Title: Emergence of compensatory mutations reveal the importance of electrostatic interactions between HIV-1 integrase and genomic RNA

Short title: Molecular basis of HIV-1 integrase-gRNA interactions

Authors: Christian Shema Mugisha[1], Tung Dinh[2], Kasyap Tenneti[1], Jenna E. Eschbach[1], Keanu Davis[1], Robert Gifford[3], Mamuka Kvaratskhelia[2], Sebla B. Kutluay[1]

[1] Department of Molecular Microbiology, Washington University School of Medicine, Saint Louis, MO 63110, USA

[2] Division of Infectious Diseases, University of Colorado School of Medicine, Aurora, CO 80045

[3] MRC-University of Glasgow Centre for Virus Research, 464 Bearsden Rd., Bearsden, Glasgow G61 1QH, UK.

Correspondence: kutluay@wustl.edu

ABSTRACT

Independent of its catalytic activity, HIV-1 integrase (IN) enzyme regulates proper particle maturation by binding to and packaging the viral RNA genome (gRNA) inside the mature capsid lattice. Allosteric integrase inhibitors (ALLINIs) and class II IN substitutions inhibit the binding of IN to the gRNA and cause the formation of non-infectious virions characterized by mislocalization of the viral ribonucleoprotein complexes between the translucent conical capsid lattice and the viral lipid envelope. To gain insight into the molecular nature of IN-gRNA interactions, we have isolated compensatory substitutions in the background of a class II IN (R269A/K273A) variant that directly inhibits IN binding to the gRNA. We found that additional D256N and D270N substitutions in the C-terminal domain (CTD) of IN restored its ability to bind gRNA and led to the formation of infectious particles with correctly matured morphology. Furthermore, reinstating the overall positive electrostatic potential of the CTD through individual D256R or D256K substitutions was sufficient to restore IN-RNA binding and infectivity for the R269A/K273A as well as the R262A/R263A class II IN mutants. The compensatory mutations did not impact functional IN oligomerization, suggesting that they directly contributed to IN binding to the gRNA. Interestingly, HIV-1 IN R269A/K273A, but not IN R262A/R263A, bearing compensatory mutations was more sensitive to ALLINIs providing key genetic evidence that specific IN residues required for RNA binding also influence ALLINI activity. Structural modeling provided further insight into the molecular nature of IN-gRNA interactions and ALLINI mechanism of action. Taken together, our findings highlight an essential role of IN-gRNA interactions for proper virion maturation and reveal the importance of electrostatic interactions between the IN CTD and the gRNA.

46 **AUTHOR SUMMARY**

47 In addition to its well-defined catalytic function, HIV-1 integrase (IN) binds to the viral RNA genome
48 and regulates proper virion maturation. Inhibition of IN binding to the HIV-1 genome through
49 mutations of positively charged residues within the C-terminal domain (CTD, i.e. R269, K273)
50 results in non-infectious particles in which the viral genomes are mislocalized in improperly
51 matured virions. Here we have isolated compensatory mutations in the background of the class
52 II IN (R269A/K273A) mutant virus that restored the ability of IN to bind RNA. We found that
53 additional substitutions of nearby acidic residues (i.e. D256 and D270), which restored the overall
54 positive charge of the CTD, rescued the ability of IN to bind RNA and thus resulted in formation
55 of correctly matured, infectious virions. These compensatory substitutions also revealed the role
56 of specific residues within the CTD that determine sensitivity to allosteric integrase inhibitors
57 (ALLINIs), a class of compounds that indirectly target IN-RNA interactions. Taken together, our
58 findings reveal the importance of the electrostatic interactions between the IN CTD and the gRNA
59 and provide key genetic evidence for a crucial role of the CTD in antiviral activity of ALLINIs.

60

61

62

63

64

65

66

67 **INTRODUCTION**

68 A defining feature of retroviruses is the reverse transcription of the viral RNA genome (gRNA) and
69 integration of the resultant linear viral DNA into a host chromosome, which establishes lifelong
70 infection. The latter reaction is mediated by the viral integrase (IN) enzyme, which catalyzes 3'
71 processing and DNA strand transfer reactions [1]. The catalytic activity of HIV-1 IN has been
72 successfully targeted by several integrase strand-transfer inhibitors (INSTIs) [2-7] that have
73 become key components of frontline anti-retroviral therapy regimens due to their high efficacy
74 and tolerance profiles [8-11]. In addition, HIV-1 IN has an emerging non-catalytic function in virus
75 replication [12-16]. Successful targeting of this second function can complement the existing
76 antiviral regimens and substantially increase the barrier to INSTI resistance.

77 HIV-1 IN consists of three independently folded protein domains: the N-terminal domain (NTD)
78 bears the conserved His and Cys residues (HHCC motif) that coordinate Zn^{2+} binding for 3-helix
79 bundle formation; the catalytic core domain (CCD) adopts an RNase H fold and harbors the
80 enzyme active site composed of an invariant DDE motif, and C-terminal domain (CTD) which
81 adopts an SH3 fold [17, 18]. Integration is facilitated by a cellular co-factor, lens epithelium-
82 derived growth factor (LEDGF/p75), which binds tightly to a site within the CCD dimer interface
83 [19, 20] and guides the preintegration complex to actively transcribed regions of the host
84 chromosome [19-24]. A group of pleotropic IN substitutions distributed throughout IN, collectively
85 known as class II mutations, disrupt viral assembly [13, 16, 25-36], morphogenesis [12, 16, 27,
86 32-34, 37, 38] and reverse transcription in target cells [12-14, 16, 29, 31, 32, 34, 36-54] often
87 without obstructing the catalytic activity of IN in vitro [13, 27, 28, 31, 40, 41, 44, 46, 55-57]. A
88 hallmark of class II IN mutant viruses is the mislocalization of the viral ribonucleoprotein
89 complexes (vRNP) outside of the viral capsid (CA) lattice, a deformation which is often referred
90 to as eccentric morphology [12, 15, 16, 27, 32-34, 37, 38, 58-60]. Although originally designed to
91 inhibit integration through preventing the binding of IN to LEDGF/p75 [58, 61-65], allosteric

92 integrase inhibitors (ALLINIs) potently inhibit proper virion maturation [12, 37, 61] and lead to the
93 formation of virions that display a similar eccentric morphology observed with class II IN mutations
94 [12].

95 We have recently shown that binding of IN to the gRNA in mature virions accounts for the non-
96 catalytic function of IN in virus replication and is required for proper encapsidation of the viral
97 ribonucleoproteins (vRNPs) inside the mature CA lattice [38]. Class II IN substitutions and
98 ALLINIs block IN-gRNA binding, thus causing formation of virions with an eccentric morphology
99 [15, 38]. IN binds to multiple distinct locations on the gRNA, including the TAR hairpin present
100 within the 5' and 3'UTRs of the gRNA and constitutes a high affinity binding site [38]. IN
101 preferentially binds to the gRNA in a tetrameric state, and many class II IN mutations block IN-
102 gRNA binding by disrupting the functional oligomerization of IN [15]. Mutation of basic residues
103 within IN-CTD (i.e. R262, R263, R269 and K273) inhibits IN-gRNA binding without altering
104 functional IN oligomerization in virions and in vitro [15, 38], suggesting that these residues are
105 directly involved in binding to the gRNA. On the other hand, the precise mechanism of how these
106 residues mediate recognition of specific sequence elements on the gRNA remains unknown. For
107 example, it is possible that the positively charged Lys and Arg residues interact with the negatively
108 charged RNA phosphate backbone in a non-specific or semi-specific manner, depending on the
109 folding and structure of the cognate RNA element, driven by electrostatic interactions [66-68].
110 Alternatively, these residues can mediate specific interactions with RNA targets through H-
111 bonding and van der Waals contacts with individual nucleobases [66-68].

112 To gain insight into the mode of IN-gRNA interactions, the non-infectious IN R269A/K273A class
113 II mutant virus was serially passaged in T-cells until the acquisition of compensatory mutations.
114 We found that two compensatory mutations, D256N and D270N, within the IN-CTD restored virion
115 infectivity, IN-RNA interactions and accurate virion morphogenesis. As the D-to-N mutations
116 resulted in loss of two negative charges, possibly overcoming the loss of two positive charges

117 with the R269A/K273A class II substitutions, we tested whether restoring the overall charge of
118 CTD through other mutations would restore IN-gRNA binding and virion infectivity. Indeed, the
119 D256R substitution alone restored virion infectivity and RNA binding for the IN R269A/K273A
120 mutant. We further extended these findings to another class II mutant, R262A/R263A, which was
121 similarly suppressed by the D256R as well as D256K substitutions. Compensatory mutations did
122 not affect the ability of IN to multimerize *in vitro* or in virions, suggesting that they restored the
123 RNA-binding ability of IN directly. Interestingly, the IN R269A/K273A, but not the IN
124 R262A/R263A, mutant viruses bearing the compensatory mutations had increased sensitivity to
125 ALLINIs, providing key genetic evidence that specific residues within the CTD, which are required
126 for RNA binding, also contribute to ALLINI mechanism of action. Together, our findings strongly
127 suggest that IN-RNA interactions are at least in part driven by electrostatic interactions between
128 the basic residues within IN-CTD and phosphate backbone of the gRNA and highlight that
129 ALLINIs engage CTD residues to inhibit RNA binding.

130 **RESULTS**

131 **Compensatory IN D256N/D270N substitutions emerge in the background of the IN**
132 **R269A/K273A class II mutant virus**

133 The R269A/K273A class II IN substitutions obstruct IN-gRNA binding directly without interfering
134 with IN multimerization and result in the formation of particles with eccentric morphology [15, 16].
135 To better understand the molecular basis of how R269 and K273 residues mediate IN-gRNA
136 binding, viruses bearing the IN R269A/K273A mutations were serially passaged in MT-4 cells until
137 the emergence of compensatory mutations at the end of passage 3, which completely restored
138 virion infectivity (**Fig. 1A**). Deep sequencing of full-length gRNA isolated from virions across the
139 three passages revealed that viruses retained the IN R269A/K273A substitutions while
140 sequentially acquiring D256N and D270N mutations in IN (**Fig. 1B**). The D256N substitution

141 emerged at the end of passage 1 whereas the D270N mutation emerged later in passage 2 (**Fig.**
142 **1B**). Both mutations were fixed by the end of passage 2 and no other mutations were observed
143 elsewhere on the viral genome.

144 The IN D256N and D270N mutations were introduced into the replication-competent pNL4-3
145 molecular clone bearing WT or R269A/K273A IN. Introduction of D256N and D270N substitutions
146 in IN had no observable effect on Gag (Pr55) expression or processing in cells or particle release
147 (**Fig. 1C**). Introduction of D256N and D270N mutations either individually or together (D2N) on
148 the WT IN backbone did not affect viral titers (**Fig. 1D**). Remarkably, while the individual D256N
149 and D270N mutations in the R269A/K273A IN backbone increased virus titers by 5-10-fold, the
150 D2N substitutions increased virus titers by 100-fold (**Fig. 1D**). Overall, these results demonstrate
151 that the combination of D256N and D270N mutations are sufficient to restore the replication
152 competency of the R269A/K273A class II IN mutant virus.

153 **D256N/D270N substitutions restore IN-gRNA binding for the R269A/K273A class II IN**
154 **mutant viruses and lead to formation of correctly matured virions**

155 We next assessed whether D256N and D270N substitutions rendered the R269A/K273A class II
156 IN mutant virus replication competent by restoring IN-gRNA binding and proper virion maturation.
157 To this end, IN-gRNA complexes were immunoprecipitated from UV-crosslinked virions and
158 visualized per CLIP protocol as described previously [38, 69, 70]. Equivalent amounts of
159 immunoprecipitated IN from WT and R269A/K273A IN viruses, or those additionally bearing the
160 IN D256N and D270N substitutions, were analyzed for their ability to bind RNA. IN-RNA
161 complexes were readily visible for WT viruses as well as D256N, D270N and D2N mutants
162 (introduced on the WT backbone), but not from the R269A/K273A class-II IN mutant viruses (**Fig.**
163 **2A**). Introduction of the D256N substitution on R269A/K273A IN backbone modestly enhanced
164 the ability of IN to bind RNA, whereas the D270N substitution had no observable impact (**Fig. 2A**).

165 In contrast, the D2N substitution substantially enhanced the ability of R269A/K273A IN to bind
166 RNA (**Fig. 2A**). The virion morphology of WT and IN mutant viruses was assessed by transmission
167 electron microscopy (TEM). As expected, more than 80% of WT particles had an electron-dense
168 condensate that represents vRNPs inside the CA lattice, whereas the majority of R269A/K273A
169 class II IN mutant virions (~68%) had a clear eccentric morphology (**Fig. 2B**). Consistent with
170 effects on virus titers and RNA-binding, the introduction of D256N/D270N substitutions restored
171 the ability of the R269A/K273A IN mutant to form properly mature virions (**Fig. 2B**). Cumulatively,
172 these data show that D256N and D270N IN substitutions restore infectivity for the R269A/K273A
173 class II IN mutant virus through reestablishing RNA binding and subsequently accurate virion
174 maturation.

175 **Electrostatic interactions are required for IN-gRNA binding.**

176 IN-CTD is decorated with several acidic and basic amino acids resulting in a net charge of +3
177 (**Fig. 3A**). The D2N substitutions in effect restored the overall charge of the IN R269A/K273A
178 CTD, suggesting that the net electrostatic charge of the IN-CTD may be a key parameter in gRNA
179 binding. To test this hypothesis, we investigated whether restoring the overall charge of IN-CTD
180 through other mutations would also restore RNA binding and infectivity for the R269A/K273A IN
181 virus. We focused our analysis on D256 and D270 residues for the following reasons: i) these
182 amino acids were amenable to substitutions during virus passaging experiments; ii) D256N and
183 D270N substitutions in the context of WT HIV-1 yielded infectious virions, suggesting that these
184 compensatory mutations do not significantly contribute to the catalytic activity of IN thus allowing
185 us to specifically probe their roles for IN-RNA interactions and virion maturation. To extend these
186 studies we introduced D256R, D270R and D256R/D270R (D2R) substitutions into the HIV-1_{NL4-3}
187 $_{IN\ R269A/K273A}$ backbone and transfected HEK293T cells with the resulting plasmids. Cell lysates and
188 cell-free virions were then analyzed for Gag processing, particle release, and infectivity. Overall,

189 D-to-R substitutions had no major effect on Gag (Pr55) expression or processing in cells (**Fig.**
190 **3B**).

191 Remarkably, the D256R substitution alone increased the titers of the R269A/K273A class II IN
192 mutant virus by 100-fold and to a level comparable to that of WT viruses (**Fig. 3C**). In contrast,
193 D270R substitution only had a modest impact on virus titers and the D2R substitution increased
194 virion infectivity at a modestly lower level than that of D256R and D2N (**Fig. 3C**). Importantly, the
195 D256R and the D2R substitutions completely restored IN-gRNA binding (**Fig. 3D**), demonstrating
196 that the increase in viral titers with the D256R and D2R substitutions correlates well with
197 enhancement of IN-gRNA binding. Further assessment of these substitutions with Vpr-IN
198 transcomplementation assays revealed that none of the IN substitutions impacted the catalytic
199 activity of IN, except for the D2R substitution (**Fig. 3E**), suggesting that the inability of the D270R
200 mutation to restore infectivity may be due to its adverse effects on the catalytic activity of IN.
201 Taken together, these findings indicate that restoring the overall charge of IN-CTD through D256R
202 mutation reestablishes virion infectivity and RNA binding for the IN R269A/K273A class II mutant.

