NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1

Thomas Güttler[1,5], Tobias Madl[2,3,5], Piotr Neumann[4,5], Danilo Deichsel[1], Lorenzo Corsini[2,3], Thomas Monecke[4], Ralf Ficner[4], Michael Sattler[2,3] & Dirk Görlich[1]

Classic nuclear export signals (NESs) confer CRM1-dependent nuclear export. Here we present crystal structures of the RanGTP–CRM1 complex alone and bound to the prototypic PKI or HIV-1 Rev NESs. These NESs differ markedly in the spacing of their key hydrophobic (Φ) residues, yet CRM1 recognizes them with the same rigid set of five Φ pockets. The different Φ spacings are compensated for by different conformations of the bound NESs: in the case of PKI, an α-helical conformation, and in the case of Rev, an extended conformation with a critical proline docking into a Φ pocket. NMR analyses of CRM1-bound and CRM1-free PKI NES suggest that CRM1 selects NES conformers that pre-exist in solution. Our data lead to a new structure-based NES consensus, and explain why NESs differ in their affinities for CRM1 and why supraphysiological NESs bind the exportin so tightly.

Nuclear export is essential for eukaryotic life. It proceeds through nuclear pore complexes (NPCs) and is typically accomplished by exportins[1,2]. The most versatile exportin is CRM1 (also called exportin 1 or Xpo1p)[3–7]. Exportins shuttle between nucleus and cytoplasm, bind cargo molecules at high RanGTP levels inside the nucleus, traverse NPCs as ternary cargo–exportin–RanGTP complexes and release their cargo upon hydrolysis of the Ran-bound GTP into the cytoplasm[4,8].

RanGTP greatly increases the affinity of exportins for their cargoes. For the exportins CAS, exportin-t and exportin 5, it is notable that RanGTP contacts not only the exportin but also the cargo[9–11]. It is likely that here the binding energy released at the cargo-RanGTP interface makes a substantial contribution to the observed cooperative effect. This mechanism is unavailable for CRM1, where RanGTP is enclosed within the toroid-like exportin molecule[12], far away from the cargo-binding site on the outside of the toroid[12,13]. It was therefore suggested that CRM1 switches between a (nuclear) high-affinity state for RanGTP and cargo and a (cytoplasmic) low-affinity conformation[12]. Cooperativity should arise if the nuclear conformation is strained and if this strain is counterbalanced by the released binding energies of the RanGTP-CRM1 and cargo-CRM1 interactions. In this scenario, RanGTP promotes cargo binding solely by stabilizing the strained nuclear conformation of CRM1. Direct evidence for such an allosteric mechanism comes from the recently solved crystal structure of a RanBP1-stabilized disassembly intermediate of a RanGTP–CRM1 complex[14].

CRM1 serves a very broad range of export substrates, and mediates, for instance, the nuclear export of small and large ribosomal subunits[15–18]. It is also required for numerous viral infection cycles.

HIV-1, for example, uses CRM1 to export its genomic RNA from nuclei[4,19,20]. The HIV-1 Rev protein is an adaptor in this process. It binds the unspliced viral RNA, recruits CRM1 and thereby triggers export of the RNA to the cytoplasm, where the RNA is assembled into the next generation of viral particles. CRM1 is also a key element in many regulatory networks. For instance, it controls the nuclear activity of protein kinase A (PKA) by expelling the PKA–PKI (PKA inhibitor) complex from nuclei[21]. In this process, PKI acts as an adaptor to CRM1. Finally, CRM1 counteracts the leakage of cytoplasmic factors into nuclei. Examples are certain translation factors[22] as well as RanBP1 (ref. 23), whose exclusive cytoplasmic localization is crucial for RanGTPase-driven nucleocytoplasmic transport[24].

The simplest CRM1-dependent nuclear export determinants are the so-called classic nuclear export signals (NESs). These are short peptides reported to comprise four spaced hydrophobic residues (denoted Φ^1–Φ^4) and to follow the consensus Φ^1-(x)$_{2–3}$-Φ^2-(x)$_{2–3}$-Φ^3-x-Φ^4, where x is an amino acid that is preferentially charged, polar or small[25]. Prototypical classic NESs are the PKI NES (**L**ALK**L**AG**L**D**I**; critical hydrophobics shown in bold), which exemplifies the most common spacing of the hydrophobic positions (Φ^1xxxΦ^2xxΦ^3xΦ^4), as well as the HIV-1 Rev NES (**L**PP**L**ER**L**T**L**), which reportedly has a Φ^1xxΦ^2xxΦ^3xΦ^4 spacing[20,21]. Even though classic NESs are also referred to as leucine-rich NESs, analyses of other export cargoes and randomization-and-selection screens showed that isoleucine, valine, methionine or phenylalanine are also permitted at the hydrophobic positions[26–28].
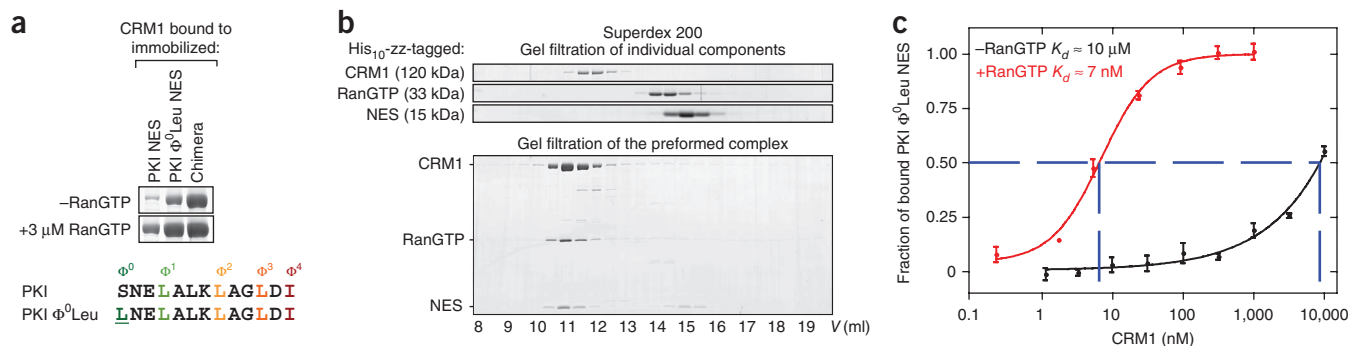
Snurportin 1 (SPN1), the nuclear import adaptor for m$_3$G-capped spliceosomal U snRNPs[29], exemplifies a CRM1 cargo with a complex export signature. Mapping experiments[30] as well as the solved

Figure 1 Enhancement of the PKI NES–CRM1 interaction by Φ^0Leu and the SPN1 fusion module. (**a**) 1 µM CRM1, pre-mixed with an *E. coli* lysate, was bound at 200 mM NaCl to beads containing either immobilized wild-type PKI NES, the Φ^0Leu variant or a chimera in which the Φ^0Leu PKI NES had been fused to SPN1 residues 15–360. The ionic strength was chosen according to the strength of the CRM1–cargo interaction: strong interactions tolerate 200 mM NaCl, whereas weak ones are evident only at 50 mM. Beads were washed and bound fractions analyzed by SDS-PAGE and Coomassie staining. (**b**) Export complex formation was analyzed by size exclusion chromatography, followed by SDS-PAGE and Coomassie staining of the obtained fractions. (**c**) CRM1 binding to the fluorescently labeled NES was quantitated at equilibrium by a fluorescence polarization assay, which exploits the fact that the rotational diffusion of the 2.7-kDa NES peptide is slowed down when assembled into an >120-kDa CRM1 complex. Apparent dissociation constants (K_d) for the CRM1-NES interaction were determined by nonlinear regression. Error bars illustrate s.d. for three independent experiments.

crystal structures of the CRM1–SPN1 (ref. 13) and the RanGTP–CRM1–SPN1 complexes[12] show that three regions of SPN1—the N terminus, the m_3G cap-binding domain and a C-terminal region—contact CRM1. The SPN1 N terminus (residues 1–14) docks into the so-called hydrophobic cleft of CRM1. In sequence, the SPN1 N terminus resembles a leucine-rich NES, but it differs from previously described classic NESs in that it contains not four but five critical hydrophobic positions (Φ^0-Φ^4) and binds CRM1 much more weakly than, for example, the PKI NES.

We now introduce a unifying structure-based NES consensus comprising five Φ positions (Φ^0, Φ^1-Φ^4), which explains differences in binding strength that were previously not understood. The general importance of a hydrophobic Φ^0 residue had so far escaped detection. However, we observed that such a residue not only causes exceptionally strong CRM1 binding of the so-called supraphysiological S1 and NS2 NESs[31,32] but also is critically important for CRM1 binding of the classic Rev NES. In fact, our data revise the Rev NES from the previously suggested $\Phi^1xx\Phi^2xx\Phi^3x\Phi^4$ spacing[25] to a completely unexpected $\Phi^0\Phi^{1\ Pro}\ x\Phi^2xx\Phi^3x\Phi^4$ pattern. It has been unresolved how CRM1 can recognize diverse NESs that differ not only in their hydrophobic Φ positions but also in the length and sequence of their inter-Φ-spacers. One possibility was that the NES-binding site flexibly adapts its conformation to match individual NES sequences. Alternatively, different Φ spacings might be compensated for by the use of additional Φ pockets[13] or by allowing Φ residues to 'slide' within the hydrophobic cleft. We have now solved the structures of RanGTP–CRM1 complexes without cargo with a PKI NES or with an HIV-1 Rev NES bound; and we found that yet another scenario holds true: the NES-binding site of RanGTP-bound CRM1 is rigid, it does not adapt to individual NES sequences and it uses the very same five binding pockets to capture the Φ residues of either the SPN1 N-terminus, PKI NES or Rev NES. These ligands bind to CRM1, however, in markedly different conformations. The CRM1-bound PKI Φ^0Leu NES is α-helical from Φ^0 to Φ^2, whereas the backbone of the shorter-spaced Rev NES is extended. Thus, our data show that the NES ligands themselves compensate for different Φ spacings by adapting their conformations to the rigid NES-binding site.

