Supplemental Information

Structural Determinants and Mechanism

of Mammalian CRM1 Allostery

Nicole Dölker, Clement E. Blanchet, Béla Voß, David Haselbach, Christian Kappel, Thomas Monecke, Dmitri I. Svergun, Holger Stark, Ralf Ficner, Ulrich Zachariae, Helmut Grubmüller, and Achim Dickmanns

<u>Inventory of Supplemental Information</u>

Supplemental Figures

Figure S1: related to figure 1 with panel A showing the changes by the radius of gyration of solely CRM1 in the various simulations. Panel B shows the little structural and positional changes observed for the acidic loop in comparison the crystal structure used as starting point.

Figure S2: related to figure 1 picturing the two significantly different conformations of the acidic loop found in crystal structures representing different stages of the transport cycle.

Figure S3: related to figure 2 and 3 showing the alignment of the two CRM1 orthologs, namely mouse and human, used.

Figure S4: related to figure 2 showing the merged SAXS patterns (original data), the *ab-initio* models obtained, as well as the backwards calculation of the *ab-initio* models.

Figure S5: related to figure 2 showing the calculated SAXS patterns of the CRM1 either alone or in complex with RanGTP and / or SPN1 derived from the crystal structure PDBid 3GJX to demonstrate their (dis-)similarity to the equivalent complexes in solution.

Figure S6: related to figure 3 showing the original data obtained in the EM experiments in panel (A) whereas the class averages and the respective back projections of the two conformers are shown in panel (B).

Figure S7: related to figure 3 showing in panel (A) the Fourier shell correlation (FSC) curves for both, the extended and compact EM density for approximation of the resolution. Panel (B) shows the eight most frequent conformers of CRM1 underlying the model of the compact conformation shown in figure 3.

Figure S8: related to figure 4, 5 and 6, showing in panel (A) the experimental setup and directions of pulling during the force probe simulations (figure 4 and 5) and in panel (B) the CRM1 molecule and the definition of parts thereof used to obtain the results shown in those figures. Panel (C) shows additional force probe simulations related to figure 6 at a slower rupture force.

Figure S9: related to figure 9, showing in panel (A) the changes in the distance between HEAT repeats, highlighting the ones important for cargo or RanGTP binding.

Figure S10: related to figure 9 and S9, showing the changes in the distances between HEAT 11 and 12 A Helices (grey), the HEATs forming the NES-binding cleft, highlighting the changes in this cleft preventing or enabling the binding of the NES (shown as stick model).

Figure S11: related to figure 9, correlating the distances of the centres of mass of the NES-binding cleft helices 11A and 12A observed in MD-simulations to the data derived from known X-ray structures, the distances of which are also referred to in the main body of the text.

Supplemental Tables

Table 1: related to Table 1, Figure 1 and 2 summarizing the radius of gyrations found in the individual experiments/simulations and the known X-ray structures.

Table 2: related to Table 1, Figure 1 and 2 summarizing the RMSD values of the molecules RanGTP and the SPN1-Cap binding domain found in the individual crystal structures to indicate their rigidity.

Table 3: related to Figure 3 indicating the flexibility of the conformation of CRM1 found in the sample of CRM1 used the SAXS measurement.

Supplemental Figures
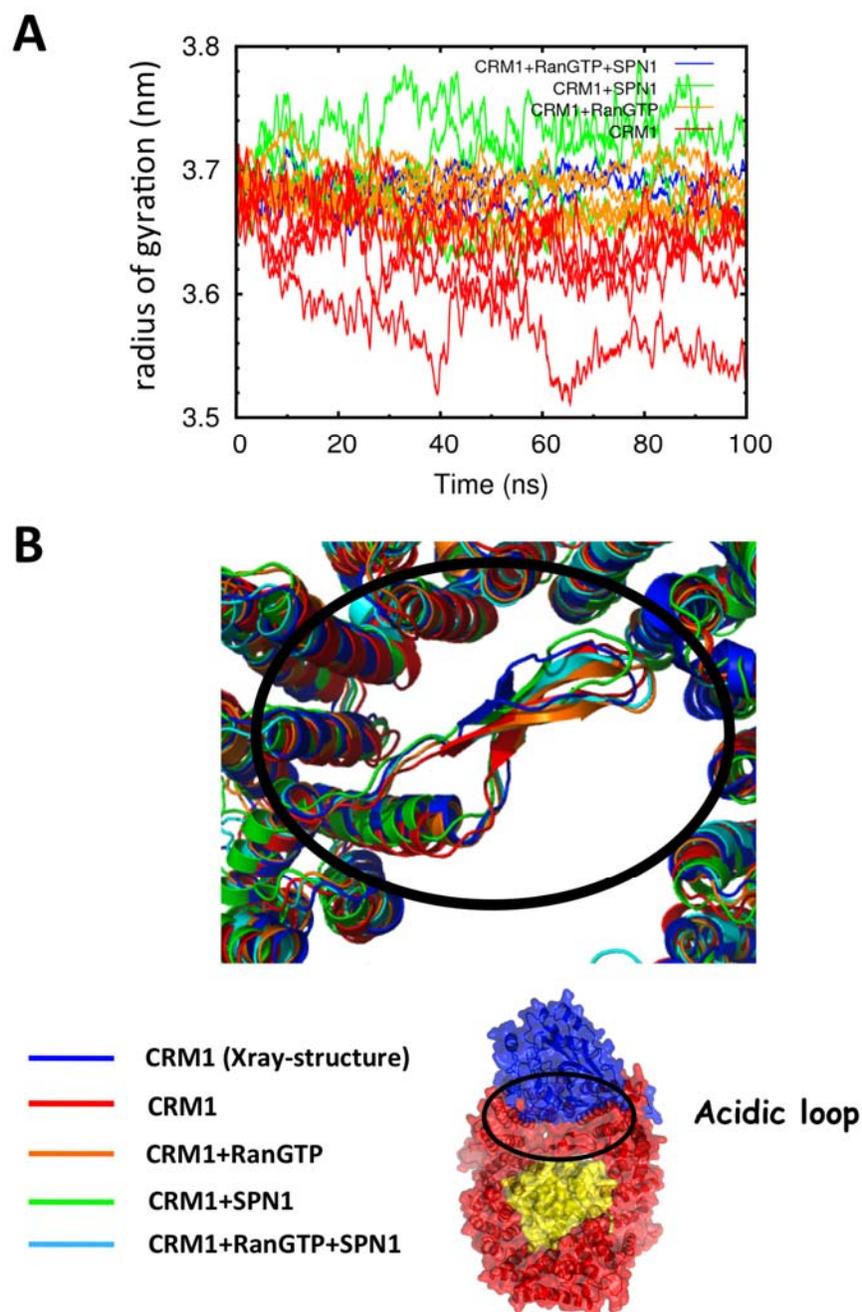
Figure S1: Changes in the radius of gyration of CRM1 do not involve the acidic loop. (A) Changes in the radius of gyration of CRM1 during MD-simulations upon removal of RanGTP and/or SPN1. The crystallographic CRM1 structure was compared to the simulation data. Color code for the individual complexes is indicated in the panel. **(B)** The acidic loop remains in a stable conformation during the MD-simulations. The AL bridges the ends of the HEAT repeat region that binds SPN1 (overview bottom right) and maintains this connection throughout the simulation as depicted by representative snapshots of the individual simulations. Colour code as indicated in bottom left **(related to Figure 1)**.
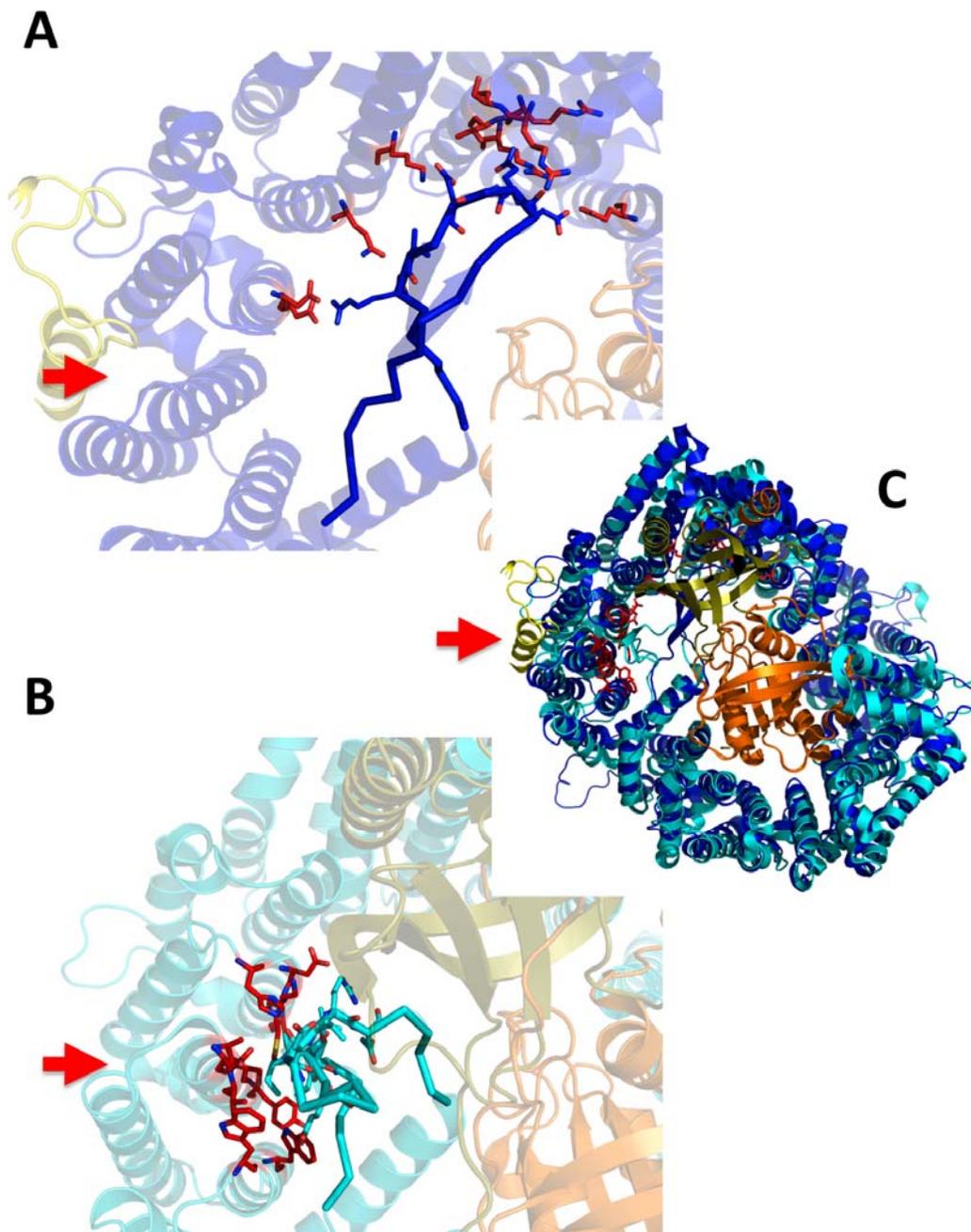
Figure S2: The acidic loop is relocalised to the backside of the NES-binding cleft upon RanBP1 binding thereby aiding in cargo release. The two positions of the AL (Ran and RanBP1 are indicated in light orange and gold, respectively, the NES helix is depicted in yellow). **(A)** Seatbelt conformation as found in 3GJX and **(B)** Disassembly conformation as found in 3M1I. **(C)** Superposition using Ran as fitting molecule and overall view of the two structures with the AL and interacting residues depicted in colors as indicated in panels A. and B. **(related to Figure 1).**