203 **D256R substitution fully restores virion infectivity and RNA binding for a separate class II**
204 **IN (R262A/R263A) mutant**

205 We have previously shown that mutation of other basic residues within the IN-CTD (i.e.
206 R262A/R263A) also directly inhibit IN-gRNA binding without compromising functional
207 oligomerization of IN [15, 38]. We next wanted to extend our observations and test how
208 substitutions of D256 and D270 residues affect IN-gRNA binding and infectivity in the background
209 of the IN R262A/R263A mutant virus. Introducing the D256N, D270N and D2N compensatory
210 mutations on the IN R262A/R263A backbone did not affect Gag expression, processing, and
211 virion release (**Fig. 4A**). While the D256N substitution increased virion infectivity by 10-fold, the
212 D270N substitution had no impact and the D2N substitution had an intermediate phenotype (**Fig.**

213 **4B**), demonstrating the context dependency of these compensatory mutations. The D256K and
214 D256R substitutions introduced in the R262A/R263A IN backbone also had minimal effects on
215 Gag expression, processing, and virion release (**Fig. 4C**). In contrast, the D256K substitution
216 increased viral titers at a greater degree than D256N and the D256R substitution completely
217 restored virion infectivity (**Fig. 4D**). In line with the titer data, D256K significantly increased RNA
218 binding whereas the D256R completely restored it in the context of the IN R262A/R263A mutant
219 virus (**Fig 4E**). D256N, D256K and D256R IN successfully transcomplemented a class I IN mutant,
220 suggesting that they did not distort the catalytic activity of IN (**Fig 4F**). Taken together, these
221 findings indicate that restoring the overall positive charge of IN-CTD through D256R or D256K
222 mutations is sufficient to restore virion infectivity and RNA binding to class II IN mutants that
223 directly inhibit RNA binding.

224 We next used reported X-ray structure [62] and molecular modelling to visualize how class II and
225 compensatory IN mutations affect electrostatic potential of the CTD surface. Mutation of R269
226 and K273 residues expectedly resulted in a substantial loss of a basic patch in IN (**Fig. 5A, B**).
227 Both the D256R and D256N substitutions resulted in more positively charged surface distal from
228 the R269/K273A residues (**Fig. 5C, D**), suggesting that compensatory mutations likely created an
229 additional interacting interface with gRNA binding. A similar outcome was observed when class II
230 R262A/R263A and compensatory D256R changes were introduced in the CTD (**Fig. 5E-G**). While
231 the D256R change is substantially distanced from R262/R263 and R269/K273 residues, we note
232 the following. R262, R263, R269, and K273 are positioned within the same highly flexible C-
233 terminal tail (aa 261-275), whereas D256 belongs to another, shorter (aa 252-257) loop (**Fig. 5H**).
234 The highly pliable nature of the tail and the loop could be crucial for IN to optimally engage with
235 cognate RNA as well as allow for emergence of compensatory mutations at alternative sites
236 positioned in these IN segments.

237 **Effects of the compensatory mutations on functional oligomerization of IN and IN-RNA**
238 **interactions.**

239 As functional IN oligomerization is a prerequisite for RNA binding [15], we next examined how the
240 compensatory substitutions affected IN oligomerization. For in virion analysis, purified HIV-1_{NL4-3}
241 virions were treated with ethylene glycol bis (succinimidyl succinate) (EGS) to covalently crosslink
242 IN in situ and virus lysates were analyzed by immunoblotting. As in WT virions, IN species that
243 migrated at molecular weights consistent with those of monomers, dimers, trimers, and tetramers
244 were readily distinguished in R269A/K273A or R262A/R263A viruses with additional
245 compensatory mutations (D256N/D270N, D256R, D256K, and D256R/D270R) but not with the
246 canonical class II IN mutant V165A that is unable to form functional oligomers (**Fig 6A-C**).
247 Complementary in vitro assessment of purified recombinant IN proteins by size exclusion
248 chromatography (SEC) revealed that D2N and D256R substitutions in the background of
249 R269A/K273A or the R262A/R263A class II mutants did not impact functional IN tetramerization
250 (**Fig. 6D-H**).

251 We have previously shown that recombinant IN binds to TAR RNA with high affinity and provides
252 a nucleation point to bridge and condense RNA [38]. We next examined the ability of class II
253 mutant INs bearing compensatory mutations to bind and bridge TAR RNA. Consistent with
254 findings from CLIP, the D2N and D256R substitutions also enhanced or restored the ability of
255 R269A/K273 and R262A/R263A mutants to bridge between RNA molecules (**Fig. 6I**). Together,
256 these data demonstrate that the compensatory mutations directly restore the ability of IN to bind
257 RNA without altering functional IN oligomerization.

258 **Sensitivity to ALLINIs is determined by distinct residues within the CTD**

259 ALLINIs potently disrupt proper virion maturation through inducing aberrant IN multimerization
260 and consequently inhibiting IN-gRNA binding [12, 38]. Recent structural studies have shown that

261 the ALLINI, GSK-1264, can directly engage residues within the IN-CTD through its tert-butoxy
262 and carboxylic acid moieties and induce open polymers [62]. These findings are in line with
263 previous biochemical and modeling studies that also showed the involvement of the IN-CTD, in
264 particular residues K264 and K266, in ALLINI-induced aberrant IN multimerization [71, 72]. Based
265 on these prior findings, we wanted to test how adjacently positioned IN R262A/R263A and
266 R269A/K273A substitutions and the compensatory mutations affected ALLINI activities.

267 To this end, we examined effects of representative quinoline-based ALLINIs, BI-D or BI-B2, on
268 the viruses bearing the class II and compensatory mutations. While the titers of WT viruses only
269 decreased at ALLINI concentrations greater than 1 μM, the titers of IN R269A/K273A viruses
270 bearing D2N, D256R and D2R substitutions were significantly reduced by ALLINIs at
271 concentrations as low as 0.1 μM (**Fig. 7A, B**). In contrast, viruses bearing the
272 R262A/R263A/D256R IN were less sensitive to low concentrations of ALLINIs (**Fig. 7D, E**),
273 suggesting that the R269A and/or K273A mutations underlie the increased sensitivity to ALLINIs.
274 The increased sensitivity to ALLINIs correlated well with inhibition of RNA binding. While 0.1 μM
275 of the ALLINI, BI-D, did affect the level of IN-gRNA binding in WT viruses, it significantly reduced
276 IN-gRNA binding in viruses with class II compensatory mutations on the R269A/K273A but not
277 R262A/R263A backbone (**Fig. 7C,F**). Taken together, our findings provide key genetic and
278 virological evidence that specific CTD residues required for RNA binding are also crucial for the
279 ALLINI mechanism of action.

280 **Characterization of IN mutations present in latently infected cells**

281 Persistence of HIV-1 in memory CD4+ T-cells as latent proviruses constitutes a major barrier to
282 HIV-1 cure. Although the majority of HIV-1 proviruses in these cells are defective [73], recent
283 evidence suggests that defective proviruses can be transcribed into RNAs that are spliced,
284 translated and can be recognized by HIV-1-specific cytotoxic T lymphocytes [74]. We decided to

285 characterize IN mutations isolated from latently infected cells, given the possibility that class II IN
286 mutations existing in latently infected cells can result in the formation of defective particles that
287 may subsequently modulate immune responses. Though relatively uncommon, we found the
288 presence of R224Q, S230N, E246K and G272R substitutions in IN-CTD (**Fig. 8A**). Of note, only
289 the R224Q substitution resulted in loss of a positive charge, whereas the E246K and G273R
290 substitutions resulted in gain of positive charge. These mutations were introduced into the NL4-3
291 proviral backbone with minimal effects on Gag expression and particle release (**Fig. 8B**). Although
292 the E246K virus was significantly less infectious (**Fig. 8C**), we did not find any evidence for
293 inhibition of IN-gRNA binding (**Fig. 8D**), suggesting that this mutant likely displays a class I
294 phenotype. Thus, we conclude that the class II mutant viruses are rarely present in the latently
295 infected cells and therefore unlikely to contribute to chronic immune activation.

296 **DISCUSSION**

297 Class II IN mutations impair virion particle maturation by blocking IN-gRNA binding in virions and
298 those within the CTD, including R269A/K273A and R262A/R263A, impede IN-gRNA binding
299 without affecting functional oligomerization of IN [15, 38]. During serial passaging experiments to
300 identify compensatory mutations of the R269A/K273A class II IN mutant, viruses acquired D256N
301 and D270N IN mutations sequentially. We initially anticipated that mutations outside of IN, such
302 as CA and NC, could also arise, given that the IN R269A/K273A mutant is still catalytically active
303 and a compensatory mutation in CA or NC could presumably allow the proper packaging of the
304 gRNA within virions. On the other hand, we did not observe any such substitutions showcasing
305 the distinct role of IN:gRNA binding in proper virion maturation. Notably, D256N and D270N
306 substitutions each arose through a single mutation (D256N: GAC→AAC, D270N:GAU→AAU)
307 and thus likely provided an easier pathway for suppression than reverting back to R269 and K273,
308 each of which would require two mutations. Though other mutations in IN could in principle restore

309 the ability of IN to bind RNA, rise of such mutations was likely constrained in part by the necessity
310 to maintain a catalytically active IN.

311 IN CTD has a net positive charge of +3. While R269A/K273A class II mutations reduced the net
312 charge to +1, the introduction of D2N substitutions restored the overall charge back to a +3. This
313 observation suggested that the net electrostatic charge of the IN-CTD may be a key parameter
314 for IN-gRNA binding. Consistent with this hypothesis, the charge reversal by the D256R
315 substitution was sufficient to restore IN-gRNA binding for the R269A/K273A IN mutant. Though
316 these findings point to an electrostatic component of IN-RNA interactions, we cannot exclude the
317 possibility that the Asn and Arg residues also mediate H-bonding and van der Waals contacts
318 with distinct nucleobases in the cognate RNA molecules. Inspection of the available X-ray
319 structure [62] indicates that all IN residues implicated in RNA binding by the present study are
320 positioned either in the highly flexible C-terminal tail (aa 261-275) or the 252-257 loop (**Fig. 5H**).
321 The pliable nature of these CTD regions could be essential for allowing IN to optimally bind to
322 cognate gRNA.

323 These observations extended to another class II mutant, R262A/R263A, whereby the D256R and
324 to a lesser extent the D256K substitutions restored RNA binding. In contrast, the D2N
325 substitutions did not enhance RNA-binding or infectivity in this setting, demonstrating a degree of
326 context dependency. It is possible that the proximity of the D270N to R269A/K273A residues may
327 explain why it was more effective in restoring RNA binding for this class II mutant but not for IN
328 R262A/R263A. The role of Arg and Lys residues in RNA-binding proteins has been noted [67, 75-
329 78]. Interestingly, Arg residues are overall more heavily involved in interactions with all bases,
330 contributing with 16%–20% of all contacts while Lys residues only provide 3%–9%, which may
331 explain why the D256R substitution restored RNA binding to a greater extent [79]. Altogether, our
332 findings suggest that IN-gRNA binding is mediated in part by electrostatic interactions between
333 the basic residues in IN CTD and the negatively charged RNA backbone.

334 Though the electrostatic component of IN-RNA interactions imply a level of non-specificity, IN
335 bound to distinct locations on the gRNA and displays high binding affinity to structured elements,
336 such as the HIV-1 TAR element [38]. Thus, it is likely that IN-gRNA interactions are mediated by
337 both the non-specific interactions of the basic residues with the RNA phosphate backbone and
338 specific interactions with the cognate RNA. For example, recognition of the TAR loop by Tat and
339 the super elongation complex is based on a complex set of interactions that results primarily in
340 the readout of the structure as opposed to the sequence by SEC, and additional interaction of the
341 Tat arginine rich motif (ARM) with the TAR bulge and the major groove through electrostatic
342 interactions with the RNA phosphates, H-bonding with specific bases, as well as multiple van der
343 Waals contacts of Cyclin T with the flipped-out G32 TAR base to strengthen the TAR complex
344 [80, 81]. Structural studies of IN in complex with cognate RNA molecules are needed to tease
345 apart the specificity determinants for IN-RNA interactions.

346 RNA binding proteins commonly encode modular RNA-binding domains (i.e. RRMs and KH
347 domains), which form specific contacts with short degenerate sequences [82-84]. Utilization of
348 multiple RRMs/KH domains is thought to create a much larger binding interface, which allows
349 recognition of longer sequences on and enhanced affinity and specificity for target RNAs [82, 83,
350 85]. IN tetramerization is required for RNA binding in vitro [15] and may serve a similar purpose
351 through generation of a larger RNA-binding surface possibly for recognition of shape, RNA-
352 backbone and base-specific interactions as discussed above.

353 IN-gRNA binding is essential for HIV-1 virion morphological maturation and infectivity, thus an
354 excellent target for novel antiretroviral compounds. Clinically used HIV-1 IN inhibitors target the
355 catalytic activity of IN by blocking the strand transfer step of integration [2-7]. Despite high barriers
356 to resistance with the second-generation INSTIs, treatment continues to select for drug resistant
357 HIV-1 variants [86-91]. ALLINI-mediated inhibition of the non-catalytic function of IN can
358 complement existing INSTI-based therapies and increase the barrier to drug resistance

359 substantially. Although some ALLINIs displayed toxicity in animals, a highly potent and safe
360 pyrrolopyridine-based ALLINI, STP0404, has advanced to a human trial [65].

361 ALLINIs potently disrupt virion maturation indirectly through inducing aberrant IN multimerization
362 resulting in inhibition of IN-gRNA binding [12, 38]. Recent findings have shown that the ALLINI,
363 GSK-1264, is buried between the CTD of one IN dimer and the CCD of another dimer resulting in
364 the formation of open, inactive IN polymers [62]. Mutation of Y226, W235, K264 and K266
365 residues within the CTD prevents ALLINI-induced aberrant multimerization of IN [62, 71, 72].
366 However, the role of these residues in ALLINI-mediated inhibition cannot be easily assessed in
367 relevant infection settings given that their mutation alters the catalytic activity as well as RNA
368 binding properties of IN [62, 92, 93]. Our findings fill this gap and provide key genetic evidence
369 that ALLINI mechanism of action indeed involves the CTD. In particular, introducing the same
370 D256R change in the background of different class II substitutions similarly restored their
371 infectivity and thus enabled us to compare effects of R262/R263 vs R269A/K273A IN substitutions
372 on ALLINI activities. While HIV-$1_{IN(R262A/R263A/D256R)}$ and WT viruses were similarly susceptible to
373 ALLINIs (**Fig. 7A,B**), HIV-$1_{IN(R269A/K273A/D256R)}$ was substantially more sensitive to these inhibitors
374 (**Fig. 7D,E**). The X-ray structure of GSK-1264 induced IN polymers [62] reveals that R262, R263,
375 K273A as well as D256 are distanced from the inhibitor binding pocket, whereas the side chain
376 of R269 points toward the inhibitor bound at the CTD-CCD dimer interface (**Fig. 7G**).
377 The guanidine group of R269 engages with Q168 of the CCD from another IN subunit and is
378 positioned within 3 Å of GSK-1264. Yet, there is no interaction between R269 and the inhibitor as
379 the positively charged guanidine group faces a hydrophobic part of the benzodihydropyran moiety
380 of GSK-1264 (**Fig. 7G**). Conversely, the R269A substitution could provide more favorable
381 hydrophobic environment for the CTD-ALLINI-CCD interactions that may explain its increased
382 sensitivity to ALLINIs.

383 Overall, our studies reveal that electrostatic interactions play an important role in mediating IN-
384 gRNA interactions and demonstrate that CTD is a key determinant of ALLINI sensitivity. Structural
385 characterization of IN-gRNA complexes will be crucial to determine the precise mechanism of IN-
386 gRNA interactions. Such studies will also help better understand the ALLINI mechanism of action
387 and aid in the development of therapeutics that directly target IN-gRNA interactions.

