PLOS | BIOLOGY

Structural Basis for Inhibitor-Induced Aggregation of HIV Integrase

Kushol Gupta[1][☯], Vesa Turkki[2][☯], Scott Sherrill-Mix[2], Young Hwang[2], Grant Eilers[2], Louis Taylor[2], Charlene McDanal[3], Ping Wang[3], David Temelkoff[3], Robert T. Nolte[4], Emile Velthuisen[3], Jerry Jeffrey[3], Gregory D. Van Duyne[1]*, Frederic D. Bushman[2]*

1 Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America, 2 Department of Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America, 3 HIV DPU, Infectious Disease Therapy Area Unit GlaxoSmithKline, Research Triangle Park, North Carolina, United States of America, 4 Protein Cellular and Structural Science, GlaxoSmithKline, Collegeville, Pennsylvania, United States of America

☯ These authors contributed equally to this work.
* vanduyne@mail.med.upenn.edu (GVD); bushman@mail.med.upenn.edu (FDB)

Abstract

The allosteric inhibitors of integrase (termed ALLINIs) interfere with HIV replication by binding to the viral-encoded integrase (IN) protein. Surprisingly, ALLINIs interfere not with DNA integration but with viral particle assembly late during HIV replication. To investigate the ALLINI inhibitory mechanism, we crystallized full-length HIV-1 IN bound to the ALLINI GSK1264 and determined the structure of the complex at 4.4 Å resolution. The structure shows GSK1264 buried between the IN C-terminal domain (CTD) and the catalytic core domain. In the crystal lattice, the interacting domains are contributed by two different dimers so that IN forms an open polymer mediated by inhibitor-bridged contacts; the N-terminal domains do not participate and are structurally disordered. Engineered amino acid substitutions at the inhibitor interface blocked ALLINI-induced multimerization. HIV escape mutants with reduced sensitivity to ALLINIs commonly altered amino acids at or near the inhibitor-bound interface, and these substitutions also diminished IN multimerization. We propose that ALLINIs inhibit particle assembly by stimulating inappropriate polymerization of IN via interactions between the catalytic core domain and the CTD and that understanding the interface involved offers new routes to inhibitor optimization.

Author Summary

A promising new class of antivirals called "ALLINIs" (allosteric inhibitors of integrase) potently inhibits HIV replication. Like other drugs, ALLINIs seem to target also the HIV-1 integrase (IN), which is crucial for the replication of this virus, but instead of acting at early phases of HIV replication, they interfere with viral particle assembly and maturation that occur at late stages and induce aggregation of IN. Despite these findings, the structural bases for the effects are still unknown. In this study, we crystallized full-length HIV-1 IN in complex with an ALLINI called GSK1264 and determined its structure to 4.4 Å.

The structure reveals for the first time the complete ALLINI-binding interface, comprised of both IN C-terminal and catalytic core domains. These domains are contributed from neighboring IN dimers, revealing an open polymeric conformation mediated by inhibitor-bridged contacts. Substitutions at this interface block ALLINI-induced multimerization, and we find that escape mutants against this class of drug lie at or near this interface. We propose that ALLINIs catalyze formation of an open IN polymer, which in turn interferes with viral particle assembly.

Introduction

Despite the success of antiretroviral therapy for HIV infection, the emergence of drug-resistant viral variants and the recognition of long-term drug toxicities leave development of new drug classes a priority [1]. Integrase strand transfer inhibitors (INSTIs) targeting the active site of the HIV-encoded integrase (IN) protein have proven highly effective [2]. An additional class of IN inhibitors, the allosteric inhibitors of integrase (ALLINIs), act at a second site on HIV IN [3–11]. ALLINIs (also referred to as LEDGINs, noncatalytic site integrase inhibitors [NCINIs], or multimodal inhibitors) are highly active against HIV replication in cell culture but have not yet been fully developed for use in patients, motivating close study to inform ongoing inhibitor development.

During the early steps of HIV infection, IN catalyzes the initial covalent attachment of the viral cDNA to host cell nuclear DNA [12,13]. IN is comprised of three independently folded domains (Fig 1A). The N-terminal domain (NTD; residues 1–50) binds Zn^{2+} via a conserved His-His-Cys-Cys (HHCC) motif. The catalytic core domain (residues 50–212) adopts an RNase H superfamily fold and contains a D,D-35-E motif that binds Mg^{2+} or Mn^{2+} ions, which mediate DNA cleaving and joining. The C-terminal domain (CTD; residues 223–268) features an Src homology domain 3 (SH3)-like fold that contributes to DNA binding and is connected to the catalytic core domain by a α-helical linker (residues 213–222).

IN function is assisted by the cellular cofactor lens epithelium-derived growth factor (LEDGF/p75), which binds tightly to a site at the catalytic core domain dimer interface [12,13]. LEDGF/p75 binding targets HIV integration to active transcription units [14–16] and promotes the integration reaction [12,13,17]. These findings and the observation that small molecules can bind this site [4,10] motivated screens for small molecules that block LEDGF/p75 binding to the catalytic core domain. Multiple compounds were identified and shown to bind the expected site on the catalytic core domain by X-ray crystallography, and many showed potent antiviral activity in cell culture [3,4,7–9,11,18,19]. However, mapping the target of ALLINI inhibition within the HIV replication cycle revealed a surprise: ALLINIs interfered only modestly with early steps of HIV replication but potently disrupted late steps, including particle assembly and maturation [6–9,18]. As a result, defective particles were produced with abnormal organization of the viral RNA and nucleocapsid (NC) proteins and greatly reduced infectivity [6,8]. It has been unclear how binding of ALLINIs to the IN catalytic core domain could interfere with particle maturation.

Here we report the crystal structure of HIV IN bound with the ALLINI GSK1264 at 4.4 Å. Prior to this study, atomic insight into the interaction of GSK1264 with IN was limited to the catalytic core domain only [7], a result which did not completely explain the mode of action of this class of compound. Here we report that the complete GSK1264 binding interface is comprised of both the CTD and the catalytic core domain, which together almost entirely bury the compound within protein. The interface is rich in residues implicated in IN oligomerization
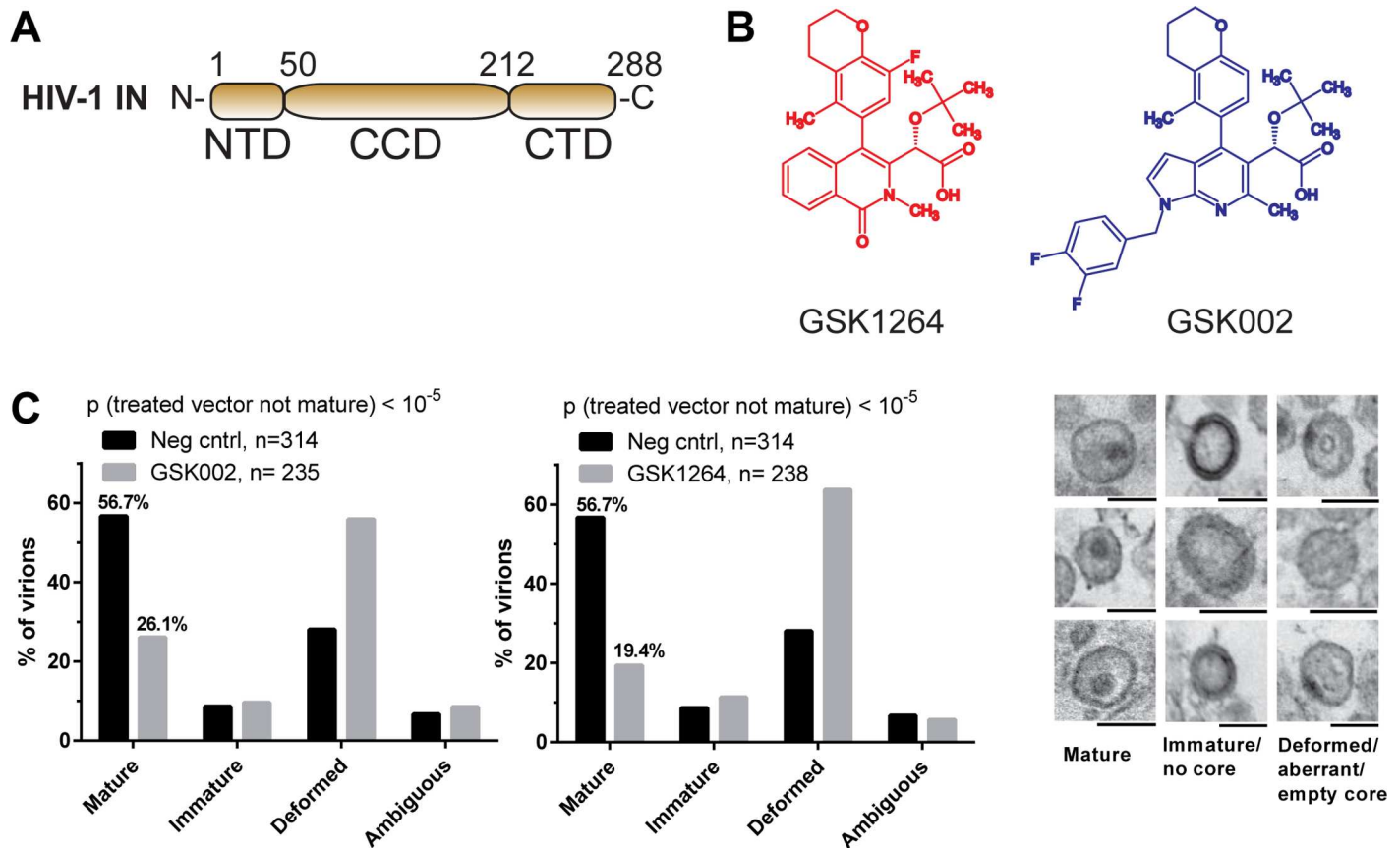
Fig 1. Allosteric inhibitors of HIV IN. (A) Domain organization of IN. (B) Chemical structures of the ALLINIs used in this study. (C) Disruption of assembly by the ALLINIs GSK1264 and GSK002. Viral particles produced in the presence of 1,000 nM GSK1264 or GSK002 were visualized by transmission electron microscopy, and morphology was scored (see S1 Data). The *p*-value is the probability of obtaining the observed (or greater) differences in numbers of nonmature particles (immature, deformed, or ambiguous) between treated and nontreated samples, given the null hypothesis of no inhibitor-induced changes.

doi:10.1371/journal.pbio.1002584.g001

and ALLINI sensitivity, indicating likely functional significance. The dimer–dimer interaction mediated by GSK1264 leads to formation of an open polymer in the crystal, a polymerization event that is readily reproduced in solution with purified components and readily attenuated by mutagenesis. To probe ALLINI function more broadly, we compared the properties of ALLINIs in biochemical, virological, and electron microscopic assays. Several IN variants, including escape mutations elicited by growth of HIV in the presence of ALLINIs, encode IN substitutions at or near the inhibitor binding site, and these substitutions also resulted in decreased IN oligomerization in vitro. The results support a mechanism in which ALLINIs disrupt viral particle maturation by promoting formation of IN polymers, as in the IN-GSK1264 crystal structure. Identification of the molecular interface responsible for polymer formation establishes a structural basis for improving the ALLINI class of HIV inhibitors.

Results

GSK1264 and GSK002 Disrupt Formation of Mature HIV Particles

To investigate the structural basis for ALLINI function, we studied GSK1264 (described in [7]) and GSK002, which is newly reported here (Fig 1B). We first examined whether these compounds disrupt HIV particle organization, as has been documented for other ALLINIs

[6,8,18], by comparing the viral particles produced from infected cells using transmission electron microscopy (Fig 1C). Inhibitor exposure significantly reduced the percentage of mature virions with proper morphology compared to untreated control cultures (100 nM and 1,000 nM doses, p-values $< 10^{-5}$). This result is consistent with the potent effect of GSK1264 late in the viral replication cycle [7] and parallels results for other ALLINIs [6–9,18].