RESULTS
Crystal structure of the classic PKI NES bound to CRM1

The PKI NES (**L**ALK**L**AG**L**D**I**) contains four critical hydrophobic positions (Φ^1-Φ^4). By means of a linker, we serendipitously introduced an additional upstream leucine into a fusion construct. This Φ^0Leu improved CRM1 binding such that it became clearly detectable even in the absence of RanGTP, when CRM1 is in its low-affinity state for cargo binding (**Fig. 1a**). Successful crystallization of classic NES–CRM1–RanGTP complexes had been hampered by their apparent instability[33]. The PKI Φ^0Leu NES–CRM1–RanGTP complex was therefore a logical candidate for further crystallization trials. Indeed, we obtained a stable PKI Φ^0Leu NES–CRM1–RanGTP complex (**Fig. 1b,c**). However, even exhaustive screening failed to identify suitable crystallization conditions for the intact complex.

Because of this, we switched strategies and proceeded on the assumption that the PKI Φ^0Leu NES docks into a binding site on CRM1 that is similar to that used by the SPN1 N terminus[12,13]. We exploited the fact that the SPN1–CRM1–RanGTP complex crystallizes readily, facilitated by the m_3G cap-binding domain of SPN1 that engages in critical crystal contacts[12]. To implement this strategy, we exchanged the N terminus of SPN1 for the PKI Φ^0Leu NES, yielding a PKI Φ^0Leu-SPN1 chimera. The addition of the SPN1 module (residues 15–360) to the NES not only stabilized the NES-CRM1 interaction (**Fig. 1a**) but also allowed the NES–CRM1–RanGTP complex to form crystals that were isomorphous to the crystals of the previously solved SPN1–CRM1–RanGTP export complex (PDB ID 3GJX, ref. 12, **Table 1**). As with this previous structure, we had to screen many crystals of the chimeric NES–CRM1–Ran complex to identify one that diffracted beyond 4-Å resolution. Here also, data collection and structure solving were complicated by high mosaicity and anisotropic diffraction that resulted in low completeness and low redundancy, especially for the highest-resolution shell (**Table 1**). By using refinement strategies optimized for datasets with low completeness (see Online Methods and ref. 34), however, we obtained a reliable model at 3.42 Å resolution, with good geometry for the PKI NES bound to CRM1 (**Fig. 2**; **Supplementary Fig. 1** in **Supplementary Results**; **Table 1**). As detailed below, the NES part of the model is supported by solution

Table 1 Data collection and refinement statistics

	PKI Φ⁰Leu NES export complex	HIV-1 Rev NES export complex I	HIV-1 Rev NES export complex II	RanGTP–CRM1 complex
Data collection				
Space group	$P2_1$	$P2_1$	$P2_1$	$C222_1$
Cell dimensions				
a, b, c (Å)	72.09, 223.73, 163.06	72.66, 224.62, 164.02	73.26, 225.90, 163.98	156.80, 216.16, 123.82
α, β, γ (°)	90.00, 100.63, 90.00	90.00, 100.83, 90.00	90.00, 100.75, 90.00	90.00, 90.00, 90.00
Resolution (Å)	38.60–3.42 (3.58–3.42)[a]	39.00–2.80 (2.90–2.80)[a]	39.00–2.90 (3.00–2.90)[a]	36.50–3.35 (3.40–3.35)[a]
R_{rim}[b]	23.30 (51.90)	15.20 (64.90)	16.20 (62.90)	7.2 (53.90)
R_{pim}[b]	18.35 (34.35)	12.17 (43.13)	13.64 (39.04)	5.2 (48.70)
$I / \sigma I$	5.68 (1.96)	7.05 (1.99)	8.97 (1.96)	18.0 (2.1)
Completeness (%)	83.9 (56.3)	90.0 (92.9)	95.4 (80.9)	94.1 (66.6)
Mosaicity (°)	1.86	0.90	0.85	0.38
Redundancy	3.0 (1.4)	2.9 (2.8)	4.2 (1.8)	4.1 (2.7)
Refinement				
Resolution (Å)	38.63–3.42 (3.47–3.42)	38.90–2.80 (2.83–2.80)	39.00–2.90 (2.93–2.90)	36.52–3.35 (3.40–3.35)
No. reflections	57482	114008	110353	28804
R_{work} (%)	25.78 (29.91)	22.62 (28.28)	24.21 (31.12)	21.87 (38.11)
R_{free} (%)	31.45 (36.81)	28.50 (34.89)	29.47 (37.78)	25.39 (41.97)
No. atoms	24087	25146	25190	9754
Protein	24021	24293	24325	9721
Ligand/ion	66	68	236	33
Water	–	785	629	–
B-factors (Å²)				
Protein	152.2/140.2[c]	52.2/34.3[c]	50.3/33.9[c]	151.3/149.7[c]
Ligand/ion	137.2	32.1	51.0	129.6
Water	–	33.8	30.3	–
R.m.s. deviations				
Bond lengths (Å)	0.013	0.006	0.003	0.002
Bond angles (°)	1.545	0.938	0.747	0.538
Corr. coeff.	0.69 / 0.73[c]	0.89 / 0.91[c]	0.88 / 0.90[c]	0.91 / 0.91[c]

[a]Values in parentheses are for highest-resolution shell. [b]R_{rim}, redundancy-independent merging R-factor[44]; R_{pim}, precision-indicating merging R-factor[44]. [c]B-factor or map correlation coefficient for NES-binding site (CRM1 residues 511–574); they indicate that the quality of maps and models for the NES regions is above average of the data. See Online Methods for further details.

NMR, whereas the CRM1–RanGTP part is supported by a total of five independently refined crystal structures (PDB ID 3GJX (ref. 12) and 3NBY, 3NBZ, 3NC0 as well as 3NC1 from this study).

As expected, the structures of RanGTP, of CRM1 and of the retained SPN1 parts showed only minor deviations from our reported structure with wild-type SPN1 (ref. 12). Crucially, however, we could now see that all experimentally confirmed Φ side chains of the PKI Φ⁰Leu NES (ref. 21 and see below) docked into the exportin's hydrophobic cleft (**Fig. 2a,b**), which is also the docking site for the authentic SPN1 N terminus (refs. 12,13 and **Supplementary Fig. 1**). The Φ²-Φ³ spacers of PKI NES and SPN1 N terminus differ in length. Yet, equivalent Φ side chains of the two NESs dock at identical positions of the CRM1 molecule (**Fig. 2c**). The shorter Φ²-Φ³ spacer of the PKI Φ⁰Leu NES is compensated for by an earlier break in the α-helix, which allows Φ³Leu11 and Φ⁴Ile13 to reach into their corresponding pockets of the hydrophobic cleft (**Fig. 2b,c**). The docking of Φ⁰Leu1 of the PKI NES into the Φ⁰ pocket explains why this additional leucine confers a stronger CRM1 binding to the PKI NES (**Fig. 1a**).

Point mutations in CRM1 that block NES binding

One might argue that the addition of the SPN1 module forced the NES into the hydrophobic cleft and that an NES without the SPN1 fusion partner would bind to a different site and/or in a different mode. We therefore probed both the atomic structure of a non-fused NES in its CRM1-bound state and the location of the NES on the exportin

molecule, using independent experimental strategies. We found that the PKI NES-SPN1 chimera competitively blocks CRM1 binding of the non-fused PKI NES (**Fig. 2d**), supporting the assumption that fused and non-fused NES contact identical sites on CRM1.

Furthermore, we mutated Ala541 of CRM1 (located at the base of the Φ³-binding pocket, refs. 12,13 and **Fig. 2a**) to a lysine that we expected to clash with the CRM1-bound PKI NES (**Supplementary Fig. 1b**). Indeed, the A541K mutation excluded binding of all tested PKI NES variants (**Fig. 2e**). Likewise, NES binding was impaired when Cys528, the residue covalently modified by the CRM1 inhibitor leptomycin B[35], was mutated. Although alanine, threonine and valine were still well tolerated at this position, NES binding was diminished in the C528S and clearly reduced in the C528W mutant (**Supplementary Fig. 5d**).

Considered together, these experiments provide, independently of the above-described crystal structure, evidence for the hydrophobic cleft being the binding site not only for the SPN1 N terminus but also for the classic PKI NES.

NMR structure of the CRM1-bound PKI Φ⁰Leu NES

To elucidate the structure of a CRM1-bound PKI Φ⁰Leu NES without the SPN1 fusion module, we performed a series of NMR experiments. We prepared several isotope-labeled variants of the NES peptide and assembled them into NES–CRM1–RanGTP complexes. Initial attempts were hampered by poor signal-to-noise in the NMR spectra, probably because of the size of this complex

