

Structure of HIV-1 RRE stem-loop II identifies two conformational states of the high-affinity Rev binding site

Received: 17 November 2023

Accepted: 22 April 2024

Published online: 17 May 2024

 Check for updatesJerricho Tipo^{1,2}, Keerthi Gottipati², Michael Slaton²,
Giovanni Gonzalez-Gutierrez² & Kyung H. Choi^{1,2,3}✉

During HIV infection, specific RNA-protein interaction between the Rev response element (RRE) and viral Rev protein is required for nuclear export of intron-containing viral mRNA transcripts. Rev initially binds the high-affinity site in stem-loop II, which promotes oligomerization of additional Rev proteins on RRE. Here, we present the crystal structure of RRE stem-loop II in distinct closed and open conformations. The high-affinity Rev-binding site is located within the three-way junction rather than the predicted stem IIB. The closed and open conformers differ in their non-canonical interactions within the three-way junction, and only the open conformation has the widened major groove conducive to initial Rev interaction. Rev binding assays show that RRE stem-loop II has high- and low-affinity binding sites, each of which binds a Rev dimer. We propose a binding model, wherein Rev-binding sites on RRE are sequentially created through structural rearrangements induced by Rev-RRE interactions.

During HIV-1 infection, the 9 kb positive-sense RNA genome is reverse-transcribed and integrated into the host cell's genome as proviral DNA^{1–3}. Transcription of the provirus DNA results in the generation of fully spliced, partially spliced, and unspliced mRNA transcripts⁴. At the early stage of HIV infection, only the intron-free fully spliced transcripts exit the nucleus for translation into HIV regulatory proteins, Tat, Rev, and Nef^{5–8}. Tat (trans-activator of transcription) recognizes its RNA-binding partner, the trans-activation response element (TAR), and enhances transcription of the full-length HIV RNA^{9–11}. However, the partially spliced and unspliced transcripts contain intronic sequences and are thus retained in the nucleus. Rev (regulator of viral expression) then specifically recognizes the Rev response element (RRE), present in the partially spliced and unspliced HIV transcripts, and facilitates nuclear export of these incompletely spliced viral transcripts^{12–15}.

The ~350 nucleotide (nt) long RRE is highly structured and predicted to fold into four or five stem-loops (SL), SLI–V^{16–19}. Two Rev-binding sites have been identified within RRE, a high-affinity site within

stem-loop II (SLII) and a secondary binding site within stem I. SLII forms a three-stem structure, IIA-C, and a purine-rich internal loop within stem IIB is predicted to be the initial high-affinity Rev-binding site (Fig. 1A)^{20,21}. The second Rev-binding site in stem I is also a purine-rich internal loop (designated IA) and is suggested to stabilize Rev multimers^{22,23}. Stem-loops III–V make relatively minor contributions to RRE activity, and even a construct that lacks all three stems (III/IV/V) retains ~33% of the wild-type (WT) RRE activity²⁴. Rev initially binds to the internal loop within stem IIB, which then triggers the oligomerization of as many as eight Rev molecules onto RRE^{20,25–27}. Rev oligomerization is required for the biological activity of the Rev–RRE complex, as binding of a single Rev molecule is insufficient for nuclear export^{28,29}. Upon formation of the homo-oligomeric Rev–RRE ribonucleoprotein (RNP), the C-terminal nuclear export signal of Rev recruits cellular Chromosomal Maintenance 1 (CRM1, also known as Exportin 1), and RAS-related nuclear protein to form the nuclear export complex^{30–35}. This nuclear export complex shuttles the intron-containing viral transcripts to the cytoplasm, where the viral RNA is

¹Department of Pharmacology and Toxicology, The University of Texas Medical Branch, Galveston, TX 77555, USA. ²Department of Molecular and Cellular Biochemistry, Indiana University, Bloomington, IN 47405, USA. ³Department of Biochemistry and Molecular Biology, Sealy Center for Structural Biology, The University of Texas Medical Branch, Galveston, TX 77555, USA. ✉ e-mail: kaychoi@iu.edu

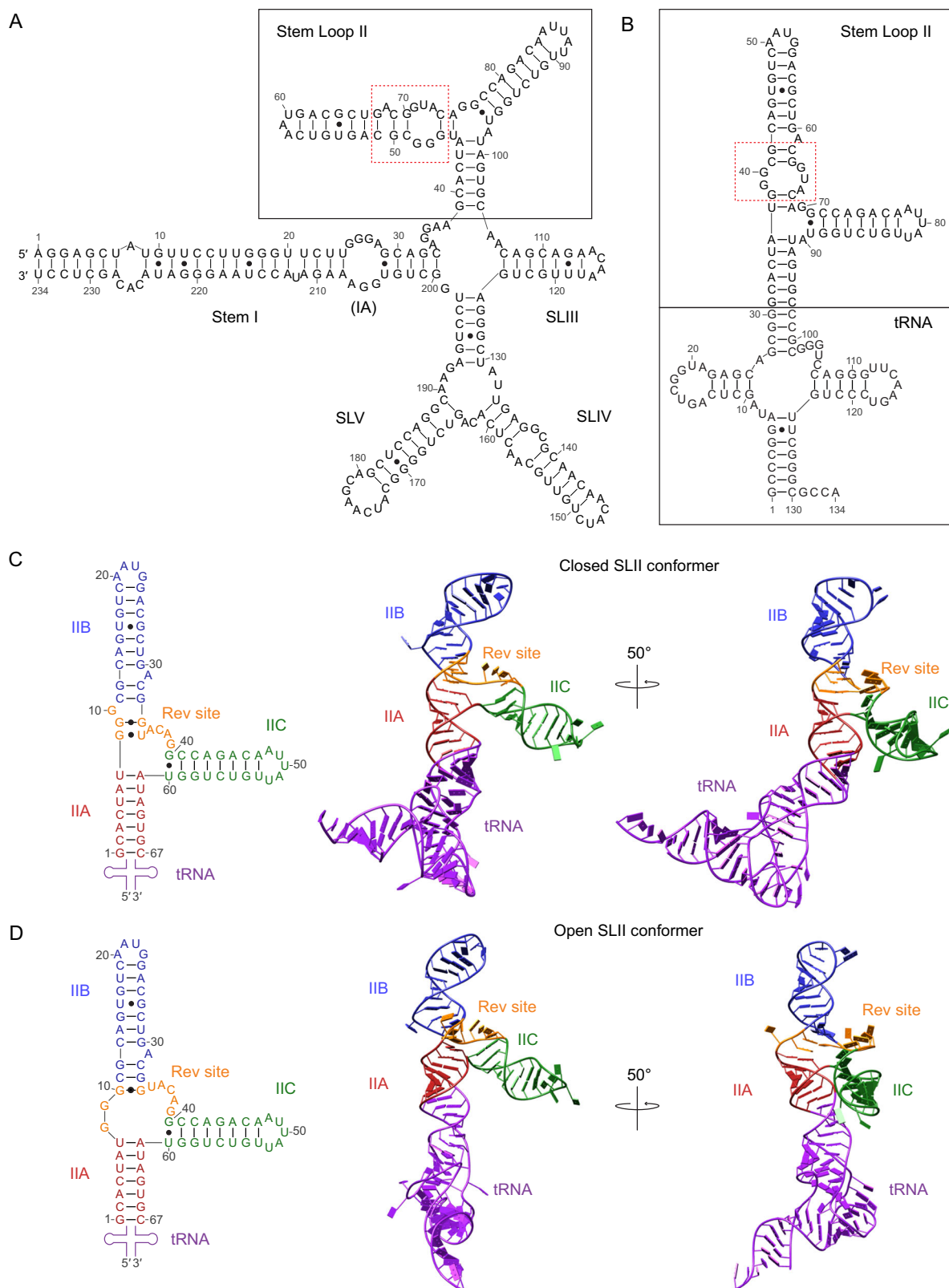


Fig. 1 | Structures of the Rev response element (RRE). **A** Secondary structure of the predicted four-stem RRE RNA. Stem-loop II (SLII) is boxed in black, and the high-affinity Rev-binding site is indicated with red dashed lines. **B** Design of SLII-tRNA. The SLII sequence is inserted into the anticodon stem of tRNA. Stem IIA is fused to the tRNA scaffold. The high-affinity Rev-binding site is indicated with red dashed lines. An additional guanine nucleotide (G53) was incorporated into stem IIB to engineer a tetraloop. **C** Structure of the SLII-tRNA in the closed conformation. The

secondary structure (left) and the crystal structure (right) are shown. SLII consists of stems IIA (red), IIB (blue), and IIC (green). The high-affinity Rev-binding site (the triple guanine motif and junction nucleotides between IIB and IIC) is shown in orange. The tRNA scaffold is shown in purple. **D** Structure of the SLII-tRNA in the open conformation. The secondary structure (left) and the crystal structure (right) are shown. The SLII stems and the high-affinity Rev-binding site are colored as in (C).

Table 1 | Anisotropic data collection and structure determination

	SLII-tRNA
Data collection	
Wavelength (Å)	1.00003
Space group	P2 ₁
Molecules/AU	2
Cell dimensions	
a, b, c (Å)	67.8, 81.3, 81.4
α, β, γ (°)	90, 99.1, 90
Resolution (Å)	57.2–2.85 (3.20–2.85) ^a
No. of unique reflections	12,674 (635)
I/σI	8.9 (1.0)
R _{meas} /R _{p.i.m}	0.10/0.05 (1.36/0.72)
CC _{1/2}	0.999 (0.487)
Completeness spherical (%)	61.5 (11.8)
Completeness ellipsoidal (%)	88.5 (68.3)
Redundancy	4.0 (3.5)
Refinement	
Resolution (Å)	57.2–2.85 (3.13–2.85)
R _{work} /R _{free}	0.22/0.25 (0.33/0.37)
No. of atoms	
Nucleotide	5743
R.M.S. deviations	
Bond length (Å)	0.003
Bond angles (°)	0.735
Mean B-factors (Å ²)	
Nucleotides	121.13

Single crystal was used for data collection.

^aValues in parentheses correspond to the highest resolution shell.

used for the translation of late-stage viral proteins and assembly of viral particles^{9–11}.