Figure S3: Alignment of human and mouse CRM1 sequences.
The alignment has been obtained using the online version of ClustalW and Espript with standard settings, shown the identical (red background) and different (white background) residues **(related to figure 2)**.
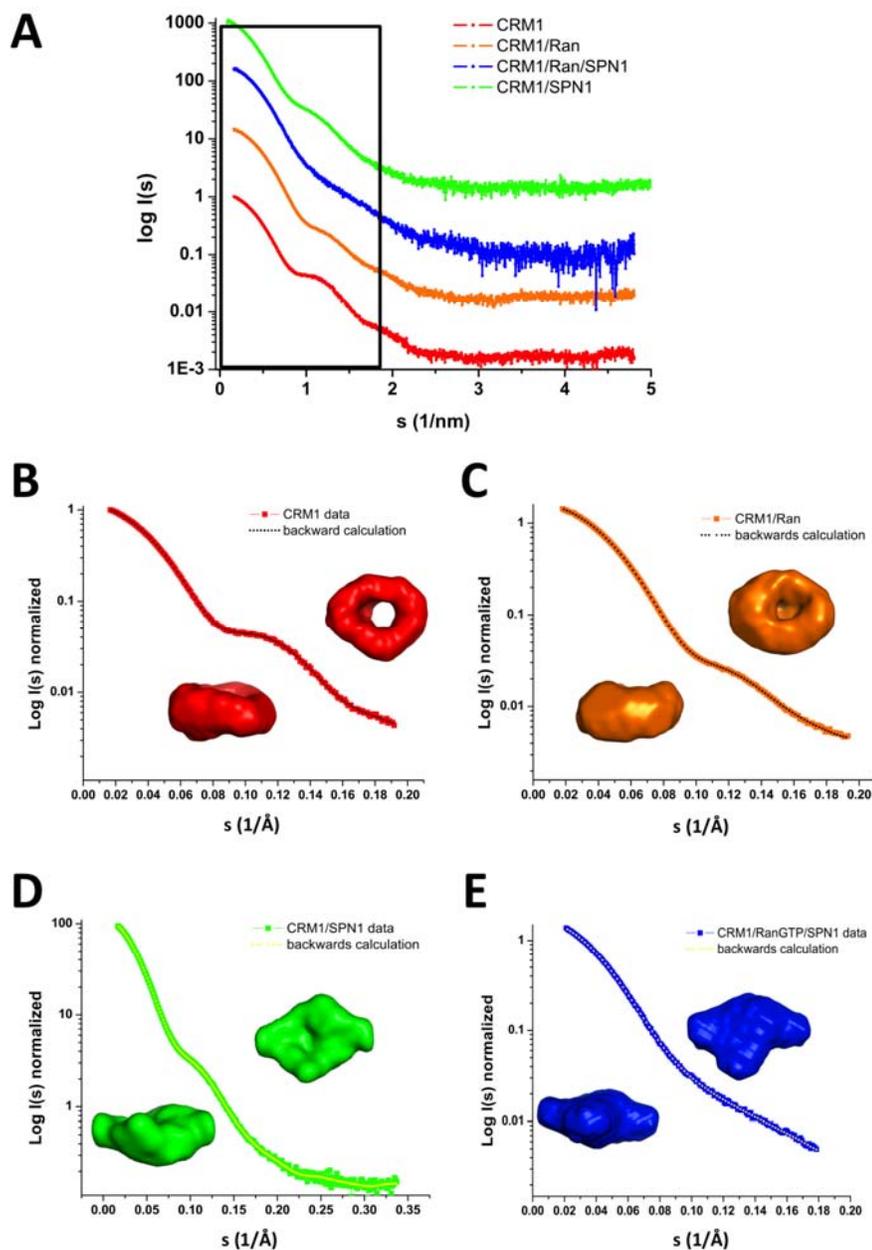
Figure S4: SAXS patterns and *ab-initio* models. (A) Processed solution scattering SAXS patterns of CRM1 (red), CRM1-RanGTP-NES (orange), CRM1-SPN1 (green) and CRM1-RanGTP-SPN1 (blue) purified as described in "Supplemental Experimental Procedures". The data have been normalized and plotted with an offset for better visualization. **(B) - (E)** Calculated *ab-initio* models calculated and their fit to the measured individual data curves. Processed solution scattering pattern from CRM1 **(A)**, CRM1-RanGTP-NES **(B)**, CRM1-SPN1 **(C)** and CRM1-RanGTP-SPN1 **(D)**. *Ab initio* models of CRM1 (red), CRM1-RanGTP-NES (orange), CRM1-SPN1 (green) and CRM1-RanGTP-SPN1 (blue) with their fit to experimental data are shown. Two views of the individual models are depicted. The curves indicate the experimental data, while the black dots show the data calculated from the obtained models, each indicating a good fit for the models obtained **(related to Figure 2)**.
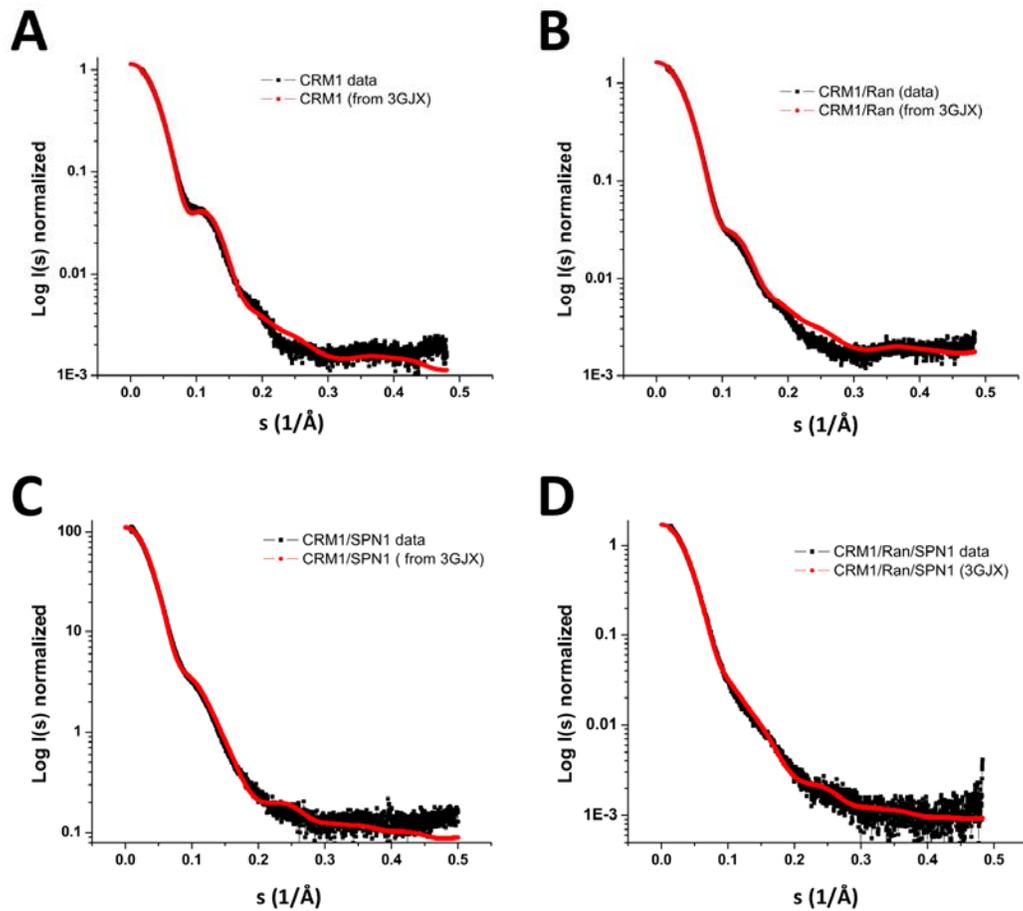
Figure S5: Multiple changes in CRM1 from the X-ray structures are required for an optimal fit to the measured SAXS data. Experimental scattering curves measured for CRM1 and the respective complexes (black) and the back projections of the crystal structures indicated in the panel (red) **(related to Figure 2)**.