Structure of an HIV IN•GSK1264 Complex

To crystallize HIV IN bound to an ALLINI, we sought substitutions that improved IN solubility but preserved some degree of ALLINI-dependent IN aggregation, allowing control of IN self-association. We used size-exclusion chromatography in line with multiangle light scattering (SEC-MALS) and analytical ultracentrifugation (AUC) to establish the solution properties of several IN variants. Our best candidate contained two amino acid substitutions: Y15A and F185H. $IN^{Y15A,F185H}$ formed discrete dimers in solution, retained the ability to bind to the LEDGF integrase binding domain (IBD), and had a reduced level of aggregation in the presence of GSK1264 and GSK002 compared to wild-type IN (S1 Fig and S1 Table). Substitutions of F185 are known to improve solubility [20], and viruses containing the F185H substitution are replication competent [21]. The Y15A substitution selects for a single conformation of the isolated NTD [22] that is correlated with diminished oligomerization of full-length IN [23]. We favored NTD substitutions for this study because IN constructs containing only the catalytic core domain and the CTD were previously shown to be sufficient for ALLINI-induced aggregation [7,24], suggesting that NTD substitutions would not interfere with understanding ALLINI action. HIV-1 containing IN^{Y15A} is replication defective, and the purified protein is not active for integration.

Crystallization trials of $IN^{Y15A,F185H}$ with GSK1264 led to a diffracting crystal form containing one IN dimer in the asymmetric unit; trials with GSK002 did not yield crystals. We determined the structure by molecular replacement using the catalytic core domain dimer structure (without ligand) from the GSK1264-bound structure previously determined (PDB 4OJR, [7]) together with the previously reported IN CTD structure [7,25]. The structure was refined using deformable elastic network (DEN) restraints [26–28]. The catalytic core domains, CTDs, and connecting linkers displayed strong electron density in simulated annealing $2F_o$-F_c composite omit maps, and thus, their positions are well defined in the structure (Fig 2A). The IN NTDs could not be positioned by molecular replacement and did not display interpretable electron density. However, IN recovered from washed crystals was full length by SDS-PAGE (S1 Fig). Hence, the NTDs are presumed to be spatially disordered. Strong difference density in simulated annealing omit maps was observed at both previously observed GSK1264 binding sites on the catalytic core domain dimer [7], allowing for placement of the ligand in the latest stages of refinement (Fig 2B). At the modest resolution available for this crystal form, we were able to reliably place domains, secondary structure elements, and the bound inhibitor, but we could not visualize detailed side chain interactions. We inferred the approximate positions of side chains from the electron density and their known locations in the high-resolution crystal structures of smaller IN fragments [7,29]. Diffraction data and refinement results are provided in S2 Table.

The GSK1264-bound HIV IN structure is shown in Fig 2C. The IN CTDs bind to the catalytic core domains and bound inhibitors in adjacent IN dimers, resulting in the formation of an open polymer of IN dimers. The interaction site resembles a hand (the CTD) grasping a ball (the catalytic core domain), with the ALLINI site cupped by the fingers (Fig 2D).

The GSK1264–IN Binding Interface

In the IN•GSK1264 complex, GSK1264 is almost entirely protected from solvent, with 500 out of 580 Å^2 of its accessible surface buried (Fig 3A–3C). Excluding bound inhibitor, over 1,100
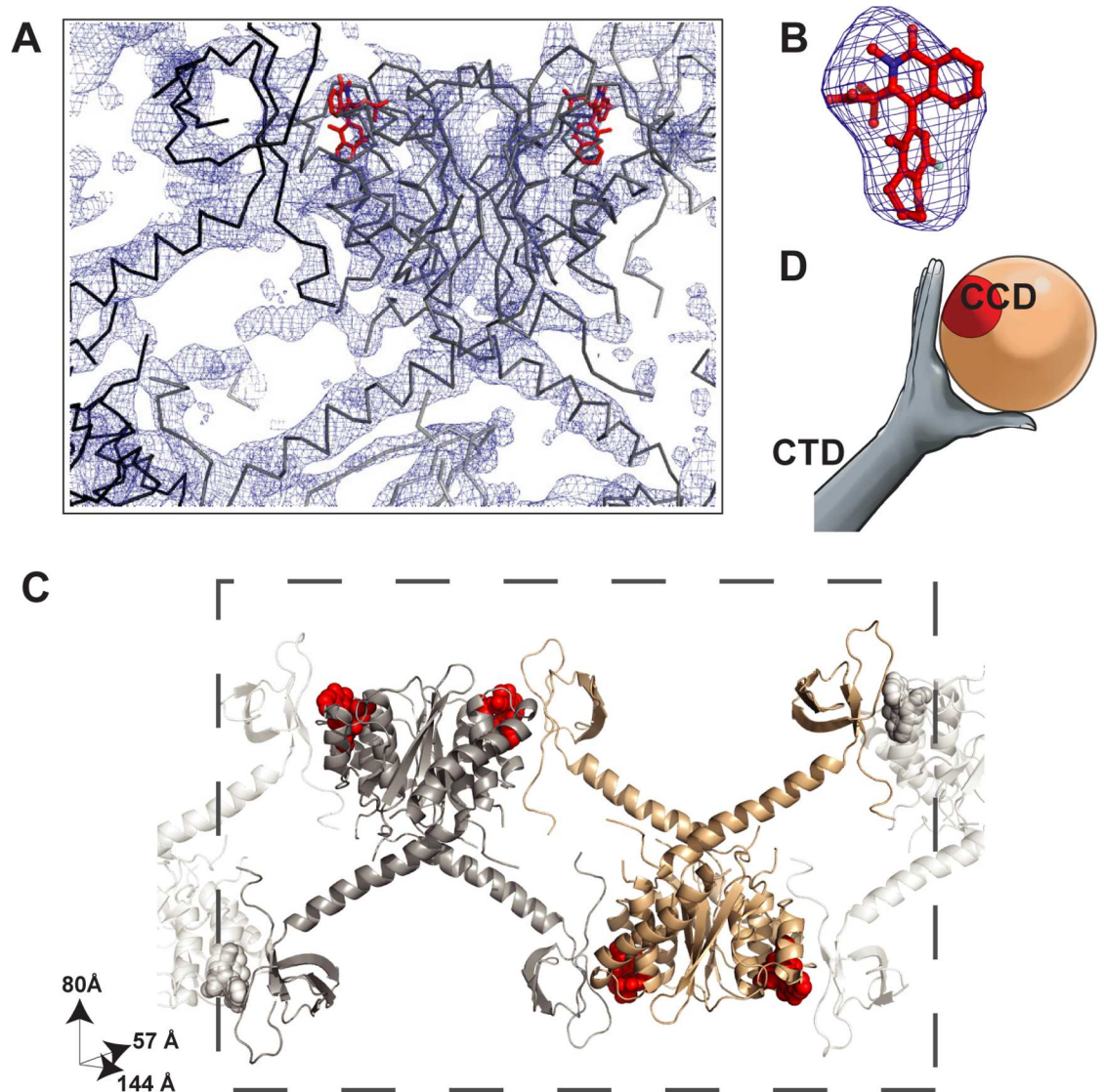
Fig 2. Crystal structure of HIV-1 IN^{Y15A,F185H}•GSK1264 complex. (A) Shown is a simulated annealing composite-omit $2F_o$-F_c electron density map contoured to 1.5 σ (shown in blue). Contiguous electron density is observed for the catalytic core domains, the CTDs, and GSK1264 (red), but the NTDs are disordered. (B) Simulated annealing F_o-F_c map density for GSK1264, contoured to 3σ (shown in blue). (C) The ALLINI-bound HIV-1 IN polymer observed in crystals. One dimer of HIV-1 IN comprises the asymmetric unit. Bound GSK1264 is shown by the red spheres, two interacting dimers are highlighted in the dotted box, and additional subunits of the open polymer are shown in light grey. (D) Cartoon of the CTD–catalytic core domain interaction, where the hand represents the CTD, the ball represents the catalytic core domain dimer, and GSK1264 is shown in red. See also S2 Fig and S2 Table.

doi:10.1371/journal.pbio.1002584.g002

Å^2 of solvent-accessible surface is buried at each CTD–catalytic core domain interface. Secondary structure elements mediating GSK1264 binding include one face of the CTD, defined by strands β1, β2, and β5, juxtaposed against the catalytic core domain dimer near helix α3 of one subunit and α4 of the other. The CTD–catalytic core domain interface has a strong electrostatic component, with the relatively acidic α4 helix of the catalytic core domain packing against the basic β5 strand of the CTD (S2 Fig). Extensive hydrophobic interactions are also formed between α3 of the catalytic core domain and the β1 and β2 strands of the CTD.
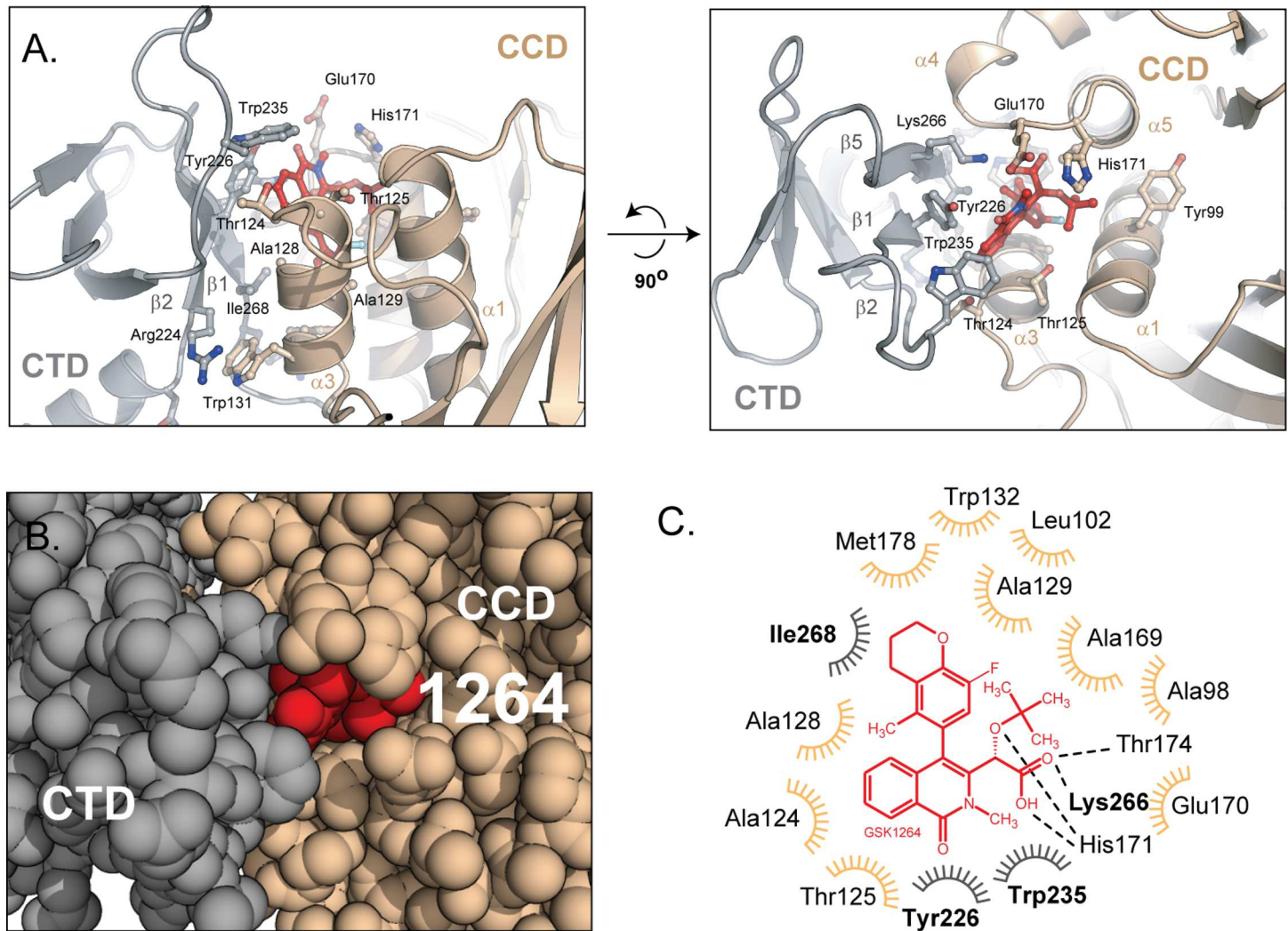
Fig 3. The GSK1264 binding interface. (A) Orthogonal views of the GSK1264-induced polymer interface. GSK1264 is shown in red, packed between the CTD (grey) and the catalytic core domain dimer (tan). The positions of the side chains rendered are inferred from the high-resolution crystal structures used in the DEN refinement procedure. (B) Sphere rendering of the protein–ALLINI-protein interface. Colored as in A. (C) Schematic of IN-GSK1264 contacts. GSK1264 (red) is predominantly buried via van der Waals contacts with 13 residues from the CTD (grey) and catalytic core domain (tan). Thr[174], Lys[266], and His[171] are predicted to hydrogen bond with the *tert*-butoxy and carboxylic acid moieties. This panel was generated using LIGPLOT [30]. See also S2 Fig, S3 Fig and S1 Table.

