FEBS Letters

FEBS PRESS science publishing by scientists

RESEARCH LETTER

Structural determinants of phosphorylation-dependent nuclear transport of HCMV DNA polymerase processivity factor UL44

Emily M. Cross[1,2], Oriano Marin[3], Daryl Ariawan[4], David Aragão[2], Giorgio Cozza[5], Enzo Di Iorio[5], Jade K. Forwood[1] and Gualtiero Alvisi[5]

1 School of Dentistry and Medical Sciences, Charles Sturt University, Wagga Wagga, Australia
2 Diamond Light Source, Didcot, UK
3 Department of Biomedical Sciences, University of Padova, Italy
4 Dementia Research Centre, Macquarie University, Sydney, Australia
5 Department of Molecular Medicine, University of Padua, Italy

Correspondence
G. Alvisi, Department of Molecular Medicine, University of Padua, 35121 Padova, Italy
Tel: +39 (0)498272353
E-mail: gualtiero.alvisi@unipd.it

(Received 30 October 2023, revised 8 December 2023, accepted 11 December 2023, available online 8 January 2024)

doi:10.1002/1873-3468.14797

Edited by John Briggs

Human cytomegalovirus DNA polymerase processivity factor UL44 is transported into the nucleus by importin (IMP) α/β through a classical nuclear localization signal (NLS), and this region is susceptible to cdc2-mediated phosphorylation at position T427. Whilst phosphorylation within and close to the UL44 NLS regulates nuclear transport, the details remain elusive, due to the paucity of structural information regarding the role of negatively charged cargo phosphate groups. We addressed this issue by studying the effect of UL44 T427 phosphorylation on interaction with several IMPα isoforms by biochemical and structural approaches. Phosphorylation decreased UL44/IMPα affinity 10-fold, and a comparative structural analysis of UL44 NLS phosphorylated and non-phosphorylated peptides complexed with mouse IMPα2 revealed the structural rearrangements responsible for phosphorylation-dependent inhibition of UL44 nuclear import.

Keywords: IMP; importins; NLS; nuclear transport; phosphorylation; regulation

Active nuclear import is a signal and energy-dependent process mediated by karyophilic proteins belonging to the importin (IMP) superfamily, whereby IMPβ1 or one of its several homologs recognize nuclear localization signals (NLSs) on the cargo protein, translocate the latter through the nuclear pore complex (NPC), and release it in the nucleoplasm upon binding the small GTPase Ran [1]. The best-characterized NLSs are highly basic sequences recognized by IMPβ1 through the adapter molecule IMPα [2], and are known as classical (c)NLSs. Human IMPαs exist as seven isoforms (hIMPα1/3/4/5/6/7/8), with the mouse homolog of hIMPα1, "mIMPα2", frequently used for structural studies [3–7]. cNLSs can either be monopartite or bipartite. Monopartite cNLSs are formed by a single stretch of basic amino acids and bind to the IMPα major binding site, which contains six main binding pockets (P0–P6) able to accommodate NLS aminoacidic side chains. Bipartite cNLSs are formed by two stretches of basic amino acids typically

Abbreviations
ARM, armadillo repeat/domain; cNLS, classical NLS; DNA, deoxyribonucleic acid; EBNA1, Epstein–Barr nuclear antigen; FITC, fluorescein isothiocyanate; FP, fluorescence polarization; HCMV, human cytomegalovirus; IBB, importin-β binding; IMP, importin; LTA, large T antigen; NLS, nuclear localization signal; NPC, nuclear pore complex; PDB, protein data bank; PTO or pT, phosphorylated threonine; RT, room temperature; SEM, standard error of the mean; SV40, Simian vacuolating virus 40; TEV, tobacco etch virus; UL44, DNA polymerase processivity factor UL44.

separated by an amino acid linker region and interact with both IMPα major and minor binding sites. The minor site contains four main binding pockets (P1′–P4′) and this additional binding interface lowers the K_d of the interaction [3,4]. A summary of the two binding pockets can be found in Fig. 1.

Nuclear transport of several cargoes is tightly regulated by a number of mechanisms, the best characterized being cargo phosphorylation [8]. However, the lack of extensive structural data leaves several unanswered questions [9]. Firstly, it is extremely challenging to predict the outcome of cargo phosphorylation on nuclear import. Indeed, cargo phosphorylation can either promote [10–15] or impair the process [11,16–20]. Secondly, although phosphorylation-dependent regulation of nuclear import has been proposed to rely on the modulation of cargo-IMP interactions, it remains uncertain whether the latter depends on a change in affinity for IMPs [13,17] or for other factors that compete with IMPs [21–23]. Additionally, phosphorylation of certain cargoes has been proposed to result in conformational changes regulating NLS exposure, and therefore IMP binding and subsequent nuclear import, as seen in hepatitis B core protein where phosphorylation likely acts as an important modulator of the viral life cycle [24].

Nuclear transport of human cytomegalovirus (HCMV) DNA polymerase processivity factor UL44 is mediated by the importin IMPα/β heterodimer, whereby IMPα recognizes a C-terminally located NLS (PNTKKQK-431) and IMPβ translocates the complex to the nucleoplasmic side of the NPC [10]. Upstream of the NLS, a stretch of serine residues (Fig. 2, indicated as red in the single letter amino acid sequence), has been shown to be the target of a phosphorylation cascade mediated by protein kinases CKII and CKI, which results in enhanced nuclear localization [11,29].

Conversely, residue T427, a target for cdk1/cycB-mediated phosphorylation *in vitro*, lies immediately upstream of the UL44-NLS basic cluster that mediates the interaction with IMPα. The phosphomimic D427 substitution has been shown to reduce both the nuclear import of exogenously expressed GFP-UL44 by live cell confocal laser scanning microscopy and *ori*Lyt-dependent DNA replication [11,29]. Although the negative regulator of nuclear import BRAP2 has been shown to decrease nuclear import of the T427D phosphomimic UL44 derivative, the impact of phosphorylation on T427 on UL44 interaction with IMPα remains unknown. We approached this question by biochemically and structurally comparing the binding properties of several IMPα isoforms to an unphosphorylated UL44 NLS peptide (UL44_410-433; 410-KEESDSEDSVTFEFVPNTKKQKCG-433) and its pT427 counterpart (UL44_410-433_pT 410-KEESDSEDSVTFEFVPNTKKQKCG-433)