Nuclear export of HIV viral transcripts is essential for progression to later stages of the HIV life cycle, yet the mechanism by which RRE is recognized by Rev is not well understood. In particular, it is not clear how Rev discriminates between RRE-containing RNA and other RNAs, or how multiple copies of Rev assemble on RRE given that only two binding sites exist on RRE. Structural studies of the Rev–RRE complexes have been limited to minimized stem IIB hairpin RNAs with Arginine-rich motif (ARM) peptide or N-terminal domain of Rev^{7,20}, which do not fully recapitulate the three-way junction structure of SLII. It is also not clear how many Rev molecules bind SLII following the initial Rev binding within stem IIB. Gel shift assays show that the full-length SLII can bind at least three Rev molecules^{36,37}, whereas native mass spectrometry shows that SLII interacts with up to five Rev ARM peptides³⁸.

To better understand how the full-length SLII structure affects Rev recognition and oligomerization, we determined the X-ray crystal structure of RRE SLII and performed binding studies with SLII RNA mutants. We observe that the native SLII structure exists in distinct “closed” and “open” conformations that differ in the non-canonical base pairs in the high-affinity Rev-binding site. We show that initial Rev dimer binding occurs at the high-affinity binding site within the three-way junction, and subsequent Rev dimer formation occurs in the adjoining stem structures of SLII. Based on our results, we propose a sequential binding model of Rev–RRE, in which the conformational state of RRE SLII acts as a checkpoint to direct Rev binding and oligomerization.

Results

RRE SLII forms a lambda (λ)-shaped three-way junction structure and the high-affinity Rev-binding site is located in the three-way junction

The crystal structure of RRE SLII was determined using the tRNA-scaffold approach^{39,40}, where the 67 nt of HIV-1 RRE SLII replaced the anticodon loop of the human tRNA^{Lys} (134 nt total). The open stem IIA was fused to the anticodon arm of tRNA, maintaining 5′ and 3′ directionality (Fig. 1A, B). To facilitate crystallization, a guanine nucleotide (G23) was added to the apical loop of stem IIB to form a tetraloop structure (5′-AAUG-3′). Mutations in the apical loop of stem IIB do not affect Rev function⁴¹. The RNA crystallized in space group P2₁ with two molecules in the asymmetric unit, and the structure was determined to 2.85 Å resolution (Fig. S1). Data collection and refinement statistics are shown in Table 1.

The two RRE SLII molecules in the crystallographic asymmetric unit adopted two different conformations, one “closed” and the other “open”. Both molecules are “lambda (λ)”-shaped and consist of stems IIA, IIB, and IIC connected by a three-way junction (Fig. 1C, D). Stem IIB is coaxially stacked on stem IIA, and stem IIC projects out from the three-way junction at a -70° angle to stem IIA. The structures belong to the family A three-way junctions, where helix 3 (i.e., stem IIC) is roughly perpendicular to the coaxially stacked helices 1 and 2 (stems IIA and IIB) and the longest junction is between helices 2 and 3⁴². The predicted secondary structure of RRE SLII (Fig. 1A) was not correct near the three-way junction. In particular, the predicted internal loop in stem IIB (Fig. 1A) does not form. First, U⁷ is a part of stem IIA, where it base pairs with A⁶¹, instead of the predicted U⁷–A³⁹ base pair in stem IIB. Thus, stem IIA is one base pair longer than predicted. Second, the predicted G⁸–C³⁷ base pair in stem IIB does not form. As a result, the ³⁶ACAG³⁹ sequence is not incorporated into an internal loop and is instead single-stranded in the three-way junction (Fig. 1C, D). Thus, the three-way junction is much longer than predicted. Lastly, the non-Watson–Crick base pairs within stem IIB differ in the closed and open conformations. The high-affinity Rev-binding sequence, the triple guanine (⁸GGG¹⁰) motif, is incorporated into dsRNA at the base of stem IIB in the closed form; in the open form, only G¹⁰ is incorporated into stem IIB while ⁸GG⁹ is single-stranded in the three-way junction (Fig. 1C, D) (see next section for details). As a result, the three-way junction in the open conformation has three more nucleotides (G⁸, G⁹, U³⁵) compared to the closed form (Fig. 1C, D). It should be noted that the secondary structures of RRE determined via SHAPE studies disagree at the high-affinity Rev-binding site in SLII, and range from a closed internal loop to an open single-stranded structure^{17,19} (Fig. S2). The observed two conformations of SLII explain this discrepancy, as the closed SLII conformer is similar to the closed secondary structure¹⁷ and the open SLII conformer resembles that of the open secondary structure¹⁹ (Fig. S2).

Non-canonical base interactions within the three-way junction of RRE SLII differ in the closed and open conformations

The high-affinity Rev-binding site in RRE has been identified by RNase protection assays in the presence of Rev⁴³. Rev protects the ⁸GGG¹⁰ motif, ³³GG³⁴, and A³⁶ nucleotides within SLII (SLII numbering, Fig. 1), all of which are located in the three-way junction of the SLII crystal structures. The ⁸GGG¹⁰ motif in RRE SLII has also been shown to be essential for Rev binding and activation of Rev function in cells⁴³. Deletion of one G from the GGG motif resulted in only ~40% of Rev response, and deletion of two or all three G’s led to a complete loss of Rev binding and Rev response⁴³. Similarly, the substitution of ³³GG³⁴ to CC resulted in a complete loss of Rev binding and Rev response in cells. The closed and open conformers differ in non-canonical base pair interactions within the initial high-affinity Rev-binding site (⁸GGCG^{12/32}CGGUACA³⁸, Figs. 1 and 2). In the closed conformation,

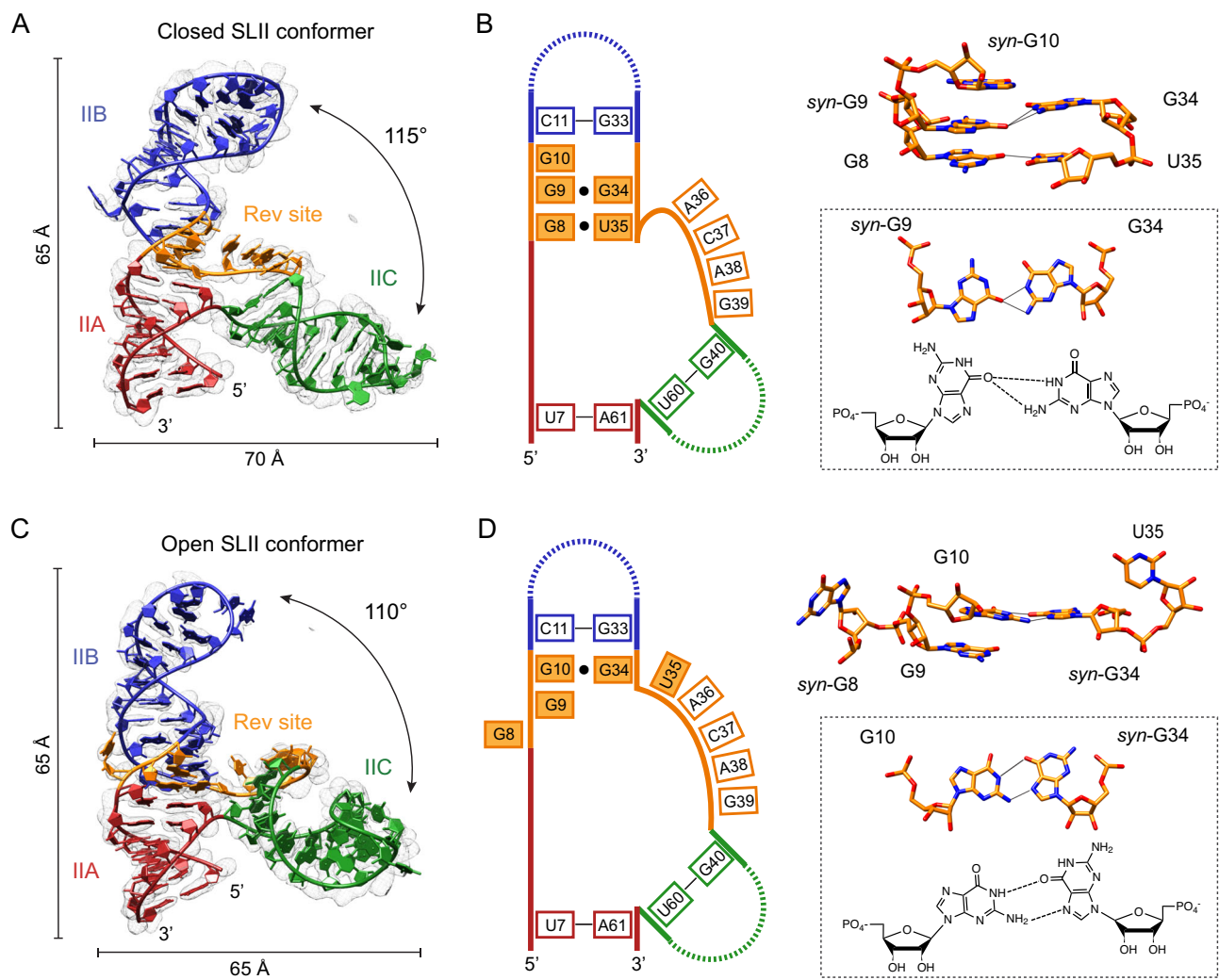


Fig. 2 | Interactions within the triple guanine (GGG) motif of the closed and open SLII conformers. **A** Closed RRE SLII conformer. The SLII closed conformer is shown with the $2F_oF_c$ electron density contoured at 1.0σ in light gray. The size of the molecule and the angle between stems IIB and IIC are indicated. **B** Tertiary interactions in the high-affinity Rev-binding site of the closed conformer. A schematic of base pair interactions in the three-way junction are shown (left). The filled boxes indicate the nucleotides involved in the interactions mediating non-canonical base pairs and the opening/closing of the three-way junction. The close-up view of the high-affinity Rev-binding site (the GGG motif, filled boxes) is shown (right). Note that closure of the three-way junction, and thus lengthening of stem IIB, is mediated by non-canonical $G^8\cdot U^{35}$ and $syn\text{-}G^9\cdot G^{34}$ base pairs. The $syn\text{-}G^{10}$ base

stacks on the $syn\text{-}G^9\cdot G^{34}$ base pair. The $syn\text{-}G^9\cdot G^{34}$ base pair interaction is shown along with chemical structures. Hydrogen bond lengths below 3.4 \AA are displayed. **C** Open RRE SLII conformer. The SLII open conformer is shown with the $2F_oF_c$ electron density contoured at 1.0σ in light gray. **D** Tertiary interactions in the high-affinity Rev-binding site of the open conformer. Schematic of base pair interactions (left) and close-up views of the high-affinity Rev-binding site (the GGG motif, right) are shown. Note that only G^{34} forms a non-canonical $G^{10}\cdot syn\text{-}G^{34}$ base pair, thus opening the three-way junction. The $G^{10}\cdot syn\text{-}G^{34}$ base pair interaction is shown along with their chemical structures. The structures and schematics are colored as in Fig. 1.

