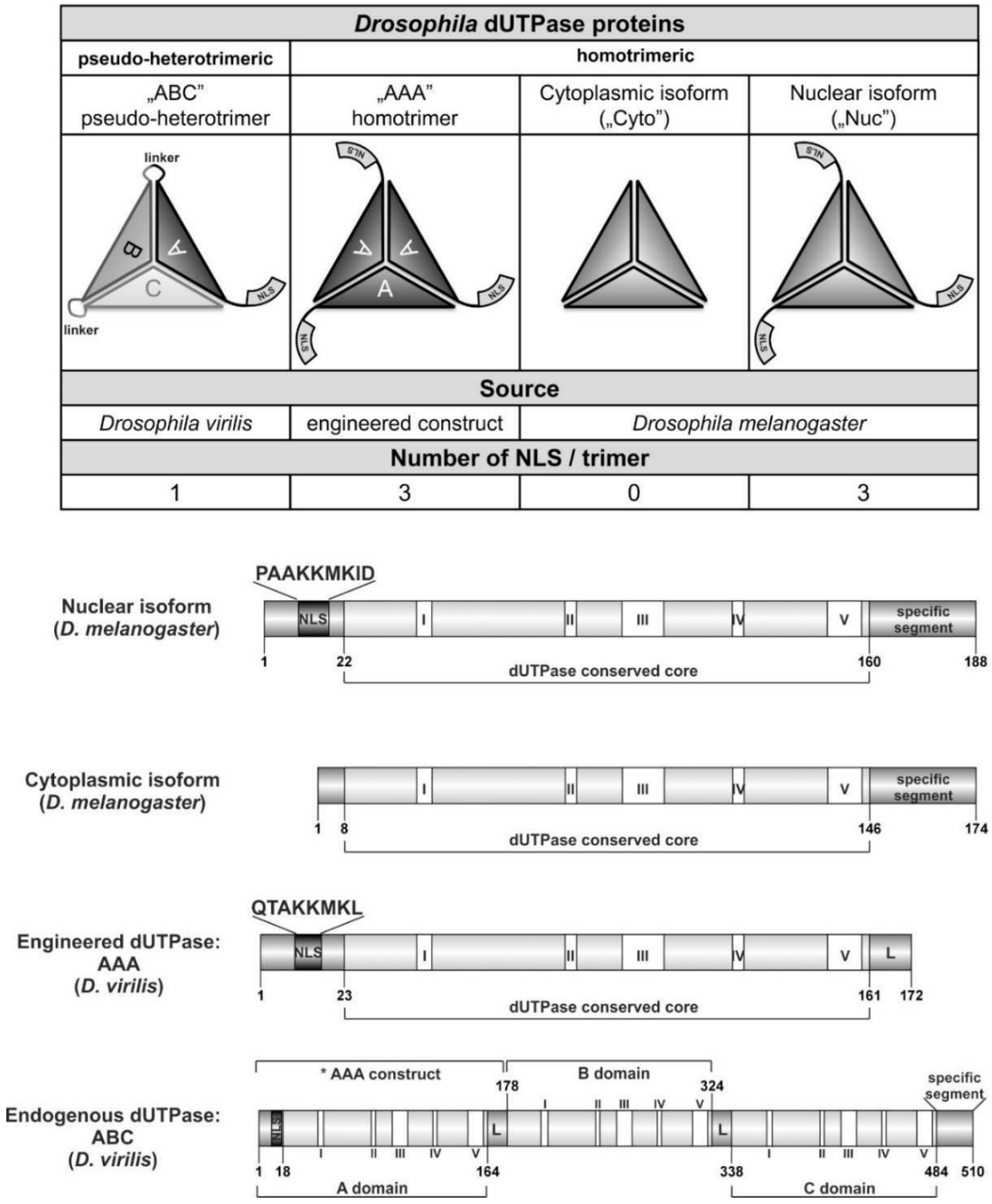


Figure 25. Schematic representation of the dUTPase constructs used in this study

Two homotrimeric dUTPase isoforms are present in *D. melanogaster*: a cytoplasmic isoform (termed "Cyto") lacking the NLS signal, and a nuclear isoform ("Nuc") containing NLSs on all three subunits. *D. virilis* is predicted to possess a unique dUTPase, consisting of three covalently linked dUTPase domains, forming a pseudo-heterotrimer (termed "ABC") harboring only one NLS located at the N-terminus, as an extension to the "A" dUTPase domain. An artificial dUTPase construct (expected to form a homotrimer, termed "AAA") was engineered, having three NLS signals. Schematics also show the position and the sequence of the NLSs along with the five characteristic conserved motifs of dUTPases (white boxes and roman numbers respectively). The specific C-terminal segment is also highlighted present only in *Drosophila* dUTPases. The linker region connecting the ABC pseudo-heterotrimer is labeled with an 'L'.

To check the validity of the annotation, we looked at the mRNA and protein levels of *D. virilis* dUTPase. *In silico* prediction of splice sites with the Human Splicing Finder (Version 2.4.1) for the dUTPase gene in *D. melanogaster* (Gene symbol: Dmel_CG4584) gives a high probability hit for a splice site at the 5' end. This is in accordance with currently available data regarding *D. melanogaster* transcript variants, with a nuclear (NM_135635.3) and a cytoplasmic (NM_001273442.1) isoform (Table A5). We have also found potential splice sites with lower probability in the *D. virilis* dUTPase gene, which could result in alternative splice variants, one lacking while the other having a single NLS (Table A5). We checked this possibility at both mRNA and protein level. The 5' RACE (used to detect possible 5' end splicing variants) products in Figure 26 show that both *D. virilis* embryos and ovary contain only one detectable dUTPase band, suggesting the presence of only one isoform. 5' RACE products were subcloned and analyzed by sequencing. Based on the sequencing results we could conclude that only one isoform is expressed *D. virilis* which harbors the single NLS sequence on its N-terminal (Table A6). This was also confirmed at protein level since the *D. virilis* cell line (WR-Dv-1) contains only one dUTPase isoform, which we term ABC dUTPase, in contrast to the two isoforms of *D. melanogaster.* The single isoform present in *D. virilis* is associated with an apparent molecular mass of 59 kDa (Figure 26), in good agreement with the calculated molecular mass of the pseudo-heterotrimeric ABC dUTPase (55.7 kDa). *D. virilis* dUTPase detection was achieved by using the polyclonal antibody raised against *D. melanogaster* dUTPase [109] which is possible presumably due to the very close similarity of these two proteins (identity 73.82%, similarity 93.08%). Antibody specificity was confirmed by using purified recombinant *D. virilis* dUTPase constructs (Figure 26).

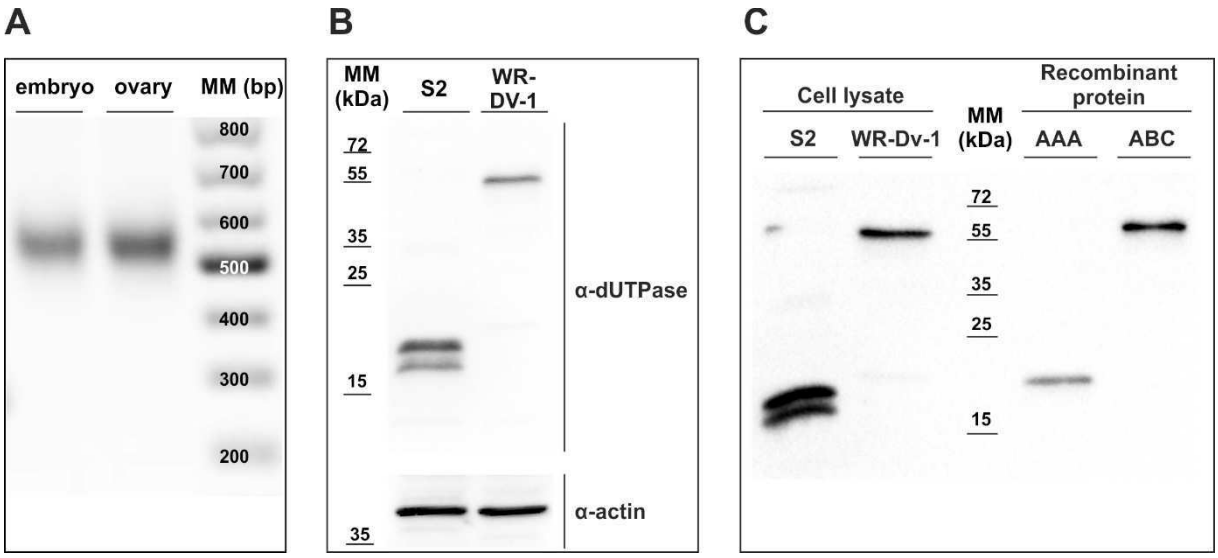


Figure 26. Detection of possible D. virilis dUTPase isoforms with 5' RACE and at protein levels

(A) In embryos and ovaries, 5' RACE results show a single dUTPase mRNA variant present. This corresponds to the pseudo-heterotrimer (ABC) sequence containing one NLS verified by sequencing (Figure S3). **(B)** Western blot analysis of *D. melanogaster* Schneider S2 and *D. virilis* WR-Dv-1 cell line. In S2 cells contain the nuclear [~23 kDa] and cytoplasmic [~21 kDa] dUTPase isoforms. WR-Dv-1 cells only have one band corresponding to an apparent molecular mass of 59 kDa as expected for the pseudo-heterotrimeric ABC dUTPase. Loading control of actin is also shown. **(C)** Polyclonal antibody raised against *D. melanogaster* dUTPase also reacts with the *D. virilis* dUTPase. The antibody specifically recognizes the endogenous ABC pseudo-heterotrimer in the WR-Dv-1 lysate since the recombinant protein gives a signal in the same height. Also the engineered AAA form (disintegrated into monomers in the gel) has a similar mass as the nuclear isoform of *D. melanogaster*.

In order to address the effect of NLS copy number on dUTPase localization we designed dUTPase constructs which mainly differed only in the number of NLSs they possessed (Figure 25). The 'A' subunit of the ABC pseudo-heterotrimer was cloned separately, which is expected to form the characteristic dUTPase homotrimer (AAA homotrimer, not present physiologically) possessing three NLSs. For comparison we used the two physiologically present isoforms of *D. melanogaster* dUTPase, corresponding to the NLS-lacking cytoplasmic and the NLS-containing nuclear isoforms, and the physiologically relevant ABC pseudo-heterotrimer of *D. virilis*. These *D. melanogaster* isoforms also form homotrimers, as shown previously [109, 112, 199, 200].

4.3.2 Nuclear targeting capability of dUTPases depend on the number of NLSs possessed

4.3.2.1 dUTPase localization pattern in WR-Dv-1 cell line

Localization studies were performed in a *D. virilis* derived cell line, WR-Dv-1 image where the same acquisition settings were applied for each construct. EGFP-labeled dUTPase constructs were transfected into cells and assayed for localization pattern (Figure 27). EGFP was applied as a control for homogenous cellular distribution, since its molecular mass of 26.94 kDa permits passive diffusion through the nuclear pore complex. The EGFP-labeled *D. melanogaster* isoforms, however, show organelle-specific distribution, either strictly in the cytoplasm or in the nucleus. The three NLS containing *D. melanogaster* isoform is exclusively nuclear. Similar strict nuclear localization could be observed for the homotrimeric *D. virilis* AAA dUTPase (also containing three NLSs). However, the one NLS containing *D. virilis* ABC pseudo-heterotrimer, is present in both compartments. Quantitative organellar distribution of the respective proteins were also determined by measuring the ratio of the relative amount of EGFP fluorescence present in the nucleus and the cytoplasm (Figure 27B). If there is no direct preference for the nuclear or the cytoplasmic compartment (homogenous distribution), the Fn/c ratio is around 1, as seen for EGFP (Fn/c ratio: 1.23) when expressed alone. The exclusive distribution of the *D. melanogaster* isoforms is also very well reflected in the Fn/c ratios of around 5.10 and 0.81 for the nuclear and cytoplasmic dUTPases, respectively. AAA construct with three NLSs is more predominant in the nucleus (Fn/c ratio: 3.58), while the ABC protein with only one NLS shows considerable presence in the cytoplasm (Fn/c ratio: 2.03) having statistically significant difference in the cellular distribution among the two. As controls we have also made NLS-lacking constructs (ΔNLS-AAA and ΔNLS-ABC). This shows that the indicated segments (identified from sequence similarities) are functional NLSs, since these constructs are less nuclear compared to their pairs with not having the deletion. Compared to the nuclear *D. melanogaster* dUTPase isoform, the AAA construct has a smaller Fn/c ratio. *D. virilis* lacks the N-terminal proline residue as well as the C-terminal isoleucine-aspartate dipeptide at the two ends of its NLS which were shown previously to have great contribution to effective nuclear targeting in *D. melanogaster* [110]. Figure 27C shows that the protein level of AAA and ABC constructs were similar in the imaged samples.

AAA and ABC proteins were also tagged with AU1 (sequence: DTYRYI), which is much smaller compared to EGFP, to exclude potential perturbing effects of the size of the tag used (Figure 28A). We saw no difference in the localization pattern of the constructs when comparing to the EGFP-tagged versions. This is also reflected in the almost same Fn/c ratios for AAA-AU1 and ABC-AU1 proteins (3.61 and 2.02 respectively) as the GFP tagged constructs (Figure 28B).

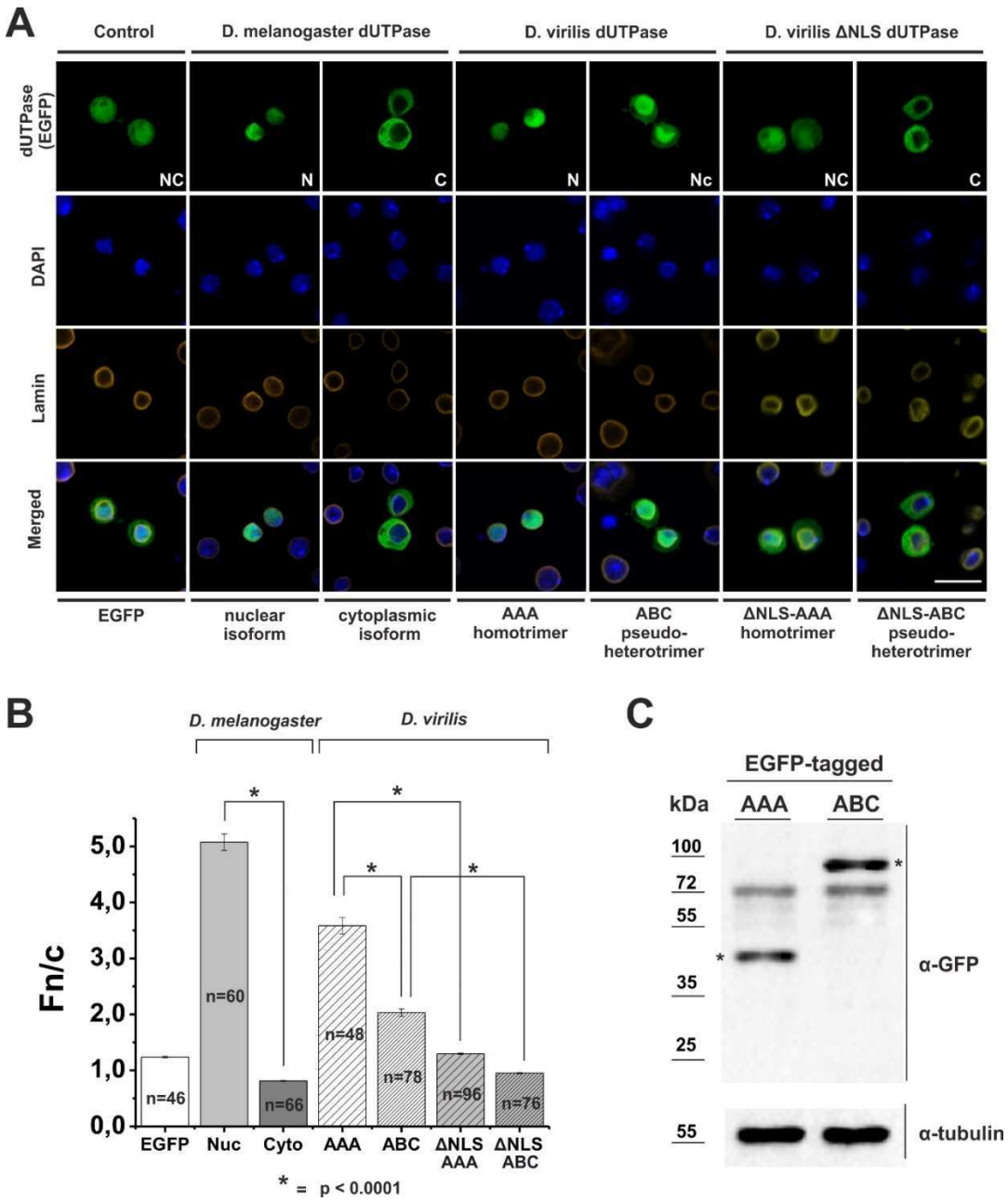


Figure 27. NLS copy number dependent localization pattern of EGFP-tagged *Drosophila* dUTPase constructs

(A) Images were taken of WR-Dv-1 cells. *D. melanogaster* isoforms are strictly nuclear (N) or cytoplasmic (C). In the case of *D. virilis*, the engineered "AAA" homotrimer with three NLS sequences is exclusively nuclear (N) while the "ABC" pseudo-heterotrimer with only one NLS is mainly present in the nucleus but is also observable in the cytoplasm (Nc). Construct lacking the NLS are less nuclear or fail to enter the nucleus. DNA was stained with DAPI and the nuclear envelope was visualized by lamin Dm0 staining. Scale bar represents 20 µm. **(B)** Relative subcellular localization of the constructs are given in the graph showing the Fn/c ratios (± standard error of the mean), where 'n' gives the number of analyzed cells. **(C)** Western blot showing similar total level of the AAA and ABC constructs used during imaging. Membrane was also developed against tubulin as loading control. Asterisks indicate the specific band (identified by the correct molecular masses).

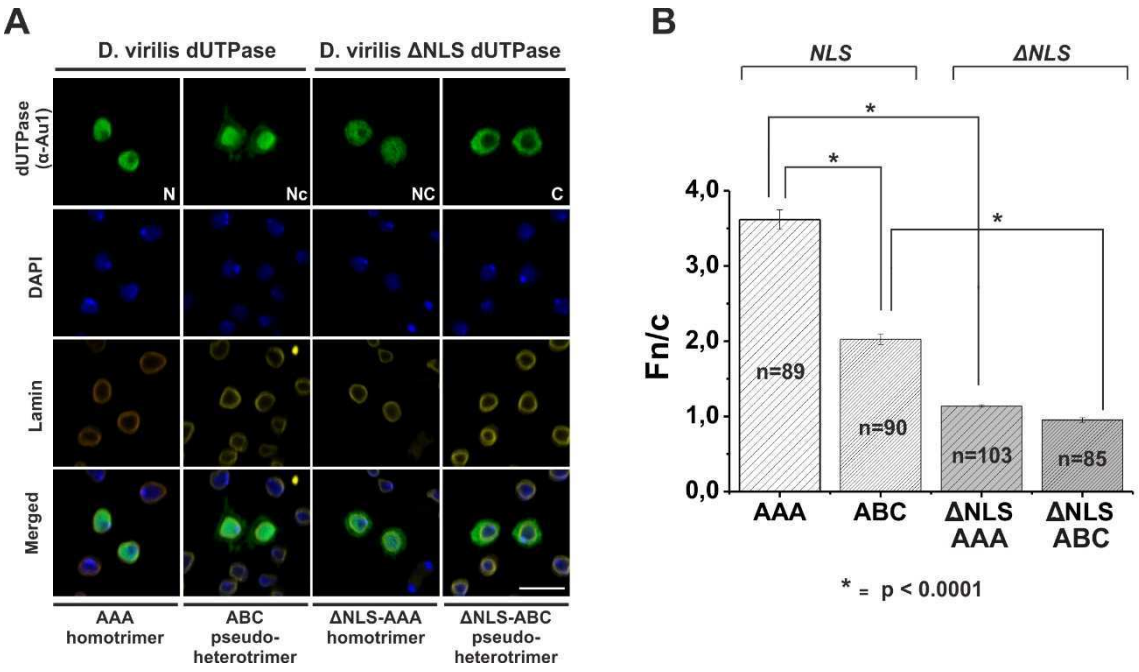


Figure 28. NLS copy number dependent localization pattern of AU1-tagged *Drosophila* dUTPase constructs

(A) The intracellular distribution of AU1-tagged *D. virilis* and *D. melanogaster* dUTPase constructs were visualized in WR-Dv-1 cells. To aid visual inspection of the localization pattern, DNA was stained with DAPI and the nuclear envelope was visualized by lamin Dm0 staining. Scale bar represents 20 µm.

(B) Image analysis to quantify relative subcellular localization of AU1-tagged constructs were also performed. Fn/c ratios (± standard error of the mean) are shown in the graph, where 'n' gives the number of analyzed cells.

4.3.2.2 Localization studies in mammalian cell lines

The cellular distribution patterns observed in insect cell lines were also investigated in three different mammalian cell lines (Figure 29). The high conservation of the nuclear import machinery [201, 202] allows us to investigate the cellular distribution patterns observed in insect cells also in mammalian cells. There is a clear advantage of using epithelial or fibroblast like mammalian cells since they have more easily observable nuclear and cytoplasmic compartments. Also we wished to address the question whether NLS copy number differences have a significant effect on nuclear targeting potency in mammals. When comparing the EGFP labeled AAA and ABC *D. virilis* constructs the same localization pattern can be seen as observed in the WR-Dv-1 cell line. This also indicates the universal character of the effect of NLS copy number variation on nuclear targeting efficiency.