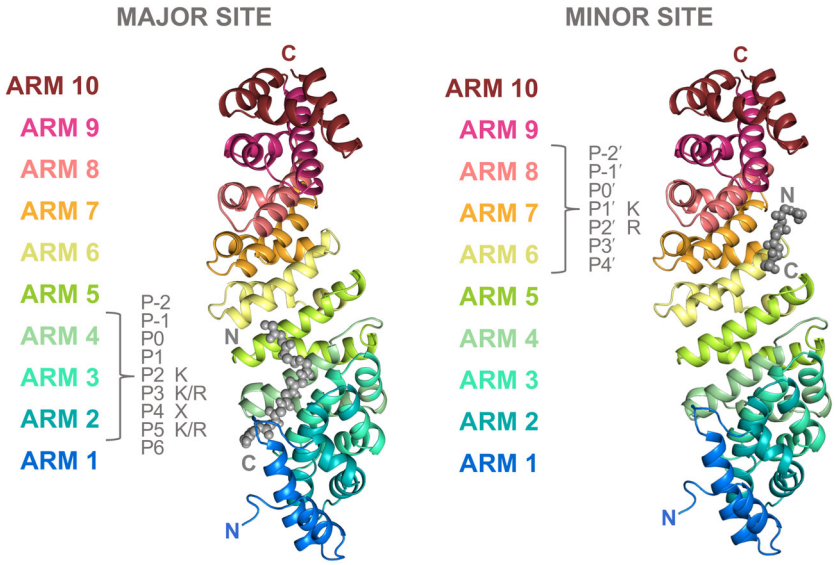


Fig. 1. Major and minor binding sites of mIMPα2. ARM repeats of mouse IMPα2 are labeled and color coded (rainbow). The major binding site is found within ARMS 2–4 where the NLS binds at P0–P6 positions, typically with a K-K/R-X-K/R consensus at the P2–P5 sites. The minor binding site is found within ARMS 6–8 where the NLS binds at P1′–P4′ positions, typically with a K-R consensus at the P1′ and P2′ sites, respectively. Bipartite NLSs span continuously across both major and minor binding sites with an amino acid linker region between each (not shown) [3,4].

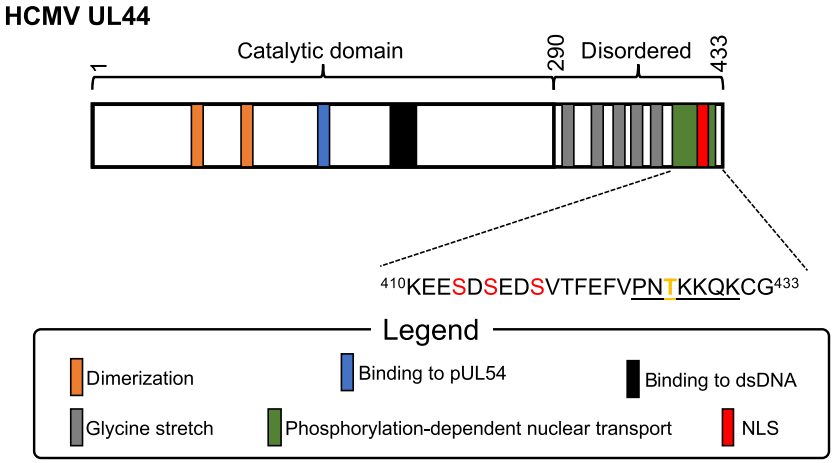


Fig. 2. Schematic representation of UL44 functional domains and determinants of phosphorylation regulated nuclear import. The HCMV UL44 sequence is represented by a horizontal white bar. Its catalytic N-terminal domain (residues 1–290) and its highly unstructured C-terminal domain (residues 291–433) are indicated. The vertical orange bars represent residues essential for dimerization (L86, L87, and F121). The vertical blue bar indicates the residue essential for binding to DNA polymerase catalytic subunit pUL54 I135 in the connector loop. Vertical black bars are residues essential for DNA binding (residues 165, 167, and 168 in the gap loop). Vertical gray bars represent glycine-rich stretches. The vertical green bar represents the phosphorylation-dependent nuclear transport region (residues 410–433) and the vertical red bar is the NLS (residues 425–431). The protein sequence encompassing the UL44 phosphorylation region and NLS (residues 410–433) is shown using the single letter amino acid code, with phosphorylation sites enhancing nuclear import depicted in red, and phosphorylation sites decreasing nuclear import depicted in yellow. The NLS is underlined [10,11,25–28].

(phosphorylated amino acid is underlined). Quantitative fluorescence polarization assays showed that phosphorylation directly affects the ability of UL44 to interact with IMPαs, decreasing binding affinity by c. 10-fold. Crystallographic analysis of peptides in complex with mIMPα2 confirmed that UL44 residues 425–431 represent a cNLS, and revealed subtle, but important differences in the IMPα interaction induced by phosphorylation of T427, thus partially explaining the molecular details of phosphorylation-dependent regulation of nuclear import. Despite the main features of the NLS–IMPα interaction being unaffected, with K428, K429, and K431 occupying the key P2, P3, and P5 positions in the mIMPα2 major binding site, a small number of structural changes were observed, including an important conformational change of mIMPα2 R238 sidechain interactions.

Materials and methods

Recombinant protein expression and purification

Representatives from each importin (IMP) superfamily (hIMPα3,5,7 and mIMPα2) were recombinantly expressed and purified for assessment of UL44 peptide binding (hIMPα3,5,7 and mIMPα2) and structure determination (mIMPα2). For this, plasmids encoding IMPs with truncated IMPβ binding (IBB) domains (first ~70 amino acids removed, summary available in Table S1), and an N-terminal His tag separated by a TEV protease cleavage site, were codon optimized and inserted into pET-30 vectors at BamHI sites (Genscript, Piscataway, NJ, USA). The plasmids were transformed into competent BL21 (DE3) pLysS *Escherichia coli* cells and expressed for 24–48 h using autoinduction. Cells were harvested and suspended in affinity buffer (50-mM phosphate buffer, 300-mM sodium chloride, 20-mM imidazole, pH 8.0). Before purification, the whole cell was freeze–thawed thrice, then lysed with lysozyme and treated with DNAse. The resulting supernatant containing soluble protein was collected and filtered, injected over a Nickel HisTrap HP 5 mL (Cytvia, Marlborough, MA, USA), washed with affinity buffer, then with elution buffer (50-mM phosphate buffer, 300-mM sodium chloride, 500-mM imidazole, pH 8.0). Peak fractions were collected and treated with TEV protease, incubated at 4 °C overnight (excluding IMPα2 which does not contain a TEV site). The sample was then injected over a HiLoad 26/600 Superdex 200 pg column (Cytvia) using a Tris buffer (50-mM Tris, 125-mM sodium chloride, pH 8.0). Purified importin protein was concentrated using Amicon Ultra-15 Centrifugal Filter Units (Merck, Millipore, Burlington, VT, USA), then aliquoted, and stored at −80 °C until further use in binding and/or crystallization experiments.