The modest resolution of our structural model precludes direct visualization of interactions but does allow us to model the approximate positions of functional groups and potential contacts mediating GSK1264 binding (Fig 3A and 3C). The Tyr[226] and Trp[235] side chains in the CTD pack against the core isoquinoline moiety, where π–π interactions between inhibitor and aromatic side chains may contribute to binding stabilization. The *tert*-butoxy moiety is buried in a hydrophobic pocket formed by the catalytic core domain dimer, without contributions from the CTD. The benzopyran group extending from the isoquinoline core is surrounded by a cradle of hydrophobics at the catalytic core domain dimer interface and, additionally, Ile[268] from the CTD. A modeled hydrogen-bonding network involving the carboxylate and *tert*-butoxy oxygen completes the IN-inhibitor interface, where Lys[266] from the CTD is positioned

to coordinate the carboxylate moiety. S2C Fig shows a surface model of interactions between each domain and GSK1264.

Functional Analysis of the ALLINI-Binding Interface

To probe the importance of the GSK1264 interface observed in this crystal form, we compared the IN aggregation properties and structural features of several compounds bound at ALLINI sites (Fig 4). Previously, we and others have reported [7,24] that truncated IN derivatives containing only the catalytic core domain and the CTD were sufficient to support aggregation by ALLINIs in vitro. In addition, crystal structures are available for the IN catalytic core domain bound to several ALLINIs and other small molecules (Fig 4A–4D). Structures of the GSK1264 [7], BI-D [8], and tetraphenylarsonium (TPA) [10] complexes were reported previously; the structure of GSK002 bound to the catalytic core domain at 1.75 Å is newly reported here (Fig 4C, S2 Table, and S4 Fig). We found that GSK1264, GSK002, and BI-D all directed potent aggregation of IN containing the catalytic core domain and the CTD, but not IN derivatives containing the NTD and the catalytic core domain or only the catalytic core domain (Fig 4). TPA, which is not an ALLINI, did not promote aggregation of any derivative even at millimolar concentrations (Fig 4H).

Analysis of the TPA-bound catalytic core domain structure superposed onto the GSK1264-bound full-length IN structure indicates that TPA would clash with the CTD and therefore block the CTD–catalytic core domain interaction seen for GSK1264 (arrows in Fig 4G). A small steric clash is also predicted for GSK002, in which the difluorobenzyl group collides with Trp^{235}. However, rotation of the Trp^{235} side chain could create enough space for the inhibitor, so we suggest that GSK002 could be accommodated in a nearly identical interface. Thus, this analysis supports the concept that CTD–ALLINI–catalytic core domain interactions are relevant to aggregation in vitro and that the shape and positioning of the ALLINI molecule is critical for the aggregation reaction.

CTD Substitutions That Block ALLINI-Induced Aggregation

IN substitutions that diminish ALLINI function have been studied on the catalytic domain side of the interface, but interactions on the CTD side are less well characterized. Biochemical studies have implicated the CTD in IN oligomerization [7,20,31,32], and CTD residues K264 and K266, when both mutated to alanine, diminished ALLINI-induced aggregation and blocked integration activity [24]. In the structure presented here, these residues contribute to the GSK1264 interface (Fig 5A), supporting the biological relevance of the observed contacts. We further probed the importance of this interface by studying substitutions at CTD residues Y226 and W235 (S1 Table, Fig 5 and S5 Fig), which are large hydrophobics comprising a substantial portion of the surface buried. Substitutions at Y226 or W235 inhibited precipitation by GSK1264 and GSK002 (Fig 5B) and abrogated IN activity (S5 Fig). Three substitutions at Y226 each resulted in IN derivatives that showed diminished oligomerization in vitro by both AUC (Fig 5C) and SEC-MALS (Fig 5D). Purified IN W235A also showed diminished oligomerization by SEC-MALS (Fig 5D) but was too unstable for sedimentation equilibrium analysis. In summary, four CTD residues in the ALLINI interface are required for ALLINI-induced aggregation of IN, supporting the biological importance of the interface shown in Fig 2, Fig 3 and Fig 5A.

Mechanism of GSK1264 and GSK002 Resistance Mutations

To identify mutations that confer resistance to GSK1264 and GSK002, we passaged HIV-1 in the presence of increasing concentrations of each inhibitor and isolated resistant viral strains.

Models of CCD-CTD Interface
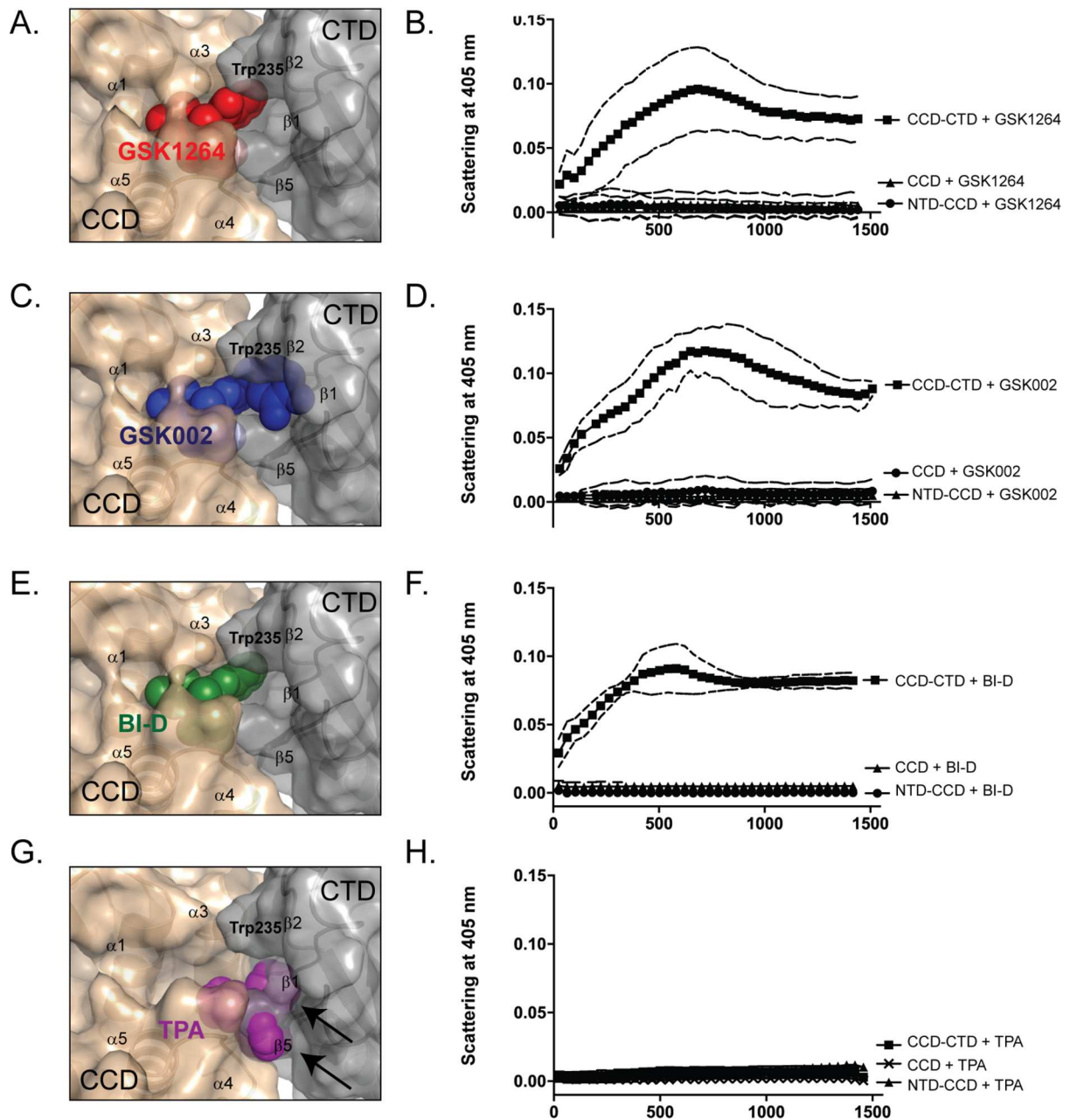
Fig 4. Experimental probing of the GSK1264 interface. Models of the catalytic domain–CTD interface with bound small molecules (A, C, E, and G) are shown beside aggregation time course assays for the same compounds (B, D, F, and H). Aggregation assays contained one and two-domain fragments of IN^{F185H} as indicated ("CCD" denotes catalytic core domain). Compounds studied are GSK1264 (A, B; red, PDB 4OJR [7]), GSK002 (C,D; blue, PDB 5HRN [this work]), BI-D (E, F; green, PDB 4ID1 [8]), and tetraphenylarsonium (TPA; G, H; magenta, PDB 1HYV [10]). Models were generated by docking the indicated CCD/ compound structures into the GSK1264/ IN^{F185H} structure studied here. Aggregation assays were carried out using light scattering at 405 nm at 25°C with 10 μM IN (see S1 Data). In panel G, arrows indicate steric clashes predicted by modeling of the CTD–catalytic core domain interface with TPA, based on the TPA binding mode in PDB 1HYV [10].

doi:10.1371/journal.pbio.1002584.g004

Three different HIV-1 strains were compared for each inhibitor: the lab-adapted HIV strain NL4-3 and Raltegravir-resistant strains 4376 and 8070. Strains were passaged 38 times in the