the A-form helical structure is maintained at the triple guanine motif (${}^8GGG^{10}$) via non-canonical base-pairing interactions of $G^8\cdot U^{35}$ and $G^9\cdot G^{34}$ (Fig. 2A). Interestingly, G^9 is in the higher-energy *syn* conformation and forms a non-canonical purine–purine base pair with G^{34} , where the Hoogsteen edge of *syn*- G^9 and the Watson–Crick face of G^{34} interact (Fig. 2A, B). The *syn*-*G-anti*-G pairing pattern has been observed in other RNA structures, as it minimizes the purine–purine clash in the center of the duplex, while utilizing its hydrogen bonding potential⁴⁴. G^{10} base is also in the *syn* conformation (*syn*- G^{10}) and stacks on *syn*- G^9 (Fig. 2B). Rotation of the guanine base about the glycosidic bond into the *syn* state is thus compensated by stabilizing the overall tertiary structure of the RNA by stacking and base pairing⁴⁵. As a result, the three-way junction of the closed form consists of single-stranded

${}^{36}ACAG^{39}$ between stems IIB and IIC. The ${}^{36}ACAG^{39}$ stretch is stacked on one another and is stabilized by interaction with the engineered tetraloop on stem IIB (${}^{20}AAUG^{23}$) of a symmetry-related molecule. The ${}^{37}CA^{38}$ base pair with ${}^{22}UG^{23}$ of the IIB tetraloop, and G^{39} (sugar edge) forms hydrogen bonds with A^{21} of the IIB tetraloop (Fig. S3).

The open conformation is marked by a non-canonical purine–purine base pair between G^{10} and G^{34} and melting of the $G^8\cdot U^{35}$ base pair interactions at the base of stem IIB (Fig. 2C, D). Thus, stem IIB in the open conformation is shorter than that in the closed form. The *syn*- G^9 and *syn*- G^{10} seen in the closed conformation are flipped into the standard *anti* conformation, whereas *anti*- G^{34} is flipped to the *syn* conformation. G^{10} of the ${}^8GGG^{10}$ motif forms a non-Watson–Crick base pair with *syn*- G^{34} , where the Watson–Crick face of

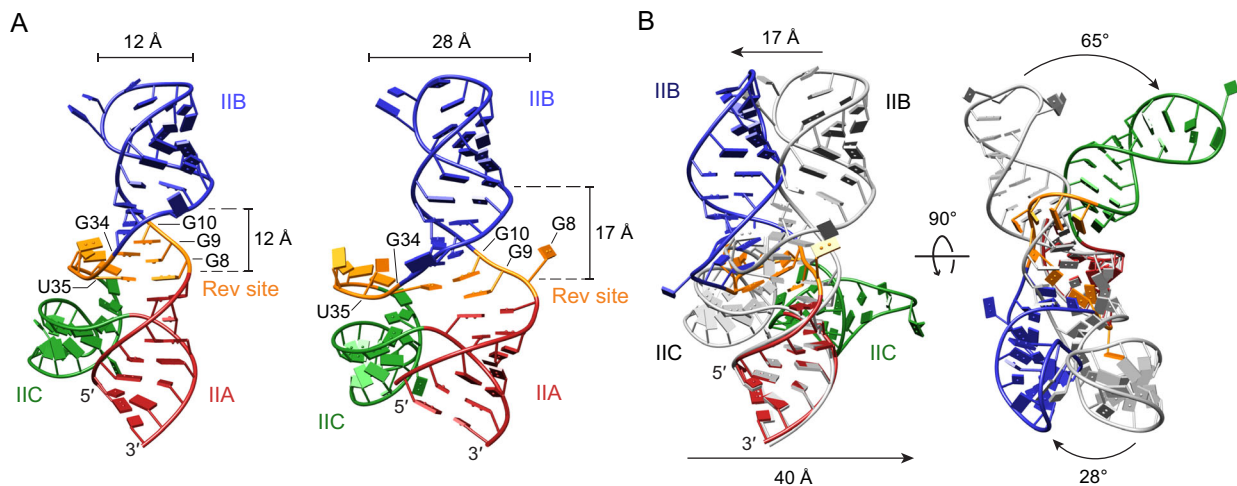


Fig. 3 | Comparison between the closed and open SLII conformers.

A Comparison of the RNA major grooves in the closed and open SLII conformers. In the closed conformer (left), the width of the major groove formed between stems IIA and IIB at the GGG motif is -12 \AA , and the distance between the C1' atoms of G8 and U35 is -12 \AA . In the open SLII conformer (right), the same major groove is

widened to -17 \AA , and the distance between the C1' atoms of G8 and U35 is widened to -28 \AA . **B** Alignment of the two conformations at stem IIA. The closed conformer is colored in gray, and the open conformer is colored as in (A). The superposition of closed and open conformers with IIA indicates that opening of the high-affinity Rev-binding site rotates stems IIB by -28° and IIC by -65° .

G¹⁰ forms hydrogen bonds with the Hoogsteen edge of *syn*-G³⁴ (Fig. 2D). The rest of the GGG motif (⁸GG⁹) and ³⁵UACAG³⁹ are single-stranded in the three-way junction. G⁸ and U³⁵ are flipped out of the helix and G⁹ is stacked between G¹⁰ and the U⁷-A⁶¹ base pair (stem IIA). The ³⁷CAG³⁹ in the single-stranded region between stems IIB and IIC interact with ²¹AUG²³ in the stem IIB tetraloop of a neighboring molecule, similar to the closed conformation (Fig. S3).

The different non-canonical interactions in the closed and open forms alter the arrangement of the three stems, pivoting stems IIB and IIC at the three-way junction. In the closed conformer, the axes of the three stems are approximately in the same plane (Fig. 3A). Opening of the three-way junction rotates stems IIB and IIC out of this planar configuration. When stem IIA of the two conformers is superposed, stem IIB in the open form is related by -28° to that of the closed form, resulting in -17 \AA translation at the tip of the apical loop (Fig. 3B). Stem IIC in the open conformer is related by 65° to that of the closed conformer, resulting in -40 \AA shift at the tip of its apical loop (Fig. 3B). The closed and open forms of full-length RRE SLII superpose with an r.m.s.d. of 6.6 \AA , yet the respective stems in the two forms are similar and superpose with r.m.s.d. values of 0.9, 2.6, and 1.4 \AA for stems IIA, IIB, and IIC, respectively (Table S1).

Opening of the three-way junction widens the major groove of the GGG motif by rearrangement of stems IIB and IIC

The closed and open forms of RRE SLII exhibit different conformations at the high-affinity Rev-binding site utilizing non-canonical base pairs. In the closed form, the GGG motif is locked into a dsRNA helix by the G⁸·U³⁵ wobble and *syn*-G⁹·G³⁴ purine-purine base pairs. In the open form, the GGG motif is more exposed by melting of the G⁸·U³⁵ and *syn*-G⁹·G³⁴ base pairs, and flipping of G⁸ and U³⁵ out of the helix. As a consequence, the C1'-C1' distance between G⁸ and U³⁵ increases from 12 \AA in the closed conformation to 28 \AA in the open conformation (Fig. 3A). The major groove formed at the GGG motif also widens from 12 \AA (between P atoms of G⁸ and G³⁰) in the closed conformer to 17 \AA (between P atoms of G⁸ and G²⁷) in the open conformer (Fig. 3A). In addition to widening of the major groove, loosening of these interactions in the open conformer increase the solvent accessible surface area of the three-way junction from 2350 to 2600 \AA^2 . Previous structural studies using stem IIB hairpin RNAs show that the major groove near the GGG motif widens from -9 \AA (P-P distance, PDB 1CSL) in the unbound form to -15 \AA (PDB 4PMI) in the bound form⁴⁶. Thus, the open

conformation of RRE SLII with an enlarged major groove (-17 \AA) and exposed surface area at the GGG motif suggests that Rev would selectively bind to the open conformation over the closed conformation (see Rev-SLII modeling section below).

RRE SLII contains two Rev recognition sites, each of which can support Rev dimer binding

Rev consists of an ARM, a bipartite oligomerization domain (OD), and a nuclear export signal, separated by linker regions^{29,47-52} (Fig. 4A). The α -helical ARM binds RNA and hosts the nuclear localization signal. The OD forms two α -helical regions and facilitates Rev-Rev dimer and oligomer formation. The C-terminal nuclear export signal is responsible for the recruitment of host nuclear export factors³³. NMR and crystal structures of the stem IIB hairpin and Rev ARM (or N-terminal domain) complexes show that Rev ARM binds across the major groove of the RNA^{46,53-56}. The Arg and Asn residues of ARM (R38, R39, N40, R41, R44, and R46) interact with the triple guanine motif by base-specific and non-specific interactions⁴⁶. However, the stem IIB RNAs used in these structural and in vitro binding studies have an altered stem IIB sequence that stabilizes the GGG motif into an internal loop and are thus different from those observed in our native RRE SLII structures (Fig. S2B). Further, N40 of Rev, the major contributor to base-specific interactions with the GGG motif interacts differently in the crystal and NMR studies. In the crystal structure, N40 interacts across G⁹ and G³⁴, whereas in the NMR structure it interacts with G⁹ and A³⁵ despite both structures having the same sequence near the GGG motif^{46,55}. Thus, the Rev interactions with the GGG motif seen in the previous structures may not represent native ARM recognition in SLII, and the GGG motif in native SLII may present different functional groups on the surface to facilitate novel base-specific Rev interactions.