A

B

Averages Backprojection

extended conformation

C

Averages Backprojection

compact conformation
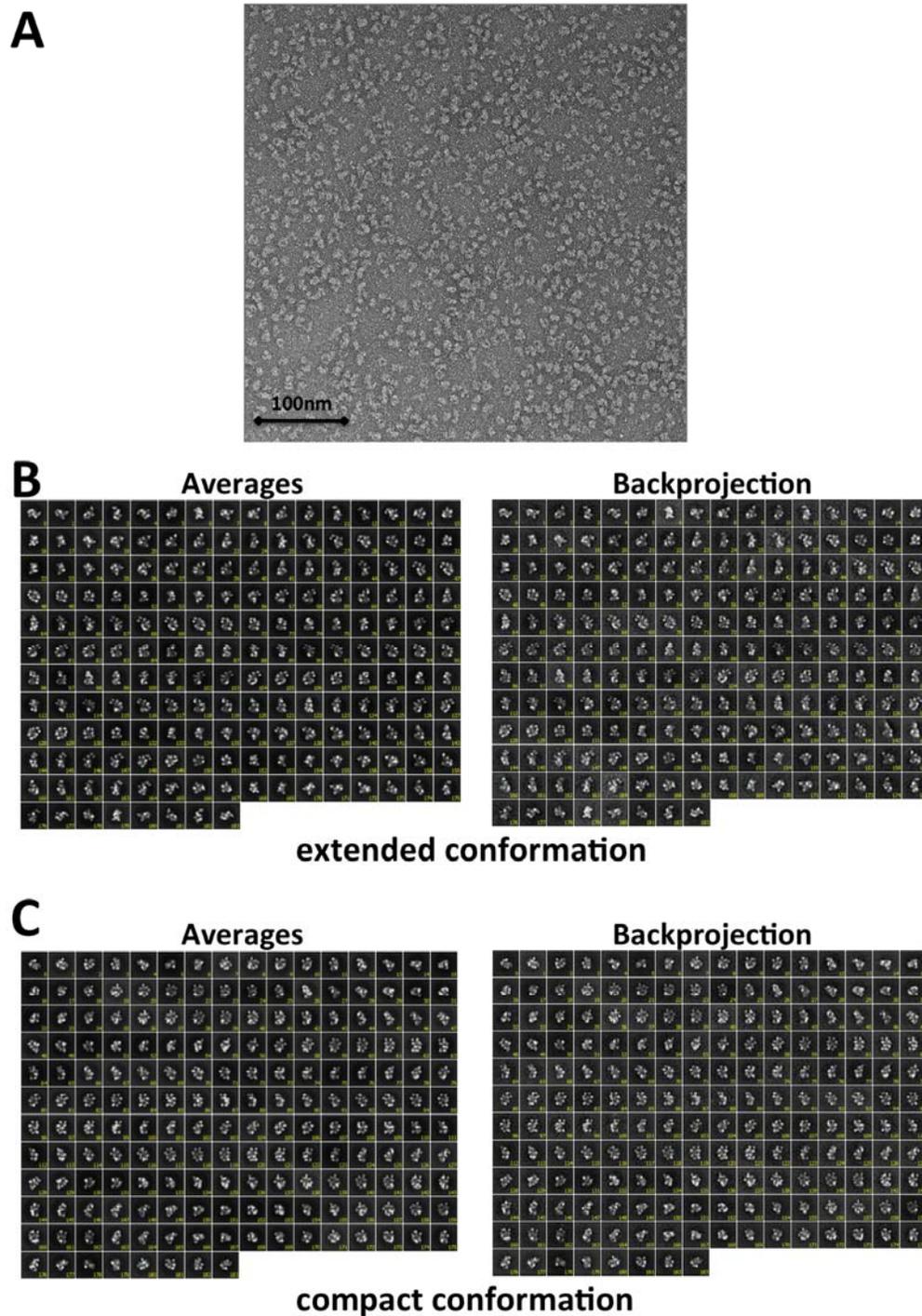
Figure S6: Electron microscopy data and processing (A). Representative EM raw image. **B. C**. Reconstruction of the extended **(B)** and compact **(C)** conformations of *hs*CRM1. The particle classes 1-184 for each reconstruction are shown. Each panel on the left contains a respective particle sum and each panel on the right the respective back projection of a given particle class **(related to Figure 3)**.
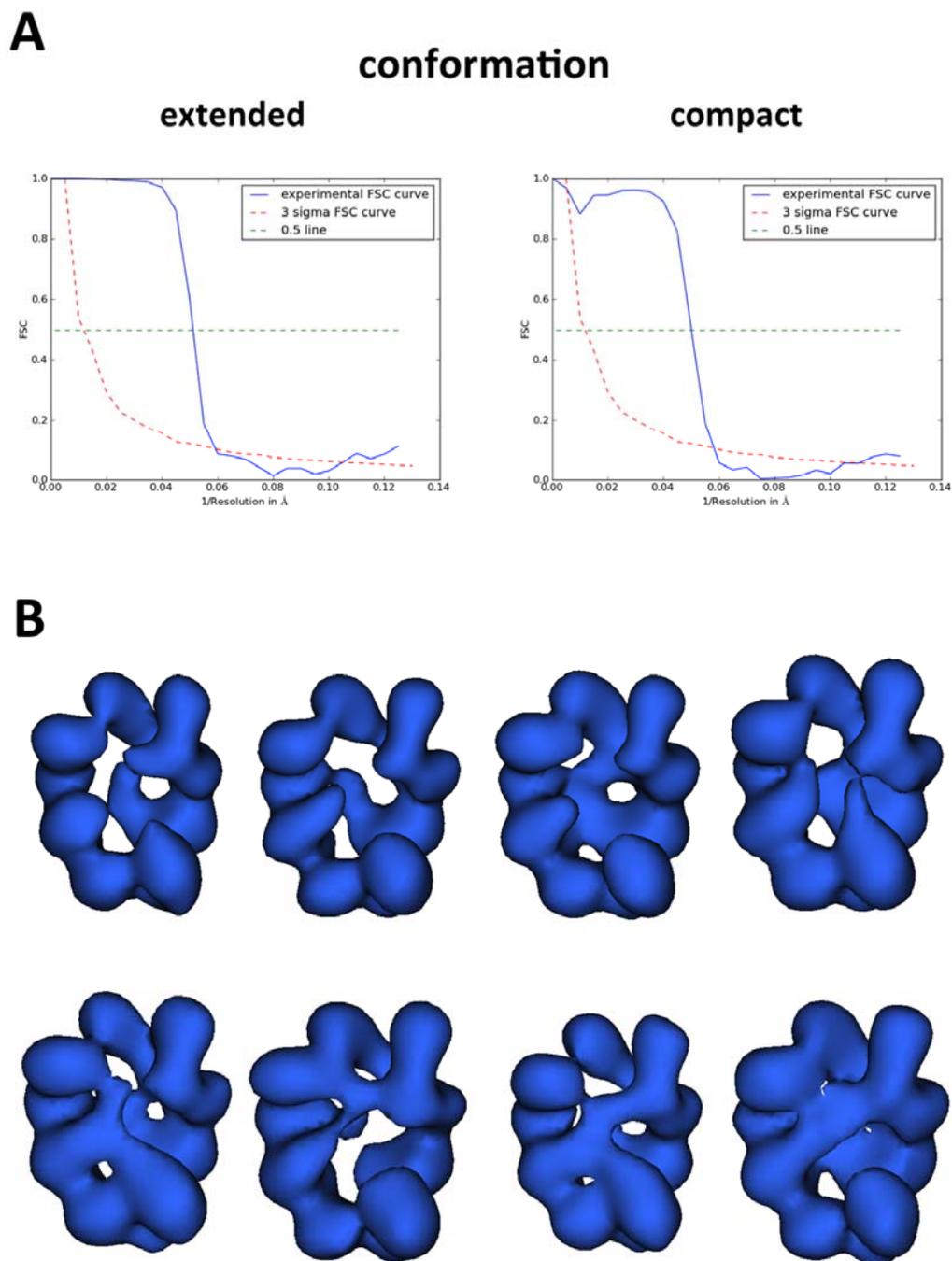
Figure S7: Electron microscopy data analysis. (A) Fourier shell correlation (FSC) curves for both, the extended **(left panel)** and compact **(right panel)** EM density maps were used to approximate the resolution of the respective models. **(B)** 3D-Classaverages after 3D-MSA. 10000 volumes from randomly chosen images, belonging to the compact conformation were reconstructed. Those noisy volumes were averaged into 150 classes. The eight most frequent conformers are shown **(related to Figure 3)**.
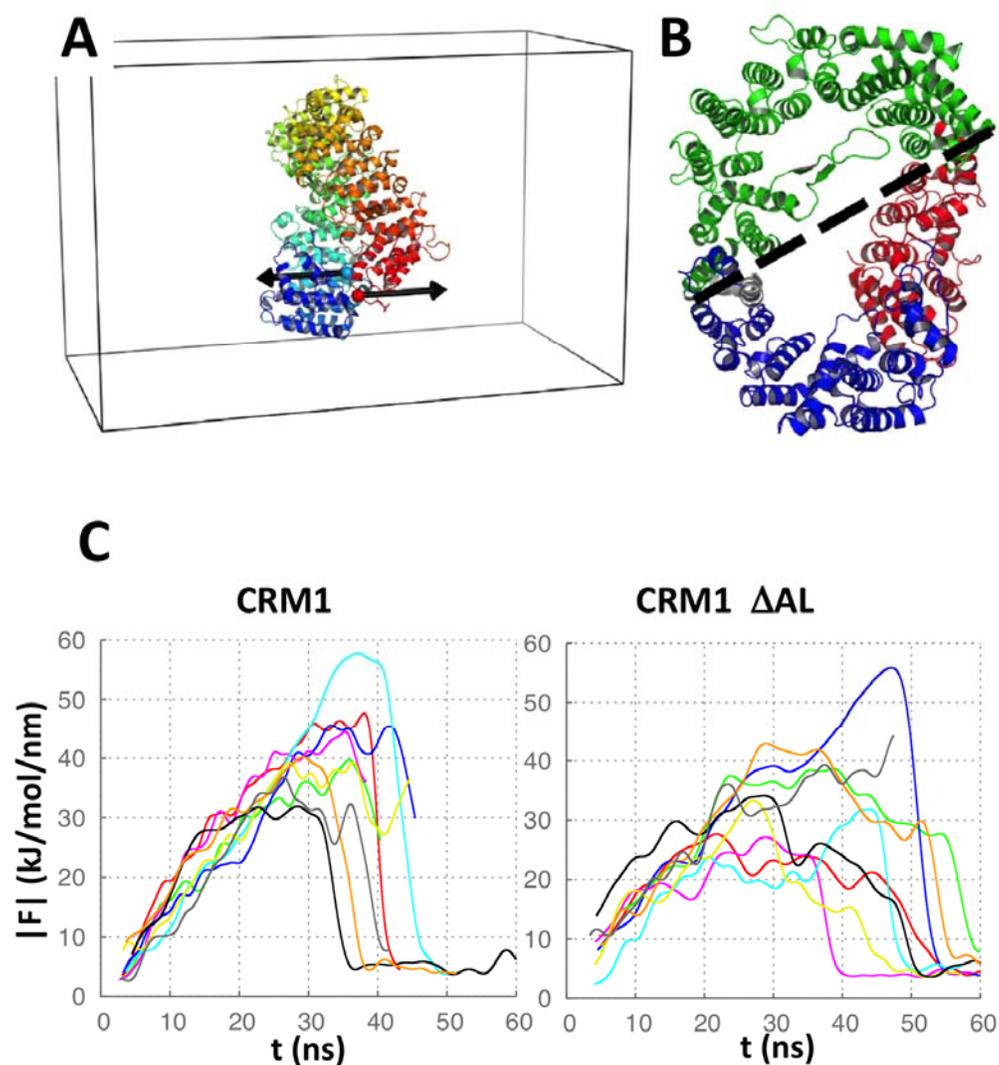
Figure S8: Setup of MD simulations and definition of regions of CRM1. Force Probe simulation set-up and simulation sub-systems of CRM1. **(A)** Simulation setup used for force probe simulations of CRM1. A side view of CRM1 is shown in cartoon representation (rainbow colours, blue and red mark the N- and C-terminus, respectively), the simulation box is shown as black lines. Two independent pulling potentials acting on the Ca-atoms of Gln185 (blue sphere) and Phe1055 (red sphere) moving in opposite directions (arrows) were applied. **(B)** Top view of CRM1. The terminal regions Ala12 to Val274 and Ile815 to Ser1055 are marked in blue and red, respectively. Residues Arg344 to Leu811 (acidic loop and connecting HEAT