Synthesis of UL44 peptides

The UL44_410-433 and UL44_410-433_pT peptides were chemically synthesized employing a solid-phase technique

performed on a fully automated peptide synthesizer (Syro II; MultiSynTechGmbh, Witten, Germany). Wang resins preloaded with the first *N*-α-Fmoc-protected amino acid were employed for stepwise assembly of the entire peptide chain. This assembly was performed according to the Fmoc standard strategy and was based on the use of HATU as the coupling reagent [30]. The side-chain protected amino acid building blocks were: *N*-α-Fmoc-β-tert-butyl-L-glutamic acid, *N*-α-Fmoc-β-tert-butyl-L-aspartic acid, *N*-α-Fmoc-*N*δ-trityl-L-glutamine, *N*-α-Fmoc-*N*γ-trityl-L-asparagine, *N*-α-Fmoc-*O*-benzyl-L-phosphoserine, *N*-α-Fmoc-*O*-tert-butyl-L-serine, *N*-α-Fmoc-*O*-tert-butyl-L-threonine, *N*-α-Fmoc-*O*-benzyl-L-phosphothreonine, *N*-α-Fmoc-*N*ε-tert-butyloxycarbonyl-L-lysine, and *N*-α-Fmoc-*S*-trityl-L-cysteine. A step of deprotection of the final peptides was conducted, followed by cleavage from the resin with a mixture of 88% (v/v) trifluoroacetic acid (TFA) with 5% phenol (w/v), 5% H_2O (v/v), and 2% (v/v) of triisopropylsilane *via* shaking at RT for 2.5 h. A step of vacuum filtration permitted the removal of the resin from the assembled peptide chains. Then, the peptides were precipitated with cold diethyl ether and transformed into pellets by a centrifugation procedure. Two washes with cold diethyl ether were performed on the precipitated peptides. The following purification steps were performed through flash chromatography (SP1; Biotage, Uppsala, Sweden) on a Biotage SNAP Ultra C18 12-g cartridge packed with Biotage HP-Sphere C18 25 μm spherical silica. A final step of molecular mass confirmation was performed by mass spectroscopy on a MALDI-TOF/TOF mass spectrometer (ABI 4800; AB Sciex, Framingham, MA, USA). N-terminally FITC-tagged peptide UL44_410-433 corresponding to the HCMV C-terminal 24 residues (410-KEESDSEDSVTFEFVPNTKKQKCG-433) and its phosphorylated variant UL44_410-433_pT (410-KEESDSEDSVTFEFVPNTKKQKCG-433) were synthesized by the Dementia Research Centre, Macquarie University and used for fluorescence polarization assays as described in Ref. [6]. For the phosphorylated variant, phosphorylated threonine amino acid (Fmoc-Thr(PO(OBzl)OH)-OH) was employed for the 427aa site. FITC tagged UL44_410-433 and UL44_410-433_pT was confirmed with Liquid Chromatography Mass Spectrometry (LCMS-8050; Shimadzu, Kyoto, Japan). LC–MS indicated a +80 Da of UL44_410-433_pT indicating the presence of phosphate group.

Fluorescence polarization assays

To quantify the effect of T427 phosphorylation on the binding affinity of UL44 peptides and mIMPα2/hIMPα3,5,7, FP assays were performed using FITC-tagged NLS peptides and bacterially expressed IMPαΔIBBs. Two-fold dilutions of 25 μM IMPα were titrated across 11 wells of a black Fluotrac microplate (Greiner Bio-One, Kremsmünster, Austria) and 100-nM FITC peptide was added to each well. The final volume per well was 200 μL using tris-buffered saline pH 8.0. A control containing no IMPα was also prepared, and this was used for gain adjustment. FP measurements were immediately recorded using a CLARIOstar Plus plate reader (BMG Labtech, Ortenberg, Germany) and the assay was repeated in triplicate for each peptide. Binding curves were plotted using nonlinear regression one site-specific binding in GRAPHPAD PRISM (Version 9.5.1): GraphPad Software, Boston, MA, USA.

Crystallization of mIMPα2 and UL44 NLSs

mIMPα2 and peptides UL44_410-433 or UL44_410-433_pT were co-crystallized for assessment of their molecular binding interfaces. For this, purified mIMPα2ΔIBB and unlabeled peptides were mixed in a 3 : 1 molar ratio and screened using hanging drop vapor diffusion techniques in known conditions (tri-sodium citrate 500–850 mM, 0.1 M HEPES pH 6.5 or 7.0 or 7.5, 0.01 M DTT) at 23 °C. Drop size was 3 μL of 1 : 1 protein/peptide mix : reservoir solution suspended over 300 μL of the screening reservoir condition. High-quality crystals used for structure determination of mIMPα2:UL44_410-433 were formed in tri-sodium citrate 0.65 M, HEPES pH 7.5, DTT 0.01 M. For mIMPα2:UL44_410-433_pT, crystals grew in tri-sodium citrate 0.65 M, HEPES pH 7.0, DTT 0.01 M. Large, thick rod-like crystals were harvested in CryoLoops (Hampton Research, Aliso Viejo, CA, USA), cryo-protected using the screen condition supplemented with 20% glycerol, and flash cooled in liquid nitrogen.

Data collection, processing, and structure determination

Crystals were diffracted on the MX2 beamline at the Australian Synchrotron [31]. Each dataset contained 3600 images from 360° crystal rotation using 0.1° oscillations. Data space group determination and integration were performed using DIALS [32], then scaled and merged in AIMLESS (CCP4 SUITE, [33]). To guarantee a truly free test set of reflections for bias assessment, a previously solved mIMPα2ΔIBB model (PDB: 7RG1) was subjected to Cartesian simulated annealing using PHENIX REFINE and the untouched test set was later used together with the output PDB as the initial start model for molecular substitutions. Iterative cycles of model building in COOT [34] and refinement in PHENIX REFINE [35] yielded a final model, validated using MOLPROBITY (in PHENIX REFINE) and the wwPDB OneDep System. Data processing and refinement statistics are listed in Table 1. Structures were deposited in the PDB and assigned the access codes 8QXW and 8QXX. Structure graphical representations were created in PYMOL (The PyMOL Molecular Graphics System, Version 2.5.2; Schrödinger, LLC, New York, NY, USA), and protein–peptide interface interactions were assessed using PDBE PISA (v1.52 [36]).

Table 1. Data processing and refinement statistics. Structure data collection: overall, innershell, outershell. Model data statistics: overall (outershell). Data collected at wavelength 0.95372.