To understand how the RNA structure mediates Rev oligomerization, we next investigated Rev interactions with RRE SLII using gel shift assays. Two Rev constructs were used to differentiate between protein-RNA and protein-protein interactions. First, a Rev construct with only the ARM (Rev_{ARM}) domain was used to identify RRE SLII recognition events (Figs. 4A and S3A). Second, a Rev construct limited to dimerization (Rev₇₀) was used to assess interaction between Rev monomers^{22,57,58}. Rev contains a dimerization interface and a higher-order oligomerization interface. In Rev₇₀, L12S and L60R mutations were introduced within the higher-order oligomerization interface, thus limiting multimerization to dimer formation (Figs. 4A and S3A)⁵⁷.

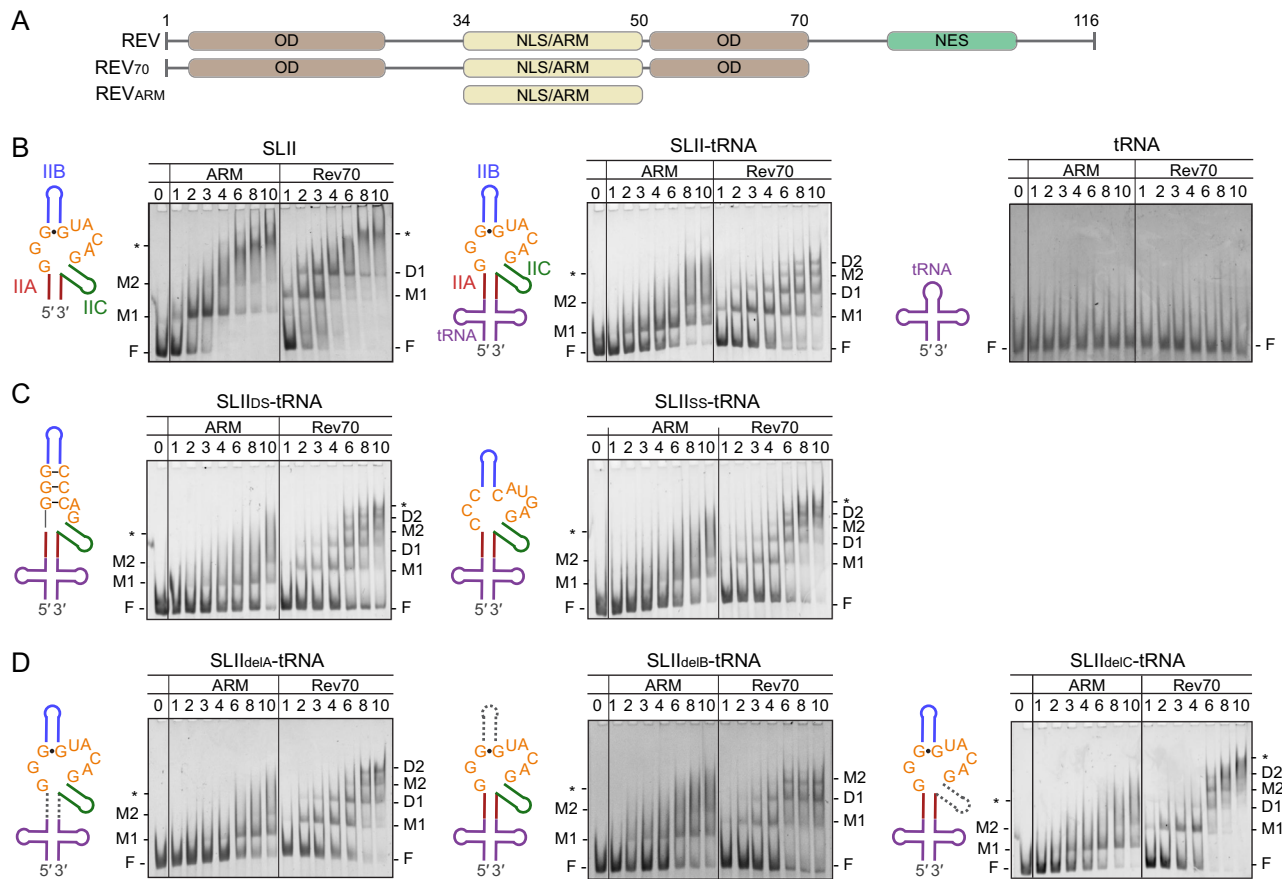


Fig. 4 | Rev and RRE SLII binding assays. **A** Schematics of Rev domains. Arginine-rich motif (ARM, beige) required for RNA interaction and oligomerization domains (OD, brown), facilitating oligomerization of Rev molecules are indicated. Rev_{ARM} was used to test RNA recognition and Rev_{70} was used to assess dimer formation. **B** Interaction of SLII, SLII-tRNA, and tRNA with Rev proteins. Binding assays were performed with 500 nM RNA and increasing molar excess of Rev proteins. A schematic of the RNA used in the binding assay is shown to the left of the gel. SLII-tRNA is colored as in Fig. 1. The sequences of RNAs are listed in Fig. S4. SLII or SLII-tRNA contains two Rev_{ARM} binding sites, M1 and M2, representing a high-affinity and a low-affinity site. The two sites individually accommodate Rev_{70}

dimers, indicated as D1 and D2. Unbound RNA is denoted by F and higher-order oligomers are denoted with an asterisk (*). **C** Interaction of SLII three-way junction mutants with Rev proteins. SLII_{DS}-tRNA mimics the closed conformation by double-stranded C-G base pairs (left) and SLII_{SS}-tRNA mimics the ssRNA three-way junction structure (right). **D** Interaction of SLII stem deletion mutants with Rev proteins. SLII_{delA}-tRNA, SLII_{delB}-tRNA, and SLII_{delC}-tRNA mutants lack stems IIA, IIB, and IIC, respectively. The deleted stems are indicated with gray dotted lines in the RNA schematics. Each binding assay was performed three times with similar results. Uncropped gel images are provided in the Source Data file.

The binding assays show that RRE SLII has two Rev_{ARM} binding sites, high- and low-affinity sites. The first Rev_{ARM} binds to RRE SLII at an equimolar ratio of RNA at a high-affinity site (M1 in Fig. 4B, left) and the complex persists even at a tenfold excess of protein. The binding of Rev_{ARM} to a second, low-affinity site (M2 in Fig. 4B, left) is seen starting with a fourfold molar excess of protein to RNA, at which point there is no longer unbound SLII. The major Rev-bound species shift to higher-order complexes upon a sixfold excess of protein, appearing as a smear, suggesting non-specific Rev_{ARM} -RNA interactions. The initial binding of Rev_{70} on SLII is similar to that of Rev_{ARM} . The first Rev_{70} complex appears at the same molar concentration as the first Rev_{ARM} -SLII complex, suggesting that Rev_{70} binds at the high-affinity binding site (M1 in Fig. 4B, left). However, the next Rev_{70} and SLII complex forms much earlier than the binding of Rev_{ARM} to the low-affinity site, suggesting that two Rev_{70} molecules dimerize at the high-affinity site (M1 and D1 in Fig. 4B, left). As the protein concentration increases, the Rev_{70} dimer and RRE SLII complex shift into non-discernible higher-order oligomeric complexes.

Next, we studied the binding of Rev_{ARM} and Rev_{70} to the SLII fused to the tRNA scaffold (SLII-tRNA). Rev_{ARM} and Rev_{70} bind SLII-tRNA with a similar stepwise binding pattern as SLII (Figs. 4B, middle, and S3B). A single Rev_{ARM} binding to the high-affinity site is the

predominant RNP species up to a sixfold molar excess of protein, at which point binding of the second Rev_{ARM} to the low-affinity site occurs (M1 and M2 in Fig. 4B, middle). In the case of Rev_{70} , the monomer and dimer complexes on the high-affinity site are formed at similar protein concentrations as Rev_{70} -SLII complexes (M1 and D1 in Fig. 4B, middle). Compared to SLII interactions, additional Rev_{70} monomer and dimer complexes at the low-affinity site were clearly visible at a sixfold molar excess of protein (M2 and D2 in Fig. 4B, middle). Rev does not interact with the tRNA scaffold, since binding of both Rev_{ARM} and Rev_{70} to tRNA alone is minimal even at a tenfold excess of protein (Fig. 4B, right).

Taken together, binding assays with monomeric Rev_{ARM} and dimer-limited Rev_{70} show that SLII has two Rev ARM binding sites, the high- and low-affinity sites. Following initial Rev ARM binding, the Rev_{70} dimer assembles independently for each site, suggesting that Rev dimer formation is driven by protein-protein interactions between Rev monomers. Interestingly, the Rev and SLII-tRNA complexes show more discrete bands than SLII alone, suggesting that the tRNA scaffold prevents aberrant oligomer formation on RRE SLII by providing structural stability. The fusion of SLII to the tRNA scaffold mimics the connectivity of SLII to the rest of the full-length RRE (Fig. 1A, B). Since the stem-loops III-V of RRE do not contribute to Rev

binding and RRE activity directly²⁴, they may function similarly to the tRNA scaffold, providing structural stability to the Rev-binding sites in RRE (stems I and II) to assemble the functional RNP.

Recognition of the initial Rev-binding site within the three-way junction requires both sequence and structural motifs

In the high-affinity purine-rich Rev-binding site, the triple guanine motif (⁸GGG¹⁰) and ³⁴GUACAG³⁹ make up the three-way junction and are responsible for the closed and open conformations. To investigate how the conformation of the three-way junction affects Rev-binding, binding assays were performed with SLII constructs, in which mutations at the three-way junction were introduced to mimic the closed and open SLII conformations. Since the tRNA scaffold provides stability to the SLII structure and allows better resolution of the resulting complexes, mutations were introduced in SLII-tRNA. To mimic the closed SLII conformer, the ⁸GGG¹⁰/³⁴GUAC³⁷ sequence was substituted with GGG/CCC such that the triple guanine motif is incorporated into dsRNA helix (SLII_{DS}), thus simulating closure by the G⁸·U³⁵ and G⁹·G³⁴ base pairs (Figs. 4C, left, and S3B). Binding assays with SLII_{DS}-tRNA show that the first Rev_{ARM} and Rev₇₀ bind the RNA with reduced binding efficiency than WT SLII-tRNA (Fig. 4C, left). Thus, the initial recognition of the high affinity, triple guanine motif by Rev_{ARM} is impaired by the closure of the SLII three-way junction, indicating that the GGG motif in the open conformation is required for primary Rev ARM binding. However, dimerization of Rev₇₀ on the RNA was not affected (Fig. 4C, left), suggesting that Rev dimerization is independent of recognition of the triple guanine motif^{37,57,59}. Similarly, Rev_{ARM} binding to the low-affinity site is also delayed, yet dimerization of Rev₇₀ at the site is comparable to the WT.

