

Article

Not peer-reviewed version

Mapping the vRNA Interaction with HIV-1 Integrase

[Jian Sun](#)^{*}, [Rahul Yadav](#), [Tolga Catmakas](#), [Luke Fisher](#), [Nicholas C. Fitzkee](#), [Jacques J. Kessl](#)^{*}

Posted Date: 8 May 2026

doi: 10.20944/preprints202605.0544.v1

Keywords: HIV; integrase; virus maturation; vRNA



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Mapping the vRNA Interaction with HIV-1 Integrase

Jian Sun ^{1,*}, Rahul Yadav ², Tolga Catmakas ¹, Luke Fisher ¹, Nicholas C. Fitzkee ² and Jacques J. Kessl ^{1,*}

¹ Department of Chemistry and Biochemistry, University of Southern Mississippi, Hattiesburg, MS

² Department of Chemistry, Mississippi State University, Mississippi State, MS

* Correspondence: Jian.Sun@usm.edu (J.S.); Jacques.kessl@usm.edu (J.J.K.)

Abstract

A series of critical interactions within the viral core between the viral RNA (vRNA) and HIV-1 Integrase (IN) has previously been reported. In these studies, contact points between vRNA and IN were identified using RNA-seq and MS-based protein foot-printing approaches. Several IN amino acids located in its C-terminal domain (CTD) were found to be essential for vRNA binding and their alanine substitution severely impacted the correct morphogenesis of the matured viral core. Here, we have extended these studies by performing a comprehensive mapping of the IN-vRNA interaction by deploying RNA crosslinking and NMR methodologies. Together, these approaches were able to identify additional contacts points between the vRNA and IN. Our results reveal several new basic amino acids located in the IN CTD critical for the vRNA-IN interaction, viral replication and correct morphology of the matured viral core.

Keywords: HIV; integrase; virus maturation; vRNA

1. Introduction

The Integrase protein (IN) of Human Immunodeficiency Virus type 1 (HIV-1) inserts the reverse transcribed viral DNA (vDNA) into the host chromosome. This catalytic activity of IN, essential for the early stage of the viral replication, has been developed as therapeutic target, and several FDA-approved inhibitors (Raltegravir, Elvitegravir, Dolutegravir, Bictegravir and Cabotegravir)[1] are now commonly used to treat HIV-1 infected individual. The IN is composed of three structurally distinct domains: The N-terminal domain (NTD), the catalytic core domain (CCD) and the C-terminal domain (CTD) [2,3]. During integration, these three domains work in conjunction to form large and stable multimeric structures called intasomes where vDNA ends are captured by IN [4–8]. The CCD subsequently inserts the vDNA insertion into host chromatin through two sequential magnesium-dependent reactions, 3'-processing and strand transfer. Several mutational studies have revealed that various amino acid substitutions across the IN sequence can influence not only vDNA integration but also other viral replication steps [9,10]. Consequently, IN mutations have been grouped into two classes: substitutions that selectively impair integration have been assigned to class I mutations and IN mutants that adversely affect multiple steps of virus replication have been grouped as class II [10]. Interestingly, some class II IN mutations can negatively impact correct virus particle maturation and infectivity during the late stage of replication. A hallmark of this phenotype is the mis-localization of the ribonucleoprotein complexes (RNPs) to an eccentric position trapped between the empty capsid (CA) core and the particle membrane in mature virions, whereas normal virions contain RNPs within the CA core. Characterisation of this viral phenotype have been performed either through electron microscopy [11,12] or sucrose gradient fractionation [12,13]. Intriguingly, allosteric IN inhibitors (ALLINIs [13–17], which are also known as LEDGINS [18], NCINIs [19], or INLAIs [20]), selectively bind at the IN CCD dimer interface, induce aberrant IN multimerization and consecutively yield non-infectious particles with eccentrically positioned RNPs similar to those seen with several class II IN mutants [11]. Together, these observations strongly suggest that IN not only directs vDNA

integration but also contributes to the maturation step during viral egress. Using an array of in vitro and ex vivo approaches; we and others have previously reported a strong and essential interaction between the viral RNA (vRNA) and HIV-1 IN within the virion [12,21]. Additionally, these studies [12] have pointed to a significant contribution of the IN CTD. Here, we have performed a comprehensive mapping of the IN-vRNA interaction using both vRNA crosslinking and NMR methodologies. These approaches are able to identify new functionally essential contacts points between the vRNA and IN. Taken together, our results reveal that several basic amino acids located in the IN CTD are critical for the vRNA-IN interaction and the correct morphology of the matured viral core.

2. Materials and Methods

DNA Constructs and Recombinant Proteins. Alanine substitutions were introduced using PCR-site directed mutagenesis (Agilent) and verified by DNA sequencing. For biochemical studies recombinant proteins 6xHis-tagged full length and CTD domain of WT and mutant INs were expressed in *E. coli* and purified by column chromatography[22,23]. The $^{13}\text{C}^{15}\text{N}$ -CTD protein was expressed in minimal media M9 supplemented with 1g/L $^{15}\text{NH}_4\text{Cl}$ and 2.5g/L $^{13}\text{C}_6$ -glucose.

IN-RNA Crosslinking for MS/MS. The 57-mer TAR vRNA substrate with a 4-thiouridine (4sU) modification at base 30 was synthesized and obtained from Horizon Discovery (UK). The IN-TAR complex was allowed to form by incubating 4 μM IN and TAR vRNA in a 1:1 ratio for 30 minutes at RT in 100 mM NaCl, 1 mM DTT, 0.8 mM CHAPS, 3% glycerol, 25mM HEPES pH 7.4. Crosslinking was performed in ice by subjecting the samples to 30 minutes 365 nm UV irradiation at 600 mJ/cm² (Analytik Jena UVP Crosslinker). Crosslinking efficiency was monitored by either PAG-SDS or western blot using monoclonal IN antibody (NIH AIDS Reagent Program). Crosslinked samples were then incubated with RNase I (Invitrogen) for 30 minutes at 37C, acetone precipitated and Trypsin digested overnight at 37C to yield peptides with RNA fragments of ~1-2 bases covalently linked.