repeats) are marked in green. **(C)** By applying a time-dependent harmonic biasing potential, CRM1 is brought from the compact into an extended conformation. The average over the maximal occurring forces during these force probing simulations, the rupture force, reflects the shape and height of the energetic barrier separating compact and extended conformation. Comparing these rupture forces for wild type (left panel) and AL deletion mutant (right panel) simulations suggests that the AL does not significantly influence this energetic barrier **(related to Figure 4)**.
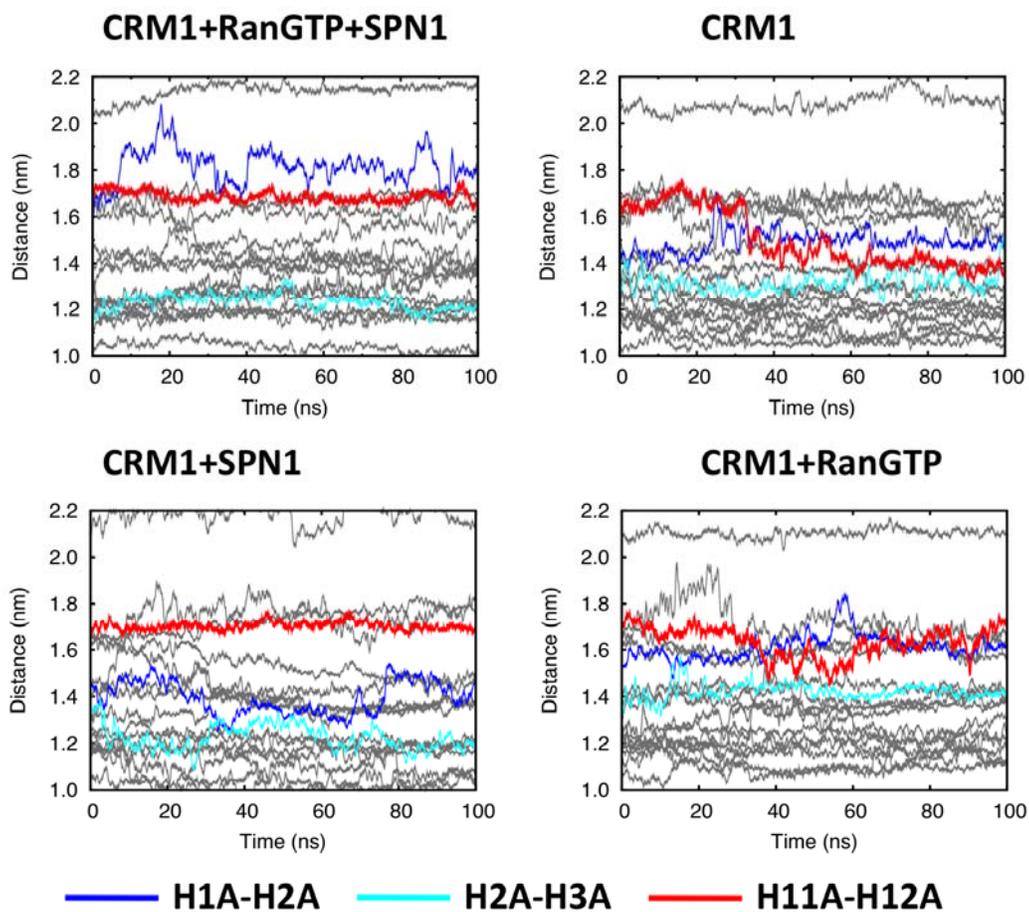
Figure S9: Crosstalk of regions involved in cargo and Ran binding correlates with changes in the distances of involved HEAT pairs and supports the idea of a sequential course of events. The distances between neighboring HEAT repeats are plotted in gray. Trajectories of distances of HEATs important for cargo or RanGTP binding are colored as indicated below the figures. Distances were determined using the centers of mass for the individual A helices of the indicated HEAT repeat pairs **(related to Figure 7)**.
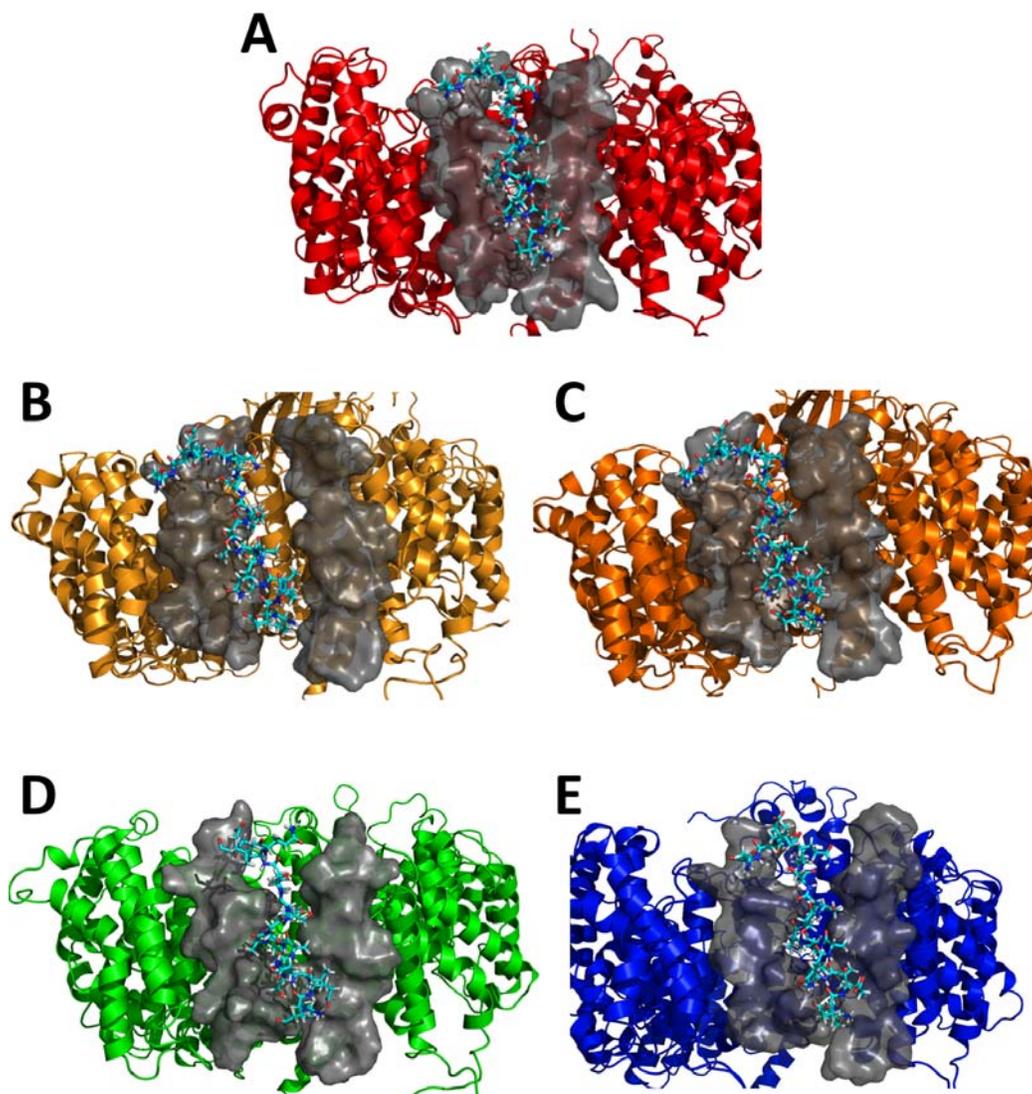
Figure S10: The distance between the centres of mass (COM) of H11A and H12A correlates with the ability to bind cargo. Randomly chosen simulation snapshots of the NES binding groove of different complexes (colour code as in Fig. 2). Surface representations of helices 11A and 12A are depicted (in gray) and the NES of SPN1 is superimposed onto the structures (in stick representation) are shown. **(A)** The NES binding-cleft conformation in free CRM1 is incompatible with NES binding. **(B)** and **(C)** RanGTP binding renders the cleft more prone for NES binding. **(D)** and **(E)** NES binding keeps the cleft in an open conformation. The colouring of CRM1 is as in **Fig S5 (related to Figure 8)**.

