Supplementary Materials for

Structural basis of second-generation HIV integrase inhibitor action and viral resistance

Nicola J. Cook, Wen Li, Dénes Berta, Magd Badaoui, Allison Ballandras-Colas, Andrea Nans, Abhay Kotecha, Edina Rosta, Alan N. Engelman*, Peter Cherepanov*

*Corresponding author. Email: peter.cherepanov@crick.ac.uk (P.C.); alan_engelman@dfci.harvard.edu (A.N.E.)

This PDF file includes:

Materials and Methods
Figs. S1 to S27
Table S1
Caption for Movie S1
References

Other Supplementary Material for this manuscript includes the following:
(available at science.sciencemag.org/cgi/content/full/science.aay4919/DC1)

Movie S1 (.mov)

Materials and Methods

Protein expression and purification

Codon-optimized DNA fragments encoding various SIV INs were cloned between *Xma*I and *Bam*HI sites of pCPH6P (*12*) to produce proteins with cleavable N-terminal hexahistidine tags. The amino acid sequence of SIVrcm IN was derived from NCBI entry HM803689 (*10*), and the A119D mutation was introduced using QuikChange procedure (Stratagene). Recombinant lentiviral INs were produced essentially as previously described (*12, 13*). Briefly, protein expression was induced in endonuclease A-negative *Escherichia coli* PC2 cells (*12*) grown in LB medium with 0.01% isopropyl β-D-1-thiogalactopyranoside (IPTG) at 25 ᵒC. Cells were lysed by sonication in buffer A containing 20 mM Tris-HCl pH 7.5, 1 M NaCl, 0.5% 3-[(3-cholamidopropyl) dimethylammonio]-1-propane sulfonate (CHAPS) supplemented with 1 mM phenylmethylsulfonyl fluoride and cOmplete EDTA-free protease inhibitor cocktail (Roche). Supernatant pre-cleared by centrifugation was incubated with NiNTA agarose (Qiagen) at 4°C in the presence of 20 mM imidazole. The resin was washed with buffer A containing 20 mM imidazole and recombinant protein was eluted in buffer A containing 200 mM imidazole. Peak fractions were supplemented with 10 mM dithiothreitol (DTT) and incubated with human rhinovirus 3C protease (1:50 w/w ratio) at 4°C overnight to remove the hexahistidine tag. The protein was diluted to 150 mM NaCl in 25 mM Tris-HCl pH 7.5, 0.5% CHAPS and loaded onto a 5-ml HiTrap Heparin column (GE Healthcare). The protein was eluted with a linear 0.15-1 M NaCl gradient; peak fractions were supplemented with 2 mM DTT and NaCl to a final concentration of 1 M. Purified SIVrcm IN, concentrated using a 10-kDa cut-off Vivaspin device (Generon) to 11-15 mg/ml, was supplemented with 10% glycerol, flash frozen in liquid nitrogen and stored at -80°C. Full-length recombinant LEDGF/p75 and HIV-1 IN (wild type (WT) or Q148H/G140S) were initially produced as hexahistidine-tagged forms, which were cleaved and removed during purification according to published procedures (*13, 26*).

Strand transfer assays

Viral DNA (vDNA) modelled on the U5 end of SIVrcm was prepared by annealing synthetic oligonucleotides 5'- GGCTAAGAAAAATCTCTACCA (pre-processed vDNA plus-strand) and 5'- AACTGGTAGAGATTTTTCTTAGCC (non-transferred minus-strand). For activity assays, 0.5 μM vDNA and 300 ng pGEM target DNA were incubated with IN (0 - 8 μM) and LEDGF/p75 (0 - 16 μM) in the presence of 40 mM KCl, 5 mM $MgSO_4$, 4 μM $ZnCl_2$, 5 mM DTT, 20 mM BisTris propane (BTP), pH 6.0 in a final volume of 40 μl. Reactions were allowed to proceed for 45 min at 37°C and stopped by addition of 0.25% sodium dodecyl sulfate (SDS) and 12.5 mM EDTA. Following ethanol precipitation, reaction products were separated by electrophoresis through 1.5% agarose gels in tris-acetate and visualized with ethidium bromide (fig. S2).

INSTIs

DTG and BIC were purchased from MedChem Express, and analogue **1** was synthetized by Giotto Biotech. For use in scintillation proximity assays (SPAs) (Fig. 3A), DTG and analogue **1** were labeled via tritium gas exchange on a hydrogen transfer catalyst to respective specific activities of 8.4 and 12 Ci/mmol and purified to radiochemical purity of >98% (Moravek).

Preparation of the intasomes

SIVrcm (WT or A119D) were assembled by incubating 8 μM IN and 8 μM LEDGF with 4 μM vDNA or 0.75 μM strand transfer product DNA (for assembly of the post-catalytic strand transfer complex) in 55-61 mM NaCl, 3 mM CaCl$_2$, 1 mM DTT, 10 μM ZnCl$_2$ and 25 mM BTP - HCl, pH 6.0 at 37°C for 10 min. The mixtures, cooled on ice for 2 min, were supplemented with NaCl to a final concentration of 350 mM, and incubated for further 10 min on ice to dissociate unproductive protein-DNA aggregates. The assembly products, concentrated on a 10-kDa cut-off Vivaspin device to a final volume of 50 μl, were separated by size exclusion chromatography through a 300 x 7.8 mm TSKgel UltraSW aggregate column (Tosoh Bioscience, part number 0022856) in 350 mM NaCl, 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) - NaOH, pH 7.0. Pre-processed vDNA was prepared by annealing oligonucleotides 5'-biotin-tetraethyleneglycol (TEG) - GTTCTAGAAGGCTAAGAAAAATCTCTACCA and 5'- AACTGGTAGAGATTTTT CTTAGCCTTCTAGAAC. Crossover DNA mimicking the product of strand transfer was made by annealing oligonucleotides 5'- biotin - TEG - GGCTAAGAAAAATCTCTAC CAggccccacgctgcctcgggtagggg, 5'- biotin - TEG - GGCTAAGAAAAATCTCTAC CAgggcccacgctgcctcgggtagggg, 5'-CCCCTACCCGAGGCAGCGTG, and 5'-AACTGGT AGAGATTTTTCTTAGCC (upper case characters represent SIVrcm U5-derived sequence, lower case characters represent the target DNA portion of the crossover construct). Addition of the 5' biotin modification to non-reactive vDNA ends appeared to enhance intasome formation in some cases. Synthetic HPLC-grade oligonucleotides for intasome assembly were obtained from Sigma-Aldrich. The INSTI complexes for cryo-EM were formed by incubation of the purified SIVrcm A119D intasome in the presence of 5 mM MgSO$_4$ and 50 μM DTG or BIC at room temperature for 15-20 min and then on ice for a further 30-60 min prior to freezing grids for cryo-EM. As alternative to size exclusion chromatography, the SIVrcm strand transfer intasome complexes were purified by ultracentrifugation through 2-ml 15-55 % (w/v) sucrose; the gradients were prepared manually by overlaying 400 μl 15, 25, 35, 45 and 55% sucrose in 25 mM 350 mM NaCl, HEPES-NaOH, pH 7.0 and incubating at room temperature for 3-4 h to allow diffusion. The intasome assembly mixture overlaid on top of the gradient was spun in a SW55Ti rotor at 42,000 rpm at 4°C for 16.75 h, and 77-μl fractions taken from the top of the gradient were processed for negative stain EM (fig. S22).

To detect strand transfer activity of the HIV-1 or SIVrcm intasomes, 19.5 μl of each size exclusion chromatography fraction was supplemented with an equal volume of 10 mM DTT, 8 μM ZnCl$_2$ and 25 mM BTP pH 6.0 (resulting in final NaCl concentration of 175 mM). The reactions were initiated by addition of 300 ng pGEM target DNA and 10 mM MgSO$_4$ in a final reaction volume of 40 μl. Samples were incubated at 37°C for 1 h and stopped by addition of 0.25% SDS and 12.5 mM EDTA. Reaction products, deproteinized by digestion with proteinase K and ethanol precipitation, were separated by agarose gel electrophoresis and visualized with ethidium bromide (fig. S3).

Electron microscopy
For negative stain analyses, 4-μl drops of freshly assembled and purified SIVrcm or HIV-1 intasomes were spotted on carbon-coated 300-mesh copper grids (EM Resolutions, catalogue #C300Cu), which were glow-discharged for 30 sec at 45 mA using a K100X instrument (EMS). Excess sample was blotted and absorbed particles were stained with 2% uranyl acetate. Grids were imaged on a Tecnai G2 Spirit LaB6 transmission 120-keV electron microscope (FEI) with an Ultrascan-1000 camera (Gatan) at 30,000x nominal magnification, resulting in a magnified pixel size of 3.45 Å, with total electron exposure of 35 e/Å^2 and

defocus range of 0.5-2.0 μm. A total of 1,376 micrographs were collected from negatively stained SIVrcm intasomes partially purified by velocity centrifugation through sucrose gradients (fig. S22) and 469 micrographs from HIV-1 intasomes isolated by size exclusion chromatography (fig. S3B).

Initial cryo-EM datasets were collected on SIVrcm intasomes spotted on carbon-coated grids under conditions previously optimized for the maedi-visna virus intasome (*14*). Briefly, ultrathin carbon-coated lacey grids (Ted Pella, catalogue #01824) were pre-treated by glow discharging for 1 min at 45 mA in the K100X instrument. Four μl intasome sample ($A_{260\,nm}$ = 0.2) freshly isolated by size exclusion chromatography was spotted on a grid at 22°C and 95% humidity. After 1 min incubation, the grid was blotted (Agar Scientific blotting paper, catalogue #47000-100) on both sides for 0.5 sec, followed by immediate application of 4 μl wash buffer (200 mM NaCl, 25 mM HEPES-NaOH, pH7.0) and second blotting on both sides for 2-3 sec prior to freezing in liquid ethane using a Vitrobot Mark IV instrument (FEI). Blot force was set to -1 for all samples. The movies were recorded on a Tecnai G2 Polara (FEI) microscope operating at 300 kV and equipped with a K2 Summit direct electron detector (Gatan) with a total electron exposure of 50 e/$Å^2$ over 40 frames (fig. S23A).

For the final high-resolution reconstructions, Apo SIVrcm intasome or its complexes with DTG or BIC, freshly isolated by size exclusion chromatography ($A_{260\,nm}$ = 0.2 - 0.7), were spotted on C-flat holey carbon grids (EMS, catalogue #CF-1.2/1.3-4Au). The use of fresh C-flat grids and omitting glow discharging was essential to observe intasomes in open holes. Following 1-min incubation at 22°C and 95% humidity, the grids were blotted (Agar Scientific blotting paper, catalogue #47000-100) on both sides for 1-2 sec and plunged into liquid ethane using Vitrobot Mark IV. Data were collected on Titan Krios microscope operating at 300 keV and equipped with a Gatan post-column energy filter, selecting a 20-eV window, on a K2 Summit direct electron detector (Gatan). A total of 8,311 (Apo SIVrcm intasome), 6,961 (DTG-bound SIVrcm intasome), 11,846 (BIC-bound SIVrcm intasome), and 6,775 (BIC-bound Q148H/G140S SIVrcm intasome) movies were recorded using single electron counting mode with a defocus range of 1.6-3.5 μm. The total electron exposure was 50.4 e/$Å^2$ spread over 30 frames for datasets with a pixel size of 1.38 Å (Apo intasome, DTG-intasome, and BIC-Q148H/G140S intasome) or 55.7 e/$Å^2$ over 40 frames for the pixel size 1.08 Å (BIC-intasome) (table S1, fig. S24A). The movies were aligned with dose weighting, as implemented in MotionCor2 (*27*), and the contrast transfer function (CTF) parameters were estimated from the frame sums using Gctf-v1.06 (*28*). After discarding images with ice contamination, the datasets contained the totals of 8,027 (Apo SIVrcm intasome), 6,838 (DTG-bound SIVrcm intasome), 11,769 (BIC-bound SIVrcm intasome), and 6,618 (BIC-bound Q148H/G140S SIVrcm intasome) movies.

Single-particle image processing and structure refinement
Negatively stained intasome particles were picked using semi-automatic mode of EMAN boxer (*29*) and subjected to reference-free 2D classifications in Relion-2.1 (*30*). For cryo-EM image processing, SIVrcm intasome particles were initially picked along with linear polymer assemblies using EMAN boxer on a subset of micrographs of carbon-coated lacey grids. These were subjected to reference-free 2D classification in Relion-2.1. The resulting 2D classes, low-pass filtered to 20 Å resolution, were used as references to pick particles on 1,989 micrographs with relion_autopick. Following three rounds of reference-free 2D classifications in Relion-2.1, 71,983 particles were selected for a 3D classification. An *ab initio* reconstruction, created using cryoSPARC (*31*), was used as a reference for 3D

classification without symmetry and classified into three classes in Relion-2.1. Refinement of the best class containing 21,108 particles and with C2 symmetry applied to the CIC assembly produced the initial reconstruction at a sub-nm resolution (~8.5 Å) (fig. S23B). These results encouraged us to further optimize sample preparation. Strong density in the target DNA pose (fig. S23B) suggested that a considerable fraction of the intasomes captured DNA from the excess of vDNA present in the assembly reactions. To avoid competition between DNA and INSTIs for the intasome active sites, the A119D variant of SIVrcm IN was used for all further cryo-EM experiments. The mutation in orthologous retroviral INs specifically blocked target DNA capture without affecting IN active site function (*16-18*).

Datasets of SIVrcm intasomes collected on C-Flat holey carbon grids were exhaustively picked with Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/) and relion_autopick using two sets of low-pass filtered 2D classes as templates. Particle sets were *in-silico* purified by two consecutive rounds of reference-free 2D classification in cryoSPARC, and those belonging to well-defined classes (fig. S24B) were subjected to a crude 3D classification in Relion into four classes without applying symmetry (fig. S25). At this point, relion_autopick and Gautomatch subsets contributing to the best 3D class were combined, and overlapping particles were pruned using the dedicated script from the LocalRec package (*32*) with proximity cut-offs of at least 35 Å (25 or 40 pixels depending on the pixel size). This procedure resulted in 829,709 (Apo intasome), 739,798 (DTG-intasome), 675,894 (BIC-intasome) and 866,839 (BIC-Q148H/G148S intasome) particles. 3D reconstructions from these particles were used to estimate positions and pick neighboring intasome repeat units along the stacks (fig. S25) by multiplying the translation vectors relating central (picked) repeat unit to the neighboring (potentially unpicked) units by the Euler matrix specific to each refined particle.

The translational symmetry-expanded Apo intasome dataset, pruned for overlaps and containing 2,061,407 particles, was subjected to 3D classification into eighteen classes without imposing symmetry (fig. S25). Combination of the six highest-resolution 3D classes (boxed in the figure) yielded 644,674 particles. At this point, local defocus parameters determined with Gctf were added, and the particles were used in a 3D reconstruction imposing 2-fold symmetry. The particles were then sorted into two 3D classes without realignment in Relion and with a soft mask focusing on a single intasome stack unit (comprising IN octamer plus associated vDNA); 171,893 particles contributing to the high-resolution class were then used for the final non-uniform 3D refinement as implemented in cryoSPARC, using default parameters (fig. S25).

Similar workflows were used to determine drug-bound SIVrcm intasome structures. The expanded DTG-intasome dataset containing 1,843,630 particles was first subjected to a crude 3D classification in Relion into four classes, and the best class containing 825,295 non-overlapping particles was further sorted into twelve 3D classes; 530,679 non-overlapping particles contributing to the best nine 3D classes were taken for per-particle motion correction (polishing) as implemented in Relion. Following 3D reconstruction, using locally determined CTF parameters, the dataset was further sorted into two 3D classes without re-alignment in Relion and with a soft mask focusing on one stack unit; the 178,295 particles contributing to the high-resolution class were used for the final non-uniform 3D refinement in cryoSPARC, using default parameters.

The expanded BIC-intasome dataset containing 1,415,850 particles was subjected to 3D classification in Relion into twelve 3D classes, and the best six classes collectively containing 643,721 non-overlapping particles were taken for polishing. Following 3D reconstruction, using locally determined CTF parameters, the dataset was sorted into two 3D classes without

realignment in Relion imposing a soft mask focusing on one stack unit; 149,779 particles contributing to the high-resolution class were used for the final non-uniform 3D refinement in cryoSPARC, using default parameters. Map sharpening and filtering was done in cryoSPARC, and Blocres (*33*) was used for local resolution estimation (fig. S8). Q148H/G140S SIVrcm intasomes assembled less efficiently forming shorter stacks, and translation symmetry expansion did not increase the number of particles. Following a 3D reconstruction using 866,839 particles of the BIC-Q148H/G140S intasome complex derived from the initial 3D classification, the dataset was sorted into two 3D classes without realignment imposing a soft mask focusing on one stack unit, and 213,635 particles contributing to the high-resolution class were used in the final non-uniform refinement in cryoSPARC, using default parameters and locally estimated determined CTF parameters.