Fig 5. CTD substitutions affect IN oligomerization and ALLINI-induced aggregation. (A) CTD residues (shown in blue) at the IN^{F185H}–ALLINI interface mapped onto the IN^{F185H}–GSK1264 structure. GSK1264 is shown in red. Mutations at K264 and K266 that affect ALLINI function were previously reported [24]; the e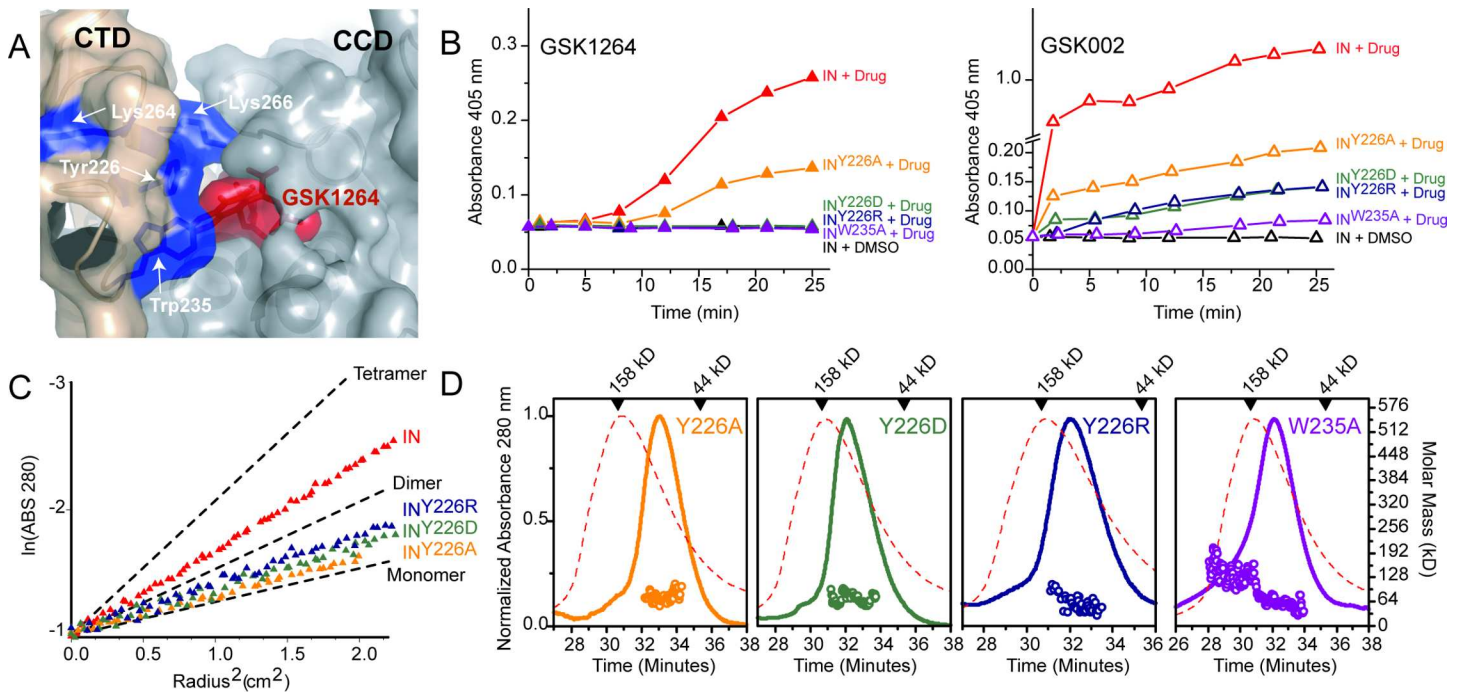ffects of mutations at Y226 and W235 are reported here. (B) Aggregation assays. The time-dependent aggregation of IN by GSK1264 and GSK002 was monitored using absorbance optics. To initiate the reaction, either DMSO or drug was added to recombinant IN^{F185H} for final concentrations of ~40–61 μM IN^{F185H} monomer and 44 μM drug at room temperature. For each panel in this figure, "IN" denotes "IN^{F185H}". (C) Analytical ultracentrifugation sedimentation equilibrium analysis. Data were recorded at 12,000 RPM, at a concentration of ~10 μM IN^{F185H}, at 4°C. Linearized radial distributions are shown. The slopes are proportional to M_w at a given value of r^2. Single-species plots with calculated slopes for idealized IN^{F185H} monomer, dimer, and tetramer are shown for the same rotor speed and temperature as black lines. (D) SEC-MALS analysis. The IN^{F185H} SEC trace is superimposed on each panel in the dotted red line. Retention times for molecular weight markers are shown at the top of each panel, and retention times for globular molecular weight standards are shown as open circles. Elution concentrations by refractive index approached ~0.1 mg/mL. Experiments were performed at room temperature using a Superdex 200 10/300 column. Data shown in panels B–D are provided in S1 Data.

doi:10.1371/journal.pbio.1002584.g005

presence of compound, resulting in an increase in the inhibitory IC$_{50}$ of >100-fold. IN coding regions were sequenced, and departures from wild-type sequence tabulated (S3 Table).

Each of the three viral strains studied showed mostly distinct escape mutations that were unique to each inhibitor (Fig 6). Multiple independent viral cultures were not compared for each ALLINI, so it is unknown whether the screens were saturated. Some of the substitutions match previously published ALLINI resistance mutations [3–5,9,11,18,19,33–35]. The IN•GSK1264 structure allows us to consider the mechanism by which these escape mutations confer resistance.

In studies with GSK1264, six different substitutions were found (Fig 6A and S3 Table). Ala^{129} and Thr^{174} are positioned to contact GSK1264 directly, so the larger side chains seen in escape mutants (A129T and T174I) may disrupt inhibitor binding via steric clashes. Y99H and L172F do not contact bound GSK1264 but may disrupt inhibitor binding via changes in occupied volumes within the catalytic core domain, resulting in displacement of α1 (Y99H) or α5 (L172F). Ala^{205} and Asn^{222} lie at the beginning and the end of the α-helical linker that bridges the catalytic core domain and the CTD of IN, so substitutions may alter the CTD position relative to the catalytic domain, thereby potentially inhibiting formation of the ALLINI complex (Fig 6A and 6C).
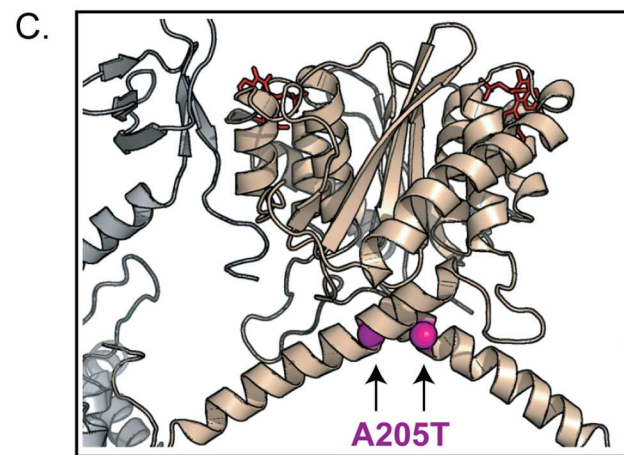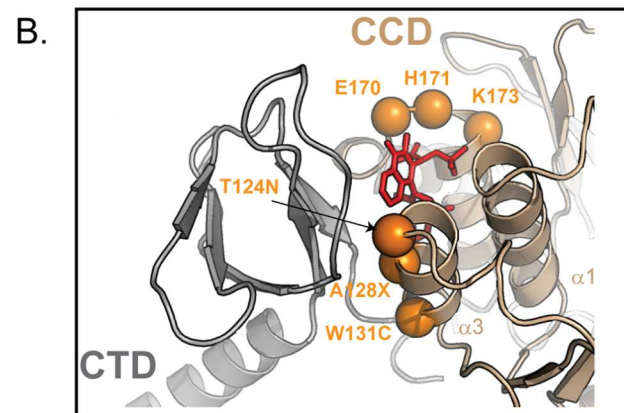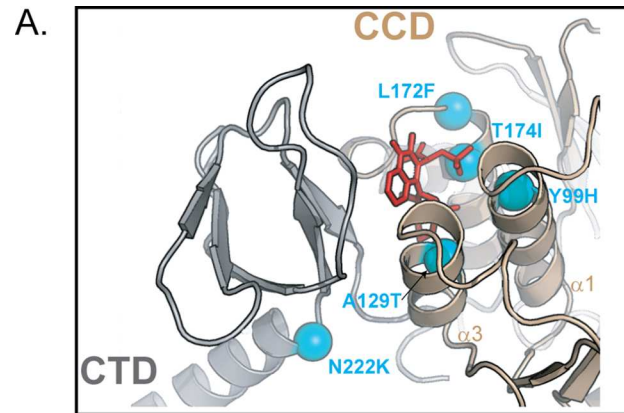
Fig 6. HIV IN mutations conferring reduced sensitivity to GSK1264 and GSK002 after serial passage.
(A) Amino acid substitutions resulting from resistance mutations identified by serial passage in the presence of GSK1264 (blue spheres). (B) Substitutions identified in the presence of GSK 002 (orange spheres). Bound GSK1264 is shown in red. (C) The A205T resistance substitutions, which arose in both experiments. S4 Fig provide structural insights into the mechanism of resistance conferred by polymorphisms at residues 124 and 125.

doi:10.1371/journal.pbio.1002584.g006

Eight mutations arose after passage in the presence of GSK002, only one of which was in common with the GSK1264 mutations (Fig 6 and S3 Table). Using the GSK1264 structure as a model, we can propose mechanisms of GSK002 resistance for several of these. IN residues 125, 128, 170, 171, and 173 are poised to contact the bound inhibitor and/or the CTD, so substitutions at these positions likely disrupt the inhibitor-mediated interface directly. A205T was seen for both GSK002 and GSK1264 resistance substitutions (described above). T124N was the only substitution to arise independently in two experiments (S3 Table), and substitutions at this position and at residue 128 have been reported as ALLINI resistance mutations previously [3–5,11,19,33–36].

Ala^{124} and Trp^{131} are the first and last residues of the α3 helix; they contribute to the oligomerization interface between the catalytic core domain and the CTD, so substitution of these residues is likely to interfere with inhibitor-mediated oligomerization. The resistance mutation encoding W131C was identified in the earliest passages and is unique to this study (Fig 6B and S3 Table). W131D-substituted IN has been reported previously to increase solubility of purified IN [29,37,38]. Trp^{131} is largely buried in the CTD interface in the IN•GSK1264 structure reported here (S3 Fig). Trp^{131} does not contribute to the ALLINI binding interface directly, but the IN^{W131D} substitution shows reduced oligomerization and reduced sensitivity to ALLINI-induced aggregation (S3 Fig and S1 Table), supporting the idea that the Trp^{131}–CTD interaction is important for the inhibitor response.
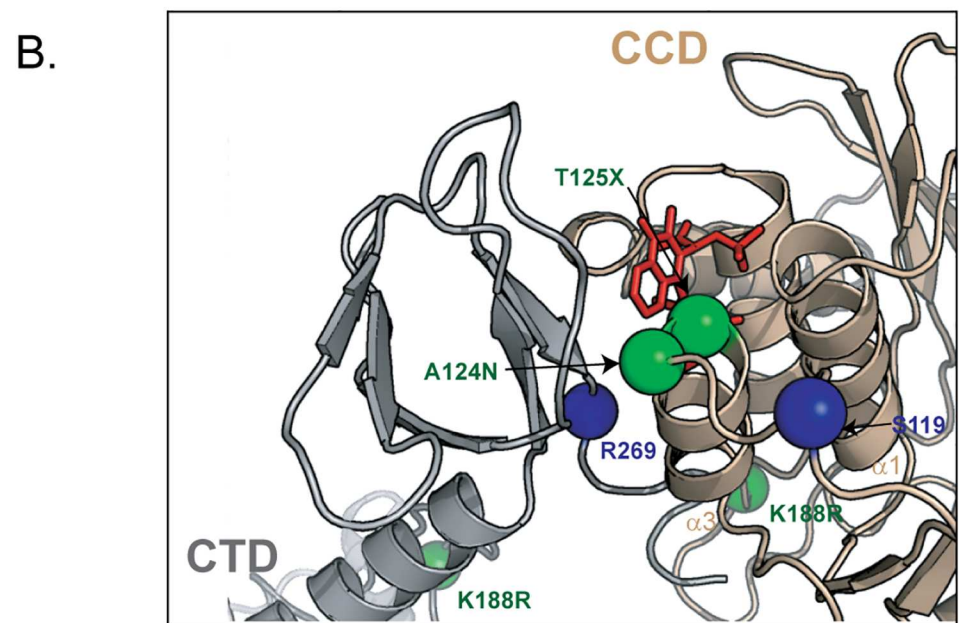
To assess cross-resistance between GSK1264 and GSK002, we examined the sensitivities of two long-term passaged variants. Thirty-eight passages in increasing concentrations of GSK002 yielded a variant with T124N, W131C, and A205T; 38 passages in GSK1264 yielded a variant with Y99H, L172F, and N222K. In both cases, these multiply mutant variants were highly resistant to both drugs (>550-fold increase in IC$_{50}$).

To investigate the mechanism of ALLINI resistance in vitro, we purified six IN variants with escape mutant substitutions and tested their oligomeric properties using sedimentation equilibrium. Each of the escape variants showed reduced oligomerization compared to IN^{F185H}, suggesting that diminished multimerization is a common characteristic of ALLINI escape substitutions (S1 Table).

We also purified IN variants containing resistance substitutions along the α3 helix and tested their ability to be aggregated by GSK1264 or GSK002 (S1 Table). For IN^{F185H}, 0.8 μM of GSK002 or 8 μM of GSK1264 was sufficient to promote aggregation. The T124N, A128T, and W131C substitutions each protected IN from aggregation by either compound at 67 μM. The T124A substitution protected IN from polymerization at 67 μM GSK1264, but not GSK002. These findings strengthen the link between HIV escape from ALLINI pressure in cell culture and diminished inhibitor-induced polymerization of IN in vitro.