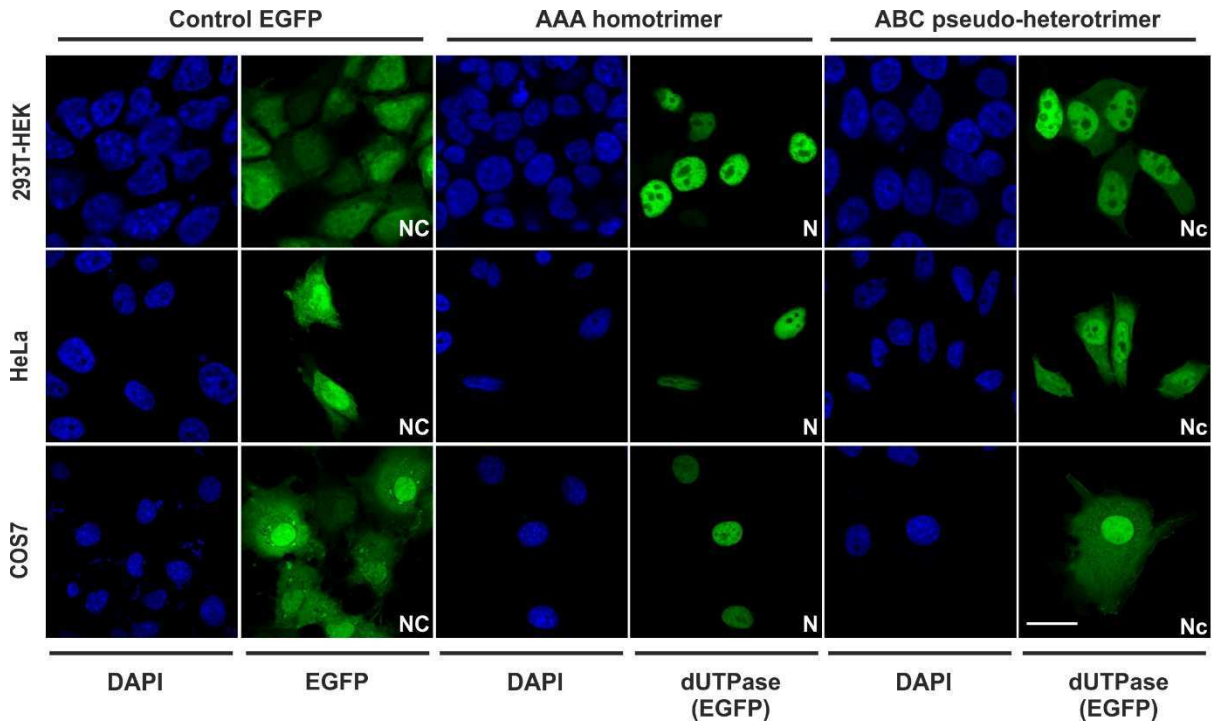


Figure 29. NLS copy number dependent localization pattern of EGFP-tagged *Drosophila* dUTPase constructs in mammalian cells

Distribution of the EGFP-tagged *Drosophila* dUTPase constructs were addressed in 293T-HEK, HeLa and COS7 cell lines. AAA dUTPase construct is solely nuclear (N), while the ABC pseudo-heterotrimeric construct is predominantly nucleular but is also observable in the cytoplasm (Nc). The control EGFP is evenly distributed between the nucleus the. Scale bar represents 20 µm.

4.3.2.3 Localization of endogenous proteins

D. virilis ABC, AAA constructs and the *D. melanogaster* dUTPase isoforms showed discrete characteristics regarding their localization pattern in cell lines overexpressing these constructs. However, we also wished to compare the localization of the physiologically occurring endogenous dUTPases (Figure 30) among *D. melanogaster* and *D. virilis* cell lines and tissues. To aid visual inspection DNA was stained with DAPI, to outline the nucleus, along with lamin (Dm0) staining, to highlight the nuclear envelope, and F-actin staining to outline the cell membrane and the cellular cytoskeleton. dUTPase is present in both the cytoplasmic and nuclear compartment with the latter being more pronounced both in tissues and cell lines. This observation is consistent among the two species however it is achieved by different mechanisms. In *D. melanogaster* two isoforms, that are strictly dedicated to the respective compartments, are required to maintain the balanced presence of dUTPase in the nucleus and the cytoplasm. In *D. virilis*, however, the single ABC dUTPase can provide enzyme presence

in both compartments. The weakened specificity of cellular distribution for the ABC dUTPase may be considered to be an advantage allowing nucleocytoplasmic distribution with just one protein isoform.

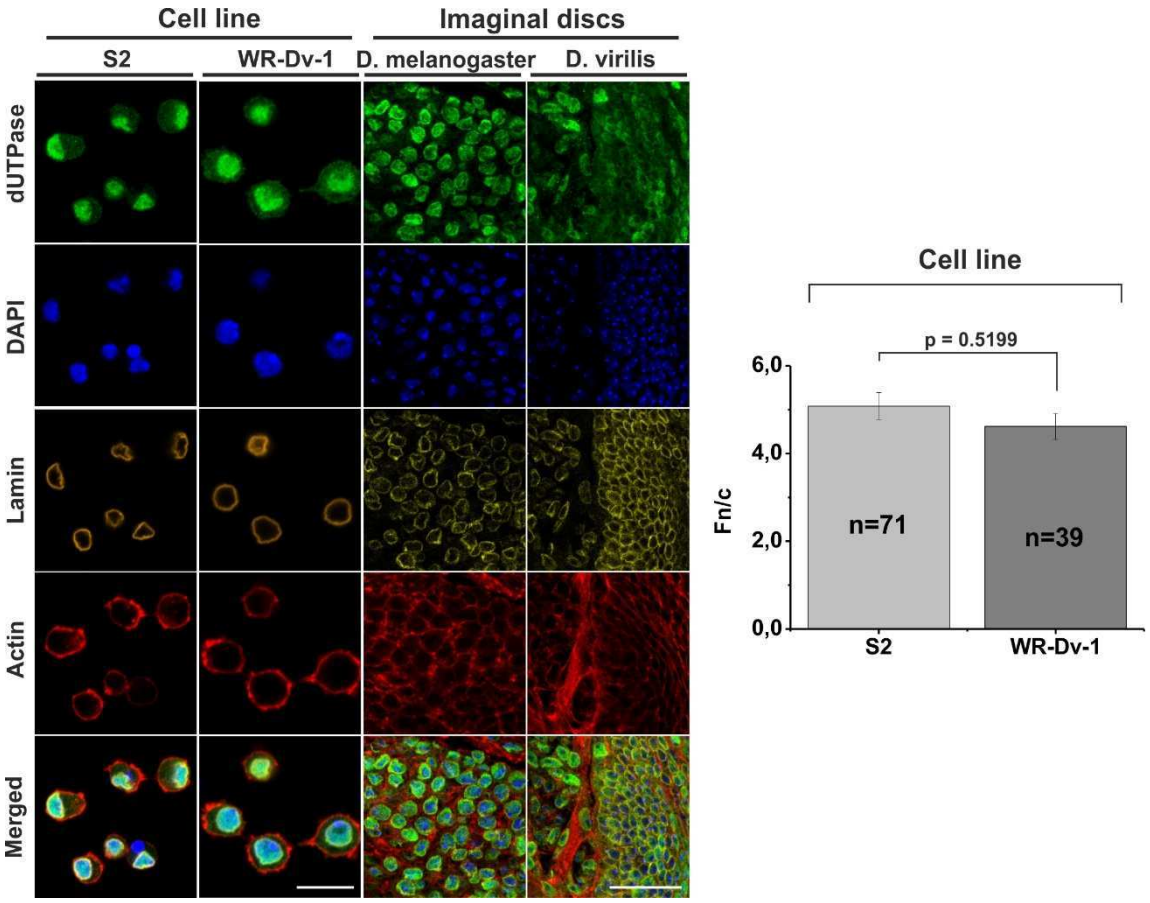


Figure 30. Comparison of localization pattern of endogenous dUTPases in D. melanogaster and D. virilis cell lines and tissues

Endogenous dUTPase isoforms were stained in both *D. melanogaster* (S2) and *D. virilis* (WR-Dv-1) derived cell lines as well as in *D. melanogaster* and *D. virilis* larval wing discs. To aid visual inspection, DNA was stained with DAPI, the nuclear envelope was visualized by lamin Dm0 staining, and cell boundaries are highlighted by phalloidin-TRITC staining for F-actin. Image analysis to quantify relative subcellular localization in the cell lines were also performed. Fn/c ratios (± standard error of the mean) are shown in the graph, 'n' denotes the number of cells analyzed. Scale bar represents 20 µm.

4.3.3 NLS copy number dictates stoichiometry in importin-α:dUTPase complexes

In case of the nuclear isoform of the human dUTPase we have seen that each NLS is capable of importin-α simultaneously. However was not evident since earlier publications based on small angle X-ray scattering measurements suggest that the three N-terminal flexible segment

of the subunits (each containing the NLS) seems to be protruding in one direction [199]. So it could be possible that not all NLSs are accessible for importins at the same time. We used different biophysical methods to characterize the interaction of AAA and ABC dUTPases with importin-α, a major karyopherin protein responsible for nuclear import of classical NLS-containing cargo proteins.

4.3.3.1 Analytical gel filtration and native-PAGE

The AAA construct indeed forms a homotrimer based on the data from size-exclusion chromatography experiments. It elutes at a similar position to ABC dUTPase and importin-α (which almost has an identical molecular weight as the trimeric dUTPases) (Figure 31A and B). A molecular species with higher molecular mass arises when AAA or ABC dUTPases and importin-α are mixed, suggesting complex formation. The chromatography profile of the two putative complexes markedly differ proposing alterations in the complex composition (Figure 31A and B). The AAA dUTPase:importin-α complex has a slightly earlier elution peak, indication of having a higher molecular mass. In addition, the chromatography profile for the AAA dUTPase:importin-α complex is highly asymmetric, hinting the possibility of formation of multiple species with different molecular masses. Observations made based on the chromatograms are also evident when looking at the SDS-PAGE analysis of the fractions (Figure 31B). When applying the protein mixtures the elution peaks contain both dUTPase and importin-α, reinforcing the suggestion for complex formation.

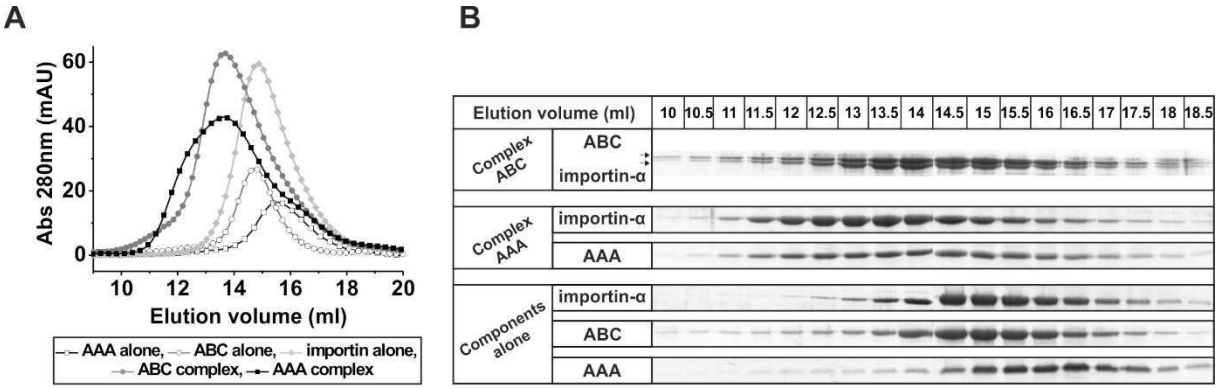


Figure 31. Size-exclusion chromatography analysis of complex formation between dUTPase constructs and importin-α

(A). Size-exclusion chromatograms of *D. virilis* AAA dUTPase, ABC dUTPase, importin-α alone and their mixtures. **(B)** Fractions of the size-exclusion chromatography were analyzed on SDS-PAGE.

The capability of AAA to form multiple species of possible complexes with importin-α were also investigated in native gel electrophoresis experiments. Figure 32A and B shows that multiple oligomeric forms are present in the AAA dUTPase:importin-α mixtures while only one species can be distinguished in the ABC dUTPase:importin-α mixtures.

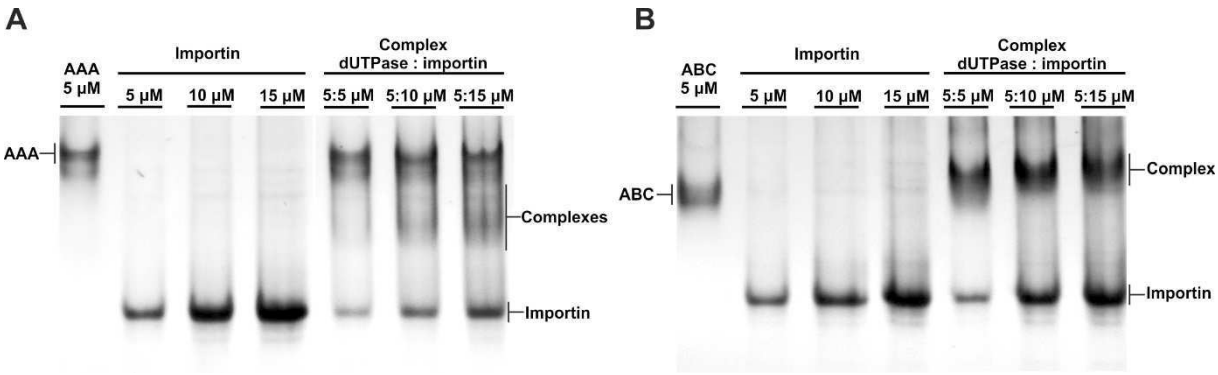
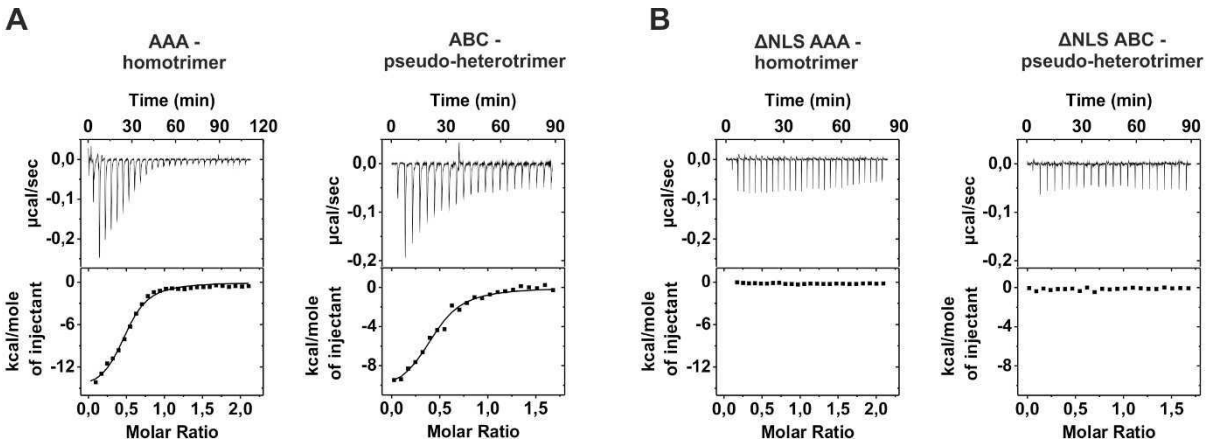


Figure 32. Native-PAGE analysis of complex formation between dUTPase constructs and importin-α

(A) AAA dUTPase or **(B)** ABC dUTPase constructs and importin-α were subjected to native gel electrophoresis either alone or in mixtures, as indicated. Note the appearance of new bands when the proteins are applied in mixtures, indicating complex formation.

4.3.3.2 Isothermal titration microcalorimetry

Isothermal titration microcalorimetry was applied to get quantitative thermodynamic insights about the complex formation between the *D. virilis* dUTPase constructs and importin-α (Figure 33A). Dissociation constant, molar enthalpy as well as the entropy of the interactions are very similar for both AAA dUTPase:importin-α and ABC dUTPase:importin-α complexes (Figure 33C). The molecular binding ratio ('n' values) indicating the stoichiometry of the complexation, is the same for the two construct, showing that each NLS could be occupied by importin-α (Figure 33C) under ideal conditions. Similarly to dUTPase all five NLSs of nucleoaplasmin could to be bound by importins [203]. The interaction between importin-α and the NLS deleted constructs (ΔNLS-AAA and ΔNLS-ABC) were fully abolished under the used conditions (Figure 33B). We could conclude that the NLS is responsible for the interaction with importin-α of these constructs, hence the most plausible explanation for the stoichiometry data is that the ABC pseudo-heterotrimeric dUTPase binds one importin-α molecule with its single NLS whereas the AAA homotrimeric dUTPase binds three importin-α molecules with the three NLSs.



C

ITC measurements *		
	AAA	ABC
N (molar binding ratio)	0.49 ± 0.01	0.44 ± 0.01
K_d (µM)	2.2 ± 0.3	1.8 ± 0.4
ΔH (kcal mol^{-1})	-15.6 ± 0.5	-11.1 ± 0.5
ΔS (cal mol^{-1} K^{-1})	-27.0	-11.6

* Reliability of the fittings are shown by the error values. Protein concentrations correspond to monomers.

Figure 33.Thermodynamic analysis of complex formation between dUTPase constructs and importin-α **(A)** shows titration data for AAA dUTPase:importin-α and ABC dUTPase:importin-α. **(B)** shows titration data for ΔNLS AAA dUTPase:importin-α and ΔNLS ABC dUTPase:importin-α. Top graphs present baseline-corrected timeline; bottom graphs show the integrated data (black square) and where binding data are observed, the fit of the binding isotherm by an "independent binding sites" model is also shown (solid line). **(C)** Values extracted are shown in the table. Note that for these calculations, the AAA concentration given corresponds to the concentration of AAA monomers.

Taking into account the similar K_d values, the difference in the cellular localization pattern between the two construct thus most likely come from the fact that the AAA construct, having three NLSs, can compete more efficiently for importin-α in the intracellular environment. We, therefore, directly show that NLS copy number in oligomers (avidity) [204] may contribute to the regulation of nucleocytoplasmic trafficking along with the affinity of the NLSs [40]. Thus NLS copy number variation can have a significant effect on nuclear transport efficiency and could provide an additional level of regulation which eukaryotes could utilize in fine tuning their transport processes.

5. Conclusion

Eukaryotes have achieved a great new potential in regulating their biochemical processes by establishing a new subcellular compartment, the nucleus. By separating their genome, replication, transcription and related processes into a newly defined compartment lead to the possibility of using means of spatial regulation not available to prokaryotes. The nuclear envelope, the barrier that separates the distinctive composition of the nucleus and the cytoplasm, is selectively permeable to macromolecules of the cell. Proteins greater than ~40 kDa can only pass through the nuclear pore complex (NPC) with active transport processes, while molecules that are smaller could theoretically enter the nucleus by passive diffusion. This selective regulation on bidirectional transport rely on active transport processes carried out by transport proteins (like the karyopherin protein family) which recognize the appropriate targeting signals on the cargo protein called the nuclear localization signal (NLS). dUTPases are no exception since their usual molecular weight (~55 kDa for the homotrimer) suggest the need of active transport mechanism to enter the nucleus. This was demonstrated in previous works addressing the sequence requirements for dUTPase nuclear transport in *Drosophila* [110] or in humans [104]. As described in section 1.2.2 in detail, vertebrates have a dedicated dUTPase isoform present in all compartments where DNA synthesis occurs, namely in the nucleus and in the mitochondria. There is no clear evidence to date that dUTPase has to be present in these dedicated compartments in order to maintain genomic integrity, however the conserved nature of the localization pattern argues for the fact that dUTPase has to be properly localized for adequate function. In this work we have seen that the nuclear localization of the human dUTPase is regulated through several aspects from limited proteolytic digestion to Cdk1 mediated phosphorylation.

The human dUTPase is a substrate of m-calpain based on *in vitro* digestion assays. The discrete cleavage sites were identified by mass spectrometry and are restricted to three positions on the N-terminal end of dUTPase which is characterized by high degree of conformational freedom. Due to this proteolytic event dUTPase might lose its NLS signal which would lead to the cytoplasmic retention of the enzyme. Cells activated by calcium have reduced dUTPase pools after 24 hours, possibly because m-calpain can only get access to the nuclear proteome during cell division. We were unable to detect fragments perhaps due to the destabilization of the enzyme and further proteolytic events as detailed in section 4.1.3, though this hypothesis needs further experimental confirmation. Our current results argue for the fact that calpain activation

might limit dUTPase nuclear availability, linking dUTPase levels to the several cellular pathways in which calpain signaling is involved.

Proteolytic activity is not the only mechanism which limits dUTPase presence in the nucleus. As we have shown Cdk1 phosphorylation, linked to the M-phase of the cell cycle, disrupts the nuclear accumulation of dUTPase. A mutant with glutamic acid substitution, mimicking the phosphorylated serine residue (S11E) in the vicinity of the NLS, is solely cytoplasmic and never enters the nucleus. Video microcopy revealed that this phosphorylation has an effect on the nuclear dUTPase pool of the daughter cells. In G1-phase dUTPase gets re-imported into the nucleus only after a significant delay compared to the non-phosphorylatable mutant (S11Q), which immediately accumulates in the nucleus once the new nuclear envelope is formed. However, the nuclear import rate for the WT is similar to the S11Q mutant dUTPase once it is licensed to enter the nucleus. Further work is needed to be done to understand the exact role why dUTPase levels are limited in the nucleus in the beginning of G1-phase by Cdk1 and to find the putative phosphatase that licenses dUTPase nuclear re-entry. We hypothesize that even though the substrate (dUTP) and products (dUMP + PPi) can freely shuttle among the compartments, dUTPase needs to maintain low levels of dUTP/dTTP ratio locally, at sites of DNA synthesis during replication and repair. It has been shown that the enzymes involved in *de novo* thymidylate biosynthesis (TYMS, DHFR, SHMT1 and SHMT2α) also have to translocate to the nucleus to prevent uracil incorporation into DNA during S-phase (as detailed in section 1.2.1.1). This might also apply for dUTPase, but this needs further experimental evidence. To investigate this issue, a dUTPase conditional knock-out cell line might be used to be rescued by dUTPase constructs differing in nuclear translocation capability. Such experiments are currently being conducted.