Gold standard FSC = 0.143 criterion (*34, 35*) was used to estimate resolutions of the 3D reconstructions (fig. S7, table S1). Heat maps of angular distributions of refined particles for the reconstructions are shown in fig. S26. The final cryo-EM maps lacked anisotropic features, and 3D FSC sphericity values (*36*) ranged between 0.960 and 0.975 (table S1).

The atomic models, initially assembled by rigid docking of partial and homologous structures (PDB IDs 1EX4, 1K6Y, 2B4J, and 5U1C) (*9, 37-39*), were manually fitted into sharpened cryo-EM maps and extended in Coot (*40*). Water molecules discernible in half-maps were included in the refinement (fig. S27); strong cryo-EM density peak underlying side chain of Tyr143 was tentatively interpreted as a Cl^- anion. The final refinement was done using the real-space engine from Phenix software package (*41*) using default restraints and automatic weight settings; to preserve 2-fold symmetry of the assembly, non-crystallographic symmetry constraints for the two halves of the intasome were enforced during refinement. Quality of the final models were assessed using MolProbity (*42*) and EMRinger (*43*) (table S1).

<u>Single-cycle HIV-1 and SIVrcm infectivity assays</u>
The SIVrcm sequence (NCBI entry HM803689) (*10*) from the upstream R region to downstream CAAGT in U5 was synthesized by BioBasic and inserted into pcDNA3.1+ vector DNA at *Bam*HI and *Xho*I restriction enzyme sites; the resulting pSIVrcm-Luc construct harboured a 1,216 bp deletion in the *env* gene that encoded for the surface subunit of the envelope glycoprotein, as well as the substitution of firefly luciferase for the Nef coding region (fig. S11A). To aid mutagenesis, the IN-coding region of pSIVrcm-Luc was flanked with *San*DI and *Bsu*36I restriction sites by introducing silent point mutations into *pol* (region encoding reverse transcriptase (RT)) and *vif*, respectively. This made it possible to shuttle *San*DI-*Bsu*36I DNA fragments encoding IN variants into the virus construct. To allow digestion with *San*DI, the viral DNA construct was isolated from methylation-negative *dam⁻/dcm⁻ E. coli* cells (New England Biolabs). For all other purposes, the DNA was propagated in Stbl3 cells (ThermoFisher Scientific) at 30°C. Single-round WT and Q148H/G140S HIV-1 luciferase reporter constructs and pCG-VSV-G were previously described (*44, 45*). Q148H/G140S DNA was mutagenized by QuickChange II site directed mutagenesis (Agilent). All viral expression constructs used for infectivity assays were verified by sequencing of the entire proviral genomes.

HEK293T cells cultured in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum, 100 IU/mL penicillin, and 100 μg/mL streptomycin (DMEM) at 37 ℃ were transfected with proviral DNA and pCG-VSV-G constructs at the ratio of 9:1 (15 μg total DNA per 10-cm dish) using PolyJet (SignaGen). Culture supernatants harvested 48 h after transfection pre-cleared for cell debris were filtered through 0.45 μm filters and

concentrated by ultracentrifugation at 4 °C for 2 h at 26,000 rpm in a SW32-Ti rotor. Virus pellets resuspended in DMEM were frozen in aliquots and stored at -80°C. Virus yield was quantified by a Taqman-based product-enhanced RT assay (Taq-PERT) as follows. Viruses diluted in water were treated with 2x lysis buffer (0.25% Triton X-100, 50 mM KCl, 100 mM Tris-HCl pH 7.4, 40% glycerol) in the presence of RNase inhibitor. Bacteriophage MS2 genomic RNA was used as the template for the reverse transcription reaction with primers AE7575 (5'-TCCTGCTCAACTTCCTGTCGA) and AE7576 (5'-CACAGGTCAAACCTCCTAGGAATG) and Taq-man probe 6[FAM]CGAGACGCTACCATGGCTATCGCTGTAG[TAMsp] (*46*). RT activity was expressed as mU/mL as determined by standard curves generated using recombinant HIV-1 RT protein (Millipore).

In preliminary experiments (fig. S11A), HEK293T cells (100,000 per well of a 24 well-plate) were infected in duplicate with 0.5 μU RT per cell. After 6-8 h, media was replaced with fresh DMEM, and cell cultures were harvested at 48 h from the start of the experiment. Cells were lysed in passive lysis buffer (Promega) by freezing overnight at -80 °C and thawing at 37 °C for 30 min. Cell lysates were centrifuged at 17,500 g for 8 min, and supernatants were analyzed in triplicate for luciferase activity. Relative light units (RLU), determined by luminometer (Berthold Technologies), were normalized to total protein concentration as measured by bicinchoninic acid protein assay kit (Thermo Fisher Scientific).

For drug sensitivity assays, HEK293T cells (10,000 per well of a 96-well plate) were infected in triplicate with WT or IN mutant SIVrcm or HIV-1 virus at 0.5 μU RT activity/cell in 200 μL DMEM medium with 8-10 concentrations of INSTIs in 4-fold dilution steps that ranged from no drug to 2 μM or 20 μM. At 6-8 h post-infection, the virus-containing media was exchanged for fresh media. At 48 h post-infection, cells were washed twice with phosphate-buffered saline and lysed in 5X passive lysis buffer (Promega) by freezing at -80 °C for 30 min and thawing at 37 °C for 20 min. Insoluble material was pelleted by centrifugation, and luciferase activity was assessed for supernatants. EC_{50} values were calculated from two independent experiments by fitting to a dose-response-inhibition (four parameters) model using GraphPad Prism 8 software. Fold changes in EC_{50} of virus relative to WT control virus were calculated using mean EC_{50} values from two independent experiments.

Tritium SPA assay

INSTI dissociation measurements were done according to published protocols (*21, 47*) with minor modifications. Biotinylated HIV-1 U5 vDNA (0.6 nmol) prepared by annealing synthetic oligonucleotides 5'-Biotin-TEG-CCAGTCAGTGTGGAAAATCTCTAGCA and 5'-ACTGCTAGAGATTTTCCACACTGACTGG was immobilized on 10 mg streptavidin-coated polyvinyl toluene beads (Perkin Elmer). The beads were washed and resuspended in 50 mM NaCl, 10 mM $MgSO_4$, 25 mM HEPES-NaOH, pH 7.0 in a total volume of 200 μl, and 150 μl suspension was diluted with 2.1 ml SPA buffer (100 mM NaCl, 25 mM $MgSO_4$, 13 mM DTT, 6% polyethylene glycol 400, 0.06% Nonidet P40, 30 mM HEPES-NaOH, pH 7.0). Bead suspensions were supplemented with 0.5 μM WT or Q148H/G140S HIV-1 IN. Upon incubation at 37°C for 80 min to form the IN-vDNA complexes, the beads were washed and resuspended in equivalent volume of SPA buffer prior to addition 0.5 μM ^{3}H-labeled DTG or analogue **1**. The inhibitors were allowed to bind overnight at room temperature. The following morning beads were transferred into white 96-well OptiPlate plates (Perkin Elmer) at 100 μl suspension per well. Dissociation was induced by addition of 100 μM unlabelled

compound, and the scintillation signal was measured using a TopCount instrument (Perkin Elmer). For consistency, the plates were kept at room temperature for the entire length of the experiment and gently mixed to suspend the SPA beads prior to each measurement. Signals from duplicate wells were averaged, and signals measured from protein-free controls were subtracted.

Molecular Dynamics (MD)
MD simulations were performed for four systems: the wild type and Q148H/G140S SIVrcm intasome bound to two types of ligands: BIC and analogue **1**. The initial models were assembled based on the experimental structures determined in this work, using N-terminal acetyl and C-terminal amide capping groups for missing residues. The ligands were parametrized using the general Amber force field (GAFF) (*48*). The atomic partial charges were obtained from electrostatic potential calculations (*49*) at the level of density functional theory (DFT) ωB97X-D/def2TZVPP as implemented in Gaussian 09 Revision E (*50-52*). The system was solvated by 100,000 - 120,000 TIP3P water molecules depending on the ligand and mutations, and neutralized with Na^+ and Cl^- ions. From the original starting cryo-EM structure, water molecules directly involved in the metal coordination were retained. The system was minimized using a standard protocol via steepest descent algorithm for a total number of 150,000 steps, followed by 10 ns equilibration with restrained heavy atoms (heavy atom of the backbone of the protein and the nucleic with an isotropic force of 1000 kJ $mol^{-1}nm^{-1}$) in constant number pressure and temperature (NPT) and constant number volume and temperature (NVT; up to 1 ns) at 300 K via standard MD procedure using NAMD 2.12 (*53*). A series of unbiased MD simulations were performed for the four systems (wild type and G140S/Q148H IN bound with BIC or analogue **1**) to obtain multiple independent trajectories. A total of 10 replicas were run for all systems, and each simulation was 100 ns long.

Quantum Mechanics/Molecular Mechanics (QM/MM)
The experimentally resolved structures were subject to QM/MM minimizations to obtain geometry-optimized structures of the active site at the *ab initio* DFT level and assess the quantitative response of the ligand bound structures to the mutations. The MM region was described by CHARMM36 force field (*54*) while the hybrid functional B3LYP (*55*) with the GTO basis set of 6-31G* (*56-58*) was utilized for the active site, as implemented in CHARMM (*59*) and Q-Chem 4.3 (*60*), respectively. QM/MM calculations were performed in a non-periodic fashion, with atoms further than 25 Å from the ligand removed while those between 20 and 25 Å were fixed. Population analysis was carried out for the QM region embedding the electrostatics of the MM atoms (*61*), using natural bond orbital (NBO) 3.1 scheme (*62, 63*), as implemented in Gaussian 09 Revision E.
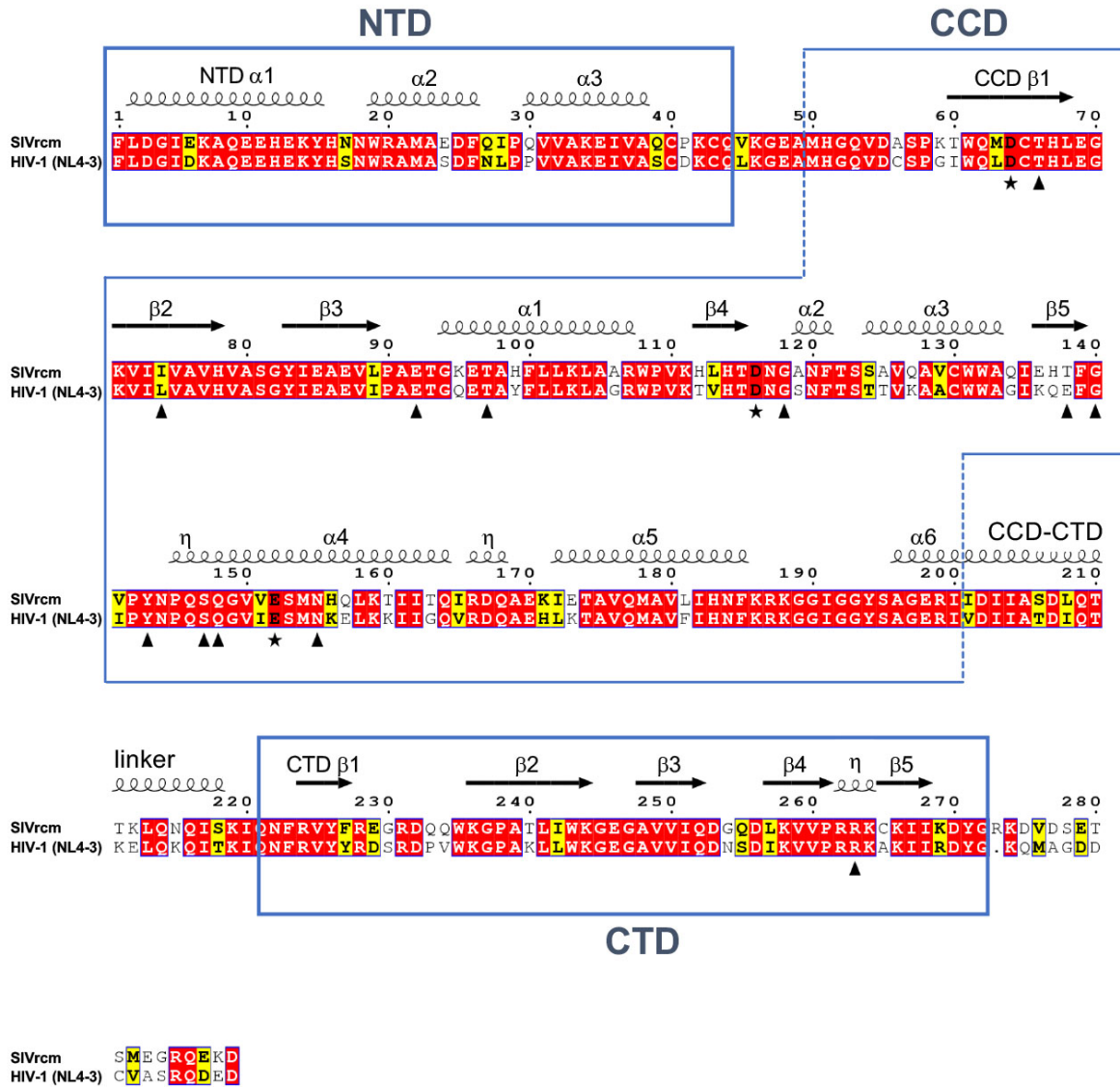
Fig. S1: Amino acid sequence alignment of SIVrcm and HIV-1 INs. Boxes highlight IN domains (NTD; catalytic core domain, CCD; and CTD); secondary structural elements are indicated atop the alignment. Filled triangles and stars respectively indicate positions corresponding to mutations associated with INSTI resistance and residues comprising the active site (metal-chelating) DDE triad. Red highlights positions of amino acid identity; yellow, chemically-related amino acids. Shown is laboratory-derived HIV-1 subtype B strain NL4-3 (NCBI entry U26942.1). Percent SIVrcm IN amino acid identity is similar to circulating HIV-1 subtypes, such as clade C transmitter-founder strain Z331M (accession code KR820323; 75.0% identical) and clade A1 strain Q23-17 (AF004885; 74.7% identical). Polymorphic position 74, substitutions at which are associated with increased INSTI resistance, is frequently occupied by Ile in HIV-1 isolates from INSTI-naïve patients (https://hivdb.stanford.edu/dr-summary/resistance-notes/INSTI/). The figure was prepared using ESPript software (*64*).
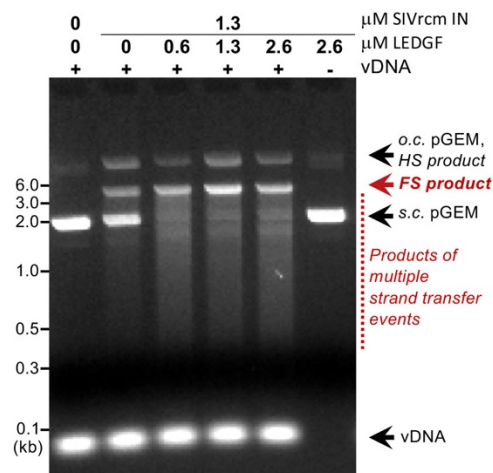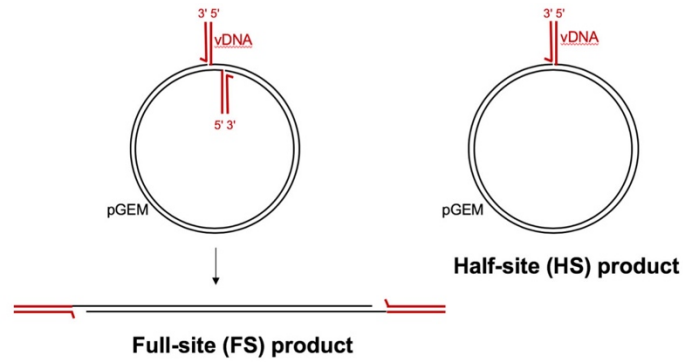
Fig. S2: Strand transfer activity of SIVrcm IN *in vitro*. Schematic of the reaction products (top). Full-site product (left) is generated by insertion of a pair of viral DNA (vDNA) ends, mimicked by oligonucleotides corresponding to pre-processed U5 vDNA (red lines) into opposing strands of target DNA (pGEM plasmid DNA, black lines). Half-site product (right) is generated by insertion of a single vDNA end. Analysis of strand transfer products in the absence or presence of increasing concentrations of LEDGF/p75 by agarose gel electrophoresis, detected by staining with ethidium bromide. The left- and the right-most lanes contained mock samples prepared in the absence of proteins and vDNA, respectively; the latter control verified the absence of nuclease contamination in recombinant IN and LEDGF/p75 protein preparations. Input concentrations of SIVrcm IN and LEDGF/p75 are shown above the gel. Migration positions of DNA size markers and specific reaction products are indicated (o.c., open circular; s.c., supercoiled). Products of multiple strand transfer events appear as a smear and become prominent when circular target DNA is expended.
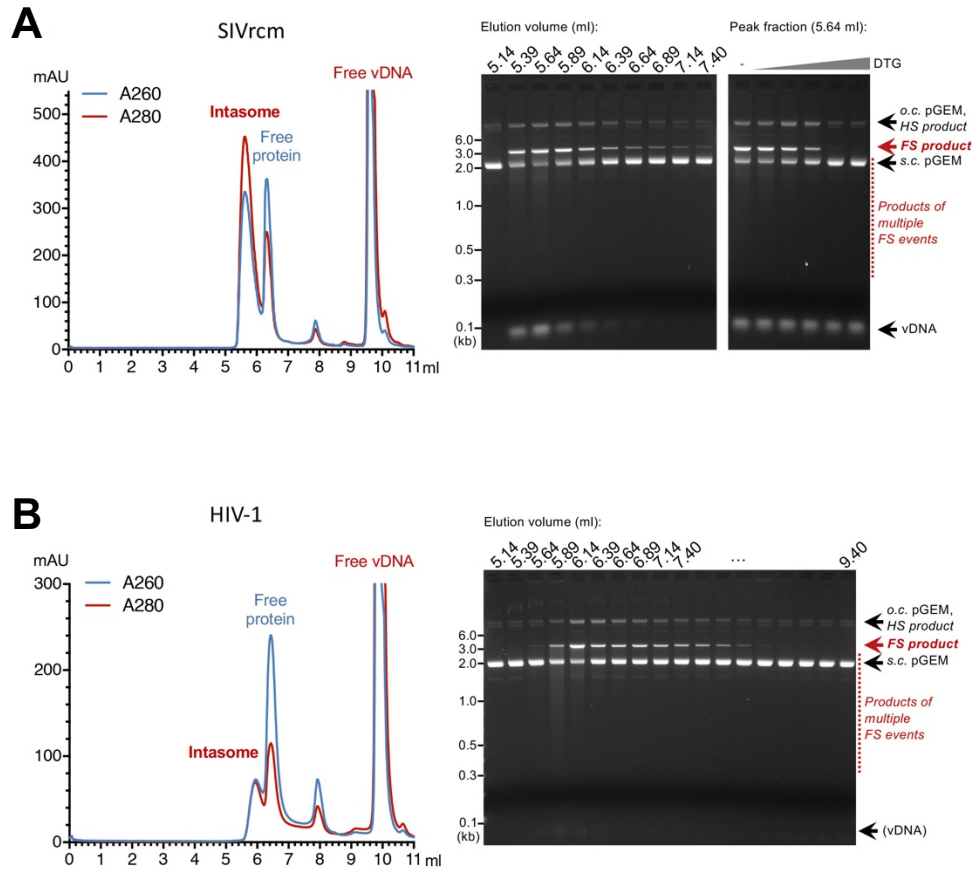
Fig. S3: Assembly and activity of SIVrcm (A) and HIV-1 (B) intasomes. Intasome assembly mixtures were separated by size exclusion chromatography (left) and specific fractions were tested for strand transfer activity using supercoiled pGEM plasmid target DNA (right). Elution positions of the intasomes, free protein, and vDNA oligonucleotides are indicated above the chromatograms. In addition, the peak fraction of the SIVrcm intasome with elution volume 5.64 ml was tested in the absence or presence of DTG (from left to right: 0, 10, 50, 100, 500, 1000 nM). Chromatograms are representative of n>10 and n=3 SIVrcm and HIV-1 intasome preparations, respectively; agarose gel images represent minimally two independent experimental replicates.
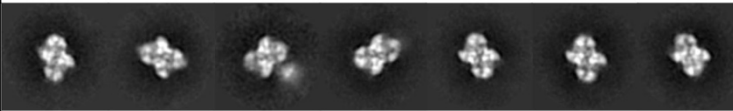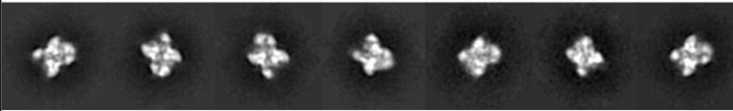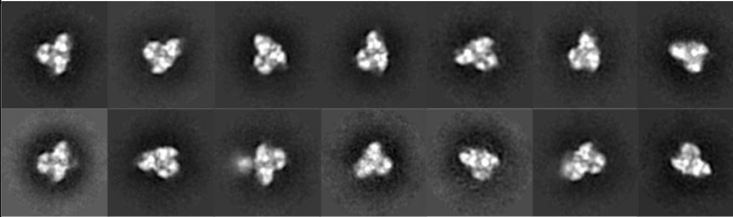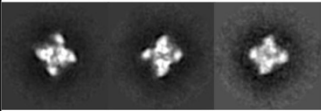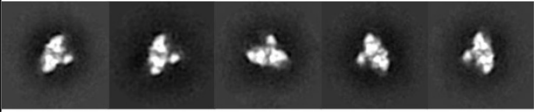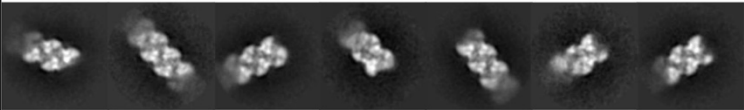
Fraction of total (excluding side views, %)	Fraction of single intasomes (%)	Apparent number of IN subunits	Single SIV intasomes:
75	19	16	
	21	14	
	32	12	
	7	12	
	21	10	
25	-	>16	**Stacks:**