Figure 2 Crystal structure of a classic PKI-type NES bound to CRM1. The structure is derived from the complex containing GTP-Ran (residues 5–180), mouse CRM1 (full length) and the PKI Φ^0Leu NES fused to the residues 15–360 of SPN1. (**a**) Picture shows the surface of the CRM1 NES-binding site (hydrophobic cleft). Blue denotes hydrophilic and white denotes hydrophobic areas. Cys528 is colored yellow, Ala541 orange. (**b**) Same structure as in **a**, but with the PKI NES (chain E) shown. The NES backbone is colored in orange. Helical regions were defined by PROMOTIF[42]. Side chains of Φ^0-Φ^4 (color-coded as in **Fig. 1a**) are presented as sticks and surfaces. See also **Supplementary Figure 1**. (**c**) Top, sequence alignment of SPN1 N terminus and PKI Φ^0Leu NES. Bottom, overlay of NES and SPN1 N terminus in their CRM1-bound forms, derived from a Cα alignment of the corresponding hydrophobic clefts. The NES is colored as above, the SPN1 N terminus (PDB-ID 3GJX, ref. 12) is colored gray. Φ side chains of the two ligands are superimposable, despite their different Φ^2-Φ^3 spacings. Both NES chains of the asymmetric unit (B and E) are shown, to illustrate allowed flexibility. (**d**) CRM1 binding to immobilized PKI Φ^0Leu NES was competitively inhibited by wild-type SPN1 or the PKI NES-SPN1 chimera, but not by the SPN1 ΔN module lacking the NES-like N terminus. (**e**) The A541K CRM1 mutation disabled binding to the indicated export ligands but not to phenyl-Sepharose, a mimic of FG repeats[43] and a probe for non-selective side effects of the mutation. Binding to NES peptides was at 100 mM NaCl, binding to SPN1 chimeras and phenyl-Sepharose was at 200 mM. See also **Supplementary Figure 5**.
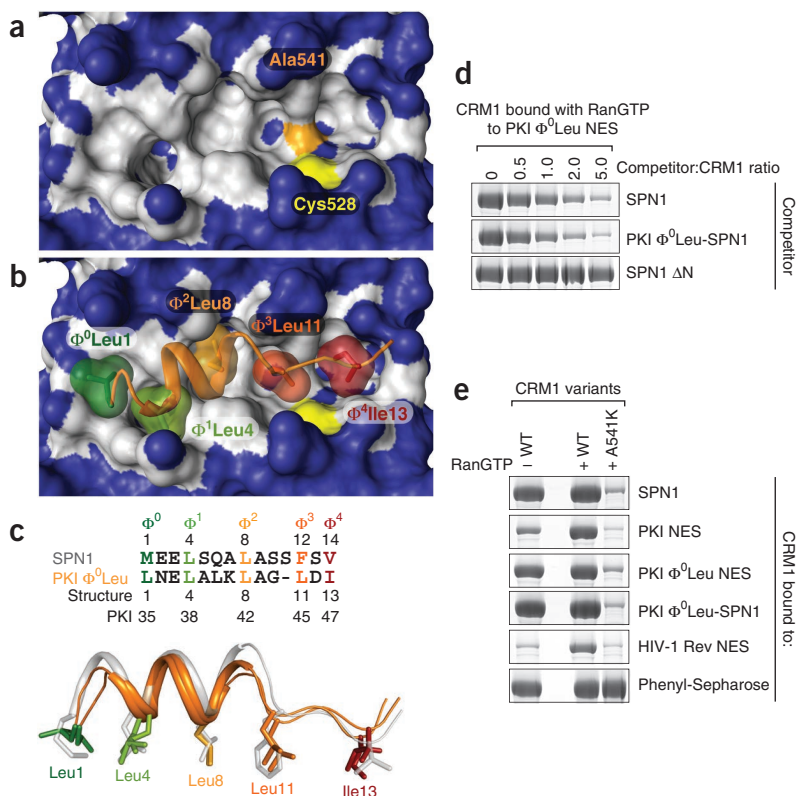
(≈150 kDa) and the extensive broadening of the NES NMR signals by relaxation with protons of the exportin. Complexes with uniformly deuterated CRM1, however, yielded high-quality NMR spectra and allowed us, in a two-step protocol, to analyze the structure of the bound NES and its docking to the exportin (**Fig. 3** and **Supplementary Fig. 4**).

First, we determined the structure of the bound NES from NOE-derived intramolecular (intrapeptide) distance restraints and chemical shift–derived restraints for the backbone torsion angles (**Table 2**). In the second step, we docked the NES onto the CRM1–RanGTP complex, which we had found to undergo surprisingly little structural change upon export ligand binding (see below, **Supplementary Fig. 3**). For docking, the following additional restraints were used: (i) NOE cross-peaks between the sulfhydryl proton of CRM1-Cys528 and NES protons from the methyl groups of Φ^3Leu11 and Φ^4Ile13 as well as from the backbone amide of NES-Asp12 (**Fig. 3c** and **Supplementary Fig. 5c**) and (ii) ambiguous distance restraints[36,37] between CRM1 amide protons and buried PKI NES methyl protons that were derived from NOEs detected in ^{13}C-edited NOESY-HMQC spectra. The docking model we obtained was further refined based on the solvent accessibility of methyl groups from the NES side chains, as probed by paramagnetic relaxation enhancement experiments (ref. 38 and **Supplementary Fig. 4b**).

The final ensemble of lowest-energy structures is well converged (pairwise backbone coordinate r.m.s. deviation of 0.9 Å) and does not show any restraint violations. Residues Φ^0Leu^1-Gly10 of the NES form an amphipathic α-helix, whereas Φ^3Leu11-Φ^4Ile13 adopt an extended conformation (for helix definition, see **Fig. 3d**). Residues N-terminal of Φ^0 and C-terminal of Φ^4 are highly flexible in solution and do not interact with CRM1. This is also indicated by the similarity of the chemical shifts of the corresponding residues in the free and

bound forms of the NES peptide (**Supplementary Fig. 4a**). Notably, the solution structure of the CRM1-bound PKI NES is fully consistent with the crystal structure of the PKI NES fused to SPN1 residues 15–360. This suggests that our crystallographic approach, using chimeras, is generally applicable for elucidating the atomic details of CRM1 binding to different NESs.

Solution conformation of the unbound NES peptide

NMR experiments also allowed us to probe the conformation of the NES peptide before exportin binding (**Fig. 3d**). The spectra revealed that peptide molecules with an α-helical conformation between residues Φ^0Leu1-Gly10, corresponding to the CRM1-bound conformation, pre-exist in solution and that this conformation is already populated to ≈31%. This suggests that recruitment of cargo to CRM1 may not solely rely on a (CRM1-) induced fit mechanism. Instead, it appears that CRM1 selects those NES molecules from a pre-existing equilibrium that are already in a 'bound conformation'. We assume that a highly populated bound conformation of the free peptide is a prerequisite for high-affinity binding to CRM1.

A consensus for optimal NES binding to CRM1

NESs vary considerably in their affinities for CRM1, but the sequence features underlying this variability are still poorly understood. We therefore addressed this issue systematically, starting with a characterization of the Φ^0 position of the PKI NES. We found that a Φ^0xxΦ^1 spacing is optimal (**Fig. 4a**), and that Φ^0Ile, Φ^0Val or Φ^0Met confer an even stronger CRM1 binding than a Φ^0Leu (see below, **Fig. 4**). In the light of the previously considered active Φ residues[26–28], it was quite surprising that even Φ^0Ala, Φ^0Pro or Φ^0Tyr enhanced CRM1 binding compared to the PKI wild-type Φ^0Ser variant.

Figure 3 Solution NMR structure of the CRM1-bound PKI Φ^0Leu NES. (**a**) Overview of the NES–CRM1–RanGTP complex. RanGTP (red) and CRM1 (blue) are shown as a surface representation. The NES-binding site is colored as in **Figure 2**. The ensemble of NMR structures of the PKI Φ^0Leu NES is shown in orange, docked onto the CRM1–RanGTP complex. (**b**) Close-up of the CRM1–PKI Φ^0Leu NES interaction. Φ side-chain carbons are colored according to **Figure 1a**; the non-interacting, solvent-accessible Leu6 is colored gray. The hydrophobic cleft, formed by HEAT repeats 11 and 12, is depicted as a blue cartoon. (**c**) Strips from a ^{13}C-edited NOESY-HMQC spectrum displaying intermolecular NOE cross-peaks between CRM1-Cys528 Hγ and methyl protons of Φ^3Leu11-Φ^4Ile13 from the NES. The ^{13}C chemical shifts for the Φ^3 and Φ^4 δ methyl groups are 26.38 and 16.32 p.p.m., respectively. See also **Supplementary Figure 5**. (**d**) Top, cartoon representation of the CRM1-bound NES as defined by the NMR secondary chemical shifts ($\Delta\delta$). Middle, difference of ^{13}Cα and ^{13}Cβ secondary chemical shifts for the residues of the CRM1-bound NES. The average over residues 1–10 (red dashed line) serves as a reference point to define the 100% bound population. Bottom, the secondary chemical shift analysis indicates that free NES molecules already populate the CRM1-bound α-helical conformation to ≈31%. See also **Supplementary Figure 4**.
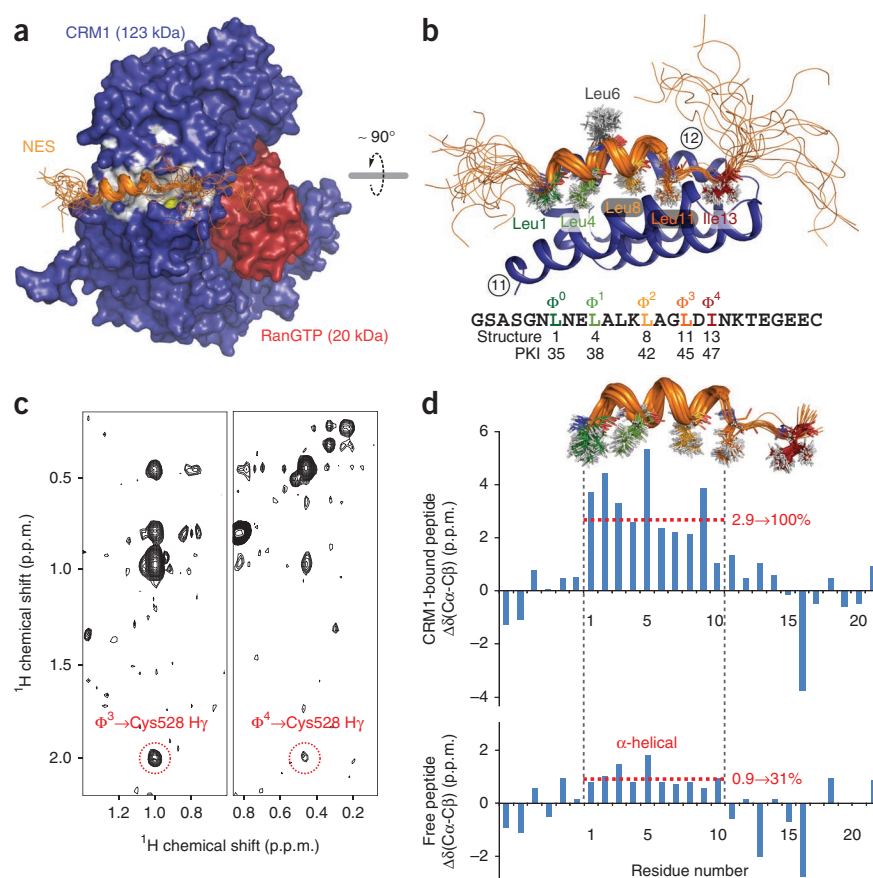
It is curious that hydrophobic Φ^0 positions have not yet been described for high-affinity NESs such as the supraphysiological S1 and NS2 NESs. We therefore re-evaluated this issue. S1 was identified through a phage display selection for 15-mer peptides that show strong CRM1 binding even in the absence of RanGTP[31]. For the sequence previously considered (**L**ARL**F**SA**LSV**, ref. 31), which lacks a Φ^0 position, we could not reproduce the high affinity that was reported. We noticed, however, a seemingly perfect Φ^0Val in the sequence stretch N-terminal to the reported 4-Φ NES. Indeed, including this valine dramatically improved CRM1 binding, whereas a Φ^0Val→Ser mutation abolished the effect (**Fig. 4b**). The NS2 NES from minute virus of mice was also published without considering a hydrophobic Φ^0 position[32], and again, we found an upstream valine from the original sequence context to be required for very tight CRM1 binding (**Fig. 4b**). We thus conclude that the NES consensus comprises not four but five hydrophobic positions (Φ^0, Φ^1–Φ^4) and that high-affinity NESs are characterized by a hydrophobic Φ^0 residue.