Sequence Polymorphisms Affecting Sensitivity to ALLINIs

To characterize naturally occuring IN variants that affect ALLINI activity, we tested a panel of 38 HIV isolates for sensitivity to GSK1264 and GSK002. Viruses of subtypes A, B, C, D, F, and A/E were compared, thereby interrogating a wide range of primary IN amino acid sequences. For GSK1264, the IC$_{50}$ values ranged from 2–100 μM (Fig 7A). In contrast, IC$_{50}$ values

Fig 7. Sensitivity of HIV strains to GSK1264 and GSK002. (A) Distributions of IC$_{50}$ values for GSK1264 and GSK002 in infections of multiple subtypes of HIV, colored by subtype and sorted by small molecule. The
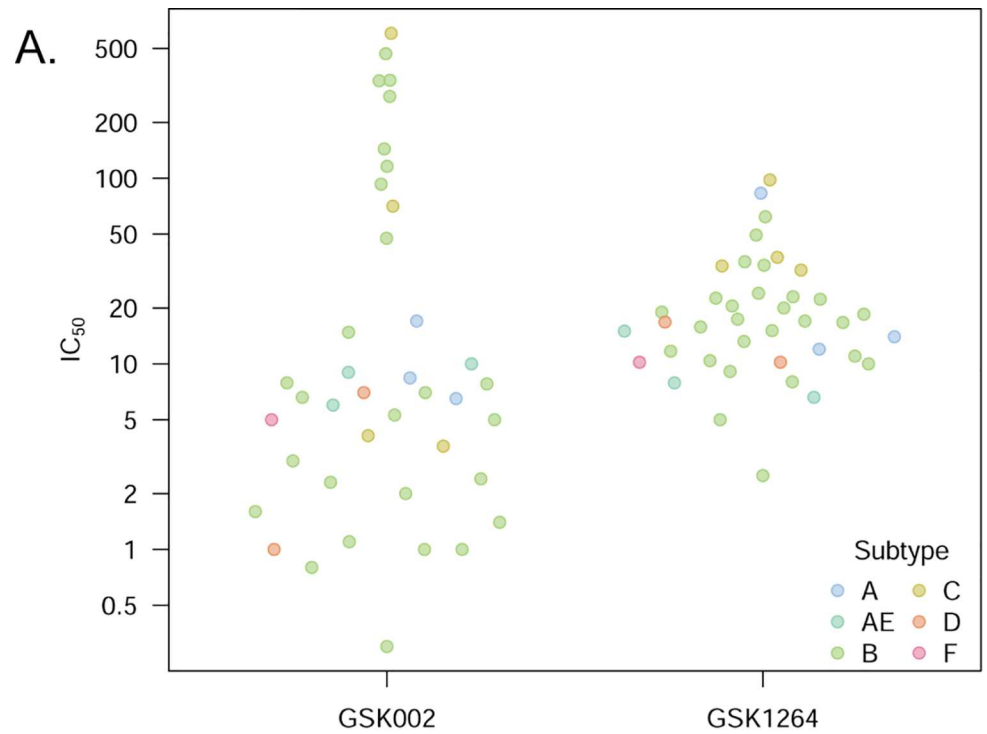
y-axis shows IC$_{50}$ values. Data are provided in tabular form in S1 Data. (B) Polymorphisms identified by lasso logistic regression linked to resistance to GSK1264 (blue spheres) or GSK002 (green spheres) mapped onto the IN-GSK1264 structure. GSK1264 is shown in red. See S3 Table.

doi:10.1371/journal.pbio.1002584.g007

determined for GSK002 against the same panel ranged from less than 0.5 μM to greater than 500 μM (Fig 7A). There was little correlation among IC$_{50}$ values for the two inhibitors ($R = 0.0492$, $R^2 = 0.0024$, $p = 0.811$), emphasizing the differential effects IN sequence variation can have on the potency of ALLINIs.

The amino-acid polymorphisms best able to predict IC$_{50}$ were extracted from the data for each inhibitor using lasso logistic regression [39]. A model predicting GSK002 activity using the top four most influential amino acid substitutions (K20R, A124N, T125X, where $X \neq T$, and K188R) performed well under cross validation, where the cross-validated mean-squared error decreased from 8.49 to 2.26. The model for GSK1264 showed weak predictive power, with no mean-squared error greater than one standard error away from random. Substitutions at Lys[20], Ala[23], Ser[119], and Arg[269], selected earliest by lasso logistic regression, may have the potential to influence IC$_{50}$ for GSK1264, though less robustly than for residues affecting GSK002.

Positions of affected residues near the inhibitor interface are shown on the IN•GSK1264 structure in Fig 7B. Substitutions at residue 124 could affect ALLINI binding or the catalytic core domain–CTD interaction. Substitutions at 125 may also affect ALLINI binding. Arg[269] lies on the CTD at the interface with the IN catalytic core domain adjacent to the ALLINI binding site, where substitutions could disrupt this interaction. We consider the possible roles of additional substitutions, including those in the NTD (K20R, A23X), in the Discussion.

Probing the Mechanism of Resistance to GSK002

To investigate the mechanism of GSK002 resistance further, we determined three high-resolution crystal structures, one of GSK002 bound to catalytic core domain-only, one of GSK002 bound to the catalytic core domain containing the Ala[124] substitution, and one of GSK002 bound to the catalytic core domain with Thr[125]. These two substitutions were chosen for study because they were the most influential in reducing GSK002 IC$_{50}$ in the study of diverse HIV isolates (Fig 7, S3 Table and S4 Fig). The protein structures were quite similar, with RMSDs ranging from 0.09 to 0.13 Å in pairwise comparisons (S4 Fig). However, the azaindole core of the bound GSK002 ligand is seen to shift in position by ~0.5 Å, depending on the identity of the substitutions at residue 125, along with movement of the difluorobenzyl group by almost 1 Å (S4A–S4C Fig). Overlay of the catalytic core domain•GSK002 structures on the IN•GSK 1264 structure suggests a further mechanism of resistance. Larger substitutions at residue 124 and 125 would create a steric clash with the polypeptide backbone of the IN CTD, disrupting the inhibitor-mediated interaction (S4D Fig). Thus, the larger side chains in the resistance substitutions appear to either clash directly with the IN CTD or displace the bound inhibitor sufficiently to result in an inhibitor clash with the CTD.

Discussion

Here we have identified the determinants of IN aggregation mediated by ALLINIs and obtained a structural model of the specific interfaces involved. In our model, IN polymers are formed by catalytic core domain–CTD interactions that are facilitated and strengthened by the bound ALLINI. We propose that ALLINIs act by causing IN to form extended polymers mediated by this interaction, which in turn leads to aggregation and inactivation. This mode of inhibition is reminiscent of microtubule-stabilizing agents such as Taxol that stimulate lateral

interactions between oligomers that modulate assembly [40,41]. Our model is supported by the properties of IN mutants that are resistant to ALLINIs, including mutants elicited during HIV passage in the presence of ALLINIs, engineered INs, and naturally occurring IN polymorphisms in circulating HIV strains. Our model is also supported by a strong correlation between in vivo inhibition and susceptibility to polymerization of purified IN proteins in vitro.

Why was the catalytic core domain–CTD interaction seen here not observed in previous experimental structures of HIV IN? We suggest that the answer lies with the use of (3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate) (CHAPS), the zwitterionic detergent used for decades to solubilize IN for in vitro experiments [29,42]. CHAPS binds to the same CTD surface used to form ALLINI-mediated polymers, providing an explanation for how the detergent prevents aggregation and for why previous structural studies would not have revealed the intermolecular interfaces reported here (S2 Fig).

How does aggregation of IN disrupt viral assembly? ALLINIs induce aggregation of IN in vitro, and escape mutations reverse this, implicating aggregation in the mechanism of inhibition. During assembly, IN is first synthesized as part of the Gag-Pol polyprotein precursor. After encapsidation, Gag-Pol is cleaved into component proteins, including IN [12,13]. Immature virions containing Gag-Pol formed in the presence of ALLINIs appear normal in electron microscopy (EM) images, but mature particles are abnormal, suggesting that IN, and not Gag-Pol, may be the ALLINI target. This model is consistent with studies in which ALLINIs were effective when IN was delivered to viral particles independently of Gag-Pol [6,8]. HIV particles can bud properly when Gag only is present to yield immature particles, but Gag-PolΔIN forms defective particles during maturation [6]. The defective particles have electron-dense material outside the normal core, as is seen after ALLINI treatment, reportedly containing NC and RNA [6]. Thus, the IN polymer seen here could act by either (1) sequestering IN and preventing a normal function important for localizing NC and the RNA during maturation or (2) creating a novel aggregate that is itself disruptive.

IN substitutions that confer resistance to ALLINIs in virus fall into three categories, the first of which likely interfere directly with inhibitor binding (e.g., A128T, E170K, and H171T). The second group perturb the catalytic core domain–CTD interaction that mediates polymer formation (e.g., A124N). These substitutions can lie at the interface between domains (e.g., K266A and W131C/D), in a "second shell" that flanks the binding site (e.g., Y99H and L192F), or in the helical connector between domains that is required for positioning the CTD (e.g., A205T and N222K). While mutations at 226, 235, 264, and 266 all confer resistance to ALLINI-induced aggregation in vitro, no strong escape mutants at the CTD side of this interface occur naturally. However, most of the residues on the CTD that are in contact with GSK1264 are absolutely conserved among HIV variants, suggesting that substitutions that block ALLINI binding are likely to be replication defective. The third group of ALLINI-resistance substitutions resides within the NTD (e.g., D6R, E11K, and Y15A). These substitutions are located on a face of the NTD that has been observed bound to the catalytic core domain in several reported crystal structures [23,38,43,44]. At present, we favor a model in which binding of the NTD to the catalytic core domain competes with binding of the CTD. Enhanced binding of the NTD as a result of amino acid substitutions could inhibit ALLINI-promoted engagement of the CTD at the catalytic core domain. Further work will be required to understand how these NTD substitutions modulate ALLINI action.

The results presented here suggest several directions for future research. We were not able to crystallize the ALLINI GSK002 with IN, and the spectrum of resistance mutations elicited is mostly different between the two inhibitors. We speculate that the IN polymer formed in the presence of GSK002 is similar to that seen for GSK1264, with side chains remodeled to accommodate the interaction, but how these small structural differences influence inhibitor potency

and escape mechanisms is just beginning to be studied. The role of the NTD in inhibition and escape is not yet understood but should also be experimentally accessible. Perhaps most importantly, our data suggest specific means of optimizing ALLINI binding in the observed interdomain pocket and specify regions of IN where additional functional groups may potentially be attached without disrupting binding. Thus, these results should be useful in optimizing the function of ALLINIs, a unique new class of HIV inhibitors that block viral replication by promoting inappropriate polymerization.

Materials and Methods

Ethics Statement

The human biological samples were sourced ethically from Gulf Coast Regional Blood Center (GCRBC). The protocol titled : IRB Specimen Procurement Protocol, Specimens to Provide to Internal or External Customers, was reviewed and approved by the GCRBC IRB (IRB approval # 06-001). The research conducted was in accord with the terms of the informed consents.

Small Molecule Inhibitors

The compound GSK1264 was previously described [7] and is referred to as compound 159 in Patent WO 2012/102985 (Compound 159); GSK002 is described in Patent WO2013/012649 (Compound 87). Raltegravir-resistant strains HIV 4376 and HIV 8070 were obtained from the NIH AIDS repository.

EM

A chronic HIV-1 producer cell line A1953 [45] was pretreated for 24 h using 100 and 1,000 nM GSK002, 100 and 1,000 nM GSK1264, or diluent control (DMSO). The cells were washed with 1 x PBS and incubated with fresh medium containing inhibitor or DMSO. After 72 h, the cells were harvested by centrifugation, and the cell pellets were fixed (2.0% paraformaldehyde, 2.5% glutaraldehyde in 0.2 M Na-CaCO, pH 7.3), embedded in Epon, and sectioned. Microscopy was performed with a JEM-1010 (JEOL, Tokyo, Japan) transmission electron microscope operated at 80 kV. Viral particles were categorized based on the viral morphology as mature, deformed (aberrant/eccentric/empty core), immature (no core), or ambiguous. For the 100 nM dose, findings were quantified by two blinded investigators.