Tandem mass spectrometry (MS/MS) analysis. Trypsin digested samples were analysed using electrospray ionization (ESI) MS/MS and RAW data files were processed with the Open search algorithm[24] to identify the crosslinked amino acid residue. As the traditional search engines for modifications require the adduct to be stable during the tandem mass spectrometry (MS/MS) and fragmentation profile of the cross-linked peptides make up complex data, where the adduct mass of the crosslinked species can change during analysis[25], open search algorithm within the MSFragger search engine[26] was utilized. This method provided a comprehensive analysis where every mass shift observed in the MS/MS portion of the data was analysed. Once the adduct identified, isolated MS/MS data peaks were exported via MZmine 3 software[27] and manually confirmed using the ProteinProspector MS-Product tool (<http://prospector.ucsf.edu>).

AlphaScreen-Based RNA Binding Assays. Direct binding of WT or mutant INs to RNA molecules was performed as previously described [12]. Briefly, 50 nM 6xHis-tagged IN was incubated with increasing concentrations of synthetic biotinylated RNAs (Integrated DNA Technologies) at 4C for 1 hr in the AlphaScreen buffer containing 100 mM NaCl, 1 mM MgCl₂, 1 mM DTT, 1 mg/ml BSA, 25 mM Tris pH 7.4. The final concentration of the nickel and streptavidin AlphaScreen beads (PerkinElmer) in each well was 20 mg/mL. The AlphaScreen signal was recorded on an EnSpire Plate Reader (PerkinElmer) with an AlphaScreen module using the instrument pre-sets. Data were fitted to a Hill equation using Origin software (OriginLab, Inc.).

NMR and Chemical Shift Perturbation Mapping of CTD-TAR Complex. All HSQC (heteronuclear single quantum coherence) NMR experiments were performed on a 600 MHz Bruker Avance III NMR system equipped with a CP-QCI cryoprobe in 100 mM NaCl, 8 mM CHAPS, 10 mM DTT, 3% glycerol, 25mM HEPES pH 7.4. The amide resonance in 2D ^1H - ^{15}N CTD spectra were assigned according to previously reported NMR data[23] and confirmed with 3D measurements for backbone assignment using the double labelled ^{15}N - ^{13}C CTD. Chemical shift perturbations were collected in the presence of an increasing amount of synthetic and unlabelled TAR RNA. After each addition, changes in chemical shifts of the protein resonances were monitored in TROSY ^1H - ^{15}N HSQC spectra. A total of

four CTD/RNA ratios (150 μ M CTD) were examined: 1:0.1, 1:0.2, 1:0.3 and 1:0.4. Experiments were recorded using standard TROSY-resolved Bruker pulse sequences[28] and spectral data were processed using NMRPipe[29] and analysed with NMRFAM-SPARKY[30].

Viruses and Cells. HIV-1_{NL4-3} molecular clone was used to generate replication competent HIV-1 virions. The mutant viruses were made by introducing the respective substitutions in the IN coding sequence by site directed mutagenesis (Agilent). HEK293T and HeLa TZM-bl cells were grown in Dulbecco's modified eagle medium (Invitrogen), supplemented with 10% fetal bovine serum (Invitrogen), penicillin (100 I.U./mL) and streptomycin (100 mg/mL) (GIBCO) at 37C and 5% CO₂.

Transfection and Virion Preparations. Transfections in HEK293T cells were performed with XtremeGENE HP (Roche) and the desired plasmid DNA (1:3 for DNA:reagent ratio). 24 hr post-transfection cells were washed, and fresh media was added (with 1mM of saquinavir or DMSO for the virion processing). Cell culture supernatants containing virions were harvested 48 hr post-transfection, filtered through a 0.45 μ m filter, and virions were concentrated by ultracentrifugation with a 25% sucrose cushion at 28,000 rpm for 2 hr at 4C in SW41 rotor (Beckman). Isolated virions were assayed by HIV-1 Gag p24 ELISA (ZeptoMetrix) to determine viral production. Levels of infectivity were monitored in TZM-bl reporter cells by assaying luciferase activity with a Luciferase Assay Kit (Promega) [13].

HIV-1 Viral Core Analyses Using Sucrose Density Gradient Fractionation. For sucrose density gradient fractionation [12,13] cell-free virions from transfected HEK293T cells were concentrated by ultracentrifugation over a 25% sucrose cushion, pelleted virions were lysed with 0.5% Triton X-100 and were centrifuged through a 30%–70% linear sucrose density gradient. Fractions were collected starting from the top of the gradient and subjected to immunodetection with commercial HIV-1 p24 monoclonal antibody (SinoBiological) to monitor the distribution of HIV-1 capsid.

3. Results

3.1. Binding Site Crosslinking of vRNA TAR Construct into Full-Length IN

WT recombinant full-length IN and a synthetic RNA construct containing a single photoreactive nucleobase were used to pinpoint the orientation and position of the vRNA during viral maturation. While several regions of the viral genome have been shown by CLIP-seq to interact with IN *ex vivo* [12], the vRNA₍₁₋₅₇₎-trans-activation response element (TAR) of the viral genome was selected as interacting ligand because of: i) its stable and simple structure and ii) it displays high *in vitro* affinity toward IN [12]. For those reasons, TAR has been used in previous studies as *in vitro* surrogate of the IN-vRNA binding activities[31–33]. In this experiment, we have designed and synthesized a TAR substrate that substitute a 4-thiouridine (4sU) at the position 30 (Figure 1A-star labelled base). This position on the vRNA has been reproducibly identified as primary crosslinking site within the TAR region during CLIP-seq experiments through the T-to-C base changes induced by the reverse transcription step[31]. Previously published K264A/K266A and R269A/K273A IN mutant constructs[31], which have shown to be TAR binding defective, were used here as negative controls (Figure 1B). Upon binding, 4sU-substituted TAR covalent crosslinking with full length recombinant IN was initiated by controlled UV irradiation[25]. The formation of the IN-TAR crosslinked adduct was monitored by either PAGE-SDS or western blot (Figure 1B) and showed a mass shift of +17kDa corresponding to the covalent addition of the RNA construct. This adduct was then subjected to RNase A/T1 treatment and trypsin proteolysis to yield peptides with RNA fragments of ~1-2 bases covalently linked. Analysis and sequencing of these IN peptides using tandem mass spectrometry (MS/MS) approaches (see Methods section) detected the formation of IN R262 residue crosslinked to the 4-thiouridine at position 30 with an added mass of 128.15 Da (Figure 1C), corresponding to 4-Thiouracil. Therefore, this sequencing data strongly suggests that within the IN-TAR complex, the nucleotide 30 of TAR is positioned in close proximity to the IN R262 residue as the RNA substrate interacts with several surface exposed and positively charged IN residues (Figure 6).

