SI Appendix: Structural basis for nuclear import of splicing factors by human Transportin 3

Goedele N. Maertens[*1], Nicola Cook[*2], Weifeng Wang[3], Stephen Hare[1], Saumya Shree Gupta[1], Ilker Öztop[3], KyeongEun Lee[4], Valerie E. Pye[2], Ophélie Cosnefroy[2], Ambrosius P. Snijders[2], Vineet N. KewalRamani[4], Ariberto Fassati[5], Alan Engelman[3] and Peter Cherepanov[†1,2]

[1]Division of Infectious Diseases, Imperial College London, St.-Mary's Campus, Norfolk Place, London, W2 1PG, UK, [2]Cancer Research UK London Research Institute, Clare Hall Laboratories, Blanche Lane, Potters Bar, EN6 3LD, UK, [3]Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, Boston, MA, 02215, USA, [4]HIV Drug Resistance Program, National Cancer Institute, Frederick, MD, 21702, USA, [5]The Wohl Virion Centre and MRC Centre for Medical & Molecular Virology, Division of Infection & Immunity, University College London, 90 Gower Street, London WC1E 6BT, UK

[*]These authors contributed equally to this work.

[†]To whom correspondence should be addressed (peter.cherepanov@cancer.org.uk).

Contents

Supplementary Materials and Methods

Recombinant proteins. Full-length human Tnpo3 was expressed as a His_6SUMO fusion using pETSUMO vector (Life Technologies). *E. coli* Rosetta2 (DE3) pLacI cells (Novagen) transformed with pETSUMO-Tnpo3 were grown at 30°C in Terrific Broth medium supplemented with 50 μg/ml kanamycin to an A_{600} of 2.0-2.5, after which the temperature was reduced to 25°C, and protein expression was induced with 0.01% (w/v) IPTG; 4-5 h post-induction, bacteria were harvested by centrifugation and stored at -80°C. Thawed bacterial pellets were disrupted by sonication in ice-cold core buffer (500 mM NaCl, 25 mM Tris-HCl, pH 7.4) supplemented with 1 mM PMSF and 0.25 mg/ml lysozyme, and insoluble debris were removed by centrifugation at 27,000 *g* for 30 min at 4°C. His_6-tagged protein was captured on NiNTA resin (Qiagen) in the presence of 20 mM imidazole, and, following washing in core buffer supplemented with 20 mM imidazole, was eluted in core buffer containing 200 mM imidazole. His_6SUMO tag was removed by overnight digestion with His_6-tagged catalytic core domain of *Saccharomyces cerevisiae* Ulp1 (1) at 4°C, and untagged Tnpo3 was purified by anion exchange chromatography. Hereto, the protein was diluted 2.5-fold in 25 mM Tris-HCl, pH 7.4 and absorbed onto a 5-ml HiTrap Q Sepharose column (GE Healthcare). Tnpo3 was eluted with a linear 0.2-1.0 M NaCl gradient in 25 mM Tris-HCl, pH 7.4 and further purified by gel filtration through a HiLoad 16/60 Superdex 200 column operated in 150 mM NaCl, 25 mM Tris-HCl, pH 7.4. The procedure typically yielded 10-20 mg of pure Tnpo3 from a 6-L shake flask culture. The protein, concentrated to 10-15 mg/ml and supplemented with 10% glycerol and 5 mM DTT, was snap frozen in liquid nitrogen and stored at -80°C. Point mutations were introduced into pET-SUMO-Tnpo3 using QuikChange procedure with Pfu Ultra II DNA polymerase (Stratagene). All Tnpo3 mutants were expressed and purified following the same strategy as for the WT protein. Selenium-labelled Tnpo3 was produced in minimal medium (Molecular Dimensions) supplemented with selenomethionine and a cocktail of amino acids required for metabolic inhibition (2) and purified as described above.

Plasmid used for expression of non-tagged human CLK1 was made by ligating a DNA fragment spanning its complete coding region between NcoI and XhoI sites of

pCDF-Duet1 (Novagen) giving pCDF-CLK1. Phosphorylated human His_6SUMO-ASF/SF2 and His_6SUMO-MBP-ASF/SF2 constructs were produced using pETSUMO vector in BL21(DE3) Codon-Plus *E. coli* cells (Stratagene) co-transformed with pCDF-CLK1. Bacteria containing ASF/SF2 and CLK1 expressing plasmids were grown in LB medium supplemented with 50 μg/ml kanamycin and 100 μg/ml spectinomycin to an A_{600} of 0.9-1.0 and induced with 0.01% IPTG at 30 ^{o}C for 3-4 h. Cells were disrupted by sonication in core buffer supplemented with 1 mM PMSF and 10 mM NaF, and the lysates were cleared by centrifugation at 27,000 *g* for 30 min at 4°C. His_6SUMO - tagged proteins were captured on NiNTA resin, extensively washed and eluted with core buffer containing 20 mM and 200 mM imidazole, respectively. The proteins were further purified by anion exchange chromatography on a 5-ml HiTrap Q sepharose column (GE Healthcare) and gel filtration through a HiLoad 16/60 Superdex 200 column in 150 mM NaCl, 25 mM Tris-HCl, pH 7.4. Peak fractions of phosphorylated His_6SUMO-ASF/SF2 proteins (0.5-1.5 mg/ml) were supplemented with 5 mM DTT and 10% glycerol and stored at -80^{o}C. For use in *in vitro* nuclear import assays, MBP-ASF/SF2 was released from His_6SUMO tag by digestion with His_6-tagged Ulp1. The tag and the protease were depleted by absorption to a HisTrap cartridge (GE Healthcare) prior to anion exchange and gel filtration chromatography.

The plasmid for expression of GST-CPSF6RS was obtained by subcloning a DNA fragment encoding residues 441-588 of human CPSF6 between BamHI and EcoRI sites of pGEX6P3 (GE Healthcare). A DNA fragment spanning the complete open reading frame of human SRPK1 was ligated between NdeI and BamHI sites of pCDF-Duet1 in frame with the sequence encoding the His_6 tag to give pCDF-His_6SRPK1. GST-CPSF6RS was produced in BL21(DE3) Codon-Plus *E. coli* cells in the presence or absence of pCDF-His_6SRPK1 and purified on glutathione sepharose followed by size exclusion chromatography on a HiLoad 16/60 Superdex-200 column in 200 mM NaCl, Tris-HCl, pH 7.4. Phosphorylation of the recombinant protein was confirmed by SDS PAGE in the presence of copolymerized PhosTag-Mn(II) complex (Wako Chemicals) (Fig. S10A).

To make constructs for bacterial expression of WT human Ran and Ran(Q69L), the DNA fragments encoding these proteins were ligated between XmaI and XhoI sites of

pCPH6P (3). PC2 *E. coli* strain (BL21(DE3), Δ*endA*::Tc^R, $T1^R$, pLysS) (3) transformed with pCPH6P-Ran or pCPH6P-Ran(Q69L) were grown at 30°C in LB medium supplemented with 120 μg/ml ampicillin to an A_{600nm} of 0.9-1.0, after which the temperature was reduced to 25°C and protein expression was induced with 0.01% IPTG for 4 h. Bacterial pellets re-suspended in ice-cold core buffer supplemented with 1 mM PMSF were disrupted by sonication, and the lysates were cleared by centrifugation at 27,000 *g* for 30 min at 4°C. Supernatants were incubated with NiNTA resin in the presence of 20 mM imidazole; after extensive washing in core buffer containing 20 mM imidazole His_6-tagged proteins were eluted in core buffer supplemented with 200 mM imidazole. When required, His_6 tags were removed by overnight digestion with human rhinovirus 3C protease at 4 ^oC. The proteins were further purified by gel filtration over a HiLoad 16/60 Superdex 200 column in 150 mM NaCl, 25 mM Tris-HCl, pH 7.4 and supplemented with 5 mM DTT. To charge Ran and Ran(Q69L) with GTP, purified proteins were incubated on ice for 30 min in the presence of 10 mM EDTA and 2 mM GTP, after which 25 mM $MgCl_2$ was added. RanGTP and Ran(Q69L)GTP complexes were supplemented with 10% glycerol, frozen in liquid nitrogen and stored at -80°C. Human nuclear transport factor 2 (NTF2) was produced in *E. coli* and purified as previously described (4).

Assembly of protein-protein complexes. To make the Tnpo3-Ran(Q69L)GTP complex, Tnpo3 and His_6-Ran(Q69L)GTP were mixed at a 1:4 molar ratio and incubated for 1 h at 18°C in binding buffer (BB: 150 mM NaCl, 25 mM Tris-HCl, pH 7.4). The complex was then bound to a 1-ml HisTrap column (GE Healthcare) and eluted in 150 mM NaCl, 200 mM imidazole, 25 mM Tris-HCl, pH 7.4. The tag was removed by incubation with His_6-tagged human rhinovirus 3C protease. The protein was diluted 5-fold in BB and filtered through a HisTrap column. The protein was further purified by size exclusion chromatography on a HiLoad 16/60 Superdex 200 column operated in 150 mM NaCl, 25 mM Tris-HCl, pH 7.4.