Protein Expression and Purification

Full-length and truncated HIV-1 IN(NL4-3) and coexpressions with LEDGF(IBD) (346–471) constructs were expressed and purified as described previously, with some modifications [46–48]. IN-only constructs encoding the F185H solubility mutation were expressed from a pET-Duet-derived (Novagen) vector in which the IN construct was inserted in frame with a C-terminal Mxe intein (New England Biolabs) containing a chitin binding domain and hexahistidine tag. The QuikChange kit (Stratagene) was used to generate the point mutation. Proteins were purified using nickel-nitrilotriacetic acid (Qiagen) and chitin (New England Biolabs) resins. After fusion proteins were liberated by intein cleavage in 50 mM dithiothreitol (DTT) overnight at 4°C, IN preparations were further purified using a Superdex 75 HiLoad 16/60 column at room temperature, eluted isocratically in 20 mM HEPES-NaOH pH 7.0, 1 M NaCl, 7 mM CHAPS, 10 uM ZnOAc$_2$, and 10 mM β-ME. Proteins were concentrated at 4°C in a YM-10 Centricon (Millipore), and aliquots were flash-frozen in liquid nitrogen for storage at −80°C. Genetically solubilized IN^{C56S,F139D,F185H,280S} (quadramutated, QM) with and without the additional W131D mutation (pentamutated, PM) were purified similarly into a final buffer of 20 mM

HEPES-NaOH, pH 7.0, 450 mM NaCl, 7 mM CHAPS, 10 uM ZnOAc$_2$, and 10 mM β-ME or 1–10 mM DTT. IN•LEDGF(IBD) coexpressions were produced from the same expression vector, except the IN-coding DNA was inserted into the first multiple cloning site (MCS) and LEDGF into the second MCS, in frame with the C-terminal Mxe intein-hexahistidine tag. These preparations were purified similarly into a final buffer of 20 mM HEPES-NaOH, pH 7.0, 1 M NaCl, 7 mM CHAPS, 10 uM ZnOAc$_2$, and 10 mM β-ME or 1–10 mM DTT. IN^{F185K}(CCD) used for crystallization was obtained by expression from the vector pET24 (Novagen, Madison, Wisconsin, United States) in BL21star (DE3) cells (Novagen) at 37°C. The sequence was inserted into the vector in frame with a N-terminal TEV-cleavable hexahistidine affinity tag. Protein was initially purified using Ni-Sepharose FF (GE Healthcare Life Sciences, Pittsburgh, Pennsylvania, US), followed by cleavage of the affinity tag using hexahistidine-tagged TEV at a 1:100 mass ratio in 25 mM HEPES-NaOH pH 7.5, 750 mM NaCl, and 25 mM Imidazole at 25°C with a 10K MWCO Jumbosep (Pall, Exton, Pennsylvania, US) centrifugal concentrator. To separate the liberated protein from uncleaved materials and protease, another Ni-Sepharose FF purification step was performed. The flow-through material was concentrated and injected onto a Superdex-75 column and eluted isocratically in 10 mM HEPES-NaOH pH 7, 500 mM NaCl, and 3 mM DTT. For biophysical analyses, samples were exchanged into 20 mM HEPES•NaOH pH 7.5, 1 M NaCl, 7 mM CHAPS, 10 mM DTT, and 10 μM Zn(OAc)$_2$.

Crystallization and Structure Determination

For crystallization, the inhibitor was prepared in 1,3-dimethyl-2-imidazolidinone (DMI, Fluka). Recombinant HIV IN^{Y15A,F185H} at 5–10 mg/mL concentrations in 20 mM HEPES pH 7.0, 0.5–1 M NaCl, 5 mM CHAPS, 10 μM ZnOAc$_2$, 1–10 mM dithiothreitol (DTT), and 0.5–1.0 mM GSK1264 was mixed 1:1 with reservoir solutions containing 30% 2-methyl-2,4 pentanediol (MPD), 0.1 M sodium citrate pH 5.6–6.5 at 21°C in a standard hanging drop format. Crystals formed within 3–4 d and were iteratively improved by macroseeding. Diffraction data were obtained at the beam lines 5.0.2 and 12.3.1 of the Advanced Light Source (ALS, Berkeley, California, US).

All data reductions were performed using the program XDS [49]. Diffraction outwards of ~4 Å was observed in the best case, with resolution dropping off quickly as a function of X-ray dose. Analysis of the final dataset by the UCLA diffraction anisotropy server [50] indicated that diffraction was significantly anisotropic along the a^*- and b^*-axes. On the basis of an F/σ (F) cutoff of 3, reflections were subjected to an anisotropic truncation of 4.5, 4.5, and 4.3 Å along a^*, b^*, and c^*, respectively, before use in refinement. Molecular replacement was performed using the program Phaser [51,52] as implemented in the Phenix software package [53]. Molecular replacement searches using a number of different combinations of monomer or dimer IN CCD$^{50-212}$ (PDB 4OJR, [7]) and CTD$^{220-270}$ (PDB 1EX4, [29]) all yielded strong molecular replacement solutions with very high LLG scores. In the initial experimental maps, strong and contiguous electron density was observed corresponding to the large helical extensions of IN$^{210:220}$ seen in the 1EX4 crystal structure of IN(CCD-CTD) and was readily traced. After an initial iteration of simulated annealing, iterative rounds of rigid body refinement and minimization were performed in Phenix alongside manual building and refinement in COOT [54]. After an IN$^{50-270}$ dimer was completely built, the structure was refined in CNS v1.3 [55,56] using rigid body refinement and the DEN method [26–28], using the high resolution IN^{F185K}(CCD)•GSK1264 (PDB 4OJR, [7]) and residues 211–270 from the available 2.4 Å IN (CCD-CTD) (PDB 1EX4, [29]) structure as reference models. Ligand was omitted from the refinement until the latest stages. Strong ligand density corresponding to GSK1264 was observed in F$_o$-F$_c$ and simulated annealing F$_o$-F$_c$ omit maps, allowing for docking of the

ligands at both binding sites at the IN(CCD) dimer interface. Composite omit maps were generated in Phenix, and figures created using PYMOL [57]. Crystallographic structures of IN (CCD)^{F185K} and related mutants were prepared and solved as previously described [7], using the beam line 21IDG at the Advanced Photon Source (Argonne, Illinois, US). Crystallographic statistics for the five structures are summarized in S2 Table. Atomic coordinates and structure factors were deposited in the Protein Data Bank under the accession codes 5HOT, 5HRN, 5HRP, 5HRR, and 5HRS.

Virological Methods

Chronically infected HIV-1 producer A1953 cells were a gift from James Hoxie. The cells were cultured at 37°C in 5% CO_2 in DMSO medium, supplemented with 10% fetal bovine serum and penicillin-streptomycin. To test the viability of HIV IN^{Y15A}, DNA encoding the substitution was built into the HIV NL4-3 backbone using Gibson assembly. The sequence of the resulting plasmid was confirmed by Illumina sequencing of Nextera-XT libraries. Viral stocks (IN^{Y15A} and wild type) were generated by transfection into 293T cells and culture supernatants containing the HIV particles that were collected. Viral stocks were normalized by the amount of p24 capsid antigen per unit volume and then applied to TZMBL indicator cells. High luciferase activity was detected on day 2 after infection for wild type (2,334 RLU), whereas luciferase expression in IN^{Y15A} samples (21 RLUs) remained close to background (7 RLUs). A long-term culture experiment was carried out in which cells infected with the IN^{Y15A} virus were monitored for 2 mo by luciferase expression and p24 assay, but no viral replication could be detected.

Virus Resistance Passage

MT4 cells infected with the HIV lab strain NL4-3 or Raltegravir-resistant viruses (NIH AIDS Research and Reference Reagent Program catalog #11842 and 11845) were cultured in the presence of suboptimal inhibitor concentrations at approximately one-half the IC$_{50}$. Viral replication kinetics were monitored by the production of RT activity in the supernatant. When the kinetics of the inhibitor-treated culture matched that of the no-inhibitor control color for three consecutive passages, the inhibitor concentration was increased. At periodic intervals, approximately every ten passages, the virus was expanded, and IN sequence determined. Drug-induced mutations were created in the proviral plasmid pNL4-3 for confirmation of inhibitor sensitivity.

HIV Replication Assays in PBMC Culture

Primary HIV-1 isolates were grown in human PBMCs. PBMCs were sedimented by layering blood obtained from donors over a Ficoll (Histopaque 1077; Sigma) cushion in 50 ml tubes. The cells were washed, pooled from various donors, and cryopreserved. PBMCs were thawed and stimulated with 2 µg/mL PHA in RPMI 1640 media supplemented with 10% heat-inactivated fetal bovine serum, 2 mM L-glutamine, 100 units/ml penicillin G, and 100 µg/ml streptomycin for 3 d prior to infection. Viral yields were determined for primary isolates 7 d post infection by a radioactive RT endpoint. To determine isolate sensitivity for various inhibitors, the appropriate titer of each isolate was used to infect PHA-stimulated PBMC pools in the presence of an inhibitor dilution series. Viral replication was measured by RT as the endpoint after 7 d in culture, and the IC$_{50}$ calculated as the inhibitor concentration required to reduce viral replication by 50% of the control.

Aggregation Assays

Assays were performed as previously described [7], with the following modification: turbidity assays were performed using VICTOR3V 1420 multilabel counter (PerkinElmer, Waltham, Massachusetts, US) by measuring the absorbance of the reaction solution at 405 nm. Final reaction conditions were 20 mM HEPES, pH 7.3, 375–505 mM NaCl, 3.75–5.05 mM CHAPS, 10 mM DTT, and 10 uM $ZnAc_2$ with inhibitor concentrations ranging from 0.08 µM to 88 µM at 24–27˚C. For the graphical representation of aggregation, the baseline (IN stock buffer + drug buffer) has been subtracted from the results.

SEC-MALS

Absolute molecular weights were determined by multiangle light scattering coupled with refractive interferometric detection (Wyatt Technology, Santa Barbara, California, US) and a Superdex 200 10/300 GL column (GE Healthcare) at room temperature as previously described [48].

Sedimentation Equilibrium Analysis

Sedimentation equilibrium AUC experiments were performed at 4˚C with an XL-A analytical ultracentrifuge (Beckman-Coulter, Brea, California, US) and a TiAn60 rotor with two-channel charcoal-filled Epon centerpieces and quartz windows. Data were collected at 4˚C with detection at 280 nm for 5, 7.5, and 10 µM samples. Linear analyses were performed by plotting the natural log of absorbance versus the square of radius, with the slope being proportional to M_w. Single-species plots with calculated slopes for idealized oligomers were also calculated at a given speed for comparison.

Oligonucleotide Assay for IN Activity

Assays for strand transfer were adapted from those described previously [58–60]. IN stocks were diluted into 20 mM HEPES•NaOH pH 7.5, 1 M NaCl, 7 mM CHAPS, 10 mM DTT, and 10 µM $Zn(OAc)_2$ for a final concentration of 1 µM IN. Processed U5 LTR substrates with a 5′ Alexafluor 488 NHS Ester label were prepared by annealing the following oligonucleotides: 5′-Alexa488-ACCCTTTTAGTCAGTGTGGAAAATCTCTAGCA-3′ and 5′-TGACGATCTCTAAAAGGTGTGACTGATTTTCCCAG-3′ (IDT). Reactions were performed with 20 mM HEPES-NaOH pH 7.6, 0.5 µM Alexaflour 488-labeled LTR substrate, 15 nM pUC19 plasmid, 10 mM DTT, 100 mM NaCl, 5 mM $MnCl_2$, 10 µM $Zn(OAc)_2$, 5% DMSO, and 12% PEG-6000. Reactions were carried out for 60 min at 37˚C and quenched using 0.5% SDS and 15 mM EDTA. After proteinase K treatment (1 mg/ml) for 30 min, reactions were separated on 1.5% Agarose gels in Tris-acetate buffer and visualized by EtBr staining or fluorescence using a Typhoon image station.