Next, the ⁸GGG¹⁰/³⁴GUAC³⁷ sequence was substituted with CCC/CAUG (SLII_{SS}). Elimination of the triple guanine motif prevents G⁹·G³⁴ and G⁸·U³⁵-mediated closure of the SLII structure, forming single-stranded RNA as in the open conformation (Figs. 4C, right, and S3B). As expected, Rev_{ARM} has a weak binding affinity for SLII_{SS}-tRNA, leading to delayed Rev binding on the initial binding site. Binding of a second Rev_{ARM} is also delayed, and a discrete band is visible at an eightfold molar excess of protein. Rev₇₀ dimerizes efficiently at both sites, and Rev₇₀ complexes quickly shift into non-specific, higher-order complexes. Thus, the closure of the three-way junction or elimination of the GGG motif interferes with specific Rev_{ARM} recognition of the initial binding site but does not affect subsequent Rev₇₀ dimerization. It was shown that substitutions of the GGG motif to CCC in RRE abolished Rev activity in cells⁴³. Thus, although Rev interacts with SLII mutants and forms multiple RNP complexes, the resulting Rev oligomer and SLII complexes are not correctly configured for a functional RNP to recruit downstream nuclear export factors.

Rev dimerization at the high- and low-affinity sites of RRE SLII requires stems IIC and IIB, respectively

Electrophoretic mobility shift assays show that Rev₇₀ can bind at least four sites on RRE SLII (Fig. 4B). Since the multimerization of Rev molecules requires protein-RNA as well as protein-protein interactions, unknown Rev-binding sites within RRE SLII branches may exist. To determine the Rev-binding sites in RRE SLII, binding assays were performed with SLII-tRNA mutants in which stems IIA, IIB, or IIC were deleted while preserving the initial high-affinity Rev-binding site (Figs. 4D and S3B). Deletion of stems IIA, IIB, and IIC (SLII_{delA}, SLII_{delB}, SLII_{delC}) resulted in similar Rev_{ARM} binding phenotypes to WT SLII-tRNA with two bands (Fig. 4D), indicating that the Rev recognition site within the three-way junction of RRE SLII is likely intact.

Binding of Rev₇₀ to stem deletion constructs, however, shows different patterns of complex formation. Binding of Rev₇₀ to the SLII_{delA}-tRNA was similar to that of the WT SLII-tRNA, suggesting that specific interaction with stem IIA is not required in stabilizing Rev multimers (Fig. 4C, left). In contrast, Rev₇₀ binding assays with

SLII_{delB}-tRNA and SLII_{delC}-tRNA suggest that stems IIB and IIC are necessary for coordinating Rev dimerization. Deletion of stem IIB does not affect initial Rev₇₀ binding or dimerization at the high-affinity site. Additional Rev₇₀ binding to the low-affinity site is also similar to the WT SLII-tRNA (Fig. 4D, middle). However, Rev₇₀ does not dimerize on the low-affinity site even at a tenfold molar excess of protein, indicating that Rev₇₀ dimerization on the low-affinity site requires stem IIB. In contrast, the deletion of stem IIC inhibits dimerization of Rev₇₀ within the high-affinity site, as the progression from Rev₇₀ monomer to dimer-bound SLII_{delC} species is stunted up to a fourfold molar excess of Rev₇₀ (Fig. 4D, right). As the protein concentration increases, the formation of Rev-bound species at the low-affinity site ensues. This suggests that the binding of the second Rev molecule at the high-affinity site requires interaction with stem IIC. Furthermore, the absence of stem IIC leads to the formation of non-specific higher-order complexes at high protein concentrations, in contrast to Rev₇₀ binding on WT SLII-tRNA or SLII_{delB}-tRNA (i.e., discrete band formation). Thus, stem IIC seems to provide structural stability to SLII to prevent non-specific Rev interactions with RNA.

Taken together, our Rev-binding assays show that SLII stabilized by the tRNA scaffold can assemble four Rev molecules as two dimers in discrete steps. The initial Rev binds at the high-affinity site (GGG site) via specific Rev ARM and RNA interaction. The second Rev then rapidly dimerizes with the first via protein-protein interactions between Rev OD. The second Rev likely binds in the major groove of the adjoining stem IIC. As the concentration of Rev protein increases, a third Rev binds at a second, low-affinity site near the three-way junction of RRE SLII, which is distinct from the initial Rev-binding site. Finally, the fourth Rev dimerizes with the third Rev molecule likely in stem IIB. Notably, the assembly of Rev on the SLII_{delB} mutant is halted at three Rev proteins even at a tenfold molar excess of protein, suggesting that the RNA structure can control the sequential Rev assembly on RRE. We propose that the conformational states of RRE SLII may act as a checkpoint to regulate its recognition by Rev (see “Discussion”).

Discussion

Model of Rev dimerization on the initial high-affinity site within the full-length RRE SLII

During HIV infection, partially spliced and unspliced viral RNAs need to be exported from the host cell nucleus to the cytoplasm to synthesize viral proteins and assemble virions. To mediate nuclear export, HIV uses specific interaction between Rev protein and an intronic RNA sequence RRE. Multimeric Rev proteins bind to RRE and recruit the nuclear export factors for cytoplasmic translocation of the viral RNAs. However, the architecture of the highly specific RRE-Rev multimeric complex is unknown, largely due to the lack of an RRE structure. In the structure of a Rev₇₀ dimer in complex with a modified stem IIB hairpin RNA (PDB code 4PMI), one Rev monomer interacts with the GGG sequence within the widened major groove and a second Rev monomer binds non-specifically along the adjacent major groove⁴⁶. Our RRE SLII structure now shows that the purine-rich, high-affinity Rev-binding site (the GGG motif) is incorporated into the three-way junction rather than the predicted internal loop in stem IIB. Thus, the specific Rev-RRE SLII interactions are likely different from those reported in the Rev and stem IIB hairpin complexes (Fig. S3).

To provide insight into RNA-directed Rev dimerization on RRE, we modeled a Rev dimer on RRE SLII. Comparison of the closed and open structures suggests that pivoting of stems IIB and IIC at the three-way junction in the open conformation increases the width of the major groove at the high-affinity Rev-binding site by ~5 Å (Fig. 3A). Since Rev binding requires a widened major groove of the RNA, we chose the open conformer of SLII for modeling the SLII-Rev dimer complex. Guided by the Rev dimer and stem IIB hairpin complex structure⁴⁶, the first Rev molecule was placed in the major groove formed by the GGG motif. Rev monomer fits well in the open form of SLII within the

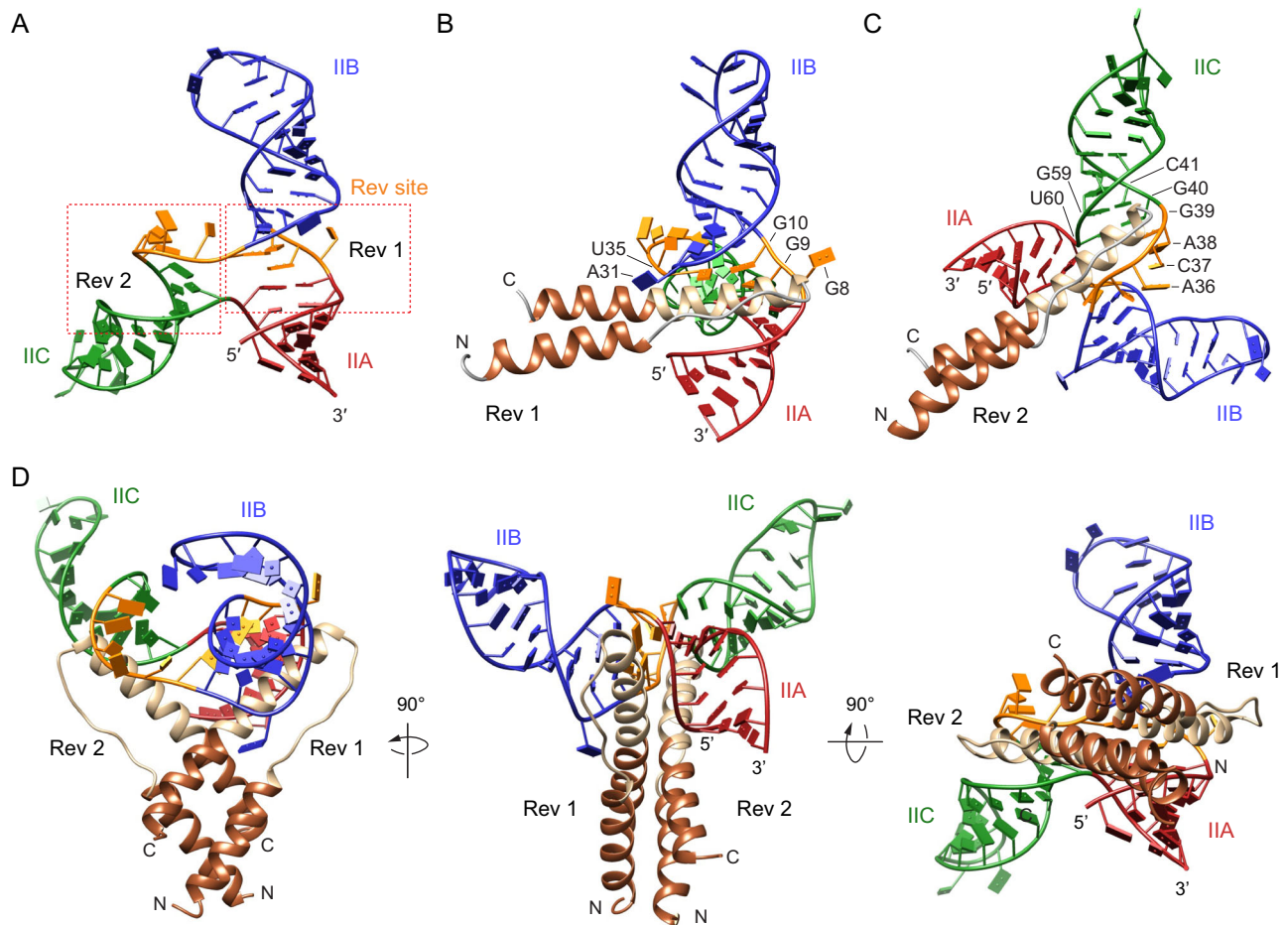


Fig. 5 | Model of Rev dimer binding onto the open conformation of RRE SLII. **A** Rev dimer binding site in RRE SLII. The first Rev binds in the high-affinity binding site, and the second Rev likely binds in the major groove next to the high-affinity site. The two Rev-binding sites are indicated with red boxes. **B** First Rev binding on the initial high-affinity binding site. Rev arginine-rich motif (ARM) and two oligomerization domains (OD) are shown in beige and brown, respectively. The ARM of

the first Rev molecule was placed within the widened major groove formed by $^8\text{GGG}^{10}$ of the open SLII conformer. **C** Second Rev molecule binding within stem IIC. The ARM of the second Rev molecule was placed within the major groove at the base of stem IIC. **D** Binding of the Rev dimer on the open SLII conformation. The Rev dimer is stabilized by the interaction of the ARM domains within the RNA major grooves as well as between neighboring Rev ODs.