Tnpo3-ASF/SF2 complexes were made by mixing Tnpo3 with 5-10 fold excess of a His_6SUMO-ASF/SF2 construct in 150 mM NaCl, 25 mM Tris-HCl, pH 7.4. The complexes were treated with His_6-tagged Ulp1 protease and injected into a HisTrap

column in the presence of 20 mM imidazole to deplete the protease and His_6SUMO. The complexes were further purified by gel filtration through a HiLoad 16/60 Superdex 200 column in 150 mM NaCl, 25 mM Tris-HCl, pH 7.4.

Crystallography

Crystals were grown at 18^{o}C by vapour diffusion in hanging drops by mixing 1 μl protein (8-12 mg/ml in 150 mM NaCl, 2 mM DTT, 25 mM Tris-HCl, pH 7.4) and 1 μl reservoir solution, which contained 18-23% (w/v) polyethylene glycol (PEG) 3,350, 0.2 M NaBr, 2 mM EDTA, 10 mM DTT and 0.1 M Bis-Tris propane-HCl, pH 8.2 (for unliganded Tnpo3(C511A)); 13-15% (w/v) PEG-3,350, 10% (w/v) benzamidine hydrochloride and 0.2 M sodium acetate, pH 6.0 (for the Tnpo3^{SeMet}-Ran(Q69L)GTP complex); and 16-19% (w/v) PEG-3,350, 0.1 M KSCN and Bis-Tris propane-HCl, pH 6.0 (for the Tnpo3-ASF/SF2(106-230) complex). Multiple rounds of streak seeding were used to improve quality of Tnpo3(C511A) and Tnpo3-ASF/SF2(12-230) crystals. Crystals of unliganded Tnpo3 and Tnpo3(C511A) were unstable, likely due to oxidation, and tended to soften up if left in mother liquor for more than 7-10 days. Crystals of Tnpo3-Ran(Q69L)GTP complex grew as clusters of very thin plates and could not be improved by seeding. Several thicker plates, nucleated at edges of crystallization drops, could be harvested following a large number of trials; setting up elongated hanging drops helped to recover larger crystals in this way. Crystals, cryoprotected with 20-25% glycerol, were flash-cooled in liquid nitrogen.

Data collection and structure determination. X-ray diffraction data for structure refinement were collected at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) beam lines ID14-4 and ID23-1. The data, integrated using XDS (5), were analysed with Pointless and scaled in Scala (6). Initially, the structure of non-liganded Tnpo3 was solved by molecular replacement using a 3.2-Å dataset collected from crystals of WT protein using Phaser (7) and fragments of Imp13 structures from PDB IDs 2X19 (8), and 2XWU (9) as search models. The model was initially re-built with Buccaneer (10) enforcing NCS constraints during refinement cycles in Refmac (11). The structure, containing four Tnpo3 chains per asymmetric unit, revealed a close

juxtaposition of Cys511 residues between pairs of neighbouring protein molecules. C511A mutation allowed to improve crystal quality, and the isomorphous structure of Tnpo3(C511A) could be refined to 2.95-Å resolution in Phenix (12) using four-fold torsional non-crystallographic symmetry (NCS) restrains.

The structures of Tnpo3-ASF/SF2(106-230) and Tnpo3-Ran(Q69L)GTP complexes were solved by molecular replacement using Phaser and fragments of the refined Tnpo3(C511A) structure, RRM domain from PDB ID 3BEG (13) and Ran from PDB ID 2X19 as search models. The structures were refined in Phenix with torsional NCS restraints and manual building in Coot (14). All phosphate groups present in the refined model of Tnpo3-ASF/SF2(106-230) complex were confirmed in a phased anomalous map calculated using X-ray diffraction data collected at a wavelength of 2.38 Å and model-based phases (Fig. S8B). The low energy X-ray diffraction data were acquired on beam line I03 at Diamond Light Source, Oxfordshire, UK; a redundant dataset with anomalous multiplicity of 6.3 was obtained by merging data from 5 crystals. Torsional external restrains generated using structure of the Tnpo3-ASF/SF2(106-230) complex (refined to 2.57 Å resolution) were used in refinement of Tnpo3-Ran(Q69L)GTP complex; translation, liberation and screw refinement was used at the final stage of refinement of this structure. Although anomalous signal from the selenomethionine derivatized crystals was insufficiently strong for phasing, it allowed to confirm positions of Tnpo3 methionine side chains in the refined structure of Tnpo3-Ran(Q69L)GTP complex (Fig. 2C). Geometry of the final structures was assessed using Molprobity (15) (Table S2).

His_6 and GST tag pull down assays. Ten μg His_6SUMO, His_6SUMO-ASF/SF2, His_6-Ran (loaded with GTP or GDP) was allowed to bind to 8 μg Tnpo3 in 700 μl Ni pull-down buffer (Ni-PDB, 200 mM NaCl, 15 mM imidazole, 2 mM $MgCl_2$, 0.5% CHAPS, 50 mM Tris-HCl, pH 7.4), containing complete EDTA-free protease inhibitors (Roche), in the presence of 20 μl NiNTA resin and 10 μg of BSA as an internal specificity control. Following 2 h of end-over-end rocking at 4°C, the beads were washed extensively in ice-cold Ni-PDB. Bound proteins were eluted by boiling the beads in 20 μl 2x Laemmli

buffer supplemented with 5 mM EDTA; 10 μl was separated on SDS PAGE gels, and the proteins were detected by staining with Coomassie Brilliant Blue. His_6 tag HIV-1 integrase pull down assays with Tnpo3 were carried out as previously described (16).

For GST pull-down assays, GST or GST-CPSF6RS was pre-bound to glutathione sepharose (1 μg of protein per μl resin). Fifteen μl loaded resin was incubated with 12 μg Tnpo3 and 10 μg BSA in GST pull down buffer (GST-PDB, 200 mM NaCl, 0.5% CHAPS, 2 mM DTT, 20 mM Tris-HCl, pH 7.4). Following 2 h of rocking at 4^{o}C, the beads were washed with three changes of 1 ml GST-PDB. The bound proteins, eluted with 2x Laemmli buffer, were separated on SDS PAGE gels and detected by staining with Coomassie.

Eukaryotic expression constructs and tissue culture. The vector pQFlag-puroR, containing a puromycin resistance cassette with an internal ribosome entry site (IRES) (17), was modified for expression of Flag-Strep tag fusion proteins by ligating annealed oligonucleotides 5'-CCGGGTGGAGCCATCCGCAGTTCGAGAAGGGC GGAA and 5'-CCGGTTCCGCCCTTCTCGAACTGCGGATGGCTCCAC into its AgeI site to obtain pQFlag-Strep2-puro^{R}. A DNA fragment encoding full-length Tnpo3 was ligated between AgeI and XhoI sites of this vector to obtain pQFlag-Strep2-Tnpo3-puro^{R}. To render the Tnpo3 mRNA insensitive to RNAi, an shRNA resistant version of Tnpo3 was made by splicing PCR, to introduce 9 silent point mutations, changing the nucleotide sequence encoding Tnpo3 residues 165-171 to AGA AGA ACG GAG ATC ATT GAG (underlined nucleotides indicate mutated positions). The resulting PCR fragment was cloned between AgeI and XhoI sites of pQFlag-hygro^{R} (made by swapping the IRES-puromycin resistance cassette in pQFlag-puroR for IRES-hygromycin B resistance cassette) giving pQFlag-Tnpo3shR-hygro^{R}. Further point mutations were introduced into this plasmid using QuikChange protocols (Stratagene).

The human embryonic kidney 293T cell line and its derivatives were cultured in Dulbecco's modified Eagles medium supplemented with 100 IU/ml penicillin, 100 μg/ml streptomycin, 10% fetal calf serum in a humidified incubator at 37°C, 5% CO_2. Retroviral transductions of 293T cells were done as described previously (17, 18). To establish the 293T-Tnpo3KD cell line stably depleted for endogenous Tnpo3 expression,

293T cells infected with a pLKO.1 derived-lentiviral vector expressing an shRNA targeting the sequence 5′-GGCGCACAGAAATTATAGAA within the coding region of Tnpo3 mRNA (MISSION, Sigma-Aldrich) were selected with 1 μg/ml puromycin. The control cell line was made using the MISSION Non-Target shRNA control vector (Sigma-Aldrich). 293T-derived cell line stably expressing Flag-Strep2-Tnpo3 and back-complemented derivatives of 293T-Tnpo3KD were made by retroviral transduction using pQFlag-Strep2-Tnpo3-puro^R and pQFlag-Tnpo3shR-hygro^R (or its mutant variants) transfer vectors, respectively. Selective pressure was maintained on the transduced cells by supplementing growth medium with 0.5 mg/ml hygromycin B and/or 0.5-1.0 μg/ml puromycin.