Free form of CRM1
4FGV — 1.32 nm
4HZK-A — 1.31 nm
4HZK-B — 1.35 nm
3VYC — 1.39 nm

RanGTP bound 3NC1 — 1.64 nm
+ RanBP1; 3M1I — 1.44 nm
+RanBP1 + SINE 4GPT — 1.60 nm
4GMX — 1.60 nm

Cargo only 3GB8 — 1.59 nm
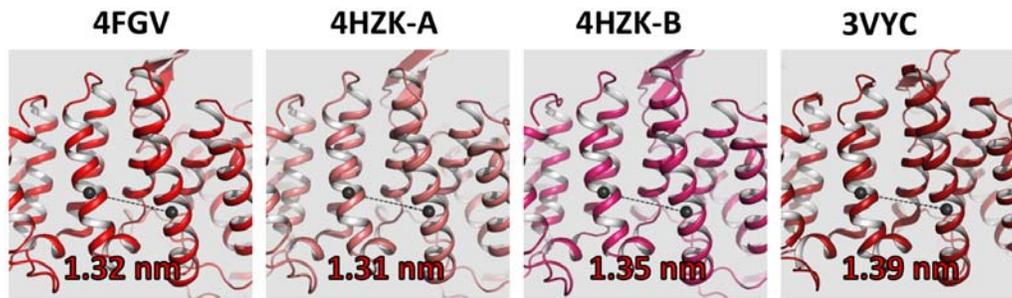
RanGTP and cargo
3GJX — 1.60 nm
3NBY — 1.63 nm
3NBZ — 1.63 nm
3NC0 — 1.64 nm
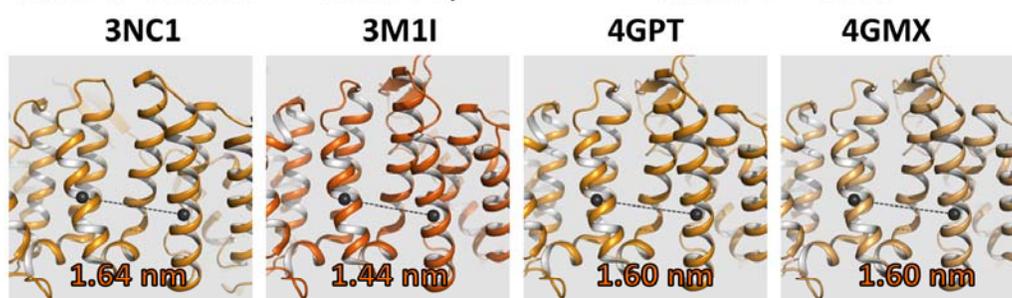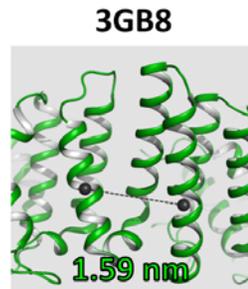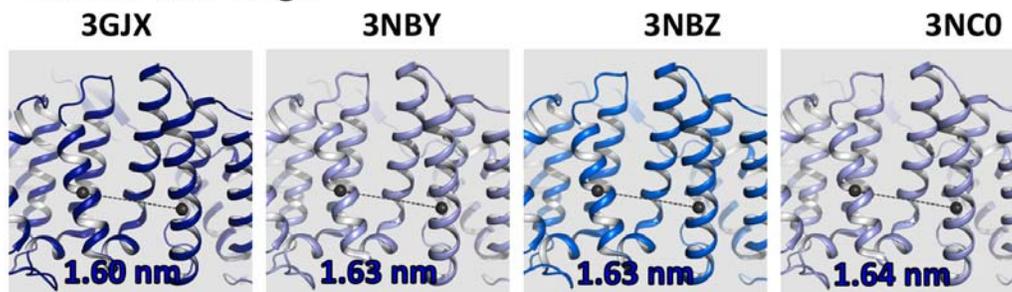
Figure S11: The size of the NES-binding cleft depends on the conformation of CRM1, which regulated by RanGTP and Cargo binding. Depicted are the available X-ray structures of CRM1 (cargoes and RanGTP have been removed for better visualisation, the respective PDBid of the protein (complex) used is shown on the left and the centers of mass for the helices 11A and 12A have been calculated and the distances between them measured. These values are used and discussed in the main body text **(related to Figure 8)**.

