Binding stoichiometry and structural model of the HIV-1 Rev/Importin β complex

Didier Spittler[1], Rose-Laure Indorato[1], Elisabetta Boeri Erba[1], Elise Delaforge[1], Luca Signor[1], Simon J. Harris[1], Isabel Garcia-Saez[1], Andrés Palencia[2], Frank Gabel[1], Martin Blackledge[1], Marjolaine Noirclerc-Savoye[1*], Carlo Petosa[1*]

[1] Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale (IBS), 71 Avenue des Martyrs, 38000 Grenoble, France

[2] Institute for Advanced Biosciences (IAB), Structural Biology of Novel Targets in Human Diseases, INSERM U1209, CNRS UMR5309, Université Grenoble Alpes, Grenoble, France

*Corresponding authors: marjolaine.noirclerc@ibs.fr, carlo.petosa@ibs.fr

ABSTRACT

HIV-1 Rev mediates the nuclear export of intron-containing viral RNA transcripts and is essential for viral replication. Rev is imported into the nucleus by the host protein Importin β (Impβ), but how Rev associates with Impβ is poorly understood. Here we report biochemical, biophysical and structural studies of the Impβ/Rev complex. Gel shift, native mass spectrometry and isothermal titration calorimetry data reveal that Impβ binds two Rev monomers through independent binding sites. Small-angle X-ray scattering (SAXS) data suggest that the HEAT repeats of Impβ retain an extended conformation upon binding Rev, which according to NMR data is primarily recognized through its helical hairpin domain. Peptide scanning data and charge-reversal mutations identify the N-terminal tip of Rev helix α2 within Rev's Arginine-Rich Motif (ARM) as a primary Impβ binding epitope. Crosslinking mass spectrometry and compensatory mutagenesis data combined with molecular docking simulations suggest a structural model in which one Rev monomer binds to the C-terminal half of Impβ with Rev helix α2 roughly parallel to the HEAT-repeat superhelical axis while the other monomer binds to the N-terminal half. These findings shed light on the molecular basis of Rev recognition by Impβ and highlight an atypical binding behaviour that distinguishes Rev from canonical cellular Impβ cargos.

INTRODUCTION

The HIV-1 protein Rev (Regulator of Expression of the Virion) is an RNA-binding protein essential for the production of mature viral particles[1,2] (reviewed in refs.[3-5]). The initial RNA molecule that results from proviral DNA transcription is differentially processed into multiply spliced transcripts that encode Rev and other early-stage viral proteins, as well as partly spliced and unspliced transcripts that encode viral structural and enzymatic proteins and provide genomic RNA for encapsidation. Host cell mechanisms rapidly export fully spliced transcripts to the cytoplasm but retain incompletely spliced molecules in the nucleus[6,7]. Rev circumvents these mechanisms by mediating the nuclear export of intron-containing viral transcripts[8-10]. Rev binds to these transcripts by recognizing the Rev-response element (RRE), a ~350-nucleotide region that adopts a complex stem-loop structure[8,9,11,12]. A first Rev monomer binds with high affinity to a small region of the RRE, stem loop IIB, triggering the cooperative binding of additional Rev monomers to the RRE[13,14]. The Rev/RRE complex is then exported to the cytoplasm by the nuclear export factor CRM1[15-17], allowing for synthesis of late-stage viral proteins as well as packaging of the replicated viral genome into new viral particles. Other reported roles for Rev include regulating the splicing[18-20], translation[21-23] and encapsidation[24-26] of viral RNAs; modulating the stability or localization of the HIV-1 proteins Tat[27] and integrase (IN)[28,29]; and interacting with host cell molecules[30], such as RNA helicases and hnRNPs[31-35].

In order to bind RRE-containing viral transcripts and perform other nuclear functions, Rev must first be imported from the cytosol, a step mediated by host cell proteins of the karyopherin/importin-β family (reviewed in refs.[36-41]). Different members of this family have been implicated as potential Rev import factors, including Importin β (Impβ), Transportin-1, Transportin-5 and Transportin-7[42-45]. One study reported that Rev is imported

by Impβ in T-lymphocyte derived (293T, Jurkat and CEM) cell lines, and by Transportin-1 in HeLa cells and in a monocytic (THP-1) and histocytic (U937) cell line, revealing that the Rev import pathway is host-cell dependent[45]. Given the central role of CD4+ T lymphocytes for HIV-1 biology, the present study focuses on the Impβ pathway.

Impβ heterodimerizes with importin α (Impα) to import proteins bearing a classical nuclear localization signal (NLS), which typically contains 1 or 2 clusters of basic residues[37]. Impα uses its C-terminal domain to recognize the NLS[46,47] and its N-terminal Impβ-binding (IBB) domain to associate with Impβ[48]. HIV-1 Rev bypasses this pathway by directly interacting with Impβ and is imported independently of Impα[42]. Impβ is regulated by Ran, a Ras-related GTPase which is predominantly bound to GTP in the nucleus and to GDP in the cytosol. Impβ associates with its cargo (directly or via Impα) in the cytosol, translocates through the nuclear pore complex via favourable interactions with FG-repeat containing nucleoporins, and releases the cargo in the nucleus upon binding RanGTP. Atomic structures are known for Impβ in the unbound state[49,50] and in complex with Ran[51-53], nucleoporins[54-56], importin7[57] and diverse macromolecular cargos[48,57-63]. Impβ is a spiral shaped molecule composed of 19 tandem helical hairpin motifs called HEAT repeats, with each hairpin comprising an outward-facing "A" helix and an inward-facing "B" helix. RanGTP and the cargo bind on the inner concave surface of the spiral, with different cargos interacting with different subsets of the HEAT repeats. The outer convex surface of Impβ interacts with nucleoporins that constitute the nuclear pore complex[54-56].

Structures are known for HIV-1 Rev and Rev-derived peptides, either alone[64-68], bound to an antibody[69], or in complex with RNA[70-75] or with CRM1 and RanGTP[76]. Standard lab isolates of Rev contain 116 amino acid residues. These define an N-terminal domain (NTD) that mediates nuclear import, RNA-binding and multimerization, and an intrinsically disordered C-terminal domain (CTD) responsible for nuclear export (**Figure 1a**). The CTD contains a leucine-rich nuclear export signal (NES; res. 75-83) which is recognized in the nucleus by CRM1 in complex with RanGTP. Residues C-terminal to the NES enhance protein stability[77,78], become partly ordered when Rev forms filaments[67], and have been predicted to interact with the NTD[79] and to regulate accessibility of the NES[80]. The NTD includes an α-helical hairpin (res. 10-65) characterized by an Arginine-rich motif (ARM; res. 35-50) that mediates RRE-binding[13,81-84] and contains the NLS recognized by Impβ[42,43]. The N-terminal 10 residues are important for protein stability[78] but appear not to constitute a significant interaction epitope[31]. Hydrophobic residues on either side of the ARM comprise an oligomerization domain responsible for Rev multimerization[13,14,85]. These residues localize to both faces of the helical hairpin (denoted "head" and "tail" surfaces) and mediate three types of homotypic interactions (**Figure 1b**). Tail-to-tail (also called B-B) interactions mediate dimer formation following the initial association of Rev with the RRE[65,74,85]. Head-to-head (or A-A) interactions mediate higher-order Rev oligomer formation[69,85]. The third type of interaction, called C-C, involves the proline-rich interhelical loop and is proposed to bridge independently bound Rev dimers on the RRE surface[67]. These three types of interface allow Rev to adapt to a wide variety of RNA sites on the RRE.

To gain insights into how Impβ associates with Rev, we analysed the Impβ/Rev complex using diverse biochemical, biophysical and structural approaches together with molecular docking simulations. Unexpectedly, our data reveal that Impβ associates with two Rev monomers through distinct binding sites. Moreover, they allow

us to localize one monomer near the N-terminal end of the HEAT-repeat superhelix and to deduce an approximate structural model of the other Rev monomer bound near the opposite, C-terminal end. The ability to engage two independent binding sites on Impβ identifies Rev as an atypical nuclear import cargo.

RESULTS

Use of point mutations to disrupt Rev oligomerization

Our initial efforts to study the Impβ/Rev interaction were hampered by the tendency of Rev to aggregate under the low to medium ionic strength conditions required for stable Impβ/Rev complex formation. Previous studies reported several point mutations that hinder Rev multimerization[13,65,85-89], including mutations V16D and I55N which destabilize the homotypic A-A and B-B interfaces, respectively (**Figure 1b**)[85,89]. In particular, the double V16D/I55N mutant was previously used for biophysical studies of Rev, including single-molecule fluorescence spectroscopy[90] and NMR[66] experiments. Accordingly, we generated Rev mutants V16D and V16D/I55N, as well as a truncated version (residues 4-69) of the V16D/I55N mutant lacking the C-terminal domain. Bacterially expressed Rev proteins were purified to homogeneity and their oligomeric state verified by size exclusion chromatography/multi-angle laser light scattering (SEC/MALLS; **Figure 1c**). Wildtype (WT) Rev eluted as two separate peaks: the first contained molecular species ranging between 330 and 500 kDa in mass (suggesting Rev multimers comprising approximately 24-36 monomers) while the second exhibited masses of 40-80 kDa (consistent with species comprising 3-6 Rev monomers), confirming the highly polydisperse nature of Rev. The V16D mutant eluted later than WT Rev, yielding a broad peak with a mass distribution (15-45 kDa) consistent with a mixture of species comprising 1-3 Rev monomers. The mass observed at the maximum of this peak (~23 kDa) indicated a prevalence of the dimeric form (27.6 kDa), in agreement with previous findings[85,89]. By contrast, both the full-length and truncated versions of the V16D/I55N mutant eluted as relatively narrow peaks with mass distributions centered around 13.3 and 7.5 kDa, respectively, consistent with a monomeric state (13.8 and 8.0 kDa, respectively). The monodisperse nature of the V16D/I55N mutant facilitated several of the interaction studies with Impβ described below. For convenience, we refer to the full-length and truncated versions of this mutant as Rev^{OD} (for oligomerization-deficient Rev) and $Rev^{OD}\Delta$, respectively.

Impβ binds up to two monomers of Rev

We next assessed the ability of Rev to associate with Impβ. SEC analysis revealed that unbound Impβ and Rev^{OD} each eluted as a single peak, whereas pre-incubating Impβ with a molar excess of Rev^{OD} led to the co-elution of both proteins within an earlier peak, indicating complex formation (**Figure 1d**). SEC/MALLS analysis revealed a molecular mass for the Impβ/Rev^{OD} complex of 105 kDa (**Figure 1e**), consistent with the mass expected for 1:1 stoichiometry (110.6 kDa). Impβ/Rev^{OD} complex formation was confirmed by native gel analysis, which revealed a single band for unbound Impβ and a discrete shift upon the addition of increasing concentrations of Rev^{OD} (**Figure 1f**, Complex 1). Surprisingly, higher Rev^{OD} concentrations resulted in a supershift, suggesting the formation of an additional species containing two or more Rev monomers (**Figure 1f**, Complex 2). The higher

Rev^{OD} concentrations required for the supershifted band and the failure to detect Complex 2 by SEC/MALLS suggest that this complex is less stable than Complex 1 and likely dissociates over the course of the chromatographic run. A glutaraldehyde crosslinking experiment followed by mass spectrometry (MS) analysis corroborated the ability of Impβ to bind two Rev monomers (**Figure S1**).

To further verify the stoichiometry of the Impβ/Rev association we performed native MS, which preserves non-covalent interactions and allows the mass of intact macromolecular complexes to be determined[91]. To ensure correct interpretation of native MS spectra, we first used liquid chromatography coupled to electrospray ionization MS (LC/ESI-MS) to measure accurate masses under denaturing conditions for the unbound Impβ and Rev^{OD} proteins; these were 97,300 Da and 13,264 Da, respectively, which are within 1 Da of the theoretical masses (**Figure S2** and **Table S5**). Native MS analysis of unbound Impβ yielded a clear set of peaks whose *m/z* values correspond to a mass of 97,296 Da (**Figure 2a** and **g**), closely matching that determined by ESI-TOF. We next incubated Impβ with Rev^{OD} and measured the masses of the resulting species (**Figure 2b**, **c** and **g**). Mixing Impβ with a 2-fold molar equivalent of Rev^{OD} led to the appearance of two additional sets of peaks: the first corresponds to a mass of 110,558 Da (**Figure 2b,** blue circles), matching the expected mass of an Impβ/Rev^{OD} heterodimer (110,563 Da), and the second to a mass of 123,829 Da (**Figure 2b,** green circles), consistent with a complex of Impβ bound to two Rev^{OD} monomers (123,827 Da). Mixing Impβ with Rev^{OD} in a 1:5 molar ratio enhanced the intensity of this second set of signals (**Figure 2c**). Mixing Impβ with the truncated construct ($Rev^{OD}\Delta$) instead of full length Rev^{OD} yielded an analogous set of peaks for Impβ/$Rev^{OD}\Delta$ complexes having a stoichiometry of 1:1 (mass: 105,330 Da) and 1:2 (113,366 Da) (**Figure 2d** and **g**). Taken together these results confirm the ability of Impβ to bind up to two monomers of Rev^{OD}.

We then investigated complex formation between Impβ and Rev^{WT}. We overcame the tendency of Rev^{WT} to oligomerise by exchanging its buffer into a high concentration of ammonium acetate prior to incubating with Impβ and using relatively low Rev^{WT} concentrations (10-20 μM). Upon mixing Impβ with a 2-fold molar equivalent of Rev^{WT}, we detected peaks consistent with the formation of a 1:1 Impβ:Rev^{WT} complex (110,476 Da; **Figure 2e**). With a 5-fold molar equivalent of Rev^{WT} we detected both 1:1 and 1:2 complexes (123,649 Da; **Figure 2f**). This confirms that Impβ can bind either one or two molecules of WT Rev.

Impβ recognizes the Rev helical hairpin through a high- and a low-affinity binding site

We next characterized the Impβ/Rev interaction using isothermal titration calorimetry (ITC). We recorded ITC profiles for the binding of Impβ to either Rev^{OD} or $Rev^{OD}\Delta$. Fitting the resulting binding isotherms using a model consisting of a single class of binding sites resulted in a poor fit, whereas a good fit was obtained with a model consisting of two non-symmetric classes of binding sites (**Figure 3a,b**), further confirming the ability of Impβ to bind two Rev monomers. These analyses yielded a dissociation constant (K_d) for the Impβ/Rev^{OD} interaction of 0.8 μM for the first site and 9.6 μM for the second site. Comparable K_d values (0.7 and 4.6 μM) were obtained for the Impβ/$Rev^{OD}\Delta$ interaction, suggesting that the C-terminal domain missing from $Rev^{OD}\Delta$ does not contribute significantly to binding affinity.

The finding that Impβ binds one Rev monomer with submicromolar affinity and the other with ~10-fold weaker affinity is consistent with the sequential appearance of complexes observed when Impβ is titrated with Rev (**Figures 1f, 2** and **S1a**) and with the inability to detect the more weakly bound Rev monomer by SEC-MALLS (**Figure 1e**). The interaction of Impβ with both the first and second Rev monomers is enthalpically driven (**Figure 3c** and **d**), suggesting electrostatic and hydrogen bond interactions that presumably take place between the highly basic ARM of Rev and the acidic cargo-binding inner surface of Impβ. Interestingly, whereas the higher affinity interaction (site 1) is associated with a negligible entropy change, the lower affinity interaction (site 2) is characterized by an unfavourable entropy term, suggesting that the binding of the second Rev monomer might significantly reduce the conformational entropy of the Impβ HEAT-repeat array, which is dynamically highly flexible in the unbound state[92].

To further probe the regions of Rev involved in Impβ binding, we performed NMR spectroscopy of ^{15}N-labelled Rev^{OD} in the presence and absence of unlabelled Impβ. Triple resonance NMR was used to assign the backbone resonances of ^{15}N-labelled Rev^{OD}. The ^{1}H-^{15}N-BEST-TROSY spectrum of free Rev^{OD} shows peaks concentrated between 7.5 and 8.5 ppm (**Figure 4a**), in agreement with a previous study[66] and consistent with the expectation that Rev contains many intrinsically disordered residues[93]. In total 67 of the 116 residues of Rev could be assigned. Assignment was hampered by resonance overlap caused by the high (14%) arginine composition of the Rev sequence and the large number (10) of prolines, as well as severe line-broadening in the C-terminus of the helical hairpin, possibly due to residual oligomerization at concentrations used for NMR measurement (90 μM). Analysis of the secondary ^{13}C chemical shifts showed a strong helical propensity for residues within the helical hairpin domain, as expected, as well as significant helical propensity for residues 84-93 (**Figure 4c**). The latter residues are immediately downstream of the NES and their helical propensity might facilitate recruitment of the NES to the binding surface of CRM1.

The addition of Impβ to Rev^{OD} had a dramatic effect on the spectrum, causing 28 peaks to disappear either completely (24 peaks) or nearly completely (4 peaks; >95% reduction in peak height) (**Figure 4b** and **d**). These peaks include all assigned peaks between residues 10 and 65, which delimit the helical hairpin domain. By contrast, peaks corresponding to N- and C-terminal regions of Rev (residues 2-9 and 73-116) were less affected by the presence of Impβ. This finding indicates that Impβ recognizes Rev primarily through its helical hairpin domain and interacts only weakly or not at all with the N-terminal extension and C-terminal domain, in agreement with the similar binding affinities measured by ITC for the interaction of Impβ with Rev^{OD} and Rev^{OD}Δ.

Impβ retains an extended conformation upon Rev binding

We next sought structural information on the Impβ/Rev complex using small-angle X-ray scattering, which was coupled with size-exclusion chromatography (SEC-SAXS) to ensure sample monodispersity. We measured data from Impβ alone or in complex with Rev^{OD}Δ and compared the observed scattering to that predicted for cargo-bound Impβ conformations (cargo coordinates removed) available from the Protein Data Bank (PDB) (**Figure 5a**). The scattering profile of unbound Impβ exhibits a shoulder at $s \approx 0.1$ Å^{-1} (**Figure 5b**), in agreement

with previous studies of unbound murine and fungal Impβ proteins[49,50,94]. The radius of gyration (R_g) and maximum dimension (D_{max}) of unbound Impβ determined from the pair distance distribution function were 37 Å and 125 Å, respectively, indicating that Impβ adopts a more extended conformation in solution compared to cargo-bound crystal structures (**Figure 5a** and **c**). Consistent with this finding, the observed scattering showed the best agreement (χ^2 values < 4) with scattering curves calculated for Impβ coordinates from PDB structures 1UKL and 3W5K, which exhibit the most elongated cargo-bound conformations (**Figure 5c**).

The binding of $Rev^{OD}\Delta$ caused only modest changes in the scattering profile relative to that of unbound Impβ (**Figure 5b**), corresponding to a small decrease in R_g and D_{max} values (to 36.7 Å and 120 Å respectively). These changes are consistent with the expected binding of $Rev^{OD}\Delta$ to the concave inner surface of the HEAT-repeat array, which appears not to undergo a dramatic conformational change relative to the unbound state. (Note that these inferences only pertain to Impβ bound to a single Rev monomer, since we were unable to identify conditions in which the second, more weakly bound Rev monomer remained associated with Impβ over the course of the SEC column run). The small mass of $Rev^{OD}\Delta$ relative to Impβ and the large impact that changes in HEAT repeat conformation have on the predicted scattering curve preclude us from deducing a reliable structural model for the 1:1 Impβ:$Rev^{OD}\Delta$ complex. Nevertheless, the data reveal that the complex retains an extended conformation resembling that of unbound Impβ and of the cargo-bound conformations 1UKL and 3W5K. Moreover, they confirm that Rev (at least the more tightly bound monomer) binds to the inner surface of Impβ.

Rev binding sites localize to N- and C-terminal regions of Impβ

To obtain additional structural insights into how Impβ recognizes Rev we performed crosslinking-mass spectrometry (XL-MS) on the Impβ/Rev complex using the amine-reactive crosslinker bissulfosuccinimidyl suberate (BS3). Although Rev contains two lysine residues (Lys20 and Lys115), BS3-mediated crosslinks with Impβ were only detected for Lys20 and for the Rev N-terminal amino group (**Figure 6a**). Conversely, peptides containing BS3-modified Lys residues were identified for 28 of the 43 lysines present in Impβ. These residues are distributed over the length of Impβ and can be classified into three groups. Group 1 formed crosslinks with both the N-terminus and Lys20 of Rev; Group 2 formed crosslinks with the flexible Rev N-terminus but not with Lys20; and Group 3 formed monolinks (where only one end of BS3 covalently bound to the protein) or crosslinks with other Impβ residues but did not form detectable crosslinks with Rev. Since Lys20 is located on Rev helix α1, Group-1 lysines on Impβ are expected to localize close to Rev's helical hairpin, whereas lysines in Groups 2 and 3 are expected to be farther away. Interestingly, whereas Group-3 residues primarily reside on the A helix of a HEAT repeat and localize to the outer convex surface of Impβ, nearly all Group-1 residues reside on a B helix and localize to the inner concave surface (**Figure 6b,c**), consistent with Rev recognition by the inner surface of Impβ.