	UL44_410-433 (8QXW)	UL44_410-433_pT (8QXX)
Resolution limit (Å)	65.75 – 2.00 (2.05 – 2.00)	61.39 – 1.90 (1.94 – 1.90)
Space group	P 21, 21, 21	P 21, 21, 21
Unit cell length (Å)	78.88, 90.09, 96.17	78.90, 89.85, 97.72
Unit cell angle (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Total observations	403 081, 4948, 27 932	243 726, 2496, 13 235
Unique reflections	47 035, 609, 3420	55 424, 589, 3510
Multiplicity	8.6, 8.1, 8.2	4.4, 4.2, 3.8
Completeness (%)	100.0, 99.3, 100.0	99.9, 99.3, 99.8
Mean I/Sigma (I)	9.8, 27.0, 0.9	9.0, 24.3, 1.1
CC (1/2)	0.998, 0.998, 0.637	0.997, 0.997, 0.480
R-pim	0.035, 0.016, 0.636	0.037, 0.015, 0.442
R-meas	0.104, 0.049, 1.845	0.080, 0.033, 0.881
R-merge	0.098, 0.046, 1.730	0.071, 0.028, 0.758
Wilson B Factor	35.14	26.82
R-work	0.19 (0.35)	0.18 (0.30)
R-free	0.22 (0.36)	0.20 (0.32)
Number of non-hydrogen atoms	3510	3777
Macromolecules	3315	3395
Ligands	0	0
Solvent	195	382
Protein residues	433	439
RMS (bonds) (Å)	0.012	0.011
RMS (angles) (°)	1.14	1.07
Ramachandran favored (%)	97.67	98.38
Ramachandran allowed (%)	2.33	1.62
Ramachandran outliers (%)	0.00	0.00
Rotamer outliers (%)	0.00	0.00
Clashscore	5.80	5.68
Average B-factor	56.22	45.20
Macromolecules	56.45	44.77
Solvent	52.23	48.98

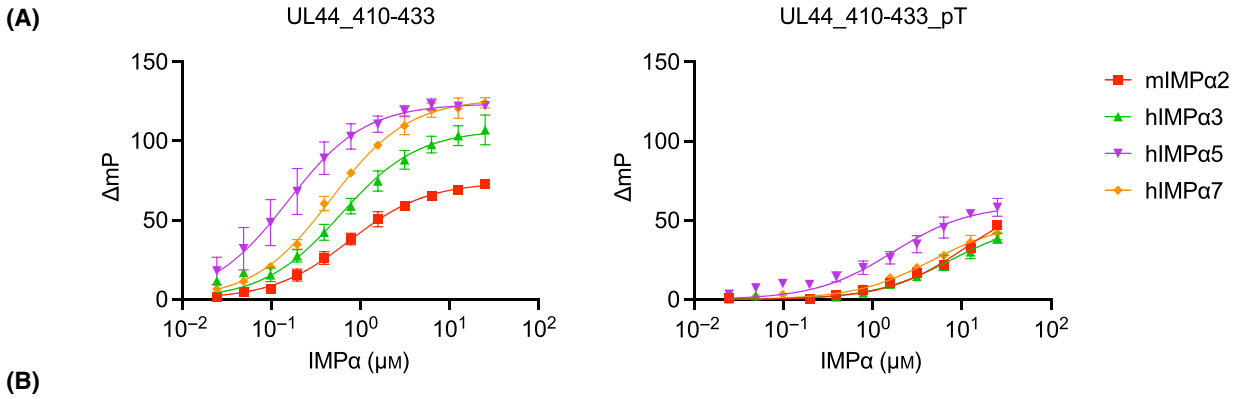
Results

Phosphorylation of HCMV UL44 T427 reduces binding to IMPα isoforms

We have previously shown that the phosphomimic substitution T427D, immediately upstream of HCMV-UL44 DNA polymerase processivity factor cNLS (426-PNTKKQK-431) strongly inhibits nuclear accumulation, and decreases *ori*Lyt-dependent DNA replication [11]. Such reduction in nuclear accumulation was speculated to depend on sequestration by the negative regulator of nuclear import BRAP2, which reduces nuclear import of UL44, but not of the phosphonull substitution derivative T427A [22]. Nonetheless, a recent study revealed that phosphorylation in the vicinity of the nuclear localization signal of human dUTPase abolishes nuclear import, and phosphomimic substitution with a negatively charged glutamic acid residue (S11E) at such position decreased affinity of the dUTPase for IMP α [17]. We, therefore, hypothesized that phosphorylation of T427 could similarly reduce binding to IMPα. To address this possibility, we employed fluorescence polarization (FP) to quantitatively evaluate the impact of T427 phosphorylation on the binding of FITC-labeled UL44 peptides to several IMPαΔIBB isoforms. In such assays, peptide UL44_410-433 bound preferentially to hIMPα5 (K_d ~ 160 nM) as compared to the other IMPα isoforms (K_d range 450–750 nM; see Fig. 3). Strikingly, peptide UL44_410-433_pT interacted with all IMPα isoforms with approximately 10-fold lower binding affinity (K_d range 1670–10 710 nM).

Phosphorylation of HCMV UL44 T427 changes interactions with mIMPα2ΔIBB

Since HCMV UL44 peptide UL44_410-433_pT bound IMPα isoforms with a 10-fold reduced affinity as compared to UL44_410-433, it suggested that the negative



(B)

UL44_410-433	K_d (μM)	B_{max} (mP)
mIMPα2	0.75 ± 0.04	73.98 ± 1.05
hIMPα3	0.60 ± 0.04	107.14 ± 1.71
hIMPα5	0.16 ± 0.01	123.82 ± 2.20
hIMPα7	0.47 ± 0.02	126.66 ± 1.21

UL44_410-433_pT	K_d (μM)	B_{max} (mP)
mIMPα2	10.71 ± 1.05	64.80 ± 2.92
hIMPα3	7.39 ± 0.98	49.48 ± 2.69
hIMPα5	1.67 ± 0.24	59.72 ± 2.41
hIMPα7	4.63 ± 0.46	49.51 ± 1.75

Fig. 3. Quantitative analysis of the effect of T427 phosphorylation on binding to mIMPα2, hIMPα3, hIMPα5, and hIMPα7 isoforms. (A) FITC-labeled peptides (100 nM) were incubated with two-fold serial dilutions of bacterially purified recombinant IMPαΔIBBs (range: 25 μM to 2.4×10^{-2} μM) and subjected to fluorescence polarization (FP) analysis. Data are shown in millipolarized units (mP) as mean ± SEM relative to three independent experiments. Data were analyzed with GRAPHPAD PRISM using the one site-specific binding least squares fit function to calculate the B_{max} and K_d relative to the peptide:IMPαΔIBBs interaction. (B) A summary table relative to data shown in (A), with mean B_{max} and K_d values relative to the interaction of indicated peptides with IMPαΔIBBs, along with the respective SEM.

charge of the phosphate moiety might affect interaction of the NLS basic residues with IMPα binding sites. To address this possibility, we solved the crystal structures of UL44_410-433 and UL44_410-433_pT in complex with mouse IMPα2ΔIBB to 2.0 Å (R_{work} = 19%, R_{free} = 22%) and 1.9 Å (R_{work} = 18%, R_{free} = 20%), respectively (Fig. 4).