Supplemental Tables

Table S1: SPN1 Cap-binding domain and RanGTP are rigid structures. The RMS-deviations of known structures of SPN1 and RanGTP from different mammals where compared with respect to structural differences. The routine SUPERPOSE of the CCP4 suite was used with the standard settings with Ran residues 10-170 and SPN1 residues 100-290 for superposition **(related to figure 2)**.

PDBid (resolution in Å)	RanGTP RMSD (Å)	SPN1 (CBD only) RMSD (Å)
3GJX (2.5)	(C)* Reference	(B) Reference
	(F) 0.13	(E) 0.51
3GB8 (2.9)		(B) 0.69
1XK5 (1.8)		(A) 0.80
3NC0 (2.9)	(C) 0.317	(B) 0.57
	(F) 0.32	(E) 0.56
3NC1 (3.35)	(C) 0.44	
3NBY (3.42)	(C) 0.38	(B) 0.54
	(F) 0.38	(E) 0.57
3NBZ (2.8)	(C) 0.32	(B) 0.56
	(F) 0.32	(E) 0.58
2BKU (2.7)	(A) 0.48	
	(C) 0.48	
1IBR (2.3)	(A) 0.40	
	(C) 0.40	
1WA5 (2.0)	(A) 0.42	
3M1I	(A) 0.44	

* The Letter in the parentheses indicates the subunit in the respective crystal structure that has been used for RMSD calculation.

Table S2: Radius of gyration determined for the available crystal structures and parts thereof as well as the final models obtained from MD simulations.

Different orthologues of CRM1 were used, either from human, mouse, *Chaetomium thermophilum* or yeast (**related to figures 1 and 2**).

	R_g (nm) X-ray	PDBid	R_g (nm) SAXS	R_g (nm) MD
CRM1	**3.62**	3GJX	3.8±0.1	3.55-8
CRM1	**3.57**	3GB8	3.9±0.1	
CRM1	**3.64**	3M1I		
CRM1	**3.62**	3NC1		
CRM1	**3.81**	4FGV		
CRM1	**3.70/3.74**	4HZK		
CRM1	**3.57**	3VYC		
CRM1+RanGTP	**3.49**	3M1I		
CRM1+RanGTP	**3.48**	3NC1		
CRM1+RanGTP	**3.44**	3GJX	3.6±0.1	3.46-50
CRM1+SPN1	**3.97**	3GJX	4.1±0.4	4.10
CRM1+SPN1	**3.93**	3GB8		
CRM1+RanGTP+SPN1	**3.81**	3GJX	4.1±0.1	3.91

Rg's of the X-ray structures calculated using CRYSOL. The Mean values of Rg of the MD simulations are an average over the last 5-10 ns ± standard deviation and were calculated using g_gyrate from the GROMACS 4.0 package. For direct comparability of the data the contribution of the protein's hydration shell has to be subtracted from the SAXS data (typically around 0.2-3 nm; (Svergun et al., 1998)). CRM1 (3GJX, *Mus musculus*) lacks residues 1-11, 67-69, 1053 to end (1071), CRM1 (3NC1) lacks 1-9, 390-400, 1051-1071 and CRM1 in 3GB8 (*Homo sapiens*) lacks residues 1-52, 58-61, 92-95, 116-120, 180-186, 388-401, 430-446 and 1062-1071. From *Saccharomyces cerevisiae*, 3M1I, which is 10 residues longer, lacks residues 377-413 (which have been deleted) and 1059-1081, 3VYC is lacking residues 1-46, 57-63 and 265-270, 377-413 and 971-984. From *Chaetomium thermophilum* 4FGV lacks the first residue, whereas 4HZK lacks 1-34, 48-52, 1073-1077.

Table S3: Ratio of extended and compact conformation of mouse and human CRM1 in the free form that best fits the SAXS data. The ratio of extended versus compact conformation was optimized using the structures of CRM1 as derived from PDBid 3GJX (compact) and 4FGV (extended) in order to obtain the lowest deviation (χ value) from the measured SAXS data **(related to figures 2, S4, S5)**.

Protein (conc in mg/ml)	χ-value	Closed (%)	Open (%)
human			
11	2.96	0.578±0.004	0.421±0.004
6	2.14	0.621±0.006	0.378±0.006
3	1.37	0.663±0.011	0.336±0.011
1	0.96	0.674±0.027	0.325±0.028
mouse			
10	5.34	0.566±0.001	0.434±0.001
5	3.04	0.589±0.003	0.411±0.003
1	1.74	0.626±0.012	0.374±0.013

Supplemental Experimental Procedures

Expression and Purification

CRM1 from *Mus musculus*, RanQ69LGTP 1-180 (referred to as RanGTP in the text) as well as Snurportin1 (SPN1) both from *Homo sapiens* were expressed and purified as described (Monecke et al., 2009; Strasser et al., 2004). The CRM1-RanGTP-SPN1 complex as well as the CRM1-RanGTP-PKI-NES complex were assembled and purified as described (Guttler et al., 2010; Monecke et al., 2009). The used PKI-NES peptide was chemically synthesized (Guttler et al., 2010). Human His$_6$-CRM1 was expressed in *Escherichia coli* TG1 as described previously (Guan et al., 2000). Cells were resuspended in lysis buffer (500 mM NaCl, 50 mM HEPES/NaOH pH 7.5, 2 mM MgCl$_2$, 30 mM imidazole and 2 mM DTT) and disrupted using a microfluidizer 110S (Microfluidics). The clarified lysate (30,000 xg, 30 min, 4 °C) was loaded onto a HisTrap column (GE Healthcare) equilibrated with lysis buffer. Unbound proteins were removed by washing with 2 column volumes (CV) of lysis buffer and bound His$_6$-CRM1 was eluted with a linear gradient of elution buffer (0-100% in 6 CV) containing additionally 400 mM imidazole. In order to remove C-terminal degradation products of His$_6$-CRM1 the pooled protein fractions were desalted (50 mM NaCl, 50 mM HEPES/NaOH pH 7.5, 2 mM MgOAc and 2 mM DTT) and loaded onto a Source 30Q anion exchange column equilibrated with the same buffer. Full length His$_6$-CRM1 was eluted with a linear gradient (0-100% in 30 CV) with a high salt buffer containing 500 mM NaCl, 50 mM HEPES/NaOH pH 7.5, 2 mM MgOAc and 2 mM DTT. The protein was further purified using a Superdex S200 (26/60) gel filtration column (GE Healthcare) in a buffer containing 50 mM NaCl, 50 mM HEPES/NaOH pH 7.5, 2 mM MgOAc and 5 mM DTT. Protein containing fractions

were concentrated to 10 mg/ml and 80 µl aliquots were frozen in liquid nitrogen and stored at -80 °C.

Molecular Dynamics Simulations

MD simulations comparing wild type and acidic loop deletion mutant behavior were carried out with Gromacs 4.5 (Hess et al., 2008; Van Der Spoel et al., 2005) using the Amber99sb force field (Hornak et al., 2006) and the SPC/E water model (Berendsen et al., 1987). All other MD simulations were carried out with the GROMACS 4 program package (Van Der Spoel et al., 2005), using the OPLS-AA force field (Friesner et al., 2001; Jorgensen et al., 1996) and the TIP4P water model (Jorgensen et al., 1983). All simulations were performed in the NpT ensemble. The temperature was kept constant using the velocity rescaling method (Bussi et al., 2007) and Berendsen coupling (Berendsen et al., 1984) at $T = 300$ K with a coupling time of $\tau_T =$ 0.1 ps in all simulations investigating the role of the acidic loop and all other simulations, respectively. The pressure was coupled to a Berendsen barostat with $\tau_p =$ 1.0 ps and an isotropic compressibility of 4.5×10^{-5} bar^{-1} in the x, y, and z directions (Berendsen et al., 1984). All bonds were constrained by using the LINCS algorithm (Hess et al., 1997). An integration time step of 2 fs was used in all but the simulations comparing acidic loop deletion and wild type systems, where fast degrees of freedom consisting of angular vibrations with at least one hydrogen atom where removed (Feenstra et al., 1999) and a time step of 4 fs was employed. Lennard-Jones interactions were calculated with a cutoff of 10 Å and at 16 Å in simulations comparing acidic loop deletion mutant and wild type systems. Electrostatic interactions were calculated explicitly at a distance smaller than 10 Å; long-range electrostatic interactions were calculated by particle-mesh Ewald summation with grid

spacing of 0.12 nm and fourth-order B-spline interpolation. Structures were written out every 1 ps for subsequent analysis.