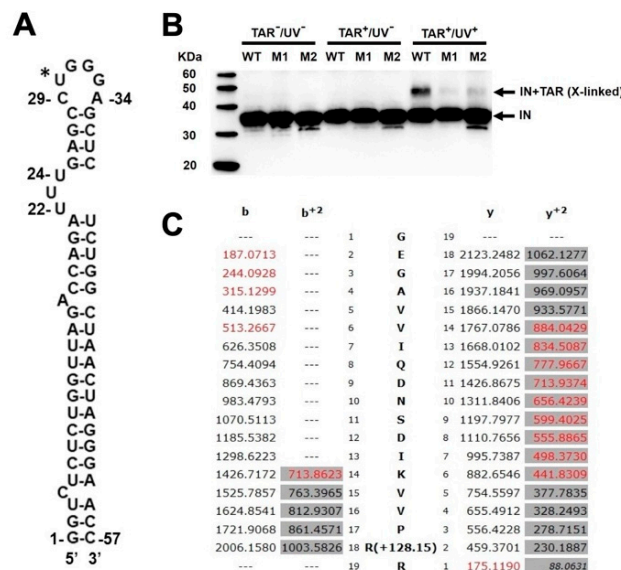


Figure 1. IN-TAR crosslinking. A: TAR sequence with star label on position 30. B: Western blot of full-length IN proteins after UV-crosslinking experiment using modified 4sU(30)-TAR vRNA construct. The R269A/K273A (M1) and K264A/K266A (M2) IN mutants were used as negative controls. C: MAs spectrometry-based sequencing of the GEGAVVIQDNSDIKVVPRR peptide showing the R262 crosslinked position.

3.2. Mapping TAR Binding to IN CTD Using Protein NMR

As the IN-TAR crosslinking study above and previously published observations[31,33] suggest, contact points between vRNA and IN point to a major role of the IN CTD. To measure in isolation the specific contribution of this particular IN domain to this interaction, we have used a modified AlphaScreen binding assay, to compare side-by-side the binding affinity of the full length IN and the CTD toward TAR (Figure 2). While this data confirms a significant binding between CTD and TAR, a lower affinity toward the RNA construct was observed with the IN domain compared to the full-length protein. This 10-folds difference strongly suggests the presence of one or several additional contact points outside the CTD. The binding properties between CTD and TAR were nevertheless compatible with the requirements for NMR detection. As NMR structures of CTD[23] have been previously reported by others, we have used published NMR amide resonance data (Figure 3A) to assign the obtained signals. In order to map the CTD residues interacting with TAR, we have monitored the chemical shift perturbations (CSP) upon the formation of the CTD-TAR complex. For this we have used a purified ^{15}N labeled CTD protein by expressing this construct in minimal media supplemented with ^{15}N -ammonium chloride as sole nitrogen source. To measure these CSPs, synthetic unlabeled TAR RNA construct was titrated into the labeled CTD protein while simultaneously acquiring 2D $^1\text{H}/^{15}\text{N}$ HSQC (Heteronuclear Single Quantum Coherence) spectrums (Figure 3B). As several backbone amide resonances shifted upon RNA additions, we were able to pinpoint several surface-exposed CTD residues (Figure 3B and Table 1) potentially contributing to the interaction with the TAR fragment. NMR resonance signals during the formation of a protein-ligand complex are known to be affected by the rate of exchange between free and bound forms relative to the frequency difference between these states[34,35]. When the rate is higher ('fast exchange'), a progressive change in peak position is observed across the titration, while when the rate is lower ('slow exchange'), separate free and bound resonances are observed with population-dependent intensities. Thus, analyses of exchange rate signals during titration (also referred to as dynamic NMR)[36] were used to reveal binding kinetics on the CTD-TAR complex formation. As high affinity complexes are usually showing "slow exchange" and low affinity complexes in "fast exchange" on the NMR chemical shift timescale[37], different CTD interacting residues were ranked accordingly using rate signal patterns observed during titration (Table 1). The residues Q221, R228

and R263 displayed slow exchange kinetics while F223, K244, R262, K264 and K266 showed fast rates (Table 1).

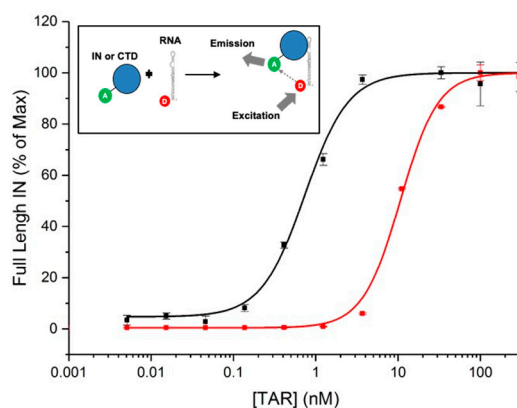


Figure 2. AlphaScreen-based binding affinity measurements of vRNA TAR fragment against full length IN (black) or IN-CTD (red). Assay schematic illustrated in top-left window.