Peptide NLS residues 425–433 were elucidated for the UL44_410-433 structure, and 421–433 for the UL44_410-433_pT structure. Our results showed that phosphorylation did not markedly alter the binding mode at the key P0-P6 pockets of the IMPα major binding site. Indeed, in both cases, peptides bound to the IMPα major binding site and displayed major site positions like the prototypical Simian Vacuolating virus 40 (SV40) large tumor antigen (LTA) NLS peptide. At the P2 site, a key residue for the NLS, the K428 makes three hydrogen bonds (with the carbonyl group of G150, the hydroxyl group of T155, and the carbonyl oxygen of D192 of IMPα) as well as a salt bridge interaction with the negatively charged IMPα D192. Notably, in the UL44_410-433 peptide, the carbonyl group of P425, accommodated in the P-1 binding pocket (Fig. 4), makes two hydrogen bonds with IMPα R238 (NH1 and NH2). Conversely, in the UL44_410-433_pT structure, IMPα R238 (NH1 and NH2) establishes electrostatic interactions *via* two hydrogen bonds with phosphorylated threonine (T427) oxygen atoms (OG1, O2P), thereby losing contact with P425. The NLS peptide electron densities and conformational change between the two peptides are highlighted in Fig. 5.

Overall, these results show that there are structural rearrangements of the NLS that occur N terminus to the P2 site lysine that could account for changes in binding affinity observed in our assays.

Discussion

This is one of the first studies to report the crystal structure of a phosphorylated-NLS residue negatively affecting binding to IMPα. Despite cargo phosphorylation being long known to regulate nuclear transport [12], its mechanistic details are still lacking, and at present it is not possible to predict the effect of phosphorylation on nuclear import [9]. This is mainly due

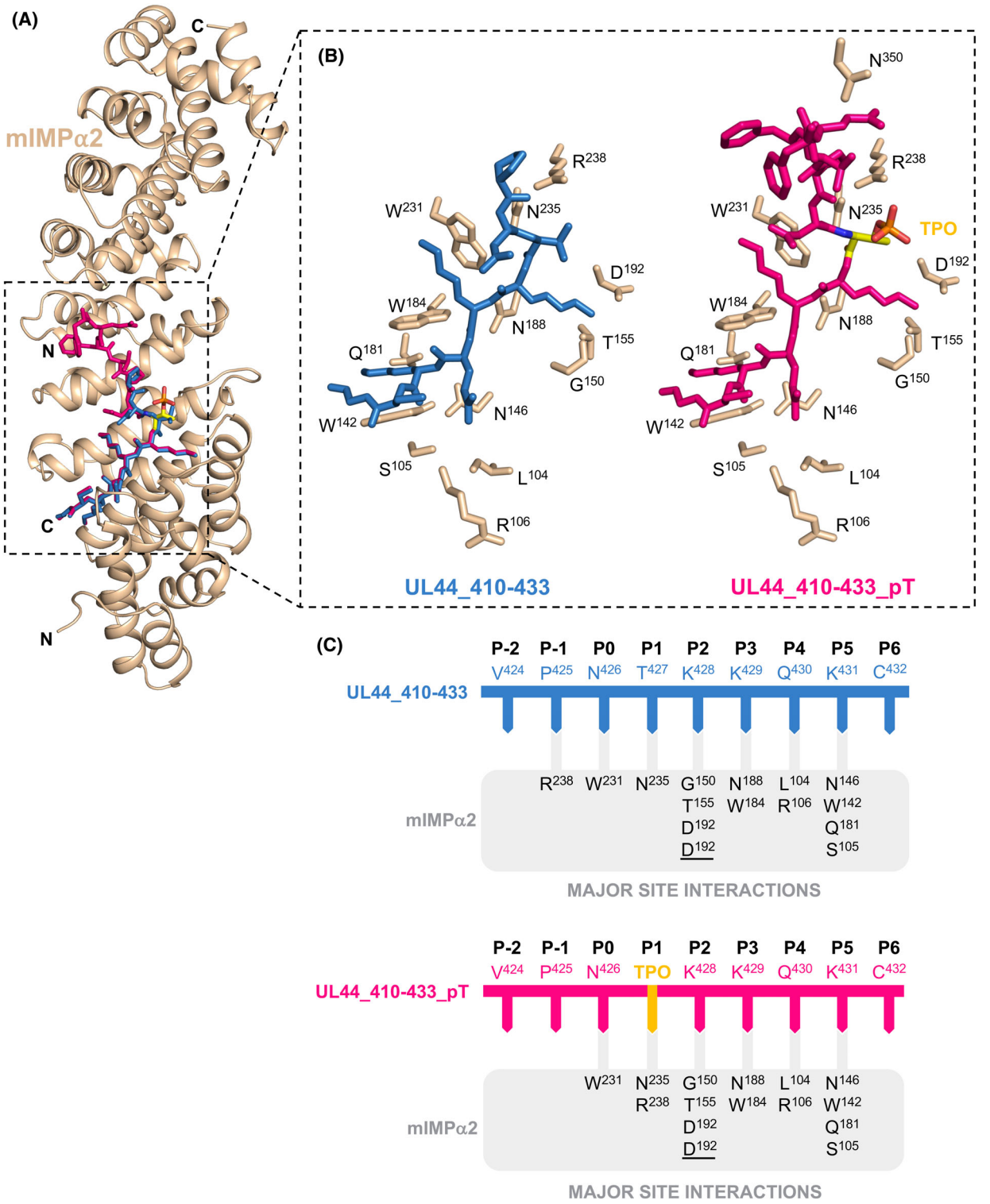


Fig. 4. Crystal structure of the UL44_410-433 and UL44_410-433_pT peptides complexed with mIMPα2ΔIBB. Structural alignment of UL44 peptides with IMPα2ΔIBB. Wheat color indicates mIMPα2ΔIBB, UL44_410-433 is blue, and UL44 UL44_410-433_pT is magenta. Hydrogen bond interactions between either UL44_410-433 (blue sticks) or UL44_410-433_pT (magenta sticks) and IMPα2ΔIBB (wheat, sidechain sticks) interactions are modeled. Below, the major site NLS interactions (P-2 to P6) displaying hydrogen bonds and the salt bridge at D192 (underlined) are indicated.