All simulations were started from crystal structures of CRM1, unbound or in complex with RanGTP, SPN1 or both based on the structure of the ternary complex (PDBid 3GJX). In the simulations with full length CRM1, residues 12 to 1055 of CRM1 were present. For the pulling simulation of the N- and C-terminal regions, two independent pulling potentials were acting on the centre of mass of residues 12-274 and 815-1055, respectively. For the acidic loop region **(Figs. S8)**, residues 344-811 were extracted from the original CRM1 simulation system. The deletion of the AL was performed as described in (Monecke et al., 2013) by removing residues Glu419 to Asp446 and replacing them by a "GGSGGSG" motif using the MODELLER 9v8 software (Sali and Blundell, 1993). Mutations were introduced using WHATIF (Vriend, 1990). Missing loops and side chains were added using the dope_loopmodel class of the MODELLER package (Sali et al., 1995). After placing the protein in an appropriately sized simulation box (see below), water molecules and sodium and chloride ions, resulting in a salt concentration of 150 mM, were added to the simulation system. Afterwards, 1000 steps of steepest descent energy minimization were performed, followed by 1 ns of equilibration with position restrained with a force constant of 1000 kJ/mol/nm^2 on all heavy protein atoms.

For the pulling simulations, CRM1 was aligned along its principal axes. Then, a rectangular box with a minimum distance of 1 nm to the protein atoms was build. This box was elongated by 5 nm, perpendicular to the plane of the CRM1 ring, in both directions. Solvent was added, followed by energy minimization and equilibration, as described. Two independent pulling potentials with a force constant of 500 kJ/mol/nm^2 moving at opposite directions perpendicular to the ring plane were

applied to specific atoms, as detailed in the results section for each individual simulation. Each pulling potential was moving at a probe velocity v. All forces presented here are the sum of the two individual forces arising from the pulling potentials.

As in Monecke et al, (Monecke et al., 2013), a common subset of Cα-atoms of residues that are identical in PDBid 4FGV and PDBid 3GJX was identified and selected and based on this subset the vector connecting the extended and compact configuration was constructed. Subsequently for force probing simulations a time dependent harmonic biasing potential was applied along this vector with a spring constant of 10 kJ/mol/nm^2 and pulling velocities of 0.5 m/s and 0.325 m/s (corresponding to 0.025 m/s and 0.0163 m/s per atom), forcing the system from the compact to the extended conformation.

For analysis of the NES-binding cleft, a subset of the Cα-atoms of the helices forming the NES-binding cleft was selected and similar to above used for construction of the difference vector between the open and closed NES-binding cleft configurations. The behavior of the NES-binding cleft was then monitored by projection of unbiased simulations of wild type and acidic loop deletion mutant onto this vector.

Small Angle X-ray Scattering

SAXS experiment and data processing

The synchrotron radiation X-ray scattering data from solutions of CRM1 alone and in complex were collected on the X33 beamline of the EMBL on the storage ring DORIS III (DESY, Hamburg, Germany) (Blanchet et al., 2012). Using a MAR345 image plate detector at a sample-detector distance of 2.7 m and a wavelength of l = 1.5 Å the range of momentum transfer $0.01 < s < 0.5$ Å^{-1} was covered (s = $4\pi \sin\theta/\lambda$, where 2θ is the scattering angle). For each construct, several solute concentrations in

the range from 1 to 10 mg/ml were measured. To monitor for the radiation damage, two successive two-minute exposures of protein solutions were compared and no significant changes were observed. The data were normalized to the intensity of the transmitted beam and radially averaged; the scattering of the buffer was subtracted and the difference curves were scaled for protein concentration. The low angle data measured at lower protein concentrations were extrapolated to infinite dilution and merged with the higher concentration data to yield the final composite scattering curves. The data processing steps were performed using the program package PRIMUS (Konarev et al., 2003).

The radius of gyration R_g was evaluated using the Guinier approximation (Guinier, 1939) assuming that at very small angles ($s < 1.3/R_g$) the intensity is represented as $I(s) = I(0) \exp(-(sR_g)^2/3)$. These parameters were also computed from the entire scattering patterns using the indirect transform package GNOM (Svergun, 1992), providing also the pair distribution function of the particle $p(r)$ and the maximum size D_{max}. The excluded volume of the hydrated particle was computed from the small angle portion of the data ($s<0.25$ Å^{-1}) using the equation (Porod, 1982):

$$V = 2\pi^2 I(0) / \int_0^\infty s^2 I_{exp}(s)ds \qquad (1)$$

Prior to this analysis an appropriate constant was subtracted from each data point to force the s^{-4} decay of the intensity at higher angles following the Porod's law (Porod, 1982) for homogeneous particles. For globular proteins, Porod (i.e. hydrated) volumes in nm^3 are about 1.6 times the MMs in kDa. Given the method dependent uncertainty in determining the protein concentration, it was difficult to estimate the molecular masses (MM) of the solutes from the forward scattering. However, these values could be assessed from the Porod volume of the individual particles in solution. For

globular proteins, the hydrated volume in $Å^3$ should be about twice the MM.

Ab initio shape determination. The "shape scattering" curve was further used to generate the low resolution ab initio shapes of CRM1 alone and in complex by the program DAMMIF (Franke and Svergun, 2009). This program represents the particle shape by an assembly of densely packed beads and employs simulated annealing to construct a compact interconnected model fitting the experimental data $I_{exp}(s)$ to minimize discrepancy:

$$\chi^2 = \frac{1}{N-1} \sum_j \left[\frac{I_{exp}(s_j) - cI_{calc}(s_j)}{\sigma(s_j)} \right]^2 \tag{2}$$

with N as the number of experimental points, c a scaling factor and $I_{calc}(s)$ and $s(s_j)$ are the calculated intensity and the experimental error at the momentum transfer s_j, respectively. Ten DAMMIF runs were performed to check the stability of solution, and the results were well superimposable with each other. These models were averaged to determine common structural features using the programs DAMAVER (Volkov and Svergun, 2003) and SUPCOMB (Svergun and Kozin, 2001). The latter program aligns two arbitrary low or high resolution models represented by ensembles of points by minimizing a dissimilarity measure called normalized spatial discrepancy (NSD). For every point (bead or atom) in the first model, the minimum value among the distances between this point and all points in the second model is found, and the same is done for the points in the second model. These distances are added and normalized against the average distances between the neighbouring points for the two models. Generally, NSD values close to unity indicate that the two models are similar. The program DAMAVER generates the average model of the set of superimposed structures and also specifies the most typical model (i.e. that having the lowest average NSD with all the other models of the set).

Ab-initio **multiphase modelling.** Low resolution shape analysis of CRM1 in complex was done using a multiphase version MONSA of the *ab-initio* program DAMMIN (Svergun, 1999) employing the range of scattering vectors up to s=0.13 Å^{-1}. The program represents the particle as a collection of M>>1 densely packed beads inside a sphere with the diameter D_{max}. Each bead can be assigned either to the solvent (index=0) or to one of distinct components in the particle (index=1,2,3 corresponding to ...). The particle is therefore represented at low resolution by NN "phases" and the structure described by a string of length M containing the phase index. Starting from a random string, simulated annealing (SA) is employed to search for a model composed by interconnected compact phases, which simultaneously fits multiple shape scattering curves from the constructs $I_k(s)$ to minimize overall discrepancy:

$$\chi^2 = \sum_k \frac{1}{N_k - 1} \sum_j \left[\frac{I_k(s_j) - c_k I_k^{calc}(s_j)}{\sigma_k(s_j)} \right]^2 \qquad (3)$$

where the index k runs over the scattering curves, N_k are the numbers of experimental points, c_k are scaling factors and $I_{calc}(s)$ and $s(s_j)$ are the intensities calculated from the subsets of the beads belonging to the appropriate phases and the experimental errors at the momentum transfer s_j, respectively.

The model of the CRM1-RanGTP-NES complex obtained by MONSA locates RanGTP directly inside the ring in an even more caged position than in the *ab-initio* model (compare **Fig. 2B** *ab-initio* versus MONSA models). This shift of Ran from a position in the vicinity to a more central position within the CRM1 ring should be taken carefully as it might be an artifact resulting solely from a conformational change of CRM1 upon binding of RanGTP. Since the curves obtained for CRM1 alone and in complex with RanGTP and the NES peptide are used for the multiphase analysis and differ in the R_g and D_{max}, RanGTP could be artificially shifted toward the

central opening of the ring due to this change of conformation. To minimize this effect on the structure determination of the ternary complex formed by addition of SPN1, the result of CRM1 alone has been omitted for the reconstruction of the ternary complex composed of CRM1, RanGTP and SPN1 **(Fig. 2D)**.