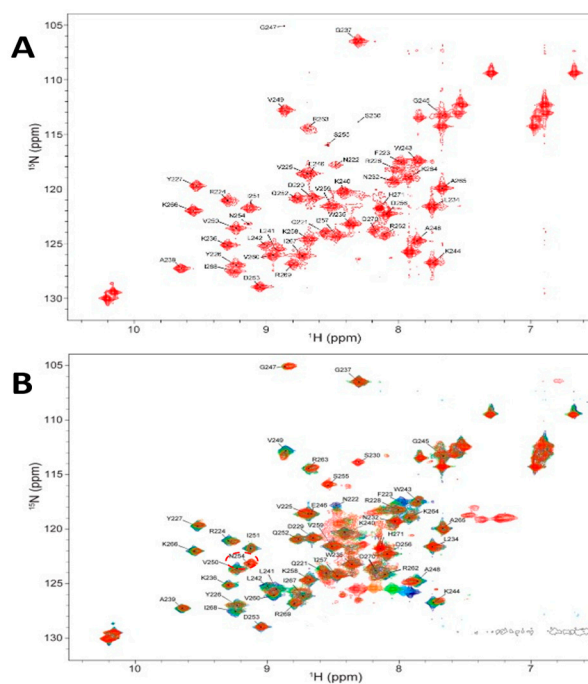


Figure 3. 2D Protein NMR of the CTD-TAR complex. A: Signal assignments confirmed by $^{15}\text{N}/^{13}\text{C}$ double labeling experiment with CTD alone. B: Spectrum of ^{15}N labeled CTD IN with chemical shift perturbations induced by TAR titration with four CTD/RNA ratios: 1:0.1 (cyan), 1:0.2 (green), 1:0.3 (orange) and 1:0.4 (red).

Table 1. NMR Exchange Rates of identified CTD residues.

CTD Residue	Exchange Rate
Q221	Slow
F223	Fast
R228	Slow
K244	Fast
R262	Fast
R263	Slow
K264	Fast
K266	Fast

3.3. In Vitro Validation of HIV-1 IN Residues Interacting with TAR

In vitro CTD-TAR and IN-TAR binding. To quantify and validate the contribution of the residues identified above (from both the 4sU crosslinking and the protein NMR experiments) to TAR binding, we have created recombinant mutated CTD and full-length IN proteins with single Alanine substitutions at these specific locations. Each mutated construct was evaluated for their binding affinity toward TAR using the AlphaScreen in vitro assay described above (Figure 2) and compared to the WT proteins. For this, WT or mutated IN proteins (Full length or CTD constructs) were incubated with biotinylated TAR before the addition of AlphaScreen beads. As the K264A/K266A double mutant IN protein has been previously shown to be defective in TAR binding[31], we have used this protein as control. To show the contribution of these 2 residues, we have also introduced the K264A/K266A double mutation into the CTD construct. These in vitro measurements revealed that all tested mutations impacted TAR binding (Figure 4A and 4B). Notably, we observed that the impact of Alanine substitution at each tested position was more severe in the CTD construct (Figure 4A) than in the context of the full length IN proteins (Figure 4B) strongly suggesting the contribution of additional contact points with the vRNA construct outside the CTD as measured earlier (Figure 2).

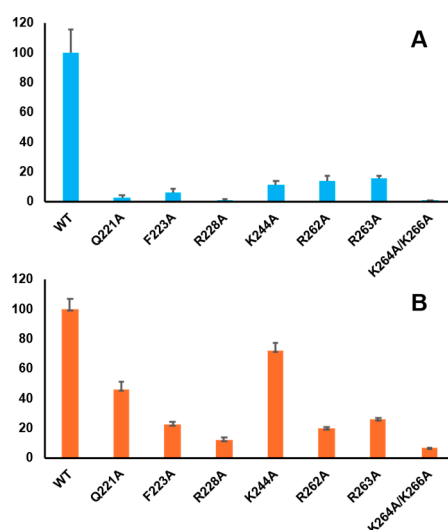


Figure 4. Effect of Alanine substitution on vRNA TAR in vitro binding to full length IN or IN-CTD using AlphaScreen-based assay. A: CTD. B: Full length protein. Values were normalized against WT proteins (100%).

3.4. IN Mutations Impaired for vRNA Binding Yield Non-Infectious Virions with Eccentric Core Morphology

To further validate the biological significance of the identified CTD residues, we performed Alanine substitutions at the positions described above in the replication-competent HIV-1_{NL4-3} construct to examine their impact on capsid core morphology and viral replication. Previous studies, including ours and others, have shown that IN mutations affecting vRNA binding influence proper virus particle maturation, leading to the production of non-infectious particles with eccentrically positioned RNPs. This mis-localization of the RNP to a position between the low-density empty capsid core and the particle membrane in mature virions has previously been observed using electron microscopy[9,21,33,38]. Additionally, the buoyant sucrose gradient technique, which measures the relative density of the viral core in mature virions, has also been validated as an effective approach to characterize this mis-localization of the RNP[13,21,31,39]. To confirm the effect of these IN mutations on virus particle maturation, we subjected cell-free and detergent-lysed progeny virions to a linear 20 to 70% (wt/vol) sucrose gradient fractionation. The CA core density was then detected by measuring the HIV-1 CA (p24) content of each sucrose fraction. We used the previously described class II V165A mutation as control as it has consistently been shown to display the eccentric core

phenotype[38]. This experiment revealed that specific CTD mutations resulted in a reduction of CA signal in high-density fractions and a simultaneously increase in low density fractions compared to the WT construct (Figure 5). This observation is consistent with the formation of empty cores due to the mis-localization of the RNPs, characteristic of IN mutants class II phenotype[9]. Among the mutations tested, the F223A exhibited a bimodal distribution of p24 along the gradient (Figure 5) indicating the presence of both eccentric and correct core densities, suggesting a weak or indirect effect on viral morphogenesis. Interestingly, the construct carrying the Q221A mutation was the only one displaying a WT profile, suggesting a marginal impact on viral maturation. To further observe the impact of these IN mutations on viral replication, the infectivity of the virions isolated from transfected cells was measured in target cells and compared to the WT construct (Figure 6). Except Q221A that retained around 60% of WT infectivity, all considered mutation had a decrease below 30%. Together, these experiments (Figures 5 and 6) confirmed that single Alanine substitutions at F223, R228, K244 and R263 positions in a replication-competent HIV-1_{NL4-3} construct significantly impact capsid core morphology and viral replication.