Supporting Information

S1 Data. Excel spreadsheet containing, in separate sheets, the underlying numerical data for Figure panels Fig 1C, Fig 4B–4H, Fig 5B–5D, Fig 7A, S1C–S1F Fig, S3B–S3D Fig and S5A Fig.
(XLSX)

S1 Fig. Properties of HIV-1 IN^Y15A,F185H. (A) The D and E-forms of HIV IN(NTD) in isolation [22]. (B) The purification scheme and SDS-PAGE analysis of purified IN (32kD). Dissolved crystals of INY15A,F185H were washed and dissolved for SDS-PAGE analysis (right). (C) Sedimentation equilibrium analysis of IN^Y15A,F185H. Data were recorded at 12,000 RPM, at

a concentration of 10 μM IN, at 4˚C. Linearized radial distributions are shown. The slopes are proportional to M_w at a given value of r^2. Single-species plots with calculated slopes for idealized IN monomer, dimer, and tetramer are shown for the same rotor speed and temperature as grey lines. (D) SEC-MALS analysis of $IN^{Y15A,F185H}$. The experiments were performed at room temperature using a Superdex 200 10/300 column. Samples were injected at 10 mg/mL. The elution concentrations by refractive index approached ~0.1 mg/mL. Both the M_w (weight-average molecular mass) from multiangle light scattering and retention times are consistent with a dimer of IN. (E) SEC-MALS analysis of $IN^{Y15A,F185H}$•LEDGF(IBD). The M_w (weight-average molecular mass) from multiangle light scattering and retention times is consistent with a 2:2 LEDGF(IBD)-bound IN dimer. Data plotted in panels C, D, and E are provided in S1 Data. (F) ALLINI-induced aggregation of HIV-1 $IN^{Y15A,F185H}$. (G) Crystals of HIV-1 $IN^{Y15A,F185H}$ (top) and the corresponding X-ray diffraction from beam line ALS 5.0.3. (TIF)

S2 Fig. CHAPS binding at the IN CTD–catalytic core domain–1264 interface. (A) Superposition of the 1EX4 structure of IN CCD-catalytic core domain ([25], blue) superposed with the IN•GSK1264 crystal structure (grey) from this work. The structures superpose with a RMSD of 2.3 Å across all atoms of the CCD and catalytic core domain, differing mostly in the CTD regions, including the rotation of the SH3 domains relative to the catalytic core domain. CHAPS molecules observed in the 1EX4 structure are shown around the CTDs, rendered as blue sticks. (B) Overlay of the IN•GSK1264 structure with the 1EX4 CHAPS-bound model. Shown in red spheres is GSK1264, and in blue spheres, CHAPS. The position of 1264 and bound CHAPS coincide. (C) Electrostatic surfaces at the GSK1264–CTD–catalytic core domain interface. Cutaway views of the electrostatic surfaces at the CTD–catalytic core domain interface are shown. GSK1264 (ball and stick) is shown for reference. The interaction is bipartite, with packing between hydrophobic surfaces of α3 and β2 and complementary electrostatics between α5 and β5. (TIF)

S3 Fig. Effects of W131D on oligomerization and ALLINI-induced effects. (A) Approximately 80% of the buried interaction between CTD and catalytic core domain occurs with the α3 helix, and Trp131 is central to this packing. Highlighted here is Trp131 at this interface. (B) SEC-MALS analysis of IN-LEDGF coexpressions. Shown is a comparison of SEC-MALS data for a genetically solubilized IN background characterized previously ($IN^{C56S, F139D, F185H, C280S}$; termed IN^{QM}) [7,48,61] coexpressed with LEDGF(IBD). IN^{QM}•LEDGF(IBD) is a 4:2 complex in solution [48]. While the W131D mutation in this background retains LEDGF binding, mostly LEDGF-bound dimers are observed, consistent with the role of the residue at the CTD–catalytic core domain interface and the model presented in Fig 3. In complementary sedimentation equilibrium experiments (not shown), a K_d for dimer-tetramer of 90 μM is modelled for $IN^{QM-W131D}$, versus 9 μM for control IN^{QM} at 4˚C. (C) Sedimentation velocity analysis of $IN^{QM-W131D}$. c(S) analysis of sedimentation velocity data is shown for full-length 30 μM IN^Q or $IN^{Q-W131D}$ at 4˚C and 20˚C in the presence of DMSO (black) or 50 μM drug (grey). Species assigned as monomers (~1.8 S), dimers (~3.4 S), and tetramers (~5.5 S) are denoted. Distributions were derived from the fitting of the Lamm equation to the experimental data collected in the first 2 h of the experiment, as implemented in the program SEDFIT [62]. (D) Concentration and time-dependent aggregation of 9 μM IN^Q alone (upper left, [48]), 9 μM IN^P alone (lower left, this work), LEDGF-bound IN^Q (upper right, [48]), or LEDGF-bound IN^P (lower right, this work). The mutation W131D attenuates drug-induced aggregation of IN. Data plotted in panels B, C, and D are provided in S1 Data. (TIF)

S4 Fig. Structural basis for polymorphism-induced resistance to GSK002. Structural basis for polymorphism-induced resistance to GSK002. (A–C) Superposition of IN^{F185K}(CCD) structures with GSK002 bound at the dimer interface. Shown in grey is IN^{F185K} bound with drug. Shown superposed are IN^{A124T,F185K} (green), IN^{A124N,T125S,F185K} (blue), and IN^{A124N,T125A,F185K} (yellow), all bound with GSK002. All structures superpose within 0.2 Å RMSD between α carbons. (D) Model of the GSK002-IN(CCD-CTD) drug binding interface. Highlighted are the predicted clashes between CTD, the difluorobenzyl moiety of GSK002, and residue 124.
(TIF)

S5 Fig. Mutations at Y226 and W235 are inactive for strand transfer. (A) SEC-MALS analysis of IN^{F185H}. Experiments were performed at room temperature using a Superdex 200 10/300 column. Data plotted here are provided in S1 Data. (B) Shown is a gel electrophoresis analysis of HIV IN mutants tested in strand transfer reactions involving preprocessed Alexaflour-488 labelled U5 substrate (35mers) mimicking the HIV LTR, as well as pUC19 plasmid (2.7 kB) serving as the integration target. Shown on the left is fluorescent imaging of products from the integration reaction and on the right the same gel visualized with ethidium bromide to show total DNA, which is predominately the plasmid target. While both half-site integrants (H.S.I) and full-site integrants (F.S.I) are seen for IN^{F185H}, point mutants of Tyr226 (Y226A, Y226D, and Y226R) and W235 (W235A) are inactive for strand transfer over a 1-h time course.
(TIF)

S1 Table. Solution properties of IN variants in this study.
(DOCX)

S2 Table. Summary of X-ray data processing and refinement statistics.
(DOCX)

S3 Table. Serial passage resistance data.
(DOC)

Acknowledgments

Author Contributions

Conceived and designed the experiments: KG JJ GDVD FDB.

Performed the experiments: KG VT YH GE LT CM PW DT.

Analyzed the data: KG VT YH GE LT CM PW SSM DT RTN EV JJ GDVD FDB.

KG VT GDVD FDB.

References

1. Arts EJ, Hazuda DJ. HIV-1 antiretroviral drug therapy. Cold Spring Harbor perspectives in medicine. 2012; 2(4):a007161. PubMed Central PMCID: PMC3312400. doi: 10.1101/cshperspect.a007161 PMID: 22474613

2. Hazuda DJ. HIV integrase as a target for antiretroviral therapy. Current opinion in HIV and AIDS. 2012; 7(5):383–9. doi: 10.1097/COH.0b013e3283567309 PMID: 22871634

3. Christ F, Shaw S, Demeulemeester J, Desimmie BA, Marchand A, Butler S, et al. Small-molecule inhibitors of the LEDGF/p75 binding site of integrase block HIV replication and modulate integrase multimerization. Antimicrobial agents and chemotherapy. 2012; 56(8):4365–74. Epub 2012/06/06. PubMed Central PMCID: PMC3421592. doi: 10.1128/AAC.00717-12 PMID: 22664975

4. Christ F, Voet A, Marchand A, Nicolet S, Desimmie BA, Marchand D, et al. Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. Nat Chem Biol. 2010; 6 (6):442–8. Epub 2010/05/18. doi: 10.1038/nchembio.370 PMID: 20473303

5. Feng L, Sharma A, Slaughter A, Jena N, Koh Y, Shkriabai N, et al. 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. Epub 2013/04/26. PubMed Central PMCID: PMC3668738. doi: 10.1074/jbc.M112.443390 PMID: 23615903

6. Fontana J, Jurado KA, Cheng N, Ly NL, Fuchs JR, Gorelick RJ, et al. Distribution and Redistribution of HIV-1 Nucleocapsid Protein in Immature, Mature, and Integrase-Inhibited Virions: a Role for Integrase in Maturation. Journal of virology. 2015; 89(19):9765–80. PubMed Central PMCID: PMC4577894. doi: 10.1128/JVI.01522-15 PMID: 26178982

7. Gupta K, Brady T, Dyer BM, Malani N, Hwang Y, Male F, et al. Allosteric inhibition of human immunodeficiency virus integrase: late block during viral replication and abnormal multimerization involving specific protein domains. The Journal of biological chemistry. 2014; 289(30):20477–88. Epub 2014/06/07. PubMed Central PMCID: PMC4110260. doi: 10.1074/jbc.M114.551119 PMID: 24904063

8. Jurado KA, Wang H, Slaughter A, Feng L, Kessl JJ, Koh Y, et al. 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. PubMed Central PMCID: PMC3666754. doi: 10.1073/pnas.1300703110 PMID: 23610442

9. Le Rouzic E, Bonnard D, Chasset S, Bruneau JM, Chevreuil F, Le Strat F, et al. Dual inhibition of HIV-1 replication by integrase-LEDGF allosteric inhibitors is predominant at the post-integration stage. Retrovirology. 2013; 10:144. Epub 2013/11/23. doi: 10.1186/1742-4690-10-144 PMID: 24261564

10. Molteni V, Greenwald J, Rhodes D, Hwang Y, Kwiatkowski W, Bushman FD, et al. Identification of a small molecule binding site at the dimer interface of the HIV integrase catalytic domain. Acta Crystallogr D Biol Crystallogr. 2001; 57:536–44. PMID: 11264582

11. Tsiang M, Jones GS, Niedziela-Majka A, Kan E, Lansdon EB, Huang W, et al. New class of HIV-1 integrase (IN) inhibitors with a dual mode of action. The Journal of biological chemistry. 2012; 287 (25):21189–203. Epub 2012/04/27. PubMed Central PMCID: PMC3375541. doi: 10.1074/jbc.M112. 347534 PMID: 22535962

12. Craigie R, Bushman FD. HIV DNA Integration. Cold Spring Harbor perspectives in medicine. 2012; 2 (7):a006890. Epub 2012/07/05. PubMed Central PMCID: PMC3385939. doi: 10.1101/cshperspect. a006890 PMID: 22762018

13. Engelman A, Cherepanov P. The structural biology of HIV-1: mechanistic and therapeutic insights. Nature reviews Microbiology. 2012; 10(4):279–90. PubMed Central PMCID: PMC3588166. doi: 10. 1038/nrmicro2747 PMID: 22421880

14. Ciuffi A, Llano M, Poeschla E, Hoffmann C, Leipzig J, Shinn P, et al. A role for LEDGF/p75 in targeting HIV DNA integration. Nat Med. 2005; 11(12):1287–9. Epub 2005/11/29. doi: 10.1038/nm1329 PMID: 16311605

15. Schroder ARW, Shinn P, Chen HM, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. Cell. 2002; 110(4):521–9. PMID: 12202041

16. Shun MC, Raghavendra NK, Vandegraaff N, Daigle JE, Hughes S, Kellam P, et al. LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. Genes & development. 2007; 21(14):1767–78. Epub 2007/07/20. PubMed Central PMCID: PMC1920171.