Immunoprecipitation and cellular fractionation. For large-scale purification of Flag-Strep2-tagged Tnpo3, three 15-cm dishes of the stably transduced 293T cell lines were grown to ~80% confluence. The parental 293T cell line was used as a negative control and treated identically. Cells, harvested by trypsinization, were washed extensively in ice-cold PBS. Proteins were extracted in 5 pellet volumes of IP buffer (150 mM NaCl, 10% glycerol, 8 mM CHAPS, 2 mM $MgCl_2$, 2 mM DTT, 10 mM Tris-HCl, pH 7.4) supplemented with 1 mM PMSF, protein phosphatase inhibitor cocktail (Sigma-Aldrich), and complete EDTA-free protease inhibitors (Roche). Extracts were cleared by centrifugation at 20,000 *g* for 30 min at 4°C and pre-cleared with protein G agarose beads (Roche). The extracts were then allowed to bind to 30 μl anti-Flag M2 agarose (Sigma-Aldrich) by end-over-end rocking for 3 h. Following extensive washes, bound proteins were eluted with 0.075 μg/ml Flag peptide in IP buffer. Eluted proteins were separated by SDS PAGE and, following in-gel trypsin digestion, were analysed by tandem mass spectrometry.

Tnpo3 knockdown and back-complemented 293T cell lines selected with puromycin and hygromycin B were plated out in 10-cm dishes and grown to 80% confluence before harvesting for small-scale immunoprecipitations. The cell lysates were treated and immunoprecipitated as described above. Two % of the extract was loaded as input and 10% of the eluted Flag-Tnpo3 and bound proteins were separated by SDS PAGE and detected by Western blotting with mouse horse radish peroxidase (HRP)

conjugated anti-Flag (M2, Sigma-Aldrich), rabbit anti-CPSF6 (ab99347, Abcam), rabbit anti-ASF/SF2 (ab38017, Abcam) or rabbit anti-TNPO3 (ab109386, Abcam) antibodies. The blots were developed using secondary antibodies conjugated to horseradish peroxidase (GE Healthcare) and Clarity ECL reagents (Bio-Rad Laboratories).

For cellular fractionation, the Tnpo3 knockdown and back-complemented cell lines grown in a 10-cm dish to ~ 80% confluence, were harvested by trypsinization and washed with ice-cold PBS. All further steps were performed on ice or at 4°C. Some minor modifications were introduced into the cellular fractionation protocol described by Lamond laboratory (http://www.lamondlab.com). The cells were washed twice with ice-cold hypotonic lysis buffer (10 mM Hepes pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT and Complete EDTA free protease inhibitor cocktail (Roche)). The cells were then re-suspended in 5 pellet volumes of hypotonic lysis buffer and kept on ice for 5 min before breaking the cells open with a pre-chilled 1 ml Dounce homogenizer (20 strokes with a tight pestle). The lysed cells were centrifuged at 228 *g* for 5 min at 4°C, and the supernatant was retained as cytoplasmic fraction. The nuclei were then resuspended in 1 ml of S1 (0.2 M sucrose, 10 mM $MgCl_2$, Complete EDTA free), layered onto 1 ml of S2 (0.35 M sucrose, 0.5 mM $MgCl_2$, Complete EDTA free) and centrifuged for 10 min at 2800 *g* at 4°C. The clean, pelleted nuclei were re-suspended in 1x RIPA buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1% NP-40, 0.5% Na-deoxycholate, Complete EDTA free) and briefly sonicated.

In vitro nuclear import assays. MBP-ASF/SF2 was fluorescently labelled by conjugation with Alexa Fluor 633 (Life Technologies) according to manufacturer's instructions. Nuclear import assays were carried out as described (19) with minor modifications. Optimization of digitonin concentration was important to avoid passive nuclear accumulation of the protein. Briefly, HeLa cells grown in Nunc Lab-Tek II chamber slides (Life Technologies) were washed once with ice cold transport buffer (20 mM Hepes pH 7.4, 110 mM potassium acetate, 2 mM magnesium acetate, 1 mM EGTA, 2 mM DTT supplemented with protease inhibitors) before treatment with 0.002 % digitonin (Sigma-Aldrich) in transport buffer on ice for 5 min. Permeabilized cells, depleted of soluble cytoplasmic proteins, were incubated in transport buffer

supplemented with 0.2 μM Alexa Fluor 633-labeled MBP-ASF/SF2, 3 μM RanGTP, 0.4 μM NTF2, 20 U/ml creatine phosphokinase, 1 mM GTP, 1 mM ATP, 10 mM sodium creatine phosphate in the presence of absence of 0.5 μM WT or 9Ala Tnpo3, for 30 min at 37^{o}C. The cells, washed in PBS and fixed with 3% paraformaldehyde were visualized by epifluorescence using a Zeiss Imager M2 instrument.

HIV-1 infectivity assays. HIV-1 infectivity assays were carried out using single-cycle reporter virus prepared by co-transfection of 293T cells with pNLENG1-ES-IRES and pCG-VSV-G using calcium phosphate (20). Five μl of virus, corresponding to an approximate multiplicity of infection of 0.3, were used to inoculate Tnpo3 knockdown and back complemented 293T cell lines in 24-well plates. The expression of enhanced green fluorescent protein in the infected cells was measured by flow cytometry 48 h post-infection.

Supplementary Table S1 | TNPO3 binding partners identified by MS/MS

Proteins copurifying with Flag-Tnpo3 identified specifically with at least 2 unique peptides detected in the condition with or without detergent (CHAPS).
Proteins identified in both conditions (highlighted in green) were considered strong candiates for being Tnpo3 binding partners.

(A) Tnpo3 binding partners containing RRM and RS (RS-like) domains

Identified Proteins[1]	Gene Name	# pept (with CHAPS)[2]	# pept (no CHAPS)[3]	RRM[4]	RS[5]
RNA-binding protein 26	RBM26	66	51	2	+
RNA-binding protein 39	RBM39	24	27	3	+
Cleavage and polyadenylation specificity factor subunit 6 (CPSF6, CF-Im-68)	CPSF6	23	11	1	+
Zinc finger protein 638	ZN638	15	20	2	+
Serine/arginine-rich splicing factor 9	SRSF9	15	17	2	+
Serine/arginine-rich splicing factor 15	SRSF15	14	4	1	+
Serine/arginine-rich splicing factor 1 (ASF/SF2)	SRSF1	13	11	2	+
RNA-binding protein 27	RBM27	12	6	1	+
Serine/arginine-rich splicing factor 6	SRSF6	12	11	2	+
Cleavage and polyadenylation specificity factor subunit 7 (CPSF7, CF-Im-59)	CPSF7	11	7	1	+
Serine/arginine-rich splicing factor 2 (SC35)	SRSF2	10	11	1	+
Serine/arginine-rich splicing factor 3	SRSF3	9	8	1	+
Serine/arginine-rich splicing factor 10	SRSF10	9	9	1	+
SR-Related CTD-Associated Factor 8	SCAF8	8	9	1	+
Peptidyl-prolyl cis-trans isomerase-like 4 (Cyclophilin-like protein PPIL4)	PPIL4	8	10	1	+
U2 small nuclear RNA auxiliary factor 2	U2AF2	8	5	3	+
RNA-binding motif protein, X-linked 2	RBMX2	8	8	1	+
Serine/arginine-rich splicing factor 11	SRSF11	7	5	1	+
Serine/arginine-rich splicing factor 5	SRSF5	7	5	1	+
Serine/arginine-rich splicing factor 7	SRSF7	6	8	1	+
Putative RNA-binding protein 3	RBM3	3	4	1	+
Probable RNA-binding protein 23	RBM23	3	4	2	+
Transformer-2 protein homolog alpha	TRA2A	2	3	1	+
U2 small nuclear RNA auxillary factor 1 isoform a	U2AF1	2	4	1	+
Serine/arginine-rich splicing factor 8	SRSF8	2	5	1	+
Serine/arginine-rich splicing factor 12	SRSF12	2	1	1	+

[1] proteins co-purifying with Flag-Tnpo3 and not from control cell extracts.
[2] number of unique peptides identified in the condition with CHAPS
[3] number of unique peptides identified in the condition without detergent
[4] number of RRM domains recognized
[5] presense (+) or absence (-) of RS or RS-like domain

Supplementary Table S1 | TNPO3 binding partners identified by MS/MS

Proteins copurifying with Flag-Tnpo3 identified specifically with at least 2 unique peptides detected in the condition with or without detergent (CHAPS).
Proteins identified in both conditions (highlighted in green) were considered strong candiates for being Tnpo3 binding partners.