Phosphorylation at the S11 residue by Cdk1 inhibits nuclear accumulation because it disrupts the interaction between dUTPase and importin-α, the adaptor karyopherin protein responsible for classical NLS dependent nuclear entry. The characteristics of the complex have been investigated with various biophysical methods as detailed in section 4.2.2. Most importantly, based on ITC measurements, there is only one order of magnitude difference between the K_d value of the WT and S11E dUTPase:importin-α complex, resulting in such drastic effects in the cell. Functional NLSs have dissociation constants for importin-α binding in the range of 10 nM to 1-3 μM based on the work of Hodel and his colleagues [38, 40]. If the affinity of a certain cargo:importin-α interaction is close to the minimal affinity limit for a functional NLS, a mutation or posttranslational modification resulting in a small difference can have a

considerable effect. On the other hand, if the K_d value is close to the high affinity limit, a much more drastic alteration of the NLS would be required to make it non-functional. The ~ 0.8 μM K_d value of the WT and the S11Q compared to the ~ 10 μM K_d value of the S11E mutant dUTPase supports this explanation. It is also worth to notice that the K_d value of the *D. virilis* ABC dUTPase and the AAA dUTPase with importin-α in around ~ 2 μM and are less nuclear compared to the human WT dUTPase, which is especially evident in case of the ABC protein having only one NLS. The amount of NLSs a macromolecule harbors is also a key feature and will be addressed later.

The structure of phosphorylation mimicking (S11E) and WT-NLS peptides derived from human dUTPase in complex with importin-α, sheds light on the molecular detail how the complex is weakened by a P-1 position phosphorylation. The S11E mutation altered the conformation of the highly conserved R15 in the P2 position of the NLS, leading to a change in the binding arrangement of the peptide in the importin-α NLS binding groove and to a loss of interaction between the P12 and R15 NLS residues and the importin-α surface (detailed in section 4.2.2.5).

Recent publications also showed that cytoplasmic retention factors (such as BRAP2), specifically recognizing phospho-NLSs, contributing to the fine-tuning of the localization of several known proteins [56, 57]. Whether dUTPase might also fall under such additional mode of regulation will have to be investigated in the future.

The universal nature of phosphorylation regulated nuclear transport motivated us to search for other putative human proteins which might have Cdk1 regulated cell-cycle specific localization pattern similar to dUTPase. Kushogi and his colleagues have found several yeast protein candidates, where nuclear import was inhibited by Cdc28 dependent phosphorylation [156]. So we aimed to find human proteins with our own screening method. By applying an *in silico* screening, searching for proteins with NLSs having a Cdk1 consensus phosphorylation site in their P0 and P-1 position, we have found 92 candidates (Table A3) falling under the criteria. Eleven of these were experimentally validated by our NLS reporter assay, all showing significantly altered nuclear accumulation if the predicted phosphorylation was mimicked by glutamic acid substitution. The selected hits vary in cellular function, ranging from transcription and cell cycle regulation, DNA damage response pathway proteins to RNA editing, showing the widespread nature of this regulation. In summary cell cycle-dependent changes in the nuclear proteome are of utmost importance in the prompt regulation of cellular events, and kinases like Cdk1 cooperate to control the cell cycle dynamics. After cell mitosis, daughter cells form their own nuclear envelope and start off with a limited set of proteins that remain strictly

adherent to the chromosomes during cytokinesis [205]. Cells with open mitosis thus have the unique opportunity of re-setting the protein composition within the nucleus of daughter cells. We conclude that Cdk1-driven phosphorylation at P-1 or P0 positions of the NLSs makes a significant contribution to this re-shaping process of the nuclear proteome after M-phase, by potentially inhibiting the transport of several substrates, while phosphorylation in the P-2 position enhances nuclear re-import.

dUTPases form homotrimers, as it is the characteristic and evolutionary conserved structure of this proten family in most organisms (detailed in section 1.2.2). This means that in case of the nuclear isoforms each subunit has an NLS. It is interesting to speculate that by modulating each of these NLSs individually, either by phosphorylation or by proteolytic cleavage, cells can fine tune the efficiency of the nuclear accumulation of the enzyme. We have used *Drosophila* dUTPases, having unique features, to try to address these questions. We have shown that the number of NLSs present on different *Drosophila* dUTPases is in correlation with the efficiency of nuclear accumulation of the different proteins. Isoforms not having NLSs could not enter the nucleus (cytoplasmic *D. melanogaster* dUTPase) since their size (~ 60 kDa) does not permit passive diffusion through the NPCs. The three NLSs containing homotrimeric dUTPases (*D. melanogaster* nuclear dUTPase isoform and the engineered *D. virilis* AAA dUTPase), are exclusively nuclear. However, the *D. virilis* ABC dUTPase (the endogenous form) harboring only one NLS, is found in both the nuclear and the cytoplasmic compartment. This feature provides dUTPase availability in the nucleus as well as in the cytoplasm relying on the single *D. virilis* isoform, whereas in *D. melanogaster*, two isoforms are present to perform this task. In the future it would be interesting to address the question how this interesting genomic arrangement of the dUTPase gene have formed during evolution.

It seems that each of the three NLSs of the *D. virilis* AAA homotrimer is capable of binding one importin-α molecule (detailed in section 4.3.3) showing that the NLS copy number dictates the stoichiometry of the dUTPase:importin-α complex formation. Previous results in the literature show that import characteristics of the cargo are mainly regulated through the affinity and the concentration of the cargo and importins [15, 18, 40]. So enhanced nuclear accumulation might be achieved because a protein with three NLSs may capture an importin-α molecule more efficiently and spend a greater fraction of time bound to it. This increases the probability of recruiting importin-β and of being delivered across the nuclear pore complex, compared to a protein having only one NLS with the same affinity. These results could apply for most know nuclear oligomers and should be taken into account when investigating their nuclear transport processes.

6. Summary

Maintaining genomic integrity relies on the precise control of intracellular dNTP levels during replication and DNA repair. dUTPases have a major role in this process by maintaining low dUTP/dTTP levels, preventing genomic uracil incorporation. Thus dUTPase function is strongly linked to sites where DNA synthesis occurs, as could be also seen for the enzymes involved in *de novo* thymidine biosynthesis. Accordingly two distinct dUTPase isoforms exists in higher order eukaryotes, one localized in the nucleus and the other usually in the mitochondria. Transport into both compartments is highly regulated for macromolecules exceeding a size limit which also applies for dUTPase. We have been investigating the nucleocytoplasmic trafficking of dUTPases. Our results show that the human nuclear isoform is phosphorylated by Cdk1 in the vicinity of the NLS in G2/M-phase. After division this leads to the delayed nuclear re-entry of dUTPase in the daughter cells in G1-phase, reaching maximum nuclear accumulation rates at S-phase when replication is initiated. Phosphorylation perturbs the complex formation among dUTPase and importin-α. We have described the molecular details of this interaction by several biophysical methods and we also managed to solve the crystal structure of the dUTPase NLS peptide:importin-α complex, showing how phosphorylation in the P-1 position disrupts complex formation. In case of dUTPase it is important to consider the fact that the protein has three NLSs due to its homotrimeric structure. Using the adequate physiological models constituted by *Drosophila* dUTPases we have shown that this significantly enhances its nuclear import due to the fact that it could compete for free importin-α more efficiently with other proteins in the cellular environment. Generally we could conclude that besides affinity among components of the transport mechanism, and their relative concentrations, NLS copy number also has an important role in the regulation of nuclear transport regulation. The nuclear presence of dUTPase might also be regulated through proteolytic events mediated by m-calpain, which cleaves off the NLS, probably initiating the degradation of the protein later on. We have also tried to generalize or findings by using an *in silico* screening to find novel Cdk1 substrates where a phosphorylation in the appropriate position (P0 and P-1) inhibits nuclear translocation. Our findings include several hits which have a major role in the regulation of cell cycle, DNA damage response and transcription. Thus Cdk1 has a general role in influencing the nuclear proteome composition of the daughter cells after division, giving cells an additional level of plasticity to re-construct their nuclear environment depending on their momentary needs.

7. Összefoglalás

A genomi integritás megőrzéséhez elengedhetetlen a nukleotid készlet helyes arányának fenntartása a replikáció és DNS hibajavítás alatt. Ebben kiemelt jelentősége van a dUTPáznak, azáltal, hogy alacsony szinten tartja a dUTP/dTTP arányt, meggátolva ezzel a genomi uracil beépülést. Funkciójánál fogva erősen köthető azokhoz a kompartmentumokhoz, ahol DNS szintézis történik. Általánosságban a magasabb rendű eukariótáknak egy magi és egy mitokondriális dUTPáz formája van. Mindkét kompartmentumba szabályozott keretek között juthatnak be a makromolekulák, ami alól a dUTPáz sem kivétel. Munkánk során a humán dUTPáz magi transzportjának szabályozását igyekeztünk feltárni. Eredményeink azt mutatják, hogy a magi izoformát az S11-es pozícióban, ami az NLS közvetlen környezetében található, a Cdk1 kináz specifikusan foszforilálja a sejtciklus G2/M-fázisában. Ennek hatására, a dUTPáz, osztódást követően időben késleltetve lesz képes újra feltölteni a sejtmagot a leánysejtek a G1-fázisában, és csupán az S-fázisra, a replikáció idejére, éri el a magi felhalmozódás a maximális értéket. A foszforiláció az importin-α-val való kölcsönhatás gyengítése révén fejti ki hatását. A kölcsönhatás molekuláris részleteit számos biofizikai módszerrel jellemeztük. Sikerült megoldanunk a dUTPáz NLS peptid:importin-α kristályszerkezetét, ami betekintést ad arra, hogy egy P-1 pozícióban lévő foszforiláció miként teszi tönkre a komplex formálódását. A dUTPáz esetén fontos figyelembe venni, hogy a fehérjecsaládra jellemző homotrimer szerkezete miatt három NLS-el rendelkezik. Ez elősegíti a hatékony magi transzportját, hiszen feltételezhetően így nagyobb hatékonysággal verseng a többi fehérjével szabad importin-α-ért a zsúfolt celluláris környezetben, amit *Drosophila* dUTPázok segítségével bizonyítottunk. Általánosan elmondható, hogy az NLS affinitás, és a transzportban résztvevő komponensek koncetrációja mellett, az NLS kópiaszámnak is kiemelt szerepe van a magi transzportfolyamatok szabályozásában. A humán dUTPáz magi jelenlétét az m-kalpain általi limitált proteolízis is befolyásolhatja, ami a dUTPáz NLS-ét eltávolítja, és feltehetőleg későbbiekben az egész fehérje degradálódik. A dUTPáz példájából kiindulva a megfigyeléseinket megpróbáltuk általánosítani is. I*n silico* szűréssel számos új humán Cdk1 szubsztrátot találtunk, melyek megfelelő pozíciójában (P-1 és P0) a foszforiláció gyengíti vagy tönkreteszi azok magba jutását, melyeket laboratóriumi kísérletekben is igazoltunk. A talált fehérjék kiemelt fontossággal bírnak a sejtciklus, DNS hibajavítás és egyéb folyamatok szabályozásban. A Cdk1 ezáltal befolyásolhatja az osztódást követően a leánysejtek magi proteom összetételét, egy sejtciklushoz köthető dinamizmust biztosítva a rendszernek, hogy a sejtek a mindenkori igényeknek megfelelően építhessék újra a magi fehérjekészletüket.

8. Reference list

1. Gorlich, D. and U. Kutay, *Transport between the cell nucleus and the cytoplasm.* Annu Rev Cell Dev Biol, 1999. **15**: p. 607-60.
2. Ryan, K.J. and S.R. Wente, *The nuclear pore complex: a protein machine bridging the nucleus and cytoplasm.* Curr Opin Cell Biol, 2000. **12**(3): p. 361-71.
3. Wente, S.R. and M.P. Rout, *The nuclear pore complex and nuclear transport.* Cold Spring Harb Perspect Biol, 2010. **2**(10): p. a000562.
4. Keminer, O. and R. Peters, *Permeability of single nuclear pores.* Biophys J, 1999. **77**(1): p. 217-28.
5. Mohr, D., S. Frey, T. Fischer, T. Guttler, and D. Gorlich, *Characterisation of the passive permeability barrier of nuclear pore complexes.* EMBO J, 2009. **28**(17): p. 2541-53.
6. Paine, P.L., L.C. Moore, and S.B. Horowitz, *Nuclear envelope permeability.* Nature, 1975. **254**(5496): p. 109-14.
7. Mosammaparast, N., Y. Guo, J. Shabanowitz, D.F. Hunt, and L.F. Pemberton, *Pathways mediating the nuclear import of histones H3 and H4 in yeast.* J Biol Chem, 2002. **277**(1): p. 862-8.
8. Mosammaparast, N., K.R. Jackson, Y. Guo, C.J. Brame, J. Shabanowitz, D.F. Hunt, and L.F. Pemberton, *Nuclear import of histone H2A and H2B is mediated by a network of karyopherins.* J Cell Biol, 2001. **153**(2): p. 251-62.
9. Pante, N. and M. Kann, *Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm.* Mol Biol Cell, 2002. **13**(2): p. 425-34.
10. Chacinska, A., C.M. Koehler, D. Milenkovic, T. Lithgow, and N. Pfanner, *Importing mitochondrial proteins: machineries and mechanisms.* Cell, 2009. **138**(4): p. 628-44.
11. Corsi, A.K. and R. Schekman, *Mechanism of polypeptide translocation into the endoplasmic reticulum.* J Biol Chem, 1996. **271**(48): p. 30299-302.
12. Hoogenraad, N.J., L.A. Ward, and M.T. Ryan, *Import and assembly of proteins into mitochondria of mammalian cells.* Biochim Biophys Acta, 2002. **1592**(1): p. 97-105.
13. Steidl, S., A. Tuncher, H. Goda, C. Guder, N. Papadopoulou, T. Kobayashi, N. Tsukagoshi, M. Kato, and A.A. Brakhage, *A single subunit of a heterotrimeric CCAAT-binding complex carries a nuclear localization signal: piggy back transport of the pre-assembled complex to the nucleus.* J Mol Biol, 2004. **342**(2): p. 515-24.
14. Ribbeck, K. and D. Gorlich, *Kinetic analysis of translocation through nuclear pore complexes.* EMBO J, 2001. **20**(6): p. 1320-30.
15. Timney, B.L., J. Tetenbaum-Novatt, D.S. Agate, R. Williams, W. Zhang, B.T. Chait, and M.P. Rout, *Simple kinetic relationships and nonspecific competition govern nuclear import rates in vivo.* J Cell Biol, 2006. **175**(4): p. 579-93.
16. Maul, G.G. and L. Deaven, *Quantitative determination of nuclear pore complexes in cycling cells with differing DNA content.* J Cell Biol, 1977. **73**(3): p. 748-60.
17. Gorlich, D., M.J. Seewald, and K. Ribbeck, *Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation.* EMBO J, 2003. **22**(5): p. 1088-100.
18. Riddick, G. and I.G. Macara, *A systems analysis of importin-{alpha}-{beta} mediated nuclear protein import.* J Cell Biol, 2005. **168**(7): p. 1027-38.
19. Yang, W. and S.M. Musser, *Nuclear import time and transport efficiency depend on importin beta concentration.* J Cell Biol, 2006. **174**(7): p. 951-61.
20. Lim, R.Y., N.P. Huang, J. Koser, J. Deng, K.H. Lau, K. Schwarz-Herion, B. Fahrenkrog, and U. Aebi, *Flexible phenylalanine-glycine nucleoporins as entropic*

barriers to nucleocytoplasmic transport. Proc Natl Acad Sci U S A, 2006. **103**(25): p. 9512-7.
21. Frey, S., R.P. Richter, and D. Gorlich, *FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties.* Science, 2006. **314**(5800): p. 815-7.
22. Gorlich, D., F. Vogel, A.D. Mills, E. Hartmann, and R.A. Laskey, *Distinct functions for the two importin subunits in nuclear protein import.* Nature, 1995. **377**(6546): p. 246-8.
23. Strom, A.C. and K. Weis, *Importin-beta-like nuclear transport receptors.* Genome Biol, 2001. **2**(6): p. REVIEWS3008.
24. Pumroy, R.A. and G. Cingolani, *Diversification of importin-alpha isoforms in cellular trafficking and disease states.* Biochem J, 2015. **466**(1): p. 13-28.
25. Moore, J.D., J. Yang, R. Truant, and S. Kornbluth, *Nuclear import of Cdk/cyclin complexes: identification of distinct mechanisms for import of Cdk2/cyclin E and Cdc2/cyclin B1.* J Cell Biol, 1999. **144**(2): p. 213-24.
26. Cingolani, G., J. Bednenko, M.T. Gillespie, and L. Gerace, *Molecular basis for the recognition of a nonclassical nuclear localization signal by importin beta.* Mol Cell, 2002. **10**(6): p. 1345-53.
27. Lam, M.H., L.J. Briggs, W. Hu, T.J. Martin, M.T. Gillespie, and D.A. Jans, *Importin beta recognizes parathyroid hormone-related protein with high affinity and mediates its nuclear import in the absence of importin alpha.* J Biol Chem, 1999. **274**(11): p. 7391-8.
28. Kalderon, D. and A.E. Smith, *In vitro mutagenesis of a putative DNA binding domain of SV40 large-T.* Virology, 1984. **139**(1): p. 109-37.
29. Robbins, J., S.M. Dilworth, R.A. Laskey, and C. Dingwall, *Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence.* Cell, 1991. **64**(3): p. 615-23.
30. Kobe, B., *Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha.* Nat Struct Biol, 1999. **6**(4): p. 388-97.
31. Conti, E., M. Uy, L. Leighton, G. Blobel, and J. Kuriyan, *Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha.* Cell, 1998. **94**(2): p. 193-204.
32. Fontes, M.R., T. Teh, D. Jans, R.I. Brinkworth, and B. Kobe, *Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha.* J Biol Chem, 2003. **278**(30): p. 27981-7.
33. Fontes, M.R., T. Teh, and B. Kobe, *Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha.* J Mol Biol, 2000. **297**(5): p. 1183-94.
34. Marfori, M., T.G. Lonhienne, J.K. Forwood, and B. Kobe, *Structural Basis of High-Affinity Nuclear Localization Signal Interactions with Importin-alpha.* Traffic, 2012.
35. Marfori, M., A. Mynott, J.J. Ellis, A.M. Mehdi, N.F. Saunders, P.M. Curmi, J.K. Forwood, M. Boden, and B. Kobe, *Molecular basis for specificity of nuclear import and prediction of nuclear localization.* Biochim Biophys Acta, 2011. **1813**(9): p. 1562-77.
36. Lange, A., R.E. Mills, C.J. Lange, M. Stewart, S.E. Devine, and A.H. Corbett, *Classical nuclear localization signals: definition, function, and interaction with importin alpha.* J Biol Chem, 2007. **282**(8): p. 5101-5.
37. Stewart, M., *Molecular mechanism of the nuclear protein import cycle.* Nat Rev Mol Cell Biol, 2007. **8**(3): p. 195-208.
38. Hodel, M.R., A.H. Corbett, and A.E. Hodel, *Dissection of a nuclear localization signal.* J Biol Chem, 2001. **276**(2): p. 1317-25.