17. Christ F, Debyser Z. The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy. Virology. 2013; 435(1):102–9. Epub 2012/12/12. doi: 10.1016/j.virol.2012.09.033 PMID: 23217620

18. Balakrishnan M, Yant SR, Tsai L, O'Sullivan C, Bam RA, Tsai A, et al. Non-catalytic site HIV-1 integrase inhibitors disrupt core maturation and induce a reverse transcription block in target cells. PLoS ONE. 2013; 8(9):e74163. PubMed Central PMCID: PMC3767657. doi: 10.1371/journal.pone.0074163 PMID: 24040198

19. Sharma A, Slaughter A, Jena N, Feng L, Kessl JJ, Fadel HJ, et al. A new class of multimerization selective inhibitors of HIV-1 integrase. PLoS Pathog. 2014; 10(5):e1004171. Epub 2014/05/31. PubMed Central PMCID: PMC4038613. doi: 10.1371/journal.ppat.1004171 PMID: 24874515

20. Jenkins TM, Engelman A, Ghirlando R, Craigie R. A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. The Journal of biological chemistry. 1996; 271(13):7712–8. PMID: 8631811

21. Engelman A, Liu Y, Chen H, Farzan M, Dyda F. Structure-based mutagenesis of the catalytic domain of human immunodeficiency virus type 1 integrase. Journal of virology. 1997; 71(5):3507–14. Epub 1997/05/01. PubMed Central PMCID: PMC191497. PMID: 9094622

22. Nomura Y, Masuda T, Kawai G. Structural analysis of a mutant of the HIV-1 integrase zinc finger domain that forms a single conformation. Journal of biochemistry. 2006; 139(4):753–9. doi: 10.1093/jb/mvj085 PMID: 16672276

23. Hare S, Di Nunzio F, Labeja A, Wang J, Engelman A, Cherepanov P. Structural basis for functional tetramerization of lentiviral integrase. PLoS Pathog. 2009; 5(7):e1000515. Epub 2009/07/18. PubMed Central PMCID: PMC2705190. doi: 10.1371/journal.ppat.1000515 PMID: 19609359

24. Shkriabai N, Dharmarajan V, Slaughter A, Kessl JJ, Larue RC, Feng L, et al. A critical role of the C-terminal segment for allosteric inhibitor-induced aberrant multimerization of HIV-1 integrase. The Journal of biological chemistry. 2014; 289(38):26430–40. Epub 2014/08/15. PubMed Central PMCID: PMC4176244. doi: 10.1074/jbc.M114.589572 PMID: 25118283

25. Chen JC, Krucinski J, Miercke LJ, Finer-Moore JS, Tang AH, Leavitt AD, et al. Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. Proceedings of the National Academy of Sciences of the United States of America. 2000; 97(15):8233–8. Epub 2000/07/13. PubMed Central PMCID: PMC26930. doi: 10.1073/pnas.150220297 PMID: 10890912

26. Brunger AT, Adams PD, Fromme P, Fromme R, Levitt M, Schroder GF. Improving the accuracy of macromolecular structure refinement at 7 A resolution. Structure. 2012; 20(6):957–66. Epub 2012/06/12. PubMed Central PMCID: PMC3380535. doi: 10.1016/j.str.2012.04.020 PMID: 22681901

27. Brunger AT, Das D, Deacon AM, Grant J, Terwilliger TC, Read RJ, et al. Application of DEN refinement and automated model building to a difficult case of molecular-replacement phasing: the structure of a putative succinyl-diaminopimelate desuccinylase from Corynebacterium glutamicum. Acta crystallographica Section D, Biological crystallography. 2012; 68(Pt 4):391–403. Epub 2012/04/17. PubMed Central PMCID: PMC3322598. doi: 10.1107/S090744491104978X PMID: 22505259

28. Schroder GF, Levitt M, Brunger AT. Deformable elastic network refinement for low-resolution macromolecular crystallography. Acta crystallographica Section D, Biological crystallography. 2014; 70(Pt 9):2241–55. Epub 2014/09/10. PubMed Central PMCID: PMC4157441. doi: 10.1107/S1399004714016496 PMID: 25195739

29. Chen Z, Yan Y, Munshi S, Li Y, Zugay-Murphy J, Xu B, et al. X-ray structure of simian immunodeficiency virus integrase containing the core and C-terminal domain (residues 50–293)—an initial glance of the viral DNA binding platform. Journal of molecular biology. 2000; 296(2):521–33. Epub 2000/02/12. doi: 10.1006/jmbi.1999.3451 PMID: 10669606

30. Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. Journal of chemical information and modeling. 2011; 51(10):2778–86. Epub 2011/09/17. doi: 10.1021/ci200227u PMID: 21919503

31. Kalpana GV, Reicin A, Cheng GS, Sorin M, Paik S, Goff SP. Isolation and characterization of an oligomerization-negative mutant of HIV-1 integrase. Virology. 1999; 259(2):274–85. Epub 1999/07/02. doi: 10.1006/viro.1999.9767 PMID: 10388652

32. Nishitsuji H, Hayashi T, Takahashi T, Miyano M, Kannagi M, Masuda T. Augmentation of reverse transcription by integrase through an interaction with host factor, SIP1/Gemin2 Is critical for HIV-1 infection. PLoS ONE. 2009; 4(11):e7825. Epub 2009/11/17. PubMed Central PMCID: PMC2771899. doi: 10.1371/journal.pone.0007825 PMID: 19915660

33. Fader LD, Malenfant E, Parisien M, Carson R, Bilodeau F, Landry S, et al. Discovery of BI 224436, a Noncatalytic Site Integrase Inhibitor (NCINI) of HIV-1. ACS medicinal chemistry letters. 2014; 5(4):422–7. Epub 2014/06/06. PubMed Central PMCID: PMC4027581. doi: 10.1021/ml500002n PMID: 24900852

34. Fenwick C, Tremblay S, Wardrop E, Bethell R, Coulombe R, Elston R, et al. Resistance studies with HIV-1 non-catalytic site integrase inhibitors. Antiviral Therapy. 2011; 16(Suppl 1):A9.

35. Slaughter A, Jurado KA, Deng N, Feng L, Kessl JJ, Shkriabai N, et al. The mechanism of H171T resistance reveals the importance of Ndelta-protonated His171 for the binding of allosteric inhibitor BI-D to HIV-1 integrase. Retrovirology. 2014; 11:100. Epub 2014/11/26. PubMed Central PMCID: PMC4251946. doi: 10.1186/s12977-014-0100-1 PMID: 25421939

36. Fenwick C, Amad M, Bailey MD, Bethell R, Bos M, Bonneau P, et al. Preclinical profile of BI 224436, a novel HIV-1 non-catalytic-site integrase inhibitor. Antimicrobial agents and chemotherapy. 2014; 58(6):3233–44. Epub 2014/03/26. PubMed Central PMCID: PMC4068430. doi: 10.1128/AAC.02719-13 PMID: 24663024

37. Goldgur Y, Dyda F, Hickman AB, Jenkins TM, Craigie R, Davies DR. Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. Proceedings of the National Academy

of Sciences of the United States of America. 1998; 95(16):9150–4. Epub 1998/08/05. PubMed Central PMCID: PMC21307. PMID: 9689049

38. Wang JY, Ling H, Yang W, Craigie R. Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. The EMBO journal. 2001; 20(24):7333–43. Epub 2001/12/18. PubMed Central PMCID: PMC125787. doi: 10.1093/emboj/20.24.7333 PMID: 11743009

39. Tibshirani R. Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society Series B (Methodological). 1996; 58(1):267–88.

40. Mitra A, Sept D. Taxol allosterically alters the dynamics of the tubulin dimer and increases the flexibility of microtubules. Biophysical journal. 2008; 95(7):3252–8. Epub 2008/07/16. PubMed Central PMCID: PMC2547448. doi: 10.1529/biophysj.108.133884 PMID: 18621813

41. Prota AE, Bargsten K, Zurwerra D, Field JJ, Diaz JF, Altmann KH, et al. Molecular mechanism of action of microtubule-stabilizing anticancer agents. Science. 2013; 339(6119):587–90. Epub 2013/01/05. doi: 10.1126/science.1230582 PMID: 23287720

42. Deprez E, Tauc P, Leh H, Mouscadet JF, Auclair C, Brochon JC. Oligomeric states of the HIV-1 integrase as measured by time-resolved fluorescence anisotropy. Biochemistry. 2000; 39(31):9275–84. Epub 2000/08/05. PMID: 10924120

43. Hare S, Gupta SS, Valkov E, Engelman A, Cherepanov P. Retroviral intasome assembly and inhibition of DNA strand transfer. Nature. 2010; 464(7286):232–6. Epub 2010/02/02. PubMed Central PMCID: PMC2837123. doi: 10.1038/nature08784 PMID: 20118915

44. Hare S, Shun MC, Gupta SS, Valkov E, Engelman A, Cherepanov P. A novel co-crystal structure affords the design of gain-of-function lentiviral integrase mutants in the presence of modified PSIP1/LEDGF/p75. PLoS Pathog. 2009; 5(1):e1000259. Epub 2009/01/10. PubMed Central PMCID: PMC2606027. doi: 10.1371/journal.ppat.1000259 PMID: 19132083

45. Inoue M, Hoxie JA, Reddy MV, Srinivasan A, Reddy EP. Mechanisms associated with the generation of biologically active human immunodeficiency virus type 1 particles from defective proviruses. Proceedings of the National Academy of Sciences of the United States of America. 1991; 88(6):2278–82. Epub 1991/03/15. PubMed Central PMCID: PMC51214. PMID: 2006168

46. Ciuffi A, Diamond TL, Hwang Y, Marshall HM, Bushman FD. Modulating target site selection during human immunodeficiency virus DNA integration in vitro with an engineered tethering factor. Human gene therapy. 2006; 17(9):960–7. Epub 2006/09/16. doi: 10.1089/hum.2006.17.960 PMID: 16972764

47. Diamond TL, Bushman FD. Role of metal ions in catalysis by HIV integrase analyzed using a quantitative PCR disintegration assay. Nucleic acids research. 2006; 34(21):6116–25. Epub 2006/11/07. PubMed Central PMCID: PMC1693899. doi: 10.1093/nar/gkl862 PMID: 17085478

48. Gupta K, Diamond T, Hwang Y, Bushman F, Van Duyne GD. Structural properties of HIV integrase. Lens epithelium-derived growth factor oligomers. The Journal of biological chemistry. 2010; 285 (26):20303–15. Epub 2010/04/22. PubMed Central PMCID: PMC2888443. doi: 10.1074/jbc.M110.114413 PMID: 20406807

49. Kabsch W. XDS. Acta Crystallogr D. 2010; 66:125–32. doi: 10.1107/S0907444909047337 PMID: 20124692

50. Strong M, Sawaya MR, Wang SS, Phillips M, Cascio D, Eisenberg D. Toward the structural genomics of complexes: Crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences of the United States of America. 2006; 103(21):8060–5. doi: 10.1073/pnas.0602606103 PMID: 16690741

51. McCoy AJ. Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr D. 2007; 63:32–41. doi: 10.1107/S0907444906045975 PMID: 17164524

52. Mccoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–74. doi: 10.1107/S0021889807021206 PMID: 19461840

53. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D. 2010; 66:213–21. doi: 10.1107/S0907444909052925 PMID: 20124702

54. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D. 2010; 66:486–501. doi: 10.1107/S0907444910007493 PMID: 20383002

55. Brunger AT. Version 1.2 of the Crystallography and NMR system. Nat Protoc. 2007; 2(11):2728–33. doi: 10.1038/nprot.2007.406 PMID: 18007608

56. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D. 1998; 54:905–21. PMID: 9757107

57. DeLano WL. PyMOL molecular viewer: Updates and refinements. Abstr Pap Am Chem S. 2009; 238.

58. Gao K, Butler SL, Bushman F. Human immunodeficiency virus type 1 integrase: arrangement of protein domains in active cDNA complexes. The EMBO journal. 2001; 20(13):3565–76. Epub 2001/07/04. PubMed Central PMCID: PMC125503. doi: 10.1093/emboj/20.13.3565 PMID: 11432843

59. Bushman FD, Fujiwara T, Craigie R. Retroviral DNA integration directed by HIV integration protein in vitro. Science. 1990; 249(4976):1555–8. Epub 1990/09/28. PMID: 2171144

60. Bushman FD, Craigie R. Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. Proceedings of the National Academy of Sciences of the United States of America. 1991; 88(4):1339–43. Epub 1991/02/15. PubMed Central PMCID: PMC51013. PMID: 1847518

61. Alian A, Griner SL, Chiang V, Tsiang M, Jones G, Birkus G, et al. Catalytically-active complex of HIV-1 integrase with a viral DNA substrate binds anti-integrase drugs. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106(20):8192–7. Epub 2009/05/07. PubMed Central PMCID: PMC2688900. doi: 10.1073/pnas.0811919106 PMID: 19416821

62. Schuck P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophysical journal. 2000; 78(3):1606–19. Epub 2000/02/29. PubMed Central PMCID: PMC1300758. doi: 10.1016/S0006-3495(00)76713-0 PMID: 10692345