Fig. S4: Distribution of negatively stained SIVrcm intasomes between 2D classes. SIVrcm intasomes were fractionated by sedimentation through sucrose gradient (fig. S22). Particles eluting at 1.5 ml of the gradient were imaged by negative-stain EM and subjected to reference-free 2D classification. Percentages of particles belonging to single and stacked intasome classes are given in the first and second columns; the apparent number of IN subunits in the classes is shown in the third column. The corresponding 2D classes are shown to the right.
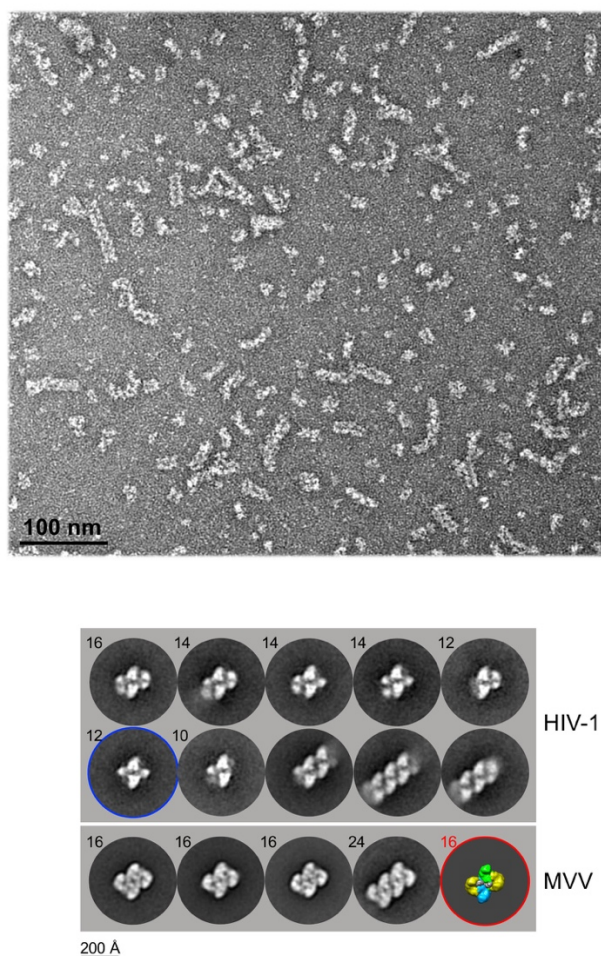
Fig. S5: Negative-stain EM of HIV-1 intasomes. A raw micrograph of HIV-1 intasomes isolated by size-exclusion chromatography and stained with uranyl acetate (top) and examples of 2D class averages (bottom). Particle distributions between 2D classes are given in fig. S6. One of the class averages that contains 12 apparent IN protomers (lower left, blue circle) resembles the dodecamer assembly reported previously for a hyperactive HIV-1 IN-Sso7d fusion protein (*9*). For comparison, 2D classes of MVV intasomes (*14*) are shown at the bottom; apparent numbers of IN subunits are indicated for non-stacked classes. Low-pass filtered envelope of the hexadecameric MVV intasome structure (red circle) is shown for comparison with central and flanking IN tetramers in blue/green and yellow, respectively.
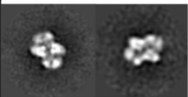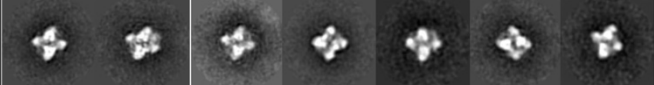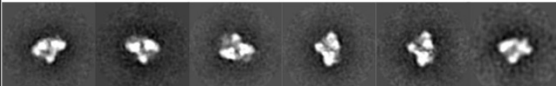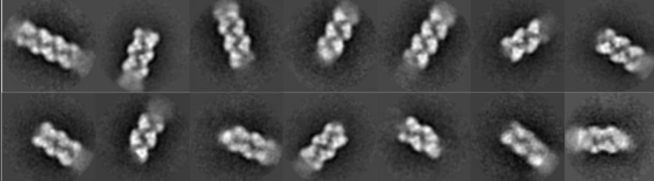
Fraction of total (excluding side views, %)	Fraction of single intasomes (%)	Apparent number of IN subunits	
			Single HIV intasomes:
42	7	16	
	37	14	
	5	12	
	27	12	
	24	10	
			Stacks:
58	-	>16	

Fig. S6: Distribution of negatively stained HIV-1 intasome between 2D classes. HIV-1 intasomes isolated by size-exclusion chromatography (fig. S3B) were imaged by negative-stain EM and subjected to reference-free 2D classification. Percentages of intasome particles belonging to single and stacked intasome classes are given in the first and second columns; the apparent number of IN subunits in the classes is shown in the third column. The corresponding 2D classes are shown to the right.
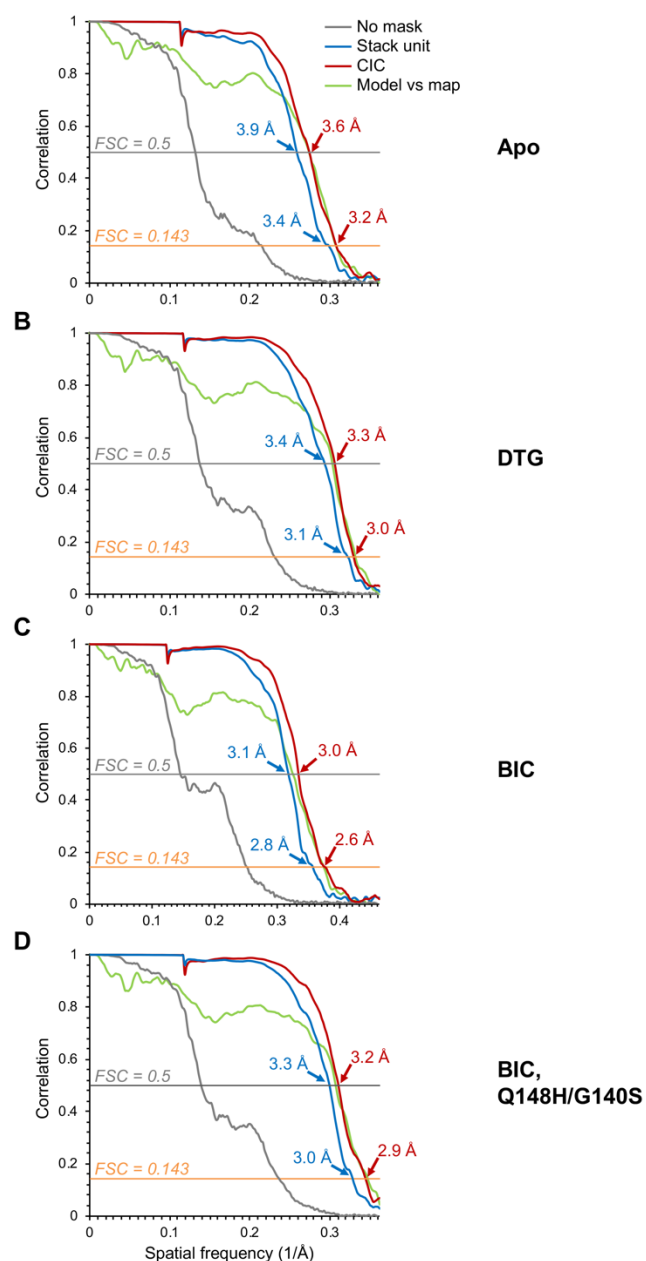
Fig. S7: Fourier shell correlation of maps from independent particle half-sets for Apo, DTG-bound, BIC-bound WT SIVrcm intasome, and BIC-bound Q148H/G140S SIVrcm intasome. Correlations were calculated without or with soft-edge masks around a single stack unit, or for the CIC (grey, blue, and red lines, respectively); corresponding resolutions at the correlation thresholds of 0.143 and 0.5 are indicated. Green lines represent FSCs between refined models and the corresponding cryo-EM maps.
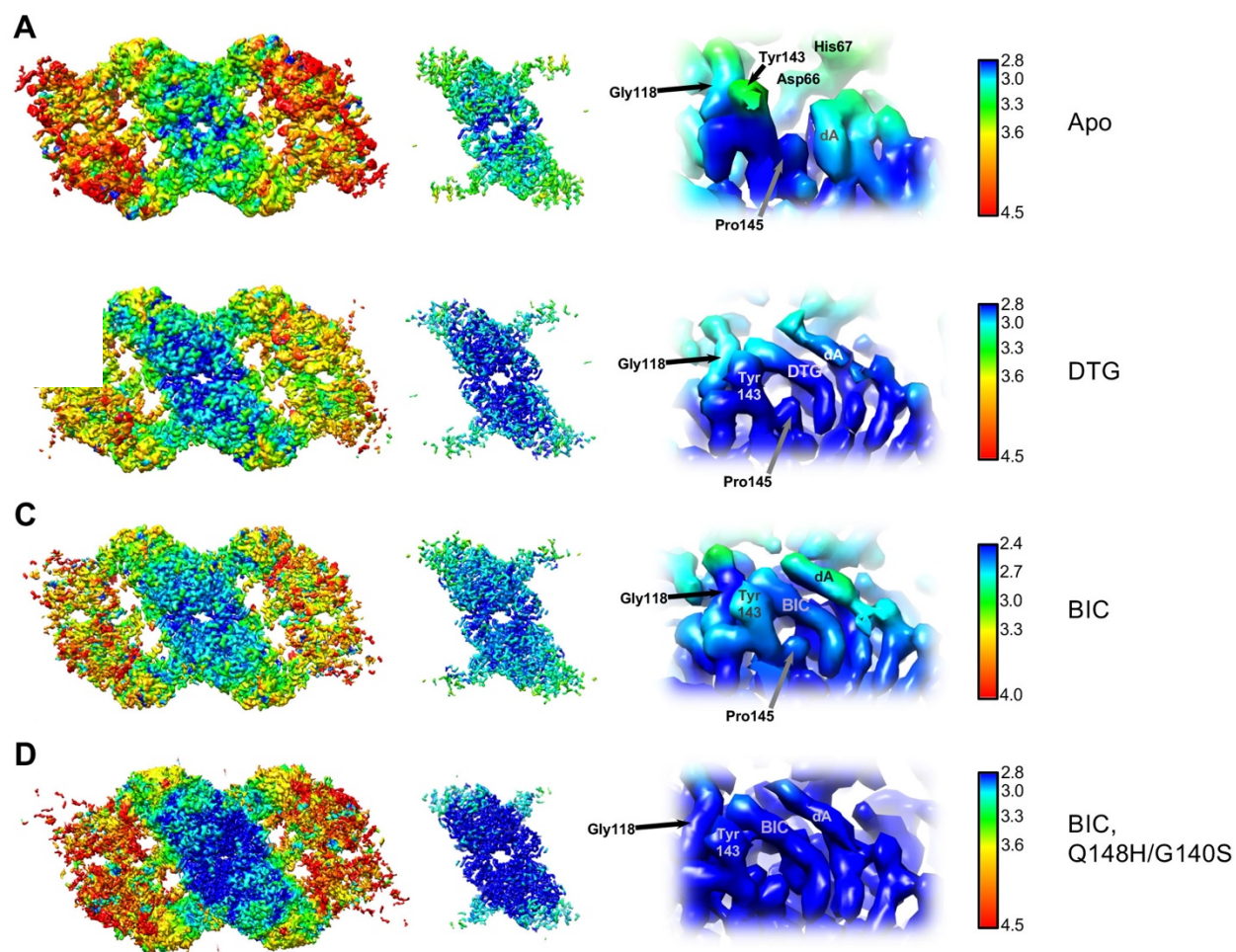
15