39. Kosugi, S., M. Hasebe, N. Matsumura, H. Takashima, E. Miyamoto-Sato, M. Tomita, and H. Yanagawa, *Six classes of nuclear localization signals specific to different binding grooves of importin alpha.* J Biol Chem, 2009. **284**(1): p. 478-85.
40. Hodel, A.E., M.T. Harreman, K.F. Pulliam, M.E. Harben, J.S. Holmes, M.R. Hodel, K.M. Berland, and A.H. Corbett, *Nuclear localization signal receptor affinity correlates with in vivo localization in Saccharomyces cerevisiae.* J Biol Chem, 2006. **281**(33): p. 23545-56.
41. Gorlich, D., N. Pante, U. Kutay, U. Aebi, and F.R. Bischoff, *Identification of different roles for RanGDP and RanGTP in nuclear protein import.* EMBO J, 1996. **15**(20): p. 5584-94.
42. Yang, W., J. Gelles, and S.M. Musser, *Imaging of single-molecule translocation through nuclear pore complexes.* Proc Natl Acad Sci U S A, 2004. **101**(35): p. 12887-92.
43. Nachury, M.V. and K. Weis, *The direction of transport through the nuclear pore can be inverted.* Proc Natl Acad Sci U S A, 1999. **96**(17): p. 9622-7.
44. Matsuura, Y. and M. Stewart, *Structural basis for the assembly of a nuclear export complex.* Nature, 2004. **432**(7019): p. 872-7.
45. Poon, I.K. and D.A. Jans, *Regulation of nuclear transport: central role in development and transformation?* Traffic, 2005. **6**(3): p. 173-86.
46. Pouton, C.W., K.M. Wagstaff, D.M. Roth, G.W. Moseley, and D.A. Jans, *Targeted delivery to the nucleus.* Adv Drug Deliv Rev, 2007. **59**(8): p. 698-717.
47. Kaffman, A., N.M. Rank, and E.K. O'Shea, *Phosphorylation regulates association of the transcription factor Pho4 with its import receptor Pse1/Kap121.* Genes Dev, 1998. **12**(17): p. 2673-83.
48. Jans, D.A., *The regulation of protein transport to the nucleus by phosphorylation.* Biochem J, 1995. **311 (Pt 3)**: p. 705-16.
49. Jans, D.A. and S. Hubner, *Regulation of protein transport to the nucleus: central role of phosphorylation.* Physiol Rev, 1996. **76**(3): p. 651-85.
50. Nardozzi, J.D., K. Lott, and G. Cingolani, *Phosphorylation meets nuclear import: a review.* Cell Commun Signal, 2010. **8**: p. 32.
51. Hubner, S., C.Y. Xiao, and D.A. Jans, *The protein kinase CK2 site (Ser111/112) enhances recognition of the simian virus 40 large T-antigen nuclear localization sequence by importin.* J Biol Chem, 1997. **272**(27): p. 17191-5.
52. Fontes, M.R., T. Teh, G. Toth, A. John, I. Pavo, D.A. Jans, and B. Kobe, *Role of flanking sequences and phosphorylation in the recognition of the simian-virus-40 large T-antigen nuclear localization sequences by importin-alpha.* Biochem J, 2003. **375**(Pt 2): p. 339-49.
53. Kosugi, S., M. Hasebe, T. Entani, S. Takayama, M. Tomita, and H. Yanagawa, *Design of peptide inhibitors for the importin alpha/beta nuclear import pathway by activity-based profiling.* Chem Biol, 2008. **15**(9): p. 940-9.
54. Riviere, Y., V. Blank, P. Kourilsky, and A. Israel, *Processing of the precursor of NF-kappa B by the HIV-1 protease during acute infection.* Nature, 1991. **350**(6319): p. 625-6.
55. Humbert-Lan, G. and T. Pieler, *Regulation of DNA binding activity and nuclear transport of B-Myb in Xenopus oocytes.* J Biol Chem, 1999. **274**(15): p. 10293-300.
56. Asada, M., K. Ohmi, D. Delia, S. Enosawa, S. Suzuki, A. Yuo, H. Suzuki, and S. Mizutani, *Brap2 functions as a cytoplasmic retention protein for p21 during monocyte differentiation.* Mol Cell Biol, 2004. **24**(18): p. 8236-43.
57. Fulcher, A.J., D.M. Roth, S. Fatima, G. Alvisi, and D.A. Jans, *The BRCA-1 binding protein BRAP2 is a novel, negative regulator of nuclear import of viral proteins,*

dependent on phosphorylation flanking the nuclear localization signal. FASEB J, 2010. **24**(5): p. 1454-66.
58. Forwood, J.K., V. Harley, and D.A. Jans, *The C-terminal nuclear localization signal of the sex-determining region Y (SRY) high mobility group domain mediates nuclear import through importin beta 1.* J Biol Chem, 2001. **276**(49): p. 46575-82.
59. Roth, D.M., G.W. Moseley, D. Glover, C.W. Pouton, and D.A. Jans, *A microtubule-facilitated nuclear import pathway for cancer regulatory proteins.* Traffic, 2007. **8**(6): p. 673-86.
60. Roth, D.M., G.W. Moseley, C.W. Pouton, and D.A. Jans, *Mechanism of microtubule-facilitated "fast track" nuclear import.* J Biol Chem, 2011. **286**(16): p. 14335-51.
61. Kohler, M., C. Speck, M. Christiansen, F.R. Bischoff, S. Prehn, H. Haller, D. Gorlich, and E. Hartmann, *Evidence for distinct substrate specificities of importin alpha family members in nuclear protein import.* Mol Cell Biol, 1999. **19**(11): p. 7782-91.
62. Echeverria, P.C. and D. Picard, *Molecular chaperones, essential partners of steroid hormone receptors for activity and mobility.* Biochim Biophys Acta, 2010. **1803**(6): p. 641-9.
63. Nordlund, P. and P. Reichard, *Ribonucleotide reductases.* Annu Rev Biochem, 2006. **75**: p. 681-706.
64. Mathews, C.K., *DNA precursor metabolism and genomic stability.* FASEB J, 2006. **20**(9): p. 1300-14.
65. Bestwick, R.K., G.L. Moffett, and C.K. Mathews, *Selective expansion of mitochondrial nucleoside triphosphate pools in antimetabolite-treated HeLa cells.* J Biol Chem, 1982. **257**(16): p. 9300-4.
66. Leeds, J.M., M.B. Slabaugh, and C.K. Mathews, *DNA precursor pools and ribonucleotide reductase activity: distribution between the nucleus and cytoplasm of mammalian cells.* Mol Cell Biol, 1985. **5**(12): p. 3443-50.
67. Prem veer Reddy, G. and A.B. Pardee, *Multienzyme complex for metabolic channeling in mammalian DNA replication.* Proc Natl Acad Sci U S A, 1980. **77**(6): p. 3312-16.
68. Anderson, D.D., C.M. Quintero, and P.J. Stover, *Identification of a de novo thymidylate biosynthesis pathway in mammalian mitochondria.* Proc Natl Acad Sci U S A, 2011. **108**(37): p. 15163-8.
69. MacFarlane, A.J., D.D. Anderson, P. Flodby, C.A. Perry, R.H. Allen, S.P. Stabler, and P.J. Stover, *Nuclear localization of de novo thymidylate biosynthesis pathway is required to prevent uracil accumulation in DNA.* J Biol Chem, 2012. **286**(51): p. 44015-22.
70. Anderson, D.D., J.Y. Eom, and P.J. Stover, *Competition between sumoylation and ubiquitination of serine hydroxymethyltransferase 1 determines its nuclear localization and its accumulation in the nucleus.* J Biol Chem, 2012. **287**(7): p. 4790-9.
71. Woeller, C.F., D.D. Anderson, D.M. Szebenyi, and P.J. Stover, *Evidence for small ubiquitin-like modifier-dependent nuclear import of the thymidylate biosynthesis pathway.* J Biol Chem, 2007. **282**(24): p. 17623-31.
72. Anderson, D.D., C.F. Woeller, E.P. Chiang, B. Shane, and P.J. Stover, *Serine hydroxymethyltransferase anchors de novo thymidylate synthesis pathway to nuclear lamina for DNA synthesis.* J Biol Chem, 2012. **287**(10): p. 7051-62.
73. Chen, Y.L., S. Eriksson, and Z.F. Chang, *Regulation and functional contribution of thymidine kinase 1 in repair of DNA damage.* J Biol Chem, 2010. **285**(35): p. 27327-35.
74. Niida, H., Y. Katsuno, M. Sengoku, M. Shimada, M. Yukawa, M. Ikura, T. Ikura, K. Kohno, H. Shima, H. Suzuki, S. Tashiro, and M. Nakanishi, *Essential role of Tip60-*

dependent recruitment of ribonucleotide reductase at DNA damage sites in DNA repair during G1 phase. Genes Dev, 2010. **24**(4): p. 333-8.

75. Somasekaram, A., A. Jarmuz, A. How, J. Scott, and N. Navaratnam, *Intracellular localization of human cytidine deaminase. Identification of a functional nuclear localization signal.* J Biol Chem, 1999. **274**(40): p. 28405-12.
76. Tye, B.K., P.O. Nyman, I.R. Lehman, S. Hochhauser, and B. Weiss, *Transient accumulation of Okazaki fragments as a result of uracil incorporation into nascent DNA.* Proc Natl Acad Sci U S A, 1977. **74**(1): p. 154-7.
77. Vertessy, B.G. and J. Toth, *Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases.* Acc Chem Res, 2009. **42**(1): p. 97-106.
78. Zhang, X. and C.K. Mathews, *Natural DNA precursor pool asymmetry and base sequence context as determinants of replication fidelity.* J Biol Chem, 1995. **270**(15): p. 8401-4.
79. Kunz, B.A., S.E. Kohalmi, T.A. Kunkel, C.K. Mathews, E.M. McIntosh, and J.A. Reidy, *International Commission for Protection Against Environmental Mutagens and Carcinogens. Deoxyribonucleoside triphosphate levels: a critical factor in the maintenance of genetic stability.* Mutat Res, 1994. **318**(1): p. 1-64.
80. Grogan, B.C., J.B. Parker, A.F. Guminski, and J.T. Stivers, *Effect of the thymidylate synthase inhibitors on dUTP and TTP pool levels and the activities of DNA repair glycosylases on uracil and 5-fluorouracil in DNA.* Biochemistry, 2011. **50**(5): p. 618-27.
81. Longley, D.B., D.P. Harkin, and P.G. Johnston, *5-fluorouracil: mechanisms of action and clinical strategies.* Nat Rev Cancer, 2003. **3**(5): p. 330-8.
82. Tinkelenberg, B.A., M.J. Hansbury, and R.D. Ladner, *dUTPase and uracil-DNA glycosylase are central modulators of antifolate toxicity in Saccharomyces cerevisiae.* Cancer Res, 2002. **62**(17): p. 4909-15.
83. Wyatt, M.D. and D.M. Wilson, 3rd, *Participation of DNA repair in the response to 5-fluorouracil.* Cell Mol Life Sci, 2009. **66**(5): p. 788-99.
84. Ladner, R.D., F.J. Lynch, S. Groshen, Y.P. Xiong, A. Sherrod, S.J. Caradonna, J. Stoehlmacher, and H.J. Lenz, *dUTP nucleotidohydrolase isoform expression in normal and neoplastic tissues: association with survival and response to 5-fluorouracil in colorectal cancer.* Cancer Res, 2000. **60**(13): p. 3493-503.
85. Koehler, S.E. and R.D. Ladner, *Small interfering RNA-mediated suppression of dUTPase sensitizes cancer cell lines to thymidylate synthase inhibition.* Mol Pharmacol, 2004. **66**(3): p. 620-6.
86. Merenyi, G., J. Kovari, J. Toth, E. Takacs, I. Zagyva, A. Erdei, and B.G. Vertessy, *Cellular response to efficient dUTPase RNAi silencing in stable HeLa cell lines perturbs expression levels of genes involved in thymidylate metabolism.* Nucleosides Nucleotides Nucleic Acids, 2011. **30**(6): p. 369-90.
87. Studebaker, A.W., W.P. Lafuse, R. Kloesel, and M.V. Williams, *Modulation of human dUTPase using small interfering RNA.* Biochem Biophys Res Commun, 2005. **327**(1): p. 306-10.
88. Wilson, P.M., W. Fazzone, M.J. LaBonte, J. Deng, N. Neamati, and R.D. Ladner, *Novel opportunities for thymidylate metabolism as a therapeutic target.* Mol Cancer Ther, 2008. **7**(9): p. 3029-37.
89. Sire, J., G. Querat, C. Esnault, and S. Priet, *Uracil within DNA: an actor of antiviral immunity.* Retrovirology, 2008. **5**: p. 45.
90. Tye, B.K., J. Chien, I.R. Lehman, B.K. Duncan, and H.R. Warner, *Uracil incorporation: a source of pulse-labeled DNA fragments in the replication of the Escherichia coli chromosome.* Proc Natl Acad Sci U S A, 1978. **75**(1): p. 233-7.

91. Griffin, J.E., J.D. Gawronski, M.A. Dejesus, T.R. Ioerger, B.J. Akerley, and C.M. Sassetti, *High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism.* PLoS Pathog, 2011. **7**(9): p. e1002251.
92. Sassetti, C.M., D.H. Boyd, and E.J. Rubin, *Genes required for mycobacterial growth defined by high density mutagenesis.* Mol Microbiol, 2003. **48**(1): p. 77-84.
93. Pecsi, I., R. Hirmondo, A.C. Brown, A. Lopata, T. Parish, B.G. Vertessy, and J. Toth, *The dUTPase enzyme is essential in Mycobacterium smegmatis.* PLoS One, 2012. **7**(5): p. e37461.
94. Gadsden, M.H., E.M. McIntosh, J.C. Game, P.J. Wilson, and R.H. Haynes, *dUTP pyrophosphatase is an essential enzyme in Saccharomyces cerevisiae.* EMBO J, 1993. **12**(11): p. 4425-31.
95. Guillet, M., P.A. Van Der Kemp, and S. Boiteux, *dUTPase activity is critical to maintain genetic stability in Saccharomyces cerevisiae.* Nucleic Acids Res, 2006. **34**(7): p. 2056-66.
96. Dengg, M., T. Garcia-Muse, S.G. Gill, N. Ashcroft, S.J. Boulton, and H. Nilsen, *Abrogation of the CLK-2 checkpoint leads to tolerance to base-excision repair intermediates.* EMBO Rep, 2006. **7**(10): p. 1046-51.
97. Erdelyi, P., E. Borsos, K. Takacs-Vellai, T. Kovacs, A.L. Kovacs, T. Sigmond, B. Hargitai, L. Pasztor, T. Sengupta, M. Dengg, I. Pecsi, J. Toth, H. Nilsen, B.G. Vertessy, and T. Vellai, *Shared developmental roles and transcriptional control of autophagy and apoptosis in Caenorhabditis elegans.* J Cell Sci, 2011. **124**(Pt 9): p. 1510-8.
98. Castillo-Acosta, V.M., A.M. Estevez, A.E. Vidal, L.M. Ruiz-Perez, and D. Gonzalez-Pacanowska, *Depletion of dimeric all-alpha dUTPase induces DNA strand breaks and impairs cell cycle progression in Trypanosoma brucei.* Int J Biochem Cell Biol, 2008. **40**(12): p. 2901-13.
99. Dubois, E., D. Cordoba-Canero, S. Massot, N. Siaud, B. Gakiere, S. Domenichini, F. Guerard, T. Roldan-Arjona, and M.P. Doutriaux, *Homologous recombination is stimulated by a decrease in dUTPase in Arabidopsis.* PLoS One, 2011. **6**(4): p. e18658.
100. Muha, V., A. Horvath, A. Bekesi, M. Pukancsik, B. Hodoscsek, G. Merenyi, G. Rona, J. Batki, I. Kiss, F. Jankovics, P. Vilmos, M. Erdelyi, and B.G. Vertessy, *Uracil-containing DNA in Drosophila: stability, stage-specific accumulation, and developmental involvement.* PLoS Genet, 2012. **8**(6): p. e1002738.
101. Baldo, A.M. and M.A. McClure, *Evolution and horizontal transfer of dUTPase-encoding genes in viruses and their hosts.* J Virol, 1999. **73**(9): p. 7710-21.
102. Ladner, R.D., D.E. McNulty, S.A. Carr, G.D. Roberts, and S.J. Caradonna, *Characterization of distinct nuclear and mitochondrial forms of human deoxyuridine triphosphate nucleotidohydrolase.* J Biol Chem, 1996. **271**(13): p. 7745-51.
103. Ladner, R.D. and S.J. Caradonna, *The human dUTPase gene encodes both nuclear and mitochondrial isoforms. Differential expression of the isoforms and characterization of a cDNA encoding the mitochondrial species.* J Biol Chem, 1997. **272**(30): p. 19072-80.
104. Tinkelenberg, B.A., W. Fazzone, F.J. Lynch, and R.D. Ladner, *Identification of sequence determinants of human nuclear dUTPase isoform localization.* Exp Cell Res, 2003. **287**(1): p. 39-46.
105. Otterlei, M., T. Haug, T.A. Nagelhus, G. Slupphaug, T. Lindmo, and H.E. Krokan, *Nuclear and mitochondrial splice forms of human uracil-DNA glycosylase contain a complex nuclear localisation signal and a strong classical mitochondrial localisation signal, respectively.* Nucleic Acids Res, 1998. **26**(20): p. 4611-7.
106. Songyang, Z., S. Blechner, N. Hoagland, M.F. Hoekstra, H. Piwnica-Worms, and L.C. Cantley, *Use of an oriented peptide library to determine the optimal substrates of protein kinases.* Curr Biol, 1994. **4**(11): p. 973-82.

107. Ladner, R.D., S.A. Carr, M.J. Huddleston, D.E. McNulty, and S.J. Caradonna, *Identification of a consensus cyclin-dependent kinase phosphorylation site unique to the nuclear form of human deoxyuridine triphosphate nucleotidohydrolase.* J Biol Chem, 1996. **271**(13): p. 7752-7.
108. Wilson, P.M., W. Fazzone, M.J. LaBonte, H.J. Lenz, and R.D. Ladner, *Regulation of human dUTPase gene expression and p53-mediated transcriptional repression in response to oxaliplatin-induced DNA damage.* Nucleic Acids Res, 2009. **37**(1): p. 78-95.
109. Bekesi, A., I. Zagyva, E. Hunyadi-Gulyas, V. Pongracz, J. Kovari, A.O. Nagy, A. Erdei, K.F. Medzihradszky, and B.G. Vertessy, *Developmental regulation of dUTPase in Drosophila melanogaster.* J Biol Chem, 2004. **279**(21): p. 22362-70.
110. Merenyi, G., E. Konya, and B.G. Vertessy, *Drosophila proteins involved in metabolism of uracil-DNA possess different types of nuclear localization signals.* FEBS J, 2010. **277**(9): p. 2142-56.
111. Muha, V., I. Zagyva, Z. Venkei, J. Szabad, and B.G. Vertessy, *Nuclear localization signal-dependent and -independent movements of Drosophila melanogaster dUTPase isoforms during nuclear cleavage.* Biochem Biophys Res Commun, 2009. **381**(2): p. 271-5.
112. Kovari, J., O. Barabas, E. Takacs, A. Bekesi, Z. Dubrovay, V. Pongracz, I. Zagyva, T. Imre, P. Szabo, and B.G. Vertessy, *Altered active site flexibility and a structural metal-binding site in eukaryotic dUTPase: kinetic characterization, folding, and crystallographic studies of the homotrimeric Drosophila enzyme.* J Biol Chem, 2004. **279**(17): p. 17932-44.
113. Goll, D.E., V.F. Thompson, H. Li, W. Wei, and J. Cong, *The calpain system.* Physiol Rev, 2003. **83**(3): p. 731-801.
114. Zatz, M. and A. Starling, *Calpains and disease.* N Engl J Med, 2005. **352**(23): p. 2413-23.
115. Farkas, A., P. Tompa, E. Schad, R. Sinka, G. Jekely, and P. Friedrich, *Autolytic activation and localization in Schneider cells (S2) of calpain B from Drosophila.* Biochem J, 2004. **378**(Pt 2): p. 299-305.
116. Jekely, G. and P. Friedrich, *Characterization of two recombinant Drosophila calpains. CALPA and a novel homolog, CALPB.* J Biol Chem, 1999. **274**(34): p. 23893-900.
117. Tompa, P., P. Buzder-Lantos, A. Tantos, A. Farkas, A. Szilagyi, Z. Banoczi, F. Hudecz, and P. Friedrich, *On the sequential determinants of calpain cleavage.* J Biol Chem, 2004. **279**(20): p. 20775-85.
118. Bozoky, Z., A. Alexa, J. Dancsok, G. Gogl, E. Klement, K.F. Medzihradszky, and P. Friedrich, *Identifying calpain substrates in intact S2 cells of Drosophila.* Arch Biochem Biophys, 2009. **481**(2): p. 219-25.
119. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.
120. Varga, B., O. Barabas, J. Kovari, J. Toth, E. Hunyadi-Gulyas, E. Klement, K.F. Medzihradszky, F. Tolgyesi, J. Fidy, and B.G. Vertessy, *Active site closure facilitates juxtaposition of reactant atoms for initiation of catalysis by human dUTPase.* FEBS Lett, 2007. **581**(24): p. 4783-8.
121. Sorg, G. and T. Stamminger, *Mapping of nuclear localization signals by simultaneous fusion to green fluorescent protein and to beta-galactosidase.* Biotechniques, 1999. **26**(5): p. 858-62.

122. Bozoky, Z., G. Rona, E. Klement, K.F. Medzihradszky, G. Merenyi, B.G. Vertessy, and P. Friedrich, *Calpain-catalyzed proteolysis of human dUTPase specifically removes the nuclear localization signal peptide.* PLoS One, 2011. **6**(5): p. e19546.
123. Pierce, M.M., C.S. Raman, and B.T. Nall, *Isothermal titration calorimetry of protein-protein interactions.* Methods, 1999. **19**(2): p. 213-21.
124. Takacs, E., V.K. Grolmusz, and B.G. Vertessy, *A tradeoff between protein stability and conformational mobility in homotrimeric dUTPases.* FEBS Lett, 2004. **566**(1-3): p. 48-54.
125. McPhillips, T.M., S.E. McPhillips, H.J. Chiu, A.E. Cohen, A.M. Deacon, P.J. Ellis, E. Garman, A. Gonzalez, N.K. Sauter, R.P. Phizackerley, S.M. Soltis, and P. Kuhn, *Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines.* J Synchrotron Radiat, 2002. **9**(Pt 6): p. 401-6.
126. Kabsch, W., *Xds.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 125-32.
127. Winn, M.D., C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, and K.S. Wilson, *Overview of the CCP4 suite and current developments.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.
128. Cowtan, K., P. Emsley, and K.S. Wilson, *From crystal to structure with CCP4.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 233-4.
129. McCoy, A.J., R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, and R.J. Read, *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.
130. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
131. Adams, P.D., P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, and P.H. Zwart, *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta crystallographica. Section D, Biological crystallography, 2010. **66**(Pt 2): p. 213-21.
132. Ellis, J.J. and B. Kobe, *Predicting protein kinase specificity: Predikin update and performance in the DREAM4 challenge.* PLoS One, 2011. **6**(7): p. e21169.
133. Mehdi, A.M., M.S. Sehgal, B. Kobe, T.L. Bailey, and M. Boden, *A probabilistic model of nuclear import of proteins.* Bioinformatics, 2011. **27**(9): p. 1239-46.
134. Uniprot, C., *Reorganizing the protein space at the Universal Protein Resource (UniProt).* Nucleic Acids Res, 2012. **40**(Database issue): p. D71-5.
135. Stark, C., B.J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, and M. Tyers, *BioGRID: a general repository for interaction datasets.* Nucleic Acids Res, 2006. **34**(Database issue): p. D535-9.
136. Desmet, F.O., D. Hamroun, M. Lalande, G. Collod-Beroud, M. Claustres, and C. Beroud, *Human Splicing Finder: an online bioinformatics tool to predict splicing signals.* Nucleic Acids Res, 2009. **37**(9): p. e67.
137. Friedrich, P. and Z. Bozoky, *Digestive versus regulatory proteases: on calpain action in vivo.* Biol Chem, 2005. **386**(7): p. 609-12.
138. Toth, J., B. Varga, M. Kovacs, A. Malnasi-Csizmadia, and B.G. Vertessy, *Kinetic mechanism of human dUTPase, an essential nucleotide pyrophosphatase enzyme.* J Biol Chem, 2007. **282**(46): p. 33572-82.
139. Gonda, D.K., A. Bachmair, I. Wunning, J.W. Tobias, W.S. Lane, and A. Varshavsky, *Universality and structure of the N-end rule.* J Biol Chem, 1989. **264**(28): p. 16700-12.