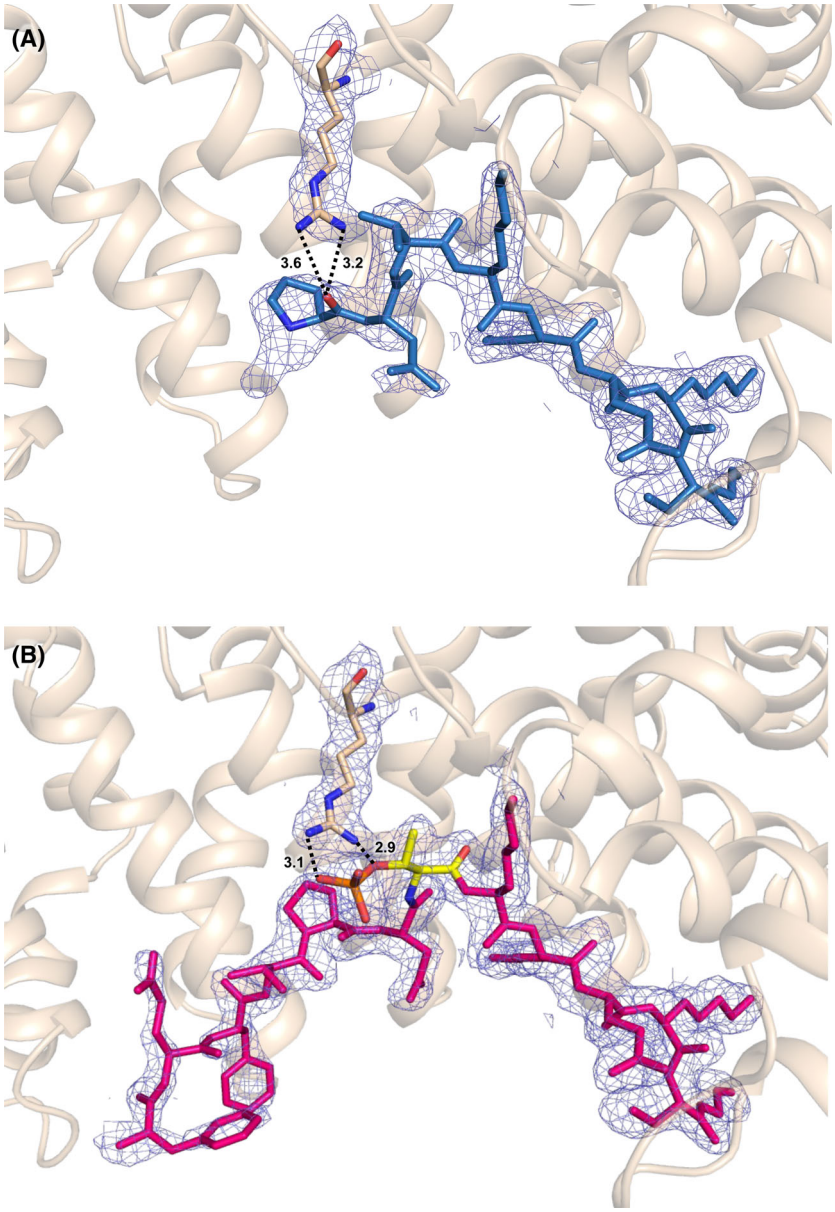


Fig. 5. Structural rearrangement of IMPα R238. (A) A close view of UL44_410-433 peptide (blue) bound to mIMPα2 (wheat, cartoon). mIMPα2 R238 (NH1 and NH2 atoms) form two hydrogen bonds with the P-1 site P425 oxygen. (B) UL44 UL44_410-433_pT (magenta) bound to mIMPα2 (wheat, cartoon). The mIMPα2 R238 (NH1 and NH2 atoms) form two hydrogen bonds with the negatively charged phosphorylated T427 (OG1 and O2P atoms) at the P1 site. Interactions were determined using PDBEPISA.

to the lack of structural comparative analysis of phosphorylated and unphosphorylated cargoes in complex with IMPαs. However, at least for cNLSs recognized by IMPα, some reports are beginning to shed light on this issue. The majority of information available comes from structural studies with cNLS bearing phosphomimic substitutions in complex with IMPα [17], or molecular dynamics simulation using the structure of IMPα with unphosphorylated cargo peptides [15]. To date, the only published study whereby the structure of a phosphorylated residue with IMPα has been reported concerns the complex between EBNA1 cNLS and mIMPα2 [14]. The peptide (379-KRPRSPSS-386) could be observed at both the major and minor IMPα binding sites. However, pS385 could only be unambiguously visualized at the minor binding site, thus providing no information regarding the role of S385 phosphorylation on

functional interaction at the physiological major IMPα binding site [14].

Overall, it appears that phosphorylation of cargoes upstream of their cNLS can either enhance or decrease nuclear import [11,13,14,37]. The mechanism of such regulation is still elusive. In the case of SV40 LTA, the binding has been proposed to rely on conformational changes induced by phosphorylation of S111/112, rather than on the ability of phosphate groups to directly interact with IMPα [11,13,37,38]. Alternatively, phosphorylation of residues that lie in close proximity or within the cNLS itself (i.e., positions from P-2 to P1) decreases nuclear import. This has been extensively shown for several viral and cellular proteins, including SV40 LTA, HCMV UL44, human dUTPase, and yeast Swi6 [2,11,16,18]. In such cases, reduction in NLS activity has been attributed to reduced affinity for IMPα/β, as well as enhanced affinity to a cytoplasmic retention factor.

Our investigation can establish a connection between the position of the phosphorylated residue in monopartite cNLSs and the mechanism underlying the inhibition of nuclear import. Indeed, in the case of SV40 LTA (124-TPPKKKRKV-132), a negative charge at the P-2 residue T124 does not decrease affinity for IMPα but promotes interaction with the cytoplasmic retention factor BRAP-2 [22]. However, for yeast Swi8 (160-SPLKKLKI) and human dUTPase (11-SPSKRARP-18), a negative charge at the P-1 residues decreases binding affinity 5- and 10-fold, respectively [17,18]. We show here that phosphorylation at the P1 residue similarly reduces binding of HCMV UL44 NLS to IMPα by 10-fold (see Fig. 3). This suggests that phosphorylation at major site residues P-1/P1 directly decreases IMPα binding, in contrast to what reported for phosphorylation at position P-2.