(B) Tnpo3 binding partners with RS (or RS-like) domain, no RRM

Identified Proteins	Gene Name	# pept (with CHAPS)	# pept (no CHAPS)	RRM	RS
Death-inducer obliterator 1	DIDO1	41	11	-	+
ATP-dependent RNA helicase DHX8	DHX8	38	29	-	+
Luc7-like protein 3	LC7L3	28	21	-	+
Splicing factor, suppressor of white-apricot homolog isoform 2	SFSWAP	24	9	-	+
Probable ATP-dependent RNA helicase DDX46	DDX46	21	21	-	+
Putative RNA-binding protein Luc7-like 2	LUC7L2	20	22	-	+
Nuclear receptor coactivator 5	NCOA5	17	16	-	+
Bcl-2-associated transcription factor 1	BCLF1	14	27	-	+
Armadillo repeat containing, X-linked 3 (ARMCX3)	ARMCX3	13	10	-	+
U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein	SNRNP27	7	4	-	+
PHD finger protein 6	PHF6	7	19	-	+
Putative RNA-binding protein Luc7-like 1	LUC7L	7	9	-	+
Serrate RNA effector molecule homolog	SRRT	6	8	-	+
Thyroid hormone receptor-associated protein 3	TR150	6	20	-	+
CDC-like kinase 2	CLK2	5	8	-	+
Arginine/serine-rich coiled-coil protein 2	RSRC2	4	2	-	+
Arginine and glutamate-rich protein 1	ARGL1	4	4	-	+
Pre-mRNA-splicing factor 38B	PR38B	3	2	-	+
Splicing factor 3a, subunit 3, 60kDa (SF3A3)	SF3A3	3	3	-	+
U4/U6.U5 tri-snRNP-associated protein 2	USP39/SNUT2	2	4	-	+
ADP-ribosylation factor-like protein 6-interacting protein 4	ARL6IP4	2	1	-	+
CDC-like kinase 3	CLK3	1	7	-	+
Pre-mRNA cleavage complex 2 protein Pcf11	PCF11	32		-	+
CLASP2 protein	B2RTR1	3		-	+
Cactin	CACTIN		8	-	+
Pre-mRNA-splicing factor 38A	PRPF38A		6	-	+
Eukaryotic translation initiation factor 3 subunit A	EIF3A		6	-	+
Histone-lysine N-methyltransferase SETD2	SETD2		6	-	+
PRP4 pre-mRNA-processing factor 4 homolog/PRP4 kinase	PRPF4B		5	-	+
Protein SCAF11/CTD-associated SR protein 11	SCAF11		4	-	+
Serine/arginine repetitive matrix protein 2	SRRM2		3	-	+
Zinc finger CCCH domain-containing protein 18	NHN1		2	-	+
Peptidyl-prolyl cis-trans isomerase G	PPIG		2	-	+

Supplementary Table S1 | TNPO3 binding partners identified by MS/MS

Proteins copurifying with Flag-Tnpo3 identified specifically with at least 2 unique peptides detected in the condition with or without detergent (CHAPS).
Proteins identified in both conditions (highlighted in green) were considered strong candiates for being Tnpo3 binding partners.

(C) Tnpo3 binding partners with RRM domain(s), no RS domain

Identified Proteins	Gene Name	# pept (with CHAPS)	# pept (no CHAPS)	RRM	RS
RNA-binding protein 14	RBM14	21	30	1	-
Poly(U)-binding-splicing factor PUF60	PUF60	16	13	3	-
RNA-binding protein 42	RBM42	2		1	-
Polypyrimidine tract-binding protein 1	PTBP1	2		4	-
Pre-mRNA-processing-splicing factor 8	PRPF8		10	1	-
HIV Tat-specific factor 1	HTATSF1		5	2	-
RNA-binding protein 4	RBM4		2	1	-

Supplementary Table S1 | TNPO3 binding partners identified by MS/MS

Proteins copurifying with Flag-Tnpo3 identified specifically with at least 2 unique peptides detected in the condition with or without detergent (CHAPS).
Proteins identified in both conditions (highlighted in green) were considered strong candiates for being Tnpo3 binding partners.

(D) Tnpo3 binding partners lacking recognizable RRM and RS-like domains

Identified Proteins	Gene Name	# pept (with CHAPS)	# pept (no CHAPS)	RRM	RS
Cleavage and polyadenylation specific factor 5 (CPSF5, CF-Im-25)	CPSF5	13	8	-	-
Forkhead box protein K1	FOXK1	8	5	-	-
Polyribonucleotide 5'-hydroxyl-kinase Clp1	CLP1	6	3	-	-
Probable ATP-dependent RNA helicase DDX41	DDX41	5	17	-	-
Splicing factor 3B subunit 1	SF3B1	4	5	-	-
U4/U6 small nuclear ribonucleoprotein Prp31	PRP31	3	4	-	-
UPF0428 protein CXorf56	CXorf56	3	8	-	-
U5 small nuclear ribonucleoprotein 200 kDa helicase	U520	3	10	-	-
26S protease regulatory subunit 4	PSMC1	3	1	-	-
Microfibrillar-associated protein 1	MFAP1	2	4	-	-
UPF0468 protein C16orf80 (transcription factor IIB)	CP080	2	2	-	-
26S protease regulatory subunit 7	PRS7	2	1	-	-
NudC domain-containing protein 1	NUDC1	2	1	-	-
Splicing factor 3B subunit 3	SF3B3	1	8	-	-
Heat shock 40 kDa protein 4	DNAJA1	1	2	-	-
Putative uncharacterized protein SP100	SP100	1	2	-	-
Cytoskeleton-associated protein 5	CKAP5	11		-	-
Rho-associated, coiled-coil containing protein kinase 1	ROCK1	10		-	-
Estradiol 17-beta-dehydrogenase 12	DHB12	4		-	-
26S proteasome non-ATPase regulatory subunit 3	PSMD3	4		-	-
Cyclin-K	CCNK	4		-	-
Leucine-rich repeat-containing protein 59	LRC59	3		-	-
Protein FAM162A	F162A	3		-	-
Coiled-coil domain-containing protein 124	CC124	2		-	-
26S protease regulatory subunit 6A	PRS6A	2		-	-
Host cell factor 1	HCFC1	2		-	-
ADP/ATP translocase 2	ADT2	2		-	-
116 kDa U5 small nuclear ribonucleoprotein component	EFTUD2		6	-	-
Pre-mRNA-processing factor 6	PRPF6		5	-	-
Splicing factor 3A subunit 1	SF3A1		4	-	-
Histone deacetylase complex subunit SAP18	SAP18		3	-	-
Retinoblastoma-associated protein RAP140	FAM208A		3	-	-
Splicing factor 3B subunit 2	SF3B2		3	-	-
Smu-1 suppressor of mec-8 and unc-52 homolog (C. elegans)	SMU1		3	-	-
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH		3	-	-
U6 snRNA-associated Sm-like protein LSm7	LSM7		2	-	-
Probable ATP-dependent RNA helicase DHX36	DHX36		2	-	-
WD repeat domain 57 (U5 snRNP specific)	WDR57		2	-	-
Multiple myeloma tumor-associated protein 2	MMTAG2		2	-	-
40S ribosomal protein S11	RPS11		2	-	-
40S ribosomal protein S13	RPS13		2	-	-
60S ribosomal protein L24	RPL24		2	-	-
60S ribosomal protein L13	RPL13		2	-	-
40S ribosomal protein S8	RPS8		2	-	-
U4/U6 small nuclear ribonucleoprotein Prp4	PRPF4		2	-	-

Supplementary Table S1 | TNPO3 binding partners identified by MS/MS

Proteins copurifying with Flag-Tnpo3 identified specifically with at least 2 unique peptides detected in the condition with or without detergent (CHAPS).
Proteins identified in both conditions (highlighted in green) were considered strong candiates for being Tnpo3 binding partners.

(E) General co-factors of nuclear import/export co-purified with Flag-Tnpo3

Identified Proteins	Gene Name	# pept (with CHAPS)	# pept (no CHAPS)	RRM	RS
RAN, member RAS oncogene family	RAN	3	6	-	-
RAN binding protein 1	RANBP1	2	1	-	-

Table S2. Crystallography data collection and refinement statistics.

	TNPO3 (C511A)	**TNPO3 -Ran(Q69L)GTP**	**TNPO3 - ASF/SF2 (106-230)**
Data collection *			
Beamline	ESRF ID14-4	ESRF ID14-4	ESRF ID23-1
Space group	P1	P1	P1
Unit cell parameters			
a, b, c (Å)	96.20, 104.02, 124.89	79.95, 93.51, 104.71	80.61, 91.10, 98.28
α, β, γ (^{o})	90.64, 97.88, 104.48	78.43, 68.29, 68.28	106.85, 100.27, 101.93
Wavelength (Å)	0.9394	0.9798	0.9730
Resolution range (Å)	40-2.95 (3.11-2.95)	40-3.42 (3.60-3.42)	47.3-2.57 (2.71-2.57)
Average multiplicity	3.7 (3.9)	5.8 (5.9)	3.4 (3.5)
Completeness (%)	98.6 (98.7)	99.2 (99.3)	99.3 (99.2)
R_{merge}	0.136 (1.13)	0.122 (1.68)	0.061 (0.71)
$\langle I \rangle / \langle \sigma(I) \rangle$	8.1 (1.4)	8.6 (1.5)	10.3 (1.6)
Wilson $\langle B \rangle$ (Å^2)	78.3	133.7	80.8
Solvent content (%)	57.3	53.2	55.7
Refinement			
Resolution	**40.0-2.95**	**40-3.42**	**40-2.57**
Reflections (total/free)	95,818/4,834	34,962/1,764	79,944/4,037
R_{work}/R_{free}	0.236/0.266	0.268/0.290	0.221/0.269
No. Atoms			
Protein	28,261	15,700	15,806
Ligand/ion	32	66	1
Water	-	-	51
Average B-factors (Å^2)			
Protein	90.8	138.0	64.6
Ligand/ion	104.5	77.7	70.5
Water	-	-	47.4
R.M.S. Deviations			
Bond lengths (Å)	0.003	0.003	0.005
Bond angles (°)	0.77	0.75	0.924
Ramachandran (%)			
Favored	96.6	97.5	97.1
Outliers	0	0	0.15
PDB ID	**4C0P**	**4C0Q**	**4C0O**

* Values in parentheses indicate the highest resolution shell.