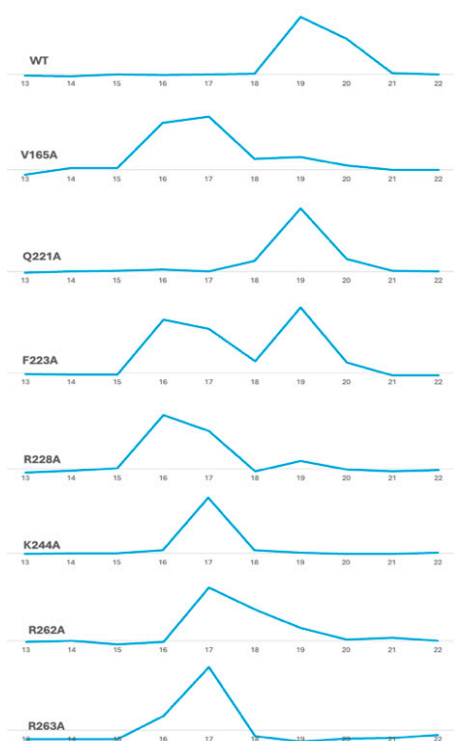


Figure 5. Viral core density profiles of IN WT and mutant virions observed by sucrose fractionation of detergent-lysed virions. The relative capsid (p24) distribution in the bottom fractions is represented.

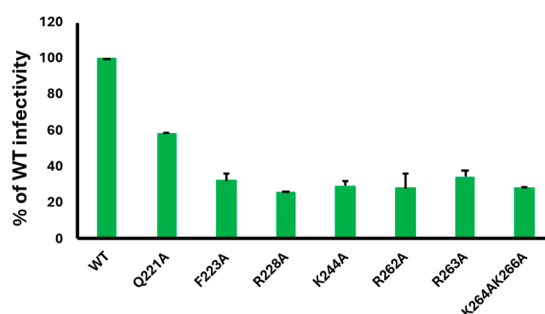


Figure 6. IN mutant virions infectivity measured in target cells. Values were normalized against WT (100%).

4. Discussion

Given its involvement observed at multiple stages of viral replication, the HIV-1 IN C-terminal domain (CTD) has been described as the virus's "Swiss army knife" [40]. With its SH3-like beta-barrel fold structure, this sub-domain appears to perform a wide variety of functions. In early studies, the domain characterized by the presence of several basic residues, has shown substantial non-specific DNA-binding properties [41–43]. Recent intasome structures have pointed to the side chains of IN residues R228, K244, R262, R263, K264, K266 and R269 as components of a basic surface patch capable of interacting with vDNA [40]. A docking study supported by biochemical analysis suggests that INI1, a subunit of the cellular SWI/SNF chromatin remodelling complex and incorporated into viral particles, also interacts with CTD [33]. Early studies have shown that various mutations within IN (class-II IN mutants) [9,10,44] could trigger the formation of a particular viral morphological defect in which the RNP condensates outside the capsid core (also described as the eccentric core phenotype). Notably, several of these IN mutants had their mutations located in the CTD [44].

These early observations strongly suggested that IN not only directs vDNA integration but also play a functional role during viral maturation. More recently, IN has been observed as interacting directly with the vRNA during viral egress [31]. The interaction occurs at specific and reproducible sites on the viral genome, mostly on structured elements such as TAR. Preventing this interaction with vRNA by mutating IN, causes the morphological defects described above.

In the present study, we have used a two-pronged in vitro-based strategy to probe and map the IN-vRNA molecular interface. We first used full length recombinant IN and a high affinity synthetic TAR 57-nucleotides construct that displays a 4sU modification at position 30 [31] to specifically crosslink and identify the IN residue located in the immediate vicinity of this nucleotide. After RNase treatment and trypsin proteolysis of the full-length protein, mass spectrometry-based sequencing reproducibly identified IN R262 residue as likely positioned adjacent to the nucleotide 30 of TAR (Figure 6, magenta position). Notably, no other crosslinked IN residues were detected. As this IN-TAR crosslinking study above and previously published observations [31,33,44] suggest, contact points between vRNA and IN point to a major contribution of the IN CTD.

Because the NMR structure of CTD [23] have been previously reported, we have used this approach to identify additional RNA interacting CTD residues by monitoring chemical shift perturbations (CSP) upon the formation of the CTD-TAR complex. Using this method, we were able to pinpoint eight surface-exposed CTD residues (Q221, F223, R228, K244, R262, R263, K264 and K266) potentially contributing to the interaction with the TAR fragment. Several of those IN residues (R228, K244 and K266) have been previously observed exhibiting class II phenotypes when mutated to Alanine [44]. Interestingly, as discussed above, some of those IN residues (R228, K244, R262, R263, K264 and K266) have also been described as components of a basic surface patch contributing to the CTD interaction with vDNA during integration [40]. The contribution to the TAR binding by each position identified in this study (Q221, F223, R228, K244, R262, R263, K264, and K266) was measured in vitro using recombinant proteins (Figure 4) and ex vivo in HIV-1_{NL4-3} constructs to examine their impact on capsid core morphology and viral replication. While Alanine replacement in the CTD construct severely impacted TAR binding at all positions (Figure 4A), the effect was mitigated for Q221A and K244A when full-length proteins were tested (Figure 4B). Additional ex vivo studies confirmed that the infectivity of virions exhibiting the Q221A mutation was marginally impacted with no observed defect on viral core morphology. On the other hand, virions with the K244A IN mutation showed poor infectivity and capsid core density defect, confirming a possible direct contribution to vRNA binding. Interestingly, while other mutated virions (R228A, K244A, R262A and R263A) were exhibiting complete shift toward low viral core densities, the F223A mutant was mixed suggesting a weaker contribution.