140. Mellgren, R.L., *Proteolysis of nuclear proteins by mu-calpain and m-calpain.* J Biol Chem, 1991. **266**(21): p. 13920-4.
141. Small, G.W., T.Y. Chou, C.V. Dang, and R.Z. Orlowski, *Evidence for involvement of calpain in c-Myc proteolysis in vivo.* Arch Biochem Biophys, 2002. **400**(2): p. 151-61.
142. Choi, Y.H., S.J. Lee, P. Nguyen, J.S. Jang, J. Lee, M.L. Wu, E. Takano, M. Maki, P.A. Henkart, and J.B. Trepel, *Regulation of cyclin D1 by calpain protease.* J Biol Chem, 1997. **272**(45): p. 28479-84.
143. Janossy, J., P. Ubezio, A. Apati, M. Magocsi, P. Tompa, and P. Friedrich, *Calpain as a multi-site regulator of cell cycle.* Biochem Pharmacol, 2004. **67**(8): p. 1513-21.
144. Raynaud, F., G. Carnac, A. Marcilhac, and Y. Benyamin, *m-Calpain implication in cell cycle during muscle precursor cell activation.* Exp Cell Res, 2004. **298**(1): p. 48-57.
145. Sherwood, S.W., A.L. Kung, J. Roitelman, R.D. Simoni, and R.T. Schimke, *In vivo inhibition of cyclin B degradation and induction of cell-cycle arrest in mammalian cells by the neutral cysteine protease inhibitor N-acetylleucylleucylnorleucinal.* Proc Natl Acad Sci U S A, 1993. **90**(8): p. 3353-7.
146. Huang, C.J., T. Gurlo, L. Haataja, S. Costes, M. Daval, S. Ryazantsev, X. Wu, A.E. Butler, and P.C. Butler, *Calcium-activated calpain-2 is a mediator of beta cell dysfunction and apoptosis in type 2 diabetes.* J Biol Chem, 2010. **285**(1): p. 339-48.
147. Del Bello, B., D. Moretti, A. Gamberucci, and E. Maellaro, *Cross-talk between calpain and caspase-3/-7 in cisplatin-induced apoptosis of melanoma cells: a major role of calpain inhibition in cell death protection and p53 status.* Oncogene, 2007. **26**(19): p. 2717-26.
148. Tarrant, M.K. and P.A. Cole, *The chemical biology of protein phosphorylation.* Annu Rev Biochem, 2009. **78**: p. 797-825.
149. Kudo, N., B. Wolff, T. Sekimoto, E.P. Schreiner, Y. Yoneda, M. Yanagida, S. Horinouchi, and M. Yoshida, *Leptomycin B inhibition of signal-mediated nuclear export by direct binding to CRM1.* Exp Cell Res, 1998. **242**(2): p. 540-7.
150. Schafer, K.A., *The cell cycle: a review.* Vet Pathol, 1998. **35**(6): p. 461-78.
151. Beeckmans, S., *Chromatographic methods to study protein-protein interactions.* Methods, 1999. **19**(2): p. 278-305.
152. Harreman, M.T., T.M. Kline, H.G. Milford, M.B. Harben, A.E. Hodel, and A.H. Corbett, *Regulation of nuclear import by phosphorylation adjacent to nuclear localization signals.* J Biol Chem, 2004. **279**(20): p. 20613-21.
153. Jans, D.A., M.J. Ackermann, J.R. Bischoff, D.H. Beach, and R. Peters, *p34cdc2-mediated phosphorylation at T124 inhibits nuclear import of SV-40 T antigen proteins.* J Cell Biol, 1991. **115**(5): p. 1203-12.
154. Geymonat, M., A. Spanos, G.P. Wells, S.J. Smerdon, and S.G. Sedgwick, *Clb6/Cdc28 and Cdc14 regulate phosphorylation status and cellular localization of Swi6.* Mol Cell Biol, 2004. **24**(6): p. 2277-85.
155. Sidorova, J.M., G.E. Mikesell, and L.L. Breeden, *Cell cycle-regulated phosphorylation of Swi6 controls its nuclear localization.* Mol Biol Cell, 1995. **6**(12): p. 1641-58.
156. Kosugi, S., M. Hasebe, M. Tomita, and H. Yanagawa, *Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs.* Proc Natl Acad Sci U S A, 2009. **106**(25): p. 10171-6.
157. Jans, D.A., T. Moll, K. Nasmyth, and P. Jans, *Cyclin-dependent kinase site-regulated signal-dependent nuclear localization of the SW15 yeast transcription factor in mammalian cells.* J Biol Chem, 1995. **270**(29): p. 17064-7.
158. Moll, T., G. Tebb, U. Surana, H. Robitsch, and K. Nasmyth, *The role of phosphorylation and the CDC28 protein kinase in cell cycle-regulated nuclear import of the S. cerevisiae transcription factor SWI5.* Cell, 1991. **66**(4): p. 743-58.

159. Jaquenoud, M., F. van Drogen, and M. Peter, *Cell cycle-dependent nuclear export of Cdh1p may contribute to the inactivation of APC/C(Cdh1).* EMBO J, 2002. **21**(23): p. 6515-26.
160. Zhou, Y., Y.P. Ching, A.C. Chun, and D.Y. Jin, *Nuclear localization of the cell cycle regulator CDH1 and its regulation by phosphorylation.* J Biol Chem, 2003. **278**(14): p. 12530-6.
161. Huynh, M.A., J. Stegmuller, N. Litterman, and A. Bonni, *Regulation of Cdh1-APC function in axon growth by Cdh1 phosphorylation.* J Neurosci, 2009. **29**(13): p. 4322-7.
162. Liku, M.E., V.Q. Nguyen, A.W. Rosales, K. Irie, and J.J. Li, *CDK phosphorylation of a novel NLS-NES module distributed between two subunits of the Mcm2-7 complex prevents chromosomal rereplication.* Mol Biol Cell, 2005. **16**(10): p. 5026-39.
163. Enquist-Newman, M., M. Sullivan, and D.O. Morgan, *Modulation of the mitotic regulatory network by APC-dependent destruction of the Cdh1 inhibitor Acm1.* Mol Cell, 2008. **30**(4): p. 437-46.
164. Hennekes, H., M. Peter, K. Weber, and E.A. Nigg, *Phosphorylation on protein kinase C sites inhibits nuclear import of lamin B2.* J Cell Biol, 1993. **120**(6): p. 1293-304.
165. Ikuta, T., Y. Kobayashi, and K. Kawajiri, *Phosphorylation of nuclear localization signal inhibits the ligand-dependent nuclear import of aryl hydrocarbon receptor.* Biochem Biophys Res Commun, 2004. **317**(2): p. 545-50.
166. Tagawa, T., T. Kuroki, P.K. Vogt, and K. Chida, *The cell cycle-dependent nuclear import of v-Jun is regulated by phosphorylation of a serine adjacent to the nuclear localization signal.* J Cell Biol, 1995. **130**(2): p. 255-63.
167. Gnad, F., J. Gunawardena, and M. Mann, *PHOSIDA 2011: the posttranslational modification database.* Nucleic Acids Res, 2011. **39**(Database issue): p. D253-60.
168. Gnad, F., S. Ren, J. Cox, J.V. Olsen, B. Macek, M. Oroshi, and M. Mann, *PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites.* Genome Biol, 2007. **8**(11): p. R250.
169. Cimprich, K.A. and D. Cortez, *ATR: an essential regulator of genome integrity.* Nat Rev Mol Cell Biol, 2008. **9**(8): p. 616-27.
170. Kim, H., J. Chen, and X. Yu, *Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response.* Science, 2007. **316**(5828): p. 1202-5.
171. Sobhian, B., G. Shao, D.R. Lilli, A.C. Culhane, L.A. Moreau, B. Xia, D.M. Livingston, and R.A. Greenberg, *RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites.* Science, 2007. **316**(5828): p. 1198-202.
172. Guerrero-Santoro, J., M.G. Kapetanaki, C.L. Hsieh, I. Gorbachinsky, A.S. Levine, and V. Rapic-Otrin, *The cullin 4B-based UV-damaged DNA-binding protein ligase binds to UV-damaged chromatin and ubiquitinates histone H2A.* Cancer Res, 2008. **68**(13): p. 5014-22.
173. Ku, W.C., S.K. Chiu, Y.J. Chen, H.H. Huang, and W.G. Wu, *Complementary quantitative proteomics reveals that transcription factor AP-4 mediates E-box-dependent complex formation for transcriptional repression of HDM2.* Mol Cell Proteomics, 2009. **8**(9): p. 2034-50.
174. Lee, Y.H., M.T. Bedford, and M.R. Stallcup, *Regulated recruitment of tumor suppressor BRCA1 to the p21 gene by coactivator methylation.* Genes Dev, 2011. **25**(2): p. 176-88.
175. Liu, H., H.C. Hew, Z.G. Lu, T. Yamaguchi, Y. Miki, and K. Yoshida, *DNA damage signalling recruits RREB-1 to the p53 tumour suppressor promoter.* Biochem J, 2009. **422**(3): p. 543-51.
176. Thiagalingam, A., A. De Bustros, M. Borges, R. Jasti, D. Compton, L. Diamond, M. Mabry, D.W. Ball, S.B. Baylin, and B.D. Nelkin, *RREB-1, a novel zinc finger protein,*

is involved in the differentiation response to Ras in human medullary thyroid carcinomas. Mol Cell Biol, 1996. **16**(10): p. 5335-45.
177. Iyer, N.G., H. Ozdag, and C. Caldas, *p300/CBP and cancer.* Oncogene, 2004. **23**(24): p. 4225-31.
178. Egawa, T. and D.R. Littman, *Transcription factor AP4 modulates reversible and epigenetic silencing of the Cd4 gene.* Proc Natl Acad Sci U S A, 2011. **108**(36): p. 14873-8.
179. Imai, K. and T. Okamoto, *Transcriptional repression of human immunodeficiency virus type 1 by AP-4.* J Biol Chem, 2006. **281**(18): p. 12495-505.
180. Jin, Q., L.R. Yu, L. Wang, Z. Zhang, L.H. Kasper, J.E. Lee, C. Wang, P.K. Brindle, S.Y. Dent, and K. Ge, *Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation.* EMBO J, 2011. **30**(2): p. 249-62.
181. Guetg, C., P. Lienemann, V. Sirri, I. Grummt, D. Hernandez-Verdun, M.O. Hottiger, M. Fussenegger, and R. Santoro, *The NoRC complex mediates the heterochromatin formation and stability of silent rRNA genes and centromeric repeats.* EMBO J, 2010. **29**(13): p. 2135-46.
182. Strohner, R., A. Nemeth, P. Jansa, U. Hofmann-Rohrer, R. Santoro, G. Langst, and I. Grummt, *NoRC--a novel member of mammalian ISWI-containing chromatin remodeling machines.* EMBO J, 2001. **20**(17): p. 4892-900.
183. Jia, S., R. Kobayashi, and S.I. Grewal, *Ubiquitin ligase component Cul4 associates with Clr4 histone methyltransferase to assemble heterochromatin.* Nat Cell Biol, 2005. **7**(10): p. 1007-13.
184. Blencowe, B.J., G. Bauren, A.G. Eldridge, R. Issner, J.A. Nickerson, E. Rosonina, and P.A. Sharp, *The SRm160/300 splicing coactivator subunits.* RNA, 2000. **6**(1): p. 111-20.
185. Yang, L., N. Li, C. Wang, Y. Yu, L. Yuan, M. Zhang, and X. Cao, *Cyclin L2, a novel RNA polymerase II-associated cyclin, is involved in pre-mRNA splicing and induces apoptosis of human hepatocellular carcinoma cells.* J Biol Chem, 2004. **279**(12): p. 11639-48.
186. Zou, Y., J. Mi, J. Cui, D. Lu, X. Zhang, C. Guo, G. Gao, Q. Liu, B. Chen, C. Shao, and Y. Gong, *Characterization of nuclear localization signal in the N terminus of CUL4B and its essential role in cyclin E degradation and cell cycle progression.* J Biol Chem, 2009. **284**(48): p. 33320-32.
187. Miranda-Carboni, G.A., S.A. Krum, K. Yee, M. Nava, Q.E. Deng, S. Pervin, A. Collado-Hidalgo, Z. Galic, J.A. Zack, K. Nakayama, K.I. Nakayama, and T.F. Lane, *A functional link between Wnt signaling and SKP2-independent p27 turnover in mammary tumors.* Genes Dev, 2008. **22**(22): p. 3121-34.
188. Higa, L.A., X. Yang, J. Zheng, D. Banks, M. Wu, P. Ghosh, H. Sun, and H. Zhang, *Involvement of CUL4 ubiquitin E3 ligases in regulating CDK inhibitors Dacapo/p27Kip1 and cyclin E degradation.* Cell Cycle, 2006. **5**(1): p. 71-7.
189. Li, H.L., T.S. Wang, X.Y. Li, N. Li, D.Z. Huang, Q. Chen, and Y. Ba, *Overexpression of cyclin L2 induces apoptosis and cell-cycle arrest in human lung cancer cells.* Chin Med J (Engl), 2007. **120**(10): p. 905-9.
190. Kondo, T., P.L. Sheets, D.A. Zopf, H.L. Aloor, T.R. Cummins, R.J. Chan, and E. Hashino, *Tlx3 exerts context-dependent transcriptional regulation and promotes neuronal differentiation from embryonic stem cells.* Proc Natl Acad Sci U S A, 2008. **105**(15): p. 5780-5.
191. Worman, H.J., *Nuclear lamins and laminopathies.* J Pathol, 2011. **226**(2): p. 316-25.

192. Dittmer, T.A. and T. Misteli, *The lamin protein family.* Genome Biol, 2011. **12**(5): p. 222.
193. Matsuoka, S., B.A. Ballif, A. Smogorzewska, E.R. McDonald, 3rd, K.E. Hurov, J. Luo, C.E. Bakalarski, Z. Zhao, N. Solimini, Y. Lerenthal, Y. Shiloh, S.P. Gygi, and S.J. Elledge, *ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage.* Science, 2007. **316**(5828): p. 1160-6.
194. Langerak, P. and P. Russell, *Regulatory networks integrating cell cycle control with DNA damage checkpoints and double-strand break repair.* Philos Trans R Soc Lond B Biol Sci, 2012. **366**(1584): p. 3562-71.
195. Dephoure, N., C. Zhou, J. Villen, S.A. Beausoleil, C.E. Bakalarski, S.J. Elledge, and S.P. Gygi, *A quantitative atlas of mitotic phosphorylation.* Proc Natl Acad Sci U S A, 2008. **105**(31): p. 10762-7.
196. Olsen, J.V., M. Vermeulen, A. Santamaria, C. Kumar, M.L. Miller, L.J. Jensen, F. Gnad, J. Cox, T.S. Jensen, E.A. Nigg, S. Brunak, and M. Mann, *Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis.* Sci Signal, 2010. **3**(104): p. ra3.
197. Stephen, A.G., J.S. Trausch-Azar, P.M. Handley-Gearhart, A. Ciechanover, and A.L. Schwartz, *Identification of a region within the ubiquitin-activating enzyme required for nuclear targeting and phosphorylation.* J Biol Chem, 1997. **272**(16): p. 10895-903.
198. Lee, M.K., W.M. Tong, Z.Q. Wang, and K. Sabapathy, *Serine 312 phosphorylation is dispensable for wild-type p53 functions in vivo.* Cell Death Differ, 2010. **18**(2): p. 214-21.
199. Takacs, E., O. Barabas, M.V. Petoukhov, D.I. Svergun, and B.G. Vertessy, *Molecular shape and prominent role of beta-strand swapping in organization of dUTPase oligomers.* FEBS Lett, 2009. **583**(5): p. 865-71.
200. Fiser, A. and B.G. Vertessy, *Altered Subunit Communication in Subfamilies of Trimeric dUTPases.* Biochem Biophys Res Commun, 2000. **279**(2): p. 534-542.
201. Cook, A., F. Bono, M. Jinek, and E. Conti, *Structural biology of nucleocytoplasmic transport.* Annu Rev Biochem, 2007. **76**: p. 647-71.
202. Feldherr, C., D. Akin, T. Littlewood, and M. Stewart, *The molecular mechanism of translocation through the nuclear pore complex is highly conserved.* J Cell Sci, 2002. **115**(Pt 14): p. 2997-3005.
203. Falces, J., I. Arregi, P.V. Konarev, M.A. Urbaneja, D.I. Svergun, S.G. Taneva, and S. Banuelos, *Recognition of nucleoplasmin by its nuclear transport receptor importin alpha/beta: insights into a complete import complex.* Biochemistry, 2010. **49**(45): p. 9756-69.
204. Dingwall, C., S.V. Sharnick, and R.A. Laskey, *A polypeptide domain that specifies migration of nucleoplasmin into the nucleus.* Cell, 1982. **30**(2): p. 449-58.
205. Swanson, J.A. and P.L. McNeil, *Nuclear reassembly excludes large macromolecules.* Science, 1987. **238**(4826): p. 548-50.
206. Chen, V.B., W.B. Arendall, 3rd, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, and D.C. Richardson, *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta crystallographica. Section D, Biological crystallography, 2010. **66**(Pt 1): p. 12-21.