388 **MATERIALS AND METHODS**

389 **Plasmids**

390 IN mutations were introduced into the HIV-1_{NL4-3} full-length proviral plasmid (pNL4-3) by overlap
391 extension PCR. Forward and reverse primers containing IN mutations were used in PCR
392 reactions with anti-sense (with EcoRI restriction endonuclease site) and sense (with AgeI
393 restriction site) outer primers. The resulting fragments containing the desired mutations were
394 mixed at a 1:1 ratio and overlapped subsequently using the sense and antisense primer pairs.
395 The overlap fragments were digested with Age-I and EcoR-I before cloning into pNL4-3 plasmids.
396 The pLR2P-vprIN plasmid expressing a Vpr-IN fusion protein has been previously described [94].
397 IN mutations were introduced in the pLR2P-VprIN plasmid using the QuikChange Site-Directed
398 Mutagenesis kit (Agilent Technologies). The presence of the desired mutations and the absence
399 of unwanted secondary changes was assessed by Sanger sequencing.

400 **Cell lines**

401 All cell lines were obtained from the American Type Culture Collection and NIH AIDS Reagents
402 where STR profiling was performed. MT-4 cells were further STR profiled at Washington
403 University School of Medicine Genome Engineering and iPSC center. The cell lines are
404 consistently inspected for mycoplasma contamination using MycoAlert mycoplasma detection kit
405 (Lonza) and checked for being free of any other contaminations. HEK293T cells (ATCC CRL-

406 11268) and Hela-derived TZM-bl cells (NIH AIDS Reagent Program) were cultured in Dulbecco's
407 modified Eagle's medium supplemented with 10% fetal bovine serum. MT-4 cells were cultured
408 in in RPMI 1640 medium supplemented with 10% fetal bovine serum.

409 **Analysis of compensatory mutations**

410 Compensatory mutations of the R269A/K273A class-II IN mutation were isolated by serial
411 passaging. In this experiment, MT4 T-cells that express GFP under the control of HIV-1 LTR
412 (MT4-LTR-GFP), were infected with HIV-1pNL4-3 carrying the R269A/K273A class II IN mutation.
413 One million cells were infected at an MOI of 2 by WT cells, input of mutant virus was normalized
414 relative to particle numbers using a reverse transcriptase (RT) activity assay [95]. Infections were
415 monitored by FACS. At the end of each passage, virions were collected and wielded to infect cells
416 in the next passage while normalizing virus input before each infection. Aliquots of infected and
417 viral particles in the cell culture supernatants were collected over the duration of each passage.
418 Virions collected over the three passages were concentrated and gRNA was isolated from virions
419 using Trizol per manufacturer's instructions. Extracted RNA was prepared for deep sequencing
420 usingthe Illumina® TruSeq® Stranded Total RNA library prep workflow kit omitting the rRNA
421 depletion step. Resulting libraries were sequenced by an Illumina HiSeq 2000 platform.

422 **Immunoblotting**

423 Viral and cell lysates were resuspended in sodium dodecyl sulfate (SDS) sample buffer,
424 separated by electrophoresis on Bolt 4-12% Bis-Tris Plus gels (Life Technologies) and transferred
425 to nitrocellulose membranes. The membranes were then probed overnight at 4°C with a mouse
426 monoclonal anti-HIV p24 antibody (183-H12-5C, NIH AIDS reagents) or a mouse monoclonal
427 anti-HIV integrase antibody [96] in Odyssey Blocking Buffer (LI-COR). Membranes were probed
428 with fluorophore-conjugated secondary antibodies (LI-COR) and scanned using an LI-COR

429 Odyssey system. IN and CA levels in virions were quantified using Image Studio software (LI-
430 COR).

431 **Vpr-IN trans-complementation experiments**

432 Viruses bearing the class I IN mutation D116N were trans-complemented with class II mutant
433 proteins as previously described [94]; two-hundred thousand HEK293T cells were co-transfected
434 with a derivative of pNL4-3-derived plasmid bearing the IN D116N mutation, VSV-G, and
435 derivatives of the pLR2P-VprIN plasmids bearing class II IN mutations (or the compensatory
436 mutations thereof) at a ratio of 6:1:3. Cell-free virions were collected from cell culture two days
437 post-transfection. MT-4 cells were infected by the virions and the integration proficiency of trans-
438 complemented class II IN mutants was measured by the yield of progeny virions in cell culture
439 supernatants over a 6-day period as described before [94]. Briefly, MT-4 cells were incubated
440 with virus inoculum in 96 V-bottom well plates for 4hr at 37°C before washing away the inoculum
441 and replacing it with fresh media. Right after the addition of fresh media and over the ensuing 6
442 days, the number of virions in culture supernatant was quantified by measuring RT activity using
443 a Q-PCR-based assay [95].

444 **CLIP experiments**

445 CLIP experiments were conducted as previously described [15, 69, 70, 97, 98]. In short, cells in
446 15 cm cell culture plates were transfected with 30 µg full-length proviral plasmid (pNL4-3) DNA
447 containing the WT sequence or indicated IN mutations. 4-thiouridine (4SU) was added to the cell
448 culture media for 16hr before virus harvest. Cell culture supernatants were filtered through 0.22
449 µm filters and pelleted by ultracentrifugation through a 20% sucrose cushion using a Beckman
450 SW32-Ti rotor at 28,000rpm for 1.5hr at 4°C. The virus pellets were resuspended in phosphate-
451 buffered saline (PBS) and UV-crosslinked. Following lysis in 1X RIPA buffer, IN-RNA complexes
452 were immunoprecipitated using a mouse monoclonal anti-IN antibody [96] . Bound RNA was end-

453 labeled with γ-^{32}P-ATP and T4 polynucleotide kinase. The isolated protein-RNA complexes were
454 separated by SDS-PAGE, transferred to nitrocellulose membranes, and exposed to
455 autoradiography films to visualize IN-RNA complexes. Lysates and immunoprecipitates were also
456 analyzed by immunoblotting using antibodies against IN.

457 **IN multimerization in virions**

458 In 10 cm dishes, HEK293T cells were transfected with 10 µg pNL4-3 plasmid DNA containing the
459 WT sequence or indicated pol mutations within IN coding sequence. Two days post-transfection,
460 cell-free virions in cell culture supernatants were pelleted through a 20% sucrose gradient using
461 a Beckman SW41-Ti rotor at 28,000 rpm for 1.5hr at 4°C. Pelleted virions were resuspended in
462 1xPBS and treated with a membrane-permeable crosslinker, EGS (ThermoFisher Scientific), at a
463 concentration of 1mM for 30 min at room temperature. Crosslinking was stopped by the addition
464 of SDS sample buffer. The cross-linked samples were then separated on 3-8% Tris-acetate gels
465 and analyzed by immunoblotting using a mouse monoclonal anti-IN antibody [96].

466 **Virus production and transmission electron microscopy**

467 HEK293T cells in a 15 cm plate were transfected with 30 µg full-length proviral plasmid (pNL4-3)
468 DNA containing the WT sequence or indicated pol mutations within IN coding sequence. Two
469 days post transfection, cell culture supernatants were filtered through 0.22 µm filters, and pelleted
470 by ultracentrifugation using a Beckman SW32-Ti rotor at 28,000 rpm for 1.5 hr at 4°C. Fixative
471 (2% paraformaldehyde/2.5% glutaraldehyde (Polysciences Inc., Warrington, PA) in 100 mM
472 sodium cacodylate buffer, pH 7.2) was gently added to resulting pellets, and samples were
473 incubated overnight at 4°C. Samples were washed in sodium cacodylate buffer and postfixed in
474 1% osmium tetroxide (Polysciences Inc.) for 1 hr. Samples were then rinsed extensively in dH_20
475 prior to en bloc staining with 1% aqueous uranyl acetate (Ted Pella Inc., Redding, CA) for 1
476 hr. After several rinses in dH_20, samples were dehydrated in a graded series of ethanol and

477 embedded in Eponate 12 resin (Ted Pella Inc.). Sections of 95 nm were cut with a Leica Ultracut
478 UCT ultramicrotome (Leica Microsystems Inc., Bannockburn, IL), stained with uranyl acetate and
479 lead citrate, and viewed on a JEOL 1200 EX transmission electron microscope (JEOL USA Inc.,
480 Peabody, MA) equipped with an AMT 8 megapixel digital camera and AMT Image Capture Engine
481 V602 software (Advanced Microscopy Techniques, Woburn, MA).

482 **Size exclusion chromatography (SEC)**

483 All of the indicated mutations were introduced into a plasmid backbone expressing His_6 tagged
484 pNL4-3-derived IN by QuikChange site directed mutagenesis kit (Agilent) [99]. His_6 tagged
485 recombinant pNL4-3 WT and mutant Ins were expressed in BL21 (DE3) *E. coli* cells followed by
486 nickel and heparin column purification as described previously [99, 100]. Recombinant WT and
487 mutant Ins were analyzed on Superdex 200 10/300 GL column (GE Healthcare) with running
488 buffer containing 20 mM HEPES (pH 7.5), 1 M NaCl, 10% glycerol and 5 mM BME at 0.3 mL/min
489 flow rate. The proteins were diluted to 10 μM with the running buffer and incubated for 1 h at 4°C
490 followed by centrifugation at 10,000g for 10 min. Multimeric form determination was based on the
491 standards including bovine thyroglobulin (670,000 Da), bovine gamma-globulin (158,000 Da),
492 chicken ovalbumin (44,000 Da), horse myoglobin (17,000 Da) and vitamin B12 (1,350 Da).
493 Retention volumes for different oligomeric forms of IN were as follows: tetramer ~12.5 mL, dimer
494 ~14 mL, monomer ~15-16 mL.

495 **Analysis of IN-RNA binding in vitro**

496 To monitor IN-RNA interactions we utilized AlphaScreen-based assay [38], which allows to
497 monitor the ability of IN to bind and bridge between two TAR RNAs. Briefly, equal concentrations
498 (1 nM) of two synthetic TAR RNA oligonucleotides labeled either with biotin or DIG were mixed
499 and then streptavidin donor and anti-DIG acceptor beads at 0.02 mg/mL concentration were
500 supplied in a buffer containing 100 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, 1 mg/mL BSA, and 25

501 mM Tris (pH 7.4). After 2 hr incubation at 4°C, 320 nM IN was added to the reaction mixture and
502 incubated further for 1.5 hr at 4°C. AlphaScreen signals were recorded with a PerkinElmer Life
503 Sciences Enspire multimode plate reader.

504 **Structural modeling of IN**

505 Electrostatic potential maps of WT and mutant IN CTDs were created by Adaptive Poisson-
506 Boltzmann Solver (APBS) program [101] with macromolecular electrostatic calculations
507 performed in PyMOL. The published crystal structure [62] (PDB ID: 5HOT) was used as a
508 template. The calculation results are displayed as an electrostatic potential molecular surface.
509 The low, mid, and high range values are -5, 0, and 5, respectively.

510 **Analysis of IN mutations from latently infected cells**

511 Sequences identified from latently infected CD4+ T-cells [73] were downloaded from NCBI
512 GenBank based on their accession numbers (KF526120-KF526339). Sequences were imported
513 into the GLUE software framework [102, 103] and aligned. Multiple sequence alignments (MSAs)
514 containing subtype B and subtype C sequences were constructed using MUSCLE, manually
515 inspected in AliView [104] and imported into a GLUE project database. Within GLUE, MSAs were
516 constrained to the pNL4-3 reference to establish a standardized coordinate space for the gene
517 being analyzed. Amino acid frequencies at each alignment position were summarized using
518 GLUE's amino-acid frequency calculation algorithm, which accounts for contingencies such as
519 missing data and incomplete codons.

520 **ACKNOWLEDGEMENTS**

521 This work was supported by NIH grants AI150497 (SBK), AI1508470 (U54 Center for HIV RNA
522 Studies, SBK, RG), R01 AI143649 (MK) and Milton Schlesinger Student Fellowship (CSM). We
523 thank all members of the Kutluay lab for critical suggestions and feedback.

524 **REFERENCES**

525 1. Engelman AN. Multifaceted HIV integrase functionalities and therapeutic strategies for
526 their inhibition. The Journal of biological chemistry. 2019. doi: 10.1074/jbc.REV119.006901.
527 PubMed PMID: 31467082.
528 2. Hazuda DJ, Felock P, Witmer M, Wolfe A, Stillmock K, Grobler JA, et al. Inhibitors of
529 strand transfer that prevent integration and inhibit HIV-1 replication in cells. Science.
530 2000;287(5453):646-50. doi: 10.1126/science.287.5453.646. PubMed PMID: 10649997.
531 3. Summa V, Petrocchi A, Bonelli F, Crescenzi B, Donghi M, Ferrara M, et al. Discovery of
532 raltegravir, a potent, selective orally bioavailable HIV-integrase inhibitor for the treatment of HIV-
533 AIDS infection. J Med Chem. 2008;51(18):5843-55. doi: 10.1021/jm800245z. PubMed PMID:
534 18763751.
535 4. Ramanathan S, Mathias AA, German P, Kearney BP. Clinical pharmacokinetic and
536 pharmacodynamic profile of the HIV integrase inhibitor elvitegravir. Clin Pharmacokinet.
537 2011;50(4):229-44. doi: 10.2165/11584570-000000000-00000. PubMed PMID: 21348537.
538 5. Min S, Song I, Borland J, Chen S, Lou Y, Fujiwara T, et al. Pharmacokinetics and safety
539 of S/GSK1349572, a next-generation HIV integrase inhibitor, in healthy volunteers. Antimicrobial
540 agents and chemotherapy. 2010;54(1):254-8. doi: 10.1128/AAC.00842-09. PubMed PMID:
541 19884365; PubMed Central PMCID: PMCPMC2798521.
542 6. Tsiang M, Jones GS, Goldsmith J, Mulato A, Hansen D, Kan E, et al. Antiviral Activity of
543 Bictegravir (GS-9883), a Novel Potent HIV-1 Integrase Strand Transfer Inhibitor with an
544 Improved Resistance Profile. Antimicrobial agents and chemotherapy. 2016;60(12):7086-97.
545 doi: 10.1128/AAC.01474-16. PubMed PMID: 27645238; PubMed Central PMCID:
546 PMCPMC5118987.
547 7. Smith SJ, Zhao XZ, Burke TR, Jr., Hughes SH. Efficacies of Cabotegravir and
548 Bictegravir against drug-resistant HIV-1 integrase mutants. Retrovirology. 2018;15(1):37. doi:
549 10.1186/s12977-018-0420-7. PubMed PMID: 29769116; PubMed Central PMCID:
550 PMCPMC5956922.
551 8. Rockstroh JK, DeJesus E, Lennox JL, Yazdanpanah Y, Saag MS, Wan H, et al. Durable
552 efficacy and safety of raltegravir versus efavirenz when combined with tenofovir/emtricitabine in
553 treatment-naive HIV-1-infected patients: final 5-year results from STARTMRK. J Acquir Immune
554 Defic Syndr. 2013;63(1):77-85. doi: 10.1097/QAI.0b013e31828ace69. PubMed PMID:
555 23412015.
556 9. Taha H, Das A, Das S. Clinical effectiveness of dolutegravir in the treatment of
557 HIV/AIDS. Infect Drug Resist. 2015;8:339-52. Epub 2015/10/01. doi: 10.2147/IDR.S68396.
558 PubMed PMID: 26491363; PubMed Central PMCID: PMCPMC4598225.
559 10. Smith SJ, Zhao XZ, Passos DO, Lyumkis D, Burke TR, Hughes SH. HIV-1 Integrase
560 Inhibitors that are active against Drug-Resistant Integrase Mutants. Antimicrob Agents
561 Chemother. 2020. Epub 2020/06/29. doi: 10.1128/AAC.00611-20. PubMed PMID: 32601157.
562 11. Brooks KM, Sherman EM, Egelund EF, Brotherton A, Durham S, Badowski ME, et al.
563 Integrase Inhibitors: After 10 Years of Experience, Is the Best Yet to Come? Pharmacotherapy.
564 2019;39(5):576-98. doi: 10.1002/phar.2246. PubMed PMID: 30860610.
565 12. Jurado KA, Wang H, Slaughter A, Feng L, Kessl JJ, Koh Y, et al. Allosteric integrase
566 inhibitor potency is determined through the inhibition of HIV-1 particle maturation. Proceedings
567 of the National Academy of Sciences of the United States of America. 2013;110(21):8690-5. doi:
568 10.1073/pnas.1300703110. PubMed PMID: 23610442; PubMed Central PMCID: PMC3666754.
569 13. Lu R, Limon A, Devroe E, Silver PA, Cherepanov P, Engelman A. Class II integrase
570 mutants with changes in putative nuclear localization signals are primarily blocked at a
571 postnuclear entry step of human immunodeficiency virus type 1 replication. J Virol.
572 2004;78(23):12735-46. doi: 10.1128/JVI.78.23.12735-12746.2004. PubMed PMID: 15542626;
573 PubMed Central PMCID: PMCPMC525011.