While our study was in late writing stage, a new approach exploring the possible morphology of the IN-RNA complex was published [45]. In this work, cryogenic electron microscopy (cryo-EM) was used to model the interactions between IN, vRNA and the capsid core. Because of the unsurmountable challenges cause by the spontaneous aggregation properties of HIV-1 IN at the

concentrations needed with this approach, the authors substituted the HIV-1 protein by the cryo-EM compatible SIVtal IN (55% identity with HIV-1 IN). Observation by cryo-EM of concentrated SIVtal IN protein vitrified in the presence of HIV RNA^{TAR} revealed the formation elongated polymers assembled from protein octamers formed around multiple extended duplexes of RNA molecules instead of distinct hairpin-like TAR structures. While the authors acknowledged that the RNA duplex arrangements may have been caused by the specific experimental conditions required for cryo-EM and unlikely to occur during maturation with the full viral genome, several CTD contact points observed with this approach (R228, K244, R262, R263, K264) confirmed our above finding.

Using the data collected in this work, we have generated an energy-minimized molecular model of the CTD-TAR complex (Figure 7) that locates the crosslinked nucleotide U30 near R262 (Figure 7, magenta residue) and predicts the relative positions of the RNA fragment according to the NMR shifts (Figure 7, yellow positions). Interestingly, the proposed model shows the TAR hairpin loop wrapping tightly around a positively charged post structure composed of R262 and R263. Alanine replacement at those positions either individually (this work) or in combination[21,44] severely impacts RNA binding, infectivity and virion morphology.

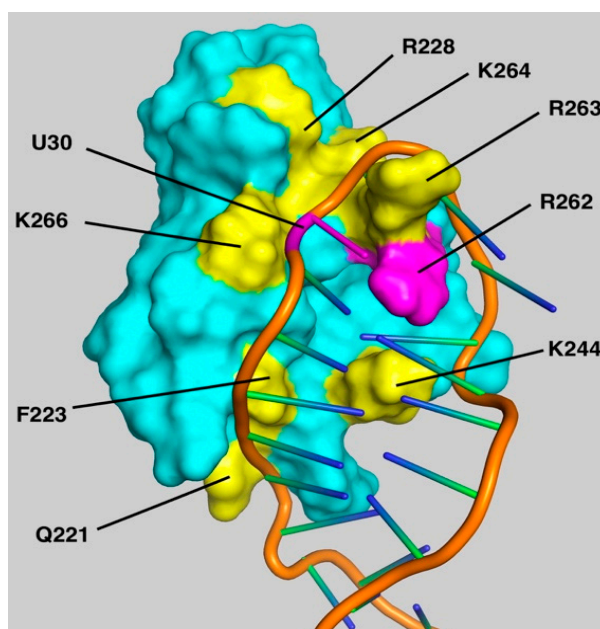


Figure 7. 3D Binding Model of the CTD-TAR Complex.

On note, both Q221 and K244 are predicted (Figure 7) to interact with the ascending and descending stem of the TAR element, which could explain why NMR and the *in vitro* measurements with the truncated CTD protein may artificially amplify the binding contribution of these two residues compared to the full-length protein. More contact points are expected to occur outside the CTD construct used in this work. Most notably is the role of K34 within the IN NTD which has been shown to contribute to vRNA binding[21]. By providing a new experimental-based model of the CTD-TAR complex (Figure 7), our study offers a molecular basis to previously observed class II mutation phenotypes such as R228A[21], double mutants R262A/R263A[21] and K264A/K266A[31].

Author Contributions: J.S., R.Y., T.C., L.F., N.C.F. and J.J.K. designed and executed experiments and interpreted results. J.S., R.Y., T.C., and J.J.K. designed the figures. J.S., N.C.F. and J.J.K. wrote the manuscript with constructive input from all authors. All authors have read and agreed to the published version of the manuscript.”.

Funding: This work was supported by the National Institutes of Health [R01AI140985 to J.J.K., N.C.F. and J.S.]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

HIV-1	Human Immunodeficiency Virus type 1
IN	Integrase
vRNA	Viral RNA
TAR	Trans-activation response element
CTD	C-terminal domain
4sU	4-Thiouridine
CA	Capsid
RNP	Ribonucleoprotein

References

1. Smith, S. J.; Zhao, X. Z.; Passos, D. O.; Lyumkis, D.; Burke, T. R., Jr.; Hughes, S. H., Integrase Strand Transfer Inhibitors Are Effective Anti-HIV Drugs. *Viruses* **2021**, *13*, (2).
2. Chiu, T. K.; Davies, D. R., Structure and function of HIV-1 integrase. *Curr Top Med Chem* **2004**, *4*, (9), 965-77.
3. Engelman, A.; Cherepanov, P., Retroviral Integrase Structure and DNA Recombination Mechanism. *Microbiol Spectr* **2014**, *2*, (6), 1-22.
4. Krishnan, L.; Engelman, A., Retroviral integrase proteins and HIV-1 DNA integration. *The Journal of biological chemistry* **2012**, *287*, (49), 40858-66.
5. Cherepanov, P.; Maertens, G. N.; Hare, S., Structural insights into the retroviral DNA integration apparatus. *Curr Opin Struct Biol* **2011**, *21*, (2), 249-56.
6. Li, X.; Krishnan, L.; Cherepanov, P.; Engelman, A., Structural biology of retroviral DNA integration. *Virology* **2011**, *411*, (2), 194-205.
7. Engelman, A.; Cherepanov, P., The structural biology of HIV-1: mechanistic and therapeutic insights. *Nat Rev Microbiol* **2012**, *10*, (4), 279-90.
8. Passos, D. O.; Li, M.; Yang, R.; Rebersburg, S. V.; Ghirlando, R.; Jeon, Y.; Shkriabai, N.; Kvaratskhelia, M.; Craigie, R.; Lyumkis, D., Cryo-EM structures and atomic model of the HIV-1 strand transfer complex intasome. *Science* **2017**, *355*, (6320), 89-92.
9. Engelman, A.; Englund, G.; Orenstein, J. M.; Martin, M. A.; Craigie, R., Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication. *J Virol* **1995**, *69*, (5), 2729-36.
10. Engelman, A., In vivo analysis of retroviral integrase structure and function. *Adv Virus Res* **1999**, *52*, 411-26.
11. Jurado, K. A.; Wang, H.; Slaughter, A.; Feng, L.; Kessl, J. J.; Koh, Y.; Wang, W.; Ballandras-Colas, A.; Patel, P. A.; Fuchs, J. R.; Kvaratskhelia, M.; Engelman, A., Allosteric integrase inhibitor potency is determined through the inhibition of HIV-1 particle maturation. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, (21), 8690-5.
12. Kessl, J. J.; Kutluay, S. B.; Townsend, D.; Rebersburg, S.; Slaughter, A.; Larue, R. C.; Shkriabai, N.; Bakouche, N.; Fuchs, J. R.; Bieniasz, P. D.; Kvaratskhelia, M., HIV-1 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. *Cell* **2016**, *166*, (5), 1257-1268 e12.
13. Sharma, A.; Slaughter, A.; Jena, N.; Feng, L.; Kessl, J. J.; Fadel, H. J.; Malani, N.; Male, F.; Wu, L.; Poeschla, E.; Bushman, F. D.; Fuchs, J. R.; Kvaratskhelia, M., A new class of multimerization selective inhibitors of HIV-1 integrase. *Plos Pathog* **2014**, *10*, (5), e1004171.