widened major groove of the three-way junction without major steric clash (Fig. 5A, B). In this orientation of the Rev monomer, the ARM domain would interact with G^9 and G^{10} , *syn*- G^{34} , and the flipped A^{31} bulge nucleotide. By virtue of dimerization, the second Rev monomer is poised to interact with the major groove of the adjacent stem-loop IIC along the three-way junction (Fig. 5C, D). The N-terminal domain of Rev (Rev₇₀) has been crystallized in various dimer forms, where the angle between the two ARM peptides ranges from 40 to 140°, suggesting that Rev may accommodate differently arranged RNA structures^{46,60,61}. Thus, to fit the Rev dimer into the three-way junction of RRE SLII, the angle between the two ARMS of the Rev dimer was increased from 50° (in the IIB hairpin complex) to 115° (Fig. 5C).

In the SLII–Rev dimer model, the Rev dimer sits astride the three-way junction with each monomer binding along the major grooves toward stems IIB and IIC (Fig. 5D). The first Rev monomer binds in the high-affinity site in the three-way junction and dimerizes with the second Rev monomer that binds in the major groove toward stem IIC. This arrangement of the Rev dimer is consistent with previous Rev-binding studies where in the presence of Rev, RRE showed reduced SHAPE activity for the nucleotides in the three-way junction (G^9 , G^{33} , A^{36} , and A^{38} ; SLII numbering)⁶². Further, our Rev-binding assays show that the initial, high-affinity Rev dimer is assembled at the three-way junction and adjoining stem IIC (Fig. 4D). In contrast to the enlarged major groove of the first Rev-binding site, the width of the major

groove in the second Rev-binding site is ~ 10 Å in both SLII conformations (Fig. 5A). Thus, binding of a second Rev would require widening of the major groove in the second Rev-binding site. It is possible that the binding of the first Rev molecule in the high-affinity site changes the conformation of the second Rev-binding site, making it more conducive for Rev binding. The G^{40} - U^{60} wobble pair at the foot of stem IIC may facilitate the widening of the major groove to allow binding of the second Rev. Previously, an RRE–Rev binding model was proposed based on small-angle X-ray scattering (SAXS) data of RRE¹⁸. Because the spacing (~ 55 Å) between the stems II and I in the “A”-shaped RRE model is similar to the spacing between the two RNA-binding domains in Rev dimers, an initial Rev dimer was proposed to bridge the stems II and I Rev-binding sites (IIB and IA sites)¹⁸. Our Rev-binding model suggests that SLII alone binds the initial Rev dimer, and is thus different from the SAXS-based model.

RRE modulates the degree of Rev binding via progressive conformational changes

RRE contains multiple Rev-binding sites, yet only two sequence-specific Rev-binding sites have been identified, one in the stem II junction and the other in stem IA, suggesting that RRE utilizes largely non-sequence-specific interactions to form the RNP complex. Additional Rev-binding sites on RRE are thus likely created sequentially through structural rearrangement of RRE as new Rev molecules

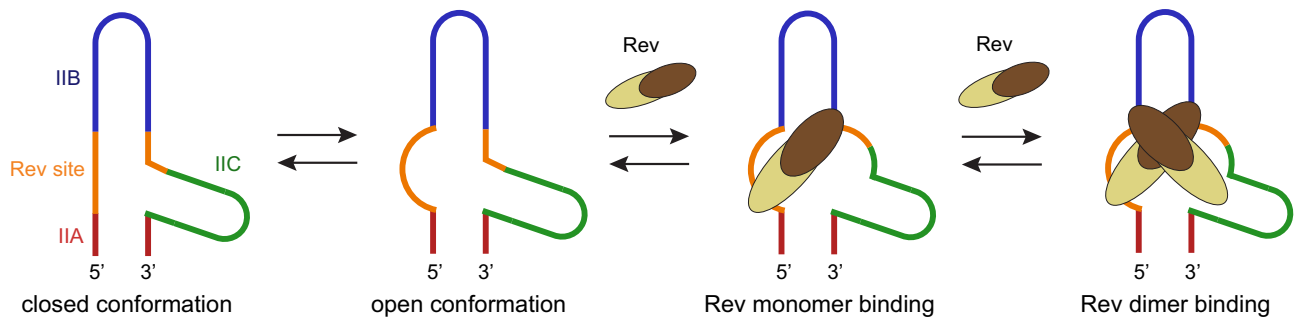


Fig. 6 | Model of progressive Rev binding on RRE SLII. The closed and open RRE SLII exist in an equilibrium. The conformational changes between the closed and open SLII depict the first step in Rev oligomerization: recognition and binding of the high-affinity Rev-binding site. A single Rev molecule binds the open SLII conformer at the high-affinity binding site and induces a conformational change in the

nearby major groove at stem-loop IIC. This allows a second Rev monomer to dimerize on the bound Rev molecule via interaction with stem IIC. The binding of additional Rev molecules continues to induce conformational changes at nearby sites on RRE until all Rev-binding sites are occupied.

interact. The native full-length SLII exists in a dynamic equilibrium of two alternative conformations, closed and open, mediated by non-canonical base pairing in the three-way junction. We propose that the closed and open states have low- and high-affinity for Rev, respectively, and that Rev preferentially binds to the open conformation (Fig. 6). Binding of the first Rev molecule then changes the conformation of the nearby Rev-binding site on RRE and increases its binding affinity for a second Rev molecule (positive cooperativity). Similar sequential conformational changes upon Rev binding to a nearby site will continue until all Rev-binding sites are occupied. This cooperative binding model suggests that RRE conformational states regulate the binding of Rev molecules in a stepwise fashion by modulating RNA-binding affinity. Previously, single-molecule FRET studies of RRE and Rev in the presence of DEAD-box helicase DDX1 showed that DDX1 accelerates the initial Rev interaction with RRE⁶³. It was proposed that DDX1 could act as an RNA folding chaperon and induce an altered RRE RNA structure that has a higher affinity for Rev. Our Rev-binding model facilitated by progressive conformational changes in RRE would be consistent with the model, where DDX1 facilitates a conformational change from the closed to open state of RRE, thus accelerating the initial Rev interaction with RRE. We show that Rev oligomerization on RRE SLII is a controlled process. Native RRE and Rev-bound RRE species exist in equilibrium and assemble discrete RNA–protein complexes, rather than exhibiting a full shift into higher-order oligomeric complexes (Figs. 4B and 6). This highly specific initial Rev binding and subsequent cooperative multimerization of Rev likely allows RRE to function as a rheostat for Rev concentrations.

Regulation of the multimeric Rev–RRE complex acts as a biological timer for the formation of the nuclear export-competent RNP complex and thus transition into late-stage HIV viral infection. The importance of this regulatory axis is demonstrated in HIV-positive individuals, in which more active Rev and RRE isolates were associated with individuals with advanced disease states^{64,65}. Due to the flexibility of RNA secondary structures, mutations within the RRE may provide more immediate modulation of the Rev–RRE axis to upregulate nuclear export under immunocompromised instances¹⁹. Since the export of RRE-containing RNAs is essential for HIV replication, the RRE and Rev complex is an attractive therapeutic target to trap HIV in a latent phase. The structure of RRE SLII would be useful to develop small molecules that bind at the three-way junction and inhibit the conformational changes required for cooperative binding of Rev molecules and the formation of the functional Rev–RRE complex.

Methods

Design and purification of tRNA-scaffolded SLII RNA constructs

The SLII–tRNA fusion constructs were designed in a pBlueScript II SK(+) vector as previously described^{39,40,66}. Briefly, the stem-loop II

sequence of HIV-1 RRE (nt 7353–7418, GenBank: AF033819.3) was inserted into the anticodon loop of a human lysine tRNA to generate the RRE SLII–tRNA chimera. Stem IIA is continuous with the anticodon stem of tRNA. An additional G54 was added to generate a stable tetraloop (⁵¹AAUG⁵⁴) in stem IIB. To test interaction with Rev proteins, deletion and substitution mutants, SLII_{DS}, SLII_{SS}, SLII_{delA}, SLII_{delB}, and SLII_{delC} with the tRNA scaffold were designed (Table S1) and generated using site-directed mutagenesis of SLII–tRNA. The secondary structure prediction using the RNAfold server within the ViennaRNA Web Services (<http://ma.tbi.univie.ac.at/>)⁶⁷ indicated that the modified RNA constructs maintain the tRNA structure.

The tRNA fusion constructs were purified as previously described³⁹. The plasmids were freshly transformed into BL21 *E. coli* cells and grown overnight in 2 L of 2× TY medium with 100 µg/mL ampicillin at 37 °C. The cells were harvested at 4000 × *g* for 30 min and resuspended in 10 mL of buffer containing 10 mM magnesium acetate and 10 mM Tris-HCl, pH 7.4. The RNA was extracted by adding 12 mL of acid phenol:chloroform (5:1) mixture and gently agitated for 1 h at room temperature. The solution was next centrifuged at 21,000 × *g* for 30 min at 20 °C. The aqueous phase was extracted and mixed with 0.1 volume of 5 M NaCl and 2 volumes of absolute ethanol for RNA precipitation. Following incubation at –20 °C overnight, the RNA was pelleted by centrifugation at 21,000 × *g* for 30 min at 4 °C, dissolved in buffer A (40 mM sodium phosphate, pH 7.0), and passed through a 0.22 µm filter. The RNA was loaded onto a HiLoad 16/20 Q ion-exchange column (Cytiva, Marlborough, MA). The desired RNA eluted between 520 and 560 mM NaCl in buffer A. The fractions were analyzed on a urea–acrylamide gel. RNA-containing fractions were then pooled and buffer-exchanged into water. Typical yield of RNA is 30–40 mg from a 2-L culture.