Further to this, our comparative structural analysis provides valuable mechanistic insights into the process, revealing that phosphorylation at UL44 T427 changes the hydrogen bonding pattern at the mIMPα2 major binding site. Regarding the human dUTPase NLS, a P-1 site phosphomimic displayed structural rearrangement where new contacts with IMPα D270 resulted in the loss of IMPα interaction with the P0 site proline and a conformational change to the P3 site arginine which resulted in a loss of interaction with IMPα N228 [17]. Phosphorylation of the P1 site in our study did not reveal major conformational changes to the P3 site K429. However, we did see a loss of two hydrogen bonds between the P-1 site P425 and mIMPα2 R238, where the R238 sidechain created two new hydrogen bonds with the phosphorylated threonine. We speculate the loss of interactions at the N-terminal region of the peptide at the major site could weaken the binding in the phosphorylated NLS, similar to what has been proposed for dUTPase.

Hence, the structural differences described here can, to some extent, explain the difference in binding affinity between IMPα and UL44_410-433 and UL44_410-433_pT. Interestingly, it was recently reported that phosphorylation of T88 within the linker of the TDP-48 bipartite NLS reduces IMPα/β binding and nuclear import by affecting the dynamics of NLS conformation. Indeed, MD simulations suggested that CKI-mediated phosphorylation of the NLS linker reduces the spontaneous transition of the NLS backbone to more kinetically favorable states, decreasing the probability of interaction with IMPα [39]. Although the NLS of UL44, Swi6, and dUTPase are monopartite and not bipartite, an effect of N-terminal phosphorylation on the conformation of their NLS cannot be excluded. Finally, it is worth mentioning that SUMOylation of UL44 residue K410, closely located to the phosphorylation-regulated NLS domain, controls UL44 subnuclear distribution and viral replication [28,40,41], and it is, therefore, possible that UL44 SUMOylation and phosphorylation are interconnected, as shown for heat shock transcription factor 4 isoform b, adding another layer of complexity to the regulation of UL44 function and nuclear transport [42].

Acknowledgements

This work was partially supported by the University of Padua (BIRD grant ALVI_SID19_01 and DOR grants to GA). This research was undertaken, in part, using the MX2 beamline at the Australian Synchrotron, part of ANSTO, and made use of the Australian Cancer Research Foundation (ACRF) detector. EMC acknowledges a joint PhD studentship co-funded by Diamond Light Source and Charles Sturt University.

Author contributions

EMC performed protein expression, purification, crystallization, structure determination, structure analysis, fluorescence polarization assays, and manuscript preparation. OM synthesized unlabeled peptides. D Ariawan synthesized labeled peptides. D Aragão assisted with structure analysis and manuscript revision. GC assisted with manuscript revision. EDI assisted with manuscript revision. JKF performed supervision, structure analysis, and manuscript revision. GA performed conceptualization, supervision, manuscript preparation, and revision.

Peer review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/1873-3468.14797.

Data accessibility

Data generated in this study will be shared after reasonable request to the corresponding author. The crystal structures were deposited in the Protein Data Bank under the accession codes 8QXW (https://doi.org/10.2210/pdb8QXW/pdb) and 8QXX (https://doi.org/10.2210/pdb8QXX/pdb). Raw data for these structures are available for download through Zenodo; https://doi.org/10.5281/zenodo.10277323 and https://doi.org/10.5281/zenodo.10276458.

References

1 Alvisi G, Jans D, Camozzi D, Avanzi S, Loregian A, Ripalti A and Palù G (2013) Regulated transport into the nucleus of herpesviridae DNA replication core proteins. *Viruses* **5**, 2210–2234.

2 Conti E, Uy M, Leighton L, Blobel G and Kuriyan J (1998) Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell* **94**, 193–204.

3 Fontes MR, Teh T and Kobe B (2000) Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha. *J Mol Biol* **297**, 1183–1194.

4 Fontes MR, Teh T, Jans D, Brinkworth RI and Kobe B (2003) Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha. *J Biol Chem* **278**, 27981–27987.

5 Christie M, Chang CW, Róna G, Smith KM, Stewart AG, Takeda AAS, Fontes MRM, Stewart M, Vértessy BG, Forwood JK *et al.* (2016) Structural biology and regulation of protein import into the nucleus. *J Mol Biol* **428** (10 Pt A), 2060–2090.

6 Alvisi G, Manaresi E, Cross EM, Hoad M, Akbari N, Pavan S, Ariawan D, Bua G, Petersen GF, Forwood J *et al.* (2023) Importin alpha/beta-dependent nuclear transport of human parvovirus B19 nonstructural protein 1 is essential for viral replication. *Antiviral Res* **213**, 105588.

7 Smith KM, di Antonio V, Bellucci L, Thomas DR, Caporuscio F, Ciccarese F, Ghassabian H, Wagstaff KM, Forwood JK, Jans DA *et al.* (2018) Contribution of the residue at position 4 within classical nuclear localization signals to modulating interaction with importins and nuclear targeting. *Biochim Biophys Acta* **1865**, 1114–1129.

8 Poon IKH and Jans DA (2005) Regulation of nuclear transport: central role in development and transformation? *Traffic* **6**, 173–186.

9 Nardozzi JD, Lott K and Cingolani G (2010) Phosphorylation meets nuclear import: a review. *Cell Commun Signal* **8**, 32.

10 Alvisi G, Jans DA, Guo J, Pinna LA and Ripalti A (2005) A protein kinase CK2 site flanking the nuclear targeting signal enhances nuclear transport of human cytomegalovirus ppUL44. *Traffic* **6**, 1002–1013.

11 Alvisi G, Marin O, Pari G, Mancini M, Avanzi S, Loregian A, Jans DA and Ripalti A (2011) Multiple phosphorylation sites at the C-terminus regulate nuclear import of HCMV DNA polymerase processivity factor ppUL44. *Virology* **417**, 259–267.

12 Rihs HP and Peters R (1989) Nuclear transport kinetics depend on phosphorylation-site-containing sequences flanking the karyophilic signal of the simian virus 40 T-antigen. *EMBO J* **8**, 1479–1484.

13 Kitamura R, Sekimoto T, Ito S, Harada S, Yamagata H, Masai H, Yoneda Y and Yanagi K (2006) Nuclear import of Epstein-Barr virus nuclear antigen 1 mediated by NPI-1 (importin alpha5) is up- and down-regulated by phosphorylation of the nuclear localization signal for which Lys379 and Arg380 are essential. *J Virol* **80**, 1979–1991.