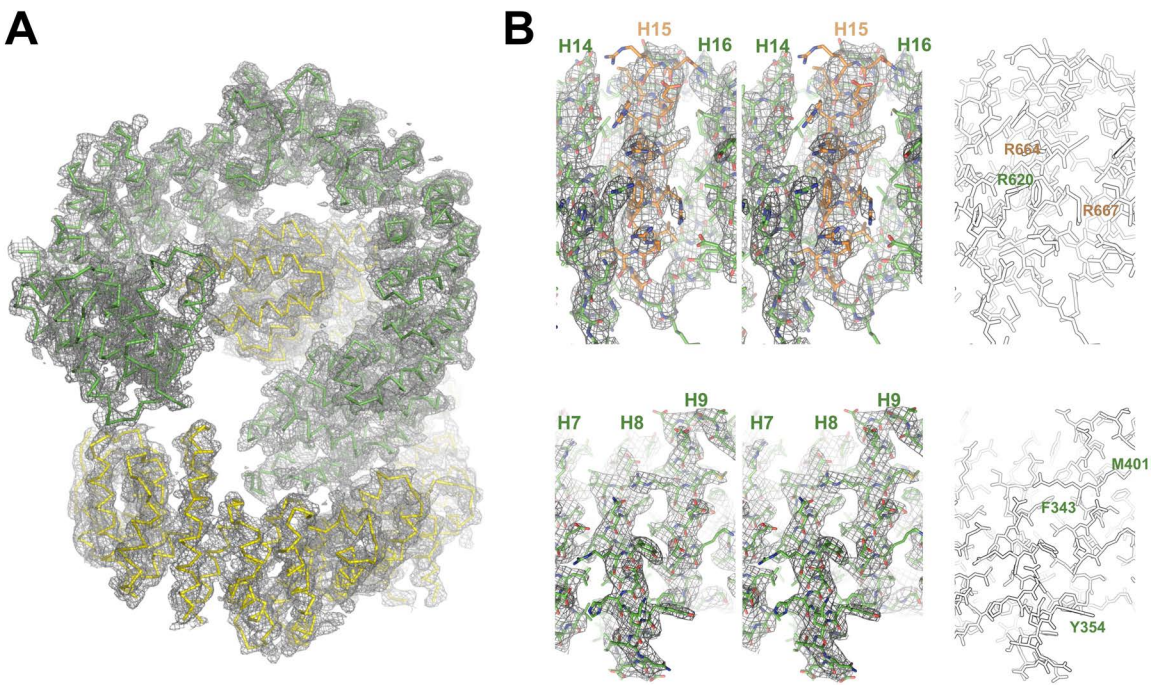

Fig. S1. Examples of electron density for the unliganded Tnpo3 structure. (*A*) Weighted *2Fo-Fc* map of the dimer (chains A and B, shown as green and yellow ribbons, respectively, and collectively representing half of the asymmetric unit), displayed as grey chicken wire and contoured at 1σ. (*B*) Stereo views of HEAT repeats 14-16 (top panels) and HEAT repeats 7-9 (bottom panels). The protein chain is shown in sticks mode, and selected amino acid residues are indicated on the black-and-white panels to the right.

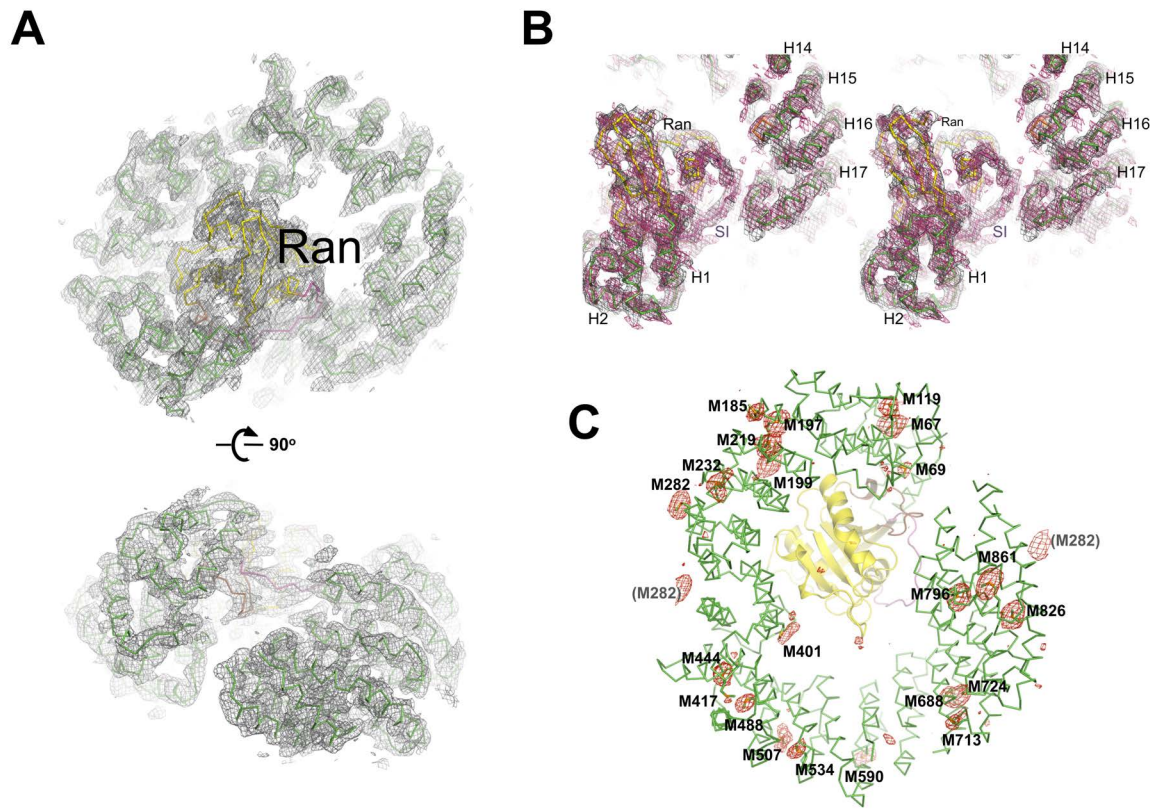

Fig. S2. Electron density and phased anomalous maps of the Tnpo3-RanGTP complex. (*A*) Weighted *2Fo-Fc* map for the refined structure (Tnpo3 chain A and Ran chain C, are shown as green and yellow cartoons, respectively) representing half of the asymmetric unit, contoured a 1σ (gray wire) and shown in two orthogonal views. (*B*) Close-up on a part of the structure (shown in same orientation as in Figs. S3B and C) with weighted *2Fo-Fc* and composite omit maps shown as gray and red chicken wires, respectively and contoured at 1σ. (*C*) Phased anomalous map calculated using diffraction data collected at the selenium edge (X ray wavelength 0.9798 Å) and contoured at 3σ (red chicken wire). Selenomethionine residues belonging to the displayed heterodimer and from the neighboring molecules in the crystal are indicated in bold black and grey print, respectively. Note that only the Tnpo3 subunit was derivatized with selenomethionine in these crystals.

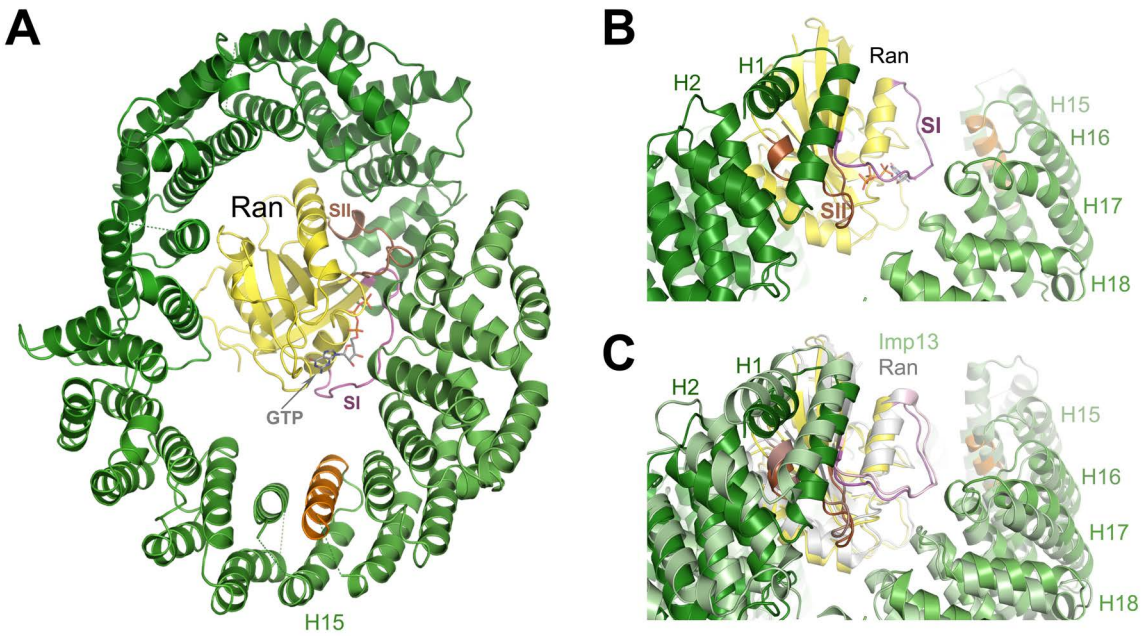

Fig. S3. (*A*) Overview of the Tnpo3-RanGTP structure with protein chains shown as cartoons. Positions of Ran Switch I (SI, magenta) and II (SII, brown), GTP, and HEAT repeat 15 of Tnpo3 are indicated. R-helix is shown in orange. (*B*) Close-up on a part of the structure showing contacts between Ran Switch I and HEAT repeats 16-18 discussed in the main text. (*C*) Same as in panel B, but superposed on Imp13-RanGTP structure (PDB ID 2X19, Bono et al., 2010). The Imp13 and Ran chains of the Imp13-RanGTP complex are shown in lighter colors.