The 11.4 Å linker arm length of BS3 allows the Cα atoms of crosslinked lysines to be up to ~24 Å apart; however, to allow for conformational dynamics, distance constraints of 30-35 Å are typically used for the purposes of 3D structural modelling[95-97]. For such modelling, the solvent-accessible surface (SAS) distance between crosslinked Lys residues is more informative than the straight-line (Euclidean) distance separating them, as the latter may

trace a path sterically inaccessible to the crosslinker[98]. Group 1 includes a cluster of lysines within HEAT repeats 1 and 2, and a second cluster within repeat 19, at the opposite extremity of Impβ. These two clusters are too far apart to form crosslinks to the same Rev K20 residue position (**Figure 6d**). For example, if one considers the Impβ conformation shown in **Figure 6e** (from PDB 1UKL), the Cα atoms of residues Lys23 and Lys873 are separated by a SAS distance of 103 Å (Euclidean distance of 91.4 Å). Even for the most compact known conformation of Impβ (from PDB 3LWW), residues Lys859, Lys867 and Lys873 are separated from Lys23 by a SAS distance of over 70 Å, twice the BS3 crosslinking distance constraint (**Figure S3a,b**). These findings are consistent with the ability of Impβ to bind two Rev molecules and suggest that the two clusters of Group-1 lysines form crosslinks with distinct Rev monomers, one bound close to the N-terminal, and the other close to the C-terminal, extremity of Impβ. We refer hereafter to these two Rev binding sites on Impβ as the N-site and C-site, respectively.

Applying a 35 Å distance constraint from each residue in the two Group-1 lysine clusters localizes the Lys20 Cα atoms of the two Rev monomers within an ellipsoidal volume next to the N- and C-terminal HEAT repeats with a positional uncertainty of ~22 Å (**Figure 6f** and **Figure S3c**). In contrast, Group-1 residues Lys206 and Lys537 could potentially crosslink with either Rev monomer. The fact that both these residues localize to the concave inner surface of Impβ confirms that at least one of the two Rev monomers is bound to this surface.

Impβ recognizes peptides derived from Rev helix α2

To delineate the region(s) of Rev recognized by Impβ we examined the ability of different Rev-derived peptides to interact with Impβ in a thermal shift assay. We used a set of 27 peptides (each comprising 15 residues) that span the entire Rev sequence, with an 11-residue overlap between consecutive peptides. We used differential scanning fluorimetry (DSF) to assess the ability of peptides to increase the thermal stability of Impβ. In the absence of peptides, the melting temperature (T_m) observed for Impβ was 35.4 ± 0.7 °C (**Figure 7a**). Whereas most Rev peptides did not significantly increase this value ($\Delta T_m < 0.5$ °C), three consecutive Rev peptides (numbered 9 to 11) strongly stabilized Impβ ($\Delta T_m > 2$ °C), with peptide 10 (^{37}ARRNRRRRWRERQRQ^{51}), which contains all but one of the ARM's 10 Arg residues, yielding the strongest effect (**Figure 7a, b**). These peptides collectively span Rev residues 33-55 on helix α2 and the α1-α2 loop and nearly perfectly match the ARM motif (res. 35-50) (**Figure 7c**).

The increased stabilization observed for peptide 10 compared to peptide 9 (^{33}GTRQARRNRRRRWRE^{47}) suggests that the Rev ^{48}RQRQ^{51} motif includes an important binding epitope. Similarly, whereas Impβ was strongly stabilized by Rev peptide 11 (^{41}RRRRWRERQRQIRSI^{55}), little or no stabilization was observed with peptide 12 (^{45}WRERQRQIRSISGWI^{59}), indicating that the tetra-arginine motif (^{41}RRRR^{44}) is critical for Impβ binding. Finally, peptide 8 (^{29}PSPEGTRQARRNRRR^{43}), which contains three of these four Arg residues and a total of six closely spaced Arg residues, only modestly stabilizes Impβ (ΔT_m=1.0 ± 0.7°C), indicating that clustered Arg residues are not sufficient for high-affinity binding and that sequence context also plays an important role.

The N-terminal tip of Rev helix α2 is a major Impβ binding epitope

We next performed mutagenesis experiments to identify individual Rev residues critical for Impβ recognition. Since Rev binding is likely to involve a significant electrostatic component, we sought to destabilize the Impβ/Rev interface by introducing charge-reversal mutations on Rev. We first made mutants in which 2 or 3 basic residues were simultaneously replaced by Asp residues (mutants R1-R5; **Figure 8a**). We used a competitive fluorescence polarization (FP) assay to compare the ability of these proteins to bind Impβ and displace a peptide cargo. In this assay the binding of Impβ to a fluorescently labelled Rev peptide (Rev-NLS) yields an FP signal that decreases when the peptide is displaced by the addition of unlabelled Rev. The half-maximal inhibitory concentration (IC_{50}) of the Rev protein tested (or its negative logarithm, pIC_{50}) provides an indirect measure of Impβ-binding affinity, with lower IC_{50} (higher pIC_{50}) values corresponding to stronger binding.

Representative assays are shown in **Figure 8b** and the results summarized in **Figure 8c-e** and **Table S1**. Wildtype (WT) Rev competes with the Rev-NLS peptide with an IC_{50} value of 0.5 μM, comparable to the K_d value of 0.8 μM obtained by ITC. Of the five mutants tested, three (R2-R4) yielded a large (10-fold) increase in IC_{50}, while the other two (R1, R5) gave more modest (3- and 1.7-fold) increases. The three mutants with the greatest effect localized to one end of the helical hairpin (the N-terminal half of helix α2 and the interhelical loop) and involved Arg clusters on both the A- and B-faces of Rev, whereas the mutants with more modest effects localized to the opposite end of the domain (helix α1 and the C-terminal half of helix α2) (**Figure 8a**), in line with our peptide-scanning data.

We then generated single point mutations of the nine Arg residues altered in mutants R2-R4. Four of these (R35D, R38D, R39D and R41D) significantly compromised the ability of Rev to compete with the Rev-NLS peptide, increasing the IC_{50} by a factor of 1.6-1.9 (versus a factor of 1.1-1.3 for the other mutants), corresponding to a drop in pIC_{50} of 0.20-0.27 (versus 0.04-0.11) (**Figure 8d,e**). Interestingly, the sum of the ΔpIC_{50} values (-0.71) for mutations R35D, R38D and R39D approached that observed for the corresponding R2 mutant (-1.0), indicating that the effect of the triple mutation primarily reflects the additive effects of the single mutations (**Figure 8e**). In contrast, the ΔpIC_{50} values for triple mutants R3 and R4 were 3-4 times larger in magnitude than the summed ΔpIC_{50} values of the corresponding single mutations, revealing that reversing the charge at these combined positions had a synergistic effect on disrupting binding. Notably, plotting the ΔpIC_{50} values of the nine R→D mutations onto the structure of Rev reveals a progressively greater effect of the mutation on Impβ binding as one proceeds along helix α2 towards its N-terminus and the interhelical loop (**Figure 8f**), consistent with the spatial trend noted above for the triple mutants. These findings point to the N-terminal tip of Rev helix α2 as a major binding epitope recognized by Impβ.

Charge-reversal mutations on Impβ and Rev show compensatory effects

We next sought to identify Impβ residues implicated in Rev recognition by introducing charge reversal mutations on the acidic inner surface of Impβ. Of the 57 Asp and Glu residues that localize to the protein's inner surface, we selected 22 that were prominently exposed or clustered into acidic patches for substitution by Arg

residues. To enhance the potential disruptive effect on Rev binding, we initially introduced these substitutions as double, triple or quadruple mutations (mutants B1-B7, **Figure 9a**). We then used our FP inhibition assay to compare the ability of WT Rev to displace the Rev-NLS peptide from WT or mutant forms of Impβ. Only minor changes in the IC_{50} value of Rev were observed when WT Impβ (IC_{50} = 0.5 μM) was replaced by mutants B1-B7 (IC_{50} = 0.43-0.55 μM) (**Figure 9b** and **Table S1**), consistent with the expectation that the Impβ mutations should not significantly alter Rev's ability to compete with the Rev-NLS peptide. We then repeated the assay using Rev double and triple mutants R1-R5. Strikingly, whereas the IC_{50} values of mutants R1 and R5 showed little variation in assays performed with WT or mutant Impβ, the IC_{50} values of mutants R2-R4 varied considerably (**Figure 9c** and **Table S1**). Most notably, these values decreased from ~5 μM when tested against WT Impβ to ~1.6 μM against mutant B2, signifying that Rev mutants R2-R4 competed with the Rev-NLS peptide with ~3 times greater efficacy when WT Impβ was replaced by the B2 mutant. More modest (1.5-1.8 fold) decreases in IC_{50} were also observed with mutants B1, B3 and B4. These findings reveal that the Asp/Glu→Arg substitutions on Impβ partly compensate for the adverse effects of the Rev Arg→Asp mutations on the ability of Impβ to bind Rev, suggesting a possible electrostatic interaction between the mutated acidic and basic residues (**Figure 9d**).

To assess the size of this compensatory effect and facilitate comparison across different combinations of Impβ and Rev mutants we used the parameter $\Delta\Delta pIC_{50}$, whose derivation is illustrated for mutants B2 and R4 in **Figure 9e**. In this figure, replacing WT Rev (solid black circles) by mutant R4 (solid red circles) causes the pIC_{50} to drop from 6.3 to 5.3, yielding a ΔpIC_{50} value of -1, whereas replacing both WT Rev and WT Impβ by mutants R4 and B2, respectively (red diamonds), results in a ΔpIC_{50} of only -0.5. Thus, compared to the former assay, the latter assay causes ΔpIC_{50} to shift from -1 to -0.5, yielding a $\Delta\Delta pIC_{50}$ value of +0.5. The $\Delta\Delta pIC_{50}$ values for selected combinations of Impβ and Rev mutants are shown in **Figures 9f** and **g**.

To pinpoint potentially interacting Impβ and Rev residues, we generated single charge-reversal mutations of the nine Impβ residues concerned by mutants B2-B4 and tested these in FP inhibition assays against Rev triple mutants R2-R4 and the corresponding Rev single point mutants shown in **Figure 8d**. The results revealed several combinations of Impβ and Rev mutants that gave notable compensatory effects (**Figure 9f,g** and **Figure S4**). These effects were generally more significant when either Impβ or Rev contained multiple mutations, as observed for example between Rev mutants R2-R4 and Impβ mutant D288R, between Rev mutant R3 and Impβ mutant D339R, and between Rev mutant R48D and Impβ mutants B3 and B4. Of the 51 pairwise combinations of single point mutants tested, the strongest compensatory effects were seen for Impβ mutant D288R with Rev mutants R42D and R46D ($\Delta\Delta pIC_{50}$ = 0.41 and 0.33, respectively) and to a lesser extent R43D ($\Delta\Delta pIC_{50}$ = 0.17) (**Figure 9g** and **Figure S4a-c**). $\Delta\Delta pIC_{50}$ values exceeding 0.1 were also observed for Impβ mutant E299R with Rev mutants R42D, R43D and R46D, for Impβ E289R with Rev R46D, and for Impβ E437R with Rev R48D (**Figure 9g** and **Figure S4d-g**). These findings identify Impβ acidic residues on the B helices of HEAT repeats 7 and 10 that are likely to be spatially proximal to basic residues on Rev helix α2 (**Figure 9h**), providing important clues into the relative binding orientations of Impβ and Rev.

Molecular docking simulations associate certain compensatory mutations with the C-site of Impβ

The mutual proximity of Rev residues Arg42 and Arg46 on helix α2 (**Figure 9h**) and the short distances (<18 Å) separating Impβ residue Asp288 from residues Glu299 and Glu437 relative to the width of the Rev helical hairpin (20-25 Å) make it likely that the compensatory effects summarized in **Figure 9h** are all mediated by the same Rev monomer. However, whether that Rev monomer localizes to the N- or C-site on Impβ is unclear. To explore this question, we performed molecular docking simulations to assess whether certain distance restraints could be attributed reliably to either site. We docked the Rev helical hairpin onto the Impβ surface (using the extended 1UKL conformation as suggested by our SAXS data) with the program HADDOCK, which exploits experimentally derived interaction restraints to guide the docking procedure[99]. (For clarity, in the following description of interaction restraints, a superscript is used to indicate whether a residue belongs to either Rev or Impβ). In initial simulations we specified distance restraints derived from compensatory mutations involving residue pairs $Asp288^{Imp\beta}$:$Arg42^{Rev}$ and $Asp288^{Imp\beta}$:$Arg46^{Rev}$ and combined these with BS3 crosslinking distance restraints for Rev bound to the N-site (experiment 1) or C-site (experiment 2) (**Table S2a**). Both simulations yielded similar scoring statistics (e.g., HADDOCK score, cluster size, buried surface area) for the top-ranked cluster of solutions (**Table S2b**). These solutions positioned Rev on the inner surface of Impβ next to HEAT repeats 1-8 (experiment 1) or repeats 7-19 (experiment 2), with Rev helix α2 oriented roughly antiparallel or parallel, respectively, to the Impβ superhelical axis (**Figure S5a,b**). (We refer hereafter to these two generic Rev orientations as "antiparallel" or "parallel", respectively). In both experiments the majority of docking solutions closely resembled the top-ranked solution with only minor variations in Rev orientation, indicating a relatively limited region of configurational space consistent with the docking parameters used (**Figure S6a,b**).

The above experiments included only two of the eight possible distance restraints suggested by our mutagenesis data (**Figure 9h**), and in particular omitted the distance restraint between residues $Glu437^{Imp\beta}$ and $Arg48^{Rev}$. Strikingly, however, experiment 2 (with Rev in the C-site) yielded solutions in which these two residues were relatively close to each another, including the 2nd ranked solution which positioned these residues <5 Å apart (**Figure S5b**). In contrast all solutions from experiment 1 (with Rev in the N-site) placed these residues far apart and sterically inaccessible to each other (**Figure S5a**). Repeating these simulations with the $Glu437^{Imp\beta}$:$Arg48^{Rev}$ interaction explicitly included as an additional distance restraint yielded significantly worse docking results for Rev bound to the N-site (experiment 3) and considerably improved results for Rev bound to the C-site (experiment 4) (**Figures S5c,d** and **S6c,d** and **Table S2**). These findings suggest that the set of electrostatic interactions used as distance restraints in experiments 3 and 4 are more compatible with the C-site than with the N-site (**Figure S5e**).

To confirm this hypothesis we performed exhaustive rigid-body sampling of Rev binding orientations on the surface of Impβ within the N-site and C-site envelopes depicted in **Figure 6f.** The Impβ structure was held fixed while the Rev helical hairpin was rotated and translated in a full 6-dimensional search that maintained $Lys20^{Rev}$ within BS3 crosslinking distance of the Group-1 lysines in HEAT repeats 1 and 2 (in the case of the N-site) or in repeat 19 (in the case of the C-site), yielding >3.5 and >2.8 million configurations, respectively. After

excluding configurations with a severe steric overlap between Impβ and Rev, the remaining configurations were used to measure Cβ-Cβ distances between specific Impβ and Rev residues. These measurements showed that the putative interactions illustrated in **Figure 9h** are individually compatible with Rev bound to either the N- or C-site, since docking configurations were identified for both sites that brought the relevant charged residues within salt-bridge distance of each other (**Figure S7a**). However, no configurations were identified for the N-site in which interactions involving residue pairs $Arg42^{Rev}$:$Asp288^{Imp\beta}$ and $Arg48^{Rev}$:$Glu437^{Imp\beta}$ could simultaneously occur (i.e., configurations in which both Cβ-Cβ distances were <14 Å), unlike the case for the C-site, where >80 such configurations were identified (**Figure S7b**). These findings confirm that the putative interactions depicted in **Figure 9h** have a much greater probability of being associated with the C-site than with the N-site.

In addition, the rigid-body sampling analysis also shed light on two of the BS3 crosslinks (between $Lys20^{Rev}$ and $Lys206^{Imp\beta}$ or $Lys537^{Imp\beta}$) that we previously were unable to assign reliably to either the N- or C-site. The minimal Cβ-Cβ distance measured between residues $Lys20^{Rev}$ and $Lys206^{Imp\beta}$ was 30.1 Å for all C-site configurations, compared to 12.5 Å for N-site configurations (**Figure S7a**), thereby assigning this crosslink with high probability to the N-site. Among the ~48,000 N-site configurations compatible with a $Lys20^{Rev}$:$Lys206^{Imp\beta}$ crosslink, only ~670 were compatible with a long crosslink between $Lys20^{Rev}$ and $Lys537^{Imp}$ (Cβ-Cβ distance between 28.9 and 30 Å) and none with a shorter crosslink (Cβ-Cβ distance <28.9 Å) (**Figure S7c**). In contrast the minimal Cβ-Cβ distance for this residue pair was 13.7 Å for C-site configurations (**Figure S7a**), making it highly likely that the $Lys20^{Rev}$:$Lys537^{Imp\beta}$ crosslink is associated with the C-site.

Structural model of Rev bound to the Impβ C-site

We next filtered the results of our rigid-body docking experiment for configurations of Rev bound to the Impβ C-site that were compatible with our BS3 crosslinking and compensatory mutagenesis data (**Figure 10a**). This resulted in a highly uniform set of parallel Rev configurations (**Figure 10b** and **Table S3).** To confirm this result, we ran HADDOCK using the BS3 crosslinking and compensatory mutagenesis data as interaction restraints to guide the docking procedure. Nearly all (97%) of the solutions obtained in the final refinement stage of docking yielded parallel configurations belonging to 7 very similar clusters (**Figure 10c** and **Table S4**) that closely resembled the configurations obtained by rigid body docking (**Figure 10d**). As expected, both docking approaches yielded models that agreed well with our BS3 crosslinking data (**Figure S8a**) and placed residues related by compensatory charge-reversal effects in close proximity (**Figure S8b**). To represent the entire ensemble of configurations obtained by rigid-body sampling and HADDOCK we calculated an "average" configuration that minimizes the change in orientation and position relative to each of the individual docking solutions obtained (**Figure 10e**). The resulting model closely resembles the top-ranked configurations obtained by rigid-body sampling and by HADDOCK (**Figure S9**). Comparing the individual docking configurations to the average model shows that the predicted location (i.e., translational coordinates of the centroid) of Rev on Impβ (**Figure 10e**, lower inset) is relatively well defined, as are two of the rotational angles (yaw and pitch), whereas the angle about the long axis of Rev (roll angle) is somewhat more variable. A SAXS curve calculated from the average model

agreed reasonably well (χ^2=1.93) with the SAXS data measured from the Impβ/Rev^{OD}Δ complex (**Figure S10a**). A near-perfect match between the calculated and observed scattering (χ^2=0.87) was obtained using a slightly more elongated conformation of Impβ, obtained by normal mode perturbation along the lowest-frequency vibrational mode (**Figure S10b**,**c**). Overall, these findings support a general binding mode whereby the Rev helical hairpin is embraced by the C-terminal half of Impβ with the long axis of Rev roughly parallel to the Impβ superhelical axis and the Rev CTD positioned near the C-terminal end of the HEAT-repeat superhelix.

Implications for Rev bound at the Impβ N-site

We next wondered how the second Rev monomer might bind to Impβ. Since Rev is known to multimerize through homotypic "A-A" and "B-B" interactions, we structurally aligned both types of Rev homodimer with our average model of Rev bound to the Impβ C-site to see whether either dimer was compatible with the model (**Figure S11a,b**). Both alignments resulted in an implausible placement of the second Rev monomer: residue Lys20 on Rev helix α1 was very far from Group-1 lysines in HEAT repeats 1 and 2, in disagreement with the BS3 crosslinking data; there was insufficient space to accommodate the Rev CTD; and, in the case of the Rev B-B homodimer, a severe steric clash occurred with Impβ (**Figure S11b**). These findings make it highly unlikely that the second Rev monomer binds Impβ by forming an A-A or B-B interface with the first bound Rev monomer. This conclusion is strongly supported by the ability of Impβ to bind two monomers of the Rev^{OD} mutant (**Figures 1f, 2b-d** and **3b**), whose ability to form A-A and B-B dimers is compromised (**Figure 1c**).