9. Publication list

PEER-REVIEWED PUBLICATIONS RELATED TO THE DOCTORAL THESIS

2014 **Rona, G.,** H.L. Palinkas, M. Borsos, A. Horvath, I. Scheer, A. Benedek, G.N. Nagy, I. Zagyva, and B.G. Vertessy,
NLS copy-number variation governs efficiency of nuclear import--case study on dUTPases. *FEBS J*, 2014. 281(24): p. 5463-78., IF: 3.99

2014 **Rona, G.,** M. Borsos, J.J. Ellis, A.M. Mehdi, M. Christie, Z. Kornyei, M. Neubrandt, J. Toth, Z. Bozoky, L. Buday, E. Madarasz, M. Boden, B. Kobe, and B.G. Vertessy
Dynamics of re-constitution of the human nuclear proteome after cell division is regulated by NLS-adjacent phosphorylation, *Cell Cycle, 2014. 13(22): p. 3551-64.,* IF: 5.24

2013 **Rona, G.**, M. Marfori, M. Borsos, I. Scheer, E. Takacs, J. Toth, F. Babos, A. Magyar, A. Erdei, Z. Bozoky, L. Buday, B. Kobe, and B.G. Vertessy,
Phosphorylation adjacent to the nuclear localization signal of human dUTPase abolishes nuclear import: structural and mechanistic insights. *Acta Crystallogr D Biol Crystallogr*, 2013. 69(Pt 12): p. 2495-505., IF: 7.23

2011 Bozoky Z*, **Rona G***, Klement E, Medzihradszky KF, Merenyi G, Vertessy BG & Friedrich P.
Calpain-catalyzed proteolysis of human dUTPase specifically removes the nuclear localization signal peptide. *PLoS One* 6(5): p. e19546, IF: 3.53
*equal contribution

OTHER PEER-REVIEWED PUBLICATIONS

2014 Szabo, J.E., V. Nemeth, V. Papp-Kadar, K. Nyiri, I. Leveles, A.A. Bendes, I. Zagyva, **G. Rona**, H.L. Palinkas, B. Besztercei, O. Ozohanics, K. Vekey, K. Liliom, J. Toth, and B.G. Vertessy,
Highly potent dUTPase inhibition by a bacterial repressor protein reveals a novel mechanism for gene expression control. *Nucleic Acids Res*, 2014. 42(19): p. 11912-20., IF: 8.81

2013 Leveles, I., Nemeth, V., Szabo, J.E., Harmat, V., Nyiri, K., Bendes, A.A., Papp-Kadar, V., Zagyva, I., **Rona, G.**, Ozohanics, O., et al.
Structure and enzymatic mechanism of a moonlighting dUTPase. *Acta crystallographica Section D, Biological crystallography 69, 2298-2308*., IF: 7.23

2013 Chiş L, Hriscu M, Bica A, Toşa M, Nagy G, **Róna G**, G Vértessy B, Dan Irimie F
Molecular cloning and characterization of a thermostable esterase/lipase produced by a novel Anoxybacillus flavithermus strain. *J Gen Appl Microbiol. 59, 119-34*., IF: 0.74

2012 D. Beke, Zs. Szekrényes, D. Pálfi, **G. Róna**, I. Balogh, P. Maák, G. Katona, Zs. Czigány, K. Kamarás, B. Rózsa, L. Buday, B. Vértessy, and A. Gali
Silicon carbide quantum dots for bioimaging. *Journal of Materials Research, DOI: 10.1557/jmr.2012.296,* IF: 1.82

2012 Muha, V., Horvath, A., Bekesi, A., Pukancsik, M., Hodoscsek, B., Merenyi, G., **Rona, G**., Batki, J., Kiss, I., Jankovics, F., Vilmos P, Erdelyi M, Vertessy BG
Uracil-containing DNA in Drosophila: stability, stage-specific accumulation, and developmental involvement. *PLoS Genetics 8, e1002738*, IF: 8.17

2011 Leveles I, **Rona G**, Zagyva I, Bendes A, Harmat V & Vertessy BG

Crystallization and preliminary crystallographic analysis of dUTPase from the phi11 helper phage of Staphylococcus aureus. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 67, *1411-1413*, IF: 0.57

2008 Kiss R, Bozoky Z, Kovacs D, **Rona G**, Friedrich P, Dvortsak P, Weisemann R, Tompa P & Perczel A
Calcium-induced tripartite binding of intrinsically disordered calpastatin to its cognate enzyme, calpain. *FEBS Lett 582, 2149-54.,* IF: 3.34

PUBLICATIONS IN PRESS AT PEER-REVIEWED JOURNALS

2015 Horváth, András; Batki, Júlia; Henn, László; Lukacsovich, Tamás; **Róna, Gergely**; Erdélyi, Miklós; Vértessy, Beáta
dUTPase expression correlates with cell division potential in Drosophila melanogaster, *FEBS J,* accepted manuscript, IF: 3.99

COMMENT ARTICLE AT PEER-REVIEWED JOURNALS

2014 **Rona, G.,** M. Borsos, B. Kobe, and B.G. Vertessy,
Factors influencing nucleo-cytoplasmic trafficking: which matter? Response to Alvisi & Jans' comment on Phosphorylation adjacent to the nuclear localization signal of human dUTPase abolishes nuclear import: structural and mechanistic insights. *Acta Crystallogr D Biol Crystallogr*, 2014. 70(Pt 10): p. 2777-8., IF: 7.23

INTERNATIONAL CONFERENCE ORAL PRESENTATIONS – PRESENTING AUTHOR UNDERLINED

2013 23rd Wilhelm Bernhard Workshop on the Cell Nucleus, Debrecen, Hungary
"Reconstitution of the human nuclear proteome after cell division through NLS modulation."
Gergely Róna, Zsuzsanna Környei, Mary Marfori, Máté Borsos, Máté Neubrandt, Ildikó Scheer, Judit Tóth, Anna Magyar, Emília Madarász, Zoltán Bozóky, Jonathan J. Ellis, Ahmed M. Mehdi, Mikael Bodén, László Buday, Bostjan Kobe, Beáta G. Vértessy

2013 9th European Biophysics Congress, Lisbon, Portugal
"Genomic integrity of a virulenece gene is preserved by a dUTPase-based molecular switch"
Judit Szabó, Veronika Németh, Veronika Kádár, Nyiri Kinga, Ibolya Leveles, Ábris Bendes, **Gergely Róna**, Hajnalka Palinkas, Balázs Besztercei, Vékey Károly, Károly Lililom, Judit Tóth, Beáta Vértessy

2012 FEBS 3+ Meeting, Opatija, Croatia
"Towards a molecular switch"
Ibolya Leveles, Imre Zagyva, Ábris Bendes, Veronika Németh, Veronika Kádár, Judit Szabó, Judit Tóth, **Gergely Róna**, Balázs Besztercei, Károly Lililom, Beáta Vértessy

2012 9th International Conference „Students for Students", Kolozsvár, Romania
"Two in one: reduction of NLS copy number provides multi-organellar availability of D.virilis dUTPase "
Palinkas HL, **Róna G**, Horváth A, Boros M, Vértessy BG

2011 8th International Conference „Students for Students", Kolozsvár, Romania
"Nuclear transport of oligomeric proteins in eukaryotes largely depends on copy number of cognate peptide signals: case of dUTPase"
Pálinkás HL, **Róna G**, Horváth A, Borsos M and Vértessy BG

2009 Spine2 Complexes 3rd Annual Congress Lisbon, Portugal

"One-way ticket to the nucleus: Phosphorylation within the nuclear localization signal of human dUTPase regulates complex formation with importin-alpha"
Gergely Róna, E. Takács, Zs. Környei, E. Madarász, and Beáta G. Vértessy

INTERNATIONAL CONFERENCE POSTER PRESENTATIONS – PRESENTING AUTHOR UNDERLINED

2013 23rd Wilhelm Bernhard Workshop on the Cell Nucleus, Debrecen, Hungary
"Zinc finger nucleases (zfns), designed to cut at specific dna sequences, are powerful tools in targeted genome editing"
Gergely Róna and Beáta G. Vértessy

2013 23rd Wilhelm Bernhard Workshop on the Cell Nucleus, Debrecen, Hungary
"Uracil exclusion from the genome of undifferentiated cells is essential for development in *Drosophila melanogaster*"
András Horváth, László Henn, Villő Muha, Júlia Batki, Péter Vilmos, **Gergely Róna**, Miklós Erdélyi, Beáta G. Vértessy

2012 János Szentágothai Memorial Conference and Student Competition, Pécs, Hungary
"The role of dUTPase in preserving genome integrity"
Júlia Batki, András Horváth, **Gergely Róna**, Péter Vilmos, Miklós Erdélyi, Beáta G. Vértessy

2012 János Szentágothai Memorial Conference and Student Competition, Pécs, Hungary
"NLS copy number variation mediates task distribution among cell organells: a case of D. virilis dUTPase"
Hajnalka Palinkas, **Gergely Róna**, Máté Borsos, András Horváth, Beáta G. Vértessy

2012 FEBS 3+ Meeting, Opatija, Croatia
"Cdk1 phosphorylation governs nuclear proteome redistribution in daughter cells after division: legacy of mother cells "
Gergely Róna, Máté Borsos, Mary Marfori, Jonathan J Ellis, Ahmed M Mehdi, Mikael Bodén, Bostjan Kobe, Beáta Vértessy

2012 FEBS 3+ Meeting, Opatija, Croatia
"dUTPase task distribution between cell organells in eukaryotes"
Hajnalka Pálinkás, **Gergely Róna**, Máté Borsos, András Horváth, Beáta Vértessy

2012 FEBS 3+ Meeting, Opatija, Croatia
"Oligomerization and cell-cycle dependent phosphorylation governs nuclear transport of dUTPases"
Máté Borsos, **Gergely Róna**, Zsuzsanna Környei, Mary Marfori, Máté, Neubrandt, Ildikó Scheer, Enikő Takács, Judit Tóth, Emília Madarász, Zoltán, Bozóky, László Buday, Bostjan Kobe, Beáta Vértessy

2011 EMBO Meeting, 2011, Vienna, Austria
"CDK1 phosphorylation governs nuclear proteome redistribution in daughter cells after division: legacy of mother cells"
Gergely Róna, Zoltán Bozóky, Enikô Takács, Ildikó Scheer, Zsuzsa Környei, Máté Neubrandt, Mary Marfori, Bostjan Kobe, Emília Madarász and Beáta G. Vértessy

2010 EMBO Meeting, 2010, Barcelona, Spain
"One way or return ticket to the nucleus? Cell-cycle dependentphosphorylation governs nuclear transport of human dUTPase"
Gergely Róna, Zoltán Bozóky, Enikô Takács, Ildikó Scheer, Zsuzsa Környei, Máté Neubrandt, Mary Marfori, Bostjan Kobe, Emília Madarász and Beáta G. Vértessy

2010 Biophysical Society 54th Annual Meeting, San Francisco, California, USA
"Phosphorylation dependent nuclear transport of human dUTPase"
Gergely Róna, Zoltán Bozóky, Enikő Takács, Zsuzsa Környei, Máté Neubrandt, Mary Marfori, BostjanKobe, Emília Madarász, Beáta Vértessy G.

HUNGARIAN CONFERENCE ORAL PRESENTATIONS – PRESENTING AUTHOR UNDERLINED

2014 Annual Meeting of the Hungarian Biochemical Society, Debrecen, Hungary
"Phosphorylation adjacent to the nuclear localization signal of human dUTPase abolishes nuclear import: structural and mechanistic insights"
Gergely Róna, Mary Marfori, Máté Borsos, Ildikó Scheer, Enikő Takács, Judit Tóth, Fruzsina Babos, Anna Magyar, Anna Erdei, Zoltán Bozóky, László Buday, Bostjan Kobe, Beáta G. Vértessy

2013 Hungarian Molecular Life Sciences 2013, Siófok, Hungary
"Cdk1 phosphorylation governs nuclear proteome redistribution in daughter cells after division: legacy of mother cells"
Gergely Róna, Máté Borsos, Mary Marfori, Jonathan J Ellis, Ahmed M Mehdi, Mikael Bodén, Zsuzsanna Környei, Máté Neubrandt, Judit Tóth, Bostjan Kobe, Beáta G. Vértessy

2012 Kálmán Erika Doctoral Conference, Mátraháza, Hungary
"Uracil containing DNA in Holometabola insects"
Scheer Ildikó, **Róna Gergely**

2012 III[rd] conference of Signal Transduction Division of Hungarian Biochemical Society, Esztergom, Hungary
"Reconstitution of the nuclear proteome after cell division through NLS modulation"
Róna Gergely; Környei Zsuzsanna; Marfori Mary; Boros Máté; Neubrandt Máté; Scheer Ildikó; Takács Enikő; Tóth Judit; Madarász Emília; J Ellis Jonathan; M. Mehdi Ahmed; Bodén Mikael; Buday László; Magyar Anna; Kobe Bostjan; Vértessy G. Beáta

2012 III[rd] conference of Signal Transduction Division of Hungarian Biochemical Society, Esztergom, Hungary
"Cellular responses induced by uracil-DNA in Drosophila"
Horváth András; Muha Villő; Békési Angéla; Batki Júlia; **Róna Gergely**; Vilmos Péter; Kiss István; Erdélyi Miklós; Vértessy G. Beáta

2010 Hungarian Biochemical Society Annual Meeting, Budapest, Hungary
"Phosphorylation governed nuclear transport of human dUTPase"
Gergely Róna, Eniko Takacs, Ildiko Scheer, Zsuzsanna Kornyei, Mate Neubrandt, Emilia Madarasz, Laszlo Buday, Judit Toth and Beata G. Vertessy

2009 Straub-days, Szeged, Hungary
"One-way or return ticket to the nucleus? Phosphorylation-induced regulation of human dUTPase nucleo-cytoplasmic trafficking"
Gergely Róna, Eniko Takacs, Ildiko Scheer, Zsuzsanna Kornyei, Mate Neubrandt, Emilia Madarasz, Laszlo Buday, Judit Toth and Beata G. Vertessy

HUNGARIAN CONFERENCE POSTER PRESENTATIONS – PRESENTING AUTHOR UNDERLINED

2012 Oláh György IX. Doctorial conference, Budapest, Hungary
"Silicon carbide quantum dots for bioimaging"
Beke D, Szekrenyes Zs, Balogh I, **Rona G**, Vertessy B, Kamaras K and Gali A

2012 Straub-days, Szeged, Hungary

"CDK1 phosphorylation governs nuclear proteome redistribution in daughter cells after division: legacy of mother cells"
Gergely Róna, Zoltán Bozóky, Enikô Takács, Ildikó Scheer, Zsuzsa Környei, Máté Neubrandt, Mary Marfori, Bostjan Kobe, Emília Madarász and Beáta G. Vértessy

2012 Straub-days, Szeged, Hungary
"Nuclear transport of oligomeric proteins NLS:importin stoichiometry in the case of dUTPase"
Hajnalka Pálinkás, **Gergely Róna**, Máté Borsos, András Horváth, Imre Zagyva, Beáta G. Vértessy

2011 Hungarian Biochemical Society Annual Meeting, Pécs, Hungary
"Nuclear transport of oligomers, a case study on dUTPase"
Róna G, Pálinkás HL, Borsos M, Horvath A, és Vértessy BG

2011 Hungarian Biochemical Society Annual Meeting, Pécs, Hungary
"Effect of cell cycle dependent phosphorylation on nuclear trasnport"
Borsos Máté, **Róna Gergely**, Környei Zsuzsa, Neubrandt Máté, Mary Marfori, Bostjan Kobe, Madarász Emília, Vértessy G. Beáta

2010 Straub-days, Szeged, Hungary
"One way or return ticket to the nucleus? Cell-cycle dependent phosphorylation governs nuclear transport of human dUTPase"
Gergely Róna, Zoltán Bozóky, Enikô Takács, Ildikó Scheer, Zsuzsa Környei, Máté Neubrandt, Mary Marfori, Bostjan Kobe, Emília Madarász and Beáta G. Vértessy

2009 Hungarian Biochemical Society Annual Meeting, Budapest, Hungary
"Phosphorlylation regulated nuclear accumulation of the human dUTPase"
Róna G., Bozóky Z., Környei Zs., Neubrandt M., Madarász E., Takács E., Vértessy G. B.

10. Appendix

Table A1. List of peptides identified by mass spectrometry of calpain digested dUTPase from gel bands 3 and 4 of Figure 6.

Peptides with non-tryptic cleavage sites, generated by calpain digestion, are shown in bold on light grey background. Start and end positions are numbered according to the human dUTPase sequence. Residues in parentheses are the N- and C-terminal neighboring residues of the protein (not present in the peptides). Error indicates the difference between the measured and calculated peptide masses. Modifications are also noted. z, charge, m/z, mass normalized to charge.

m/z	z	Error/Da	Peptide	Start	End
529.8700	**2**	**0.098**	**(S)EETPAISPSK(R)**	**5**	**14**
607.9600	**2**	**0.14**	**(S)EETPAISPSKR(A)**	**5**	**15**
428.3200	**2**	**0.064**	**(T)PAISPSKR(A)**	**8**	**15**
434.3700	3	0.14	(R)ARPAEVGGM(Oxidation)QLR(F)	16	27
642.9900	2	0.14	(R)ARPAEVGGMQLR(F)	16	27
537.3800	2	0.11	(R)PAEVGGM(Oxidation)QLR(F)	18	27
529.3600	2	0.083	(R)PAEVGGMQLR(F)	18	27
361.6500	3	0.13	(R)LSEHATAPTR(G)	31	40
485.3300	**2**	**0.089**	**(L)SEHATAPTR(G)**	**32**	**40**
1092.6900	2	0.19	(R)AAGYDLYSAYDYTIPPM(Oxidation)EK(A)	45	63
723.5100	3	0.18	(R)AAGYDLYSAYDYTIPPMEK(A)	45	63
776.1300	2	0.24	(K)TDIQIALPSGC(Carbamidomethyl)YGR(V)	68	81
853.6000	2	0.19	(K)HFIDVGAGVIDEDYR(G)	92	106
626.0900	2	0.24	(R)GNVGVVLFNFGK(E)	107	118
502.0000	2	0.23	(R)IAQLIC(Carbamidomethyl)ER(I)	129	136
1034.2900	2	0.29	(R)IFYPEIEEVQALDDTER(G)	137	153

Table A2. Crystallographic data collection and refinement statistics

	dUTPase NLS WT	**dUTPase S11E mutant**
PDB ID	4FDR	4FDS
Data Collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters		
a, b, c (Å)	77.88, 89.99, 99.73	78.03, 89.95, 99.44
Resolution (Å)	99.73- 2.10 (2.21-2.10) [a]	99.44-1.88 (1.98-1.88) [a]
Total No. of reflections	238847 (34586) [a]	421273 (61134) [a]
No. of unique reflections	41431 (5965) [a]	57564 (8321) [a]
Redundancy	5.8 (5.8) [a]	7.3 (7.3) [a]
Completeness (%)	99.4 (100.0) [a]	99.9 (100.0) [a]
$\langle I/\sigma(I)\rangle$	10.8 (2.3) [a a]	25.8 (2.4) [a]
R_{merge} [b]	0.101 (0.794)	0.049 (0.806) [a]
Refinement		
No. of Impα/ peptide molecules in the asymmetric unit	1/ 2	1/ 2
No. of protein atoms	3244	3252
No. of dUTPase peptide atoms/ water	150/ 220	141/ 378
R_{cryst}/R_{free} (%) [c]	19.9/ 23.4	17.4/ 19.4
Luzzati plot coordinate error (Å)	0.286	0.221
Average B factors (Å^2)		
Wilson B factor (Å^2)	34.48	31.53
Protein/ peptide/ water	49.05/ 58.13/ 48.35	38.84/ 49.28/ 45.2
R.m.s deviations from ideal values		
Bond lengths (Å)	0.007	0.014
Bond angles (°)	1.049	1.376
Ramachandaran plot (%) [d]		
Favoured region	99.3	99.09
Allowed region	0.7	0.91
Disallowed region	0.0	0.0

[a] Values in parenthesis are for the highest resolution shell.

[b] $R_{merge} = \sum_{hkl}(\sum_i(|I_{hkl,i}-\langle I_{hkl}\rangle|))/\sum_{hkl,i}\langle I_{hkl}\rangle$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices h, k and l, and $\langle I_{hkl}\rangle$ is the mean intensity of that reflection. Calculated for $I > -3\sigma(I)$.

[c] $R_{work} = \sum_{hkl}(||Fobs_{hkl}|-|Fcalc_{hkl}||)/|Fobs_{hkl}|$, where $|Fobs_{hkl}|$ and $|Fcalc_{hkl}|$ are the observed and calculated structure factor amplitudes. R_{free} is equivalent to R_{work} but calculated with reflections (5 %) omitted from the refinement process.

[d] Calculated by the program Molprobity [206].

Table A3. List of all the hits from the screening with proteins harboring a predicted cNLS and a predicted Cdk1 phosphorylation site at the P0 position or P-1 position.

The table excludes isoforms. Proteins for which the indicated phosphorylation site was experimentally confirmed according to the Phosida database (http://www.phosida.com/) are indicated in italics.

P0 position					
Uniprot	**Phosphorylation position**	**NLS sequence**	**Protein name**	**Gene name**	**SNPs at P0 position**
Q9H9L7	22	**S**PKRRRCA	Akirin-1	**AKIRIN1** C1orf108	
Q53H80	21	**S**PKRRRCA	*Akirin-2*	**AKIRIN2** C6orf166	
Q15699	129	**S**SKKRRHR	ALX homeobox protein 1	**ALX1** CART1	
Q13535	428	**S**PKRRRLS	Serine/threonine-protein kinase ATR	**ATR** FRP1	
Q9NYF8	114	**S**PKRRSVS	Bcl-2-associated transcription factor 1	**BCLAF1** BTF KIAA0164	
Q9H0E9	77	**T**PKRKRGE	Bromodomain-containing protein 8	**BRD8** SMAP SMAP2	
Q96S94	400	**S**PKRRKSD	Cyclin-L2	**CCNL2** SB138	
O14646	1328	**S**SKRRKAR	*Chromodomain-helicase-DNA-binding protein 1*	**CHD1**	
P49759	27	**S**HKRRKRS	Dual specificity protein kinase CLK1	**CLK1** CLK	
Q8N684	417	**S**RKRHRSR	*Cleavage and polyadenylation specificity factor*	**CPSF7**	
Q13620	53	**S**AKKRKLN	Cullin-4B	**CUL4B** KIAA0695	
Q09472	12	**S**AKRPKLS	Histone acetyltransferase p300	**EP300** P300	
P17096	53	**T**PKRPRGR	*High mobility group protein HMG-I/HMG-Y*	**HMGA1** HMGIY	
P52926	44	**S**PKRPRGR	High mobility group protein HMGI-C	**HMGA2** HMGIC	
Q9NV88	564	**S**GKKRKRV	Integrator complex subunit 9	**INTS9** RC74	
O43679	254	**T**TKRRKRK	LIM domain-binding protein 2	**LDB2** CLIM1	
P43364	16	**S**IKRKKKR	Melanoma-associated antigen 11	**MAGEA11** MAGE11	
Q02078	496	**S**VKRMRMD	Myocyte-specific enhancer factor 2A	**MEF2A** MEF2	
Q06413	461	**S**VKRMRLS	Myocyte-specific enhancer factor 2C	**MEF2C**	
Q8N5Y2	334	**T**PKRRKAE	Male-specific lethal 3 homolog	**MSL3** MSL3L1	
Q659A1	757	**S**KKRKKIR	NMDA receptor-regulated protein 2	**NARG2** BRCC1 UNQ3101/PRO10100	
O00712	268	**S**SKRPKTI	Nuclear factor 1 B-type	**NFIB**	rs146765479
O00567	556	**S**KKKRKFS	Nucleolar protein 56	**NOP56** NOL5A	
Q14207	1366	**T**TKKRKIE	Protein NPAT	**NPAT** CAND3 E14	