574 14. Engelman A. In vivo analysis of retroviral integrase structure and function. Adv Virus
575 Res. 1999;52:411-26. Epub 1999/06/29. PubMed PMID: 10384245.
576 15. Elliott JL, Eschbach JE, Koneru PC, Li W, Puray-Chavez M, Townsend D, et al.
577 Integrase-RNA interactions underscore the critical role of integrase in HIV-1 virion
578 morphogenesis. Elife. 2020;9. Epub 2020/09/23. doi: 10.7554/eLife.54311. PubMed PMID:
579 32960169; PubMed Central PMCID: PMCPMC7671690.
580 16. Engelman A, Englund G, Orenstein JM, Martin MA, Craigie R. Multiple effects of
581 mutations in human immunodeficiency virus type 1 integrase on viral replication. J Virol.
582 1995;69(5):2729-36. Epub 1995/05/01. PubMed PMID: 7535863; PubMed Central PMCID:
583 PMC188965.
584 17. Engelman A, Cherepanov P. Retroviral Integrase Structure and DNA Recombination
585 Mechanism. Microbiol Spectr. 2014;2(6). doi: 10.1128/microbiolspec.MDNA3-0024-2014.
586 PubMed PMID: 26104441.
587 18. Engelman AN, Cherepanov P. Retroviral intasomes arising. Curr Opin Struct Biol.
588 2017;47:23-9. doi: 10.1016/j.sbi.2017.04.005. PubMed PMID: 28458055; PubMed Central
589 PMCID: PMCPMC5660667.
590 19. Craigie R, Bushman FD. HIV DNA integration. Cold Spring Harb Perspect Med.
591 2012;2(7):a006890. Epub 2012/07/05. doi: 10.1101/cshperspect.a006890. PubMed PMID:
592 22762018; PubMed Central PMCID: PMCPMC3385939.
593 20. Engelman A, Cherepanov P. The structural biology of HIV-1: mechanistic and
594 therapeutic insights. Nat Rev Microbiol. 2012;10(4):279-90. doi: 10.1038/nrmicro2747. PubMed
595 PMID: 22421880; PubMed Central PMCID: PMC3588166.
596 21. Ciuffi A, Llano M, Poeschla E, Hoffmann C, Leipzig J, Shinn P, et al. A role for
597 LEDGF/p75 in targeting HIV DNA integration. Nat Med. 2005;11(12):1287-9. doi:
598 10.1038/nm1329. PubMed PMID: 16311605.
599 22. Shun MC, Raghavendra NK, Vandegraaff N, Daigle JE, Hughes S, Kellam P, et al.
600 LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific
601 HIV-1 integration. Genes Dev. 2007;21(14):1767-78. doi: 10.1101/gad.1565107. PubMed PMID:
602 17639082; PubMed Central PMCID: PMC1920171.
603 23. Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the
604 human genome favors active genes and local hotspots. Cell. 2002;110(4):521-9. Epub
605 2002/08/31. PubMed PMID: 12202041.
606 24. Christ F, Debyser Z. The LEDGF/p75 integrase interaction, a novel target for anti-HIV
607 therapy. Virology. 2013;435(1):102-9. Epub 2012/12/12. doi: 10.1016/j.virol.2012.09.033.
608 PubMed PMID: 23217620.
609 25. Ansari-Lari MA, Donehower LA, Gibbs RA. Analysis of human immunodeficiency virus
610 type 1 integrase mutants. Virology. 1995;213(2):680. PubMed PMID: 7491794.
611 26. Bukovsky A, Gottlinger H. Lack of integrase can markedly affect human
612 immunodeficiency virus type 1 particle production in the presence of an active viral protease. J
613 Virol. 1996;70(10):6820-5. Epub 1996/10/01. PubMed PMID: 8794322; PubMed Central
614 PMCID: PMC190728.
615 27. Jenkins TM, Engelman A, Ghirlando R, Craigie R. A soluble active mutant of HIV-1
616 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. The
617 Journal of biological chemistry. 1996;271(13):7712-8. PubMed PMID: 8631811.
618 28. Kalpana GV, Reicin A, Cheng GS, Sorin M, Paik S, Goff SP. Isolation and
619 characterization of an oligomerization-negative mutant of HIV-1 integrase. Virology.
620 1999;259(2):274-85. doi: 10.1006/viro.1999.9767. PubMed PMID: 10388652.
621 29. Leavitt AD, Robles G, Alesandro N, Varmus HE. Human immunodeficiency virus type 1
622 integrase mutants retain in vitro integrase activity yet fail to integrate viral DNA efficiently during
623 infection. J Virol. 1996;70(2):721-8. PubMed PMID: 8551608; PubMed Central PMCID:
624 PMCPMC189872.

625 30. Liao WH, Wang CT. Characterization of human immunodeficiency virus type 1 Pr160
626 gag-pol mutants with truncations downstream of the protease domain. Virology.
627 2004;329(1):180-8. doi: 10.1016/j.virol.2004.08.010. PubMed PMID: 15476885.
628 31. Lu R, Ghory HZ, Engelman A. Genetic analyses of conserved residues in the carboxyl-
629 terminal domain of human immunodeficiency virus type 1 integrase. J Virol. 2005;79(16):10356-
630 68. doi: 10.1128/JVI.79.16.10356-10368.2005. PubMed PMID: 16051828; PubMed Central
631 PMCID: PMCPMC1182625.
632 32. Nakamura T, Masuda T, Goto T, Sano K, Nakai M, Harada S. Lack of infectivity of HIV-1
633 integrase zinc finger-like domain mutant with morphologically normal maturation. Biochem
634 Biophys Res Commun. 1997;239(3):715-22. doi: 10.1006/bbrc.1997.7541. PubMed PMID:
635 9367835.
636 33. Quillent C, Borman AM, Paulous S, Dauguet C, Clavel F. Extensive regions of pol are
637 required for efficient human immunodeficiency virus polyprotein processing and particle
638 maturation. Virology. 1996;219(1):29-36. doi: 10.1006/viro.1996.0219. PubMed PMID: 8623542.
639 34. Shin CG, Taddeo B, Haseltine WA, Farnet CM. Genetic analysis of the human
640 immunodeficiency virus type 1 integrase protein. J Virol. 1994;68(3):1633-42. PubMed PMID:
641 8107224; PubMed Central PMCID: PMCPMC236621.
642 35. Taddeo B, Haseltine WA, Farnet CM. Integrase mutants of human immunodeficiency
643 virus type 1 with a specific defect in integration. J Virol. 1994;68(12):8401-5. PubMed PMID:
644 7966634; PubMed Central PMCID: PMCPMC237311.
645 36. Wu X, Liu H, Xiao H, Conway JA, Hehl E, Kalpana GV, et al. Human immunodeficiency
646 virus type 1 integrase protein promotes reverse transcription through specific interactions with
647 the nucleoprotein reverse transcription complex. J Virol. 1999;73(3):2126-35. PubMed PMID:
648 9971795; PubMed Central PMCID: PMCPMC104457.
649 37. Fontana J, Jurado KA, Cheng N, Ly NL, Fuchs JR, Gorelick RJ, et al. Distribution and
650 Redistribution of HIV-1 Nucleocapsid Protein in Immature, Mature, and Integrase-Inhibited
651 Virions: a Role for Integrase in Maturation. J Virol. 2015;89(19):9765-80. doi:
652 10.1128/JVI.01522-15. PubMed PMID: 26178982; PubMed Central PMCID: PMCPMC4577894.
653 38. Kessl JJ, Kutluay SB, Townsend D, Rebensburg S, Slaughter A, Larue RC, et al. HIV-1
654 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. Cell.
655 2016;166(5):1257-68 e12. doi: 10.1016/j.cell.2016.07.044. PubMed PMID: 27565348; PubMed
656 Central PMCID: PMCPMC5003418.
657 39. Ao Z, Fowke KR, Cohen EA, Yao X. Contribution of the C-terminal tri-lysine regions of
658 human immunodeficiency virus type 1 integrase for efficient reverse transcription and viral DNA
659 nuclear import. Retrovirology. 2005;2:62. doi: 10.1186/1742-4690-2-62. PubMed PMID:
660 16232319; PubMed Central PMCID: PMCPMC1277849.
661 40. Busschots K, Voet A, De Maeyer M, Rain JC, Emiliani S, Benarous R, et al. Identification
662 of the LEDGF/p75 binding site in HIV-1 integrase. Journal of molecular biology.
663 2007;365(5):1480-92. doi: 10.1016/j.jmb.2006.10.094. PubMed PMID: 17137594.
664 41. Engelman A, Liu Y, Chen H, Farzan M, Dyda F. Structure-based mutagenesis of the
665 catalytic domain of human immunodeficiency virus type 1 integrase. J Virol. 1997;71(5):3507-
666 14. PubMed PMID: 9094622; PubMed Central PMCID: PMCPMC191497.
667 42. Limon A, Devroe E, Lu R, Ghory HZ, Silver PA, Engelman A. Nuclear localization of
668 human immunodeficiency virus type 1 preintegration complexes (PICs): V165A and R166A are
669 pleiotropic integrase mutants primarily defective for integration, not PIC nuclear import. J Virol.
670 2002;76(21):10598-607. doi: 10.1128/jvi.76.21.10598-10607.2002. PubMed PMID: 12368302;
671 PubMed Central PMCID: PMCPMC136649.
672 43. Lloyd AG, Ng YS, Muesing MA, Simon V, Mulder LC. Characterization of HIV-1
673 integrase N-terminal mutant viruses. Virology. 2007;360(1):129-35. doi:
674 10.1016/j.virol.2006.10.007. PubMed PMID: 17109911; PubMed Central PMCID:
675 PMCPMC2094525.

676 44. Lu R, Vandegraaff N, Cherepanov P, Engelman A. Lys-34, dispensable for integrase
677 catalysis, is required for preintegration complex function and human immunodeficiency virus
678 type 1 replication. J Virol. 2005;79(19):12584-91. doi: 10.1128/JVI.79.19.12584-12591.2005.
679 PubMed PMID: 16160186; PubMed Central PMCID: PMCPMC1211547.
680 45. Masuda T, Planelles V, Krogstad P, Chen IS. Genetic analysis of human
681 immunodeficiency virus type 1 integrase and the U3 att site: unusual phenotype of mutants in
682 the zinc finger-like domain. J Virol. 1995;69(11):6687-96. PubMed PMID: 7474078; PubMed
683 Central PMCID: PMCPMC189578.
684 46. Rahman S, Lu R, Vandegraaff N, Cherepanov P, Engelman A. Structure-based
685 mutagenesis of the integrase-LEDGF/p75 interface uncouples a strict correlation between in
686 vitro protein binding and HIV-1 fitness. Virology. 2007;357(1):79-90. doi:
687 10.1016/j.virol.2006.08.011. PubMed PMID: 16959283.
688 47. Riviere L, Darlix JL, Cimarelli A. Analysis of the viral elements required in the nuclear
689 import of HIV-1 DNA. J Virol. 2010;84(2):729-39. doi: 10.1128/JVI.01952-09. PubMed PMID:
690 19889772; PubMed Central PMCID: PMCPMC2798343.
691 48. Tsurutani N, Kubo M, Maeda Y, Ohashi T, Yamamoto N, Kannagi M, et al. Identification
692 of critical amino acid residues in human immunodeficiency virus type 1 IN required for efficient
693 proviral DNA formation at steps prior to integration in dividing and nondividing cells. J Virol.
694 2000;74(10):4795-806. doi: 10.1128/jvi.74.10.4795-4806.2000. PubMed PMID: 10775618;
695 PubMed Central PMCID: PMCPMC112002.
696 49. Wiskerchen M, Muesing MA. Human immunodeficiency virus type 1 integrase: effects of
697 mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA
698 templates, and sustain viral propagation in primary cells. J Virol. 1995;69(1):376-86. PubMed
699 PMID: 7983732; PubMed Central PMCID: PMCPMC188585.
700 50. Zhu K, Dobard C, Chow SA. Requirement for integrase during reverse transcription of
701 human immunodeficiency virus type 1 and the effect of cysteine mutations of integrase on its
702 interactions with reverse transcriptase. J Virol. 2004;78(10):5045-55. doi:
703 10.1128/jvi.78.10.5045-5055.2004. PubMed PMID: 15113886; PubMed Central PMCID:
704 PMCPMC400327.
705 51. De Houwer S, Demeulemeester J, Thys W, Rocha S, Dirix L, Gijsbers R, et al. The HIV-
706 1 integrase mutant R263A/K264A is 2-fold defective for TRN-SR2 binding and viral nuclear
707 import. The Journal of biological chemistry. 2014;289(36):25351-61. doi:
708 10.1074/jbc.M113.533281. PubMed PMID: 25063804; PubMed Central PMCID:
709 PMCPMC4155696.
710 52. Johnson BC, Metifiot M, Ferris A, Pommier Y, Hughes SH. A homology model of HIV-1
711 integrase and analysis of mutations designed to test the model. Journal of molecular biology.
712 2013;425(12):2133-46. doi: 10.1016/j.jmb.2013.03.027. PubMed PMID: 23542006.
713 53. Mohammed KD, Topper MB, Muesing MA. Sequential deletion of the integrase (Gag-
714 Pol) carboxyl terminus reveals distinct phenotypic classes of defective HIV-1. J Virol.
715 2011;85(10):4654-66. doi: 10.1128/JVI.02374-10. PubMed PMID: 21367893; PubMed Central
716 PMCID: PMC3126176.
717 54. Shehu-Xhilaga M, Hill M, Marshall JA, Kappes J, Crowe SM, Mak J. The conformation of
718 the mature dimeric human immunodeficiency virus type 1 RNA genome requires packaging of
719 pol protein. J Virol. 2002;76(9):4331-40. doi: 10.1128/jvi.76.9.4331-4340.2002. PubMed PMID:
720 11932399; PubMed Central PMCID: PMCPMC155102.
721 55. Engelman A, Craigie R. Identification of conserved amino acid residues critical for
722 human immunodeficiency virus type 1 integrase function in vitro. J Virol. 1992;66(11):6361-9.
723 PubMed PMID: 1404595; PubMed Central PMCID: PMCPMC240128.
724 56. Lutzke RA, Plasterk RH. Structure-based mutational analysis of the C-terminal DNA-
725 binding domain of human immunodeficiency virus type 1 integrase: critical residues for protein