For additional insights we further exploited our rigid-body docking data. We structurally aligned the average model of Rev bound to the Impβ C-site with all Rev docking configurations that were compatible with BS3 crosslinks involving the N-site (including the K20^{Rev}-K206^{Impβ} crosslink; ~48,000 configurations; **Figure S7c**) and retained those exhibiting no, or at most only a mild, steric overlap between the two Rev monomers (~18,000 configurations). As expected, the resulting Rev monomers all localize to the N-terminal half of Impβ, with approximately one-third located within contact distance of the Rev monomer bound at the C-site (**Figure S11c**). The latter subset is characterized by two predominant configurations in which the two Rev monomers interact through their α1-α2 loops, with the second Rev monomer adopting either an anti-parallel or a roughly perpendicular ("transverse") orientation (**Figure S11d**). Interestingly, in the latter case, the relative orientation of the two Rev monomers closely resembles the C-C homodimer arrangement recently described for Rev[67]. Indeed, aligning the structure of a Rev C-C homodimer (PDB 5DHV) with our average model of Rev bound to the C-site yields only a minor steric clash with Impβ that can be relieved by a small rigid-body rotation of the dimer (**Figure S11e**), suggesting that Impβ could plausibly bind two Rev monomers sharing a C-C interface. However, additional data are required to establish the specific binding configuration of the second Rev monomer and whether it interacts with the first.

DISCUSSION

In this study we investigated the interaction of Impβ with HIV-1 Rev using diverse experimental and computational approaches. Gel shift and native MS experiments show that Impβ can bind either one or two Rev monomers, with ITC measurements reporting an approximately 10-fold difference in binding affinity between the two monomers. Crosslinking/MS data reveal that the two Rev binding sites localize to N- and C-terminal regions (denoted the "N-site" and "C-site") of the HEAT-repeat superhelix, which according to SAXS data adopts an extended conformation upon binding Rev. ITC and NMR analyses indicate that Impβ primarily recognizes the helical hairpin domain of Rev, while peptide scanning and mutagenesis results identify the N-terminal tip of Rev helix α2 and preceding loop, constituting the first half of the ARM, as a key binding epitope. Charge-reversal (Arg→Asp) mutations made in this region impair the ability of Rev to displace a Rev-NLS peptide from Impβ, and this effect is countered by converse (Asp/Glu→Arg) point mutations on Impβ, suggesting close spatial proximity of the mutated Impβ and Rev residues. Molecular docking simulations informed by the above data suggest a specific structural model for Rev bound at the C-site of Impβ and provide hints about the possible binding mode of Rev at the N-site.

Impβ is conformationally highly flexible, allowing it to adapt to and recognize diverse cargos through different regions of its HEAT repeat array. Whereas a C-terminal Impβ region (approximately repeats 7-19) recognizes the IBB domains of Impα[48] and Snurportin1[61] and the DNA-binding domain of SREBP-2[59], a middle region (repeats 5-14) recognizes Snail1, and an N-terminal region (approximately repeats 1-11) recognizes cyclin B1[100], ribosomal protein L23[101], and the parathyroid hormone-related protein PTHrP[58]. The *in vitro* observation that cyclin B1 and the IBB domain of Impα can both bind to the same Impβ molecule demonstrates that Impβ can use distinct N- and C-terminal regions to bind two molecular cargos simultaneously[100]. Our finding that Impβ contains two Rev binding sites involving opposite ends of the HEAT-repeat superhelix is consistent with this observation and agrees with a study reporting that Rev interacts with both an N-terminal (res. 1-396) and a C-terminal (res. 304-876) fragment of Impβ[44]. Interestingly, the alternative Rev import receptor Transportin also contains two independent Rev binding sites[44]. Conceivably, the ability of these import receptors to bind two Rev molecules might accelerate the nuclear import rate of Rev or increase its half-life in the cell[102], which could potentially modulate nuclear or cytosolic functions of Rev.

The ability of Impβ to engage two Rev monomers is in striking contrast with previously reported structures of Impβ/cargo complexes. These show Impβ loaded with only a single copy of its cargo[48,58,61,63], except for SREBP-2, which binds Impβ as a pre-formed dimer[59]. Unlike SREBP-2, the independent binding of Rev monomers revealed by gel shift and ITC analysis show that Impβ does not associate with a pre-formed Rev dimer but binds one Rev molecule at a time, reminiscent of how the RRE binds Rev[65,90]. Our structural model of Rev bound at the C-site of Impβ overlaps sterically with that of RanGTP bound to the N-terminal half of Impβ (**Figure S12a**; presumably a Rev monomer bound at the N-site would also strongly overlap with RanGTP), accounting for the ability of RanGTP to compete with Rev for binding to Impβ[43,44]. Previous studies identified residues within the ARM of Rev (res. 35-50) as important for binding Impβ[42-44]. Our mutagenesis and peptide-scanning experiments are consistent with

these reports and further delineate the first half of the ARM (res. 35-41) as being particularly important for binding affinity. These residues mediate key interactions with stem IIB and with an adjacent site on the RRE[70,74] as well as with stem IA[88], explaining why Impβ and the RRE bind Rev in a mutually exclusive fashion[42,43,103]. Our compensatory mutagenesis experiments identify four acidic residues of Impβ (Asp288, Glu289, Glu299 and Glu437) as being potentially in direct contact with the Rev ARM. Subsets of these residues also mediate recognition of the Impα IBB domain (Asp288), SREBP-2 (Glu299, Glu437) and Snail1 (Asp288, Glu289, Glu437), suggesting that Rev exploits the same Impβ residues used to import cellular cargos.

Among the structurally characterized Impβ/cargo complexes, the predicted binding mode of Rev at the C-site most closely resembles that of the Impα IBB domain (**Figure S12b**). The long C-terminal helix of the IBB domain adopts a similar parallel orientation as that predicted for Rev helix α2 but is shifted by 25 Å towards the C-terminal HEAT repeats, where it makes salt bridge interactions with the B helices of repeats 12-18. Besides its C-terminal helix, the IBB also contains an N-terminal extended moiety that makes extensive van der Waals and hydrogen-bonding interactions with the B helices of HEAT repeats 7-11. N-terminal truncations that remove one or both of the highly conserved Arg13 and Lys18 residues in this moiety severely abrogate binding to Impβ[104,105]. The Arg residues at the tip of Rev helix α2 in our structural model overlap closely with the N-terminal IBB moiety **(Figure S12c)**, accounting for the ability of the IBB domain to compete with Rev for binding to Impβ[43] and raising the possibility that key basic residues in these two molecular cargos mediate similar interactions with Impβ.

Rev is known to multimerize through hydrophobic A-A and B-B interfaces[69,73,85,106], which above a critical Rev concentration (~6 μM) mediate the *in vitro* assembly of helical filaments[84,107-109]. Our molecular docking simulations and the observation that Impβ binds two monomers of the oligomerization-defective Rev^{OD} mutant argue against the formation of an A-A or B-B interface between the two Impβ-bound Rev monomers. Interestingly, however, our data are compatible with the two bound Rev monomers sharing a weaker C-C interface, previously observed in Rev filaments and in crystallization contacts[67,68]. In the Rev-RRE assembly pathway, a C-C interaction between two RRE-bound Rev dimers has been proposed to mediate a "four-Rev" specificity checkpoint that precedes the association of additional Rev molecules[67,110]. This raises the speculative possibility that a C-C interface mediating recruitment of the second Rev monomer to Impβ might compete with the checkpoint C-C interaction and potentially regulate Rev-RRE assembly/disassembly.

In conclusion, our findings reveal that the interaction of Rev with Impβ is unlike that of cellular Impβ cargos in that it can interact through two independent binding sites that localize to the N- and C-terminal halves of the import receptor. Additional studies are needed to uncover the atomic details underlying Rev recognition by Impβ and the functional implications of the atypical binding stoichiometry.

METHODS

Protein expression and purification

Impβ. N-terminally His-tagged Impβ was expressed from a pETM-11 based plasmid in *E. coli* strain BL21 (DE3). Cells were grown in LB medium containing kanamycin (50 μg/ml) at 37°C until an OD_{600} of 0.5 - 0.6, and protein

expression was induced with 0.5 mM IPTG for 5 h at 30°C. Harvested cells were lysed by sonication at 4°C in lysis buffer [50 mM HEPES pH 7.5, 150 mM NaCl, 10 mM imidazole, 10 mM $MgCl_2$, 2 mM β-mercaptoethanol (BME)] supplemented with nuclease and protease inhibitors [10 µg/ml DNase I, 10 µg/ml RNase A, 2 mM PMSF and 1 tablet/80 ml of complete EDTA-free protease inhibitor cocktail (Roche)]. The soluble fraction was recovered by centrifugation (50,000 g, 20 min, 4°C) and applied to a HisTrap FF NiNTA column (GE Healthcare). After washing with buffer A (50 mM HEPES pH 7.5, 200 mM NaCl, 2 mM BME) containing 40 mM imidazole, proteins were eluted in the same buffer containing 300 mM imidazole. For untagged Impβ the His-tag was then removed by overnight incubation with His-tagged TEV protease (1:20 w/w TEV/Impβ protein ratio) supplemented with 2 mM BME and 2 mM PMSF and recovered in the flow-through of a new NiNTA column. Both untagged and His-tagged Impβ were further purified using a Superdex 200 16/60 gel filtration column (GE Healthcare) in 50 mM HEPES pH 7.5, 100 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP). Fractions containing pure Impβ were pooled, concentrated on an Amicon centrifugal filter (30 kDa cutoff, Millipore) and stored as aliquots at -80°C. The Impβ protein concentration was determined using a molar extinction coefficient of 79,051 $M^{-1}cm^{-1}$, which was experimentally determined by quantitative amino acid analysis.

Impβ charge reversal point mutations were generated by a PCR-based protocol adapted from the QuikChange site-directed mutagenesis method (Agilent Technologies). Impβ mutant proteins were expressed and purified as the WT and their molecular masses verified by LC/ESI mass spectrometry (**Table S5**).

HIV-1 Rev. Rev proteins were expressed from a pET-28a based plasmid as an N-terminal His-GB1-TEV fusion as previously described[111]. Protein expression was obtained from *E. coli* BL21 (DE3) cells grown in Auto-Induction Medium[112] containing kanamycin (50 µg/mL) for 12 h at 27°C. Harvested cells were lysed by sonication at 4°C, in lysis buffer (25 mM HEPES pH 7.5, 200 mM NaCl, 100 mM Na_2SO_4, 10 mM $MgCl_2$, 10 mM imidazole, 0.1 % Tween 20) containing nuclease and protease inhibitors (as described above for Impβ). The clarified lysate was incubated with RNAse T1 (20 U/µL, Roche), RNase A (20 µg/ml, Euromedex) and 2 M NaCl for 1.5 h at room temperature. After centrifugation (50,000 g, 10 min, 4°C), the soluble fraction was applied to a HisTrap FF NiNTA column (GE Healthcare) and washed first with high-salt buffer B (2 M NaCl, 50 mM Tris pH 8, 0.1 % Tween 20, 2 mM BME and 10 mM imidazole) followed by low-salt buffer B (including 250 mM NaCl and no Tween 20) containing 40 mM imidazole. Proteins were eluted in buffer B containing 300 mM imidazole, and diluted with an equal volume of 50 mM Tris pH 8 before loading onto HiTrap Heparin HP columns (GE Healthcare). After extensive washing with 50 mM Tris pH 8, 125 mM NaCl, elution was performed using 50 % High Salt Buffer (HSB) (50 mM Tris pH 8, 200 mM NaCl, 400 mM $(NH4)_2SO_4$, 100 mM Na_2SO_4, 2 mM BME) and 50 % of 2 M NaCl. The His-GB1 tag was then removed from the Rev protein by an overnight incubation with His-tagged TEV protease (1:20 w/w TEV/Rev protein ratio) supplemented with 2 mM BME and 2 mM PMSF. The untagged Rev protein recovered in the flow-through after passing the sample through a new NiNTA column, was then diluted with 5 volumes of 50 mM Tris pH 8 and then concentrated by repeating the heparin chromatography step described above. Heparin elution fractions containing Rev were applied onto a Superdex 75 16/60 gel filtration column (GE Healthcare) in HSB

containing 1 mM TCEP. Fractions containing pure Rev were pooled, concentrated on an Amicon centrifugal filter (10 kDa cutoff, Millipore) and stored at -80°C. Rev protein concentration was determined using a molar extinction coefficient of 9476 $M^{-1}cm^{-1}$, which was experimentally determined by quantitative amino acid analysis.

$Rev^{OD}\Delta$ (Rev4-69 V16D I55N) and Rev charge-reversal point mutations were generated by a PCR-based protocol adapted from the QuikChange site-directed mutagenesis method (Agilent Technologies). Rev mutant proteins were expressed and purified as the WT and their molecular masses verified by LC/ESI mass spectrometry (**Table S5**).

$^{15}N,^{13}C$-labeled Rev^{OD} protein. *E. coli* BL21 (DE3) cells expressing Rev^{OD} were grown in M9/H_2O medium containing 10 g/L $Na_2HP0_4.7H_2O$, 3 g/L KH_2PO_4, 0.5 g/L NaCl, 1 g/L $^{15}NH_4Cl$, 2 g/L ^{13}C D-glucose, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.01 mM $MnCl_2$, 0.05 mM $ZnS0_4$, 1X MEM vitamin solution (ThermoFisher Scientific) and kanamycin (50 µg/ml). Precultures were started from a glycerol stock in LB + kanamycin, 180 rpm, 37 °C. After 8 h, cells were collected for 15 min at 3,000 rpm and inoculated in M9/^{15}N/^{13}C medium and grown overnight at 37 °C. The following day, precultures were inoculated in M9/^{15}N/^{13}C medium (1:35) + kanamycin and grown at 37 °C until OD_{600} reach 0.7. Protein induction was obtained by adding 1 mM IPTG and growing the culture overnight at 16 °C. Cells were harvested by centrifugation and resuspended in denaturating lysis buffer (8 M urea, 50 mM NaH_2PO_4/Na_2HPO_4 pH 7.4, 500 mM NaCl, 5 mM β-mercaptoethanol, 0.02 % NaN_3, 25 mM imidazole) supplemented with 1 mM PMSF and one tablet of complete EDTA free (Roche). After sonication at 4°C, lysed cells were centrifuged at 50,000 g, 4 °C for 30 min. The supernatant containing $^{15}N,^{13}C$ -labeled His-GFP-TEV-Rev^{OD} denaturated protein was loaded on His Trap FF nickel column. After washing with denaturating lysis buffer at 5 ml/min, the flow rate was decreased to 1 ml/min and Rev refolding obtained by applying a linear gradient from 100 % lysis buffer to 100 % refolding buffer (50 mM NaH_2PO_4/Na_2HPO_4 pH 7.4, 500 mM NaCl, 5 mM β-mercaptoethanol, 0.02 % NaN_3). After an additional wash of refolding buffer, refolded protein was eluted with elution buffer (50 mM NaH_2PO_4/Na_2HPO_4 pH 7.4, 500mM NaCl, 5mM β-mercaptoethanol, 0.02 % azide, 500 mM imidazole). The His-GFP-TEV fusion was then removed from the Rev^{OD} protein using His-tagged TEV protease (1:20 w/w TEV:Rev protein ratio), by overnight incubation at 4 °C in dialysis buffer (50 mM NaH_2PO_4/Na_2HPO_4 pH 7.4, 150 mM KCl, 1 mM DTT, 1 mM EDTA, 0.02 % NaN_3) using 10 kDa cutoff dialysis membrane. The untagged $^{15}N,^{13}C$-labeled Rev^{OD} protein was recovered in the flow-through after passing the sample through a new His Trap FF column, centrifugated (50,000 g, 30 min, 4 °C) and concentrated by ultrafiltration (Amicon, 10 kDa cutoff,) before further purifying on a Superdex S75 16/60 gel filtration column (GE Healthcare) in 25 mM NaH_2PO_4/Na_2HPO_4 pH 7.4, 150 mM KCl, 1 mM DTT, 1 mM EDTA, 0.02 % NaN_3. The Rev^{OD} protein was dissolved in 90 % H_2O/10 % D_2O with gel filtration buffer and the protein concentration was adjusted to 90 µM in 150 µL in the absence or presence of Impβ (at equimolar ratio)

TEV protease. N-terminally His-tagged TEV protease bearing the point mutation S219V was expressed and purified as previously described[111].

SEC/MALLS

SEC was performed on a high-pressure liquid chromatography system equipped with an LC-20AD pump, an autosampler SIL20-ACHT storing the samples at 4°C, a communication interface CBM-20A (Shimadzu, Kyoto, Japan), using a Superdex 75 10/300 GL column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden), thermostated at 20°C in an oven XL-Therm (WynSep, Sainte Foy d'Aigrefeuille, France), and equilibrated with 200 mM NaCl, 400 mM $(NH_4)_2SO_4$, 100 mM Na_2SO_4, 50 mM Tris pH 8 and 1 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP). Volumes of 50 µl of each sample (60 - 710 µM) were injected at a flow rate of 0.5 ml/min. The elution profiles were followed on-line at 280 nm (SPD-M20A Shimadzu) by refractive index, static and dynamic light scattering using a laser emitting at 658 nm (Optilab rEX, miniDAWN TREOS and Dynapro Nanostar detectors, respectively, Wyatt Technology, Santa-Barbara, USA). Data were analysed using ASTRA V 5.3.4.20 software (Wyatt Technology). The basic equation for light scattering, restricted to diluted solutions of small particles (< 20 nm), reduces to I = KcM × (∂n/∂c)2, where I is the excess intensity of scattered light, K is an optical parameter, c is the weight concentration, and ∂n/∂c is the refractive index increment. The concentration was evaluated on-line from the refractive index signal, combined, for two-component analysis, with the 280 nm absorbance. From the amino acid composition and using Sedfit (https://sedfitsedphat.nibib.nih.gov/), we obtained ∂n/∂c of 0.184 mL g^{-1} and extinction coefficients of 0.647 L g^{-1} cm^{-1}.

Differential scanning fluorimetry

HIV-1 consensus B Rev (15-mer) peptides were obtained from NIH AIDS Research and Reference Reagent Program (Cat.# 6445). Peptides were dissolved at 5 mM in water or 1-80 % DMSO, as recommended. DSF experiments were performed on a CFX96 touch real-time PCR detection system (Bio-Rad) using 0.2 ml 8-tube PCR strips (#TBS0201) and optical Flat 8-cap strips (#TCS0803). Assay samples were done over a temperature range starting from 20 °C up to 100 °C, in increments of 0.5 °C, with 15 µM Impβ mixed with 75 µM Rev-peptide and 1x SYPRO Orange (Molecular Probes) diluted in H_2O. The melting temperature (T_m) values of Impβ in presence of each Rev-peptide were determined from the fluorescence first derivative of the melting curves using CFX Maestro software. The wavelengths used for SYPRO Orange excitation and emission were 470 nm and 570 nm, respectively.

Native gel shift assay

Impβ and Rev proteins were initially diluted in phosphate buffered saline (PBS) to 25 µM, and incubated in a 1:1 to 1:5 Impβ:Rev ratio (3.7 µM Impβ, 10 µL final volume) for 5 min. After adding 1 µL of native loading buffer (62.8 mM Tris HCl pH 6.8, 40 % glycerol, 0.01 % bromophenol blue), 2.5 µL of samples were analysed on a 10 % TGX (Biorad) gel, run under native conditions (0.5X TBE buffer, 4°C, 100 V, 120 min). Gels were stained with Coomassie blue and scanned on a ChemiDoc MP gel imaging system (Biorad).

Isothermal titration calorimetry

Calorimetric experiments were performed on a MicroCal iTC200 calorimeter (GE Healthcare, PA) at 20 °C while stirring at 800 rpm. All proteins were buffer exchanged by dialysis into 50 mM HEPES pH 7.5, 200 mM NaCl. Typically, 35-50 µM Rev^{OD} or $Rev^{OD}\Delta$ and 100-130 µM Impβ proteins were placed in the cell and syringe, respectively. Titrations consisted of 26 identical injections of 1.5 µl made at time intervals of 3 min. ITC data were analysed with NITPIC, SEDPHAT and GUSSI public-domain software packages[113] using a single or two binding site models. The first data point was excluded from the analysis. Experiments were done at least in duplicates, and the variability was estimated to be less than 5% in the binding enthalpy and 10% in both the binding affinity and the number of sites.