Our analysis of the An3 NES[39] documents another critical feature of NESs with high affinity for CRM1. This NES contains an optimally spaced Φ^0Val. Here, however, a change to serine had surprisingly little effect on CRM1 binding (**Fig. 4b**). This correlated with another difference: whereas the Φ^0-Φ^1 spacer is acidic in the SPN1 N terminus as well as in the PKI Φ^0Leu, NS2 and S1 NESs, it consists of two neutral residues in the An3 NES. Indeed, changing these neutral residues to acidic ones greatly improved the CRM1 interaction. In such a sequence context, the Φ^0Val→Ser mutant had a striking negative effect on CRM1 binding (**Fig. 4b**). Acidic residues from the Φ^0-Φ^1 spacer engage in electrostatic interactions with Lys522 and Lys560 of CRM1 (refs. 12,13). This explains their positive effect on exportin binding. Apparently, they also position the Φ^0 residue within the hydrophobic cleft.

To complete our systematic analysis, we probed, in the context of the PKI Φ^0Leu NES, all five Φ-binding pockets for their side chain preferences. The result of our survey is summarized in **Figure 4c** (see also **Supplementary Fig. 6** and discussion below). Leucine is well accepted by all five Φ pockets but only in Φ^1 is it clearly preferred over all other hydrophobics. Notably, the Φ^2 pocket has the preference phenylalanine ≈ methionine > leucine, and it even accepts tyrosine. These data suggest that the five Φ-binding pockets are non-equivalent and that the strength of NES binding to CRM1 can be adjusted at each Φ position by the choice of the hydrophobic residue.

Rational design of a supraphysiological NES
With respect to maximum CRM1 binding, our PKI Φ^0Leu NES was still suboptimal at Φ^0 (leucine instead of isoleucine), Φ^2 (leucine instead of phenylalanine) and Φ^4 (isoleucine instead of leucine). We therefore constructed a fully Φ-optimized version and found that this Super-PKI-1 NES bound CRM1 nearly as strongly as the supraphysiological S1 NES (**Fig. 5a**).

Around their Φ^0 positions, the NS2 and S1 NESs are more negatively charged than the PKI NES (**Fig. 5a**). This is noteworthy because a negatively charged Φ^0-Φ^1 spacer confers stronger CRM1 binding to the An3 NES (**Fig. 4b**). We therefore gradually adjusted the original NΦ^0NEΦ^1 sequence of the Super-PKI-1 NES to more acidic patterns. The effect was striking. The final Super-PKI-3 NES had a DΦ^0DEΦ^1 pattern and a CRM1 affinity that clearly exceeded the affinity of S1 and matched that of the NS2 NES (**Fig. 5a**). Structurally, this effect can be attributed to additional electrostatic interactions of the acidic Φ^0 neighbors with Lys514 and Lys560 of CRM1 (**Supplementary Figs. 1b** and **2b**).

Supraphysiological NESs naturally occur in certain viral proteins[32]. They strongly compete CRM1-mediated nuclear export and thus

might even represent a viral strategy to disable a critical element of the host's cellular infrastructure. Our data suggest that such high CRM1 affinity results if an NES sequence closely matches our consensus for PKI-type NESs with five Φ positions and an acidic Φ^0 context (see discussion below).

NES attenuation

Cellular NESs are apparently not optimized to maximum CRM1 binding[25,32] because too strong an NES–CRM1 interaction would cause the Ran-free form of CRM1 to stably bind the NES in the cytoplasm and thus to reimport such cargo into the nucleus. A typical cellular NES will therefore be attenuated, compared to the strongest possible ligand of the hydrophobic cleft. An extreme example of such attenuation is the NES-like N terminus of SPN1: its Φ residues occupy virtually

Table 2 NMR and refinement statistics for the CRM1-bound PKI Φ^0Leu NES

	PKI Φ^0Leu NES bound to CRM1–Ran complex
NMR restraints	
Distance restraints[a]	
Total NOE	37
Intra-residue	0
Inter-residue (NES)	
Sequential ($|i − j| = 1$)	0
Medium-range ($|i − j| \leq 4$)	9
Long-range ($|i − j| > 4$)	10
Intermolecular (NES-CRM1)	
Unambiguous	3
Ambiguous	15
Dihedral angle restraints (NES)[b]	
ϕ	9
ψ	10
Solvent PREs (NES)	11
Structure statistics	
Violations (mean ± s.d.)	
Distance restraints (Å)	0.331 ± 0.106
Dihedral angle restraints (°)	0.000 ± 0.000
Max. dihedral angle violation (°)	0.00
Max. distance restraint violation (Å)	0.13
Deviations from idealized geometry	
Bond lengths (Å)	0.00344
Bond angles (°)	0.57480
Impropers (°)	1.46510
Average pairwise r.m.s. deviation (Å)	
Backbone	0.89 ± 0.25
Heavy	1.62 ± 0.22
Ramachandran plot (PROCHECK)[c]	
Most favored regions	88.3%
Allowed regions	8.3%
Generously allowed regions	3.3%
Disallowed regions	0.0%

Statistics are given for the 10 lowest-energy structures out of 1,000 calculated for the PKI NES docking to CRM1. These structures were obtained after scoring against the solvent PRE data as described in the Online Methods. The CNS E_{repel} function was used to simulate van der Waals interactions with an energy constant of 25 kcal mol^{-1} Å^{-4} using van der Waals radii as defined by PROLSQ[45]. r.m.s. deviation and PROCHECK values apply for residues 1–11 and 1–13, for the PKI NES peptide and the docking to CRM1, respectively.
[a]Distance restraints were used with a soft square-well potential applying an energy constant of 50 kcal mol^{-1} Å^{-2}. [b]Torsion angle restraints derived from TALOS[46] were applied to ϕ, ψ backbone angles using energy constants of 200 kcal mol^{-1} rad^{-2}. [c]PROCHECK[47] was used to determine the quality of the structure.

the same positions within the hydrophobic cleft as the equivalent residues of the PKI Φ^0Leu NES (**Fig. 2c**), yet the SPN1 N terminus binds CRM1 ≈100 times more weakly. Our analysis (**Fig. 5b**) revealed that the SPN1 N terminus is down tuned in at least three regards as compared to the PKI NES, namely by containing suboptimal Φ^3Phe and Φ^4Val residues and by having a Φ^2xxxΦ^3 spacing instead of the optimal Φ^2xxΦ^3 spacing. Possibly the Φ^2xxxΦ^3 spacing results in tension within the hydrophobic cleft and/or the NES peptide.

Unconventionally spaced NESs derived from the 5-Φ consensus

The allowed variability of the Φ^2-Φ^3 spacing is also notable because it creates a greater diversity of functional NESs. An even greater diversity arises from the observation that not all hydrophobic positions are necessarily essential, especially if the other Φ positions are sufficiently strong. In the context of the Φ-optimized Super-PKI-1 NES (**Fig. 5a**), for example, each of the Φ residues can be exchanged for alanine without abolishing CRM1 binding and export activity (**Fig. 6a,b**). These Φ→Ala mutations lead to 4-Φ NESs with very unusual hydrophobic spacings, such as ΦxxΦxxxΦxxxxΦ, in the case of a Φ^3→Ala mutant. Likewise, the Φ^4→Ala mutation generates a peptide with a ΦxxΦxxxΦxxΦ spacing that corresponds to an exotic "class 3 NES"[28]. These considerations suggest that the number of peptide sequences with CRM1-dependent NES activity is far larger than was previously thought. The data also show that an optimized 5-Φ NES tolerates losses at hydrophobic positions considerably better than the wild-type 4-Φ version of the PKI NES[21]. This tolerance certainly reaches its limits, however, when the peptide contains residues that are incompatible with the required NES conformation (such as proline in the Φ^1-Φ^3 helical region) or cause steric exclusion from the central Φ^1-Φ^3 pockets.

"4-Φ" Rev NES docks in unanticipated manner into all 5 Φ pockets

The HIV-1 Rev NES was initially identified as the Rev-activation domain[19] and later shown to mediate nuclear export of the genomic HIV-1 RNA[20] in a CRM1-dependent manner[4]. Its recognition by the exportin is notable because in this case the postulated Φ^1 is separated from Φ^2 by a double-proline spacer that should break any helicity. Therefore, its mode of CRM1 binding cannot be explained by the SPN1-PKI paradigm described above.