Molecular modelling. The scattering from the high-resolution models was calculated with the program CRYSOL (Svergun et al., 1995). Given the atomic coordinates, the program minimizes discrepancy in the fit to the experimental intensity by adjusting the excluded volume of the particle and the contrast of the hydration layer.

Electron microscopy preparation and image processing

Purified human CRM1 was stabilized using the GraFix protocol (Kastner et al., 2008) In brief, the sample was loaded on a 5-20% (w/v) sucrose gradient in standard buffer containing a 0-1% (v/v) glutaraldehyde and was centrifuged in a TH660 rotor (Sorvall) for 20 hours with 42,000 rpm at -10°C. Gradients were fractionated from the bottom and fixing reaction was quenched by adding 25 mM aspartate (pH 7.4). Complexes were then bound to a thin carbon film and transferred to an electron microscopic grid covered with a perforated carbon film. The bound molecules were stained with 2% (w/v) uranyl formate and air-dried. Images were recorded at a magnification of 155,000-fold on a 4k x 4k CCD camera (TVIPS GmbH) using two fold pixel binning (1.85 Å/pixel) in a Phillips CM200 FEG electron microscope (Philips/FEI) operated at 160 kV acceleration voltage.

42,108 particle images (144 x 144 pixels) were selected with the custom written software JOHNHENRY (manuscript in preparation). CTF correction was performed on the particle images based on their classified power spectra (Sander et al., 2003a). Further image processing was done in IMAGIC (van Heel et al., 1996).

CTF corrected images were coarsened by a factor of 2 to a pixel size of 3.7 Å/pixel and reference free-alignment was performed. After several rounds of exhaustive multi-reference alignment based on resampling in polar coordinates (Sander et al., 2003b) and multivariate statistical analysis (van Heel, 1984) stable class averages were obtained **(Fig. S6B, C)**. Starting models were generated using angular reconstitution facilitated by a voting algorithm (Singer et al., 2010). The handedness was determined from the solved crystal structures. Two distinct conformations could be detected in the dataset and it was split based in the cross correlation towards both models. To crosscheck those two models, angles for class averages from one model were determined using the other model (Monecke et al., 2013). By iterating this procedure models relaxed from their initial state. Models were further refined to yield a resolution of approximately 20 Å by projection matching. Resolution was judged by Fourier shell correlation employing a cut-off of 0.5 **(Fig. S7A)**. To further subdivide the compact structure into subpopulations resampling followed by 3D-MSA was performed **(Fig. S7B)** (Fischer et al., 2010).

References

Berendsen, H.J.C., Grigera, J.R., and Straatsma, T.P. (1987). The missing term in effective pair potentials. The Journal of Physical Chemistry *91*, 6269-6271.

Berendsen, H.J.C., Postma, J.P.M., Vangunsteren, W.F., Dinola, A., and Haak, J.R. (1984). Molecular-Dynamics with Coupling to an External Bath. Journal of Chemical Physics *81*, 3684-3690.

Blanchet, C.E., Zozulya, A.V., Kikhney, A.G., Franke, D., Konarev, P.V., Shang, W., Klaering, R., Robrahn, B., Hermes, C., Cipriani, F*., et al.* (2012). Instrumental setup for high-throughput small- and wide-angle solution scattering at the X33 beamline of EMBL Hamburg. Journal of Applied Crystallography *45*, 489-495.

Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. J Chem Phys *126*, 014101.

Feenstra, K.A., Hess, B., and Berendsen, H.J.C. (1999). Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. Journal of Computational Chemistry *20*, 786-798.

Fischer, N., Konevega, A.L., Wintermeyer, W., Rodnina, M.V., and Stark, H. (2010). Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. Nature *466*, 329-333.

Franke, D., and Svergun, D.I. (2009). DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. Journal of Applied Crystallography *42*, 342-346.

Friesner, R.A., Kaminski, G.A., Tirado-Rives, J., and Jorgensen, W.L. (2001). Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. Journal of Physical Chemistry B *105*, 6474-6487.

Guan, T., Kehlenbach, R.H., Schirmer, E.C., Kehlenbach, A., Fan, F., Clurman, B.E., Arnheim, N., and Gerace, L. (2000). Nup50, a nucleoplasmically oriented nucleoporin with a role in nuclear protein export. Mol Cell Biol *20*, 5619-5630.

Guinier, A. (1939). La diffraction des rayons X aux très petits angles : application a l'étude de phénomènes ultramicroscopiques. Ann. Phys. (Paris) *25*, 495-503.

Guttler, T., Madl, T., Neumann, P., Deichsel, D., Corsini, L., Monecke, T., Ficner, R., Sattler, M., and Gorlich, D. (2010). NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. Nat Struct Mol Biol *17*, 1367-1376.

Hess, B., Bekker, H., Berendsen, H.J.C., and Fraaije, J.G.E.M. (1997). LINCS: A linear constraint solver for molecular simulations. Journal of Computational Chemistry *18*, 1463-1472.

Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008). GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. Journal of Chemical Theory and Computation *4*, 435-447.

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins *65*, 712-725.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water J Chem Phys *79*, 926-935.

Jorgensen, W.L., Maxwell, D.S., and TiradoRives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. Journal of the American Chemical Society *118*, 11225-11236.

Kastner, B., Fischer, N., Golas, M.M., Sander, B., Dube, P., Boehringer, D., Hartmuth, K., Deckert, J., Hauer, F., Wolf, E.*, et al.* (2008). GraFix: sample preparation for single-particle electron cryomicroscopy. Nat Methods *5*, 53-55.

Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. Journal of Applied Crystallography *36*, 1277-1282.

Monecke, T., Guttler, T., Neumann, P., Dickmanns, A., Gorlich, D., and Ficner, R. (2009). Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. Science *324*, 1087-1091.

Monecke, T., Haselbach, D., Voß, B., A., R., Neumann, A., Thomson, E., Hurt, E., Zachariae, U., Stark, H., Grubmüller, H.*, et al.* (2013). Structural basis for cooperativity of CRM1 export complex formation. Proc Natl Acad Sci U S A *110*, 960-965.

complex formation. Proc Natl Acad Sci U S A *published ahead of print December 31, 2012, doi:10.1073/pnas.1215214110*.

Porod, G. (1982). General theory. Small-Angle X-Ray Scattering. London, Academic Press. Edited by O. Glatter and O. Kratky,, 17-51.

Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol *234*, 779-815.

Sali, A., Potterton, L., Yuan, F., van Vlijmen, H., and Karplus, M. (1995). Evaluation of comparative protein modeling by MODELLER. Proteins *23*, 318-326.

Sander, B., Golas, M.M., and Stark, H. (2003a). Automatic CTF correction for single particles based upon multivariate statistical analysis of individual power spectra. J Struct Biol *142*, 392-401.

Sander, B., Golas, M.M., and Stark, H. (2003b). Corrim-based alignment for improved speed in single-particle image processing. J Struct Biol *143*, 219-228.

Singer, A., Coifman, R.R., Sigworth, F.J., Chester, D.W., and Shkolnisky, Y. (2010). Detecting consistent common lines in cryo-EM by voting. J Struct Biol *169*, 312-322.

Strasser, A., Dickmanns, A., Schmidt, U., Penka, E., Urlaub, H., Sekine, M., Luhrmann, R., and Ficner, R. (2004). Purification, crystallization and preliminary crystallographic data of the m(3)G cap-binding domain of human snRNP import factor snurportin 1. Acta Crystallogr D Biol Crystallogr *60*, 1628-1631.

Svergun, D., Barberato, C., and Koch, M.H.J. (1995). CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. Journal of Applied Crystallography *28*, 768-773.

Svergun, D.I. (1992). Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. Journal of Applied Crystallography *25*, 495-503.

Svergun, D.I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys J *76*, 2879-2886.

Svergun, D.I., and Kozin, M.B. (2001). Automated matching of high- and low-resolution structural models. Journal of Applied Crystallography *34*, 33-41.

Svergun, D.I., Richard, S., Koch, M.H., Sayers, Z., Kuprin, S., and Zaccai, G. (1998). Protein hydration in solution: experimental observation by x-ray and neutron scattering. Proc Natl Acad Sci U S A *95*, 2267-2272.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., and Berendsen, H.J. (2005). GROMACS: fast, flexible, and free. J Comput Chem *26*, 1701-1718.

van Heel, M. (1984). Multivariate statistical classification of noisy images (randomly oriented biological macromolecules). Ultramicroscopy *13*, 165-183.

van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., and Schatz, M. (1996). A new generation of the IMAGIC image processing system. J Struct Biol *116*, 17-24.

Volkov, V.V., and Svergun, D.I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. Journal of Applied Crystallography *36*, 860-864.

Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. J Mol Graph *8*, 52-56, 29.