Glutaraldehyde crosslinking

5 µM Impβ was mixed with 0-20 µM Rev in 50 mM HEPES pH 7.5, 100 mM NaCl. Subsequently, 0.004% (v/v) of glutaraldehyde (Sigma, ref. G7776) was added and the reaction was incubated for 5 min at RT. The crosslinking reaction was stopped by adding 90 mM Tris pH 8 and the mixture analysed by native and denaturing gel electrophoresis and mass spectrometry.

Mass spectrometry

Liquid Chromatography/Electrospray Ionization Mass Spectrometry (LC/ESI-MS)

LC/ESI-MS was performed on a 6210 LC-TOF spectrometer coupled to an HPLC system (Agilent Technologies). All solvents used were HPLC grade (Chromasolv, Sigma-Aldrich). Trifluoroacetic acid (TFA) was from Acros Organics (puriss., p.a.). Solvent A was 0.03 % TFA in water; solvent B was 95 % acetonitrile-5 % water-0.03 % TFA. Immediately before analysis, protein samples were diluted to a final concentration of 5 µM with solvent A and then desalted on a reverse phase-C8 cartridge (Zorbax 300SB-C8, 5 µm, 300 µm ID 5 mm, Agilent Technologies) at a flow rate of 50 µl/min for 3 min with 100 % solvent A and subsequently eluted with 70 % solvent B for MS detection. MS acquisition was carried out in the positive ion mode in the 300-3200 *m/z* range. MS spectra were acquired and the data processed with MassHunter workstation software (v. B.07.00, Agilent Technologies) and with GPMAW software (v. 7.00b2, Lighthouse Data, Denmark).

Matrix-Assisted Laser Desorption/Ionisation-Time-of-Flight Mass Spectrometry (MALDI-TOF MS)

MALDI-TOF mass spectra were measured with an Autoflex mass spectrometer (Bruker Daltonics, Bremen, Germany) operated in linear positive ion mode. External mass calibration of the instrument, for the *m/z* range of interest, was carried out using the monomeric (66.4 kDa) and dimeric (132.8 kDa) molecular ions of BSA (reference 7030, Sigma Aldrich) as calibrants. Before MS analysis the protein samples (concentration around 10 µM; before and after crosslinking) were submitted to buffer exchange against 20 mM Tris pH 7.5, 50 mM NaCl, using Vivaspin devices (Sartorius, cut-off of 30 kDa for Impβ alone and 10 kDa for Impβ/Rev). The buffer-

exchanged protein samples were then mixed in a 1:2, 1:5 or 1:10 (*v/v*) ratio with sinapinic acid matrix (Sigma; 10 mg/mL in water/acetonitrile/trifluoroacetic acid, 50/50/0.1, *v/v/v*) and 1 to 2 µL were deposited on the target and allowed to air dry at RT. Mass spectra data were processed with Flexanalysis software (v.3.0, Bruker Daltonics).

Native Mass Spectrometry

Impβ and Rev proteins were individually buffer exchanged by gel filtration using a Superdex200 5/150 (Impβ) or Superdex75 5/150 (Rev) column pre-equilibrated with either 250 mM (Impβ), 400 mM (Rev^{OD} and $Rev^{OD}\Delta$) or 800 mM (Rev WT) ammonium acetate. The individual proteins or mixtures of Impβ with Rev were incubated at 20°C for 15 min prior to MS analysis. Protein ions were generated using a nanoflow electrospray (nano-ESI) source. Nanoflow platinum-coated borosilicate electrospray capillaries were bought from Thermo Electron SAS (Courtaboeuf, France). MS analyses were carried out on a quadrupole time-of-flight mass spectrometer (Q-TOF Ultima, Waters Corporation, Manchester, U.K.) modified for the detection of high masses[114-116]. The following instrumental parameters were used: capillary voltage = 1.2-1.3 kV, cone potential = 40 V, RF lens-1 potential = 40 V, RF lens-2 potential = 1 V, aperture-1 potential = 0 V, collision energy = 30-140 V, and microchannel plate (MCP) = 1900 V. All mass spectra were calibrated externally using a solution of cesium iodide (6 mg/mL in 50% isopropanol) and processed with the Masslynx 4.0 software (Waters Corporation, Manchester, U.K.) and Massign software package[117].

LC-MS/MS analysis and BS3 crosslink identification

4.1 µM of His-tagged Impβ was mixed with 8.2 µM of WT Rev in 50 mM HEPES pH 7.5, 150 mM NaCl before adding 25 µM of bis(sulfosuccinimidyl)suberate (BS3) (Thermo Fisher, ref. 21580). The reaction was incubated for 30 min at room temperature and stopped with 25 mM NH_4HCO_3. The reaction mixture was then subjected to SDS-PAGE and the band corresponding to the crosslinked complex was excised from the Coomassie-stained gel. Disulfide bridges were reduced using dithiothreitol (Roche). Free thiol groups were subsequently alkylated with iodoacetamide. Proteins were digested with trypsin (Promega) at 37°C overnight. Reactions were stopped with formic acid (FA) and peptides were extracted from the gel with 5% formic acid (FA) in 50% acetonitrile. Acetonitrile content was reduced in a vacuum centrigufe prior to LC-MS/MS analysis. Peptides were analysed on an UltiMate 3000 HPLC RSLC nano system coupled to a Q Exactive HF-X quadrupole-orbitrap mass spectrometer, equipped with a Nanospray Flex ion source (all Thermo Fisher Scientific). Peptides were trapped on a PepMap C18 cartridge (5 mm × 300 µm ID, 5 µm particles, 100 Å pore size) and separated on a PepMap C18 column (500 mm × 75 µm ID, 2 µm, both Thermo Fisher Scientific) applying a linear gradient from 2% to 35% solvent B (80% acetonitrile, 0.08% FA; solvent A 0.1% FA) at a flow rate of 230 nL/min over 120 min. The mass spectrometer was operated in data-dependent mode: survey scans were obtained in a mass range of 350-1600 m/z with lock mass on, at a resolution of 60,000 at 200 m/z and an AGC target value of 3E6. The 10 most intense ions were collected subsequently with an isolation width of 1.6 Th for maximal 250 ms, and fragmented at 28% normalized collision

energy. Spectra were recorded at a target value of 2E4 and a resolution of 60,000. Peptides with a charge of +1,+2 or >+7 were excluded from fragmentation, the peptide match feature was set to preferred, the exclude isotope feature was enabled, and selected precursors were dynamically excluded from repeated sampling for 20 seconds within a mass tolerance of 5 ppm. Raw data were searched with pLink 2.3.5[118] against the sequences of the 16 most abundant protein hits from the MaxQuant search (version 1.6.0.16)[119], as well as the sequences of target proteins and common contaminant sequences. BS3 was selected as the crosslinking chemistry. Carbamidomethyl on Cys was set as fixed, oxidation of Met and protein N-terminal acetylation as variable modifications. Enzyme specificity was selected according to the protease used for digestion. Search results were filtered for 1% FDR on the PSM level limiting the precursor mass deviation to 5 ppm. To remove low quality peptide spectrum matches an additional e-Value cutoff of < 0.001 was applied.

NMR

Assignment was obtained using ^{15}N,^{13}C-labeled samples (90 μM) using BEST-TROSY three-dimensional experiments HNCO, intra-residue HN(CA)CO, HN(CO)CA, intra-residue HNCA, HN(COCA)CB, and intra-residue HN(CA)CB[120] recorded on a Bruker spectrometer equipped with a cryoprobe operating at 20 °C and a ^{1}H frequency of 600 and 700 MHz. Spectra were processed using NMRPipe[121] and analyzed in Sparky (Goddard, T. & Kneller, D. SPARKY 3. University of California, San Francisco). MARS[122] was used for spin system identification, combined with manual verification.

SEC-SAXS

Size exclusion chromatography coupled with small-angle x-ray scattering (SEC-SAXS) experiments were conducted on the BM29 beamline at the European Synchrotron Radiation Facility (Grenoble, France) using a Superdex 200 Increase 5/150 GL column. 30 μL of protein sample (containing 25 μM of ImpB with or without 125 μM of Rev), previously spun (21 000 *g*, 10 min), were injected and the flow rate maintained at 0.35 ml/min using 50 mM Tris buffer, pH 8, 100 mM NaCl. Acquisitions (1 sec / frame) were recorded during 10 and 1000 s for buffer and protein, respectively. Individual frames were processed using the software BsxCUBE, yielding individual radially averaged curves of normalized intensity versus scattering angle. Time frames corresponding to each elution peak were combined to give the average scattering curve for each measurement. The averaged spectrum of the buffer was subtracted from the averaged spectrum of protein before data analysis. Primary data reduction was performed and model-independent parameters R_g and D_{max} were determined using the program PRIMUS[123]. The program CRYSOL[124] was used to calculate the theoretical scattering from individual crystal structures. Normal mode perturbed models of Impβ (PDB 1UKL) were calculated for the five lowest-frequency vibrational modes using the ElNemo web server[125] with parameter values of DQMIN=-400 and DQMAX=400 for the amplitude range and of DQSTEP=20 for the increment.

Fluorescence polarization (FP) assays

Fluorescence polarization assays were performed as previously described[111] in black 384-well plates (Greiner ref. 781076, Greiner Bio-One) on a Clariostar plate reader (BMG Labtech) using a 540 ± 20 nm excitation filter, a 590 ± 20 nm emission filter and an LP565 dichroic mirror. The inhibition assays were carried out in a total volume of 10 µL, with 195 nM Impβ, 200 nM Rev-NLS-TAMRA peptide and 4.3 nM - 35 µM Rev variants. Curves obtained with each Impβ and Rev protein were independently fitted and the IC_{50} values were determined with GraphPad Prism 9 software using the equation $F=F_{min} + (F_{max} - F_{min})/(1+10^{((LogIC_{50}-X)*HillSlope)})$, where F_{min} *and* F_{max} are the fluorescence polarization values at saturating Rev concentration and in the absence of Rev, respectively, and X is the common logarithm of the Rev concentration.

Molecular Docking with HADDOCK

The helical hairpin domain of Rev (PDB 2X7L chain M) was docked onto Impβ (PDB 1UKL chain A) using the HADDOCK2.2 webserver[99]. HADDOCK uses biochemical information on interacting residues as ambiguous interaction restraints (AIRs) to drive the docking. Residues directly implicated or potentially involved in mediating the interaction are defined as "active" and "passive", respectively. Each AIR is defined between an active residue of one molecule and specific active or passive residues of the other. Passive residues for Rev were defined as all residues (res. 9-65) except those designated as active, while passive residues for Impβ were defined as all solvent-accessible residues on the concave inner surface of the protein. (See footnote 2 of **Table S2** for the explicit list of residues). The set of AIRs and corresponding cutoff distances for each docking experiment are listed in **Tables S2A** and **S4A**. For BS3 crosslinking restraints the cutoff distance between the Cβ atoms of crosslinked lysine residues was set at 30 Å. For AIRs between Rev Arg residues and Impβ Asp or Glu residues that yielded compensatory mutagenesis effects the Cβ-Cβ cutoff distance was set to a value between 5 and 8 Å depending on the associated $\Delta\Delta pIC_{50}$ value (**Figure 9f**), with shorter distances assigned to residue pairs yielding larger $\Delta\Delta pIC_{50}$ values. The HADDOCK docking protocol consists of randomization of orientations and rigid body energy minimization, semi-rigid simulated annealing in torsion angle space, and final refinement in Cartesian space with explicit solvent. For each docking experiment, a total of 1000 complex configurations were calculated in the randomization stage. The best 200 solutions were then used in the semi-rigid and final refinement stages, in which the side chain and backbone atoms of Impβ and Rev were allowed to move. The final structures were ranked according to HADDOCK score and clustered by pairwise backbone RMSD. Clusters were then ranked according to the average of the four best-scoring structures within each cluster. In all docking experiments the top-ranked cluster scored substantially better than the next ranked cluster and was also the largest in size.

Rigid body docking

Rigid body sampling was performed with the Rev helical hairpin domain (PDB 2X7L chain M) and Impβ (PDB 1UKL chain A) using in-house Unix shell and Fortran scripts combined with the program Lsqkab from the CCP4 suite[126].

Two independent docking jobs were run for Rev bound either to the N-site or the C-site of Impβ. For each job, the Impβ structure was held fixed while Rev was rotated and translated in a 6-dimensional search. The translational search was performed on a 4 Å grid spacing over grid points located within 35 Å of Cα atoms of the Impβ lysine residues that formed BS3 crosslinks with Rev Lys20 (Group-1 residues in **Figure 6b**; Lys residues 23, 62 and 68 In the case of the N-site, and Lys residues 854, 857, 859, 867 and 873 in the case of the C-site). In total 879 and 714 grid points were used for the N-site and C-site docking jobs, respectively. Rotations of Rev were performed with the Cβ atom of Rev Lys20 held fixed at the origin of the rotation, which was centered on each grid point. The rotational search employed a 15° polar angle step size (4056 orientations per grid point), resulting in >3.5 and >2.8 million configurations for the N- and C- sites, respectively.

For each docking configuration, the steric overlap between Rev and Impβ was estimated using a mask overlap procedure, as follows. A mask for Impβ was generated at 5 Å sampling resolution by (i) rounding the coordinates of each atom to the nearest multiple of 5 Å, (ii) sorting the resulting file on the rounded coordinates (i.e., sorting for increasing value of *x* then *y* then *z*) and (iii) removing duplicate coordinates from the sorted file. These same steps were then used to generate a mask for Rev for each docking configuration. For each configuration a combined mask covering both the Impβ and Rev structures was generated by concatenating the individual Impβ and Rev mask files, sorting the resulting file on the mask coordinates, and removing duplicate coordinates from the sorted file. The overlap between the Impβ and Rev masks was then given by $N_{overlap} = N_{\mathrm{Imp}\beta} + N_{\mathrm{Rev}} - N_{comb}$, where $N_{\mathrm{Imp}\beta}$, N_{Rev} and N_{comb} are the number of mask points in the Impβ, Rev and combined masks, respectively. N_{overlap} values above 15 were found to correspond to a large steric overlap between Impβ and Rev that could not be relieved by simply reorienting the side chains of these two proteins. Docking configurations yielding an $N_{\mathrm{overlap}} > 15$ were thus rejected from further analysis. The number of configurations that survived this steric overlap criterion was approximately 1.5 and 1.9 million for the N- and C-site, respectively. For each surviving configuration, Cβ-Cβ distances were calculated for the ten Impβ/Rev residue pairs indicated in **Figure S7A** and used to generate the boxplots and distance distribution diagrams shown in **Figures S7A** and **B.**

For the docking of Rev to the C-site of Impβ, configurations were subsequently filtered according to the distance cutoff criteria indicated in **Table S3A**, resulting in the 15 configurations shown in **Figure 10b**. These configurations were then ranked as follows. For each of the 8 Impβ/Rev residue pairs mediating compensatory mutation effects (left column of **Table S3A**) the minimal distance (D_{min}) between the Impβ and Rev residues was calculated over all sterically allowed side chain rotamers. (In a few cases where the two closest Impβ and Rev rotamers sterically overlapped the value of D_{min} was set to 2.3 Å). Each D_{min} value was then weighted by the corresponding $\Delta\Delta p\mathrm{IC}_{50}$ value (**Figure 9g**) and the weighted mean ($<\Delta\Delta p\mathrm{IC}_{50}*D_{\mathrm{min}}>$) was calculated over all eight residue pairs. Docking configurations were then ranked in increasing value of $<\Delta\Delta p\mathrm{IC}_{50}*D_{\mathrm{min}}>$, which ranged from 4.6 Å for the top-ranked solution to 7.31 Å for the solution ranked 15th.

Average Docking Model of Rev bound to the C-site of Impβ

The average configuration of Rev bound to the Impβ C-site was calculated from a total of 43 individual docking configurations. These consist of the 15 configurations obtained by rigid body docking (indicated in **Figure 10b** and **Table S3**) and 28 configurations (the 4 lowest energy structures from each of the 7 clusters) from the final HADDOCK experiment (**Figure 10c** and **Table S4**). The 43 configurations were superimposed onto a reference Impβ structure and the centroid of all 43 Rev structures was calculated. The 43 Rev structures were then translated so as to have their individual centroids coincident at this common centroid position. A Rev search model (PDB 2X7L chain M) was then centered on this position in an arbitrary orientation and a rotational search was performed using a 10° polar angle step size (12,996 orientations). For each orientation a distance score was calculated between representative atoms located along the three principal axes of the Rev search model and the corresponding atoms of the 43 individual Rev structures. The orientation yielding the lowest score was selected as the average docking configuration.

ACKNOWLEDGEMENTS

We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities. We thank M.E. Brennich and M. Tully for assistance in using beamline BM29, L. Shamseddine for help with protein purification, A. Le Roy for SEC/MALLS analysis, J.P. Andrieu for amino acid analysis, H. Dawi for initial HADDOCK experiments, A. Sinz, C. Ihling and M. Schäfer for exploratory XL-MS experiments, and T. Gossenreiter, D. Anrather and M. Hartl from the Mass Spectrometry Facility at Max Perutz Labs for BS3 crosslinking analyses, which were performed using the VBCF instrument pool. This study was supported by grants (to CP) from the ANRS and Fondation de France and PhD funding (to DS) from the ANRS and CEA. This work used the platforms of the Grenoble Instruct-ERIC center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-0005-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003).

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

REFERENCES

1 Sodroski, J., Goh, W. C., Rosen, C., Dayton, A., Terwilliger, E. & Haseltine, W. A second post-transcriptional trans-activator gene required for HTLV-III replication. *Nature* **321**, 412-417 (1986). doi.org/10.1038/321412a0

2 Feinberg, M. B., Jarrett, R. F., Aldovini, A., Gallo, R. C. & Wong-Staal, F. HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. *Cell* **46**, 807-817 (1986). doi.org/10.1016/0092-8674(86)90062-0

3 Pollard, V. W. & Malim, M. H. The HIV-1 Rev protein. *Annu Rev Microbiol* **52**, 491-532 (1998). doi.org/10.1146/annurev.micro.52.1.491

4 Karn, J. & Stoltzfus, C. M. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb Perspect Med* **2**, a006916 (2012). doi.org/10.1101/cshperspect.a006916