To investigate this problem, we used the chimera strategy, crystallized the corresponding Rev NES–CRM1–RanGTP complex and solved its structure. Considering that the Rev NES had so far been assumed to contain four Φ residues, all leucines, the structure revealed a fully unexpected mode of binding, namely that the Rev NES docks with five residues into all five Φ pockets of CRM1 (**Fig. 7a–c** and **Supplementary Fig. 2**). The structure shows that the traditional alignment between the Rev and PKI-type NESs was correct only for the stretch from Φ^2Leu9 to Φ^4Leu14—that is, in the region that is indeed analogous between the two NES types (**Fig. 7c**). The critical Leu6, however, is not the Φ^1 residue but docks into the Φ^0 pocket. Unexpectedly, the Φ^1 pocket accommodates the following Pro7. The spacer to Φ^2Leu9 does not comprise three residues, as in the PKI NES or the SPN1 N terminus, but rather only a single proline. These shorter spacers between Φ^0 and Φ^2 are compensated for by an extended structure instead of an α-helical one. Notably, this allows the five newly defined Φ residues of the Rev NES to occupy positions in three-dimensional (3D) space that are nearly identical to those of the equivalent hydrophobic residues of SPN1 or PKI (**Fig. 7c**; compare **Fig. 2b,c** and **Fig. 7a–c** and **Supplementary Figs. 1** and **2**).

Our Rev NES–CRM1 structure is supported by biochemical data. The notion that the Rev NES docks into the hydrophobic cleft, for
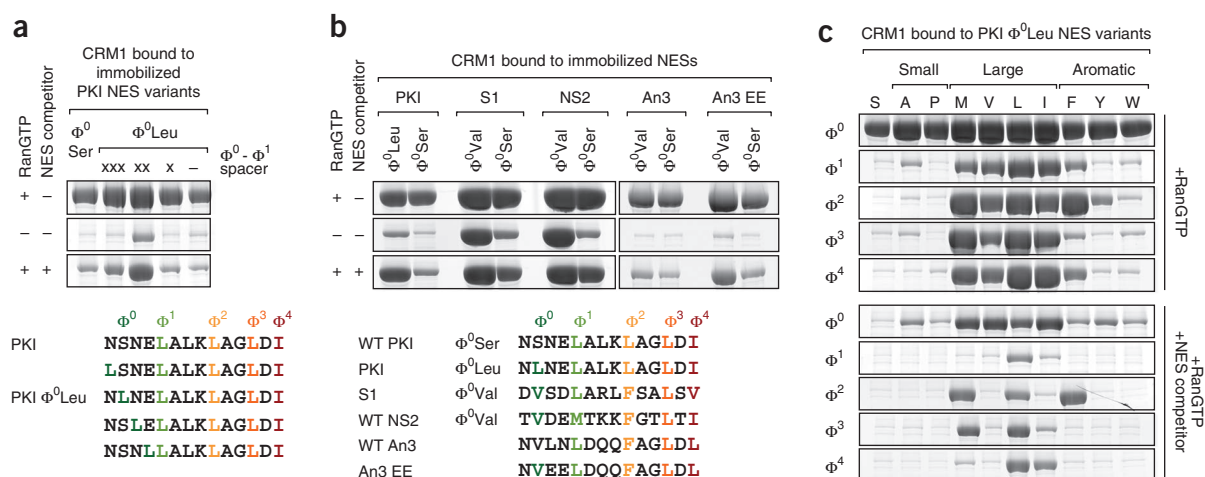
Figure 4 Φ preferences in PKI-type NESs. (**a**) 1 μM CRM1 was bound at 3 μM RanGTP and 100 mM NaCl to 1 μM immobilized NESs. At this NES concentration, the assay barely discriminates between low and medium nanomolar dissociation constants. Binding was therefore performed also in the absence of Ran ("−RanGTP") or with RanGTP and 2 μM of competing free PKI NES ("+RanGTP+NES competitor"). The $\Phi^{0\,Leu}xx\Phi^{1\,Leu}$-spaced NES recruited CRM1 most efficiently. (**b**) Assay was performed as in **a** to probe the impact of hydrophobic Φ^0 residues on CRM1 binding of the supraphysiological S1 and NS2 NESs, as well as of the An3 NES. (**c**) Starting with the PKI Φ^0Leu NES, all five Φ positions were systematically mutated to indicated residues. These mutants were tested for CRM1 binding as in **a**, using the conditions "+RanGTP" and "+RanGTP+NES competitor". The control without Ran is shown in **Supplementary Figure 6**.

example, is fully consistent with the inhibition of the Rev NES–CRM1 interaction by competing PKI NES or SPN1 (data not shown), by leptomycin B[4], by the A541K and C528W mutations in CRM1 (**Fig. 2e** and **Supplementary Fig. 5d**) and by the I521A, L525A, F561A, F572A quadruple mutation in CRM1 (ref. 13).

Mutation of the critical Φ leucines of the Rev NES to serine reduced CRM1 binding to background levels (**Fig. 7d**). The very strong effect of the Φ^0Leu6→Ser mutation is particularly notable because it classifies the Rev NES as an export signal for which a hydrophobic Φ^0 position not only tightens the interaction with CRM1 but appears essential for CRM1 binding. The critical Φ^1Pro7 turned out to be very intolerant toward mutations (**Fig. 7d**). Changes to glycine or alanine abolished CRM1 binding. Even the Φ^1Pro→Leu mutant showed strongly reduced CRM1 binding, emphasizing that Pro7 is not just a hydrophobic ligand but also favors the special backbone conformation that positions other Φ residues within the hydrophobic cleft. Indeed, this mutant reveals an unexpectedly strong interdependence of Φ preference and backbone conformation: Φ^1Pro is optimal in the Rev NES context but kills the PKI NES, whereas Φ^1Leu disables the Rev NES but is optimal for PKI (summarized in **Fig. 8**). The change of the spacer Pro8 to glycine also impaired the interaction with CRM1, whereas a change to alanine had no negative effect (**Fig. 7d**). This probably reflects a negative entropic impact of the very flexible glycine residue, and it emphasizes once again the great influence of the NES backbone conformation on recognition by CRM1.

**The NES-binding site does not adapt to individual cargoes**
We also obtained a crystal structure of the binary RanGTP–CRM1 complex. It lacked a cargo molecule and crystallized in a different space group than the RanGTP–CRM1–cargo complexes (**Table 1** and **Supplementary Fig. 3**). Nevertheless, Ran and CRM1 showed only negligible deviations from the structures containing cargo. In particular, the backbones of the NES-binding sites of all four RanGTP–CRM1 complexes (containing SPN1, the PKI NES or Rev NES fusions, or lacking cargo) are essentially indistinguishable (**Fig. 7e**). Differences appear merely restricted to conformations of the

side chains. Thus, the NES-binding site in RanGTP-bound CRM1 is rigid: that is, it does not adapt to individual NES sequences. Instead, the NES ligands themselves compensate for different Φ spacings by adapting their conformations to the rigid NES-binding site.

The fact that RanGTP alone suffices to force the CRM1 molecule into the nuclear conformation emphasizes that Ran is the master regulator of the cargo-binding site. It further predicts that
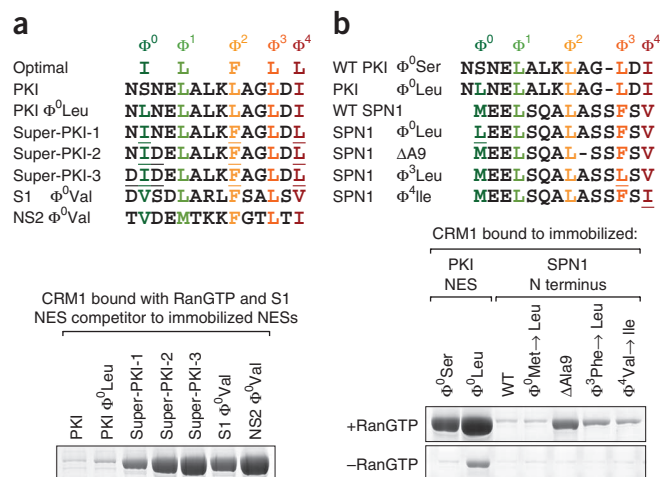
Figure 5 Supraphysiological and attenuated NESs. (**a**) Conversion of the PKI NES into a supraphysiological NES. To maximize CRM1 affinity of the PKI Φ^0Leu NES, all Φ positions were changed to preferred residues, yielding the Super-PKI-1 NES. Subsequently, aspartate residues were introduced next to Φ^0. The final Super-PKI-3 NES matched the CRM1 affinity of NS2 and outperformed the S1 NES. Binding assay was performed as in **Figure 4a**, using 2 μM of the supraphysiological S1 NES as a competitor. (**b**) Mutants that render the weakly CRM1-binding SPN1 N terminus more similar to the PKI NES were tested for CRM1 binding at 50 mM NaCl. The greatest gain in CRM1 binding was observed for the ΔAla9 mutant that adjusted the Φ^2xxxΦ^3 to the PKI-like Φ^2xxΦ^3 spacing. The Φ^3Phe→Leu and Φ^4Val→Ile mutations introduced more optimal Φ residues for CRM1 binding.
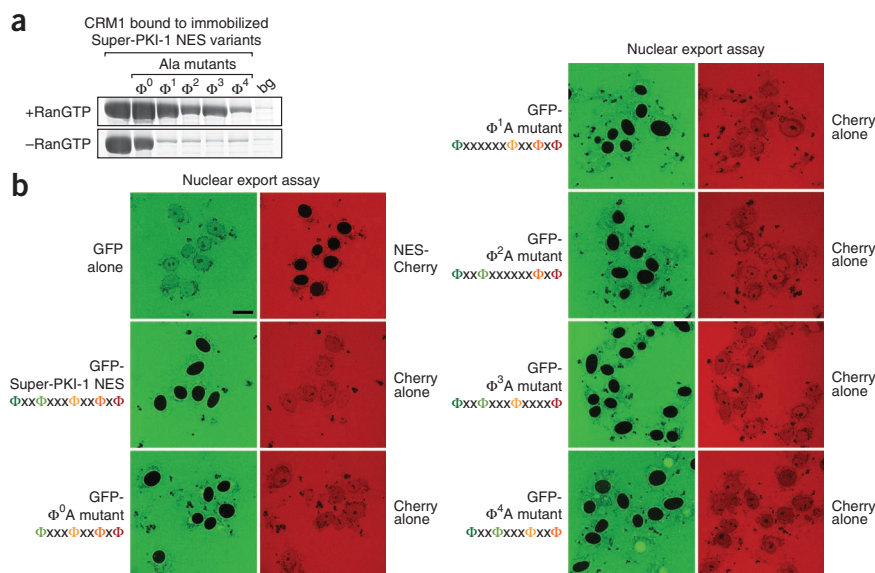
Figure 6 Export activity of NESs with unconventionally spaced Φ residues. (**a**) Starting from the Super-PKI-1 NES, each of the Φ positions was mutated to alanine, generating 4-Φ NESs with unconventional Φ spacing. The CRM1 binding assay was at 100 mM NaCl. The term "bg" denotes background binding to empty beads. (**b**) PKI NES variants from **a** were fused to GFP and assayed for nuclear export activity. The readout was steady-state exclusion from HeLa cell nuclei incubated in *Xenopus* egg extract. The assay exploits the fact that these small fusion proteins passively diffuse into nuclei. Nuclear exclusion results only if the reporter is fused to a functional NES. Confocal images were taken 3 h after addition of the indicated fluorescent proteins. As internal controls, we used mCherry (which enters nuclei but lacks CRM1 binding) or a PKI Φ^0Leu NES-mCherry fusion. Scale bar, 20 μm.