14. Kessler, J. J.; Jena, N.; Koh, Y.; Taskent-Sezgin, H.; Slaughter, A.; Feng, L.; de Silva, S.; Wu, L.; Le Grice, S. F.; Engelman, A.; Fuchs, J. R.; Kvaratskhelia, M., Multimode, cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. *The Journal of biological chemistry* **2012**, *287*, (20), 16801-11.
15. Feng, L.; Sharma, A.; Slaughter, A.; Jena, N.; Koh, Y.; Shkriabai, N.; Larue, R. C.; Patel, P. A.; Mitsuya, H.; Kessler, J. J.; Engelman, A.; Fuchs, J. R.; Kvaratskhelia, M., The A128T resistance mutation reveals aberrant protein multimerization as the primary mechanism of action of allosteric HIV-1 integrase inhibitors. *The Journal of biological chemistry* **2013**, *288*, (22), 15813-20.
16. Patel, P. A.; Kvaratskhelia, N.; Mansour, Y.; Antwi, J.; Feng, L.; Koneru, P.; Kobe, M. J.; Jena, N.; Shi, G.; Mohamed, M. S.; Li, C.; Kessler, J. J.; Fuchs, J. R., Indole-based allosteric inhibitors of HIV-1 integrase. *Bioorg Med Chem Lett* **2016**, *26*, (19), 4748-52.
17. Jentsch, N. G.; Hart, A. P.; Hume, J. D.; Sun, J.; McNeely, K. A.; Lama, C.; Pigza, J. A.; Donahue, M. G.; Kessler, J. J., Synthesis and Evaluation of Aryl Quinolines as HIV-1 Integrase Multimerization Inhibitors. *ACS Med Chem Lett* **2018**, *9*, (10), 1007-1012.
18. Christ, F.; Voet, A.; Marchand, A.; Nicolet, S.; Desimmie, B. A.; Marchand, D.; Bardiot, D.; Van der Veken, N. J.; Van Remoortel, B.; Strelkov, S. V.; De Maeyer, M.; Chaltin, P.; Debyser, Z., Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nature chemical biology* **2010**, *6*, (6), 442-8.
19. Tsiang, M.; Jones, G. S.; Niedziela-Majka, A.; Kan, E.; Lansdon, E. B.; Huang, W.; Hung, M.; Samuel, D.; Novikov, N.; Xu, Y.; Mitchell, M.; Guo, H.; Babaoglu, K.; Liu, X.; Geleziunas, R.; Sakowicz, R., New class of HIV-1 integrase (IN) inhibitors with a dual mode of action. *The Journal of biological chemistry* **2012**, *287*, (25), 21189-203.
20. Bonnard, D.; Le Rouzic, E.; Eiler, S.; Amadori, C.; Orlov, I.; Bruneau, J. M.; Brias, J.; Barbion, J.; Chevreuil, F.; Spohner, D.; Chasset, S.; Ledoussal, B.; Moreau, F.; Saib, A.; Klaholz, B. P.; Emiliani, S.; Ruff, M.; Zamborlini, A.; Benarous, R., Structure-function analyses unravel distinct effects of allosteric inhibitors of HIV-1 integrase on viral maturation and integration. *The Journal of biological chemistry* **2018**, *293*, (16), 6172-6186.
21. Elliott, J. L.; Eschbach, J. E.; Koneru, P. C.; Li, W.; Puray-Chavez, M.; Townsend, D.; Lawson, D. Q.; Engelman, A. N.; Kvaratskhelia, M.; Kutluay, S. B., Integrase-RNA interactions underscore the critical role of integrase in HIV-1 virion morphogenesis. *Elife* **2020**, *9*, e54311.
22. Kessler, J. J.; Jena, N.; Koh, Y.; Taskent-Sezgin, H.; Slaughter, A.; Feng, L.; de Silva, S.; Wu, L.; Le Grice, S. F.; Engelman, A.; Fuchs, J. R.; Kvaratskhelia, M., Multimode, cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. *J. Biol. Chem.* **2012**, *287*, (20), 16801-11.
23. Lodi, P. J.; Ernst, J. A.; Kuszewski, J.; Hickman, A. B.; Engelman, A.; Craigie, R.; Clore, G. M.; Gronenborn, A. M., Solution structure of the DNA binding domain of HIV-1 integrase. *Biochemistry* **1995**, *34*, (31), 9826-33.
24. Yu, F.; Teo, G. C.; Kong, A. T.; Haynes, S. E.; Avtonomov, D. M.; Geiszler, D. J.; Nesvizhskii, A. I., Identification of modified peptides using localization-aware open search. *Nat Commun* **2020**, *11*, (1), 4065.
25. Kramer, K.; Hummel, P.; Hsiao, H. H.; Luo, X.; Wahl, M.; Urlaub, H., Mass-spectrometric analysis of proteins cross-linked to 4-thio-uracil- and 5-bromo-uracil-substituted RNA. *Int J Mass Spectrom* **2011**, *304*, (2-3), 184-194.
26. Kong, A. T.; Leprevost, F. V.; Avtonomov, D. M.; Mellacheruvu, D.; Nesvizhskii, A. I., MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat Methods* **2017**, *14*, (5), 513-520.
27. Schmid, R.; Heuckeroth, S.; Korf, A.; Smirnov, A.; Myers, O.; Dyrland, T. S.; Bushuiev, R.; Murray, K. J.; Hoffmann, N.; Lu, M.; Sarvepalli, A.; Zhang, Z.; Fleischauer, M.; Duhrkop, K.; Wesner, M.; Hoogstra, S. J.; Rudt, E.; Mokshyna, O.; Brungs, C.; Ponomarov, K.; Mutabdzija, L.; Damiani, T.; Pudney, C. J.; Earll, M.; Helmer, P. O.; Fallon, T. R.; Schulze, T.; Rivas-Ubach, A.; Bilbao, A.; Richter, H.; Nothias, L. F.; Wang, M.; Oresic, M.; Weng, J. K.; Bocker, S.; Jeibmann, A.; Hayen, H.; Karst, U.; Dorrestein, P. C.; Petras, D.; Du, X.; Pluskal, T., Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat Biotechnol* **2023**, *41*, (4), 447-449.

28. Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wuthrich, K., TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proc Natl Acad Sci U S A* **1998**, *95*, (23), 13585-90.
29. 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* **1995**, *6*, (3), 277-93.
30. Lee, W.; Tonelli, M.; Markley, J. L., NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **2015**, *31*, (8), 1325-7.
31. Kessler, J. J.; Kutluay, S. B.; Townsend, D.; Rebensburg, S.; Slaughter, A.; Larue, R. C.; Shkriabai, N.; Bakouche, N.; Fuchs, J. R.; Bieniasz, P. D.; Kvaratskhelia, M., HIV-1 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. *Cell* **2016**, *166*, (5), 1257-1268.
32. Patacchini, E.; Madia, V. N.; Albano, A.; Ruggieri, G.; Messori, A.; Ialongo, D.; Saccoliti, F.; Scipione, L.; Cosconati, S.; Koneru, P. C.; Haney, R.; Kvaratskhelia, M.; Di Santo, R.; Costi, R., Quinolinonyl Derivatives as Dual Inhibitors of the HIV-1 Integrase Catalytic Site and Integrase-RNA interactions. *ACS Med Chem Lett* **2024**, *15*, (9), 1533-1540.
33. Dixit, U.; Bhutoria, S.; Wu, X.; Qiu, L.; Spira, M.; Mathew, S.; Harris, R.; Adams, L. J.; Cahill, S.; Pathak, R.; Rajesh Kumar, P.; Nguyen, M.; Acharya, S. A.; Brenowitz, M.; Almo, S. C.; Zou, X.; Steven, A. C.; Cowburn, D.; Girvin, M.; Kalpana, G. V., INI1/SMARCB1 Rpt1 domain mimics TAR RNA in binding to integrase to facilitate HIV-1 replication. *Nat Commun* **2021**, *12*, (1), 2743.
34. Bain, A. D., Chemical exchange in NMR. *Prog Nucl Mag Res Sp* **2003**, *43*, (3-4), 63-103.
35. Williamson, M. P., Using chemical shift perturbation to characterize ligand binding. *Prog Nucl Magn Reson Spectrosc* **2013**, *73*, 1-16.
36. Waudby, C. A.; Ramos, A.; Cabrita, L. D.; Christodoulou, J., Two-Dimensional NMR Lineshape Analysis. *Sci Rep* **2016**, *6*, 24826.
37. Latham, M. P.; Zimmermann, G. R.; Pardi, A., NMR chemical exchange as a probe for ligand-binding kinetics in a theophylline-binding RNA aptamer. *J Am Chem Soc* **2009**, *131*, (14), 5052-3.
38. Jurado, K. A.; Wang, H.; Slaughter, A.; Feng, L.; Kessler, J. J.; Koh, Y.; Wang, W.; Ballandras-Colas, A.; Patel, P. A.; Fuchs, J. R.; Kvaratskhelia, M.; Engelman, A., Allosteric integrase inhibitor potency is determined through the inhibition of HIV-1 particle maturation. *Proc. Natl. Acad. Sci. USA*. **2013**, *110*, (21), 8690-5.
39. Sun, J.; Patel, K.; Hume, J.; Pigza, J. A.; Donahue, M. G.; Kessler, J. J., Optimized binding of substituted quinoline ALLINIs within the HIV-1 integrase oligomer. *J Biol Chem* **2021**, 100363.
40. Rocchi, C.; Gouet, P.; Parissi, V.; Fiorini, F., The C-Terminal Domain of HIV-1 Integrase: A Swiss Army Knife for the Virus? *Viruses* **2022**, *14*, (7).
41. Lutzke, R. A.; Vink, C.; Plasterk, R. H., Characterization of the minimal DNA-binding domain of the HIV integrase protein. *Nucleic Acids Res* **1994**, *22*, (20), 4125-31.
42. Woerner, A. M.; Marcus-Sekura, C. J., Characterization of a DNA binding domain in the C-terminus of HIV-1 integrase by deletion mutagenesis. *Nucleic Acids Res* **1993**, *21*, (15), 3507-11.
43. Khan, E.; Mack, J. P.; Katz, R. A.; Kulkosky, J.; Skalka, A. M., Retroviral integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Res* **1991**, *19*, (4), 851-60.
44. Lu, R.; Ghory, H. Z.; Engelman, A., Genetic analyses of conserved residues in the carboxyl-terminal domain of human immunodeficiency virus type 1 integrase. *J Virol* **2005**, *79*, (16), 10356-68.
45. Singer, M. R.; Li, Z.; Rey, J. S.; Hope, J.; Chenavier, F.; Cook, N. J.; Punch, E.; Smith, J.; Zhou, Z.; Maslen, S.; Masino, L.; Nans, A.; Skehel, M.; Taylor, I. A.; Zanetti, G.; Zhang, P.; Perilla, J. R.; Engelman, A. N.; Cherepanov, P., Integrase anchors viral RNA to the HIV-1 capsid interior. *Nature* **2026**.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.