5 Fernandes, J. D., Booth, D. S. & Frankel, A. D. A structurally plastic ribonucleoprotein complex mediates post-transcriptional gene regulation in HIV-1. *Wiley Interdiscip Rev RNA* **7**, 470-486 (2016). doi.org/10.1002/wrna.1342
6 Wegener, M. & Muller-McNicoll, M. Nuclear retention of mRNAs - quality control, gene regulation and human disease. *Semin Cell Dev Biol* **79**, 131-142 (2018). doi.org/10.1016/j.semcdb.2017.11.001
7 Williams, T., Ngo, L. H. & Wickramasinghe, V. O. Nuclear export of RNA: Different sizes, shapes and functions. *Semin Cell Dev Biol* **75**, 70-77 (2018). doi.org/10.1016/j.semcdb.2017.08.054
8 Malim, M. H., Hauber, J., Le, S. Y., Maizel, J. V. & Cullen, B. R. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **338**, 254-257 (1989). doi.org/10.1038/338254a0
9 Emerman, M., Vazeux, R. & Peden, K. The rev gene product of the human immunodeficiency virus affects envelope-specific RNA localization. *Cell* **57**, 1155-1165 (1989). doi.org/10.1016/0092-8674(89)90053-6
10 Felber, B. K., Hadzopoulou-Cladaras, M., Cladaras, C., Copeland, T. & Pavlakis, G. N. rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc Natl Acad Sci U S A* **86**, 1495-1499 (1989). doi.org/10.1073/pnas.86.5.1495
11 Hadzopoulou-Cladaras, M., Felber, B. K., Cladaras, C., Athanassopoulos, A., Tse, A. & Pavlakis, G. N. The rev (trs/art) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a cis-acting sequence in the env region. *J Virol* **63**, 1265-1274 (1989).
12 Dayton, E. T., Powell, D. M. & Dayton, A. I. Functional analysis of CAR, the target sequence for the Rev protein of HIV-1. *Science* **246**, 1625-1629 (1989). doi.org/10.1126/science.2688093
13 Malim, M. H. & Cullen, B. R. HIV-1 structural gene expression requires the binding of multiple Rev monomers to the viral RRE: implications for HIV-1 latency. *Cell* **65**, 241-248 (1991). doi.org/10.1016/0092-8674(91)90158-u
14 Mann, D. A., Mikaelian, I., Zemmel, R. W., Green, S. M., Lowe, A. D., Kimura, T., Singh, M., Butler, P. J., Gait, M. J. & Karn, J. A molecular rheostat. Co-operative rev binding to stem I of the rev-response element modulates human immunodeficiency virus type-1 late gene expression. *J Mol Biol* **241**, 193-207 (1994). doi.org/10.1006/jmbi.1994.1488
15 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).
16 Fukuda, M., Asano, S., Nakamura, T., Adachi, M., Yoshida, M., Yanagida, M. & Nishida, E. CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* **390**, 308-311 (1997).
17 Neville, M., Stutz, F., Lee, L., Davis, L. I. & Rosbash, M. The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr Biol* **7**, 767-775 (1997). doi.org/10.1016/s0960-9822(06)00335-6
18 Kjems, J., Frankel, A. D. & Sharp, P. A. Specific regulation of mRNA splicing in vitro by a peptide from HIV-1 Rev. *Cell* **67**, 169-178 (1991). doi.org/10.1016/0092-8674(91)90580-r
19 Kjems, J. & Sharp, P. A. The basic domain of Rev from human immunodeficiency virus type 1 specifically blocks the entry of U4/U6.U5 small nuclear ribonucleoprotein in spliceosome assembly. *J Virol* **67**, 4769-4776 (1993).
20 Powell, D. M., Amaral, M. C., Wu, J. Y., Maniatis, T. & Greene, W. C. HIV Rev-dependent binding of SF2/ASF to the Rev response element: possible role in Rev-mediated inhibition of HIV RNA splicing. *Proc Natl Acad Sci U S A* **94**, 973-978 (1997).
21 Arrigo, S. J. & Chen, I. S. Rev is necessary for translation but not cytoplasmic accumulation of HIV-1 vif, vpr, and env/vpu 2 RNAs. *Genes Dev* **5**, 808-819 (1991). doi.org/10.1101/gad.5.5.808
22 D'Agostino, D. M., Felber, B. K., Harrison, J. E. & Pavlakis, G. N. The Rev protein of human immunodeficiency virus type 1 promotes polysomal association and translation of gag/pol and vpu/env mRNAs. *Mol Cell Biol* **12**, 1375-1386 (1992). doi.org/10.1128/mcb.12.3.1375
23 Perales, C., Carrasco, L. & Gonzalez, M. E. Regulation of HIV-1 env mRNA translation by Rev protein. *Biochim Biophys Acta* **1743**, 169-175 (2005). doi.org/10.1016/j.bbamcr.2004.09.030
24 Brandt, S., Blissenbach, M., Grewe, B., Konietzny, R., Grunwald, T. & Uberla, K. Rev proteins of human and simian immunodeficiency virus enhance RNA encapsidation. *PLoS Pathog* **3**, e54 (2007).
25 Blissenbach, M., Grewe, B., Hoffmann, B., Brandt, S. & Uberla, K. Nuclear RNA export and packaging functions of HIV-1 Rev revisited. *J Virol* **84**, 6598-6604 (2010). doi.org/10.1128/JVI.02264-09

26 Grewe, B., Ehrhardt, K., Hoffmann, B., Blissenbach, M., Brandt, S. & Uberla, K. The HIV-1 Rev protein enhances encapsidation of unspliced and spliced, RRE-containing lentiviral vector RNA. *PLoS One* **7**, e48688 (2012). doi.org/10.1371/journal.pone.0048688

27 Lata, S., Ali, A., Sood, V., Raja, R. & Banerjea, A. C. HIV-1 Rev downregulates Tat expression and viral replication via modulation of NAD(P)H:quinine oxidoreductase 1 (NQO1). *Nat Commun* **6**, 7244 (2015). doi.org/10.1038/ncomms8244

28 Rosenbluh, J., Hayouka, Z., Loya, S., Levin, A., Armon-Omer, A., Britan, E., Hizi, A., Kotler, M., Friedler, A. & Loyter, A. Interaction between HIV-1 Rev and integrase proteins: a basis for the development of anti-HIV peptides. *J Biol Chem* **282**, 15743-15753 (2007). doi.org/10.1074/jbc.M609864200

29 Levin, A., Hayouka, Z., Friedler, A. & Loyter, A. Nucleocytoplasmic shuttling of HIV-1 integrase is controlled by the viral Rev protein. *Nucleus* **1**, 190-201 (2010). doi.org/10.4161/nucl.1.2.11300

30 Jager, S., Cimermancic, P., Gulbahce, N., Johnson, J. R., McGovern, K. E., Clarke, S. C., Shales, M., Mercenne, G., Pache, L., Li, K. *et al.* Global landscape of HIV-human protein complexes. *Nature* **481**, 365-370 (2011). doi.org/10.1038/nature10719

31 Hadian, K., Vincendeau, M., Mausbacher, N., Nagel, D., Hauck, S. M., Ueffing, M., Loyter, A., Werner, T., Wolff, H. & Brack-Werner, R. Identification of a heterogeneous nuclear ribonucleoprotein-recognition region in the HIV Rev protein. *J Biol Chem* **284**, 33384-33391 (2009). doi.org/10.1074/jbc.M109.021659

32 Robertson-Anderson, R. M., Wang, J., Edgcomb, S. P., Carmel, A. B., Williamson, J. R. & Millar, D. P. Single-molecule studies reveal that DEAD box protein DDX1 promotes oligomerization of HIV-1 Rev on the Rev response element. *J Mol Biol* **410**, 959-971 (2011). doi.org/10.1016/j.jmb.2011.04.026

33 Naji, S., Ambrus, G., Cimermancic, P., Reyes, J. R., Johnson, J. R., Filbrandt, R., Huber, M. D., Vesely, P., Krogan, N. J., Yates, J. R., 3rd *et al.* Host cell interactome of HIV-1 Rev includes RNA helicases involved in multiple facets of virus production. *Mol Cell Proteomics* **11**, M111 015313 (2012). doi.org/10.1074/mcp.M111.015313

34 Lamichhane, R., Hammond, J. A., Pauszek, R. F., 3rd, Anderson, R. M., Pedron, I., van der Schans, E., Williamson, J. R. & Millar, D. P. A DEAD-box protein acts through RNA to promote HIV-1 Rev-RRE assembly. *Nucleic Acids Res* **45**, 4632-4641 (2017). doi.org/10.1093/nar/gkx206

35 Hammond, J. A., Zhou, L., Lamichhane, R., Chu, H. Y., Millar, D. P., Gerace, L. & Williamson, J. R. A Survey of DDX21 Activity During Rev/RRE Complex Formation. *J Mol Biol* **430**, 537-553 (2018). doi.org/10.1016/j.jmb.2017.06.023

36 Pemberton, L. F. & Paschal, B. M. Mechanisms of receptor-mediated nuclear import and nuclear export. *Traffic* **6**, 187-198 (2005).

37 Stewart, M. Molecular mechanism of the nuclear protein import cycle. *Nat Rev Mol Cell Biol* **8**, 195-208 (2007).

38 Lott, K. & Cingolani, G. The importin beta binding domain as a master regulator of nucleocytoplasmic transport. *Biochim Biophys Acta* **1813**, 1578-1592 (2011). doi.org/10.1016/j.bbamcr.2010.10.012

39 Chook, Y. M. & Suel, K. E. Nuclear import by karyopherin-betas: recognition and inhibition. *Biochim Biophys Acta* **1813**, 1593-1606 (2011).

40 Kimura, M. & Imamoto, N. Biological significance of the importin-beta family-dependent nucleocytoplasmic transport pathways. *Traffic* **15**, 727-748 (2014). doi.org/10.1111/tra.12174

41 Christie, M., Chang, C. W., Rona, G., Smith, K. M., Stewart, A. G., Takeda, A. A., Fontes, M. R., Stewart, M., Vertessy, B. G., Forwood, J. K. *et al.* Structural Biology and Regulation of Protein Import into the Nucleus. *J Mol Biol* **428**, 2060-2090 (2016). doi.org/10.1016/j.jmb.2015.10.023

42 Truant, R. & Cullen, B. R. The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importin beta-dependent nuclear localization signals. *Mol Cell Biol* **19**, 1210-1217 (1999).

43 Henderson, B. R. & Percipalle, P. Interactions between HIV Rev and nuclear import and export factors: the Rev nuclear localisation signal mediates specific binding to human importin-beta. *J Mol Biol* **274**, 693-707 (1997).

44 Arnold, M., Nath, A., Hauber, J. & Kehlenbach, R. H. Multiple importins function as nuclear transport receptors for the Rev protein of human immunodeficiency virus type 1. *J Biol Chem* **281**, 20883-20890 (2006).

45 Gu, L., Tsuji, T., Jarboui, M. A., Yeo, G. P., Sheehy, N., Hall, W. W. & Gautier, V. W. Intermolecular masking of the HIV-1 Rev NLS by the cellular protein HIC: novel insights into the regulation of Rev nuclear import. *Retrovirology* **8**, 17 (2011).
46 Conti, E., Uy, M., Leighton, L., Blobel, G. & Kuriyan, J. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell* **94**, 193-204 (1998).
47 Conti, E. & Kuriyan, J. Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. *Structure* **8**, 329-338 (2000).
48 Cingolani, G., Petosa, C., Weis, K. & Muller, C. W. Structure of importin-beta bound to the IBB domain of importin-alpha. *Nature* **399**, 221-229 (1999).
49 Forwood, J. K., Lange, A., Zachariae, U., Marfori, M., Preast, C., Grubmuller, H., Stewart, M., Corbett, A. H. & Kobe, B. Quantitative structural analysis of importin-beta flexibility: paradigm for solenoid protein structures. *Structure* **18**, 1171-1183 (2010).
50 Tauchert, M. J., Hemonnot, C., Neumann, P., Koster, S., Ficner, R. & Dickmanns, A. Impact of the crystallization condition on importin-beta conformation. *Acta Crystallogr D Struct Biol* **72**, 705-717 (2016). doi.org/10.1107/S2059798316004940
51 Vetter, I. R., Arndt, A., Kutay, U., Gorlich, D. & Wittinghofer, A. Structural view of the Ran-Importin beta interaction at 2.3 A resolution. *Cell* **97**, 635-646 (1999).
52 Lee, S. J., Matsuura, Y., Liu, S. M. & Stewart, M. Structural basis for nuclear import complex dissociation by RanGTP. *Nature* **435**, 693-696 (2005).
53 Forwood, J. K., Lonhienne, T. G., Marfori, M., Robin, G., Meng, W., Guncar, G., Liu, S. M., Stewart, M., Carroll, B. J. & Kobe, B. Kap95p binding induces the switch loops of RanGDP to adopt the GTP-bound conformation: implications for nuclear import complex assembly dynamics. *J Mol Biol* **383**, 772-782 (2008). doi.org/10.1016/j.jmb.2008.07.090
54 Bayliss, R., Littlewood, T. & Stewart, M. Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. *Cell* **102**, 99-108 (2000).
55 Bayliss, R., Littlewood, T., Strawn, L. A., Wente, S. R. & Stewart, M. GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta. *J Biol Chem* **277**, 50597-50606 (2002).
56 Liu, S. M. & Stewart, M. Structural basis for the high-affinity binding of nucleoporin Nup1p to the Saccharomyces cerevisiae importin-beta homologue, Kap95p. *J Mol Biol* **349**, 515-525 (2005). doi.org/10.1016/j.jmb.2005.04.003
57 Ivic, N., Potocnjak, M., Solis-Mezarino, V., Herzog, F., Bilokapic, S. & Halic, M. Fuzzy Interactions Form and Shape the Histone Transport Complex. *Mol Cell* **73**, 1191-1203 e1196 (2019). doi.org/10.1016/j.molcel.2019.01.032
58 Cingolani, G., Bednenko, J., Gillespie, M. T. & Gerace, L. Molecular basis for the recognition of a nonclassical nuclear localization signal by importin beta. *Mol Cell* **10**, 1345-1353 (2002).
59 Lee, S. J., Sekimoto, T., Yamashita, E., Nagoshi, E., Nakagawa, A., Imamoto, N., Yoshimura, M., Sakai, H., Chong, K. T., Tsukihara, T. *et al.* The structure of importin-beta bound to SREBP-2: nuclear import of a transcription factor. *Science* **302**, 1571-1575 (2003).
60 Wohlwend, D., Strasser, A., Dickmanns, A. & Ficner, R. Structural basis for RanGTP independent entry of spliceosomal U snRNPs into the nucleus. *J Mol Biol* **374**, 1129-1138 (2007). doi.org/10.1016/j.jmb.2007.09.065
61 Mitrousis, G., Olia, A. S., Walker-Kopp, N. & Cingolani, G. Molecular basis for the recognition of snurportin 1 by importin beta. *J Biol Chem* **283**, 7877-7884 (2008). doi.org/10.1074/jbc.M709093200
62 Bhardwaj, A. & Cingolani, G. Conformational selection in the recognition of the snurportin importin beta binding domain by importin beta. *Biochemistry* **49**, 5042-5047 (2010). doi.org/10.1021/bi100292y
63 Choi, S., Yamashita, E., Yasuhara, N., Song, J., Son, S. Y., Won, Y. H., Hong, H. R., Shin, Y. S., Sekimoto, T., Park, I. Y. *et al.* Structural basis for the selective nuclear import of the C2H2 zinc-finger protein Snail by importin beta. *Acta Crystallogr D Biol Crystallogr* **70**, 1050-1060 (2014). doi.org/10.1107/S1399004714000972
64 Scanlon, M. J., Fairlie, D. P., Craik, D. J., Englebretsen, D. R. & West, M. L. NMR solution structure of the RNA-binding peptide from human immunodeficiency virus (type 1) Rev. *Biochemistry* **34**, 8242-8249 (1995). doi.org/10.1021/bi00026a005
65 Daugherty, M. D., Liu, B. & Frankel, A. D. Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. *Nat Struct Mol Biol* **17**, 1337-1342 (2010).

66 Casu, F., Duggan, B. M. & Hennig, M. The arginine-rich RNA-binding motif of HIV-1 Rev is intrinsically disordered and folds upon RRE binding. *Biophys J* **105**, 1004-1017 (2013). doi.org/10.1016/j.bpj.2013.07.022
67 DiMattia, M. A., Watts, N. R., Cheng, N., Huang, R., Heymann, J. B., Grimes, J. M., Wingfield, P. T., Stuart, D. I. & Steven, A. C. The Structure of HIV-1 Rev Filaments Suggests a Bilateral Model for Rev-RRE Assembly. *Structure* **24**, 1068-1080 (2016). doi.org/10.1016/j.str.2016.04.015
68 Watts, N. R., Eren, E., Zhuang, X., Wang, Y. X., Steven, A. C. & Wingfield, P. T. A new HIV-1 Rev structure optimizes interaction with target RNA (RRE) for nuclear export. *J Struct Biol* **203**, 102-108 (2018). doi.org/10.1016/j.jsb.2018.03.011
69 DiMattia, M. A., Watts, N. R., Stahl, S. J., Rader, C., Wingfield, P. T., Stuart, D. I., Steven, A. C. & Grimes, J. M. Implications of the HIV-1 Rev dimer structure at 3.2 A resolution for multimeric binding to the Rev response element. *Proc Natl Acad Sci U S A* **107**, 5810-5814 (2010).
70 Battiste, J. L., Mao, H., Rao, N. S., Tan, R., Muhandiram, D. R., Kay, L. E., Frankel, A. D. & Williamson, J. R. Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science* **273**, 1547-1551 (1996). doi.org/10.1126/science.273.5281.1547
71 Ye, X., Gorin, A., Ellington, A. D. & Patel, D. J. Deep penetration of an alpha-helix into a widened RNA major groove in the HIV-1 rev peptide-RNA aptamer complex. *Nat Struct Biol* **3**, 1026-1033 (1996).
72 Ye, X., Gorin, A., Frederick, R., Hu, W., Majumdar, A., Xu, W., McLendon, G., Ellington, A. & Patel, D. J. RNA architecture dictates the conformations of a bound peptide. *Chem Biol* **6**, 657-669 (1999).
73 Daugherty, M. D., Booth, D. S., Jayaraman, B., Cheng, Y. & Frankel, A. D. HIV Rev response element (RRE) directs assembly of the Rev homooligomer into discrete asymmetric complexes. *Proc Natl Acad Sci U S A* **107**, 12481-12486 (2010). doi.org/10.1073/pnas.1007022107
74 Jayaraman, B., Crosby, D. C., Homer, C., Ribeiro, I., Mavor, D. & Frankel, A. D. RNA-directed remodeling of the HIV-1 protein Rev orchestrates assembly of the Rev-Rev response element complex. *Elife* **3**, e04120 (2014). doi.org/10.7554/eLife.04120
75 Dearborn, A. D., Eren, E., Watts, N. R., Palmer, I. W., Kaufman, J. D., Steven, A. C. & Wingfield, P. T. Structure of an RNA Aptamer that Can Inhibit HIV-1 by Blocking Rev-Cognate RNA (RRE) Binding and Rev-Rev Association. *Structure* **26**, 1187-1195 e1184 (2018). doi.org/10.1016/j.str.2018.06.001
76 Guttler, T., Madl, T., Neumann, P., Deichsel, D., Corsini, L., Monecke, T., Ficner, R., Sattler, M. & Gorlich, D. NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nat Struct Mol Biol* **17**, 1367-1376 (2010).
77 Faust, O., Grunhaus, D., Shimshon, O., Yavin, E. & Friedler, A. Protein Regulation by Intrinsically Disordered Regions: A Role for Subdomains in the IDR of the HIV-1 Rev Protein. *Chembiochem* **19**, 1618-1624 (2018). doi.org/10.1002/cbic.201800192
78 Jayaraman, B., Fernandes, J. D., Yang, S., Smith, C. & Frankel, A. D. Highly Mutable Linker Regions Regulate HIV-1 Rev Function and Stability. *Sci Rep* **9**, 5139 (2019). doi.org/10.1038/s41598-019-41582-7
79 Toth-Petroczy, A., Palmedo, P., Ingraham, J., Hopf, T. A., Berger, B., Sander, C. & Marks, D. S. Structured States of Disordered Proteins from Genomic Sequences. *Cell* **167**, 158-170 e112 (2016). doi.org/10.1016/j.cell.2016.09.010
80 Behrens, R. T., Aligeti, M., Pocock, G. M., Higgins, C. A. & Sherer, N. M. Nuclear Export Signal Masking Regulates HIV-1 Rev Trafficking and Viral RNA Nuclear Export. *J Virol* **91** (2017). doi.org/10.1128/JVI.02107-16
81 Olsen, H. S., Cochrane, A. W., Dillon, P. J., Nalin, C. M. & Rosen, C. A. Interaction of the human immunodeficiency virus type 1 Rev protein with a structured region in env mRNA is dependent on multimer formation mediated through a basic stretch of amino acids. *Genes Dev* **4**, 1357-1364 (1990). doi.org/10.1101/gad.4.8.1357
82 Hope, T. J., McDonald, D., Huang, X. J., Low, J. & Parslow, T. G. Mutational analysis of the human immunodeficiency virus type 1 Rev transactivator: essential residues near the amino terminus. *J Virol* **64**, 5360-5366 (1990).
83 Bohnlein, E., Berger, J. & Hauber, J. Functional mapping of the human immunodeficiency virus type 1 Rev RNA binding domain: new insights into the domain structure of Rev and Rex. *J Virol* **65**, 7051-7055 (1991).
84 Zapp, M. L., Hope, T. J., Parslow, T. G. & Green, M. R. Oligomerization and RNA binding domains of the type 1 human immunodeficiency virus Rev protein: a dual function for an arginine-rich binding motif. *Proc Natl Acad Sci U S A* **88**, 7734-7738 (1991). doi.org/10.1073/pnas.88.17.7734