Fig. S8: Local resolution maps of Apo, DTG-, and BIC-bound WT SIVrcm intasomes and BIC-bound Q148H/G140S SIVrcm intasome. The left and middle images show the full map at a low and a high-density threshold, respectively. Views focused on the active sites with color scale are shown to the right. Locations of selected protein residues, 3' vDNA nucleotide (dA), and drug molecules are indicated.
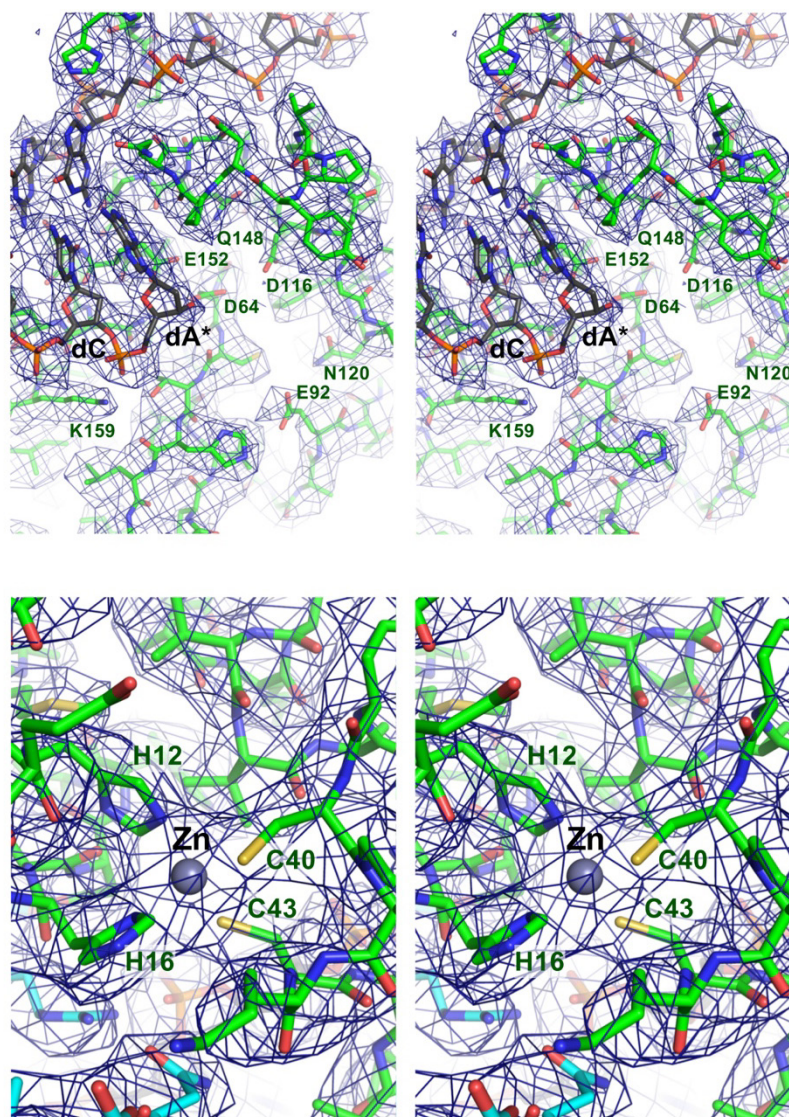
Fig. S9: Examples of cryo-EM density of the Apo SIVrcm intasome. Stereo views (wall-eye format) of IN chain A active site region and Zn-binding cluster of the NTD are shown on top and bottom, respectively. Cryo-EM density is shown as blue chicken wire. Protein and DNA residues are shown as sticks with carbon atoms in green and grey, respectively; the grey sphere represents a Zn^{2+} ion coordinated by the invariant His and Cys residues of the NTD. Selected amino acid residues and the nucleotides comprising the invariant dCdA 3' dinucleotide of the vDNA plus-strand are indicated on the stereo-views; 3' deoxyadenosine, which is displaced by INSTIs in drug-bound structures, is highlighted with an asterisk.
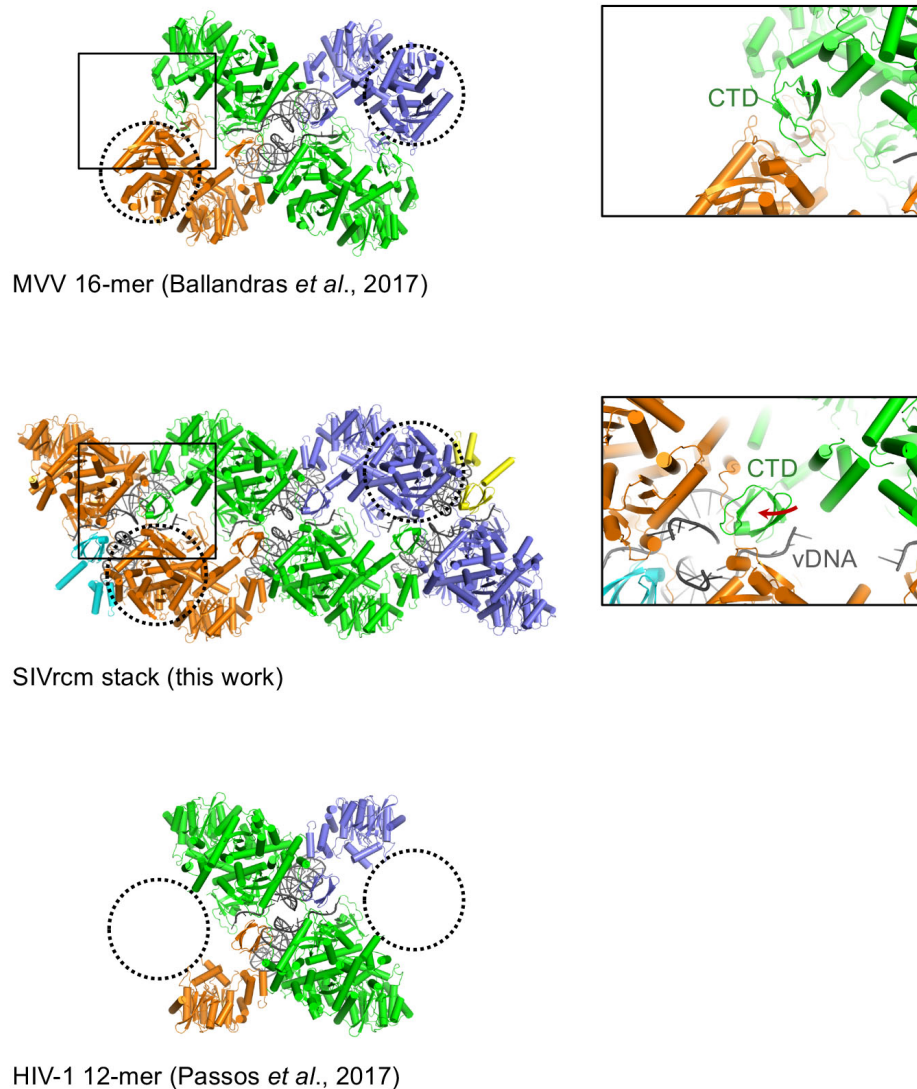
MVV 16-mer (Ballandras *et al.*, 2017)

SIVrcm stack (this work)

HIV-1 12-mer (Passos *et al.*, 2017)

Fig. S10: Comparison of hexadecameric MVV intasome, SIVrcm stack (middle), and dodecameric HIV-1 intasome (bottom) reconstructions. MVV intasome (EMDB-4138, PDB ID 5m0q; top) (*14*), SIVrcm stack (this work; middle), and HIV-1 intasome (EMDB-8483; bottom) (*9*) are shown as cartoons with helices as tubes. The IN tetramers providing the active sites in MVV and HIV intasomes are green, and flanking structures, which donate the synaptic CTDs, are orange and purple; vDNA is grey. The structures are shown in the same orientation and scale. A pair of IN dimers belonging to the flanking tetramers (indicated with dotted circles on the MVV intasome structure) is missing in the reported HIV-1 intasome. Corresponding IN dimers of the flanking repeat units in the case of the SIVrcm stack contribute to the CIC structure. Expansion of the SIVrcm stack is aided by the mobility (red arrow in boxed views to the right) of the two CTDs belonging to an adjacent central tetramer (chains C and K; orange), which allows them to assume synaptic positions within the neighboring stack units. Cyan and yellow are outermost NTDs and CTDs from adjacently stacked leftward and rightward IN-vDNA units, respectively. Note the striking similarity of the common elements among the three structures.
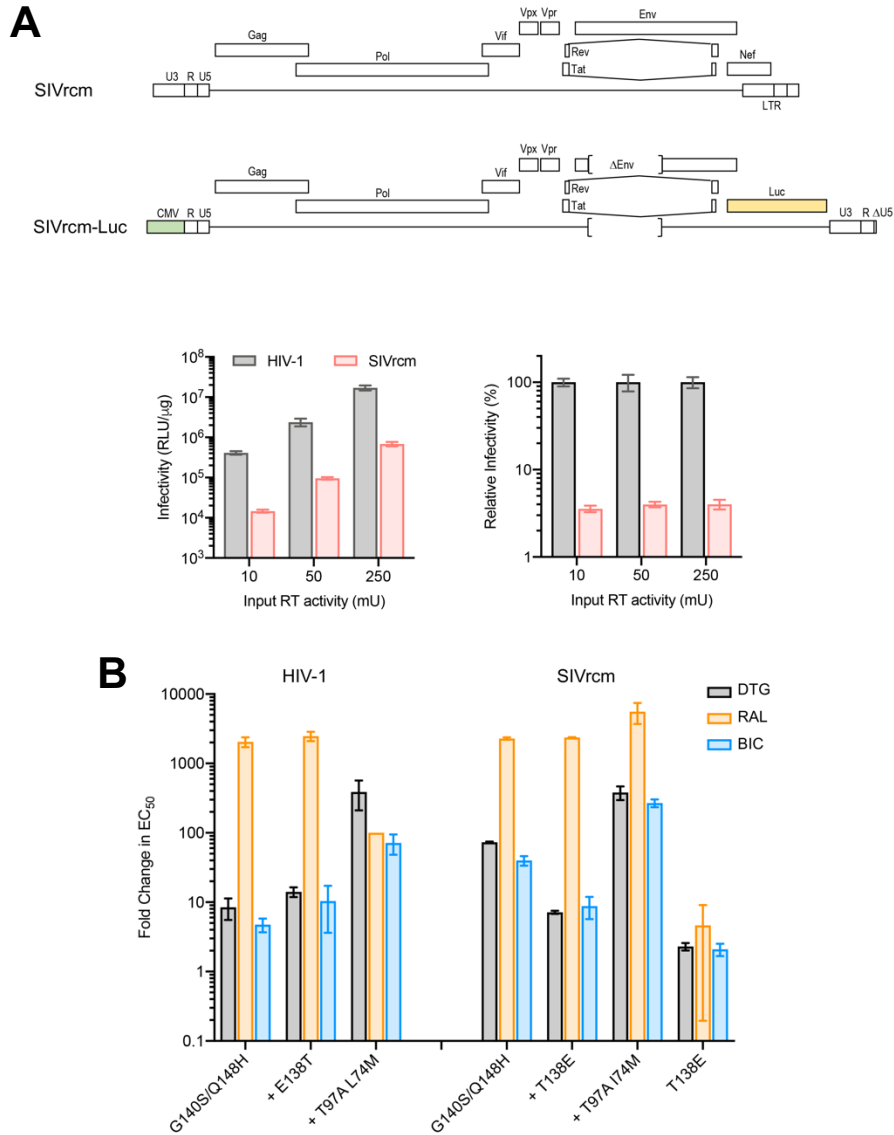
18