Q7Z417	220	**T**PKKRKAR	*Nuclear fragile X mental retardation-interacting protein 2*	**NUFIP2** KIAA1321 PIG1	
Q96QT6	240	**S**SKRRRKE	PHD finger protein 12	**PHF12** KIAA1523	
Q8TF01	551	**S**PKRKKRH	Arginine/serine-rich protein PNISR	**PNISR** C6orf111 SFRS18 SRRP130 HSPC261 HSPC306	
Q8NAV1	245	**S**PKRRSPS	Pre-mRNA-splicing factor 38A	**PRPF38A**	
Q13523	349	**S**PKRRSLS	*Serine/threonine-protein kinase PRP4 homolog*	**PRPF4B** KIAA0536 PRP4 PRP4H PRP4K	
Q8NDT2	19	**S**AKRPRER	Putative RNA-binding protein 15B	**RBM15B** OTT3	
Q96LT9	115	**S**EKKKRSD	RNA-binding protein 40	**RNPC3** KIAA1839 RBM40 RNP	
Q13127	541	**T**KKKKKVE	RE1-silencing transcription factor	**REST** NRSF XBR	
O15446	459	**S**TKKRKKQ	DNA-directed RNA polymerase I subunit RPA34	**CD3EAP** ASE1 CAST PAF49	
Q14690	41	**S**TKRKKSQ	*Protein RRP5 homolog*	**PDCD11** KIAA0185	
Q9Y6X0	620	**T**KKRKRRR	SET-binding protein	**SETBP1** KIAA0437	
O95104	442	**S**PKRRRSR	Arginine/serine-rich splicing factor 15	**SCAF4** KIAA1172 SFRS15	
O15042	991	**T**PKRSRRS	U2 snRNP-associated SURP motif- containing protein	**U2SURP** KIAA0332 SR140	
Q08170	391	**S**KKKKKED	Serine/arginine-rich splicing factor 4	**SRSF4** SFRS4 SRP75	
Q16629	217	**S**PKRSRSP	Serine/arginine-rich splicing factor 7	**SRSF7** SFRS7	
Q8IWZ8	378	**T**VKRKRKS	SURP and G-patch domain-containing protein 1	**SUGP1** SF4	
P54274	347	**S**TKKKKES	Telomeric repeat-binding factor 1	**TERF1** PIN2 TRBF1 TRF TRF1	
Q12789	1214	**S**QKRKRLK	General transcription factor 3C polypeptide 1	**GTF3C1**	
Q01664	124	**S**PKRRRAE	Transcription factor AP-4	**TFAP4** BHLHC41	
Q15583	162	**S**GKRRRRG	Homeobox protein TGIF1	**TGIF1** TGIF	
Q13769	5	**S**SKKRKPK	THO complex subunit 5 homolog	**THOC5** C22orf19 KIAA0983	
O15405	243	**T**PKKKKKK	TOX high mobility group box family member 3	**TOX3** CAGF9 TNRC9	
Q9NPG3	187	**S**PKKRKLK	Ubinuclein-1	**UBN1**	
Q5T4S7	3366	**S**TKKSKKE	E3 ubiquitin-protein ligase UBR4	**UBR4** KIAA0462 KIAA1307 RBAF600 ZUBR1	
Q96RL1	29	**S**VKRKRRL	BRCA1-A complex subunit RAP80	**UIMC1** RAP80 RXRIP110	
Q96JG9	3786	**S**TKRKKGQ	Zinc finger protein 469	**ZNF469** KIAA1858	
P-1 positions					
Uniprot	**Phosphorylation position**	**NLS sequence**	**Protein name**	**Gene names**	**SNPs at P-1 position**
Q9UIF9	1783	**S**PSKRRRL	Bromodomain adjacent to zinc finger domain protein 2A	**BAZ2A** KIAA0314 TIP5	
Q9NSI6	905	**S**PPKRRRK	Bromodomain and WD repeat-containing protein	**BRWD1** C21orf107 WDR9	

Q8N684	416	**S**SRKRHRS	*Cleavage and polyadenylation specificity factor*	**CPSF7**	
Q13620	52	**T**SAKKRKL	Cullin-4B	**CUL4B** KIAA0695	
O75618	175	**S**QRKRRKS	Death effector domain-containing protein	**DEDD** DEDPRO1 DEFT KE05	
Q9NPF5	459	**S**SVKKAKK	DNA methyltransferase 1-associated protein 1	**DMAP1** KIAA1425	
Q03001	1382	**S**PVKRRRM	Dystonin	**DST** BP230 BP240 BPAG1 DMH DT KIAA0728	
Q92522	31	**S**PSKKRKN	*Histone H1x*	**H1FX**	
P02545	414	**S**VTKKRKL	*Prelamin-A/C*	**LMNA** LMN1	
Q15788	30	**S**TEKRRRE	Nuclear receptor coactivator 1	**NCOA1** BHLHE74 SRC1	
Q14207	1365	**S**TTKKRKI	Protein NPAT	**NPAT** CAND3 E14	
O94913	475	**S**TRKRSRS	Pre-mRNA cleavage complex 2 protein Pcf11	**PCF11** KIAA0824	
Q8TF01	550	**S**SPKRKKR	Arginine/serine-rich protein PNISR	**PNISR** C6orf111 SFRS18 SRRP130 HSPC261 HSPC306	
Q8NDT2	18	**S**SAKRPRE	Putative RNA-binding protein 15B	**RBM15B** OTT3	
Q13127	540	**S**TKKKKKV	RE1-silencing transcription factor	**REST** NRSF XBR	
Q92766	161	**S**PLKRRRL	Ras-responsive element-binding protein 1	**RREB1** FINB	
Q9UQ35	2599	**T**PAKRKRR	Serine/arginine repetitive matrix protein 2	**SRRM2** KIAA0324 SRL300 SRM300 HSPC075	
Q08170	390	**S**SKKKKKE	Serine/arginine-rich splicing factor 4	**SRSF4** SFRS4 SRP75	
Q01664	123	**S**SPKRRRA	Transcription factor AP-4	**TFAP4** BHLHC41	
O43763	153	**T**PPKRKKP	T-cell leukemia homeobox protein 2	**TLX2** HOX11L1 NCX	
O43711	162	**T**PPKRKKP	T-cell leukemia homeobox protein 3	**TLX3** HOX11L2	rs139496015
Q5T4S7	3365	**S**STKKSKK	E3 ubiquitin-protein ligase UBR4	**UBR4** KIAA0462 KIAA1307 RBAF600 ZUBR1	

Table A4. Gene Ontology term enrichment report for hits from the screening with proteins harboring a predicted cNLS and a predicted Cdk1 phosphorylation site at the P0 position or P-1 position.

The table shows the output of a GO enrichment analysis of proteins with predicted cNLSs and phosphorylation sites (72 proteins (isoforms excluded) from UniProtKB with 1310 unique biological process terms) relative to the complete human proteome (20246 proteins retrieved from UniProtKB with 10726 unique biological process terms) using Fisher's exact test. The *E*-value is a Bonferroni corrected *p*-value. We only list GO terms with *E*-value < 0.01. For each significant term, we also specify the number of assigned "hit" proteins vs. the "total" number found in the complete background set.

E-value	*p*-value	GO term	Hits/Total (with GO term)	GO term description
1.50E-17	1.40E-21	GO:0090304	39/1924	Nucleic acid metabolic process
8.40E-14	7.83E-18	GO:0006139	39/2453	Nucleobase-containing compound metabolic process
1.50E-13	1.40E-17	GO:0016070	30/1298	RNA metabolic process
7.50E-12	6.99E-16	GO:0045934	24/847	Negative regulation of nucleobase-containing compound metabolic process
9.30E-12	8.67E-16	GO:0051172	24/855	Negative regulation of nitrogen compound metabolic process
1.50E-11	1.40E-15	GO:0010467	30/1538	Gene expression
1.90E-11	1.77E-15	GO:0034641	39/2871	Cellular nitrogen compound metabolic process
3.20E-11	2.98E-15	GO:0051252	39/2916	Regulation of RNA metabolic process
6.20E-11	5.78E-15	GO:0051253	22/741	Negative regulation of RNA metabolic process
7.20E-11	6.71E-15	GO:0006807	39/2986	Nitrogen compound metabolic process
2.30E-10	2.14E-14	GO:0044260	44/4028	Cellular macromolecule metabolic process
3.90E-10	3.64E-14	GO:0006396	20/630	RNA processing
4.80E-10	4.48E-14	GO:0019219	40/3341	Regulation of nucleobase-containing compound metabolic process
9.10E-10	8.48E-14	GO:0010468	39/3221	Regulation of gene expression
1.10E-09	1.03E-13	GO:0051171	40/3422	Regulation of nitrogen compound metabolic process
2.30E-09	2.14E-13	GO:0010605	24/1098	Negative regulation of macromolecule metabolic process
3.40E-09	3.17E-13	GO:2000113	21/802	Negative regulation of cellular macromolecule biosynthetic process
3.70E-09	3.45E-13	GO:0031324	24/1123	Negative regulation of cellular metabolic process
3.70E-09	3.45E-13	GO:0006397	16/386	mRNA processing
6.50E-09	6.06E-13	GO:0010558	21/830	Negative regulation of macromolecule biosynthetic process
2.10E-08	1.96E-12	GO:0009892	24/1217	Negative regulation of metabolic process
2.70E-08	2.52E-12	GO:0008380	14/300	RNA splicing
3.60E-08	3.36E-12	GO:0031327	21/908	Negative regulation of cellular biosynthetic process
4.20E-08	3.92E-12	GO:0060255	41/4013	Regulation of macromolecule metabolic process
4.40E-08	4.10E-12	GO:0045892	19/715	Negative regulation of transcription, DNA-dependent
4.90E-08	4.57E-12	GO:0009890	21/923	Negative regulation of biosynthetic process
7.30E-08	6.81E-12	GO:0006351	16/470	Transcription, DNA-dependent
9.80E-08	9.14E-12	GO:0043170	44/4749	Macromolecule metabolic process

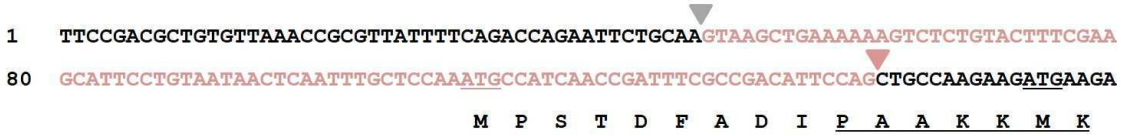
2.10E-07	1.96E-11	GO:0080090	41/4214	Regulation of primary metabolic process
2.50E-07	2.33E-11	GO:0010629	19/790	Negative regulation of gene expression
4.30E-07	4.01E-11	GO:0031323	41/4302	Regulation of cellular metabolic process
7.10E-07	6.62E-11	GO:0006355	33/2829	Regulation of transcription, DNA-dependent
8.40E-07	7.83E-11	GO:2001141	33/2846	Regulation of RNA biosynthetic process
9.10E-07	8.48E-11	GO:0052472	6/21	Modulation by host of symbiont transcription
9.10E-07	8.48E-11	GO:0043921	6/21	Modulation by host of viral transcription
9.20E-07	8.58E-11	GO:0032774	16/557	RNA biosynthetic process
1.10E-06	1.03E-10	GO:2000112	34/3051	Regulation of cellular macromolecule biosynthetic process
1.20E-06	1.12E-10	GO:0052312	6/22	Modulation of transcription in other organism involved in symbiotic interaction
1.80E-06	1.68E-10	GO:0010556	34/3111	Regulation of macromolecule biosynthetic process
2.10E-06	1.96E-10	GO:0045893	19/896	Positive regulation of transcription, DNA-dependent
3.20E-06	2.98E-10	GO:0016071	16/606	mRNA metabolic process
3.50E-06	3.26E-10	GO:0048523	30/2478	Negative regulation of cellular process
5.40E-06	5.03E-10	GO:0051254	19/947	Positive regulation of RNA metabolic process
6.20E-06	5.78E-10	GO:0051851	6/28	Modification by host of symbiont morphology or physiology
6.70E-06	6.25E-10	GO:0010628	19/960	Positive regulation of gene expression
7.20E-06	6.71E-10	GO:0045935	20/1083	Positive regulation of nucleobase-containing compound metabolic process
8.60E-06	8.02E-10	GO:0031326	34/3291	Regulation of cellular biosynthetic process
9.50E-06	8.86E-10	GO:0019222	41/4731	Regulation of metabolic process
1.00E-05	9.32E-10	GO:0043922	5/14	Negative regulation by host of viral transcription
1.00E-05	9.32E-10	GO:0051173	20/1106	Positive regulation of nitrogen compound metabolic process
1.10E-05	1.03E-09	GO:0009889	34/3319	Regulation of biosynthetic process
1.20E-05	1.12E-09	GO:0016568	13/392	Chromatin modification
1.20E-05	1.12E-09	GO:0051702	6/31	Interaction with symbiont
1.70E-05	1.58E-09	GO:0006325	14/489	Chromatin organization
1.90E-05	1.77E-09	GO:0010557	19/1023	Positive regulation of macromolecule biosynthetic process
2.00E-05	1.86E-09	GO:0006366	12/330	Transcription from RNA polymerase II promoter
2.20E-05	2.05E-09	GO:0032897	5/16	Negative regulation of viral transcription
2.30E-05	2.14E-09	GO:0006357	19/1034	Regulation of transcription from RNA polymerase II promoter
3.20E-05	2.98E-09	GO:0048519	30/2715	Negative regulation of biological process
4.30E-05	4.01E-09	GO:0044419	13/436	Interspecies interaction between organisms
1.50E-04	1.40E-08	GO:0006369	6/46	Termination of RNA polymerase II transcription
1.50E-04	1.40E-08	GO:0031328	19/1160	Positive regulation of cellular biosynthetic process
1.70E-04	1.58E-08	GO:0044237	44/5889	Cellular metabolic process
1.90E-04	1.77E-08	GO:0051817	6/48	Modification of morphology or physiology of other organism involved in symbiotic interaction
1.90E-04	1.77E-08	GO:0035821	6/48	Modification of morphology or physiology of other organism
2.00E-04	1.86E-08	GO:0009891	19/1179	Positive regulation of biosynthetic process
3.10E-04	2.89E-08	GO:0000377	9/194	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile
3.10E-04	2.89E-08	GO:0000398	9/194	Nuclear mRNA splicing, via spliceosome

3.10E-04	2.89E-08	GO:0044403	7/89	Symbiosis, encompassing mutualism through parasitism
3.80E-04	3.54E-08	GO:0000375	9/199	RNA splicing, via transesterification reactions
4.90E-04	4.57E-08	GO:0043484	6/56	Regulation of RNA splicing
5.10E-04	4.75E-08	GO:0051276	14/640	Chromosome organization
7.00E-04	6.53E-08	GO:0048024	5/30	Regulation of nuclear mRNA splicing, via spliceosome
7.30E-04	6.81E-08	GO:0034645	19/1281	Cellular macromolecule biosynthetic process
8.00E-04	7.46E-08	GO:0044238	44/6177	Primary metabolic process
8.70E-04	8.11E-08	GO:0016570	9/219	Histone modification
9.70E-04	9.04E-08	GO:0016569	9/222	Covalent chromatin modification
9.80E-04	9.14E-08	GO:0009059	19/1305	Macromolecule biosynthetic process
1.00E-03	9.32E-08	GO:0010604	20/1454	Positive regulation of macromolecule metabolic process
1.80E-03	1.68E-07	GO:0046782	6/69	Regulation of viral transcription
2.30E-03	2.14E-07	GO:0031124	6/72	mRNA 3'-end processing
2.60E-03	2.42E-07	GO:0031325	20/1536	Positive regulation of cellular metabolic process
2.80E-03	2.61E-07	GO:0050684	5/39	Regulation of mRNA processing
3.40E-03	3.17E-07	GO:0048524	6/77	Positive regulation of viral reproduction
5.30E-03	4.94E-07	GO:0006353	6/83	Transcription termination, DNA-dependent
5.60E-03	5.22E-07	GO:0009893	20/1613	Positive regulation of metabolic process
6.60E-03	6.15E-07	GO:0031123	6/86	RNA 3'-end processing
8.90E-03	8.30E-07	GO:2000242	5/49	Negative regulation of reproductive process
9.20E-03	8.58E-07	GO:0090343	3/6	Positive regulation of cell aging

Table A5. *In silico* splice site prediction for *D. melanogaster* and *D. virilis* dUTPase

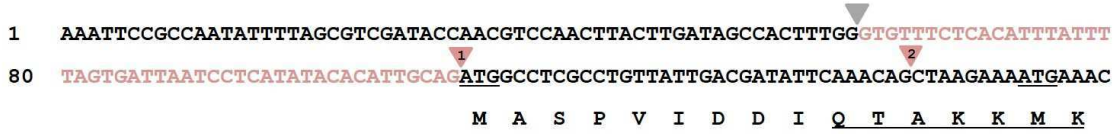
Intronic sequences are in lower case and exonic sequences are in upper case in the *'new potential splice site'* table column. Donor and acceptor site pairs with highest consensus values are highlighted in grey and magenta respectively. NLS sequences are underlined in the N-terminal of the dUTPases. Splicing at the second acceptor site would result in a second dUTPase isoform in *D. virilis*, lacking its N-terminal NLS sequence.

***D. melanogaster in silico* splice site prediction**



Position	Splice site type	Motif	New potential splice site	Consensus value
20	Donor	CGCGTTATT	CGCgttatt	66.79
22	Acceptor	CGTTATTTTCAGAC	cgttattttcagAC	86.78
27	Acceptor	TTTTCAGACCAGAA	ttttcagaccagAA	86.48
38	Acceptor	GAATTCTGCAAGTA	gaattctgcaagTA	70.38
42	Acceptor	TCTGCAAGTAAGCT	tctgcaagtaagCT	71.23
46	Donor	CAAGTAAGC	CAAgtaagc	87.25
52	Acceptor	AGCTGAAAAAAGTC	agctgaaaaaagTC	65.11
69	Acceptor	GTACTTTCGAAGCA	gtactttcgaagCA	74.94
85	Donor	CCTGTAATA	CCTgtaata	69.73
129	Acceptor	CCGACATTCCAGCT	ccgacattccagCT	83.22
137	Acceptor	CCAGCTGCCAAGAA	ccagctgccaagAA	72.52
140	Acceptor	GCTGCCAAGAAGAT	gctgccaagaagAT	71.77

***D. virilis in silico* splice site prediction**



Position	Splice site type	Motif	New potential splice site	Consensus value
9	Acceptor	CCAATATTTTAGCG	ccaatattttagCG	73.55
40	Acceptor	CTTACTTGATAGCC	cttacttgatagCC	77.59
58	Donor	TGGGTGTTT	TGGgtgttt	70.88
71	Acceptor	CATTTATTTTAGTG	catttattttagTG	77.76
79	Donor	TTAGTGATT	TTAgtgatt	69.21
98	Acceptor	TACACATTGCAGAT	tacacattgcagAT	85.18
119	Donor	CCTGTTATT	CCTgttatt	67.04
132	Acceptor	ATATTCAAACAGCT	atattcaaacagCT	81.15
137	Acceptor	CAAACAGCTAAGAA	caaacagctaagAA	67.52
141	Donor	CAGCTAAGA	CAGctaaga	70.83

Table A6. 5’ RACE sequencing results from embryos

Final RACE products were cloned into the SalI/EcoRI sites of pBluescript SK+ vector. Sequencing was performed using the M13 uni(-43) primer. Table shows the sequencing result of 20 clones from the SalI restriction site to the first ATG codon (highlighted in bold and colored gray). Although clones somewhat differ in their 5’ ends, translation starts from the same position from all of them, resulting in only one isoform in *D. virilis*, which contains an NLS signal. Splicing did not occur at either *in silico* predicted splice sites in *D. virilis*.