85 Jain, C. & Belasco, J. G. Structural model for the cooperative assembly of HIV-1 Rev multimers on the RRE as deduced from analysis of assembly-defective mutants. *Mol Cell* **7**, 603-614 (2001).
86 Thomas, S. L., Oft, M., Jaksche, H., Casari, G., Heger, P., Dobrovnik, M., Bevec, D. & Hauber, J. Functional analysis of the human immunodeficiency virus type 1 Rev protein oligomerization interface. *J Virol* **72**, 2935-2944 (1998).
87 Madore, S. J., Tiley, L. S., Malim, M. H. & Cullen, B. R. Sequence requirements for Rev multimerization in vivo. *Virology* **202**, 186-194 (1994). doi.org/10.1006/viro.1994.1334
88 Daugherty, M. D., D'Orso, I. & Frankel, A. D. A solution to limited genomic capacity: using adaptable binding surfaces to assemble the functional HIV Rev oligomer on RNA. *Mol Cell* **31**, 824-834 (2008). doi.org/10.1016/j.molcel.2008.07.016
89 Edgcomb, S. P., Aschrafi, A., Kompfner, E., Williamson, J. R., Gerace, L. & Hennig, M. Protein structure and oligomerization are important for the formation of export-competent HIV-1 Rev-RRE complexes. *Protein Sci* **17**, 420-430 (2008). doi.org/10.1110/ps.073246608
90 Pond, S. J., Ridgeway, W. K., Robertson, R., Wang, J. & Millar, D. P. HIV-1 Rev protein assembles on viral RNA one molecule at a time. *Proc Natl Acad Sci U S A* **106**, 1404-1408 (2009). doi.org/10.1073/pnas.0807388106
91 Boeri Erba, E. & Petosa, C. The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes. *Protein Sci* **24**, 1176-1192 (2015). doi.org/10.1002/pro.2661
92 Halder, K., Dolker, N., Van, Q., Gregor, I., Dickmanns, A., Baade, I., Kehlenbach, R. H., Ficner, R., Enderlein, J., Grubmuller, H. *et al.* MD simulations and FRET reveal an environment-sensitive conformational plasticity of importin-beta. *Biophys J* **109**, 277-286 (2015). doi.org/10.1016/j.bpj.2015.06.014
93 Kragelj, J., Ozenne, V., Blackledge, M. & Jensen, M. R. Conformational propensities of intrinsically disordered proteins from NMR chemical shifts. *Chemphyschem* **14**, 3034-3045 (2013). doi.org/10.1002/cphc.201300387
94 Fukuhara, N., Fernandez, E., Ebert, J., Conti, E. & Svergun, D. Conformational variability of nucleo-cytoplasmic transport factors. *J Biol Chem* **279**, 2176-2181 (2004).
95 Fischer, L., Chen, Z. A. & Rappsilber, J. Quantitative cross-linking/mass spectrometry using isotope-labelled cross-linkers. *J Proteomics* **88**, 120-128 (2013). doi.org/10.1016/j.jprot.2013.03.005
96 Merkley, E. D., Rysavy, S., Kahraman, A., Hafen, R. P., Daggett, V. & Adkins, J. N. Distance restraints from crosslinking mass spectrometry: mining a molecular dynamics simulation database to evaluate lysine-lysine distances. *Protein Sci* **23**, 747-759 (2014). doi.org/10.1002/pro.2458
97 Orban-Nemeth, Z., Beveridge, R., Hollenstein, D. M., Rampler, E., Stranzl, T., Hudecz, O., Doblmann, J., Schlogelhofer, P. & Mechtler, K. Structural prediction of protein models using distance restraints derived from cross-linking mass spectrometry data. *Nat Protoc* **13**, 478-494 (2018). doi.org/10.1038/nprot.2017.146
98 Bullock, J., Schwab, J., Thalassinos, K. & Topf, M. The Importance of Non-accessible Crosslinks and Solvent Accessible Surface Distance in Modeling Proteins with Restraints From Crosslinking Mass Spectrometry. *Mol Cell Proteomics* **15**, 2491-2500 (2016). doi.org/10.1074/mcp.M116.058560
99 van Zundert, G. C. P., Rodrigues, J., Trellet, M., Schmitz, C., Kastritis, P. L., Karaca, E., Melquiond, A. S. J., van Dijk, M., de Vries, S. J. & Bonvin, A. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol* **428**, 720-725 (2016). doi.org/10.1016/j.jmb.2015.09.014
100 Moore, J. D., Yang, J., Truant, R. & Kornbluth, S. Nuclear import of Cdk/cyclin complexes: identification of distinct mechanisms for import of Cdk2/cyclin E and Cdc2/cyclin B1. *J Cell Biol* **144**, 213-224 (1999). doi.org/10.1083/jcb.144.2.213
101 Jakel, S. & Gorlich, D. Importin beta, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. *Embo J* **17**, 4491-4502 (1998). doi.org/10.1093/emboj/17.15.4491
102 Kubota, S., Duan, L., Furuta, R. A., Hatanaka, M. & Pomerantz, R. J. Nuclear preservation and cytoplasmic degradation of human immunodeficiency virus type 1 Rev protein. *J Virol* **70**, 1282-1287 (1996). doi.org/10.1128/JVI.70.2.1282-1287.1996
103 Fineberg, K., Fineberg, T., Graessmann, A., Luedtke, N. W., Tor, Y., Lixin, R., Jans, D. A. & Loyter, A. Inhibition of nuclear import mediated by the Rev-arginine rich motif by RNA molecules. *Biochemistry* **42**, 2625-2633 (2003). doi.org/10.1021/bi0206199

104 Gorlich, D., Henklein, P., Laskey, R. A. & Hartmann, E. A 41 amino acid motif in importin-alpha confers binding to importin-beta and hence transit into the nucleus. *Embo J* **15**, 1810-1817 (1996).
105 Weis, K., Ryder, U. & Lamond, A. I. The conserved amino-terminal domain of hSRP1 alpha is essential for nuclear protein import. *Embo J* **15**, 1818-1825 (1996).
106 Vercruysse, T. & Daelemans, D. HIV-1 Rev multimerization: mechanism and insights. *Curr HIV Res* **11**, 623-634 (2013). doi.org/10.2174/1570162x12666140307094603
107 Heaphy, S., Finch, J. T., Gait, M. J., Karn, J. & Singh, M. Human immunodeficiency virus type 1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich "bubble" located within the rev-responsive region of viral mRNAs. *Proc Natl Acad Sci U S A* **88**, 7366-7370 (1991). doi.org/10.1073/pnas.88.16.7366
108 Watts, N. R., Misra, M., Wingfield, P. T., Stahl, S. J., Cheng, N., Trus, B. L., Steven, A. C. & Williams, R. W. Three-dimensional structure of HIV-1 Rev protein filaments. *J Struct Biol* **121**, 41-52 (1998). doi.org/10.1006/jsbi.1998.3964
109 Wingfield, P. T., Stahl, S. J., Payton, M. A., Venkatesan, S., Misra, M. & Steven, A. C. HIV-1 Rev expressed in recombinant Escherichia coli: purification, polymerization, and conformational properties. *Biochemistry* **30**, 7527-7534 (1991). doi.org/10.1021/bi00244a023
110 Bai, Y., Tambe, A., Zhou, K. & Doudna, J. A. RNA-guided assembly of Rev-RRE nuclear export complexes. *Elife* **3**, e03656 (2014). doi.org/10.7554/eLife.03656
111 Ben Fadhel, N., Signor, L., Petosa, C. & Noirclerc-Savoye, M. Phosphomimetic mutations modulate the ability of HIV-1 Rev to bind human Importin β in vitro. *Matters* (2019). doi.org/10.19185/matters.201903000032
112 Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-234 (2005).
113 Brautigam, C. A., Zhao, H., Vargas, C., Keller, S. & Schuck, P. Integration and global analysis of isothermal titration calorimetry data for studying macromolecular interactions. *Nat Protoc* **11**, 882-894 (2016). doi.org/10.1038/nprot.2016.044
114 Sobott, F., Hernandez, H., McCammon, M. G., Tito, M. A. & Robinson, C. V. A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal Chem* **74**, 1402-1407 (2002). doi.org/10.1021/ac0110552
115 van den Heuvel, R. H., van Duijn, E., Mazon, H., Synowsky, S. A., Lorenzen, K., Versluis, C., Brouns, S. J., Langridge, D., van der Oost, J., Hoyes, J. *et al.* Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Anal Chem* **78**, 7473-7483 (2006). doi.org/10.1021/ac061039a
116 Boeri Erba, E., Signor, L., Oliva, M. F., Hans, F. & Petosa, C. Characterizing Intact Macromolecular Complexes Using Native Mass Spectrometry. *Methods Mol Biol* **1764**, 133-151 (2018). doi.org/10.1007/978-1-4939-7759-8_9
117 Morgner, N. & Robinson, C. V. Massign: an assignment strategy for maximizing information from the mass spectra of heterogeneous protein assemblies. *Anal Chem* **84**, 2939-2948 (2012). doi.org/10.1021/ac300056a
118 Chen, Z. L., Meng, J. M., Cao, Y., Yin, J. L., Fang, R. Q., Fan, S. B., Liu, C., Zeng, W. F., Ding, Y. H., Tan, D. *et al.* A high-speed search engine pLink 2 with systematic evaluation for proteome-scale identification of cross-linked peptides. *Nat Commun* **10**, 3404 (2019). doi.org/10.1038/s41467-019-11337-z
119 Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301-2319 (2016). doi.org/10.1038/nprot.2016.136
120 Solyom, Z., Schwarten, M., Geist, L., Konrat, R., Willbold, D. & Brutscher, B. BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *J Biomol NMR* **55**, 311-321 (2013). doi.org/10.1007/s10858-013-9715-0
121 Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995). doi.org/10.1007/BF00197809
122 Jung, Y. S. & Zweckstetter, M. Mars -- robust automatic backbone assignment of proteins. *J Biomol NMR* **30**, 11-23 (2004). doi.org/10.1023/B:JNMR.0000042954.99056.ad
123 Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS - a Windows-PC based system for small-angle scattering data analysis. *J Appl Crystallogr* **36**, 1277-1282 (2003).

124 Svergun, D., Barberato, C. & Koch, M. H. CRYSOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* **28**, 768-773 (1995).

125 Suhre, K. & Sanejouand, Y. H. ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Res* **32**, W610-614 (2004). doi.org/10.1093/nar/gkh368

126 Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242 (2011). doi.org/10.1107/S0907444910045749

127 Marsh, J. A., Singh, V. K., Jia, Z. & Forman-Kay, J. D. Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci* **15**, 2795-2804 (2006). doi.org/10.1110/ps.062465306

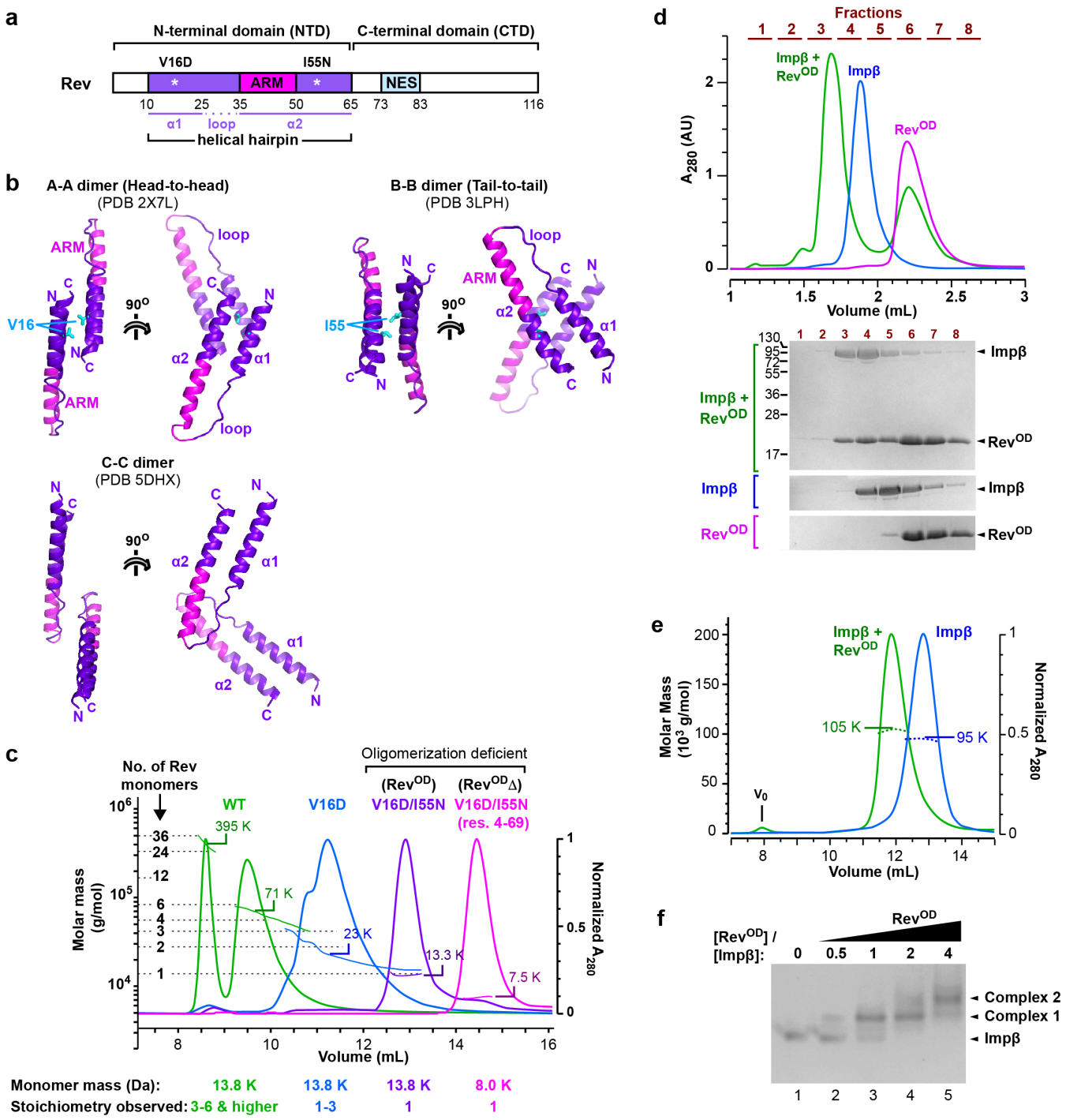
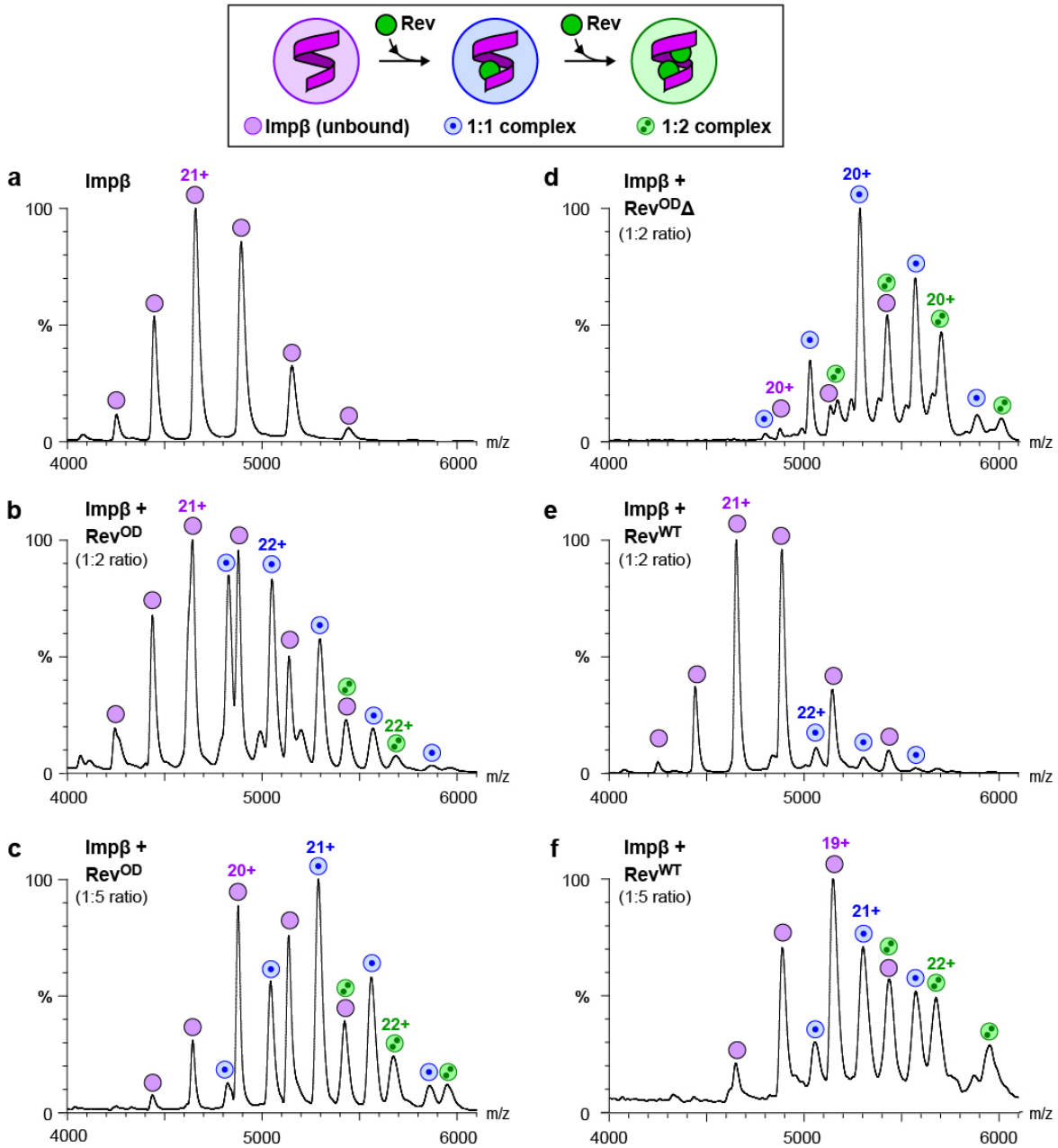


Figure 1. HIV-1 Rev forms a stable complex with Impβ. a. Domain organization of Rev. **b.** Oligomerization interfaces reported for Rev. Residues mutated to destabilize the A-A and B-B interfaces are indicated in cyan. **c.** Analysis of different Rev constructs by SEC/MALLS using a Superdex 75 10/300 GL column. Elution curves recorded at 280 nm and molar mass distributions determined by MALLS are shown for the following Rev constructs: WT (green), V16D (blue), V16D/I55N (purple) and Rev^{4-69}(V16D/I55N)(pink). The molar masses detected at the elution peaks are indicated. **d.** SEC analysis of an Impβ/Rev^{OD} complex. *Top*, elution profiles of samples containing Impβ (blue), Rev^{OD} (magenta) or a mixture of Impβ and Rev^{OD} (green). Fractions collected are indicated in brown. Chromatography was performed using a Superdex 200 5/150 column. *Bottom*, SDS-PAGE analysis of the indicated fractions. **e.** SEC/MALLS analysis of Impβ in the presence and absence of Rev. Elution curves recorded at 280 nm and molar mass distributions determined by MALLS are shown for Impβ in the absence (blue) and presence (green) of Rev^{OD}. Chromatography was performed using a Superdex 200 10/30 GL column. Injected concentrations were 40 μM Impβ and 80 μM Rev^{OD}. **f.** Native gel analysis showing the association of Impβ with one or more monomers of Rev^{OD}.



g

Stoichiometry		Experimental mass (Da)	Theoretical mass (Da)	Δ mass (Da)
Impβ	Rev^{OD}			
1	-	97,296 ± 3	97,299	-3
-	1	13,263 ± 1	13,264	-1
1	1	110,558 ± 8	110,563	-5
1	2	123,829 ± 6	123,827	+2
Impβ	$Rev^{OD}\Delta$			
-	1	8,031 ± 1	8,030	+1
1	1	105,330 ± 3	105,330	0
1	2	113,366 ± 6	113,361	+5
Impβ	Rev^{WT}			
-	1	13,173 ± 2	13,174	-1
1	1	110,476 ± 4	110,473	+3
1	2	123,649 ± 5	123,648	+1

Figure 2. Native MS reveals that Impβ binds up to two Rev monomers. Peaks labelled by magenta circles correspond to unbound Impβ. Peaks labelled by blue circles with a single dot or by green circles with two dots correspond to Impβ/Rev complexes with 1:1 or 1:2 stoichiometry, respectively. **a.** Spectrum of Impβ in the absence of Rev. **b,c.** Spectra of Impβ incubated with a b) 2-fold or c) 5-fold molar equivalent of Rev^{OD}. **d.** Spectra of Impβ in the presence of the truncated construct $Rev^{OD}\Delta$. **e,f.** Spectra of Impβ incubated with a e) 2-fold or f) 5-fold molar equivalent of Rev^{WT}. **g.** Summary of masses observed by native MS.