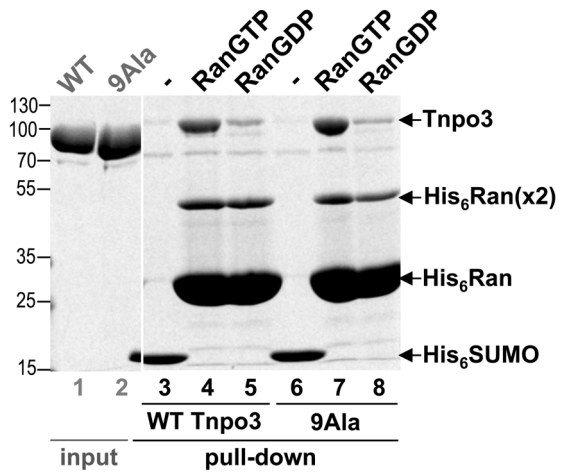

Fig. S4. The interaction of WT Tnpo3 and 9Ala mutant with His_6-Ran in the presence of GTP (lanes 4 and 7) or GDP (lanes 5 and 8) in a His_6 tag pull down assay. His_6SUMO was used as a negative control (lanes 3 and 6); input quantities of WT and 9Ala Tnpo3 are shown in lanes 1 and 2.

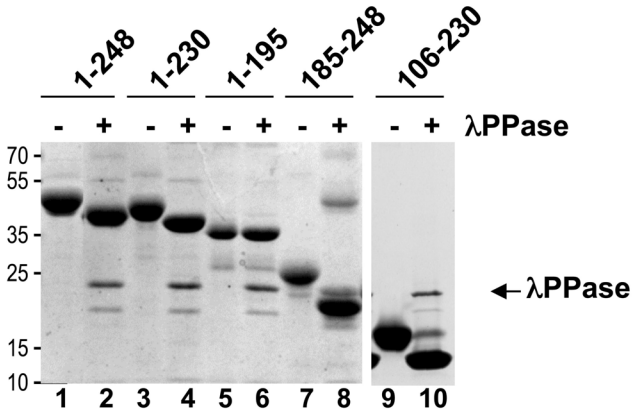

Fig. S5. Treatment of ASF/SF2 constructs with bacteriophage λ phosphatase. His_6SUMO fusions of full-length (lanes 1-2) or truncated (lanes 3-10) ASF/SF2 produced in the presence of CLK1 were separated in 11% SDS PAGE gel prior (odd numbered lanes) or after (even numbered lanes) treatment with bacteriophage λ phosphatase (λPPase).

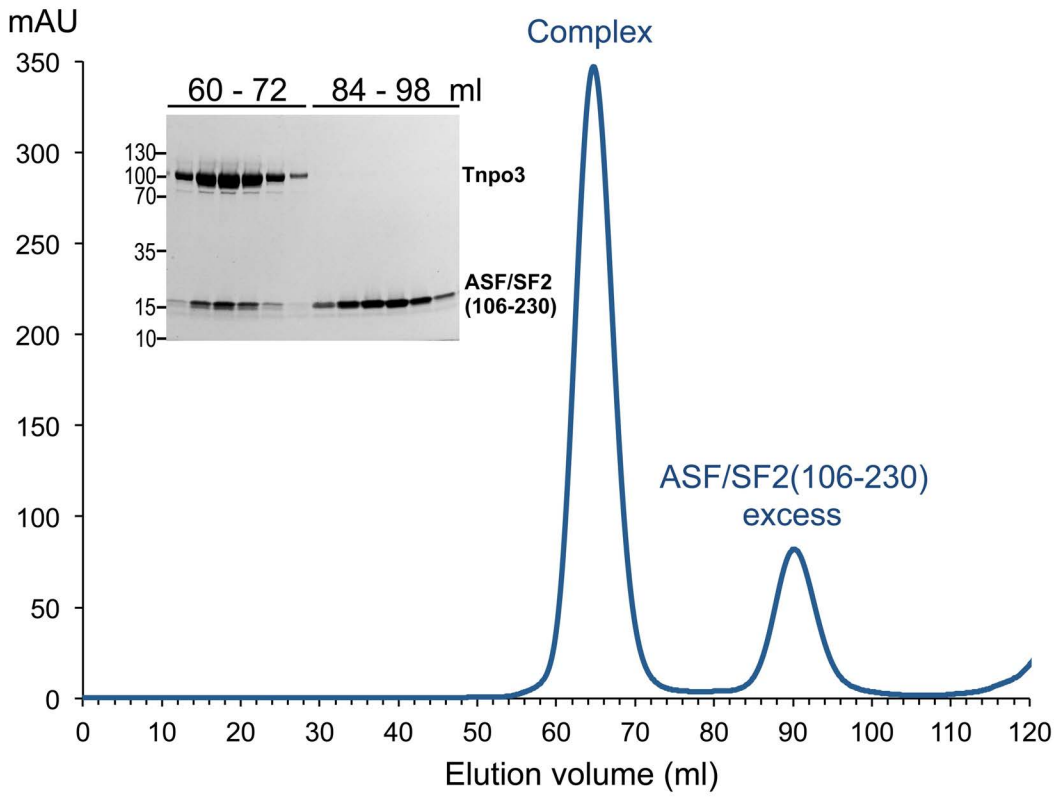

Fig. S6. Isolation of the Tnpo3-ASF/SF2(106-230) complex by gel filtration, detected by absorbance at 280 nm. The inset shows separation of the peak fractions corresponding to elution volumes 60-72 ml and 84-98 ml by SDS PAGE.

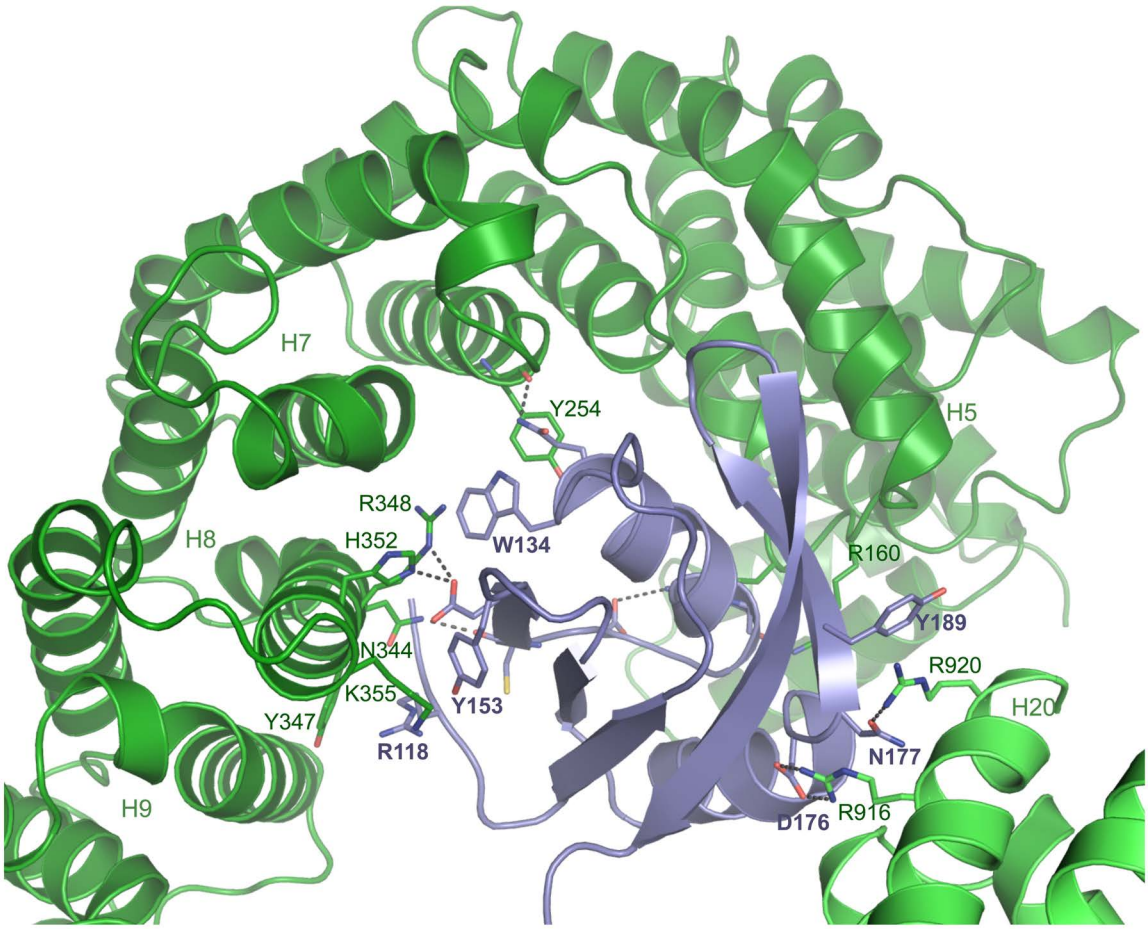

Fig. S7. Details of the interactions between Tnpo3 and ASF/SF2 RRM2. Protein chains are shown as cartoons with selected amino acids in sticks and indicated.