high-affinity state of the hydrophobic cleft has a very defined conformation that does not adapt to different NESs. We also found no indication of a sliding of Φ residues within the hydrophobic cleft or of alternative Φ-binding sites. Instead, we discovered that different Φ spacings are compensated for by allowing NESs to dock their Φ residues with an α-helical, near α-helical or an extended backbone conformation into the Φ pockets. This, along with the facts that single Φ residues are dispensable (provided the others are close to optimal) and that each Φ-binding pocket accepts various hydrophobic residues, accounts for the observed very wide range of specifically recognized NESs.

NESs of the PKI class are characterized by a $\Phi^1 xxx\Phi^2$ spacing[25]. The traditional NES consensus, however, also included a deviating $\Phi^1 xx\Phi^2$

the conformation of CRM1 in export complexes with other cargoes will be very similar to the structures we observed.
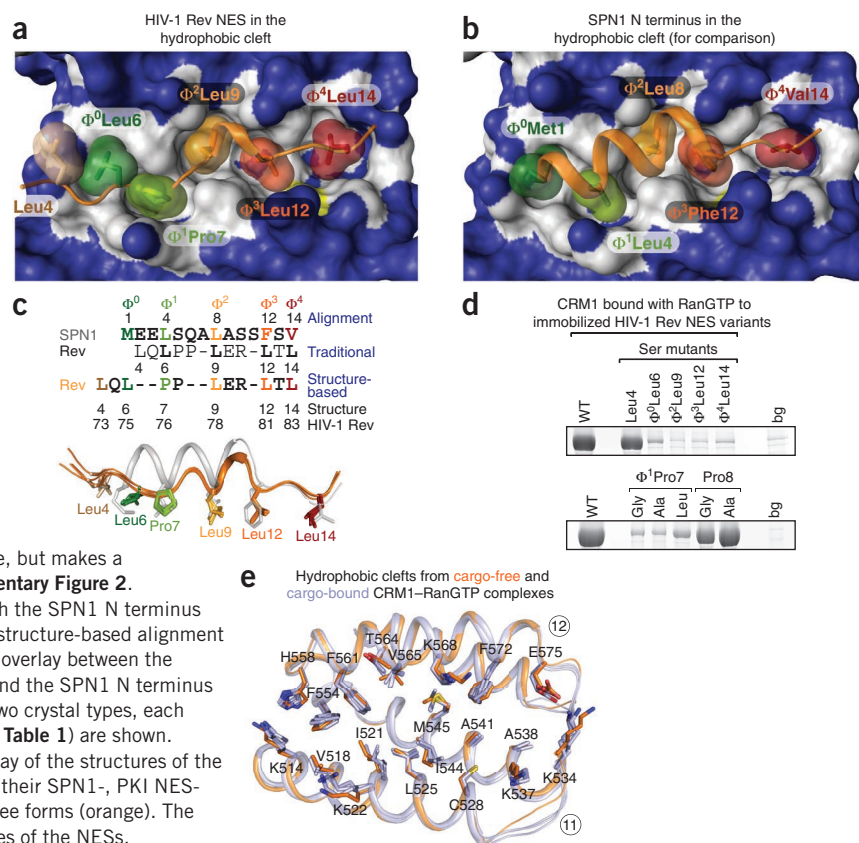
DISCUSSION

The great diversity of functional leucine-rich NESs poses an intriguing question, namely, how a single receptor, CRM1, can recognize all of these different peptides. In particular, it was difficult to comprehend how the exportin can cope with different spacings of the key Φ residues. To address this recognition problem, we developed and validated a general crystallographic approach (based on NES-SPN1 chimeras) that allowed us to elucidate the structures of NES–CRM1–RanGTP complexes. We applied this approach to the export signals of the protein kinase A inhibitor PKI and the HIV-1 Rev protein, which not only represent the two prototypical NESs but also mark (together with the SPN1 N terminus) the known extremes of inter-Φ spacing.

Notably, we found that CRM1 uses the very same set of Φ pockets for the recognition of each of these diverse NES peptides. Thus, the

Figure 7 Structure of the CRM1-bound HIV-1 Rev NES. The complex of GTP-Ran (residues 5–180), CRM1 and the NES fused to SPN1 (residues 15–360) was crystallized under two different conditions (**Supplementary Table 1**, **Table 1**). (**a**) Illustration of how the Rev NES docks with Leu6, Pro7, Leu9, Leu12 and Leu14 into the five Φ-pockets. Leu4 is not a Φ residue, but makes a weak hydrophobic contact to Leu6. See also **Supplementary Figure 2**. (**b**) For comparison, the NES-binding site is shown with the SPN1 N terminus bound. (**c**) Top, traditional sequence-based and novel structure-based alignment of SPN1 N terminus and Rev NES. Bottom, structural overlay between the CRM1-bound forms of the Rev NES (colored as in **a**) and the SPN1 N terminus (gray). All four Rev NES structures (derived from the two crystal types, each containing two export complexes per asymmetric unit; **Table 1**) are shown. (**d**) CRM1 binding assay was at 75 mM NaCl. (**e**) Overlay of the structures of the hydrophobic clefts from RanGTP–CRM1 complexes in their SPN1-, PKI NES- and Rev NES-bound (light blue backbones) or cargo-free forms (orange). The side chains shown are those that contact the Φ residues of the NESs.

Figure 8 Redefinition of the NES consensus. (**a**,**b**) Our structural and biochemical analyses revealed that five and not just four Φ positions need to be considered. In addition, the rules for Φ^0-Φ^2 spacing differ markedly between PKI-class NESs (**a**) and Rev-class NESs (**b**). Φ preferences for PKI-class NESs were elucidated in **Figure 4c** and **Supplementary Figure 6**. A given Φ residue was scored as active if the tested NES bound CRM1 in a RanGTP-stimulated manner and stronger than the corresponding serine variant.

spacing[25]. We now have reason to assume that this latter Φ spacing is highly disfavored because it is too short to bridge the distance between the Φ^1 and Φ^2 pockets in a PKI-typical α-helical conformation and too long for a favorable extended conformation. The $\Phi^1 xx\Phi^2$ pattern was historically derived from NESs containing a Rev-like Φ-Pro-x-Φ motif (see ref. 25 and citations therein). Our crystal structure reveals that proline is not part of an inter-Φ spacer but rather is the essential Φ^1 residue of the NES, whereas the preceding leucine is Φ^0. In other words, Rev-class NESs are characterized by a $\Phi^0\Phi^{1\,Pro} x\Phi^2$ and not a $\Phi^1 xx\Phi^2$ spacing (**Fig. 8b**).

NES-like sequence patterns occur rather frequently in proteins, even in those that are not recognized by CRM1. This phenomenon has been studied in detail for the Abl tyrosine kinase[40], whose actin-binding domain contains a perfect 4-Φ PKI-type NES stretch (**L**ENN**L**REL**QI**, ref. 41). When taken out of its original protein context, this NES indeed confers efficient CRM1-dependent nuclear export. However, the complete domain cannot be recruited as a CRM1 cargo because the three leucines of the 'NES' are buried in the hydrophobic core and are thus not accessible for the exportin[40]. Functional NESs should hence lack an appropriate 'packing partner' in their sequence context, and operational classic NESs will not occur within compactly folded domains but rather at the N terminus, at the C terminus or within an unstructured region of an export cargo. The latter consideration is obviously crucial for a precise prediction of export signals. From a protein-folding perspective, the docking of an NES into CRM1's hydrophobic cleft can be seen as a *trans*-complementation that creates a new hydrophobic core.