726 oligomerization and DNA binding. J Virol. 1998;72(6):4841-8. PubMed PMID: 9573250; PubMed
727 Central PMCID: PMCPMC110031.
728 57. Lutzke RA, Vink C, Plasterk RH. Characterization of the minimal DNA-binding domain of
729 the HIV integrase protein. Nucleic Acids Res. 1994;22(20):4125-31. doi:
730 10.1093/nar/22.20.4125. PubMed PMID: 7937137; PubMed Central PMCID: PMCPMC331899.
731 58. Balakrishnan M, Yant SR, Tsai L, O'Sullivan C, Bam RA, Tsai A, et al. Non-catalytic site
732 HIV-1 integrase inhibitors disrupt core maturation and induce a reverse transcription block in
733 target cells. PloS one. 2013;8(9):e74163. Epub 2013/09/17. doi: 10.1371/journal.pone.0074163
734 PONE-D-13-26050 [pii]. PubMed PMID: 24040198; PubMed Central PMCID: PMC3767657.
735 59. Desimmie BA, Schrijvers R, Demeulemeester J, Borrenberghs D, Weydert C, Thys W, et
736 al. LEDGINs inhibit late stage HIV-1 replication by modulating integrase multimerization in the
737 virions. Retrovirology. 2013;10:57. Epub 2013/06/01. doi: 1742-4690-10-57 [pii]
738 10.1186/1742-4690-10-57. PubMed PMID: 23721378; PubMed Central PMCID: PMC3671127.
739 60. Sharma A, Slaughter A, Jena N, Feng L, Kessl JJ, Fadel HJ, et al. A new class of
740 multimerization selective inhibitors of HIV-1 integrase. PLoS pathogens. 2014;10(5):e1004171.
741 doi: 10.1371/journal.ppat.1004171. PubMed PMID: 24874515; PubMed Central PMCID:
742 PMC4038613.
743 61. Gupta K, Brady T, Dyer BM, Malani N, Hwang Y, Male F, et al. Allosteric inhibition of
744 human immunodeficiency virus integrase: late block during viral replication and abnormal
745 multimerization involving specific protein domains. The Journal of biological chemistry.
746 2014;289(30):20477-88. doi: 10.1074/jbc.M114.551119. PubMed PMID: 24904063; PubMed
747 Central PMCID: PMC4110260.
748 62. Gupta K, Turkki V, Sherrill-Mix S, Hwang Y, Eilers G, Taylor L, et al. Structural Basis for
749 Inhibitor-Induced Aggregation of HIV Integrase. PLoS Biol. 2016;14(12):e1002584. doi:
750 10.1371/journal.pbio.1002584. PubMed PMID: 27935939; PubMed Central PMCID:
751 PMCPMC5147827 contained in this manuscript to improve therapeutics for HIV infection. The
752 other authors declare that they have no conflict of interest.
753 63. Christ F, Shaw S, Demeulemeester J, Desimmie BA, Marchand A, Butler S, et al. Small-
754 molecule inhibitors of the LEDGF/p75 binding site of integrase block HIV replication and
755 modulate integrase multimerization. Antimicrobial agents and chemotherapy. 2012;56(8):4365-
756 74. doi: 10.1128/AAC.00717-12. PubMed PMID: 22664975; PubMed Central PMCID:
757 PMC3421592.
758 64. Christ F, Voet A, Marchand A, Nicolet S, Desimmie BA, Marchand D, et al. Rational
759 design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication.
760 Nature chemical biology. 2010;6(6):442-8. doi: 10.1038/nchembio.370. PubMed PMID:
761 20473303.
762 65. Maehigashi T, Ahn S, Kim UI, Lindenberger J, Oo A, Koneru PC, et al. A highly potent
763 and safe pyrrolopyridine-based allosteric HIV-1 integrase inhibitor targeting host LEDGF/p75-
764 integrase interaction site. PLoS pathogens. 2021;17(7):e1009671. Epub 2021/07/23. doi:
765 10.1371/journal.ppat.1009671. PubMed PMID: 34293041; PubMed Central PMCID:
766 PMCPMC8297771 following competing interests: S.A., U.K., and K.K. are employees of ST
767 Pharm, B.K. is consultant of ST Pharm, and A.N.E. consults for ViiV Healthcare, Co.. All other
768 authors declare no competing interests.
769 66. Serganov A, Patel DJ. Towards deciphering the principles underlying an mRNA
770 recognition code. Curr Opin Struct Biol. 2008;18(1):120-9. Epub 2008/02/08. doi:
771 10.1016/j.sbi.2007.12.006. PubMed PMID: 18255277; PubMed Central PMCID:
772 PMCPMC3744885.
773 67. Bahadur RP, Zacharias M, Janin J. Dissecting protein-RNA recognition sites. Nucleic
774 Acids Res. 2008;36(8):2705-16. Epub 2008/03/21. doi: 10.1093/nar/gkn102. PubMed PMID:
775 18353859; PubMed Central PMCID: PMCPMC2377425.

776 68. Stefl R, Skrisovska L, Allain FH. RNA sequence- and shape-dependent recognition by
777 proteins in the ribonucleoprotein particle. EMBO Rep. 2005;6(1):33-8. Epub 2005/01/12. doi:
778 10.1038/sj.embor.7400325. PubMed PMID: 15643449; PubMed Central PMCID:
779 PMCPMC1299235.
780 69. Shema Mugisha C, Tenneti K, Kutluay SB. Clip for studying protein-RNA interactions
781 that regulate virus replication. Methods. 2020;183:84-92. Epub 2019/11/26. doi:
782 10.1016/j.ymeth.2019.11.011. PubMed PMID: 31765715; PubMed Central PMCID:
783 PMCPMC6963001.
784 70. Bieniasz PD, Kutluay SB. CLIP-related methodologies and their application to
785 retrovirology. Retrovirology. 2018;15(1):35. Epub 2018/05/03. doi: 10.1186/s12977-018-0417-2.
786 PubMed PMID: 29716635; PubMed Central PMCID: PMCPMC5930818.
787 71. Deng N, Hoyte A, Mansour YE, Mohamed MS, Fuchs JR, Engelman AN, et al. Allosteric
788 HIV-1 integrase inhibitors promote aberrant protein multimerization by directly mediating inter-
789 subunit interactions: Structural and thermodynamic modeling studies. Protein Sci.
790 2016;25(11):1911-7. Epub 2016/10/26. doi: 10.1002/pro.2997. PubMed PMID: 27503276;
791 PubMed Central PMCID: PMCPMC5079246.
792 72. Shkriabai N, Dharmarajan V, Slaughter A, Kessl JJ, Larue RC, Feng L, et al. A critical
793 role of the C-terminal segment for allosteric inhibitor-induced aberrant multimerization of HIV-1
794 integrase. J Biol Chem. 2014;289(38):26430-40. Epub 2014/08/15. doi:
795 10.1074/jbc.M114.589572. PubMed PMID: 25118283; PubMed Central PMCID:
796 PMCPMC4176244.
797 73. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DI, et al. Replication-
798 competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell.
799 2013;155(3):540-51. Epub 2013/11/19. doi: 10.1016/j.cell.2013.09.020. PubMed PMID:
800 24243014; PubMed Central PMCID: PMCPMC3896327.
801 74. Pollack RA, Jones RB, Pertea M, Bruner KM, Martin AR, Thomas AS, et al. Defective
802 HIV-1 Proviruses Are Expressed and Can Be Recognized by Cytotoxic T Lymphocytes, which
803 Shape the Proviral Landscape. Cell Host Microbe. 2017;21(4):494-506 e4. Epub 2017/04/14.
804 doi: 10.1016/j.chom.2017.03.008. PubMed PMID: 28407485; PubMed Central PMCID:
805 PMCPMC5433942.
806 75. Bayer TS, Booth LN, Knudsen SM, Ellington AD. Arginine-rich motifs present multiple
807 interfaces for specific binding by RNA. RNA. 2005;11(12):1848-57. Epub 2005/11/30. doi:
808 10.1261/rna.2167605. PubMed PMID: 16314457; PubMed Central PMCID: PMCPMC1370873.
809 76. Jones S, Daley DT, Luscombe NM, Berman HM, Thornton JM. Protein-RNA interactions:
810 a structural analysis. Nucleic Acids Res. 2001;29(4):943-54. Epub 2001/02/13. doi:
811 10.1093/nar/29.4.943. PubMed PMID: 11160927; PubMed Central PMCID: PMCPMC29619.
812 77. Jeong E, Kim H, Lee SW, Han K. Discovering the interaction propensities of amino acids
813 and nucleotides from protein-RNA complexes. Mol Cells. 2003;16(2):161-7. Epub 2003/12/04.
814 PubMed PMID: 14651256.
815 78. Kim OT, Yura K, Go N. Amino acid residue doublet propensity in the protein-RNA
816 interface and its application to RNA interface prediction. Nucleic Acids Res. 2006;34(22):6450-
817 60. Epub 2006/11/30. doi: 10.1093/nar/gkl819. PubMed PMID: 17130160; PubMed Central
818 PMCID: PMCPMC1761430.
819 79. Kruger DM, Neubacher S, Grossmann TN. Protein-RNA interactions: structural
820 characteristics and hotspot amino acids. RNA. 2018;24(11):1457-65. Epub 2018/08/11. doi:
821 10.1261/rna.066464.118. PubMed PMID: 30093489; PubMed Central PMCID:
822 PMCPMC6191724.
823 80. Schulze-Gahmen U, Hurley JH. Structural mechanism for HIV-1 TAR loop recognition by
824 Tat and the super elongation complex. Proceedings of the National Academy of Sciences of the
825 United States of America. 2018;115(51):12973-8. Epub 2018/12/06. doi:

826 10.1073/pnas.1806438115. PubMed PMID: 30514815; PubMed Central PMCID:
827 PMCPMC6305006.
828 81. Schulze-Gahmen U, Echeverria I, Stjepanovic G, Bai Y, Lu H, Schneidman-Duhovny D,
829 et al. Insights into HIV-1 proviral transcription from integrative structure and dynamics of the
830 Tat:AFF4:P-TEFb:TAR complex. Elife. 2016;5. Epub 2016/10/21. doi: 10.7554/eLife.15910.
831 PubMed PMID: 27731797; PubMed Central PMCID: PMCPMC5072841.
832 82. Han SP, Tang YH, Smith R. Functional diversity of the hnRNPs: past, present and
833 perspectives. Biochem J. 2010;430(3):379-92. doi: 10.1042/BJ20100396. PubMed PMID:
834 20795951.
835 83. Jeong S. SR Proteins: Binders, Regulators, and Connectors of RNA. Mol Cells.
836 2017;40(1):1-9. doi: 10.14348/molcells.2017.2319. PubMed PMID: 28152302; PubMed Central
837 PMCID: PMCPMC5303883.
838 84. Lunde BM, Moore C, Varani G. RNA-binding proteins: modular design for efficient
839 function. Nat Rev Mol Cell Biol. 2007;8(6):479-90. Epub 2007/05/03. doi: 10.1038/nrm2178.
840 PubMed PMID: 17473849; PubMed Central PMCID: PMCPMC5507177.
841 85. Ban T, Zhu JK, Melcher K, Xu HE. Structural mechanisms of RNA recognition:
842 sequence-specific and non-specific RNA-binding proteins and the Cas9-RNA-DNA complex.
843 Cell Mol Life Sci. 2015;72(6):1045-58. Epub 2014/11/30. doi: 10.1007/s00018-014-1779-9.
844 PubMed PMID: 25432705.
845 86. Wijting IEA, Lungu C, Rijnders BJA, van der Ende ME, Pham HT, Mesplede T, et al.
846 HIV-1 Resistance Dynamics in Patients With Virologic Failure to Dolutegravir Maintenance
847 Monotherapy. J Infect Dis. 2018;218(5):688-97. doi: 10.1093/infdis/jiy176. PubMed PMID:
848 29617822.
849 87. Radzio-Basu J, Council O, Cong ME, Ruone S, Newton A, Wei X, et al. Drug resistance
850 emergence in macaques administered cabotegravir long-acting for pre-exposure prophylaxis
851 during acute SHIV infection. Nat Commun. 2019;10(1):2005. Epub 2019/05/01. doi:
852 10.1038/s41467-019-10047-w. PubMed PMID: 31043606; PubMed Central PMCID:
853 PMCPMC6494879.
854 88. Anstett K, Brenner B, Mesplede T, Wainberg MA. HIV drug resistance against strand
855 transfer integrase inhibitors. Retrovirology. 2017;14(1):36. doi: 10.1186/s12977-017-0360-7.
856 PubMed PMID: 28583191; PubMed Central PMCID: PMCPMC5460515.
857 89. Rossouw TM, Hitchcock S, Botes M. The end of the line? A case of drug resistance to
858 third-line antiretroviral therapy. South Afr J HIV Med. 2016;17(1):454. Epub 2016/05/31. doi:
859 10.4102/sajhivmed.v17i1.454. PubMed PMID: 29568612; PubMed Central PMCID:
860 PMCPMC5843205.
861 90. Zhang WW, Cheung PK, Oliveira N, Robbins MA, Harrigan PR, Shahid A. Accumulation
862 of Multiple Mutations In Vivo Confers Cross-Resistance to New and Existing Integrase
863 Inhibitors. J Infect Dis. 2018;218(11):1773-6. doi: 10.1093/infdis/jiy428. PubMed PMID:
864 30010985.
865 91. Ndashimye E, Li Y, Reyes PS, Avino M, Olabode AS, Kityo CM, et al. High-level
866 resistance to bictegravir and cabotegravir in subtype A- and D-infected HIV-1 patients failing
867 raltegravir with multiple resistance mutations. J Antimicrob Chemother. 2021. Epub 2021/08/29.
868 doi: 10.1093/jac/dkab276. PubMed PMID: 34453542.
869 92. Kessl JJ, Sharma A, Kvaratskhelia M. Methods for the Analyses of Inhibitor-Induced
870 Aberrant Multimerization of HIV-1 Integrase. Methods Mol Biol. 2016;1354:149-64. doi:
871 10.1007/978-1-4939-3046-3_10. PubMed PMID: 26714710; PubMed Central PMCID:
872 PMCPMC4793731.
873 93. Madison MK, Lawson DQ, Elliott J, Ozanturk AN, Koneru PC, Townsend D, et al.
874 Allosteric HIV-1 Integrase Inhibitors Lead to Premature Degradation of the Viral RNA Genome
875 and Integrase in Target Cells. J Virol. 2017;91(17). doi: 10.1128/JVI.00821-17. PubMed PMID:
876 28615207; PubMed Central PMCID: PMCPMC5553177.