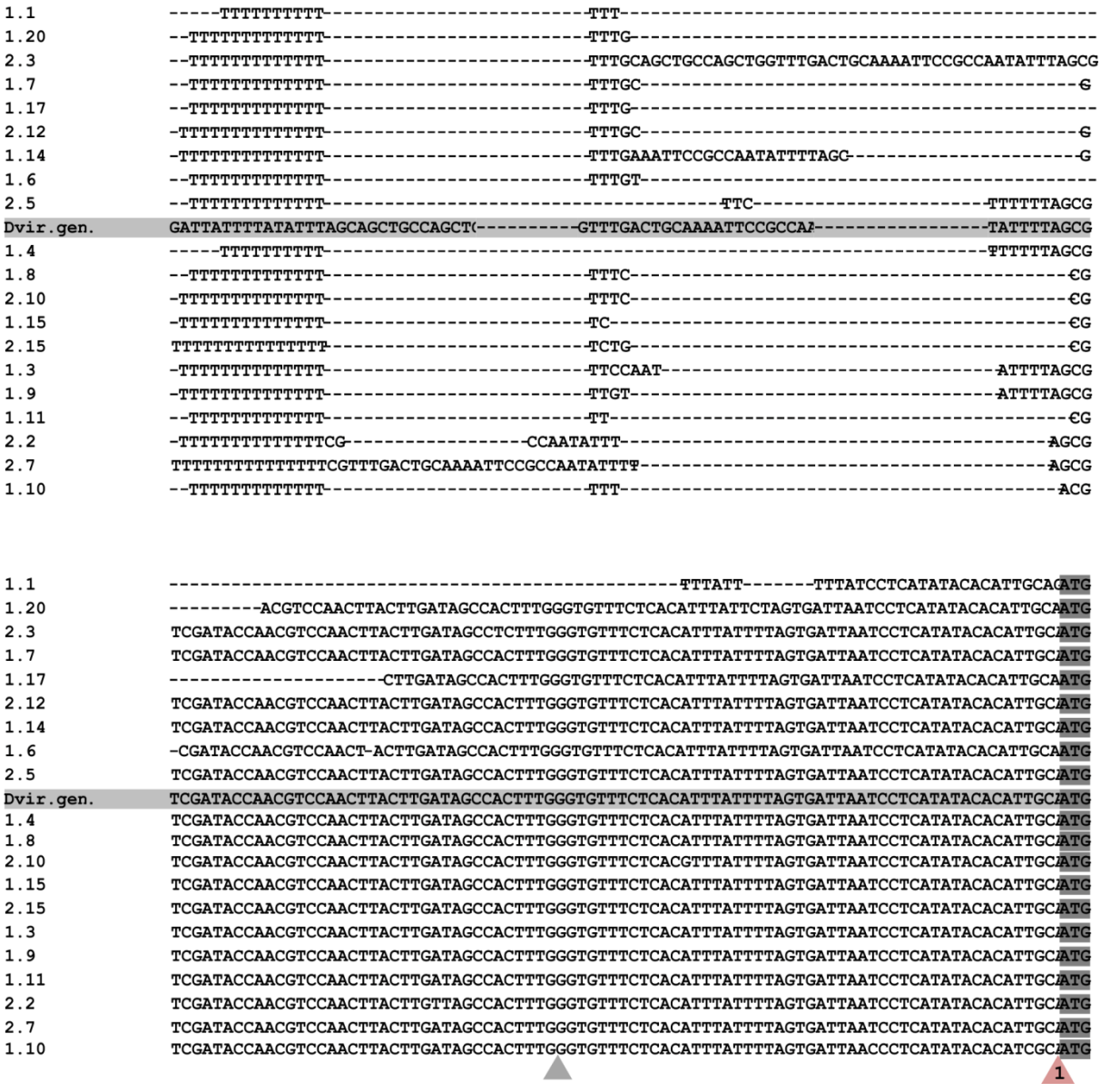


Table A7. Oligonucleotides used in the study

Aim	Phospho	Oligo name	Oligonucleotides 5'-3'
Cloning into pGal-DsRed reporter	testing NLS function on pGal-DsRed	SV40_WT_F	CTAGCATGGGAGCTTCACCCAAGAAGAAGAGAAAGGTGGG
		SV40_WT_R	AATTCCCACCTTTCTCTTCTTCTTGGGTGAAGCTCCCATG
	S160	Swi6_WT_F	CTAGCATGGGAGCTTCACCCCTGAAGAAGCTGAAGATCGACGG
		Swi6_WT_R	AATTCCGTCGATCTTCAGCTTCTTCAGGGGTGAAGCTCCCATG
		Swi6_P-1E_F	CTAGCATGGGAGCTGAACCCCTGAAGAAGCTGAAGATCGACGG
		Swi6_P-1E_R	AATTCCGTCGATCTTCAGCTTCTTCAGGGGTTCAGCTCCCATG
		Swi6_P-2E_F	CTAGCATGGGAGCTGAACCCGCCCTGAAGAAGCTGAAGATCGACGG
		Swi6_P-2E_R	AATTCCGTCGATCTTCAGCTTCTTCAGGGCGGGTTCAGCTCCCATG
		Swi6_P0_F	CTAGCATGGGAGCTGAACCCAAGAAGCTGAAGATCGACGG
		Swi6_P0_R	AATTCCGTCGATCTTCAGCTTCTTGGGTTCAGCTCCCATG
	S428	ATR_WT_F	CTAGCATGGGAGCTATCAGCCCCAAGAGAAGAAGACTGGG
		ATR_WT_R	AATTCCCAGTCTTCTTCTCTTGGGGCTGATAGCTCCCATG
		ATR_E_F	CTAGCATGGGAGCTATCGAACCCAAGAGAAGAAGACTGGG
		ATR_E_R	AATTCCCAGTCTTCTTCTCTTGGGTTCGATAGCTCCCATG
	S400	CCLN2_WT_F	CTAGCATGGGAGCTGCCAGCCCCAAGAGAAGAAAGAGCGG
		CCLN2_WT_R	AATTCCGCTCTTTCTTCTCTTGGGGCTGGCAGCTCCCATG
		CCLN2_E_F	CTAGCATGGGAGCTGCCGAACCCAAGAGAAGAAAGAGCGG
		CCLN2_E_R	AATTCCGCTCTTTCTTCTCTTGGGTTCGGCAGCTCCCATG
	S124 (P0) S123 (P-1)	TFAP4_WT_F	CTAGCATGGGAGCTAGCAGCCCCAAGAGAAGAAGAGCCGG
		TFAP4_WT_R	AATTCCGGCTCTTCTTCTCTTGGGGCTGCTAGCTCCCATG
		TFAP4_EP0_F	CTAGCATGGGAGCTAGCGAGCCCAAGAGAAGAAGAGCCGG
		TFAP4_EP0_R	AATTCCGGCTCTTCTTCTCTTGGGCTCGCTAGCTCCCATG
		TFAP4_EP-1_F	CTAGCATGGGAGCTGAGAGCCCCAAGAGAAGAAGAGCCGG
		TFAP4_EP-1_R	AATTCCGGCTCTTCTTCTCTTGGGGCTCTCAGCTCCCATG
	S53 (P0) T52 (P-1)	CUL4B_WT_F	CTAGCATGGGAGCTACCAGCGCCAAGAAGAGAAAGCTGGG
		CUL4B_WT_R	AATTCCCAGCTTTCTCTTCTTGGCGCTGGTAGCTCCCATG
		CUL4B_EP0_F	CTAGCATGGGAGCTACCGAGGCCAAGAAGAGAAAGCTGGG
		CUL4B_EP0_R	AATTCCCAGCTTTCTCTTCTTGGCCTCGGTAGCTCCCATG
		CUL4B_EP-1_F	CTAGCATGGGAGCTGAGAGCGCCAAGAAGAGAAAGCTGGG
		CUL4B_EP-1_R	AATTCCCAGCTTTCTCTTCTTGGCGCTCTCAGCTCCCATG
	S12	EP300_WT_F	CTAGCATGGGAGCTCCCAGCGCCAAGAGACCCAAGCTGGG
		EP300_WT_R	AATTCCCAGCTTGGGTCTCTTGGCGCTGGGAGCTCCCATG
		EP300_E_F	CTAGCATGGGAGCTCCCGAAGCCAAGAGACCCAAGCTGGG
		EP300_E_R	AATTCCCAGCTTGGGTCTCTTGGCTTCGGGAGCTCCCATG
	S29	UIMC1_WT_F	CTAGCATGGGAGCTGTGAGCGTGAAGAGAAAGAGAAGAGG
		UIMC1_WT_R	AATTCCTCTTCTCTTTCTCTTCACGCTCACAGCTCCCATG
		UIMC1_E_F	CTAGCATGGGAGCTGTGGAAGTGAAGAGAAAGAGAAGAGG
		UIMC1_E_R	AATTCCTCTTCTCTTTCTCTTCACTTCCACAGCTCCCATG
	S161	RREB1_WT_F	CTAGCATGGGAGCTAGCCCCCTGAAGAGAAGAAGACTGGG
		RREB1_WT_R	AATTCCCAGTCTTCTTCTCTTCAGGGGGCTAGCTCCCATG
		RREB1_E_F	CTAGCATGGGAGCTGAACCCCTGAAGAGAAGAAGACTGGG
		RREB1_E_R	AATTCCCAGTCTTCTTCTCTTCAGGGGTTCAGCTCCCATG
	S162	TLX3_WT_F	CTAGCATGGGAGCTACCCCCCCCAAGAGAAAGAAGCCCGG

		TLX3_WT_R	AATTCCGGGCTTCTTTCTCTTGGGGGGGGTAGCTCCCATG
		TLX3_E_F	CTAGCATGGGAGCTGAGCCCCCCAAGAGAAAGAAGCCCGG
		TLX3_E_R	AATTCCGGGCTTCTTTCTCTTGGGGGGCTCAGCTCCCATG
	S2599	SRRM2_WT_F	CTAGCATGGGAGCTACCCCCGCCAAGAGAAAGAGAAGAGG
		SRRM2_WT_R	AATTCCTCTTCTCTTTCTCTTGGCGGGGGTAGCTCCCATG
		SRRM2_E_F	CTAGCATGGGAGCTGAGCCCGCCAAGAGAAAGAGAAGAGG
		SRRM2_E_R	AATTCCTCTTCTCTTTCTCTTGGCGGGCTCAGCTCCCATG
	S4	UBA E1_WT_F	CTAGCATGGGAGCTTCGCCGCTGTCCAAGAAACGTCGCGTGGG
		UBA E1_WT_R	AATTCCCACGCGACGTTTCTTGGACAGCGGCGAAGCTCCCATG
		UBA E1_EP-2_F	CTAGCATGGGAGCTGAACCGCTGTCCAAGAAACGTCGCGTGGG
		UBA E1_EP-2_R	AATTCCCACGCGACGTTTCTTGGACAGCGGTTCAGCTCCCATG
		UBA E1_EP-1_F	CTAGCATGGGAGCTGAACCGTCCAAGAAACGTCGCGTGGG
		UBA E1_EP-1_R	AATTCCCACGCGACGTTTCTTGGACGGTTCAGCTCCCATG
	S14	UNG_WT_F	CTAGCATGGGAGCTTCACCCGCCAGGAAGCGACACGCCCCCGG
		UNG_WT_R	AATTCCGGGGGCGTGTCGCTTCCTGGCGGGTGAAGCTCCCATG
		UNG_EP-2_F	CTAGCATGGGAGCTGAACCCGCCAGGAAGCGACACGCCCCCGG
		UNG_EP-2_R	AATTCCGGGGGCGTGTCGCTTCCTGGCGGGTTCAGCTCCCATG
		UNG_EP-1_F	CTAGCATGGGAGCTGAACCCAGGAAGCGACACGCCCCCGG
		UNG_EP-1_R	AATTCCGGGGGCGTGTCGCTTCCTGGGTTCAGCTCCCATG
	K18N	UNG_K18N_F	CTAGCATGGGAGCTAGCCCCGCCAGGAACCGACACGCCCCCGG
		UNG_K18N_R	AATTCCGGGGGCGTGTCGGTTCCTGGCGGGGCTAGCTCCCATG
	S315	p53_WT_F	CTAGCATGGGAGCTAAGCGAGCACTGCCCAACAACACCAGCTCCTCTCCCCAGCCAAA GAAGAAACCACTGGG
		p53_WT_E	AATTCCCAGTGGTTTCTTCTTTGGCTGGGGAGAGGAGCTGGTGTTGTTGGGCAGTGCTC GCTTAGCTCCCATG
		p53_EP-2_F	CTAGCATGGGAGCTAAGCGAGCACTGCCCAACAACACCAGCTCCGAACCCCAGCCAAA GAAGAAACCACTGGG
		p53_EP-2_R	AATTCCCAGTGGTTTCTTCTTTGGCTGGGGTTCGGAGCTGGTGTTGTTGGGCAGTGCTC GCTTAGCTCCCATG
		p53_EP-1_F	CTAGCATGGGAGCTAAGCGAGCACTGCCCAACAACACCAGCTCCGAACCCCCAAAGAA GAAACCACTGGG
		p53_EP-1_R	AATTCCCAGTGGTTTCTTCTTTGGGGGTTCGGAGCTGGTGTTGTTGGGCAGTGCTCGCT TAGCTCCCATG
Cloning into pHM830 reporter	S11	DUT_WT_F	TTAAGGGCCTTGCCTCACCCAGTAAGCGGGCCCGGCCTGCGT
		DUT_WT_R	CTAGACGCAGGCCGGGCCCGCTTACTGGGTGAGGCAAGGCCC
		DUT_E_F	TTAAGGGCCTTGCCGAACCCAGTAAGCGGGCCCGGCCTGCGT
		DUT_E_F	CTAGACGCAGGCCGGGCCCGCTTACTGGGTTCGGCAAGGCCC
	S414	LMNA_WT_F	TTAAGGGCCTTGCCAGCGTGACCAAGAAGAGAAAGCTGT
		LMNA_WT_R	CTAGACAGCTTTCTCTTCTTGGTCACGCTGGCAAGGCCC
		LMNA_E_F	TTAAGGGCCTTGCCGAAGTGACCAAGAAGAGAAAGCTGT
		LMNA_E_R	CTAGACAGCTTTCTCTTCTTGGTCACTTCGGCAAGGCCC
	S1783	BAZ2A_WT_F	TTAAGGGCCTTGCCAGCCCCAGCAAGAGAAGAAGACTGT
		BAZ2A_WT_R	CTAGACAGTCTTCTTCTCTTGCTGGGGCTGGCAAGGCCC
		BAZ2A_E_F	TTAAGGGCCTTGCCGAACCCAGCAAGAGAAGAAGACTGT
		BAZ2A_WT_R	CTAGACAGTCTTCTTCTCTTGCTGGGTTCGGCAAGGCCC
Cloning dUTPase into pDsRed-Monomer-N1		dutN1F	GATCTCGAGATGCCCTGCTCTGAAGAGAC
		dutN1R	ACGGTACCGCATTCTTTCCAGTGGAACCAAAAC
Cloning β-gal into pDsRed-Monomer-N1		galN1F	GCAAGAATTCCAGCATCGTTTACTTTGACCAACAAGAACG
		galN1R	AATTGGTACCGCTTTTTGACACCAGACCAACTGGTAATGG

Cloning dUTPase into pET20b		dutpETF	GGAATTCCATATGCCCTGCTCTGAAGAGACAC
		dutpETR	CCGCTCGAGCTGGGAGCCGGAGTGG
Cloning p53 into pDsRed-Monomer-C1		p53_F	AGCTCTCGAGATGGAGGAGCCGCAGTCAGATC
		p53_R	CACTGGATCCTCAGTCTGAGTCAGGCCCTTCTG
Cloning UBA1 into pDsRed-Monomer-N1		UBA1_F	CAGCTGGTACCATGTCCAGCTCGCCGCTGTCC
		UBA1_R	AGATGGATCCGCTCCGCGGATGGTGTATCGGACATAGG
Cloning UNG2 into pDsRed-Monomer-N1		UNG2_F	CGATCTCGAGATGATCGGCCAGAAGACGC
		UNG2_R	ACGGTACCGCGATGTACCTGTAGGTGTCCAGC
Mutagenesis oligos on full length proteins	**dUTPase**	S11E_F	GAGACACCCGCCATTGAACCCAGTAAGCGGGC
		S11E_R	GCCCGCTTACTGGGTTCAATGGCGGGTGTCTC
		S11Q_F	GAGACACCCGCCATTCAACCCAGTAAGCGGGC
		S11Q_R	GCCCGCTTACTGGGTTGAATGGCGGGTGTCTC
	p53	S315E_F (P-2)	CCCAACAACACCAGCTCCGAACCCCAGCCAAAGAAGAAACC
		S315E_R (P-2)	GGTTTCTTCTTTGGCTGGGGTTCGGAGCTGGTGTTGTTGGG
		Q317Δ_F (P-1)	ACACCAGCTCCGAACCCCCAAAGAAGAAACCACTGGATG
		Q317Δ_F (P-1)	CATCCAGTGGTTTCTTCTTTGGGGGTTCGGAGCTGGTGT
	UBA1	S4E_F (P-2)	CGACGGTACCATGTCCAGCGAACCGCTGTCCAAGAAACGTC
		S4E_R (P-2)	GACGTTTCTTGGACAGCGGTTCGCTGGACATGGTACCGTCG
		L6Δ_F (P-1)	GTACCATGTCCAGCGAACCGTCCAAGAAACGTCGCGTGTC
		L6Δ_R (P-1)	GACACGCGACGTTTCTTGGACGGTTCGCTGGACATGGTAC
	UNG2	S14E_F (P-2)	CTACTCCTTTTTCTCCCCCGAACCCGCCAGGAAGCGAC
		S14E_R (P-2)	GTCGCTTCCTGGCGGGTTCGGGGGAGAAAAAGGAGTAG
		A16Δ_F (P-1)	CCTTTTTCTCCCCCGAACCCAGGAAGCGACACGCCCC
		A16Δ_R (P-1)	GGGGCGTGTCGCTTCCTGGGTTCGGGGGAGAAAAAGG
Cloning into pEGFP-N1 *(Drosophila DUT)*		ABC_R	AATTCCCGGGCTGTGGAAATGGGTGTAGCATCATTTTCG
		A_F	AATTGGTACCATGGCCTCGCCTGTTATTGACG
		A_R	AATTCCCGGGCAGTTACAATCTGCTGCTTGTCTCCTGG
Cloning into pET15b *(Drosophila DUT)*		DvirDF	GGAATTCCATATGGCCTCGCCTGTTATTGACG
		DvirDR3	CGAAGATCTCTATGTGGAAATGGGTGTAGCATCATTTTC
		DvirDRs1	CGAAGATCTAAGTTACAATCTGCTGCTTGTCTCCTGG
5'Race *(Drosophila DUT)*		SP1 green	CGACCTGTTCCAGCTCCGG
		Oligo dT-anchor primer	GACCACGCGTATCGATGTCGACTTTTTTTTTTTTTTTV
		SP2	CGATGGATTCCCAGCTCCGGAATGGATATGC
		PCR anchor primer	GACCACGCGTATCGATGTCGAC
		SP3	GGTAGAATTCGCGATCACCGCGCTTGACC
		M13 uni(-43)	AGGGTTTTCCCA GTCACGACGTT
NLS deletion *(Drosophila DUT)*		NLSdelF	CCTCGCCTGTTATTGACGATATTGATAAGAAATGCGTTTTGCGTTATG
		NLSdelR	CATAACGCAAAACGCATTTCTTATCAATATCGTCAATAACAGGCGAGG
Au1 tagging *(Drosophila DUT)*		A_Au1_F	GATCCAGGCGCTAGCCCGGTCGCCACCGACACCTACCGCTACATCGGCGCTTAAGC
		A_Au1_R	GGCCGCTTAAGCGCCGATGTAGCGGTAGGTGTCGGTGGCGACCGGGCTAGCGCCTG
		ABC_Au1_F	AATTGGTACCATGGCCTCGCCTGTTATTGACG
		ABC_Au1_R	ATGAATTCTTAAGCGCCGATGTAGCGGTAGGTGTCGGTGGCGACCGGGCTAGCGCCTGGATCCCGGGCTGTGG

[38]ADATLAP
a doktori értekezés nyilvánosságra hozatalához

I. A doktori értekezés adatai

A szerző neve: **Róna Gergely**
MTMT-azonosító: **10027694**
A doktori értekezés címe és alcíme:
Mechanism and regulation of dUTPase nucleocytoplasmic transport with an outlook on cell cycle dependent nuclear proteome reconstruction
DOI-azonosító[39]: **10.15476/ELTE.2015.025**
A doktori iskola neve: **ELTE TTK Biológia Doktori Iskola**
A doktori iskolán belüli doktori program neve: **Szerkezeti Biokémia Doktori Program**
A témavezető neve és tudományos fokozata: **Vértessy G. Beáta, Ph.D., D.Sc., Professzor**
A témavezető munkahelye: **MTA TTK Enzimológiai Intézet valamint az Alkalmazott Biotechnológia és Élelmiszertudományi Tanszék, BME**

II. Nyilatkozatok

A doktori értekezés szerzőjeként[40]
a) hozzájárulok, hogy a doktori fokozat megszerzését követően a doktori értekezésem és a tézisek nyilvánosságra kerüljenek az ELTE Digitális Intézményi Tudástárban. Felhatalmazom a Természettudományi Kar Tudományszervezési és Egyetemközi Kapcsolatok Osztályának ügyintézőjét ……, hogy az értekezést és a téziseket feltöltse az ELTE Digitális Intézményi Tudástárba, és ennek során kitöltse a feltöltéshez szükséges nyilatkozatokat.
b) kérem, hogy a mellékelt kérelemben részletezett szabadalmi, illetőleg oltalmi bejelentés közzétételéig a doktori értekezést ne bocsássák nyilvánosságra az Egyetemi Könyvtárban és az ELTE Digitális Intézményi Tudástárban;[41]
c) kérem, hogy a nemzetbiztonsági okból minősített adatot tartalmazó doktori értekezést a minősítés *(dátum*)-ig tartó időtartama alatt ne bocsássák nyilvánosságra az Egyetemi Könyvtárban és az ELTE Digitális Intézményi Tudástárban;[42]
d) kérem, hogy a mű kiadására vonatkozó mellékelt kiadó szerződésre tekintettel a doktori értekezést a könyv megjelenéséig ne bocsássák nyilvánosságra az Egyetemi Könyvtárban, és az ELTE Digitális Intézményi Tudástárban csak a könyv bibliográfiai adatait tegyék közzé. Ha a könyv a fokozatszerzést követőn egy évig nem jelenik meg, hozzájárulok, hogy a doktori értekezésem és a tézisek nyilvánosságra kerüljenek az Egyetemi Könyvtárban és az ELTE Digitális Intézményi Tudástárban.[43]
2. A doktori értekezés szerzőjeként kijelentem, hogy
a) az ELTE Digitális Intézményi Tudástárba feltöltendő doktori értekezés és a tézisek saját eredeti , önálló szellemi munkám és legjobb tudomásom szerint nem sértem vele senki szerzői jogait;
b) a doktori értekezés és a tézisek nyomtatott változatai és az elektronikus adathordozón benyújtott tartalmak (szöveg és ábrák) mindenben megegyeznek.
3. A doktori értekezés szerzőjeként hozzájárulok a doktori értekezés és a tézisek szövegének plágiumkereső adatbázisba helyezéséhez és plágiumellenőrző vizsgálatok lefuttatásához.

Kelt: **2015. február 4.**

……
a doktori értekezés szerzőjének aláírása

[38] Beiktatta az Egyetemi Doktori Szabályzat módosításáról szóló CXXXIX/2014. (VI. 30.) Szen. sz. határozat. Hatályos: 2014. VII.1. napjától.
[39] A kari hivatal ügyintézője tölti ki.
[40] A megfelelő szöveg aláhúzandó.
[41] A doktori értekezés benyújtásával egyidejűleg be kell adni a tudományági doktori tanácshoz a szabadalmi, illetőleg oltalmi bejelentést tanúsító okiratot és a nyilvánosságra hozatal elhalasztása iránti kérelmet.
[42] A doktori értekezés benyújtásával egyidejűleg be kell nyújtani a minősített adatra vonatkozó közokiratot.
[43] A doktori értekezés benyújtásával egyidejűleg be kell nyújtani a mű kiadásáról szóló kiadói szerződést.