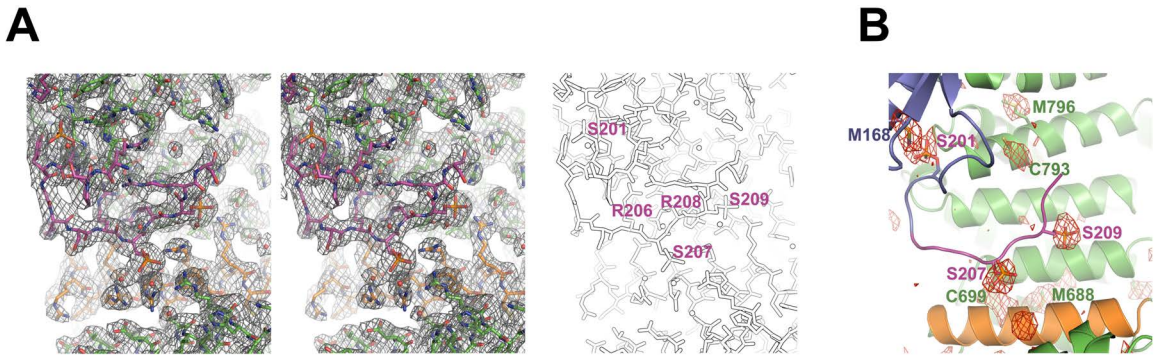

Fig. S8. An example of the electron density map for the refined Tnpo3-ASF/SF2 structure and validation of the phosphorylation state of ASF/SF2 within the crystallized complex by anomalous scattering. (*A*) Stereoview of a section of weighted *2Fo-Fc* map for the Tnpo3-ASF/SF2 complex, contoured at 1σ (grey wire). The protein is shown in sticks mode, and selected amino acid residues are indicated on the black-and-white image to the right. (*B*) Phased anomalous map calculated using diffraction data collected at an X-ray wavelength of 2.38 Å and model-based phases, contoured at 3σ (red wire). The peaks identify positions of the phosphate groups attached to Ser residues 201, 207 and 209. Sulfur atoms, also possessing anomalous signal at this X ray energy can likewise be identified in this map. The protein is shown as cartoons with selected amino acid residues as sticks and indicated. The color scheme is conserved from the figures in the article.

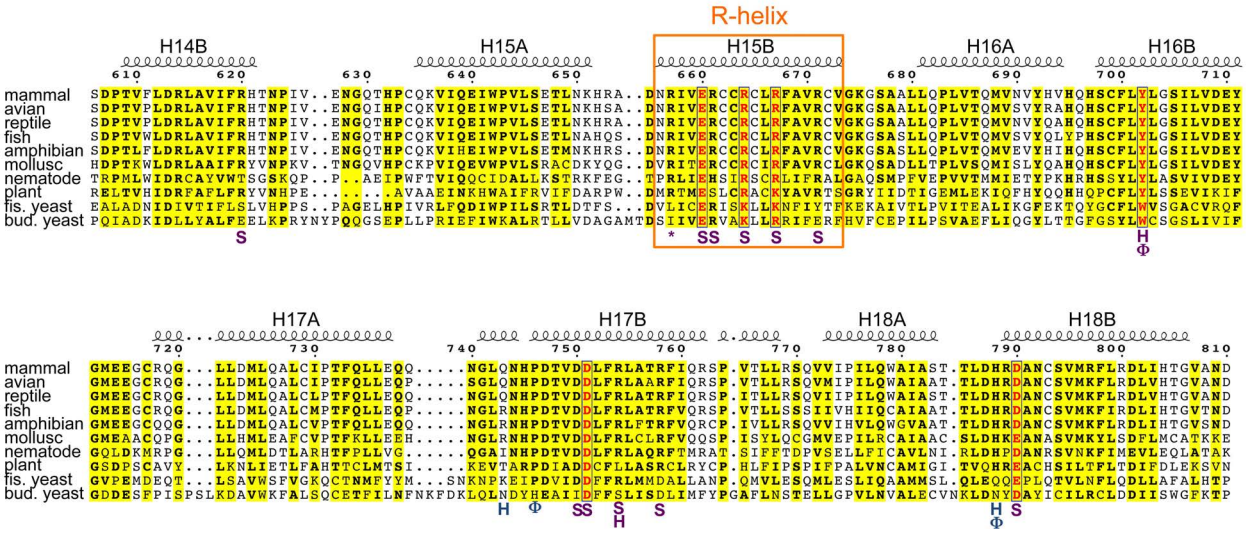

Fig. S9. Amino acid sequence alignment of the human Tnpo3 RS binding region with predicted or known orthologous proteins. The protein sequences were from *Homo sapiens* (mammal), *Gallus gallus* (avian), *Chrysemys picta bellii* (reptile), *Danio rerio* (fish), *Xenopus laevis* (amphibian), *Crassostrea gigas* (mollusc), *Caenorhabditis elegans* (nematode), *Arabidopsis thaliana* (plant), *Schizosaccharomyces pombe* (fiss. yeast), and *Saccharomyces cerevisiae* (bud. yeast). The secondary structure elements indicated atop the alignment and the residue numbering correspond to the human Tnpo3 structure. Select invariant/conserved residues are shown in bold red print and boxed; other residues conserved in most of the orthologs are highlighted in yellow. The R-helix is indicated with an orange box. Human Tnpo3 residues involved in salt bridges, hydrogen bonds and hydrophobic interactions with ASF/SF2 are indicated with the respective symbols S, H, and Φ, below the alignment (magenta symbols indicate Tnpo3 residues interacting with the RS domain, and blue symbols indicate those making contacts with the remainder of the ASF/SF2 molecule).

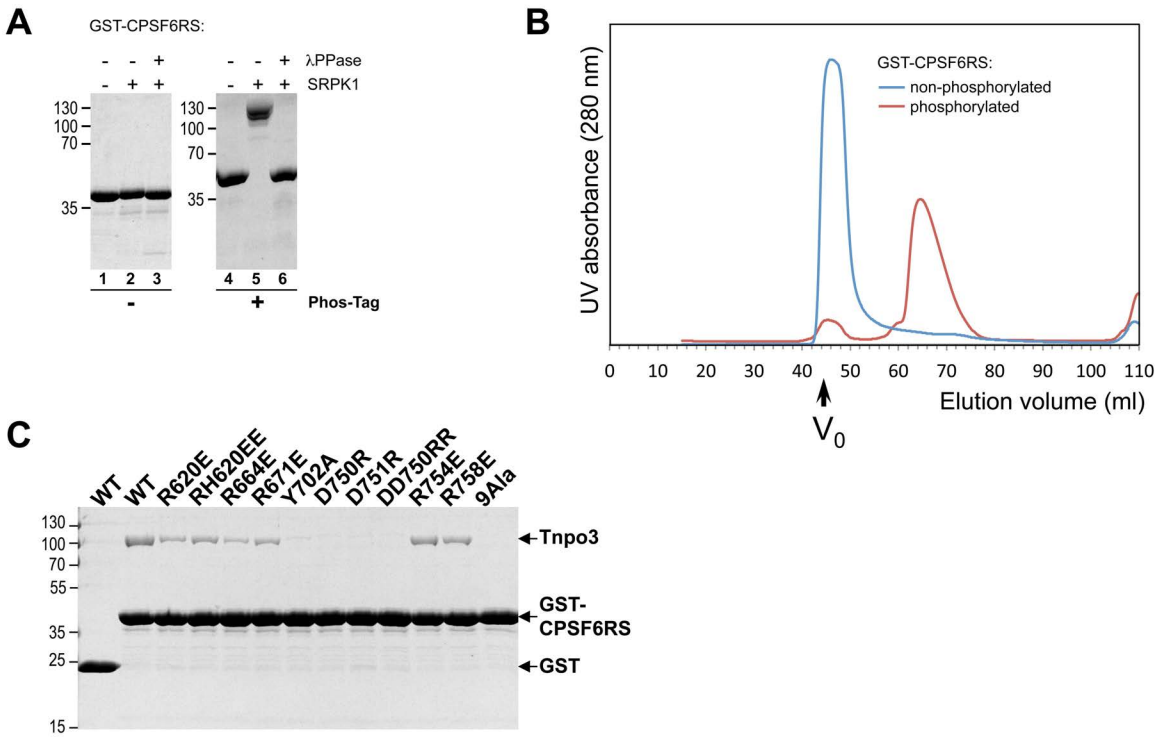

Fig. S10. (*A*) Separation of GST-CPSF6RS in 10% SDS PAGE gels in the absence (lanes 1-3) or presence (lanes 4-6) of co-polymerized PhosTag-Mn(II) complex (lanes 4-6). The protein was expressed in the absence (lanes 1 and 4) or presence (lanes 2, 3, 5 and 6) of SRPK1. The material in lanes 3 and 6 was treated with bacteriophage λ phosphatase. Note the phosphorylation-dependent migration shift in the presence of PhosTag. (*B*) Elution of GST-CPSF6RS expressed in the absence (blue trace) or presence (red trace) of SRPK1 from a Superdex 200 size exclusion chromatography column, detected by absorbance at 280 nm. Void volume (V_0) is indicated. (*C*) Pull down of Tnpo3 and its mutant variants with non-phosphorylated GST-CPSF6RS. The conditions are otherwise identical to those in the experiment shown in Fig. 4E.