877 94. Liu H, Wu X, Xiao H, Conway JA, Kappes JC. Incorporation of functional human
878 immunodeficiency virus type 1 integrase into virions independent of the Gag-Pol precursor
879 protein. J Virol. 1997;71(10):7704-10. PubMed PMID: 9311854; PubMed Central PMCID:
880 PMCPMC192121.
881 95. Pizzato M, Erlwein O, Bonsall D, Kaye S, Muir D, McClure MO. A one-step SYBR Green
882 I-based product-enhanced reverse transcriptase assay for the quantitation of retroviruses in cell
883 culture supernatants. J Virol Methods. 2009;156(1-2):1-7. doi: 10.1016/j.jviromet.2008.10.012.
884 PubMed PMID: 19022294.
885 96. Bouyac-Bertoia M, Dvorin JD, Fouchier RA, Jenkins Y, Meyer BE, Wu LI, et al. HIV-1
886 infection requires a functional integrase NLS. Molecular cell. 2001;7(5):1025-35. PubMed PMID:
887 11389849.
888 97. Kutluay SB, Bieniasz PD. Analysis of HIV-1 Gag-RNA Interactions in Cells and Virions
889 by CLIP-seq. Methods Mol Biol. 2016;1354:119-31. doi: 10.1007/978-1-4939-3046-3_8.
890 PubMed PMID: 26714708; PubMed Central PMCID: PMCPMC6548315.
891 98. Kutluay SB, Zang T, Blanco-Melo D, Powell C, Jannain D, Errando M, et al. Global
892 changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. Cell.
893 2014;159(5):1096-109. Epub 2014/11/25. doi: 10.1016/j.cell.2014.09.057. PubMed PMID:
894 25416948; PubMed Central PMCID: PMCPMC4247003.
895 99. Kessl JJ, Jena N, Koh Y, Taskent-Sezgin H, Slaughter A, Feng L, et al. Multimode,
896 cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. The Journal of
897 biological chemistry. 2012;287(20):16801-11. doi: 10.1074/jbc.M112.354373. PubMed PMID:
898 22437836; PubMed Central PMCID: PMC3351293.
899 100. Cherepanov P. LEDGF/p75 interacts with divergent lentiviral integrases and modulates
900 their enzymatic activity in vitro. Nucleic Acids Res. 2007;35(1):113-24. doi: 10.1093/nar/gkl885.
901 PubMed PMID: 17158150; PubMed Central PMCID: PMCPMC1802576.
902 101. Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, et al. Improvements to the
903 APBS biomolecular solvation software suite. Protein Sci. 2018;27(1):112-28. Epub 2017/08/25.
904 doi: 10.1002/pro.3280. PubMed PMID: 28836357; PubMed Central PMCID: PMCPMC5734301.
905 102. Singer JB, Thomson EC, Hughes J, Aranday-Cortes E, McLauchlan J, da Silva Filipe A,
906 et al. Interpreting Viral Deep Sequencing Data with GLUE. Viruses. 2019;11(4). Epub
907 2019/04/17. doi: 10.3390/v11040323. PubMed PMID: 30987147; PubMed Central PMCID:
908 PMCPMC6520954.
909 103. Singer JB, Thomson EC, McLauchlan J, Hughes J, Gifford RJ. GLUE: a flexible software
910 system for virus sequence data. BMC Bioinformatics. 2018;19(1):532. Epub 2018/12/20. doi:
911 10.1186/s12859-018-2459-9. PubMed PMID: 30563445; PubMed Central PMCID:
912 PMCPMC6299651.
913 104. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets.
914 Bioinformatics. 2014;30(22):3276-8. Epub 2014/08/07. doi: 10.1093/bioinformatics/btu531.
915 PubMed PMID: 25095880; PubMed Central PMCID: PMCPMC4221126.
916

917

918

919

920 **FIGURE LEGENDS:**

921 **Figure 1. D256N and D270N substitutions in HIV-1 IN suppresses the replication defect of**
922 **R269A/K273A class-II IN mutant virus. (A)** MT4-LTR-GFP cells were infected with equal particle
923 numbers of either WT or R269A/K273A IN mutant HIV-1 (NL4-3). The R269A/K273A IN viruses
924 were serially passaged for three times until the emergence of compensatory mutations. The
925 graphs represent the percentage of GFP positive cells as assessed by FACS over three passages
926 at the indicated days post-infection (dpi). **(B)** HIV-1 genomic RNA was isolated from viruses
927 collected from cell culture supernatants over the three passages (i.e. P1, P2, P3) and at the
928 indicated days post-infection (i.e. D29, D40, etc.). Whole-genome deep sequencing revealed the
929 acquisition of D256N and D270N compensatory mutations. Heatmap shows the percentage of
930 mutations at the indicated passages and days post-infection (dpi). **(C)** HEK293T cells were
931 transfected with full-length pNL4-3 expression plasmids carrying the D256N, D270N and
932 D256N/D270N (D2N) IN mutations introduced on the WT IN and IN R269A/K273A backbones.
933 Cell lysates and virions were purified two days post transfection and analyzed by immunoblotting
934 for CA and IN. The image is representative of five independent experiments. **(D)** HEK293T cells
935 were transfected as in C and cell culture supernatants containing viruses were titered on TZM-bl
936 indicator cells. The titers are presented relative to WT (set to 1). The columns represent the
937 average of five independent experiments and the error bars represent SEM (****$p<0.0001$, by
938 one-way ANOVA with Dunnett's multiple comparison test).

939 **Figure 2. D256N and D270N substitutions restore IN-gRNA binding and accurate virion**
940 **maturation for the R269A/K273A class II IN mutant virus. (A)** Autoradiogram of IN-RNA
941 adducts immunoprecipitated from virions bearing the indicated substitutions in IN. Immunoblots
942 below show the amount of IP'ed IN or CA protein in lysates. Data shown are representative of at
943 least three independent experiments. **(B)** Examination of virion maturation in WT and mutant IN
944 viruses by thin section electron microscopy (TEM). The graph quantifies virion morphologies;

945 each column is the average of two independent experiments and error bars represent the SEM
946 (****p<0.0001, by repeated measures one-way ANOVA).

947 **Figure 3. Restoring the net charge of IN-CTD restores RNA binding and infectivity for the**
948 **R269A/K273A class II mutant. (A)** Schematic diagram of IN and sequence of CTD residues with
949 basic and acidic amino acids highlighted in blue and red, respectively. **(B, C)** HEK293T cells were
950 transfected with full-length proviral HIV-1_{NL4-3} expression plasmids carrying *pol* mutations
951 encoding for the indicated IN substitutions. **(B)** Cell lysates and purified virions were harvested
952 two days post transfection and analyzed by immunoblotting for CA and IN. Representative image
953 of one of five independent experiments is shown. **(C)** Infectious titers of WT or IN mutant HIV-
954 1_{NL4-3} viruses in cell culture supernatants were determined on TZM-bl indicator cells and
955 normalized relative to particle number based on RT activity. Titer values are expressed relative
956 to WT (set to 1). The columns represent the average of four independent experiments and the
957 error bars represent SEM (****p<0.0001, by one-way ANOVA with Dunnett's multiple comparison
958 test). **(D)** Autoradiogram of IN-RNA adducts IP'ed from WT or IN mutant HIV-1_{NL4-3} virions. The
959 amount of immunoprecipitated IN protein was assessed by the immunoblot below. Results are
960 representative of four independent replicates. **(E)** A representative growth curve of HIV-1 $_{NL4-3\ IN}$
961 $_{(D116N)}$ viruses that were trans-complemented with the indicated Vpr-IN mutant proteins in cell
962 culture. Y-axis indicates fold increase in virion yield over day 0 as measured by RT activity in
963 culture supernatants. Error bars show SEM from 3 independent many replicates.

964 **Figure 4. D256R and D256K substitutions restore IN-RNA binding and infectivity for the**
965 **HIV-$1_{NL4-3\ IN(R262A/R263A)}$ virus. (A-E)** HEK293T cells were transfected with proviral HIV-1_{NL4-3}
966 expression plasmids carrying *pol* mutations for the indicated IN substitutions. **(A, C)** Cell lysates
967 and virions were purified two days post transfection and analyzed by immunoblotting for CA and
968 IN. **(B, D)** WT or IN mutant HIV-1_{NL4-3} viruses in cell culture supernatants were titered on TZM-bl
969 indicator cells. The titer values are presented relative to WT (set to 1). The columns represent the

970 average of two-three independent experiments, and the error bars represent SEM (****p<0.0001,
971 by one-way ANOVA with Dunnett's multiple comparisons). **(E)** Autoradiogram of IN-RNA adducts
972 immunoprecipitated from WT or IN mutant HIV-1_{NL4-3} virions. The amount of immunoprecipitated
973 IN was assessed by the immunoblot shown below. Immunoblots and CLIP autoradiographs are
974 representative of four independent replicates. **(F)** A representative growth curve of HIV-$1_{NL4-3\ IN}$
975 $_{(D116N)}$ viruses that were trans-complemented with the indicated IN mutant proteins in cell culture.
976 Y-axis indicates fold increase in virion yield over day 0 as measured by RT activity in culture
977 supernatants. Error bars show SEM from 3 independent many replicates.

978 **Figure 5. Electrostatic potential maps of HIV-1 IN bearing class II and compensatory**
979 **mutations. (A-G)** Electrostatic potential maps of the indicated HIV-1 IN mutants derived from the
980 crystal structure of Gupta et al. [62] is depicted. Calculation results are displayed as an
981 electrostatic potential molecular surface. The low, mid, and high range values are -5, 0, and 5,
982 respectively. (**H**) A cartoon view of the CTD structure [62] is shown with the C-terminal tail (aa
983 261-275) and the loop (aa 252-257) colored in cyan and orange, respectively. Side chains of
984 indicated residues are shown.

985 **Figure 6. Assessing multimerization properties of IN in mutant viruses. (A)** Purified WT or
986 IN mutant HIV-1_{NL4-3} virions were treated with 1 mM EGS, and virus lysates analyzed by
987 immunoblotting using antibodies against IN following separation on 6% Tris-acetate gels. The
988 position of monomers (M), dimers (D), and tetramers (T) are indicated by arrows in a
989 representative western blot. **(B,C)** Quantification of IN multimerization in virions from experiments
990 conducted as in A. Error bars show the SEM from three independent experiments. **(D-I)**
991 Biochemical analysis of IN mutants multimerization and their interactions with HIV-1 TAR RNA.
992 **(D-H)** Representative SEC traces for indicated recombinant IN proteins. The X-axis indicates
993 elution volume (mL) and Y-axis indicates the intensity of absorbance (mAU). Tetramers (T),
994 dimers (D), and monomers (M) are indicated. **(I)** Summary of mutant INs bridging TAR RNA

995 compared to WT IN. Alpha screen counts at 320 nM for each protein is shown. The graphs show
996 average values of three independent experiments and the error bars indicate standard deviation.

997 **Figure 7. Secondary site suppressors of the IN R269A/K273A mutant increase**
998 **susceptibility to ALLINIs. (A, B, D, E)** Titers of viruses bearing the indicated substitutions in IN
999 produced from HEK293T cells at different concentrations of ALLINIs, BI-B2 or BI-D. WT or IN
1000 mutant HIV-1_{NL4-3} in cell culture supernatants were titered on TZM-bl indicator cells. The titer
1001 values are represented relative to the mock control of each mutant (set to 1). Data are from two
1002 independent biological replicates. **(C, F)** Autoradiogram of IN-RNA adducts immunoprecipitated
1003 from WT or IN mutant HIV-1_{NL4-3} virions produced from HEK293T cells in the presence of 0.1μM
1004 of BI-D. The amount of immunoprecipitated IN protein was visualized by the immunoblot shown
1005 below. Immunoblots and CLIP autoradiographs results are a representative of three independent
1006 replicates. **(G)** The GSK-1264 binding pocket between CCD-CCD (brown) and CTD (light blue)
1007 from different IN subunits is shown [62] (PDB ID: 5HOT). GSK-1264 is in green with nitrogen and
1008 oxygen atoms colored blue and red, respectively. The side chain of R269 is shown with nitrogen
1009 atoms colored in dark blue. Interactions between CTD R269 and CCD Q168 are indicated by
1010 dashed lines.

1011 **Figure 8. Characterization of IN mutations present in latently infected CD4+ T-cells. (A)**
1012 **Alignment from Gifford- (B)** HEK293T cells were transfected with proviral HIV-1_{NL4-3} expression
1013 plasmids carrying the R224Q, S230N, E246K, and G272R IN mutations. Cell lysates and virions
1014 were purified two days post transfection and analyzed by immunoblotting for CA and IN. The
1015 image is representative of two independent experiments. **(C)** WT or IN mutant HIV-1L4-3 viruses
1016 in cell culture supernatants were titered on TZM-bl indicator cells. The titers are presented relative
1017 to WT (set to 1). The columns represent the average of three independent experiments and the
1018 error bars represent SEM (****$p<0.0001$, by one-way ANOVA with Dunnett's multiple comparison

1019 test). **(D)** Autoradiogram of IN-RNA adducts immunoprecipitated from virions bearing the

1020 indicated substitutions in IN. Immunoblots below show the amount of immunoprecipitated IN.

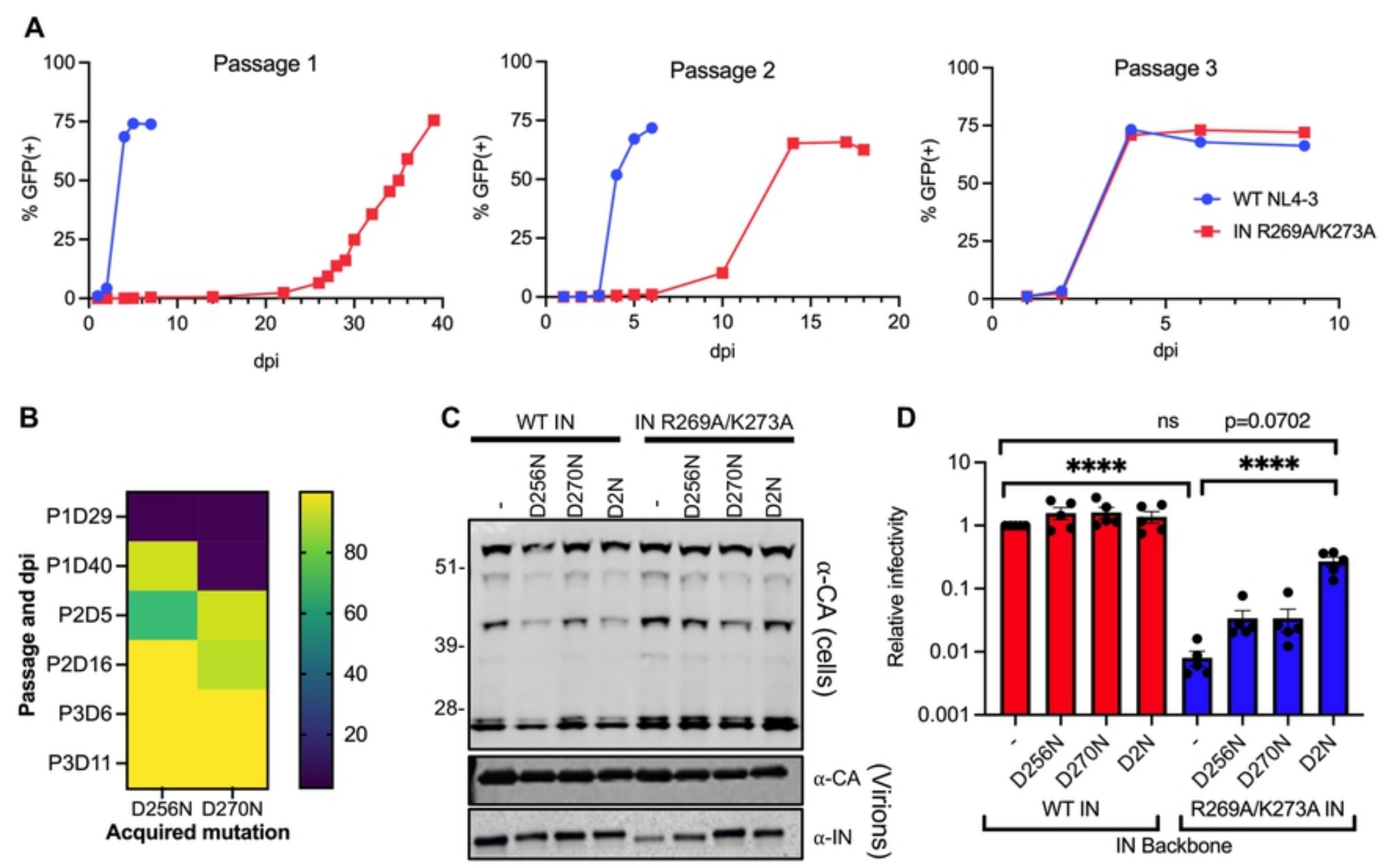
A
Passage 1
Passage 2
Passage 3
% GFP(+)
dpi
WT NL4-3
IN R269A/K273A
B
Passage and dpi
P1D29
P1D40
P2D5
P2D16
P3D6
P3D11
D256N
D270N
Acquired mutation
C
WT IN
IN R269A/K273A
D256N
D270N
D2N
51-
39-
28-
α-CA (cells)
α-CA
α-IN
(Virions)
D
ns
p=0.0702
****
****
Relative infectivity
WT IN
R269A/K273A IN
IN Backbone