Fig. S11: Reconstruction and characterization of SIVrcm recombinant virus. (**A**) To-scale genomic map of pSIVrcm-Luc compared with parental SIVrcm provirus. Heterologous elements (CMV promoter in place of upstream U3; firefly luciferase in place of Nef) are noted. Lower, Infectivity of RT mU-matched SIVrcm-Luc (with silent mutations in *pol* and *vif*) and HIV1-Luc at three different levels of viral inocula. Results are averages and standard deviations of minimally n = 2 independent experiments, with each experiment conducted in duplicate. (**B**) Fold changes in EC_{50} values of noted IN mutants versus respective WT HIV-1 and SIVrcm viruses. Results are averages and standard deviations of n = 2 experiments with each experiment conducted in triplicate. The data for G140S/Q148H/T97A/L74M HIV-1 is from ref. (*7*).
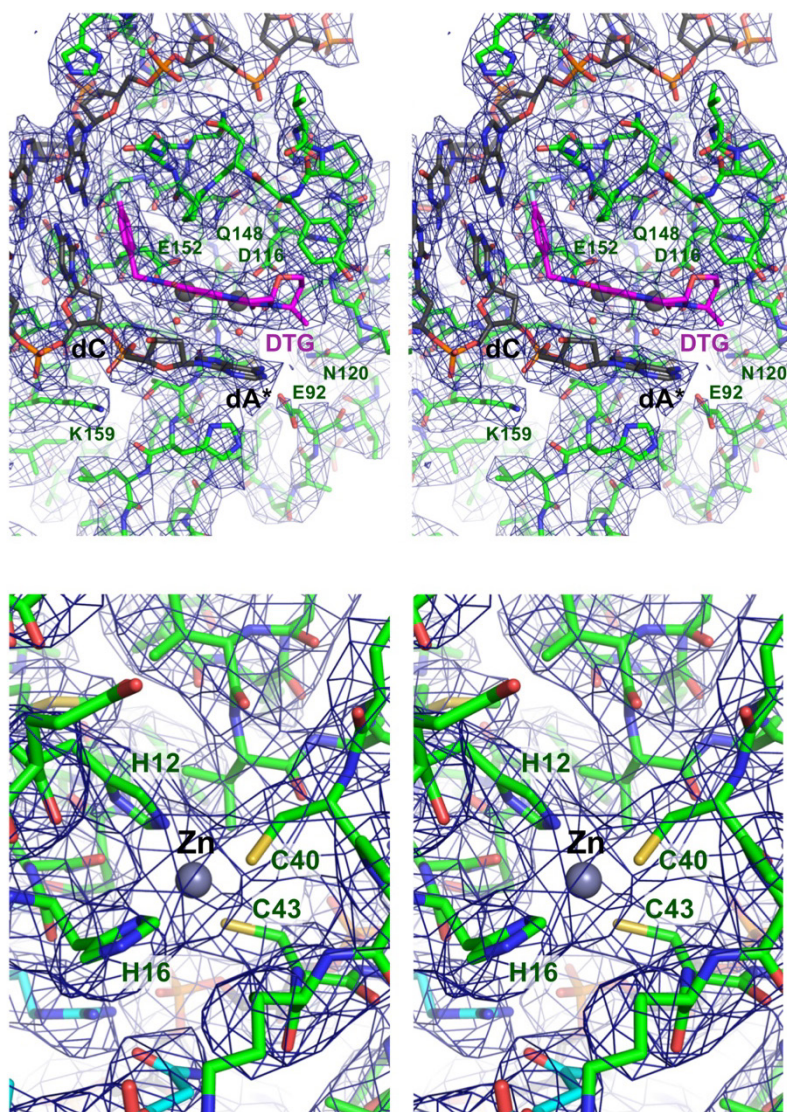
Fig. S12: Examples of cryo-EM density of DTG-bound SIVrcm intasome. Stereo views of IN chain A active site region and Zn-binding cluster of the NTD are shown on top and bottom, respectively. The cryo-EM density and the refined model are depicted and labeled as in fig. S9. In addition, the INSTI (DTG) molecule is shown as sticks with carbon atoms in magenta; red and grey spheres are water molecules and metal ions, respectively.
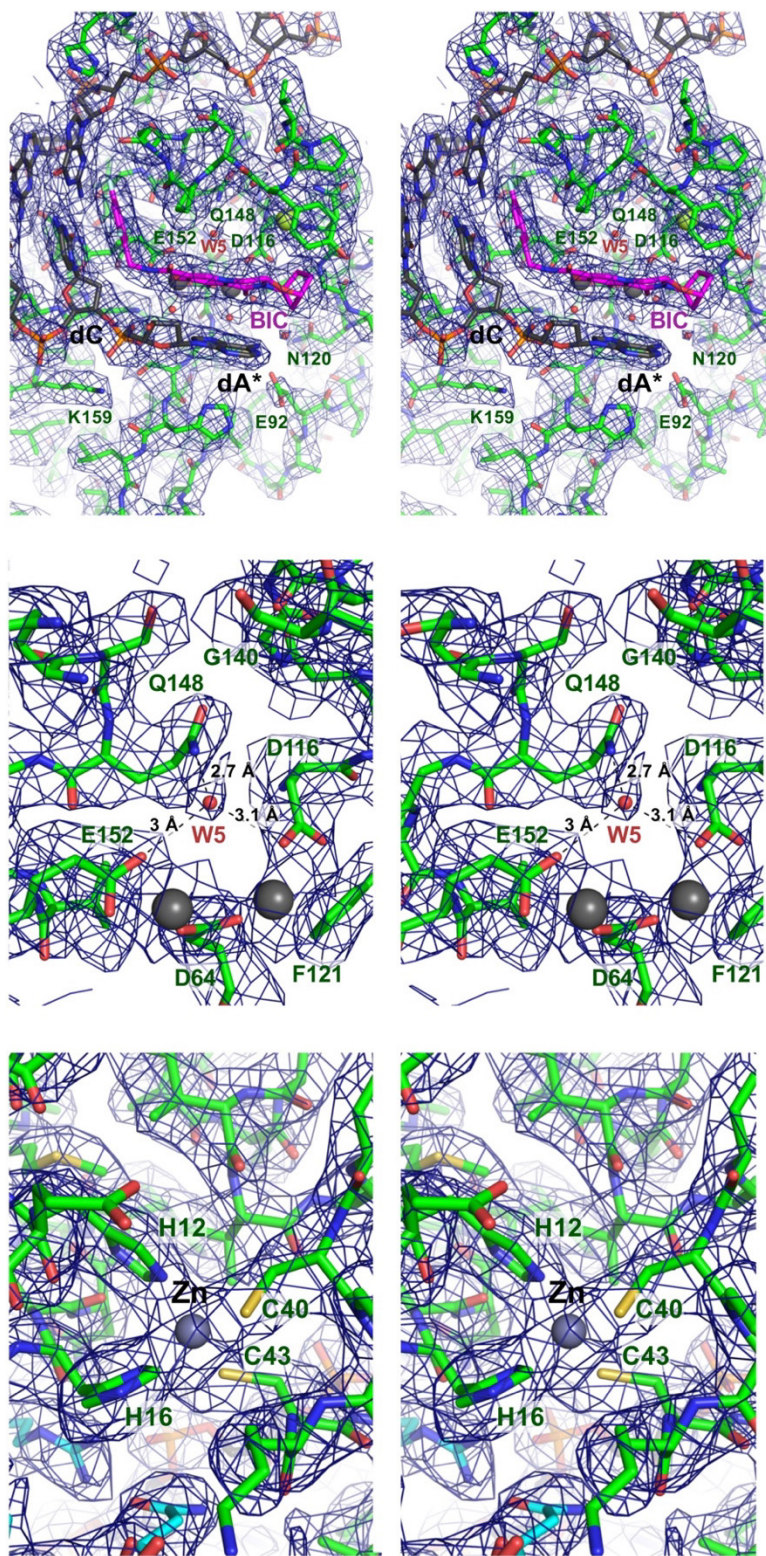
Fig. S13: Examples of cryo-EM density of BIC-bound SIVrcm intasome. Stereo views show IN chain A active site region (top and middle) and Zn-binding cluster of the NTD (bottom). The cryo-EM density and the refined model are depicted as in figs S9 and S12. The key W5 water molecule with corresponding distances is shown in the middle stereo view. See fig. S21 for a view focusing on Ile74 and Thr97.
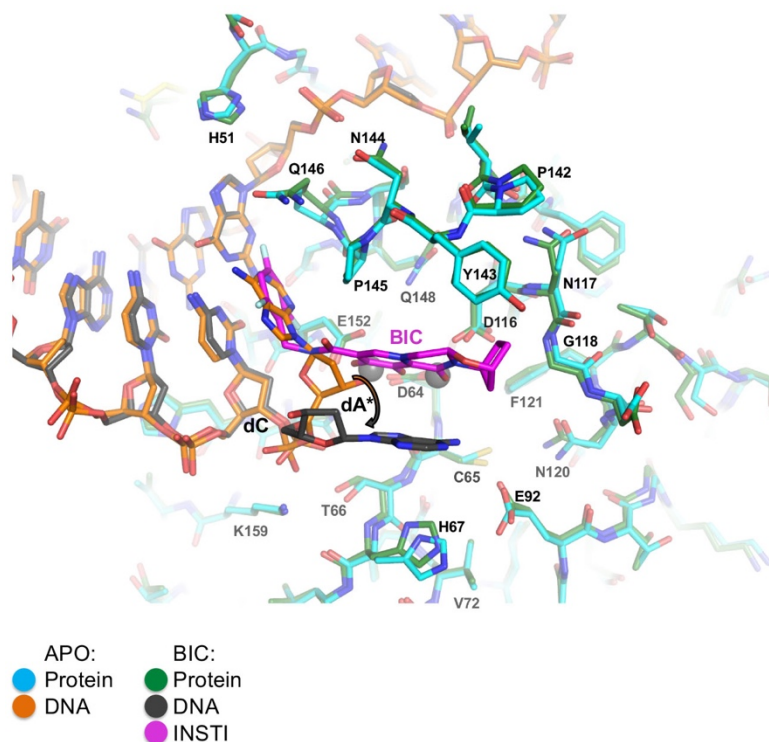
21

Fig. S14: Rearrangement of the intasome active site from INSTI binding. Active site regions of the APO and BIC-bound WT SIVrcm intasomes are shown as sticks with carbon atoms colored as in the inset (bottom left). Grey spheres are Mg^{2+} ions. Selected IN amino acid residues and the invariant 3' vDNA dCdA dinucleotide are indicated. Note the displacement of 3' vDNA deoxyadenosine (asterisk) in the INSTI-bound structure as indicated by the curved arrow.
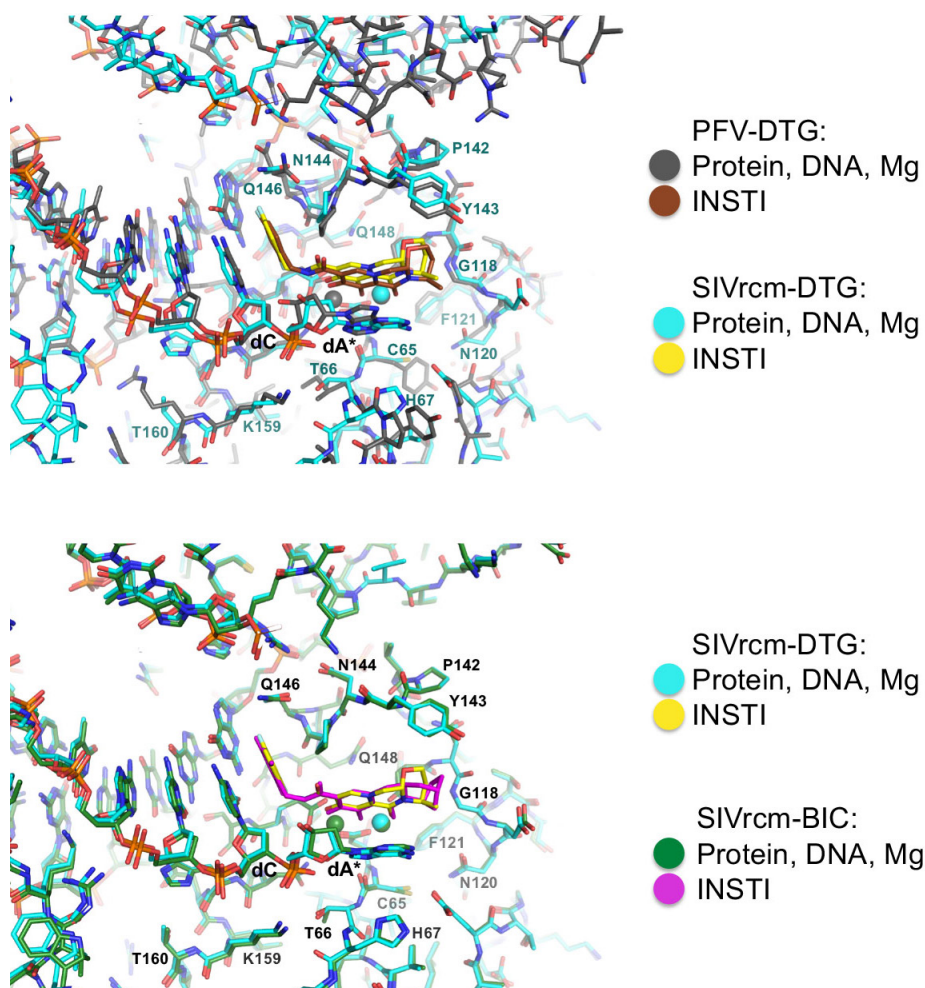
Fig. S15: Comparison of SIVrcm intasome-DTG with PFV intasome-DTG (top) and SIVrcm intasome-BIC (bottom) complexes. PFV intasome-DTG is an X-ray structure (PDB entry 3S3M) refined at 2.5 Å resolution. The structures, which were superposed by Cα atoms of the invariant active site residues, are shown as sticks with carbon atoms and Mg^{2+} ions (spheres) colored as indicated on the inset to the right. Selected SIVrcm IN amino acid residues and the invariant 3' vDNA dCdA dinucleotide are labeled; the deoxyadenosine displaced by the bound inhibitor is indicated with an asterisk. One crucial difference between the PFV and SIVrcm structures is the replacement of Gln148 with a Ser in the spumavirus, which cannot coordinate a water molecule with carboxylate Asp and Glu sidechains as does Gln148 in the SIVrcm intasome. The r.m.s. deviation between eleven Cα atoms of the residues in the immediate vicinity of the bound drugs is 0.58 Å. Thus, the overall conformations of the PFV and SIVrcm active sites are sufficiently discordant to affect geometric fit of the bound drug molecule. Furthermore, the packing of the displaced vDNA 3' adenosine against the metal chelating core of DTG differ between the SIVrcm and PFV structures (top panel). Finally, in accordance with the scarce overall amino acid sequence identity, the SIVrcm and PFV structures become highly divergent immediately outside of the active sites, where many secondary amino acid substitutions that enhance the levels of INSTI-resistance occur. By contrast, the differences between DTG and BIC-bound SIVrcm intasomes are subtle (bottom panel).
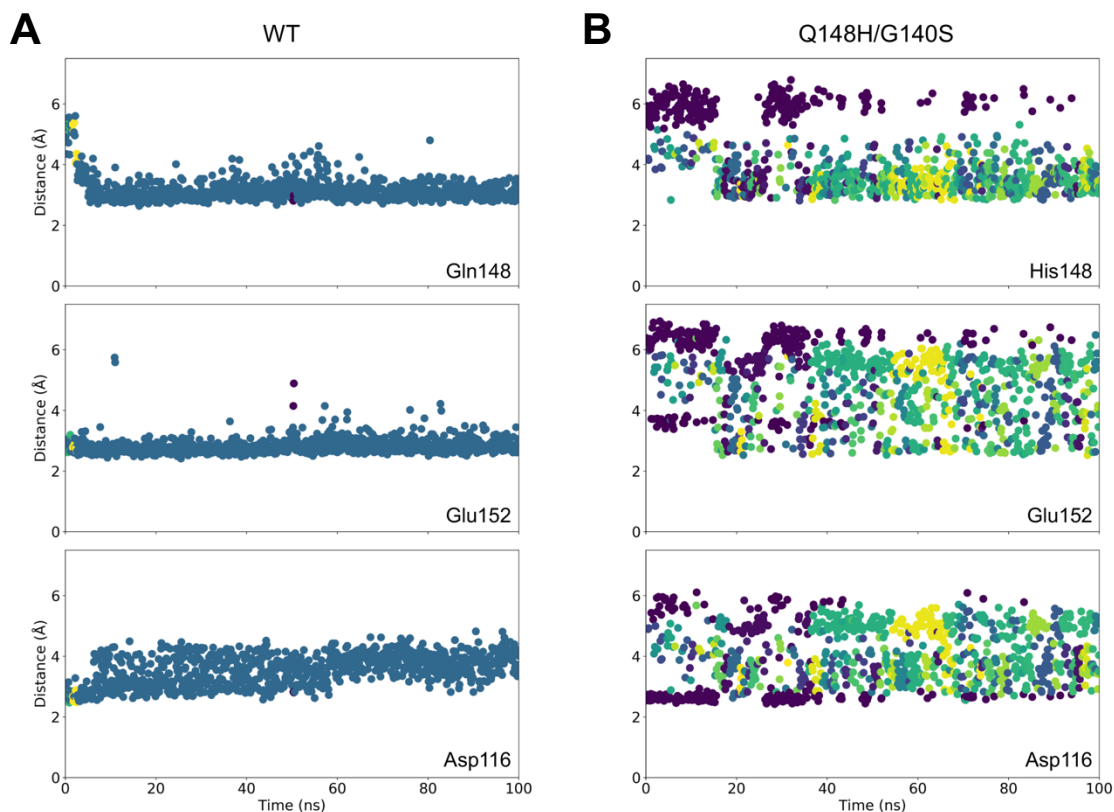
23

Fig. S16: Behavior of water molecules shared by Gln148 or His148 and active site carboxylates. (**A**) The structure of WT SIVrcm intasome bound to BIC was stripped of water molecules not directly coordinated to metal ions, embedded in bulk solvent, and subjected to 100 ns of molecular dynamics. The bulk solvent-derived water molecule closest to Gln148 Nε2 and the carboxylates of Glu152 and Asp116 was identified in each frame of the simulation. The dot plots report the corresponding distances every 0.1 ns of the molecular dynamics. The position corresponding to that occupied by W5 in the structure becomes occupied by a stably bound water molecule during the initial 5 ns of the simulation. Alterations of dot colors correspond to exchanges of the water molecule with bulk solvent. (**B**) A similar analysis with the Q148H/G140S SIVrcm BIC structure. Here, a water molecule closest to His148 Nε2 and the carboxylates of Glu152 and Asp116 was identified in each frame of the simulation. Note frequent exchanges with bulk solvent; ~4 Å is considered to be the upper limit for hydrogen bonding.
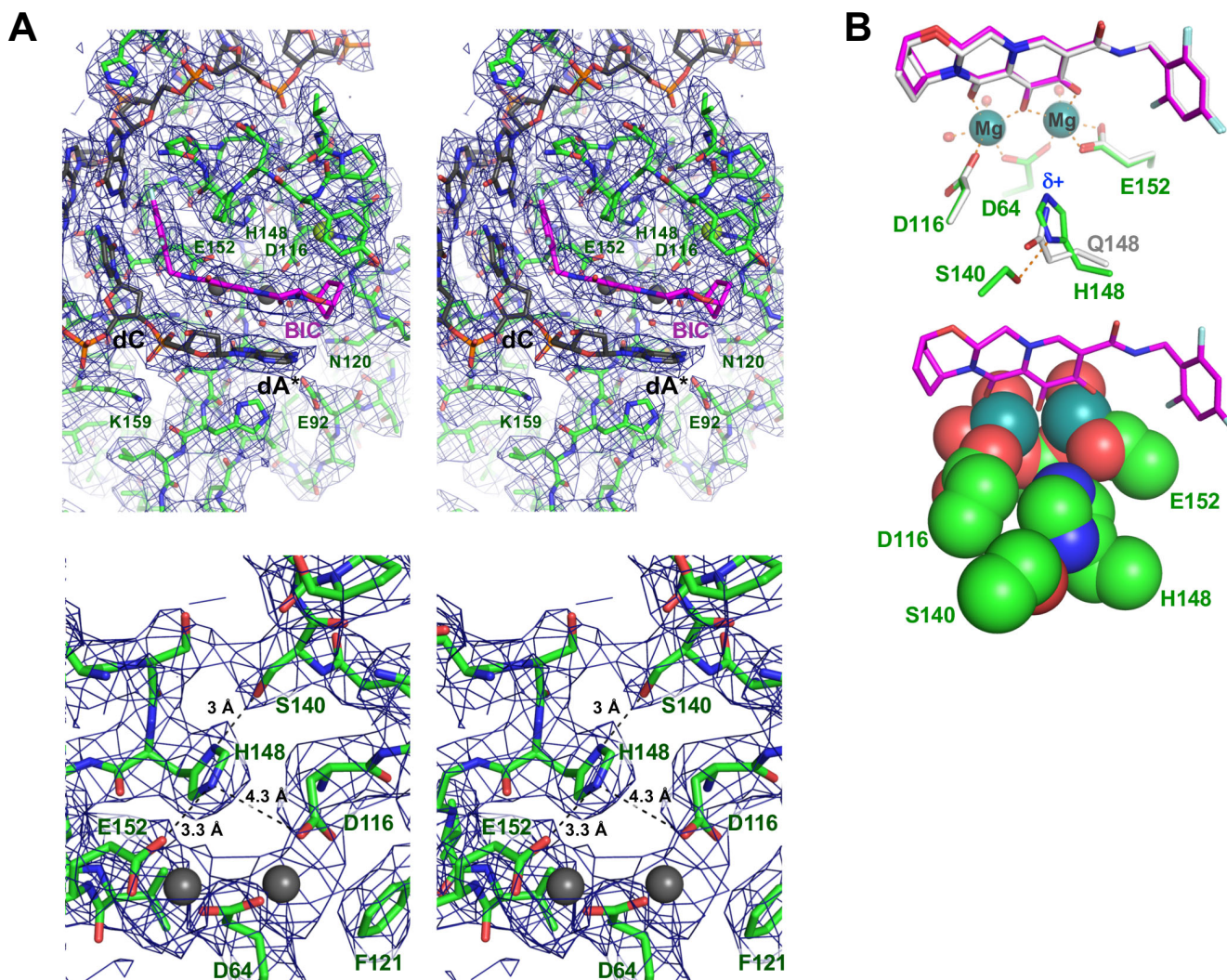
24