In vitro transcription and purification of SLII and tRNA control

The DNA templates for SLII and tRNA^{Lys} were individually generated using PCR from the SLII–tRNA plasmid. RNA was in vitro transcribed using standard T7 RNA polymerase transcription methods⁶⁸. The transcription reaction was prepared by combining 10 mM rNTPs, 20 nM DNA template, and 200 µg/mL T7 RNA polymerase⁶⁹ in buffer containing 40 mM Tris-HCl, pH 8.3, 20 mM MgCl₂, 2.5 mM spermidine, 10 mM DTT, and 0.01% Triton X-100. In vitro transcribed RNA was diluted in buffer A (40 mM sodium phosphate, pH 7.0) and loaded onto a MonoQ 5/50 GL ion-exchange column (Cytiva, Marlborough, MA). The desired RNA eluted between 560 and 600 mM NaCl in buffer A. Fractions were analyzed on a urea–polyacrylamide gel, and RNA-containing fractions were pooled and buffer-exchanged into water. Typical yield is ~20 µg RNA from 1 mL T7 RNA transcription reaction.

Expression and purification of Rev proteins

The design of the HIV Rev₇₀ protein is as previously described^{22,70}. Rev₇₀ includes only the ordered N-terminal domain (NTD, amino acids 1–70, UniProt ID: P69718). Rev₇₀ was designed as a GB1 (B1 domain of Streptococcal Protein G)-fusion protein within the pET-28 vector and contains E47A, L12S, and L60R mutations, the latter two of which limit Rev higher-order oligomerization to dimer formation^{57,60}. Rev_{ARM} (amino acids 34–50) was generated using site-directed mutagenesis of the Rev₇₀ plasmid. The entire sequences of the GB1–Rev₇₀ and GB1–Rev_{ARM} constructs are listed in the Source Data file. To purify Rev₇₀, the plasmid was transformed into *E. coli* BL21(DE3) cells, and cells were grown overnight at 37 °C in 1 L of TB medium with 30 µg/mL of kanamycin. Protein expression was induced with 1 mM IPTG for 4 h at 37 °C. The cells were harvested at 4000 × *g* and resuspended in lysis buffer containing 50 mM Tris-HCl, pH 7.4, 500 mM NaCl, 100 mM (NH₄)₂SO₄, 0.1% Tween20, and cOmplete EDTA-free Protease Inhibitor Cocktail (Roche, Indianapolis, IN). The cells were lysed by sonication and centrifuged at 21,000 × *g* for 30 min, and the supernatant was collected. The protein was purified using TALON immobilized metal affinity chromatography resin (Takara, San Jose, CA) equilibrated with 40 mM Tris-HCl, pH 7.4, 500 mM NaCl, and 5 mM imidazole. Protein was eluted using a gradient of 5–200 mM imidazole, and fractions were analyzed using a 15% SDS-polyacrylamide gel. Fractions containing protein were concentrated and dialyzed into a buffer containing 40 mM Tris-HCl, pH 7.4, 250 mM NaCl, and 400 mM (NH₄)₂SO₄. Rev_{ARM} was similarly expressed and purified except that all buffers do not contain (NH₄)₂SO₄.

Crystallization, X-ray data collection, and structure determination

For crystallization, SLII-tRNA was reconstituted at 20 mg/mL in 10 mM Tris pH 7.0, 50 mM NaCl, and 10 mM MgCl₂. Initial crystals of SLII-tRNA appeared in the PACT high throughput screen condition F2, containing 200 mM sodium bromide, 100 mM Bis-Tris propane pH 6.5, and 20% polyethylene glycol (PEG) 3350. Following optimization of the condition, diffraction quality crystals were obtained using the hanging-drop vapor diffusion method at 18 °C by mixing the RNA with an equal volume of reservoir solution containing 100 mM S.P.G. (buffer mixture of succinic acid, sodium dihydrogen phosphate, and glycine in a 2:7:7 molar ratio) pH 7.0, 20–24% PEG-6000, and 200 mM sodium fluoride⁷¹. Prior to sealing the drops, acetone was added to 8% into the reservoir solution, as previously described⁴⁰. Crystals appeared within 3 days and grew to full size within 2 weeks. For data collection, crystals were passed through mother liquor with 20% (v/v) glycerol, cryoprotected in paratone-N oil, and flash frozen in liquid nitrogen. Diffraction data to 2.85 Å resolution were collected at the Advanced Light Source beamline 8.2.2 (Lawrence Berkeley National Laboratory, California). The dataset was processed using XDS with anisotropic correction⁷². The crystal belonged to space group P2₁ with unit cell dimensions of $a = 67.6$ Å, $b = 81.2$ Å, $c = 81.1$ Å, and $\beta = 99.1^\circ$, and contained two molecules in the asymmetric unit. The initial solutions were found by molecular replacement with tRNA (PDB 7LYG) and small dsRNA helix (PDB 3ND3) as search models using the Phaser program within the Phenix application suite⁷³. The final model contained two SLII-tRNA molecules with R and R_{free} of 0.23 and 0.26, respectively. Atomic coordinates for SLII-tRNA are deposited to PDB with accession number 8UO6. For structural analysis, chimera was used to determine atomic distances, angles, and surface areas⁷⁴.

Modeling of Rev dimer on RRE SLII

The crystal structure of the Rev dimer and RRE stem-loop IIB hairpin complex (PDB 4PMI) was used to model Rev binding onto the open RRE SLII conformer. First, the Rev-RRE stem-loop IIB complex was superposed with the open RRE SLII conformer by aligning the GGG

motif and stem IIB hairpin. This places the first Rev molecule within the high-affinity site at the three-way junction of the open SLII conformer. The position of the first Rev molecules was manually adjusted to better fit into the widened major groove formed at the GGG motif of SLII. Placement of the first Rev molecule orients the ARM of the second Rev molecule of the dimer toward stem IIC. Since the angle between the two ARMs of Rev dimer was shown to range between 40 and 140°^{46,60,61}, the angle between the two ARMs was adjusted to place the second Rev molecule into the major groove at the base of stem IIC.

RNA-protein binding gel shift assays

Electrophoretic mobility shift assays were used to measure the interaction between Rev₇₀ or Rev_{ARM} with native SLII and tRNA fused SLII constructs, SLII, SLII_{DS}, SLII_{SS}, SLII_{delA}, SLII_{delB}, and SLII_{delC}. tRNA^{Lys} was also included as a negative control. Both Rev_{ARM} and Rev₇₀ have the GB1 linker to allow for better resolution of distinct RNP species. RNA (500 nM) was incubated with increasing molar amounts of Rev protein (500 nM to 5 µM) in binding buffer (20 mM Tris-HCl, pH 7.4, 100 mM NaCl, and 10 mM MgCl₂) for 15 min at room temperature and loaded on an 8% native polyacrylamide gel. The gel was stained with ethidium bromide to visualize RNA. The binding assays were performed in triplicate with similar results. Uncropped gel images are provided in the Source Data file.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data supporting the findings of this study are available from the corresponding author upon request. Atomic coordinates and structure factors for SLII-tRNA have been deposited in the RCSB Protein Data Bank under the accession code 8UO6. Accession codes with hyperlinks used in this manuscript are as follows: PDB 7LYG, PDB 3ND3, PDB 1CSL, PDB 4PMI, UniProt P69718, GenBank AF033819.3 [<https://www.ncbi.nlm.nih.gov/nuccore/AF033819>]. Source data are provided with this paper.

References

- Mölling, K. et al. Association of viral reverse transcriptase with an enzyme degrading the RNA moiety of RNA-DNA hybrids. *Nat. N. Biol.* **234**, 240–243 (1971).
- Esposito, D. & Craigie, R. Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein-DNA interaction. *EMBO J.* **17**, 5832–5843 (1998).
- Frankel, A. D. & Young, J. A. HIV-1: fifteen proteins and an RNA. *Annu. Rev. Biochem.* **67**, 1–25 (1998).
- Purcell, D. F. & Martin, M. A. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J. Virol.* **67**, 6365–6378 (1993).
- Rosen, C. A. Tat and Rev: positive modulators of human immunodeficiency virus gene expression. *Gene Expr.* **1**, 85–90 (1991).
- Anand, K., Schulte, A., Vogel-Bachmayr, K., Scheffzek, K. & Geyer, M. Structural insights into the cyclin T1-Tat-RNA transcription activation complex from EIAV. *Nat. Struct. Mol. Biol.* **15**, 1287–1292 (2008).
- Rausch, J. W. & Le Grice, S. F. HIV Rev assembly on the Rev response element (RRE): a structural perspective. *Viruses* **7**, 3053–3075 (2015).
- Karn, J. & Stoltzfus, C. M. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb. Perspect. Med.* **2**, a006916 (2012).
- Laspia, M. F., Rice, A. P. & Mathews, M. B. HIV-1 Tat protein increases transcriptional initiation and stabilizes elongation. *Cell* **59**, 283–292 (1989).