The most conserved part of CRM1 (and probably of all nuclear transport components) comprises residues 495–595 and includes the hydrophobic cleft as well as flanking residues (**Supplementary Fig. 7**). In humans, this part is 100% identical to those of fish (*Danio rerio*), 94% to *Drosophila melanogaster*, 91% to *Penicillium chrysogenum*, 82% to *Oryza sativa* and even very similar to those of distant protozoa such as *Toxoplasma gondii* (76% identity) or *Tetrahymena thermophila* (53% identity). This extreme conservation readily explains why the Rev and PKI NESs are functional from yeast to human. Notably, however, the conservation is not restricted to the residues that contact the Φ side chains. Instead, strict evolutionary constraints were evidently imposed onto the entire HEAT repeats 11 and 12, which harbor the hydrophobic cleft, as well as onto

flanking regions, in particular the acidic loop that contacts RanGTP. This probably reflects the fact that the NES-binding site and the acidic loop cooperate in a highly optimized manner when CRM1 is switched between its cytoplasmic state (which is the form with low affinity for Ran and cargo) and its nuclear high-affinity state[12,14]. The necessary rigid coupling between RanGTP binding and export cargo loading probably requires the hydrophobic cleft to be rigid in its nuclear state. As a consequence, the NES-binding site cannot adapt to different Φ spacings. For this reason, peptides will display CRM1-dependent NES activity only if a favorable conformation of their backbone can place a sufficient number of Φ side chains into the rigid arrangement of Φ pockets.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/nsmb/.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

AUTHOR CONTRIBUTIONS
T.G. conceived biochemical and crystallization experiments, prepared protein constructs (with the help of D.D.), performed the biochemical experiments and export assays, crystallized the export complexes, contributed to crystallographic data collection and structure refinement, prepared the samples for the NMR study, interpreted data and contributed to the writing of the manuscript (including the figures). T. Madl conceived and performed NMR experiments, interpreted the NMR data and contributed to the writing of the manuscript (including the NMR figures). L.C. conducted initial NMR experiments. P.N. performed crystallographic data collection, data processing and structure refinement, and contributed to the interpretation of the crystallographic data. T. Monecke carried out crystallographic data collection. R.F. conducted molecular replacement for the binary CRM1–RanGTP complex. M.S. conceived NMR experiments, interpreted NMR data and contributed to the writing of the manuscript. D.G. conceived biochemical and crystallization experiments, prepared home-made affinity matrices, interpreted data and wrote the manuscript. All authors commented on the paper.

1. Görlich, D. & Kutay, U. Transport between the cell nucleus and the cytoplasm. *Annu. Rev. Cell Dev. Biol.* **15**, 607–660 (1999).

2. Cook, A.G. & Conti, E. Nuclear export complexes in the frame. *Curr. Opin. Struct. Biol.* **20**, 247–252 (2010).

3. Adachi, Y. & Yanagida, M. Higher order chromosome structure is affected by cold-sensitive mutations in a *Schizosaccharomyces pombe* gene crm1+ which encodes a 115-kD protein preferentially localized in the nucleus and its periphery. *J. Cell Biol.* **108**, 1195–1207 (1989).

4. Fornerod, M., Ohno, M., Yoshida, M. & Mattaj, I.W. Crm1 is an export receptor for leucine rich nuclear export signals. *Cell* **90**, 1051–1060 (1997).

5. Stade, K., Ford, C.S., Guthrie, C. & Weis, K. Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell* **90**, 1041–1050 (1997).

6. Nishi, K. *et al.* Leptomycin B targets a regulatory cascade of crm1, a fission yeast nuclear protein, involved in control of higher order chromosome structure and gene expression. *J. Biol. Chem.* **269**, 6320–6324 (1994).

7. Wolff, B., Sanglier, J.J. & Wang, Y. Leptomycin B is an inhibitor of nuclear export: inhibition of nucleo-cytoplasmic translocation of the human immunodeficiency virus type 1 (HIV-1) Rev protein and Rev-dependent mRNA. *Chem. Biol.* **4**, 139–147 (1997).

8. Kutay, U., Bischoff, F.R., Kostka, S., Kraft, R. & Görlich, D. Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. *Cell* **90**, 1061–1071 (1997).

9. Matsuura, Y. & Stewart, M. Structural basis for the assembly of a nuclear export complex. *Nature* **432**, 872–877 (2004).

10. Cook, A.G., Fukuhara, N., Jinek, M. & Conti, E. Structures of the tRNA export factor in the nuclear and cytosolic states. *Nature* **461**, 60–65 (2009).

11. Okada, C. *et al.* A high-resolution structure of the pre-microRNA nuclear export machinery. *Science* **326**, 1275–1279 (2009).

12. Monecke, T. *et al.* Crystal Structure of the Nuclear Export Receptor CRM1 in Complex with Snurportin1 and RanGTP. *Science* **324**, 1087–1091 (2009).

13. Dong, X. *et al.* Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature* **458**, 1136–1141 (2009).

14. Koyama, M. & Matsuura, Y. An allosteric mechanism to displace nuclear export cargo from CRM1 and RanGTP by RanBP1. *EMBO J.* **29**, 2002–2013 (2010).

15. Ho, J.H., Kallstrom, G. & Johnson, A.W. Nmd3p is a Crm1p-dependent adapter protein for nuclear export of the large ribosomal subunit. *J. Cell Biol.* **151**, 1057–1066 (2000).

16. Gadal, O. *et al.* Nuclear export of 60s ribosomal subunits depends on Xpo1p and requires a nuclear export sequence-containing factor, Nmd3p, that associates with the large subunit protein Rpl10p. *Mol. Cell. Biol.* **21**, 3405–3415 (2001).

17. Moy, T.I. & Silver, P.A. Requirements for the nuclear export of the small ribosomal subunit. *J. Cell Sci.* **115**, 2985–2995 (2002).

18. Thomas, F. & Kutay, U. Biogenesis and nuclear export of ribosomal subunits in higher eukaryotes depend on the CRM1 export pathway. *J. Cell Sci.* **116**, 2409–2419 (2003).

19. Malim, M.H., McCarn, D.F., Tiley, L.S. & Cullen, B.R. Mutational definition of the human immunodeficiency virus type 1 Rev activation domain. *J. Virol.* **65**, 4248–4254 (1991).

20. Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W. & Lührmann, R. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82**, 475–483 (1995).

21. Wen, W., Meinkoth, J., Tsien, R. & Taylor, S. Identification of a signal for rapid export of proteins from the nucleus. *Cell* **82**, 463–473 (1995).

22. Bohnsack, M.T. *et al.* Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm. *EMBO J.* **21**, 6205–6215 (2002).

23. Richards, S.A., Lounsbury, K.M., Carey, K.L. & Macara, I.G. A nuclear export signal is essential for the cytosolic localization of the Ran binding protein, RanBP1. *J. Cell Biol.* **134**, 1157–1168 (1996).

24. Izaurralde, E., Kutay, U., von Kobbe, C., Mattaj, I.W. & Görlich, D. The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. *EMBO J.* **16**, 6535–6547 (1997).

25. Kutay, U. & Guttinger, S. Leucine-rich nuclear-export signals: born to be weak. *Trends Cell Biol.* **15**, 121–124 (2005).

26. Bogerd, H.P., Fridell, R.A., Benson, R.E., Hua, J. & Cullen, B.R. Protein sequence requirements for function of the human T-cell leukemia virus type 1 Rex nuclear export signal delineated by a novel in vivo randomization-selection assay. *Mol. Cell. Biol.* **16**, 4207–4214 (1996).

27. Zhang, M.J. & Dayton, A.I. Tolerance of diverse amino acid substitutions at conserved positions in the nuclear export signal (NES) of HIV-1 Rev. *Biochem. Biophys. Res. Commun.* **243**, 113–116 (1998).

28. Kosugi, S., Hasebe, M., Tomita, M. & Yanagawa, H. Nuclear export signal consensus sequences defined using a localization-based yeast selection system. *Traffic* **9**, 2053–2062 (2008).

29. Huber, J. *et al.* Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *EMBO J.* **17**, 4114–4126 (1998).

30. Paraskeva, E. *et al.* CRM1-mediated recycling of snurportin 1 to the cytoplasm. *J. Cell Biol.* **145**, 255–264 (1999).

31. Engelsma, D., Bernad, R., Calafat, J. & Fornerod, M. Supraphysiological nuclear export signals bind CRM1 independently of RanGTP and arrest at Nup358. *EMBO J.* **23**, 3643–3652 (2004).

32. Engelsma, D. *et al.* A supraphysiological nuclear export signal is required for parvovirus nuclear export. *Mol. Biol. Cell* **19**, 2544–2552 (2008).

33. Cook, A., Bono, F., Jinek, M. & Conti, E. Structural biology of nucleocytoplasmic transport. *Annu. Rev. Biochem.* **76**, 647–671 (2007).

34. Rice, L.M. & Brunger, A.T. Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. *Proteins* **19**, 277–290 (1994).

35. Kudo, N. *et al.* Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. *Proc. Natl. Acad. Sci. USA* **96**, 9112–9117 (1999).

36. Nilges, M. A calculation strategy for the structure determination of symmetrical dimers by H-1-NMR. *Proteins* **17**, 297–309 (1993).

37. Nilges, M. & O'Donoghue, S.I. Ambiguous NOEs and automated NOE assignment. *Prog. Nucl. Magn. Reson. Spectrosc.* **32**, 107–139 (1998).

38. Madl, T., Bermel, W. & Zangger, K. Use of relaxation enhancements in a paramagnetic environment for the structure determination of proteins using NMR spectroscopy. *Angew. Chem. Int. Ed.* **48**, 8259–8262 (2009).

39. Askjaer, P. *et al.* RanGTP-regulated interactions of CRM1 with nucleoporins and a shuttling DEAD-box helicase. *Mol. Cell. Biol.* **19**, 6276–6285 (1999).

40. Hantschel, O. *et al.* Structural basis for the cytoskeletal association of Bcr-Abl/c-Abl. *Mol. Cell* **19**, 461–473 (2005).

41. Taagepera, S. *et al.* Nuclear-cytoplasmic shuttling of C-ABL tyrosine kinase. *Proc. Natl. Acad. Sci. USA* **95**, 7457–7462 (1998).

42. Hutchinson, E.G. & Thornton, J.M. PROMOTIF—a program to identify and analyze structural motifs in proteins. *Protein Sci.* **5**, 212–220 (1996).