Fig. S17: Examples of cryo-EM density of BIC-bound Q148H/G140S SIVrcm intasome and details of the interactions in the active site. (**A**) The stereo views show IN chain A active site region (top) and details of the key involving His148 and Ser140 residues with corresponding interatomic distances in Ångstroms (bottom). The cryo-EM density and the refined model are depicted as in figs S9 and S12. See fig. S20 for a view focusing on Thr138 and His114. (**B**) Packing of His148 side chain in the active site of the Q148H/G140S intasome. Select residues of the Q148H/G148S (green with magenta BIC) and WT (grey residues and drug) structures are shown as overlaid sticks (top) or Van der Waals spheres (bottom; only Q148H/G140S structure shown); δ+ indicates increased electropositivity of the His148 Nε2 proton.
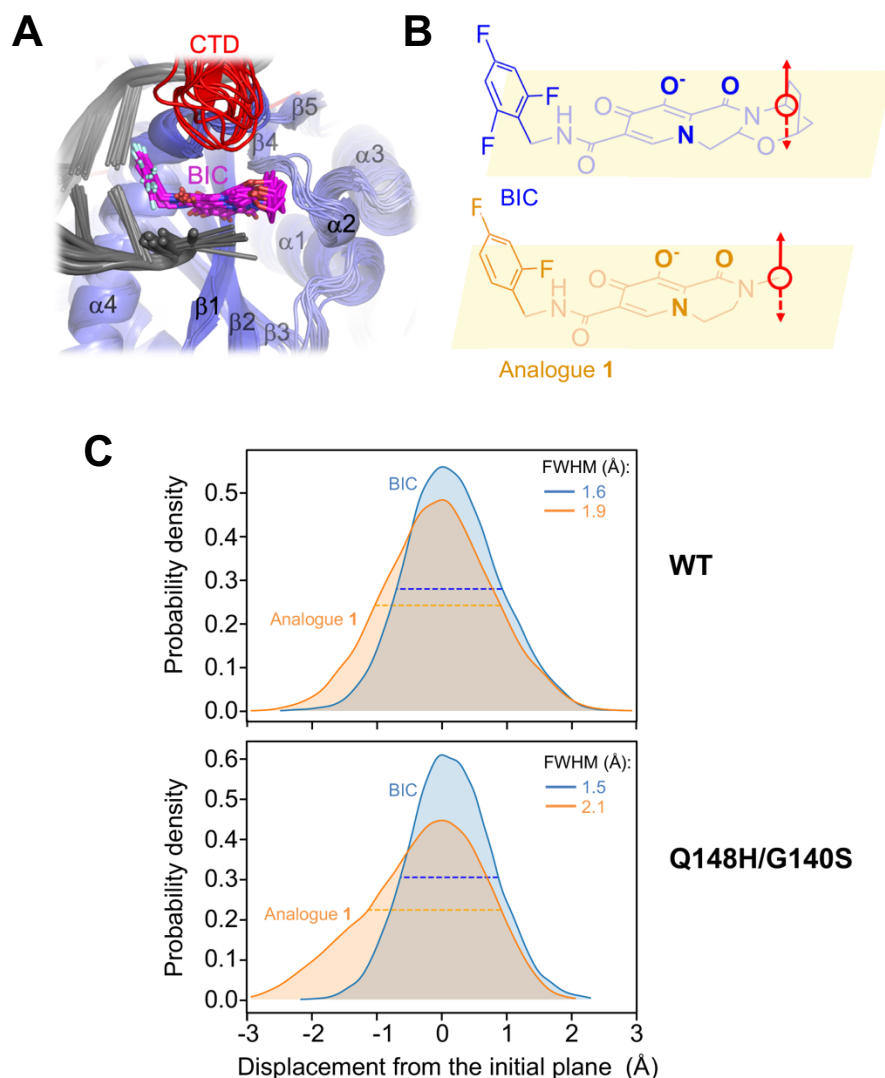
Fig. S18: Dynamics of BIC and analogue-1 in the intasome active site. WT and Q148H/G140S SIVrcm intasomes bound to BIC or analogue **1** were subjected to MD simulation. The resulting frames (12,500 structures derived from a total of 250 ns simulation per condition) were aligned by Cα atoms of intasome active site residues. (**A**) Subset of 10 WT intasome-BIC complex frames separated by 10 ns of simulation. Protein and DNA shown as cartoons and BIC as sticks. DNA is colored gray and protein is colored according to r.m.s. deviation from the initial position (blue, small displacement; red large displacement); the IN CTD and visible secondary structure elements of the CCD are indicated. (**B**) BIC and analogue **1** with the common carbon atoms closest to the β4-α2 connector when bound to the intasome active site indicated with red circles; arrowheads show direction of displacement chosen for the analysis. (**C**) Probability density for a given displacement of the chosen BIC (blue) or analogue-**1** (orange) carbon atom from the initial plane defined by bolded atoms in panel B in complex with WT (top) or Q148H/G140S (bottom) intasome. Full width at half maximum (FWHM) is listed for each distribution. Note a wider distribution of the atomic displacements in the case of analogue **1**.
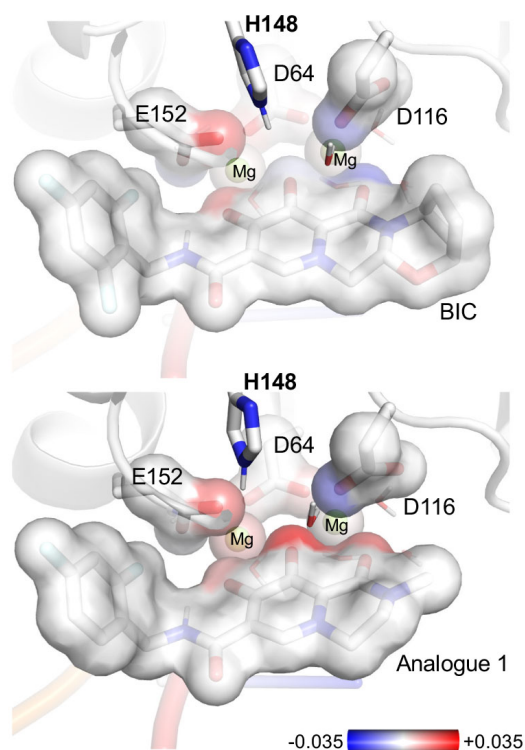
26

Fig. S19: Active site polarizability changes due to Q148H/G140S substitutions for BIC (top) and analogue 1 (bottom). Natural bond orbital analysis results are shown for the active site Mg^{2+}-ligand cluster. Protein residues, bound ligands, and metal ions from QM/MM minimized structures are represented as sticks and semi-transparent space-fill spheres. Colors indicate changes in charge distributions between the Q148H/G140S mutant and the wild type. Note the increased change in polarization of the metal chelating atoms of analogue **1** due to the amino acid substitutions.
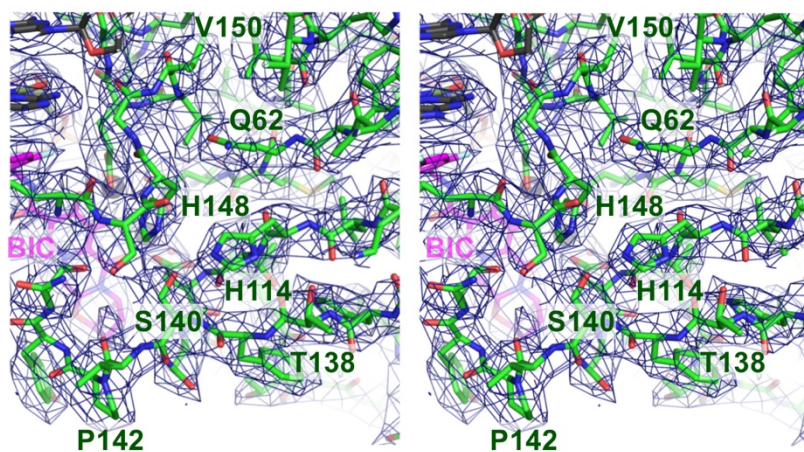
27