Figure 1

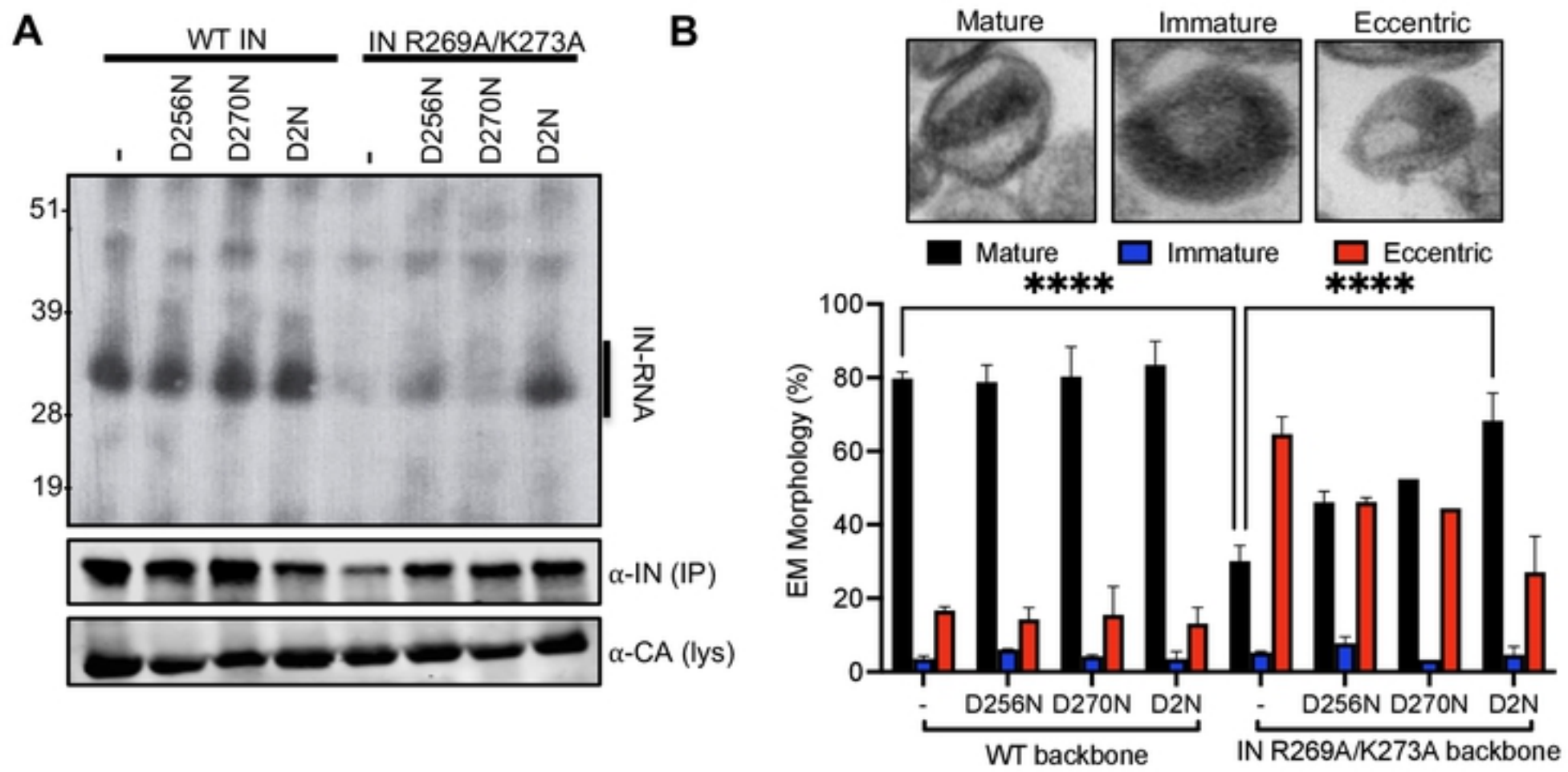
A
WT IN
IN R269A/K273A
-
D256N
D270N
D2N
-
D256N
D270N
D2N
51
39
28
19
IN-RNA
α-IN (IP)
α-CA (lys)
B
Mature
Immature
Eccentric
Mature
Immature
Eccentric
****
****
EM Morphology (%)
100
80
60
40
20
0
-
D256N
D270N
D2N
-
D256N
D270N
D2N
WT backbone
IN R269A/K273A backbone

Figure 2

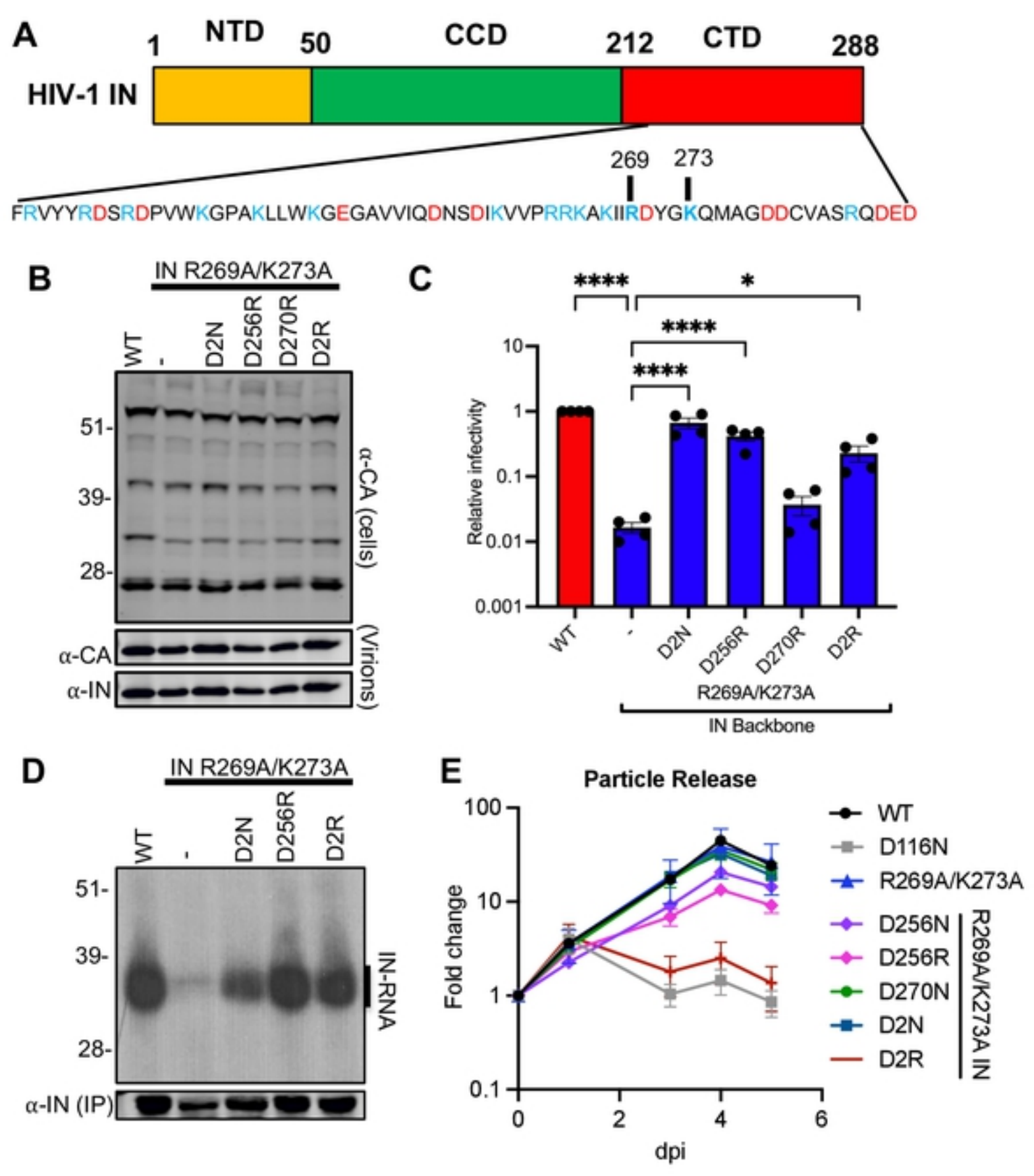
A
HIV-1 IN
1
NTD
50
CCD
212
CTD
288
269
273
FRVYYRDSRDPVWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED
B
IN R269A/K273A
WT
-
D2N
D256R
D270R
D2R
51-
39-
28-
α-CA (cells)
α-CA
α-IN
(Virions)
C
Relative infectivity
10
1
0.1
0.01
0.001
****
*
WT
-
D2N
D256R
D270R
D2R
R269A/K273A
IN Backbone
D
IN R269A/K273A
WT
-
D2N
D256R
D2R
51-
39-
28-
IN-RNA
α-IN (IP)
E
Particle Release
Fold change
100
10
1
0.1
0
2
4
6
dpi
WT
D116N
R269A/K273A
D256N
D256R
D270N
D2N
D2R
R269A/K273A IN

Figure 3

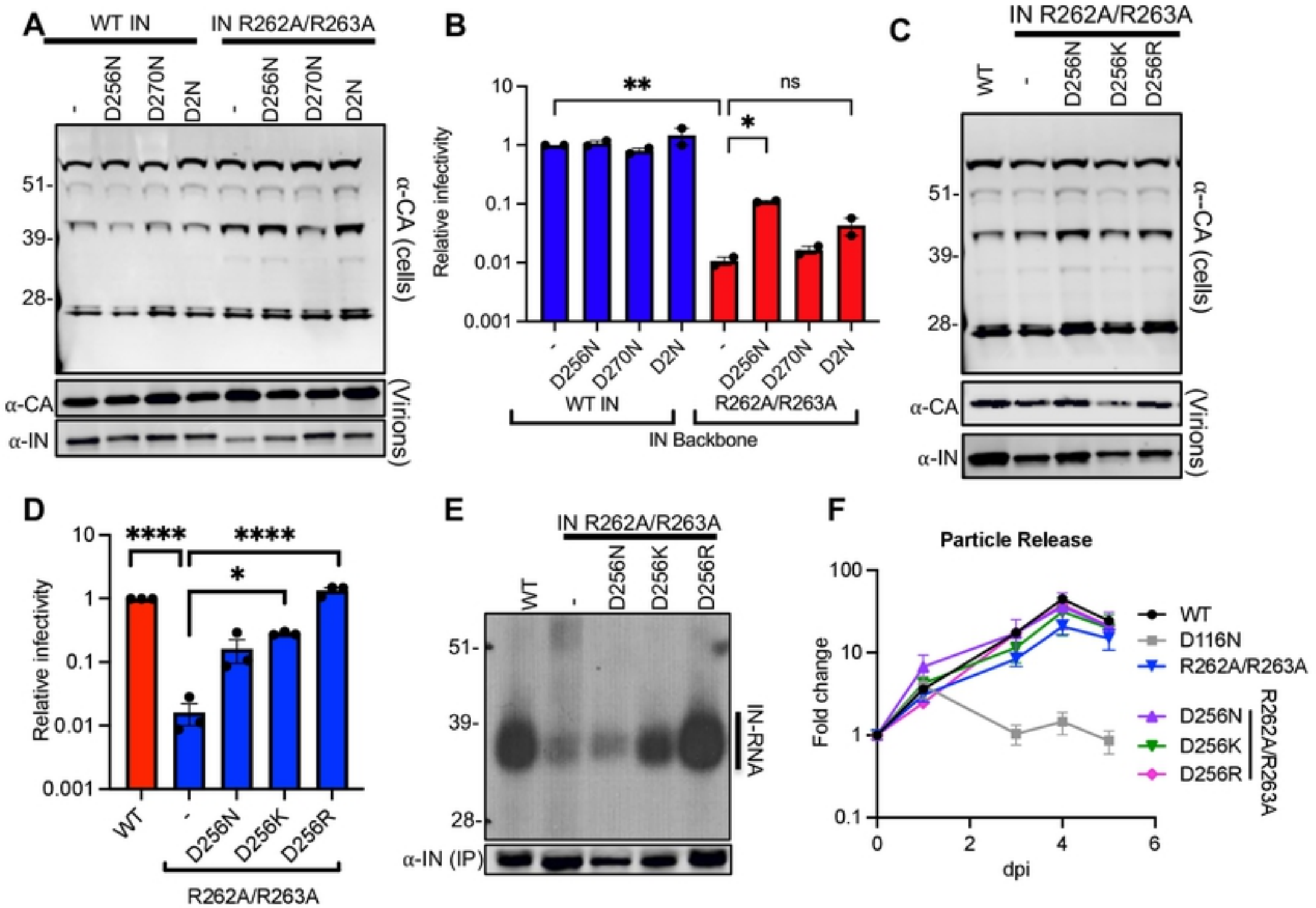
A
WT IN
IN R262A/R263A
-
D256N
D270N
D2N
51-
39-
28-
α-CA (cells)
α-CA
α-IN
(Virions)
B
Relative infectivity
10
1
0.1
0.01
0.001
**
ns
*
WT IN
R262A/R263A
IN Backbone
C
IN R262A/R263A
WT
D256N
D256K
D256R
D
****
****
*
R262A/R263A
E
IN R262A/R263A
α-IN (IP)
IN-RNA
F
Particle Release
Fold change
dpi
WT
D116N
R262A/R263A
D256N
D256K
D256R