10. Marciniak, R. A., Calnan, B. J., Frankel, A. D. & Sharp, P. A. HIV-1 Tat protein trans-activates transcription in vitro. *Cell* **63**, 791–802 (1990).
11. Feinberg, M. B., Baltimore, D. & Frankel, A. D. The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation. *Proc. Natl Acad. Sci. USA* **88**, 4045–4049 (1991).
12. Hammarskjöld, M. L. et al. Regulation of human immunodeficiency virus env expression by the rev gene product. *J. Virol.* **63**, 1959–1966 (1989).
13. Pollard, V. W. & Malim, M. H. The HIV-1 Rev protein. *Annu. Rev. Microbiol.* **52**, 491–532 (1998).
14. Böhnlein, 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).
15. 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).
16. Watts, J. M. et al. Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* **460**, 711–716 (2009).
17. Bai, Y., Tambe, A., Zhou, K. & Doudna, J. A. RNA-guided assembly of Rev-RRE nuclear export complexes. *Elife* **3**, e03656 (2014).
18. Fang, X. et al. An unusual topological structure of the HIV-1 Rev response element. *Cell* **155**, 594–605 (2013).
19. Legiewicz, M. et al. Resistance to RevM10 inhibition reflects a conformational switch in the HIV-1 Rev response element. *Proc. Natl Acad. Sci. USA* **105**, 14365–14370 (2008).
20. Cook, K. S. et al. Characterization of HIV-1 REV protein: binding stoichiometry and minimal RNA substrate. *Nucleic Acids Res.* **19**, 1577–1583 (1991).
21. 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. USA* **88**, 7366–7370 (1991).
22. 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).
23. Mann, D. A. et al. 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).
24. O’Carroll, I. P. et al. Contributions of individual domains to function of the HIV-1 Rev response element. *J. Virol.* **91**, e00746–17 (2017).
25. Daly, T. J., Cook, K. S., Gray, G. S., Maione, T. E. & Rusche, J. R. Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element in vitro. *Nature* **342**, 816–819 (1989).
26. Tiley, L. S., Malim, M. H., Tewary, H. K., Stockley, P. G. & Cullen, B. R. Identification of a high-affinity RNA-binding site for the human immunodeficiency virus type 1 Rev protein. *Proc. Natl Acad. Sci. USA* **89**, 758–762 (1992).
27. Daly, T. J. et al. Biochemical characterization of binding of multiple HIV-1 Rev monomeric proteins to the Rev responsive element. *Biochemistry* **32**, 10497–10505 (1993).
28. Malim, M. H., Böhnlein, S., Hauber, J. & Cullen, B. R. Functional dissection of the HIV-1 Rev trans-activator-derivation of a trans-dominant repressor of Rev function. *Cell* **58**, 205–214 (1989).
29. 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).
30. Daelemans, D., Costes, S. V., Lockett, S. & Pavlakis, G. N. Kinetic and molecular analysis of nuclear export factor CRM1 association with its cargo in vivo. *Mol. Cell Biol.* **25**, 728–739 (2005).
31. 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).
32. 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).
33. Fischer, U., Huber, J., Boelens, W. C., Mattaj, I. W. & Lührmann, R. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82**, 475–483 (1995).
34. Booth, D. S., Cheng, Y. & Frankel, A. D. The export receptor Crm1 forms a dimer to promote nuclear export of HIV RNA. *Elife* **3**, e04121 (2014).
35. Askjaer, P., Jensen, T. H., Nilsson, J., Englmeier, L. & Kjems, J. The specificity of the CRM1-Rev nuclear export signal interaction is mediated by RanGTP. *J. Biol. Chem.* **273**, 33414–33422 (1998).
36. Zimmel, R. W., Kelley, A. C., Karn, J. & Butler, P. J. Flexible regions of RNA structure facilitate co-operative Rev assembly on the Rev-response element. *J. Mol. Biol.* **258**, 763–777 (1996).
37. Tanamura, S., Terakado, H. & Harada, K. Cooperative dimerization of a stably folded protein directed by a flexible RNA in the assembly of the HIV Rev dimer-RRE stem II complex. *J. Mol. Recognit.* **29**, 199–209 (2016).
38. Schneeberger, E. M. et al. Native mass spectrometry reveals the initial binding events of HIV-1 rev to RRE stem II RNA. *Nat. Commun.* **11**, 5750 (2020).
39. Lee, E. et al. Structures of flavivirus RNA promoters suggest two binding modes with NS5 polymerase. *Nat. Commun.* **12**, 2530 (2021).
40. Gottipati, K., McNamee, S. C., Tipo, J., White, M. A. & Choi, K. H. Structural basis for cloverleaf RNA-initiated viral genome replication. *Nucleic Acids Res.* **51**, 8850–8863 (2023).
41. Bartel, D. P., Zapp, M. L., Green, M. R. & Szostak, J. W. HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **67**, 529–536 (1991).
42. Lescoute, A. & Westhof, E. Topology of three-way junctions in folded RNAs. *RNA* **12**, 83–93 (2006).
43. Holland, S. M., Chavez, M., Gerstberger, S. & Venkatesan, S. A specific sequence with a bulged guanosine residue(s) in a stem-bulge-stem structure of Rev-responsive element RNA is required for trans activation by human immunodeficiency virus type 1 Rev. *J. Virol.* **66**, 3699–3706 (1992).
44. Rypniewski, W., Adamiak, D. A., Milecki, J. & Adamiak, R. W. Non-canonical G(syn)-G(anti) base pairs stabilized by sulphate anions in two X-ray structures of the (GUGGUCUGAUGAGGCC) RNA duplex. *RNA* **14**, 1845–1851 (2008).
45. Sokoloski, J. E., Godfrey, S. A., Dombrowski, S. E. & Bevilacqua, P. C. Prevalence of syn nucleobases in the active sites of functional RNAs. *RNA* **17**, 1775–1787 (2011).
46. Jayaraman, B. et al. RNA-directed remodeling of the HIV-1 protein Rev orchestrates assembly of the Rev-Rev response element complex. *Elife* **3**, e04120 (2014).
47. Tan, R., Chen, L., Buettner, J. A., Hudson, D. & Frankel, A. D. RNA recognition by an isolated alpha helix. *Cell* **73**, 1031–1040 (1993).
48. Zhang, G., Zapp, M. L., Yan, G. & Green, M. R. Localization of HIV-1 RNA in mammalian nuclei. *J. Cell Biol.* **135**, 9–18 (1996).
49. Perkins, A., Cochrane, A. W., Ruben, S. M. & Rosen, C. A. Structural and functional characterization of the human immunodeficiency virus rev protein. *J. Acquir. Immune Defic. Syndr.* **2**, 256–263 (1989).

50. Zapp, M. L. & Green, M. R. Sequence-specific RNA binding by the HIV-1 Rev protein. *Nature* **342**, 714–716 (1989).
51. Bogerd, H. & Greene, W. C. Dominant negative mutants of human T-cell leukemia virus type I Rex and human immunodeficiency virus type 1 Rev fail to multimerize in vivo. *J. Virol.* **67**, 2496–2502 (1993).
52. Madore, S. J., Tiley, L. S., Malim, M. H. & Cullen, B. R. Sequence requirements for Rev multimerization in vivo. *Virology* **202**, 186–194 (1994).
53. Ippolito, J. A. & Steitz, T. A. The structure of the HIV-1 RRE high affinity rev binding site at 1.6 Å resolution. *J. Mol. Biol.* **295**, 711–717 (2000).
54. Hung, L. W., Holbrook, E. L. & Holbrook, S. R. The crystal structure of the Rev binding element of HIV-1 reveals novel base pairing and conformational variability. *Proc. Natl Acad. Sci. USA* **97**, 5107–5112 (2000).
55. Battiste, J. L. et al. Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science* **273**, 1547–1551 (1996).
56. 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).
57. 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).
58. 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. USA* **88**, 7734–7738 (1991).
59. 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).
60. 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).
61. Watts, N. R. et al. A new HIV-1 Rev structure optimizes interaction with target RNA (RRE) for nuclear export. *J. Struct. Biol.* **203**, 102–108 (2018).
62. Hammond, J. A., Lamichhane, R., Millar, D. P. & Williamson, J. R. A DEAD-Box helicase mediates an RNA structural transition in the HIV-1 Rev response element. *J. Mol. Biol.* **429**, 697–714 (2017).
63. Lamichhane, R. et al. A DEAD-box protein acts through RNA to promote HIV-1 Rev-RRE assembly. *Nucleic Acids Res.* **45**, 4632–4641 (2017).
64. Jackson, P. E., Tebit, D. M., Rekosh, D. & Hammarskjöld, M. L. Rev-RRE functional activity differs substantially among primary HIV-1 isolates. *AIDS Res. Hum. Retroviruses* **32**, 923–934 (2016).
65. Sloan, E. A. et al. Limited nucleotide changes in the Rev response element (RRE) during HIV-1 infection alter overall Rev-RRE activity and Rev multimerization. *J. Virol.* **87**, 11173–11186 (2013).
66. Ponchon, L. & Dardel, F. Recombinant RNA technology: the tRNA scaffold. *Nat. Methods* **4**, 571–576 (2007).
67. Kerpedjiev, P., Hammer, S. & Hofacker, I. L. Forna (force-directed RNA): simple and effective online RNA secondary structure diagrams. *Bioinformatics* **31**, 3377–3379 (2015).
68. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **15**, 8783–8798 (1987).
69. Guillerez, J., Lopez, P. J., Proux, F., Launay, H. & Dreyfus, M. A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc. Natl Acad. Sci. USA* **102**, 5958–5963 (2005).
70. 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. USA* **107**, 12481–12486 (2010).
71. Newman, J. et al. Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 1426–1431 (2005).
72. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
73. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).
74. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

Acknowledgements

We thank Dr. Mark White of Sealy Center for Structure Biology at UTMB, Juan Nysschen of IU Macromolecular Crystallography Facility, and Dr. Jay Nix at Advanced Light Source for help with X-ray data collection and helpful discussions. Beamline 8.2.2 of the Advanced Light Source, a DOE Office of Science User Facility under Contract No. DE-AC02-05CH11231, is supported in part by the ALS-ENABLE program funded by the National Institutes of Health, National Institute of General Medical Sciences, grant P30 GM124169-01. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant O85P1000817). This work was supported by NIH grants R01AI187856 and U19AI171413 (to K.H.C.), Jeane B. Kempner pre-doctoral fellowship (to J.T.) and Institute for Human Infections and Immunity (IHII) Data acquisition grant (to K.G.).

Author contributions

K.H.C., K.G. and J.T. conceived the study. J.T. and M.S. prepared various RNAs and crystallized. J.T. collected X-ray diffraction data, and J.T., G.G. and K.H.C. processed the data. J.T. determined the crystal structure. J.T. and K.G. prepared Rev proteins and J.T. performed binding assays. J.T., K.G. and K.H.C. analyzed the structure and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-48162-y>.

Correspondence and requests for materials should be addressed to Kyung H. Choi.

Peer review information *Nature Communications* thanks Gino Cingolani and the other, anonymous, reviewer for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024