Fig. S20: The extended hydrogen bond network coupling Thr138 to His148 in the Q148H/G140S SIVrcm intasome. Cryo-EM map of Q148H/G140S SIVrcm intasome-BIC centered on the region shown in Fig. 2E. The cryo-EM density and the refined model are depicted as in figs S9 and S12.
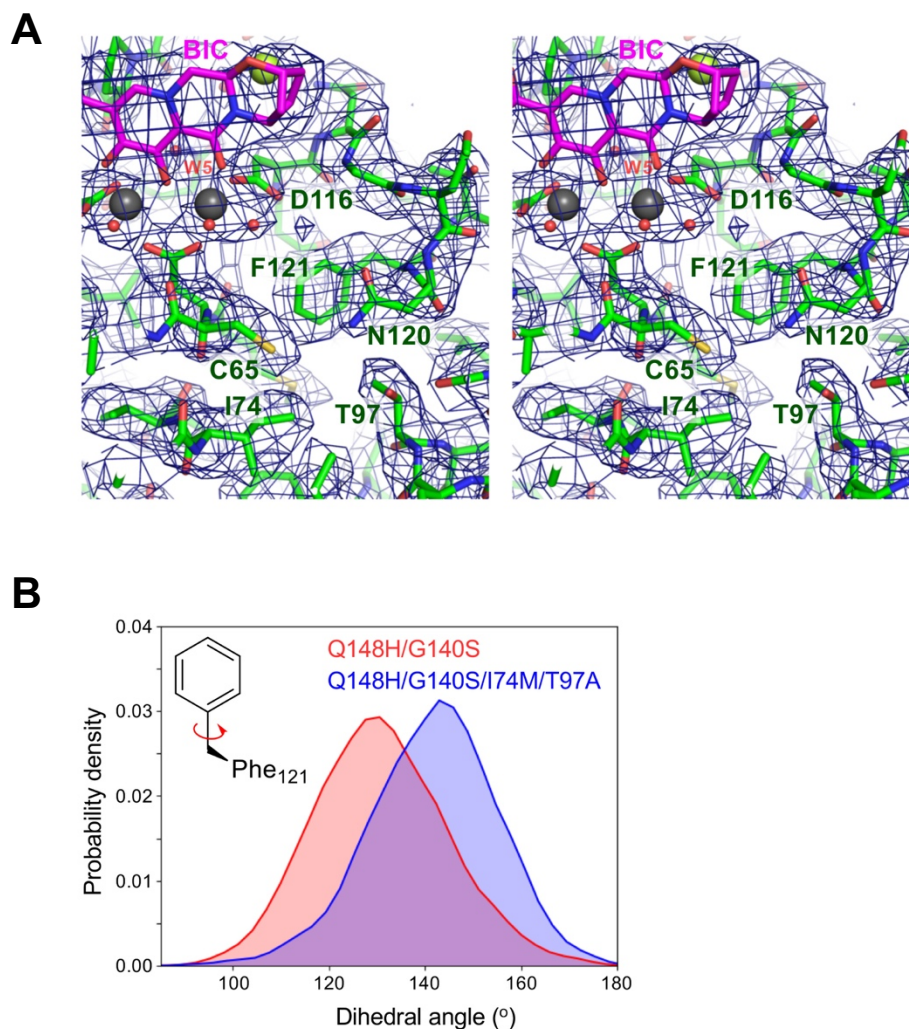
Fig. S21: Long-range interactions of Ile74 and Thr97 with the metal chelating cluster of the IN active site. (**A**) Cryo-EM map of WT SIVrcm intasome-BIC centered in the region shown in Fig. 2F, revealing definition of the side chain rotamers. The cryo-EM density and the refined model are depicted as in figs S9 and S12. (**B**) Effect of I74M/T97A substitutions on the conformation of Phe121. Models of Q148H/G140S SIVrcm-BIC intasome with and without the I74M/T97A substitutions were subjected to 100 ns of molecular dynamics. The graphs illustrate a shift in Phe121 Cβ-Cγ dihedral angle distributions in response to the I74M/T97A changes. Because Phe121 side chain intimately interacts with the Asp116 carboxylate (panel A and Fig. 2F), its conformation and dynamics are expected to directly affect the metal coordination cluster.
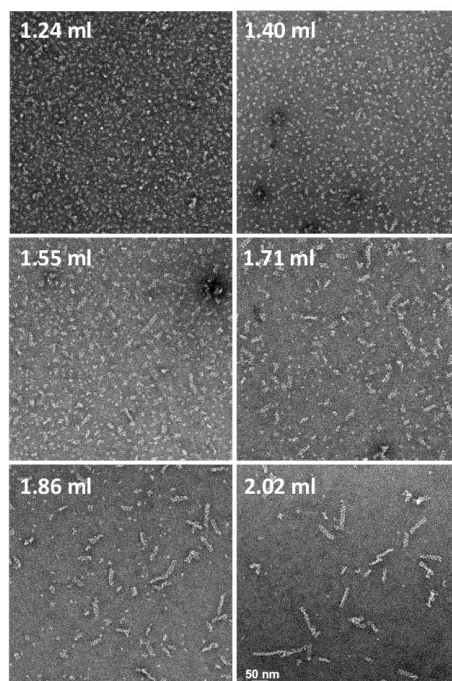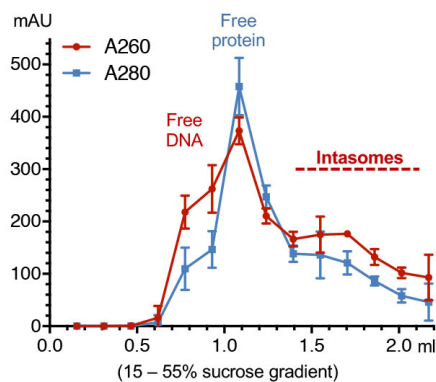
Fig. S22: Fractionation of SIVrcm intasomes by velocity sedimentation. SIVrcm strand transfer intasome assembly reactions were separated by ultracentrifugation through a 15-55% sucrose gradient; mean A_{260} (red) and A_{280} (blue) values for each fraction are plotted (error bars are from three independent experiments), and positions of free DNA oligonucleotide, protein, and intasomes are indicated. Selected fractions were analyzed by negative-stain EM (bottom).
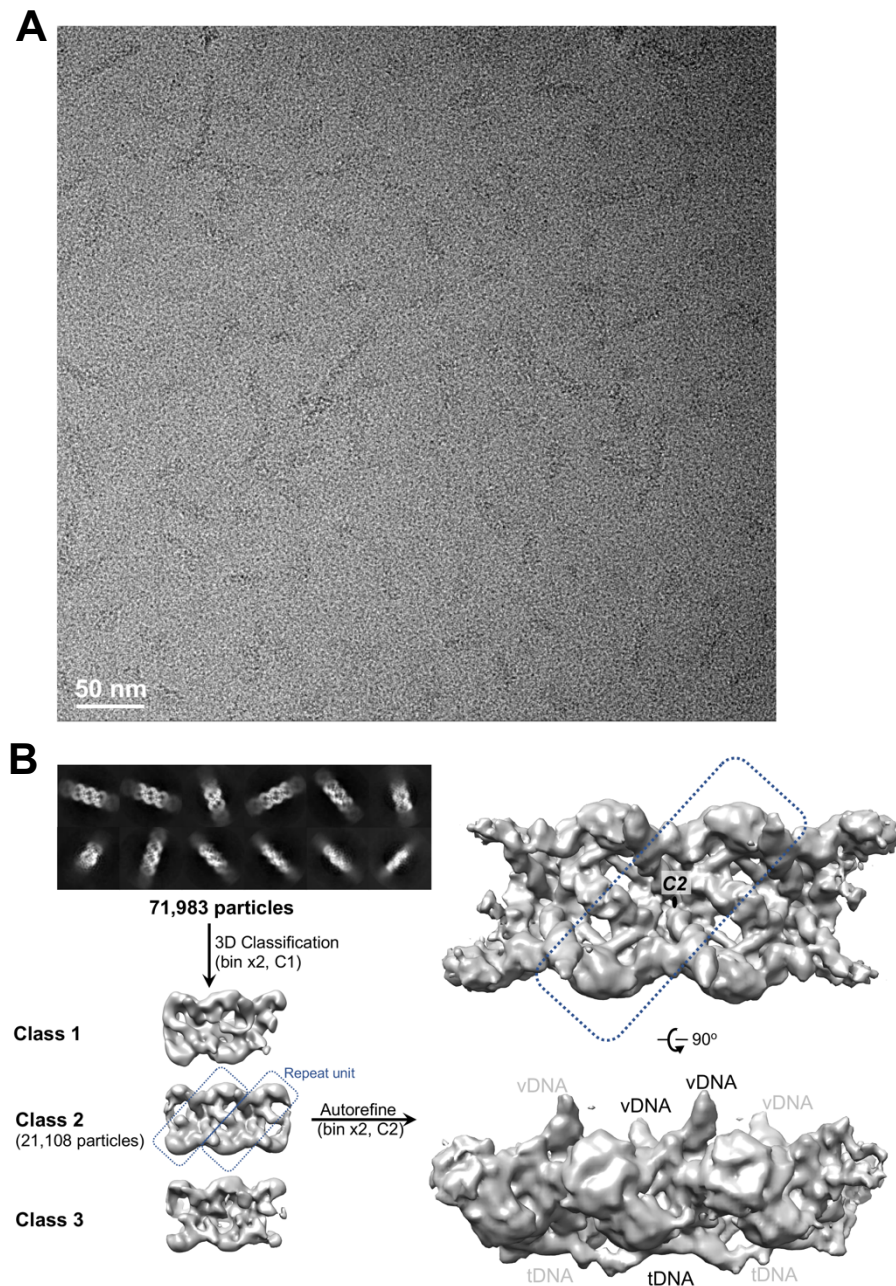
Fig. S23: Preliminary cryo-EM analysis of SIVrcm intasomes on ultrathin carbon-coated lacey grids. A. Example of a raw micrograph. **B.** Results of 2D and 3D classifications, and 3D reconstruction; tDNA, target DNA.
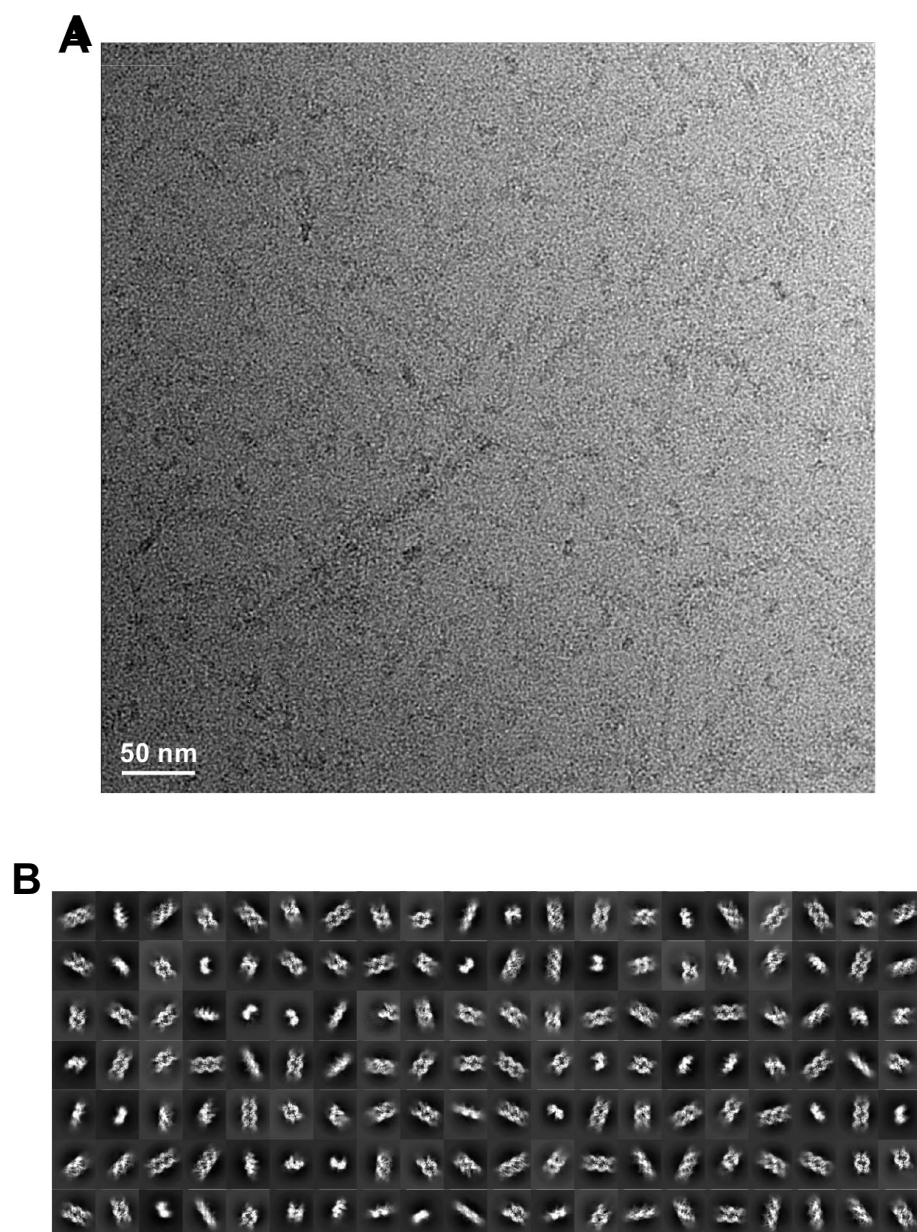
Fig. S24: Example of a raw micrograph (A) and 2D class averages (B) of SIVrcm intasomes in vitreous ice on holey carbon grids.

Relion_autopick
(Template set 1)

Relion_autopick
(Template set 2)

**Gautomatch**
(Template set 1)

Gautomatch
(Template set 2)

2D classification
(bin x4)

2D classification
(bin x4)

2D classification
(bin x4)

2D classification
(bin x4)

1,270,025 particles   **1,200,591 particles**   **835,959 particles**   **1,200,581 particles**

3D classification
(bin x2 ,C1, 4 classes)

3D classification
(bin x2 ,C1, 4 classes)

3D classification
(bin x2 ,C1, 4 classes)

3D classification
(bin x2 ,C1, 4 classes)

Combine high-resolution 3D classes,
remove overlapping particles

829,709 particles

3D Refinement
(bin x2, C2)

C2 V_R

V_L

Apply Euler matrix to V_L and V_R for each refined particle
to estimate coordinates of the neighboring repeat units
on the micrographs, merge and remove overlapping
particles

2,061,407 particles

(continued on the next page)

3D classification
(bin x2, C1, 18 classes)

Combine high-resolution
classes, refine, prune
overlapping particles

644,674 particles

Per-particle CTF estimation
3D refinement, unbinned particles, C2

3D classification with soft mask
focused on a single repeat unit
(unbinned particles, C1, 2 classes)

171,893 particles

Non-uniform refinement in cryoSPARC
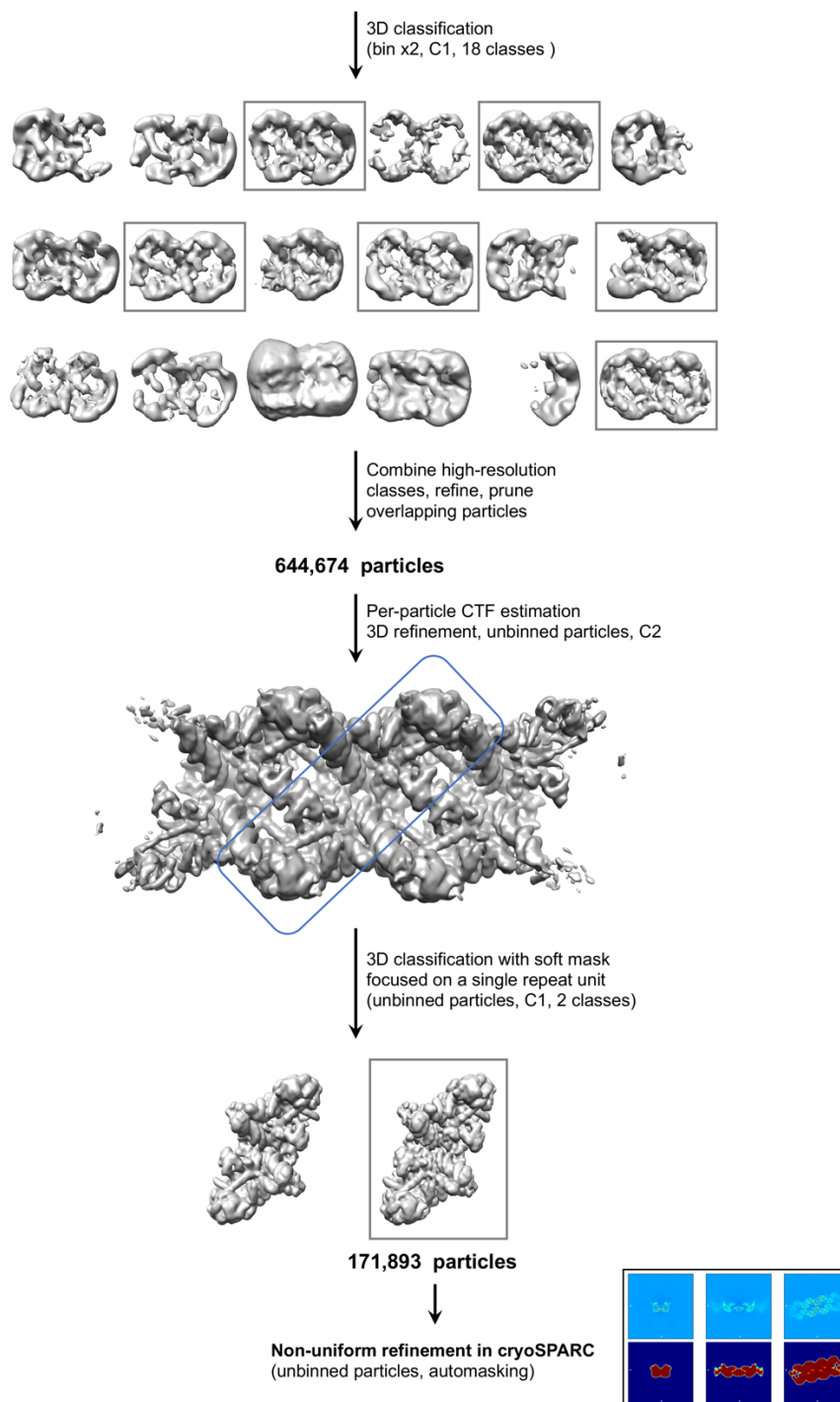(unbinned particles, automasking)

Fig. S25: Schematic of image processing and reconstruction of the Apo SIVrcm intasome. Inset on the bottom right shows orthogonal real space slices of the refined map (top) and auto-mask applied during non-uniform refinement in cryoSPARC (bottom). Similar flow charts applied to data acquisition and reconstitution of DTG-bound and BIC bound wild-type SIVrcm intasomes, as well as BIC-bound Q148H/G140S intasome; see Materials and Methods for details.
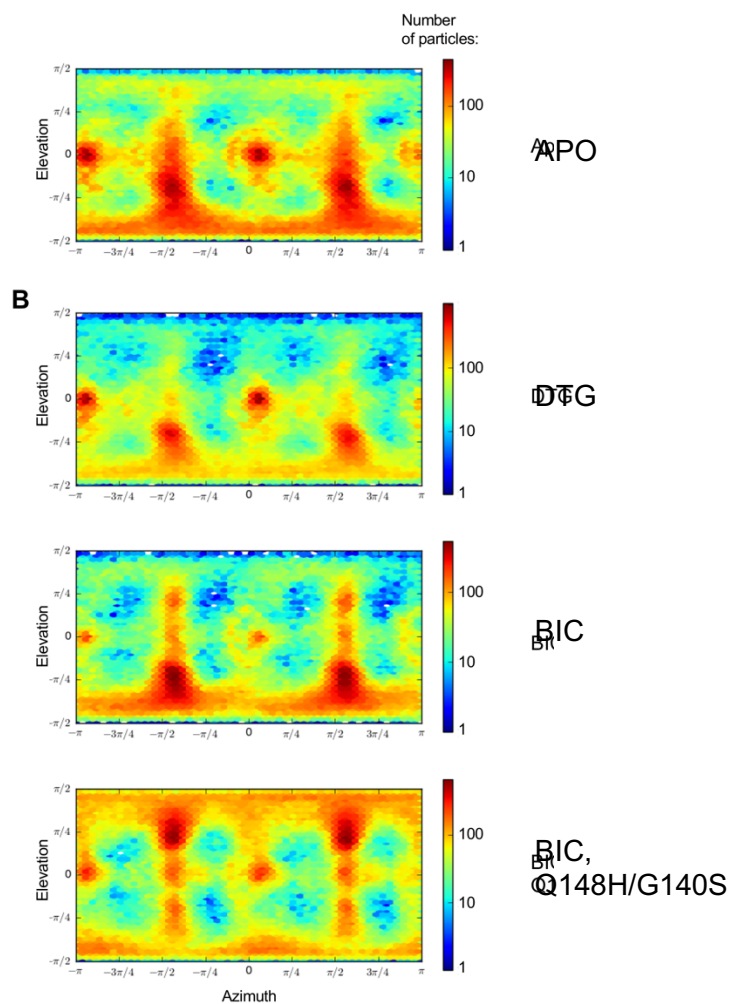
34

Fig. S26: Viewing direction distributions for the refined structures of Apo, DTG-bound, and BIC-bound WT SIVrcm intasomes and BIC-bound Q148H/G140S SIVrcm intasome. Heat maps represent particle number for each combination of refined azimuth and elevation, as indicated on the color scales to the right.
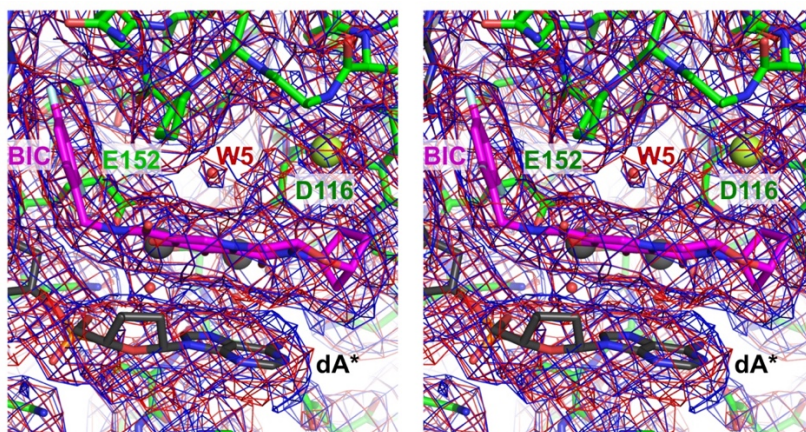
Fig. S27: Half-maps for the active site region of the WT SIVrcm intasome in complex with BIC. The maps obtained by refinement of two independent particle half-sets (generated during the final non-uniform refinement in cryoSPARC and auto-sharpened in Phenix) are displayed as blue and red chicken wire. Small red spheres are water molecules. Protein, DNA, and BIC are shown as sticks, with carbon atoms in green, grey and magenta, respectively. Note the presence of density for the refined water molecules, including W5, in both half-maps.