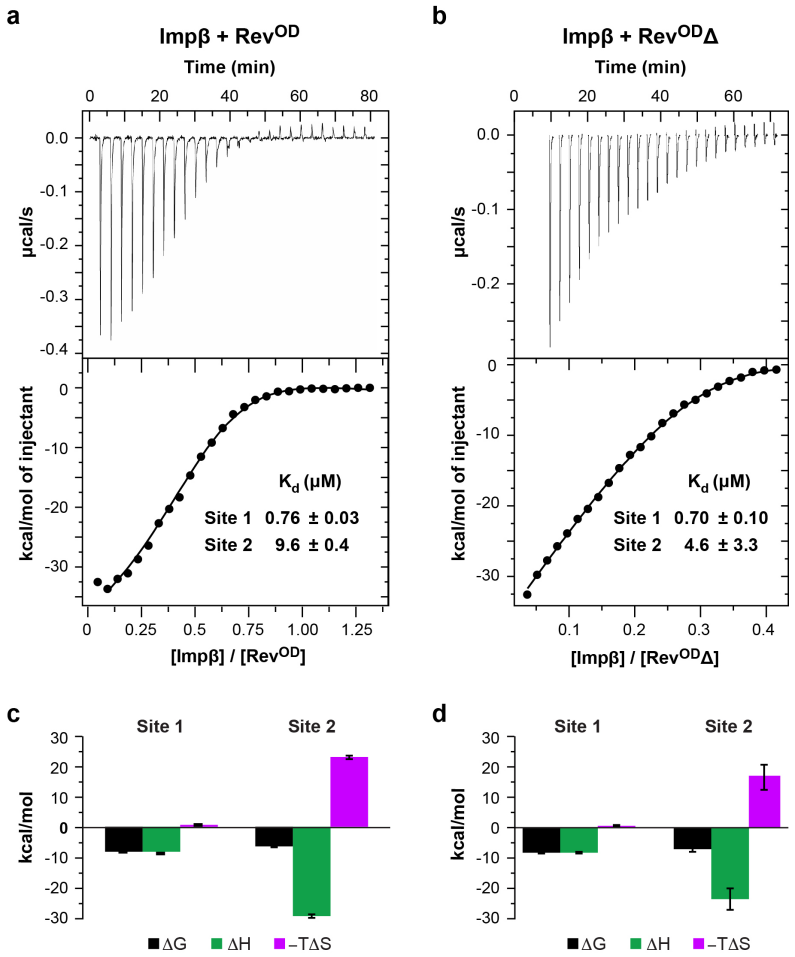


Figure 3. Rev binds Impβ at two sites with sub- and low micromolar affinity. a,b. ITC profiles of the binding of Impβ to a) Rev^{OD} and b) $Rev^{OD}\Delta$. *Top*, Differential power time course of raw injection heats for a titration of Impβ into the Rev protein solutions. *Bottom*, Normalized binding isotherms corrected for the heat of dilution of Impβ into buffer. The solid line represents a nonlinear least squares fit using a model consisting of two non-symmetric classes of binding sites. **c,d.** Thermodynamic values obtained from ITC data for the binding of Impβ to c) Rev^{OD} and d) $Rev^{OD}\Delta$. Binding to site 1 is characterized by a favourable enthalpy ($\Delta H \approx$ -8 kcal/mol) and a negligible entropy change for both Rev^{OD} and $Rev^{OD}\Delta$, whereas binding to site 2 is associated with an unfavourable entropy change that is offset by a large negative enthalpy change. Thermodynamic parameters (in kcal/mol) are as follows: Rev^{OD}: ΔG = -8.3 ± 0.1 and -6.8 ± 0.1, ΔH = -8.8 ± 0.2 and -29.8 ± 0.9, $-T\Delta S$ = 0.6 ± 0.3 and 23.1 ±0.8 for sites 1 and 2, respectively; $Rev^{OD}\Delta$: ΔG= 8.4 ± 0.1 and -7.5 ± 0.7, ΔH = -8.6 ± 0.5 and -24.1 ± 4.9, $-T\Delta S$= 0.2 ±0.6 and 16.6 ± 5.7 for sites 1 and 2, respectively.

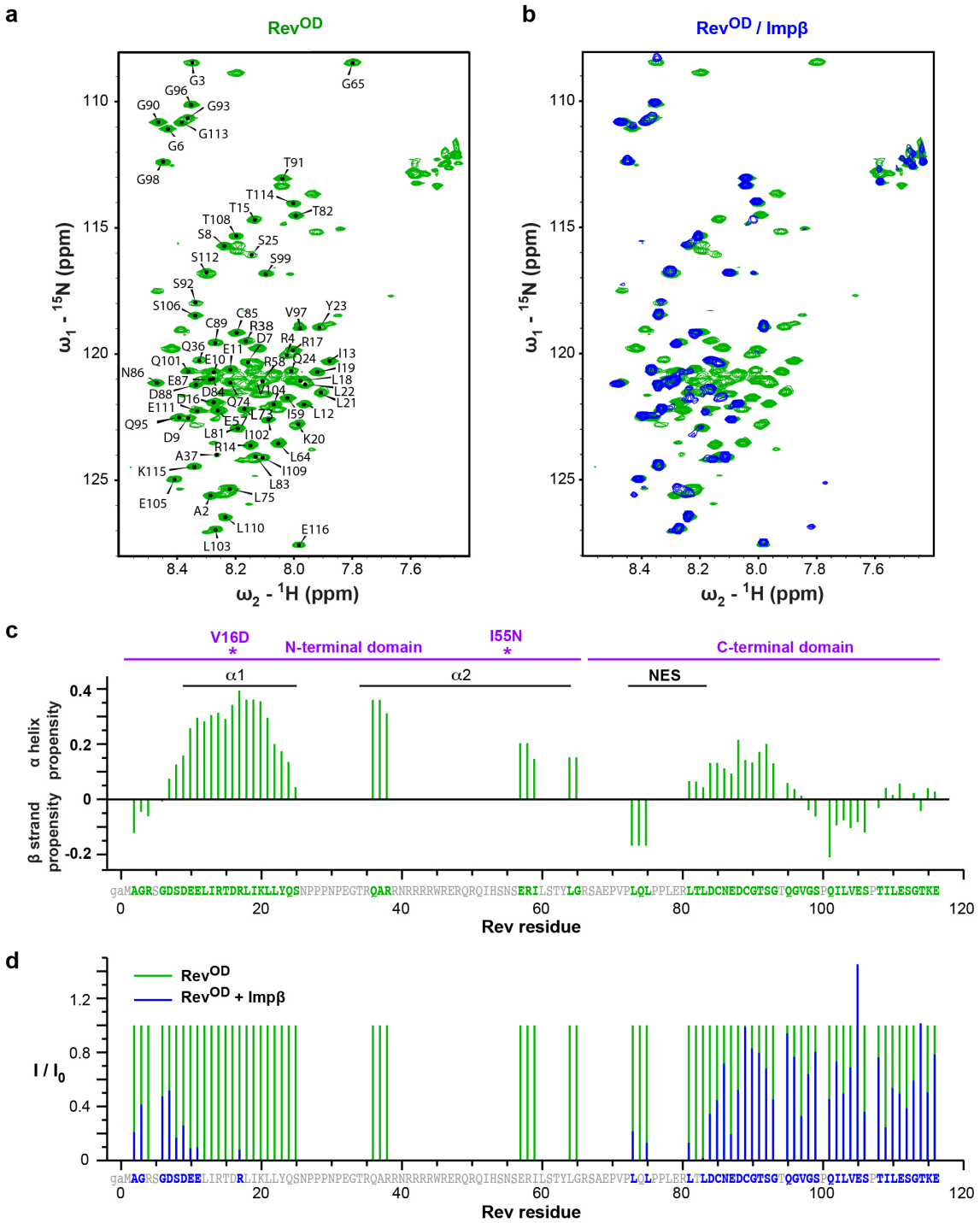
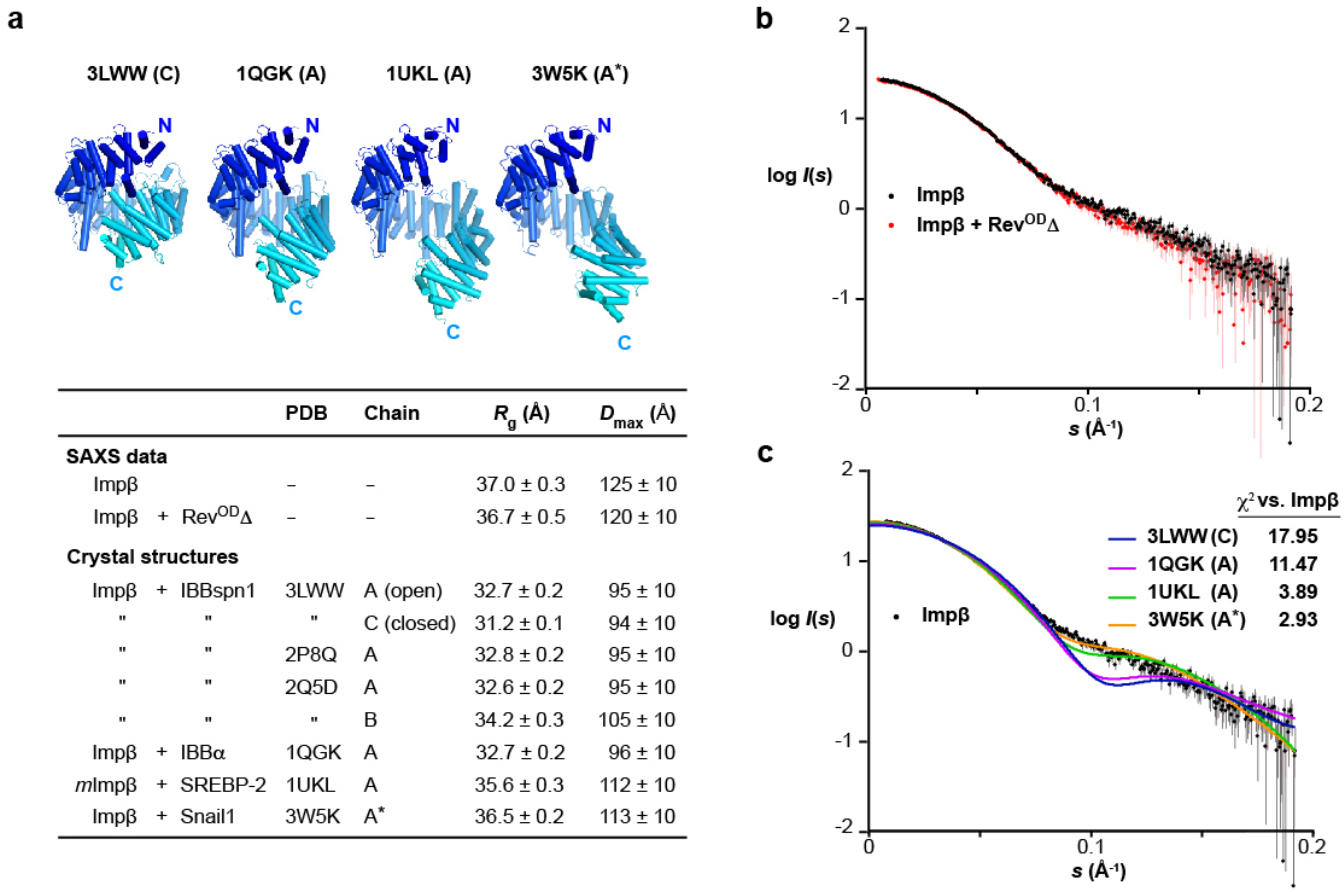


Figure 4. NMR analysis of Rev binding by Impβ. a. ^{1}H,^{15}N HSQC spectrum of free Rev^{OD} (600MHz, 283K). **b.** ^{1}H,^{15}N-HSQC spectrum Impβ-bound Rev^{OD} (blue) superimposed on that of unbound Rev^{OD} (green). **c.** Secondary chemical shifts[127] of unbound Rev^{OD}. The two point mutations are indicated by asterisks. **d.** Ratio of peak intensities for Rev residues in the presence and absence of Impβ (600MHz, 283K).



	PDB	Chain	R_g (Å)	D_{max} (Å)
SAXS data				
Impβ	–	–	37.0 ± 0.3	125 ± 10
Impβ + Rev^{OD}Δ	–	–	36.7 ± 0.5	120 ± 10
Crystal structures				
Impβ + IBBspn1	3LWW	A (open)	32.7 ± 0.2	95 ± 10
" "	"	C (closed)	31.2 ± 0.1	94 ± 10
" "	2P8Q	A	32.8 ± 0.2	95 ± 10
" "	2Q5D	A	32.6 ± 0.2	95 ± 10
" "	"	B	34.2 ± 0.3	105 ± 10
Impβ + IBBα	1QGK	A	32.7 ± 0.2	96 ± 10
*m*Impβ + SREBP-2	1UKL	A	35.6 ± 0.3	112 ± 10
Impβ + Snail1	3W5K	A*	36.5 ± 0.2	113 ± 10

Figure 5. Impβ retains an extended conformation upon binding Rev. a. *Top.* Representative crystal structures of Impβ illustrating different degrees of compaction of the HEAT-repeat array. *Bottom*. Model-independent parameters obtained from SAXS data compared to values calculated for available Impβ crystal structures using the programs CRYSOL[124] and PRIMUS[123]. All structures are of human Impβ except for 1UKL which is of murine Impβ. Chain A in 3W5K lacks coordinates for Impβ residues 1-15. For the calculation of R_g, D_{max} and χ^2 values this chain was extended to include these residues (denoted A*) by replacing HEAT repeat 1 (residues 1-31) by the corresponding residues from 1UKL following local alignment of the two structures. **b.** Scattering data from unbound Impβ (black) and the Impβ/Rev^{OD}Δ complex (red). **c.** Scattering data from unbound Impβ (black) compared to profiles calculated from representative cargo-bound conformations exhibiting different degrees of elongation (colored lines).

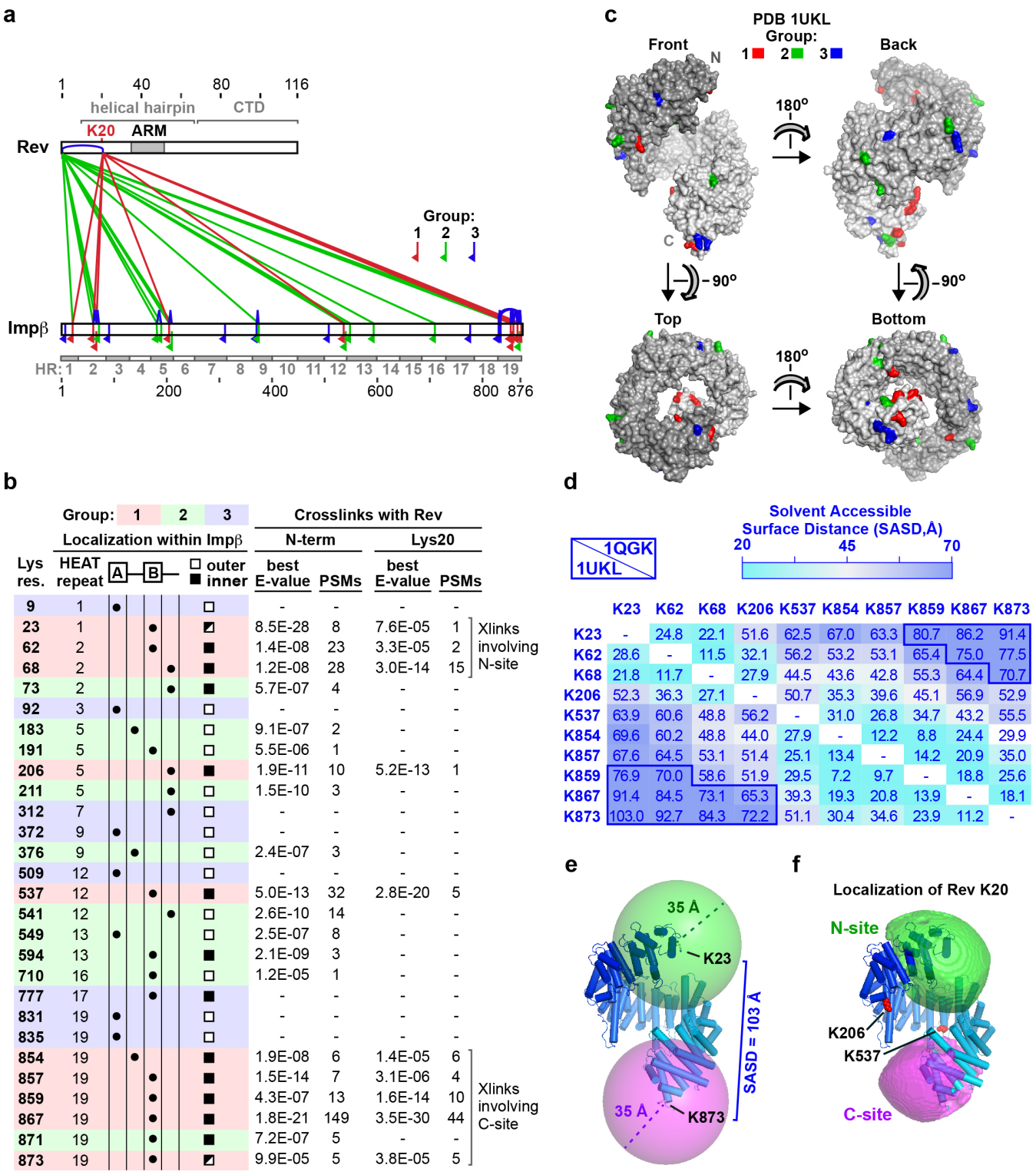


Figure 6. Crosslinking-MS localizes two Rev-binding regions on Impβ. a. Graphical summary of crosslinks. Impβ-Rev crosslinks involving the N-terminus or Lys20 residue of Rev are shown in green and red, respectively. Rev-Rev and Impβ-Impβ crosslinks and monolinks are shown in blue. Red, green and blue inverted flags indicate Impβ residues that form crosslinks with Rev Lys20 (Group 1), with both Rev Lys20 and the Rev N-terminus (Group 2), or for which no crosslinks with Rev were detected (Group 3), respectively. The 19 HEAT repeats (HR) of Impβ are indicated. **b.** Impβ Lys residues modified by BS3 and detected in crosslinks or monolinks. Group 1, 2 and 3 lysines are highlighted in red, green and blue, respectively. For Group-1 and -2 lysines the number of peptide spectrum matches (PSMs) and the best pLink E-value score is indicated. **c**. Surface representation of Impβ showing the location of crosslinked Group-1, -2 and -3 Lys residues, colored red, green and blue, respectively. Shading is from dark to light gray from N- to C-terminus. **d.** Solvent-accessible surface distances (SASDs) between pairs of Impβ

Group-1 Lys residues. The upper and lower triangles show distances for the conformations of Impβ bound to the Importin α IBB domain (PDB 1QGK) and to SREBP-2 (PDB 1UKL), respectively. SASDs over 70 Å are outlined in dark blue. Distances were calculated using the Jwalk webserver[98]. The SAS distances shown here and in **Figure S3b** suggest that Lys23, Lys62 and Lys68 crosslink with Rev at the N-site, and that Lys854, Lys857, Lys859, Lys867 and Lys873 crosslink with Rev at the C-site. **e.** Spheres of radius 35 Å centered on the Cα atoms of Group-1 residues Lys23 and Lys873 show that BS3 molecules bound to these two lysines cannot crosslink to the same Rev Lys20 position. **f.** Localization of the Cα atom of Rev Lys20. Green and magenta volumes show the N- and C-terminal regions of space within crosslinking distance of Group-1 residues in either HEAT repeats 1 and 2 (K23, K62, K68) or repeat 19 (K854, K857, K859, K867 and K873), respectively, defined by the intersection of 35 Å spheres centered on the Cα atoms of these residues. If the centroid of each envelope is used to estimate the position of the Rev K20 Cα atom, then the positional uncertainty, calculated as the rmsd of each grid point within the envelope (sampled on a 2 Å grid) relative to the centroid, is 22.8 Å for the N-site and 21.6 Å for the C-site for the 1UKL conformation. The Impβ conformation shown in panels c, e and f is that of SREBP-2-bound Impβ (PDB 1UKL).