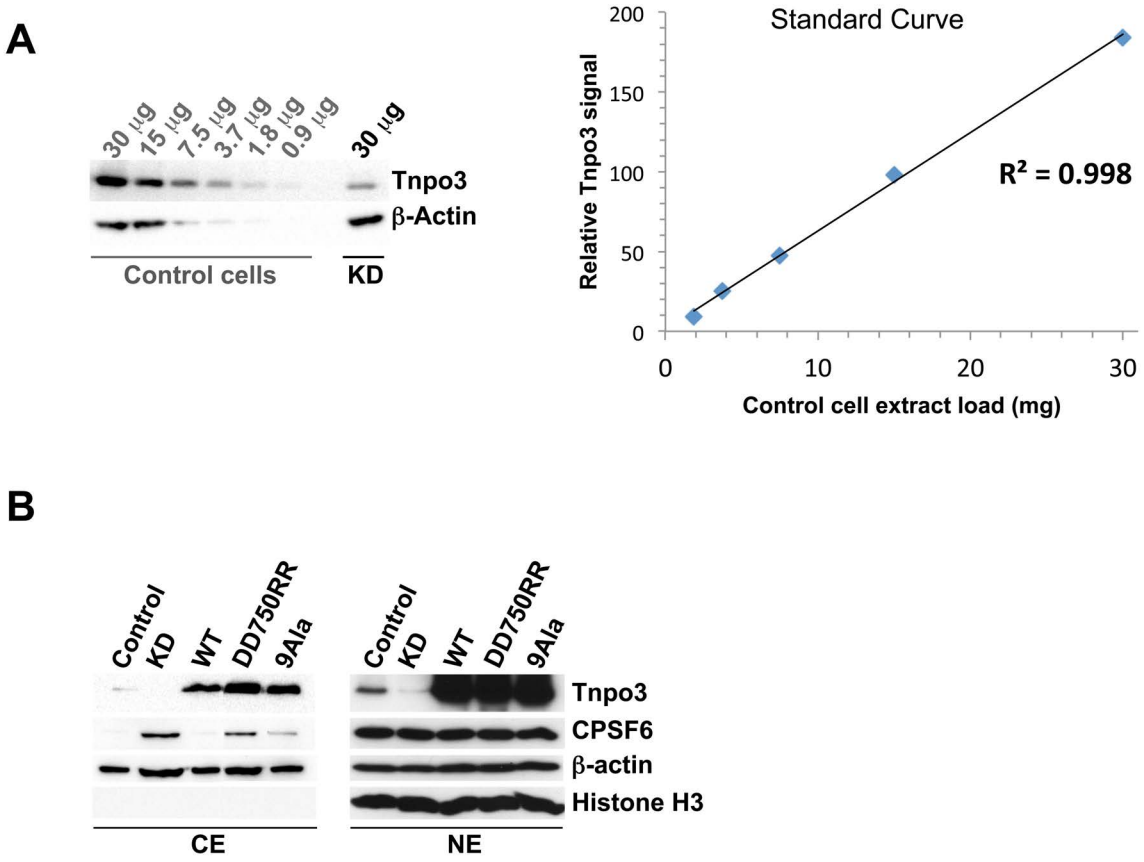

Fig. S11. (*A*) Western blot titration experiment used to determine the level of Tnpo3 depletion upon knockdown. Integrated volume chemiluminescence intensities of a standard curve showing linear correlation between Tnpo3 signal and the amount of loaded control cell lysate is shown to the right of the blot. Tnpo3 signal in control and knockdown cells were normalized by the β-actin signal. (*B*) Detection of Tnpo3, CPSF6, β-actin and histone H3 in cytoplasmic (CE) and nuclear (NE) extracts of control and Tnpo3 knockdown cells before and after expression of WT, DD750RR or 9Ala Flag-Tnpo3.

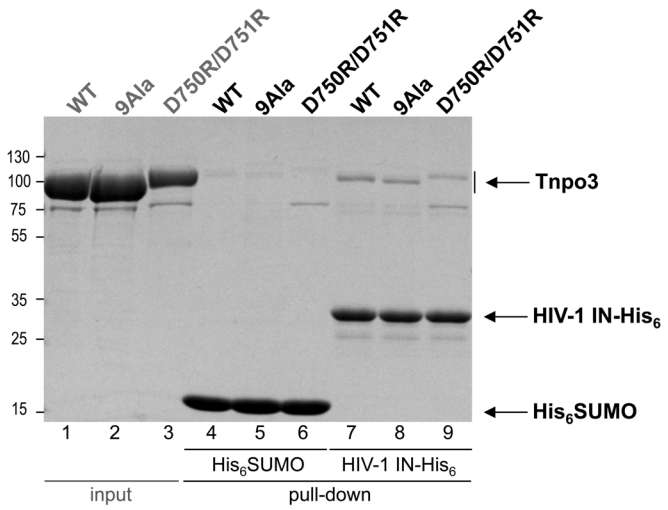

Fig. S12. His_6 tag pull down of WT (lanes 4 and 7), 9Ala (lanes 5 and 8) and DD750RR (lanes 6 and 9) Tnpo3 with His_6SUMO (as a negative control, lanes 4-6) or His_6-tagged HIV-1 integrase (lanes 7-9) on NiNTA agarose resin. Lanes 1-3 contained input quantities of Tnpo3 proteins.

References

1. Mossessova E & Lima CD (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol Cell* 5(5):865-876.
2. Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, & Clardy J (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* 229(1):105-124.
3. Cherepanov P (2007) LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. *Nucleic Acids Res* 35(1):113-124.
4. Kutay U, Izaurralde E, Bischoff FR, Mattaj IW, & Gorlich D (1997) Dominant-negative mutants of importin-beta block multiple pathways of import and export through the nuclear pore complex. *EMBO J* 16(6):1153-1163.
5. Kabsch W (2010) Xds. *Acta Crystallographica. Section D: Biological Crystallography* 66(Pt 2):125-132.
6. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62(Pt 1):72-82.
7. McCoy AJ*, et al.* (2007) Phaser crystallographic software. *J. Appl. Cryst.* 40:658-674.
8. Bono F, Cook AG, Grunwald M, Ebert J, & Conti E (2010) Nuclear import mechanism of the EJC component Mago-Y14 revealed by structural studies of importin 13. *Mol Cell* 37(2):211-222.
9. Grunwald M & Bono F (2011) Structure of Importin13-Ubc9 complex: nuclear import and release of a key regulator of sumoylation. *EMBO J* 30(2):427-438.
10. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica. Section D: Biological Crystallography* 62(Pt 9):1002-1011.
11. Murshudov GN, Vagin AA, & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53(Pt 3):240-255.

12. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213-221.
13. Ngo JC*, et al.* (2008) A sliding docking interaction is essential for sequential and processive phosphorylation of an SR protein by SRPK1. *Mol Cell* 29(5):563-576.
14. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132. .
15. Chen VB*, et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12-21.
16. Krishnan L*, et al.* (2010) The requirement for cellular transportin 3 (TNPO3 or TRN-SR2) during infection maps to human immunodeficiency virus type 1 capsid and not integrase. *J Virol* 84(1):397-406.
17. Maertens GN, El Messaoudi-Aubert S, Elderkin S, Hiom K, & Peters G (2010) Ubiquitin-specific proteases 7 and 11 modulate Polycomb regulation of the INK4a tumour suppressor. *EMBO J* 29(15):2553-2565.
18. Matreyek KA & Engelman A (2011) The requirement for nucleoporin NUP153 during human immunodeficiency virus type 1 infection is determined by the viral capsid. *J Virol* 85(15):7818-7827.
19. Melchior F (1998) Nuclear protein import in a permeabilized cell assay. *Methods Mol Biol* 88:265-273.
20. Levy DN, Aldrovandi GM, Kutsch O, & Shaw GM (2004) Dynamics of HIV-1 recombination in its natural target cells. *Proc Natl Acad Sci U S A* 101(12):4204-4209.