14 Nakada R, Hirano H and Matsuura Y (2017) Structural basis for the regulation of nuclear import of Epstein-Barr virus nuclear antigen 1 (EBNA1) by phosphorylation of the nuclear localization signal. *Biochem Biophys Res Commun* **484**, 113–117.

15 Koyama M and Matsuura Y (2017) Crystal structure of importin-alpha3 bound to the nuclear localization signal of Ran-binding protein 3. *Biochem Biophys Res Commun* **491**, 609–613.

16 Jans DA, Ackermann MJ, Bischoff JR, Beach DH and Peters R (1991) p34cdc2-mediated phosphorylation at T124 inhibits nuclear import of SV-40 T antigen proteins. *J Cell Biol* **115**, 1203–1212.

17 Rona G, Marfori M, Borsos M, Scheer I, Takács E, Tóth J, Babos F, Magyar A, Erdei A, Bozóky Z *et al.* (2013) Phosphorylation adjacent to the nuclear localization signal of human dUTPase abolishes nuclear import: structural and mechanistic insights. *Acta Crystallogr D Biol Crystallogr* **69** (Pt 12), 2495–2505.

18 Harreman MT, Kline TM, Milford HG, Harben MB, Hodel AE and Corbett AH (2004) Regulation of nuclear import by phosphorylation adjacent to nuclear localization signals. *J Biol Chem* **279**, 20613–20621.

19 Zhang F, White RL and Neufeld KL (2000) Phosphorylation near nuclear localization signal regulates nuclear import of adenomatous polyposis coli protein. *Proc Natl Acad Sci USA* **97**, 12577–12582.

20 Zheng W, Li J, Wang S, Cao S, Jiang J, Chen C, Ding C, Qin C, Ye X, Gao GF *et al.* (2015) Phosphorylation controls the nuclear-cytoplasmic shuttling of influenza A virus nucleoprotein. *J Virol* **89**, 5822–5834.

21 Fulcher AJ, Dias MM and Jans DA (2010) Binding of p110 retinoblastoma protein inhibits nuclear import of simian virus SV40 large tumor antigen. *J Biol Chem* **285**, 17744–17753.
22 Fulcher AJ, Roth DM, Fatima S, Alvisi G and Jans DA (2010) 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* **24**, 1454–1466.
23 Sekimoto T, Fukumoto M and Yoneda Y (2004) 14-3-3 suppresses the nuclear localization of threonine 157-phosphorylated p27(Kip1). *EMBO J* **23**, 1934–1942.
24 Kann M, Sodeik B, Vlachou A, Gerlich WH and Helenius A (1999) Phosphorylation-dependent binding of hepatitis B virus core particles to the nuclear pore complex. *J Cell Biol* **145**, 45–55.
25 Alvisi G, Ripalti A, Ngankeu A, Giannandrea M, Caraffi SG, Dias MM and Jans DA (2006) Human cytomegalovirus DNA polymerase catalytic subunit pUL54 possesses independently acting nuclear localization and ppUL44 binding motifs. *Traffic* **7**, 1322–1332.
26 Sinigalia E, Alvisi G, Mercorelli B, Coen DM, Pari GS, Jans DA, Ripalti A, Palù G and Loregian A (2008) Role of homodimerization of human cytomegalovirus DNA polymerase accessory protein UL44 in origin-dependent DNA replication in cells. *J Virol* **82**, 12574–12579.
27 Alvisi G, Roth DM, Camozzi D, Pari GS, Loregian A, Ripalti A and Jans DA (2009) The flexible loop of the human cytomegalovirus DNA polymerase processivity factor ppUL44 is required for efficient DNA binding and replication in cells. *J Virol* **83**, 9567–9576.
28 Sinigalia E, Alvisi G, Segré CV, Mercorelli B, Muratore G, Winkler M, Hsiao HH, Urlaub H, Ripalti A, Chiocca S *et al.* (2012) The human cytomegalovirus DNA polymerase processivity factor UL44 is modified by SUMO in a DNA-dependent manner. *PLoS One* **7**, e49630.
29 Silva LA, Strang BL, Lin EW, Kamil JP and Coen DM (2011) Sites and roles of phosphorylation of the human cytomegalovirus DNA polymerase subunit UL44. *Virology* **417**, 268–280.
30 Fields GB and Noble RL (1990) Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int J Pept Protein Res* **35**, 161–214.
31 Aragao D, Aishima J, Cherukuvada H, Clarken R, Clift M, Cowieson NP, Ericsson DJ, Gee CL, Macedo S, Mudie N *et al.* (2018) MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron. *J Synchrotron Radiat* **25** (Pt 3), 885–891.
32 Winter G, Waterman DG, Parkhurst JM, Brewster AS, Gildea RJ, Gerstel M, Fuentes-Montero L, Vollmar M, Michels-Clark T, Young ID *et al.* (2018) DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr D Struct Biol* **74** (Pt 2), 85–97.
33 Evans PR and Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69** (Pt 7), 1204–1214.
34 Emsley P and Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60** (Pt 12 Pt 1), 2126–2132.
35 Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ *et al.* (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75** (Pt 10), 861–877.
36 Krissinel E and Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
37 Hübner S, Xiao CY and Jans DA (1997) 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* **272**, 17191–17195.
38 Xiao CY, Jans P and Jans DA (1998) Negative charge at the protein kinase CK2 site enhances recognition of the SV40 large T-antigen NLS by importin: effect of conformation. *FEBS Lett* **440**, 297–301.
39 Doll SG, Meshkin H, Bryer AJ, Li F, Ko YH, Lokareddy RK, Gillilan RE, Gupta K, Perilla JR and Cingolani G (2022) Recognition of the TDP-43 nuclear localization signal by importin alpha1/beta. *Cell Rep* **39**, 111007.
40 Chen J, Li G, He H, Li X, Niu W, Cao D and Shen A (2021) Sumoylation of the carboxy-terminal of human cytomegalovirus DNA polymerase processivity factor UL44 attenuates viral DNA replication. *Front Microbiol* **12**, 652719.
41 Lai S, Xu M, Wang Y, Li R, Xia C, Xia S and Chen J (2021) Site-specific SUMOylation of viral polymerase processivity factor: a way of localizing to ND10 subnuclear domains for restricted and self-controlled reproduction of herpesvirus. *Virulence* **12**, 2883–2901.
42 Hietakangas V, Anckar J, Blomster HA, Fujimoto M, Palvimo JJ, Nakai A and Sistonen L (2006) PDSM, a motif for phosphorylation-dependent SUMO modification. *Proc Natl Acad Sci USA* **103**, 45–50.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Importin α truncation derivatives used in this study.