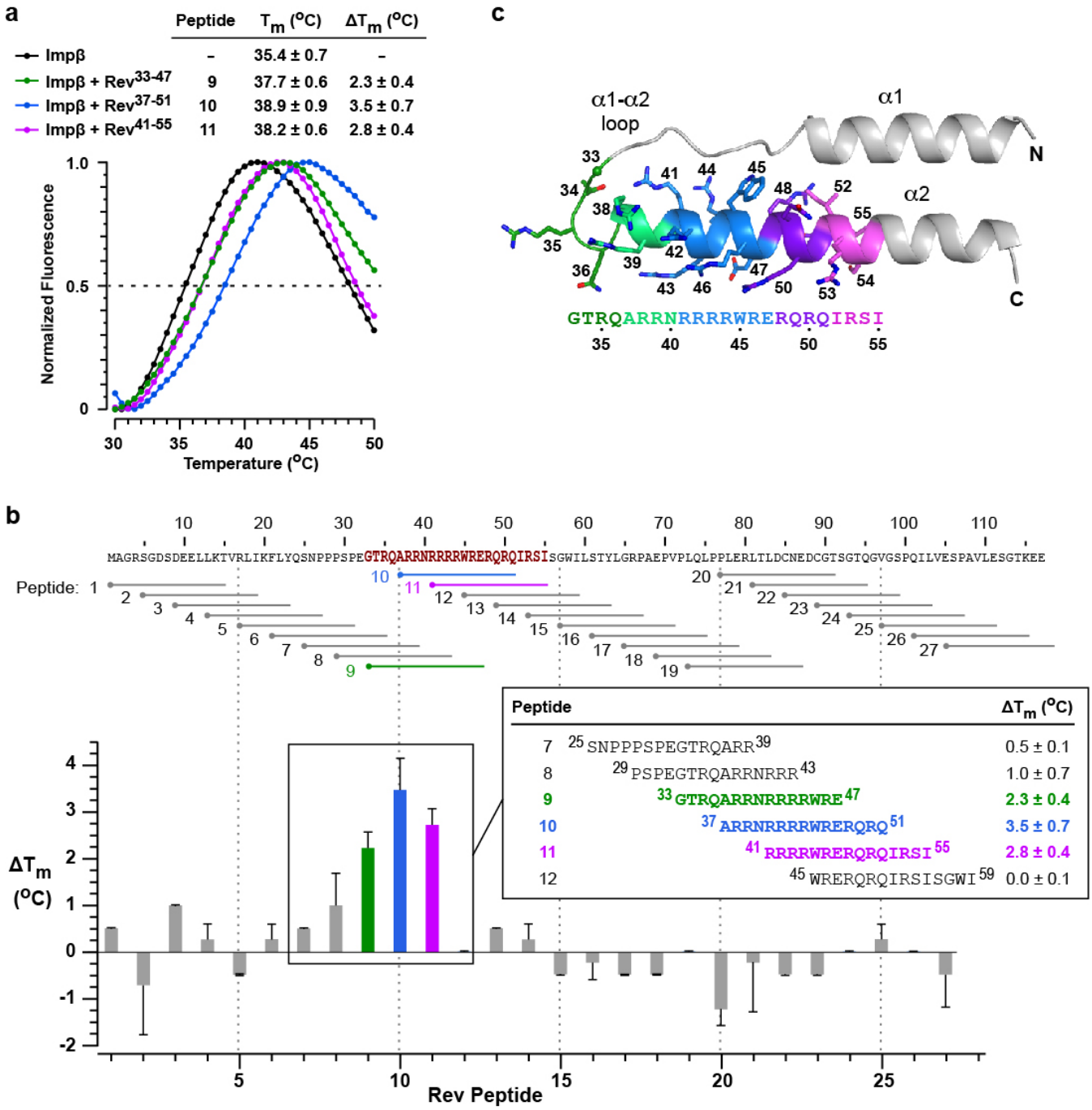
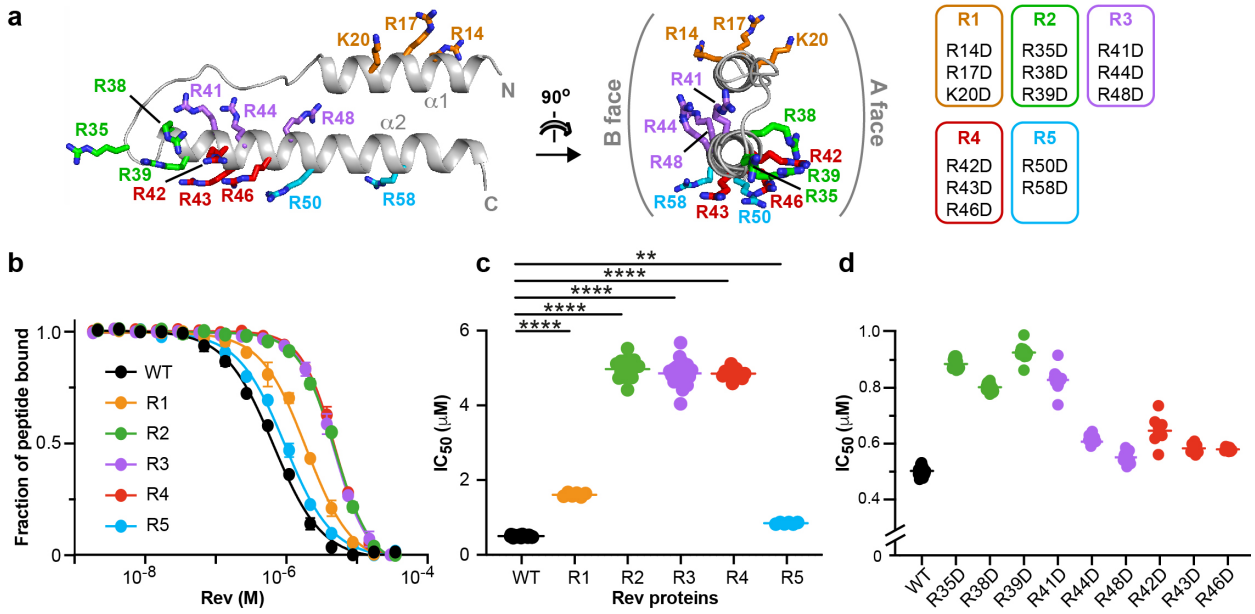


Figure 7. Impβ recognizes Rev peptides derived from helix α2 and the α1-α2 loop. a. Examples of thermal denaturation curves measured by differential scanning fluorimetry (DSF) of Impβ in the presence and absence of Rev peptides. Data are shown for Rev peptides 9-11. The melting temperature (T_m) and difference in T_m compared to unbound Impβ (ΔT_m) are listed as mean values ± SD from three independent experiments. **b.** Summary of ΔT_m values determined by DSF analysis of Impβ for all 27 peptides spanning the Rev sequence. Details are shown for peptides 7-12. **c**. Structure of the Rev N-terminal domain highlighting the residues in helix α2 and the α1-α2 loop spanned by Rev peptides 9-11. Atomic coordinates are taken from PDB 2X7L[69].



e

Rev	IC_{50} (μM)	*N*	pIC_{50}	ΔpIC_{50}
WT	0.502 ± 0.022	34	6.30	—
R1	1.609 ± 0.052	6	5.79	-0.51
R2	4.971 ± 0.268	16	5.30	-1.00
R35D	0.883 ± 0.028	10	6.05	-0.25
R38D	0.802 ± 0.023	10	6.10	-0.20
R39D	0.928 ± 0.048	10	6.03	-0.27
R3	4.860 ± 0.325	22	5.31	-0.99
R41D	0.828 ± 0.017	15	6.08	-0.22
R44D	0.607 ± 0.021	14	6.22	-0.08
R48D	0.551 ± 0.030	15	6.26	-0.04
R4	4.852 ± 0.134	13	5.31	-0.99
R42D	0.646 ± 0.079	8	6.19	-0.11
R43D	0.583 ± 0.027	8	6.23	-0.06
R46D	0.579 ± 0.011	8	6.24	-0.06
R5	0.846 ± 0.023	6	6.07	-0.23

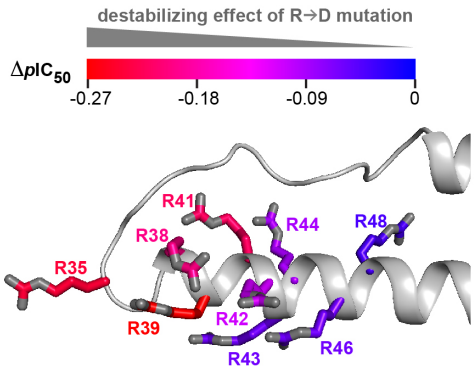


Figure 8. Charge-reversal mutations identify an Impβ-binding epitope on Rev. a. Rev double and triple R/K→D substitution mutations. **b.** Competitive fluorescence polarization (FP) inhibition assays showing the ability of WT and mutant forms of Rev to displace a fluorescently labelled Rev-NLS peptide from Impβ. A representative experiment is shown for each protein. Data shown are mean and s.d. values from two technical replicates. **c.** Plot of IC_{50} values derived from FP inhibition assays for WT Rev and double and triple mutants. ****$P \leq 0.001$; ** $P \leq 0.01$. *P*-values were determined using an ordinary ANOVA test. **d.** Plot of IC_{50} values for WT Rev and single R→D point mutants. **e.** Summary of IC_{50} (mean ± s.d.) and corresponding pIC_{50} values. *N* represents the number of biological replicates. ΔpIC_{50} values are reported relative to WT Rev. **f.** View of the Rev helical hairpin showing the Arg residues selected for single point mutations. Carbon atoms are colored from blue to red in order of increasing ability of the R→D substitution to compromise Impβ recognition, as measured by ΔpIC_{50} values.

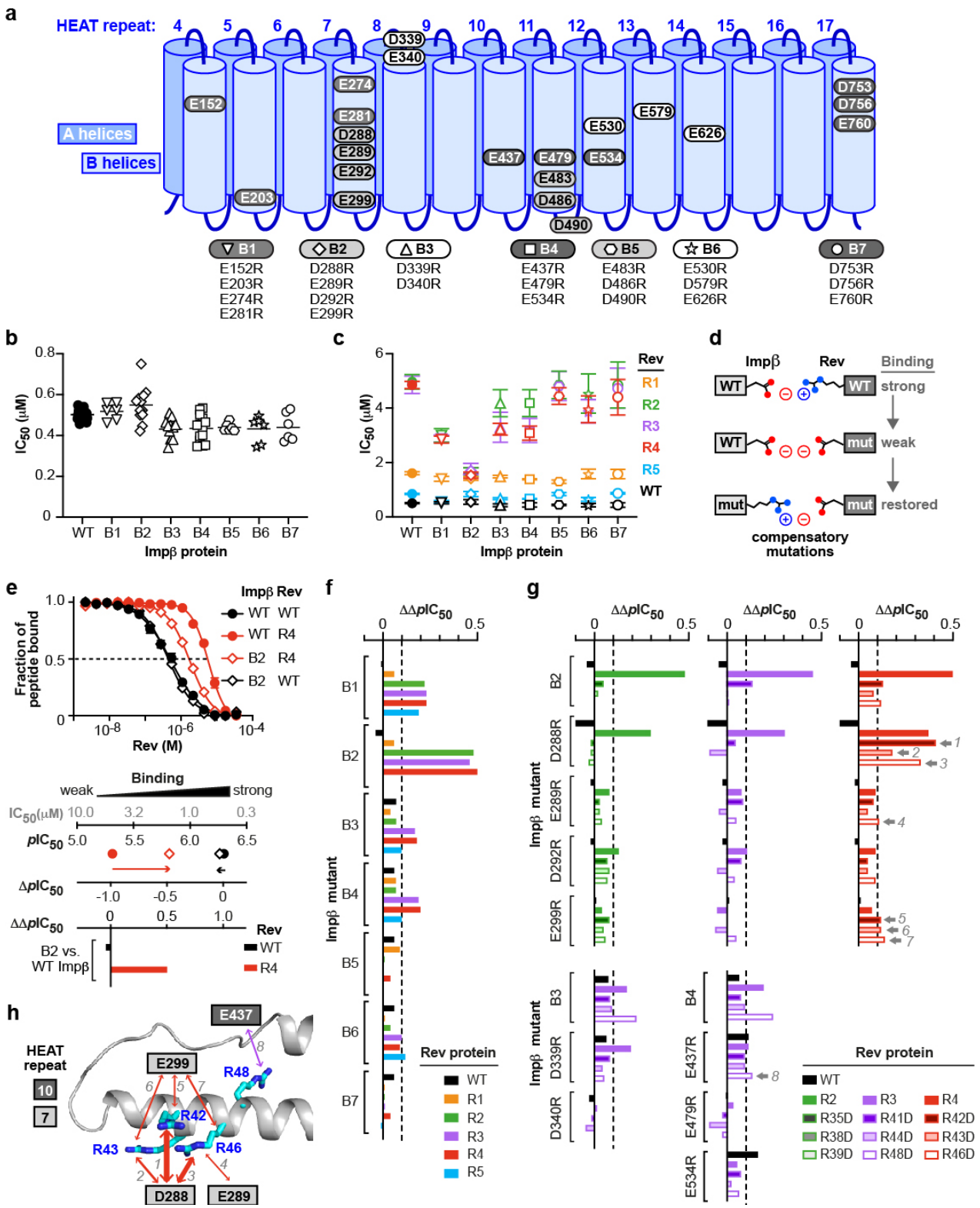


Figure 9. Compensatory effects between charge-reversal mutants of Impβ and Rev. a. Multiple point mutants of Impβ in which 2-4 Asp or Glu residues are replaced by Arg residues. **b.** IC_{50} values obtained from FP inhibition assays measuring the ability of WT Rev to displace a Rev-NLS peptide from WT or mutant forms of Impβ. Results from individual assays are shown together with the mean. **c.** The same assay performed with WT and charge-reversal mutant forms of Rev. Data for Rev mutants are colored as in **Figure 8**. Data shown are mean and s.d. values from 4-14 biological replicates. (See also **Table S1**). **d.** Compensatory effects between charge-reversal mutants. An interaction between an acidic residue on Impβ and an Arg residue on Rev is disrupted by a charge reversal (R→D) mutation of Rev. The interaction is restored by a charge reversal (D/E→R) mutation on Impβ. **e.** Derivation of $\Delta\Delta pIC_{50}$ values. *Top.* FP inhibition assays performed with WT Rev (black curves) or the R4 mutant

(red curves) together with WT Impβ (circles) or the B2 mutant (diamonds). *Bottom.* The ability of WT or mutant Rev to displace the Rev-NLS peptide from WT or mutant Impβ is plotted as pIC_{50} values and as the shift (ΔpIC_{50}) relative to the value observed when WT Rev is assayed with WT Impβ. The $\Delta\Delta pIC_{50}$ value represents the shift in ΔpIC_{50} observed when the assay is performed with an Impβ mutant instead of WT Impβ. A positive value of $\Delta\Delta pIC_{50}$ indicates that the tested Rev mutant more potently displaces the Rev-NLS peptide from the indicated Impβ mutant than from WT Impβ. **f**,**g.** Summary of $\Delta\Delta pIC_{50}$ values for the indicated combinations of Impβ and Rev proteins. The dotted line at $\Delta\Delta pIC_{50} = 0.1$ (corresponding to a 21% decrease in IC_{50} between the mutant and WT Impβ) indicates the threshold used to identify compensatory mutations. Compensatory effects satisfying this criterion involving single point mutants of Impβ and Rev are marked by a grey arrow and numbered 1 to 8. For all eight of these mutant combinations, the mean shift in IC_{50} value observed with the Impβ single point mutant relative to WT Impβ was statistically significant according to a Dunnett's multiple comparisons test (*p* values for mutant combinations 1-8 were <0.0001, 0.0002, 0.0001, 0.0022, 0.0368, 0.0084, <0.0001 and 0.0127, respectively). **h**. Summary of compensatory interactions involving single Impβ and Rev residues. The numbered arrows correspond to the mutant combinations numbered 1-8 in **g**. The thickness of arrows is proportional to the $\Delta\Delta pIC_{50}$ values of the interacting mutants.

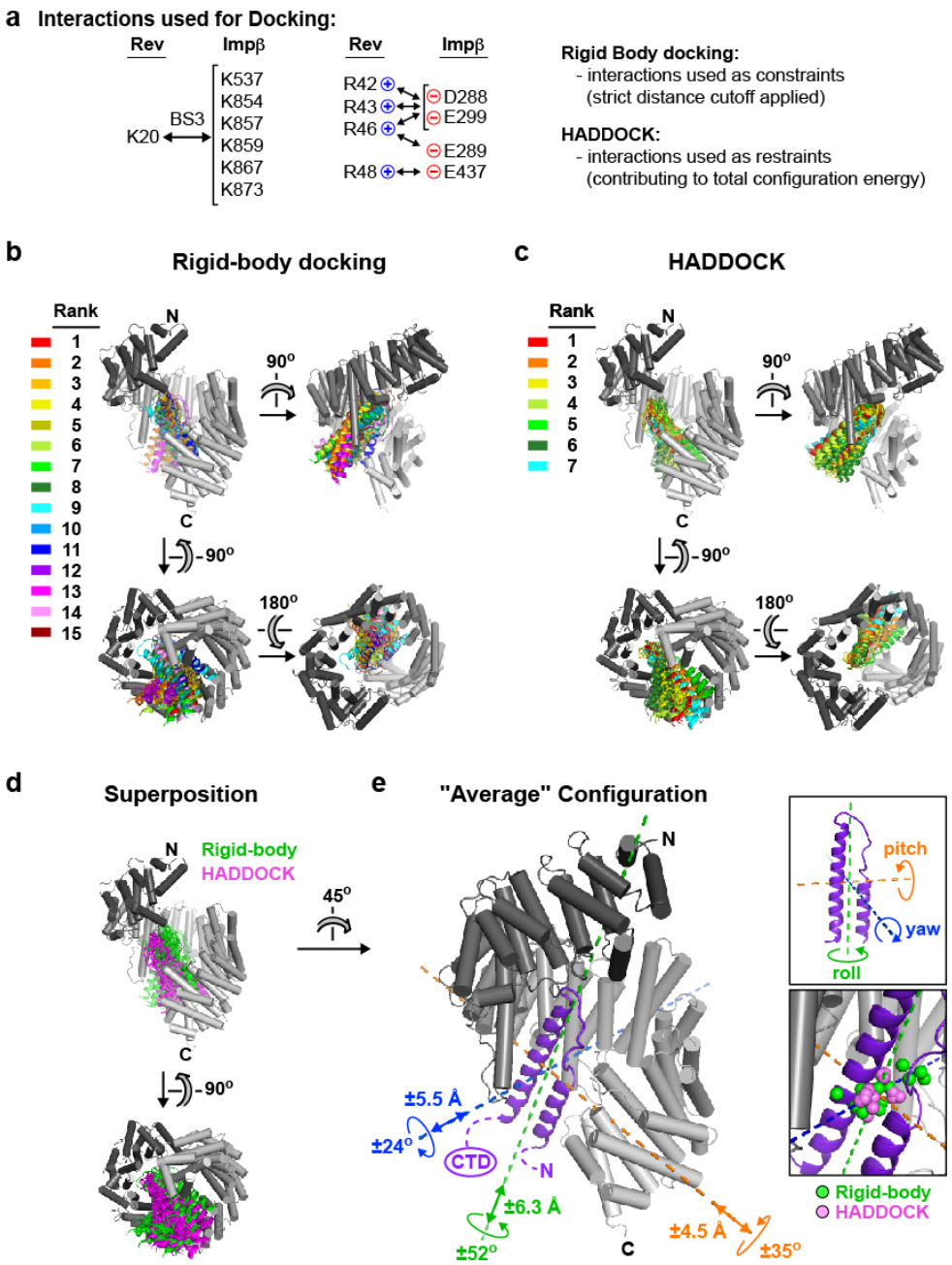


Figure 10. Structural model of Rev bound to Impβ at the C-site. a. BS3 crosslinks and electrostatic interactions used as distance constraints for rigid body docking and as interaction restraints in program HADDOCK. **b.** Configurations of the Impβ/Rev complex obtained by rigid body docking. The mean variation in Rev orientation and centroid position are 33° and <5 Å, respectively, relative to the top-ranked configuration (see also **Table S3**). **c.** Configurations of the Impβ/Rev complex obtained using HADDOCK. For clarity, only 28 (the four lowest-energy structures in each of the 7 clusters) of the 194 clustered solutions are shown. Solutions are colored according to the rank of the corresponding cluster. The mean variation in Rev orientation and centroid position are 34° and 4 Å, respectively (see also **Table S4**). **d.** Superposition showing agreement between docking solutions obtained by rigid body sampling and by HADDOCK. **e.** "Average" configuration of the Impβ/Rev complex that minimizes the deviation from models obtained by rigid-body and HADDOCK docking experiments. Compared to the average configuration, the Rev orientation in the individual rigid-body and HADDOCK configurations shows the greatest variation (by a maximum of 52°) in the angle about the long axis of Rev (roll angle) and smaller variations (up to 25-35°) about the two orthogonal angles (yaw and pitch). The approximate location of the Rev CTD is indicated. *Upper inset.* Principal axes and corresponding rotational angles of Rev. *Lower inset.* Centroids of Rev monomers obtained in rigid-body (green spheres) and HADDOCK (violet spheres) docking experiments. These centroids differ from that of the average model by a maximum of 5 or 6 Å along each of the principal axes of Rev.