Figure 4

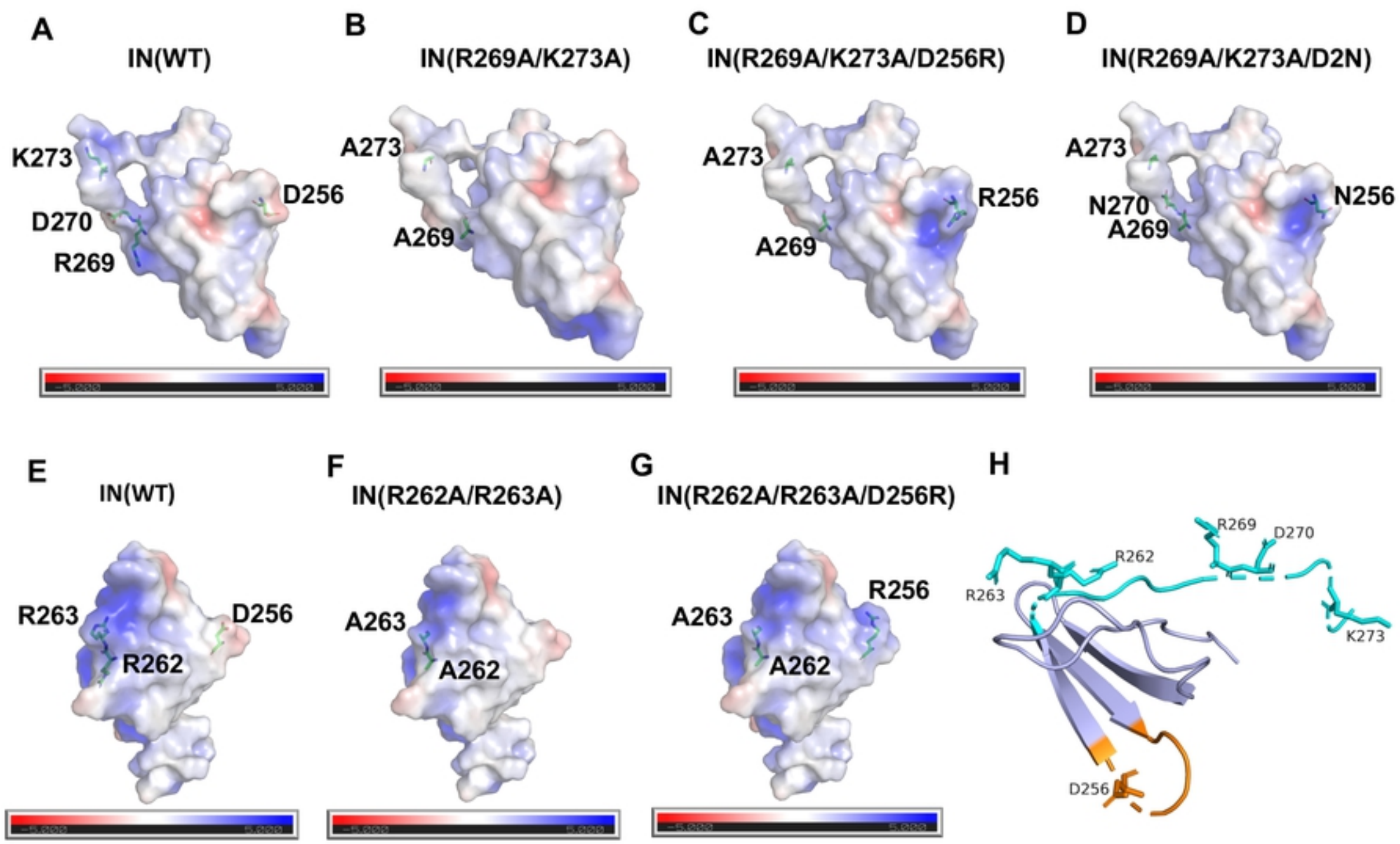
A
IN(WT)
K273
D256
D270
R269
B
IN(R269A/K273A)
A273
A269
C
IN(R269A/K273A/D256R)
A273
R256
A269
D
IN(R269A/K273A/D2N)
A273
N270
A269
N256
E
IN(WT)
R263
D256
R262
F
IN(R262A/R263A)
A263
A262
G
IN(R262A/R263A/D256R)
R256
A263
A262
H
R269
D270
R262
R263
K273
D256

Figure 5

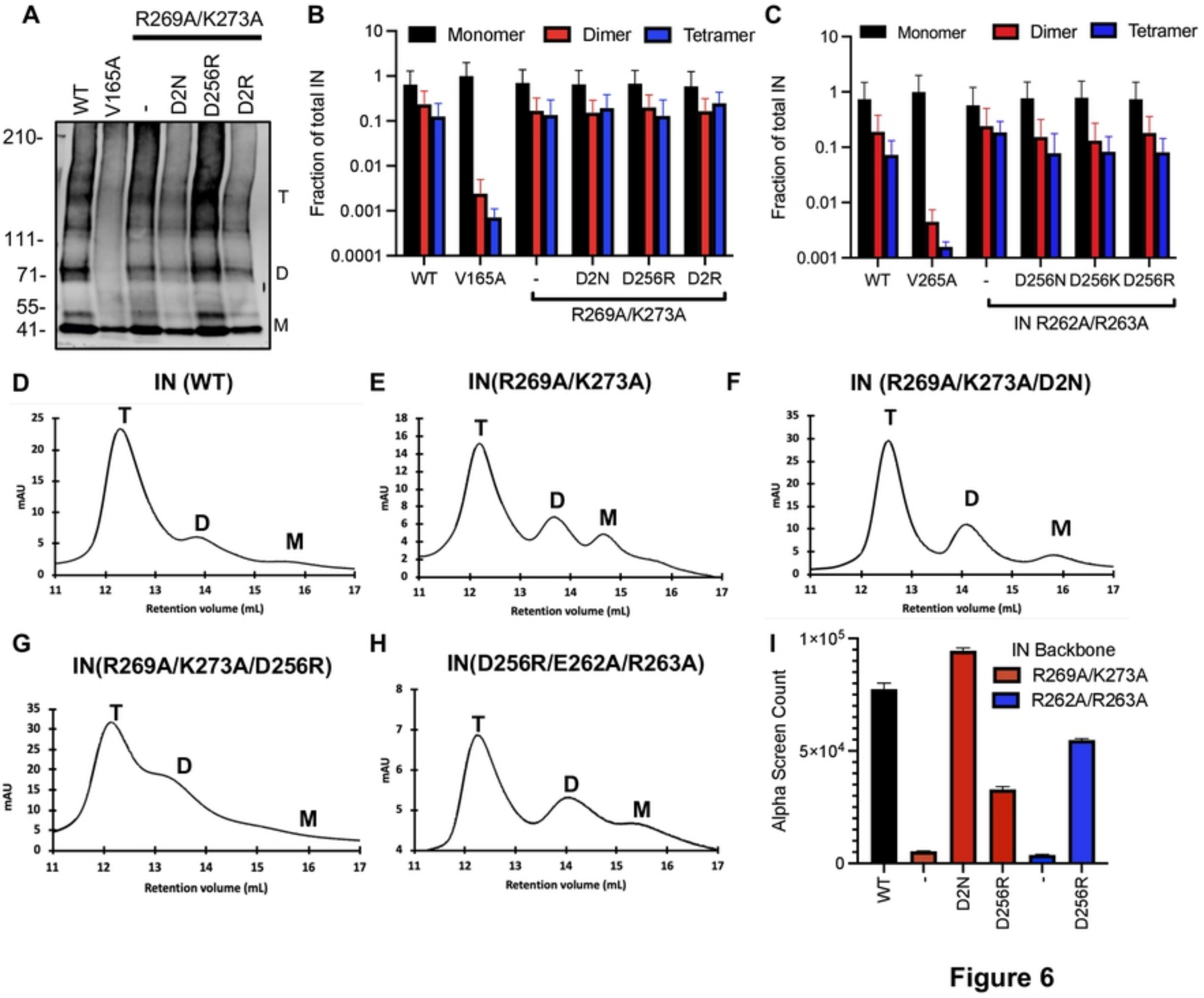
A
R269A/K273A
WT
V165A
-
D2N
D256R
D2R
210-
111-
71-
55-
41-
T
D
M
B
Monomer
Dimer
Tetramer
Fraction of total IN
WT
V165A
-
D2N
D256R
D2R
R269A/K273A
C
Monomer
Dimer
Tetramer
Fraction of total IN
WT
V265A
-
D256N
D256K
D256R
IN R262A/R263A
D
IN (WT)
E
IN(R269A/K273A)
F
IN (R269A/K273A/D2N)
G
IN(R269A/K273A/D256R)
H
IN(D256R/E262A/R263A)
mAU
Retention volume (mL)
I
IN Backbone
R269A/K273A
R262A/R263A
Alpha Screen Count
WT
-
D2N
D256R
-
D256R

Figure 6

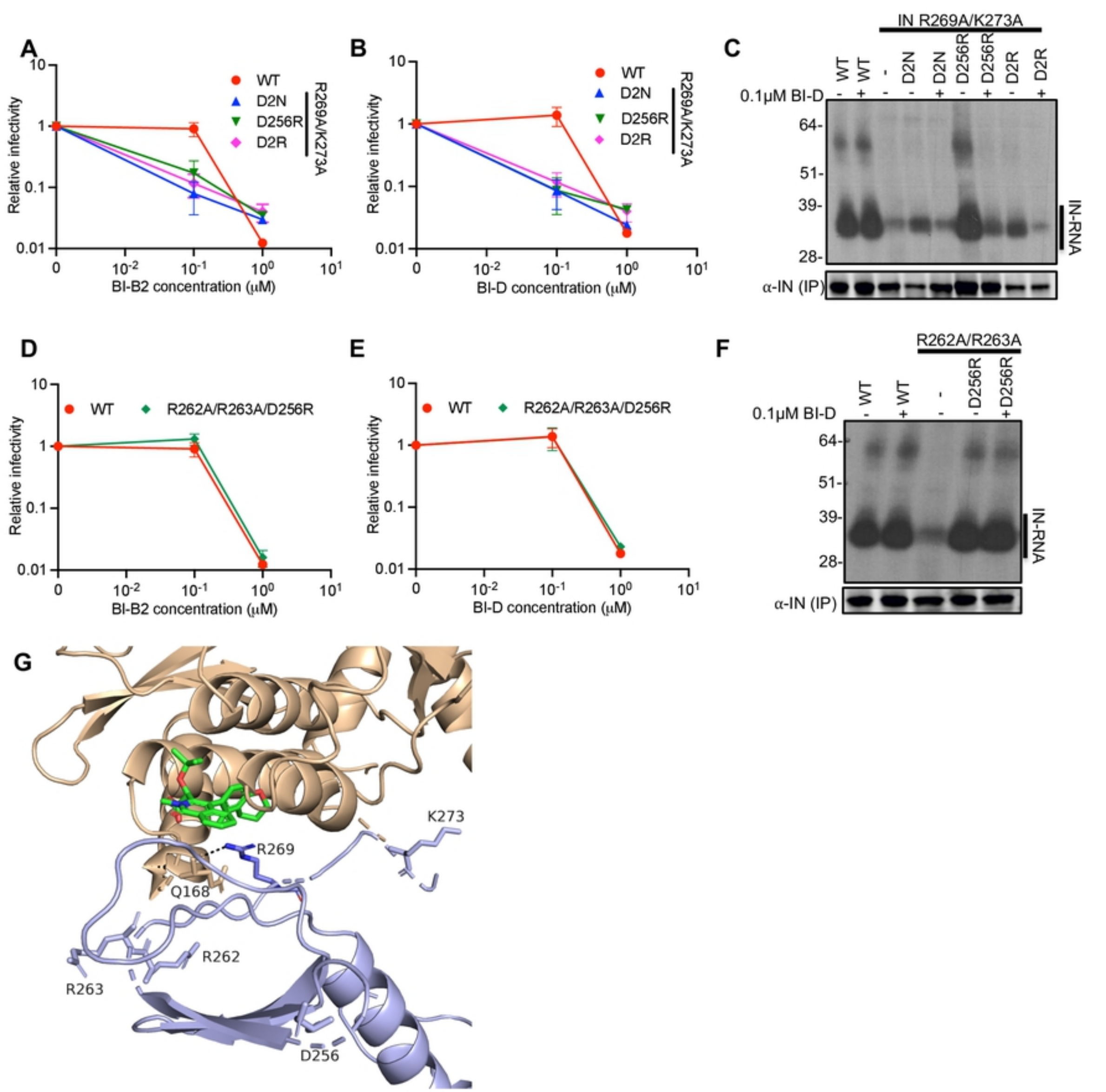
A
B
C
D
E
F
G
WT
D2N
D256R
D2R
R269A/K273A
Relative infectivity
BI-B2 concentration (μM)
BI-D concentration (μM)
IN R269A/K273A
0.1μM BI-D
IN-RNA
α-IN (IP)
R262A/R263A/D256R
R262A/R263A
K273
R269
Q168
R262
R263
D256

Figure 7

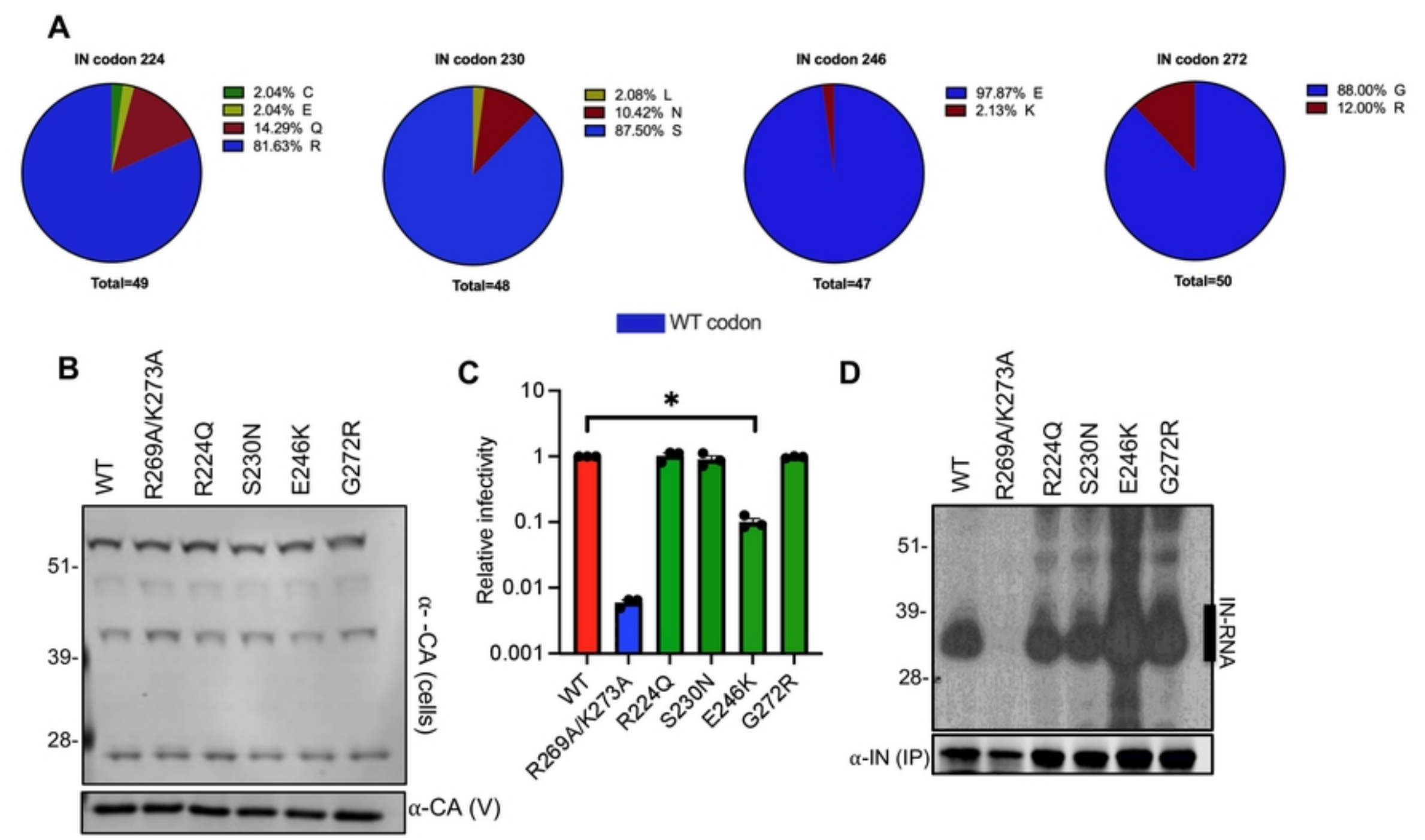
A
IN codon 224
2.04% C
2.04% E
14.29% Q
81.63% R
Total=49
IN codon 230
2.08% L
10.42% N
87.50% S
Total=48
IN codon 246
97.87% E
2.13% K
Total=47
IN codon 272
88.00% G
12.00% R
Total=50
WT codon
B
WT
R269A/K273A
R224Q
S230N
E246K
G272R
51-
39-
28-
α-CA (cells)
α-CA (V)
C
Relative infectivity
10
1
0.1
0.01
0.001
*
WT
R269A/K273A
R224Q
S230N
E246K
G272R
D
WT
R269A/K273A
R224Q
S230N
E246K
G272R
51-
39-
28-
IN-RNA
α-IN (IP)

Figure 8