43. Ribbeck, K. & Görlich, D. The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J.* **21**, 2664–2671 (2002).

44. Weiss, M.S. & Hilgenfeld, R. On the use of the merging R factor as a quality indicator for X-ray data. *J. Appl. Cryst.* **30**, 203–205 (1997).

45. Linge, J.P., Williams, M.A., Spronk, C.A., Bonvin, A.M. & Nilges, M. Refinement of protein structures in explicit solvent. *Proteins* **50**, 496–506 (2003).

46. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223 (2009).

47. Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R. & Thornton, J.M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486 (1996).

ONLINE METHODS

Crystallization and structure determination. Protein expression and purification are described in the **Supplementary Methods**. All complexes were crystallized by vapor diffusion (**Supplementary Table 1**). X-ray oscillation photographs were integrated and scaled using XDS[48].

Crystals containing the NES-SPN1 chimeras were isomorphous with the described crystals of the SPN1–CRM1–RanGTP complex[12] (PDB-ID 3GJX). Iterative cycles of refinement were done using CNS[49]. To avoid overfitting, refinement was based on slow-cooling simulated annealing (SA) torsion angle dynamics (TAD) combined with standard minimization and individually restrained B-factor refinement. The structures were refined using strong non-crystallographic symmetry (NCS) restraints between equivalent residues, which were later relaxed for regions showing different conformations. The final refinement steps were performed with PHENIX[50], using the NCS restraints and the set of reflections for calculating R_{free} that had also been used for CNS refinement. Translation-libration-screw (TLS) refinement was applied in order to derive the anisotropic motions of the domains. Water molecules for the HIV-1 Rev models had been added manually, based on stringent criteria: that is, only if both difference electron density maps ($2F_o − F_c$ contoured at 1σ and $F_o − F_c$ contoured at 3σ) showed a peak within a distance ranging between 2.2 and 3.5 Å to oxygen or nitrogen atoms of the protein molecules. Deletion of these water molecules increased both R and R_{free} factors by a similar value, indicating that the water molecules included in the current model are justified and are not simply masking errors.

The structure of the binary RanGTP–CRM1 complex was solved by molecular replacement using PHASER[51], with the crystal structures of Ran (residues 9–179) and CRM1 (residues 12–1055) (PDB-ID 3GJX, ref. 12) as separated search models. Refinement was conducted using the same strategy as was used for the cargo complexes (refinement with CNS, followed by use of PHENIX, with grouped B-factor refinement and TLS).

Models were built in COOT[52] based on SIGMA$_A$-weighted SA omit maps and/or difference electron density maps. Molecular contacts were analyzed with the CONTACTS program of the CCP4 suite (version 6.1.3, ref. 53). Map correlation coefficients in **Table 1** were calculated with PHENIX (against $2F_o − F_c$ map). All main figures were prepared with PyMOL (http://www.pymol.org).

NMR spectroscopy and NMR structure calculation. Protein sample preparation is described in the **Supplementary Methods**. NMR spectra were recorded at 298 K using Avance 900 and Avance III 600 Bruker NMR spectrometers, equipped with cryogenic triple resonance gradient probes. Assignments were carried out using a combination of CRINEPT[54], Methyl-TROSY[55], NOE-based (mixing times 30–140 ms) and ^{13}C direct-detected[56,57] NMR experiments (see **Supplementary Methods**). ^{13}C direct-detected spectra were recorded at 298 K using Avance III 600 (equipped with a TCI probehead) and Avance III 700 (equipped with a CP-TXO probehead) spectrometers. Spectra were processed with NMRPipe/Draw[58] and analyzed with Sparky 3 (T.D. Goddard and D.G. Kneller, University of California, San Francisco, California, USA) and NMRView[59].

For measurement of solvent PREs, Gd(DTPA-BMA) was added to the protein sample to final concentrations of 0.5–4.5 mM. Proton T_1 relaxation times were obtained from a series of saturation-recovery 2D ^{1}H,^{13}C-HMQC and ^{1}H,^{15}N-HMQC spectra. Typically, eight such 2D datasets were acquired with recovery delays of 0.2, 0.5, 1.0, 1.5, 2.0, 4.0, 6.0 and 8.0 s. For two delays, duplicates were acquired for statistical purposes. To obtain the relaxation times, the peak intensities were fitted to

$$I = I_0 \left(1 - e^{-\frac{t}{T_1}} \right)$$

using the NMRViewJ (version 8.0) rate analysis tool. PREs were calculated as reported[38,60] and are represented by the slope of the relaxation rate as a function of the concentration of the paramagnetic agent. PREs were back calculated and averaged for the ten lowest-energy structures by numerical integration of the paramagnetic environment[38].

Structures were calculated with the standard ARIA/CNS setup[61,62] using NOEs and backbone dihedral angles derived from ^{13}C secondary chemical shifts (TALOS+)[46]. The peptide was docked to the CRM1–RanGTP crystal structure (PDB-ID 3GJX, ref. 14) using all intrapeptide restraints, ambiguous PKI NES–CRM1

NOEs[36,37] and unambiguous PKI NES–CRM1 Cys528 NOEs. The protein backbone of CRM1 was fixed while CRM1 side chains were left flexible for optimal accommodation of the PKI NES in the hydrophobic cleft. From the 1,000 structures generated, the 100 structures of lowest energy were evaluated against solvent PRE data (**Supplementary Fig. 4b**) and ranked according to the correlation between measured and back-calculated PREs. The 10 best structures in terms of the correlations are displayed and used to calculate the structural statistics (**Table 2**).

Nuclear export assay. HeLa cell were prepared as described[63,64], with some changes of the protocol. Briefly, cells were grown to ≈95% confluency, washed with PBS, detached with a Citrate-EDTA solution and washed three times with ice-cold transport buffer (20 mM HEPES-KOH pH 7.5, 110 mM potassium acetate, 5 mM magnesium acetate, 0.5 mM EGTA, 250 mM sucrose). Cell membranes were permeabilized with 25 µg ml^{-1} digitonin (in ice-cold transport buffer) and nuclei recovered by centrifugation. After two washes, nuclei were resuspended in transport buffer (without sucrose) containing 5% (w/v) glycerol and 0.5 M trehalose and were slowly frozen at −80 °C. We used *Xenopus laevis* egg extract[65] for nuclear export assays, as it stabilizes the nuclei over a period of hours[64]. Nuclei were mixed with egg extract (supplemented with 0.5 µM CRM1 and a GTP-regenerating system) and incubated at 22 °C for 30 min before addition of 2 µM of the indicated fluorescent proteins. Imaging was performed with a confocal laser-scanning microscope system (SP5; Leica) using the 488- and 561-nm laser lines and a 63×–numerical aperture 1.3 Plan-Apochromat glycerol objective (Leica).

Fluorescence polarization (FP) assay. 10 nM of the PKI Φ^0Leu NES (fluorescein maleimide–labeled via its C-terminal cysteine) were incubated for 4 h with the indicated concentrations of CRM1 and GTP-Ran (Q69L). The assay buffer contained 20 mM Tris-HCl, pH 7.5, 130 mM NaCl, 2 mM Mg(OAc)$_2$, 1 mM DTT, 0.005% (w/v) digitonin, 0.1 mg ml^{-1} BSA and a RanGTP-regenerating system (0.5 mM GTP, 10 mM PEP, 0.5 µM pyruvate kinase, 0.5 µM RanGEF/RCC1). FP measurements were performed with a Victor 3 multilabel plate reader (PerkinElmer). Apparent equilibrium dissociation constants were derived by nonlinear regression using Prism (GraphPad Software, Inc.).

48. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795–800 (1993).
49. Brunger, A.T. Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2**, 2728–2733 (2007).
50. Adams, P.D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948–1954 (2002).
51. McCoy, A.J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 32–41 (2007).
52. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
53. Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
54. Riek, R., Wider, G., Pervushin, K. & Wüthrich, K. Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. *Proc. Natl. Acad. Sci. USA* **96**, 4918–4923 (1999).
55. Tugarinov, V., Hwang, P.M., Ollerenshaw, J.E. & Kay, L.E. Cross-correlated relaxation enhanced 1H[bond]13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J. Am. Chem. Soc.* **125**, 10420–10428 (2003).
56. Bermel, W., Bertini, I., Felli, I.C., Piccioli, M. & Pierattelli, R. ^{13}C-detected protonless NMR spectroscopy of proteins in solution. *Prog. Nucl. Magn. Reson. Spectrosc.* **48**, 25–45 (2006).
57. Oh, B.H., Westler, W.M., Darba, P. & Markley, J.L. Protein C-13 spin systems by a single two-dimensional nuclear magnetic-resonance experiment. *Science* **240**, 908–911 (1988).
58. Delaglio, F. *et al.* Nmrpipe—a multidimensional spectral processing system based on unix pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
59. Johnson, B.A. & Blevins, R.A. NMRView: a computer-program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**, 603–614 (1994).
60. Pintacuda, G. & Otting, G. Identification of protein surfaces by NMR measurements with a paramagnetic Gd(III) chelate. *J. Am. Chem. Soc.* **124**, 372–373 (2002).
61. Linge, J.P., Habeck, M., Rieping, W. & Nilges, M. ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics* **19**, 315–316 (2003).
62. Linge, J.P., O'Donoghue, S.I. & Nilges, M. Automated assignment of ambiguous nuclear overhauser effects with ARIA. *Methods Enzymol.* **339**, 71–90 (2001).
63. Adam, S.A., Marr, R.S. & Gerace, L. Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. *J. Cell Biol.* **111**, 807–816 (1990).
64. Stüven, T., Hartmann, E. & Görlich, D. Exportin 6: a novel nuclear export receptor that is specific for profilin.actin complexes. *EMBO J.* **22**, 5928–5940 (2003).
65. Leno, G.H. & Laskey, R.A. DNA replication in cell-free extracts from Xenopus laevis. *Methods Cell Biol.* **36**, 561–579 (1991).