Table S1: Cryo-EM data collection and refinement.

	Apo	DTG	BIC	BIC Q148H/G140S
Database accession codes				
EMDB	EMD-10041	EMD-10043	EMD-10042	EMD-10044
RCSB	6RWL	6RWN	6RWM	6RWO
Data collection				
Microscope	Titan Krios	Titan Krios	Titan Krios	Titan Krios
Operating voltage (kV)	300	300	300	300
Detector	Gatan K2	Gatan K2	Gatan K2	Gatan K2
Magnification nominal/calibrated	105,000/36,232	105,000/36,232	130,000/46,296	105,000/36,232
Pixel size (Å)	1.38	1.38	1.08	1.38
Defocus range (µm)	1.6-3.5	1.6-3.5	1.6-3.5	1.6-3.5
Number of frames per movie	30	30	40	30
Electron exposure rate (e/$Å^2$/s)	4.2	4.2	5.6	4.2
Total electron exposure (e/$Å^2$)	50.4	50.4	55.7	50.4
Energy filter slit width (eV)	20	20	20	20
Automation software	EPU	EPU	EPU	EPU
Total movies acquired/used	8,311/8,027	6,961/6,838	11,846/11,769	6,775/6,618
Movie alignment software	MotionCor2	MotionCor2	MotionCor2	MotionCor2
Classification and 3D reconstruction				
Software for 2D classification	Cryosparc-1	Cryosparc-1	Cryosparc-1	Cryosparc-2
Software for 3D classification	Relion-2.1	Relion-2.1	Relion-2.1	Relion-3.0
Software for reconstruction	Cryosparc-2	Cryosparc-2	Cryosparc-2	Cryosparc-2
Number of extracted particles	6,235,062	5,370,417	6,426,769	5,642,300
Number of refined particles	171,893	178,294	149,778	217,635
Symmetry imposed	*C2*	*C2*	*C2*	*C2*
Map resolution (Å) [a]				
Unmasked volume	4.6 (7.6)	4.3 (7.2)	3.9 (6.8)	4.2 (7.1)
Repeat unit	3.4 (3.9)	3.1 (3.4)	2.8 (3.1)	2.9 (3.3)
CIC	3.2 (3.6)	3.0 (3.3)	2.6 (3.0)	3.0 (3.2)
Active site local resolution	2.8 - 3.0	2.8 - 2.9	2.4 - 2.7	2.8 - 2.9
3D FSC sphericity	0.960	0.969	0.975	0.971
Model refinement				
Map sharpening B-factor ($Å^2$)	-170	-110	-95	-160
Software for real-space refinement	Phenix	Phenix	Phenix	Phenix
Model composition				
Non-hydrogen atoms	18,272	18,346	18,450	18,366
Protein and DNA residues	2,165	2,165	2,165	2,159
Metal ions (Zn, Mg)	8	10	10	10
Water molecules	0	8	108	6
Ligands (INSTI)	0	2	2	2
Mean B factors ($Å^2$)				
Protein and DNA residues	65.5	25.1	34.6	54.1
Water	-	12.0	18.3	23.6
Ligands (INSTI)	-	10.3	14.1	26.0
Real-space correlation coefficient	0.8027	0.7974	0.7844	0.7961
R.m.s. deviations				
Bonds (Å)	0.003	0.004	0.006	0.006
Angles (º)	0.504	0.560	0.675	0.640
Validation [b]				
MolProbity clash score	2.63	2.62	2.40	2.29
Rotamer outliers (%)	2.86	3.66	3.20	4.10
Cβ deviations (%)	0	0	0	0
CαBLAM outliers (%)	1.31	1.21	1.31	1.51
EMRinger score	2.24	3.05	3.78	3.08
Ramachandran plot quality (%)				
Favored	96.95	96.85	96.11	96.65
Disallowed	0	0	0	0

[a] Based on the FSC of 0.143 between half-sets (values for FSC of 0.5 are given in parentheses).
[b] Assessed using MolProbity (*42*) and EMRinger (*43*).

Movie S1
Atomic model of BIC-bound SIVrcm intasome fitted into the experimental cryo-EM density. Labelling indicates IN and vDNA residues, as well as bound Mg^{2+} ions and waters (turquoise and red spheres, respectively). Additional solute density was tentatively modelled as a chloride ion (chartreuse sphere).

References and Notes

1. K. Anstett, B. Brenner, T. Mesplede, M. A. Wainberg, HIV drug resistance against strand transfer integrase inhibitors. *Retrovirology* **14**, 36 (2017). [doi:10.1186/s12977-017-0360-7](doi:10.1186/s12977-017-0360-7) Medline

2. B. A. Johns, T. Kawasuji, J. G. Weatherhead, T. Taishi, D. P. Temelkoff, H. Yoshida, T. Akiyama, Y. Taoda, H. Murai, R. Kiyama, M. Fuji, N. Tanimoto, J. Jeffrey, S. A. Foster, T. Yoshinaga, T. Seki, M. Kobayashi, A. Sato, M. N. Johnson, E. P. Garvey, T. Fujiwara, Carbamoyl pyridone HIV-1 integrase inhibitors 3. A diastereomeric approach to chiral nonracemic tricyclic ring systems and the discovery of dolutegravir (S/GSK1349572) and (S/GSK1265744). *J. Med. Chem.* **56**, 5901–5916 (2013). [doi:10.1021/jm400645w](doi:10.1021/jm400645w) Medline

3. M. Oliveira, R.-I. Ibanescu, K. Anstett, T. Mésplède, J.-P. Routy, M. A. Robbins, B. G. Brenner; Montreal Primary HIV (PHI) Cohort Study Group, Selective resistance profiles emerging in patient-derived clinical isolates with cabotegravir, bictegravir, dolutegravir, and elvitegravir. *Retrovirology* **15**, 56 (2018). [doi:10.1186/s12977-018-0440-3](doi:10.1186/s12977-018-0440-3) Medline

4. S. J. Smith, X. Z. Zhao, T. R. Burke Jr., S. H. Hughes, Efficacies of Cabotegravir and Bictegravir against drug-resistant HIV-1 integrase mutants. *Retrovirology* **15**, 37 (2018). [doi:10.1186/s12977-018-0420-7](doi:10.1186/s12977-018-0420-7) Medline

5. H. T. Pham, L. Labrie, I. E. A. Wijting, S. Hassounah, K. Y. Lok, I. Portna, M. E. Goring, Y. Han, C. Lungu, M. E. van der Ende, B. G. Brenner, C. A. Boucher, B. J. A. Rijnders, J. J. A. van Kampen, T. Mesplède, M. A. Wainberg, The S230R Integrase Substitution Associated With Virus Load Rebound During Dolutegravir Monotherapy Confers Low-Level Resistance to Integrase Strand-Transfer Inhibitors. *J. Infect. Dis.* **218**, 698–706 (2018). doi:10.1093/infdis/jiy175 Medline

6. I. E. A. Wijting, C. Lungu, B. J. A. Rijnders, M. E. van der Ende, H. T. Pham, T. Mesplede, S. D. Pas, J. J. C. Voermans, R. Schuurman, D. A. M. C. van de Vijver, P. H. M. Boers, R. A. Gruters, C. A. B. Boucher, J. J. A. van Kampen, HIV-1 Resistance Dynamics in Patients With Virologic Failure to Dolutegravir Maintenance Monotherapy. *J. Infect. Dis.* **218**, 688–697 (2018). doi:10.1093/infdis/jiy176 Medline

7. W. W. Zhang, P. K. Cheung, N. Oliveira, M. A. Robbins, P. R. Harrigan, A. Shahid, Accumulation of Multiple Mutations In Vivo Confers Cross-Resistance to New and Existing Integrase Inhibitors. *J. Infect. Dis.* **218**, 1773–1776 (2018). doi:10.1093/infdis/jiy428 Medline

8. S. Hare, S. S. Gupta, E. Valkov, A. Engelman, P. Cherepanov, Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* **464**, 232–236 (2010). doi:10.1038/nature08784 Medline

9. D. O. Passos, M. Li, R. Yang, S. V. Rebensburg, R. Ghirlando, Y. Jeon, N. Shkriabai, M. Kvaratskhelia, R. Craigie, D. Lyumkis, Cryo-EM structures and atomic model of the HIV-1 strand transfer complex intasome. *Science* **355**, 89–92 (2017). doi:10.1126/science.aah5163 Medline

10. S. Ahuka-Mundeke, F. Liegeois, A. Ayouba, Y. Foupouapouognini, E. Nerrienet, E. Delaporte, M. Peeters, Full-length genome sequence of a simian immunodeficiency virus (SIV) infecting a captive agile mangabey (Cercocebus agilis) is closely related

to SIVrcm infecting wild red-capped mangabeys (Cercocebus torquatus) in Cameroon. *J. Gen. Virol.* **91**, 2959–2964 (2010). doi:10.1099/vir.0.025767-0 Medline

11. P. M. Sharp, G. M. Shaw, B. H. Hahn, Simian immunodeficiency virus infection of chimpanzees. *J. Virol.* **79**, 3891–3902 (2005). doi:10.1128/JVI.79.7.3891-3902.2005 Medline

12. P. Cherepanov, LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. *Nucleic Acids Res.* **35**, 113–124 (2007). doi:10.1093/nar/gkl885 Medline

13. S. Hare, M.-C. Shun, S. S. Gupta, E. Valkov, A. Engelman, P. Cherepanov, 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.* **5**, e1000259 (2009). doi:10.1371/journal.ppat.1000259 Medline

14. A. Ballandras-Colas, D. P. Maskell, E. Serrao, J. Locke, P. Swuec, S. R. Jónsson, A. Kotecha, N. J. Cook, V. E. Pye, I. A. Taylor, V. Andrésdóttir, A. N. Engelman, A. Costa, P. Cherepanov, A supramolecular assembly mediates lentiviral DNA integration. *Science* **355**, 93–95 (2017). doi:10.1126/science.aah7002 Medline

15. A. S. Espeseth, P. Felock, A. Wolfe, M. Witmer, J. Grobler, N. Anthony, M. Egbertson, J. Y. Melamed, S. Young, T. Hamill, J. L. Cole, D. J. Hazuda, HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11244–11249 (2000). doi:10.1073/pnas.200139397 Medline

16. W. M. Konsavage Jr., S. Burkholder, M. Sudol, A. L. Harper, M. Katzman, A substitution in rous sarcoma virus integrase that separates its two biologically relevant enzymatic activities. *J. Virol.* **79**, 4691–4699 (2005). doi:10.1128/JVI.79.8.4691-4699.2005 Medline

17. M. G. Nowak, M. Sudol, N. E. Lee, W. M. Konsavage Jr., M. Katzman, Identifying amino acid residues that contribute to the cellular-DNA binding site on retroviral integrase. *Virology* **389**, 141–148 (2009). doi:10.1016/j.virol.2009.04.014 Medline

18. G. N. Maertens, S. Hare, P. Cherepanov, The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature* **468**, 326–329 (2010). doi:10.1038/nature09517 Medline

19. R. W. Shafer, Rationale and uses of a public HIV drug-resistance database. *J. Infect. Dis.* **194** (Suppl. 1), S51–S58 (2006). doi:10.1086/505356 Medline

20. J. M. George, S. S. Kuriakose, N. Dee, P. Stoll, T. Lalani, R. Dewar, M. A. Khan, M. T. Rehman, Z. Grossman, F. Maldarelli, A. K. Pau, Rapid Development of High-Level Resistance to Dolutegravir With Emergence of T97A Mutation in 2 Treatment-Experienced Individuals With Baseline Partial Sensitivity to Dolutegravir. *Open Forum Infect. Dis.* **5**, ofy221 (2018). doi:10.1093/ofid/ofy221 Medline

21. K. E. Hightower, R. Wang, F. Deanda, B. A. Johns, K. Weaver, Y. Shen, G. H. Tomberlin, H. L. Carter 3rd, T. Broderick, S. Sigethy, T. Seki, M. Kobayashi, M. R. Underwood, Dolutegravir (S/GSK1349572) exhibits significantly slower dissociation than raltegravir and elvitegravir from wild-type and integrase inhibitor-resistant HIV-1 integrase-DNA complexes. *Antimicrob. Agents Chemother.* **55**, 4552–4559 (2011). doi:10.1128/AAC.00157-11 Medline

22. D. Blow, More of the catalytic triad. *Nature* **343**, 694–695 (1990). doi:10.1038/343694a0 Medline

23. J. C. Marx, J. Poncin, J. P. Simorre, P. W. Ramteke, G. Feller, The noncatalytic triad of α-amylases: A novel structural motif involved in conformational stability. *Proteins* **70**, 320–328 (2008). doi:10.1002/prot.21594 Medline

24. M. E. Maguire, J. A. Cowan, Magnesium chemistry and biochemistry. *Biometals* **15**, 203–210 (2002). doi:10.1023/A:1016058229972 Medline

25. M. M. Harding, Geometry of metal-ligand interactions in proteins. *Acta Crystallogr. D* **57**, 401–411 (2001). doi:10.1107/S0907444900019168 Medline

26. F. Turlure, G. Maertens, S. Rahman, P. Cherepanov, A. Engelman, A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. *Nucleic Acids Res.* **34**, 1653–1665 (2006). doi:10.1093/nar/gkl052 Medline

27. S. Q. Zheng, E. Palovcak, J. P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017). doi:10.1038/nmeth.4193 Medline

28. K. Zhang, Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016). doi:10.1016/j.jsb.2015.11.003 Medline

29. G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, S. J. Ludtke, EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007). doi:10.1016/j.jsb.2006.05.009 Medline

30. D. Kimanius, B. O. Forsberg, S. H. Scheres, E. Lindahl, Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, e18722 (2016). doi:10.7554/eLife.18722 Medline

31. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017). doi:10.1038/nmeth.4169 Medline

32. S. L. Ilca, A. Kotecha, X. Sun, M. M. Poranen, D. I. Stuart, J. T. Huiskonen, Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes. *Nat. Commun.* **6**, 8843 (2015). doi:10.1038/ncomms9843 Medline

33. J. B. Heymann, Guidelines for using Bsoft for high resolution reconstruction and validation of biomolecular structures from electron micrographs. *Protein Sci.* **27**, 159–171 (2018). doi:10.1002/pro.3293 Medline

34. P. B. Rosenthal, R. Henderson, Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003). doi:10.1016/j.jmb.2003.07.013 Medline

35. S. H. Scheres, S. Chen, Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012). doi:10.1038/nmeth.2115 Medline

36. Y. Z. Tan, P. R. Baldwin, J. H. Davis, J. R. Williamson, C. S. Potter, B. Carragher, D. Lyumkis, Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017). doi:10.1038/nmeth.4347 Medline

37. J. C. Chen, J. Krucinski, L. J. Miercke, J. S. Finer-Moore, A. H. Tang, A. D. Leavitt, R. M. Stroud, Crystal structure of the HIV-1 integrase catalytic core and C-terminal

domains: A model for viral DNA binding. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8233–8238 (2000). [doi:10.1073/pnas.150220297](#) [Medline](#)

38. J. Y. Wang, H. Ling, W. Yang, R. Craigie, Structure of a two-domain fragment of HIV-1 integrase: Implications for domain organization in the intact protein. *EMBO J.* **20**, 7333–7343 (2001). [doi:10.1093/emboj/20.24.7333](#) [Medline](#)

39. P. Cherepanov, A. L. Ambrosio, S. Rahman, T. Ellenberger, A. Engelman, Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 17308–17313 (2005). [doi:10.1073/pnas.0506924102](#) [Medline](#)

40. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004). [doi:10.1107/S0907444904019158](#) [Medline](#)

41. P. V. Afonine, B. K. Poon, R. J. Read, O. V. Sobolev, T. C. Terwilliger, A. Urzhumtsev, P. D. Adams, Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018). [doi:10.1107/S2059798318006551](#) [Medline](#)

42. C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall 3rd, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson, D. C. Richardson, MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018). [doi:10.1002/pro.3330](#) [Medline](#)

43. B. A. Barad, N. Echols, R. Y. Wang, Y. Cheng, F. DiMaio, P. D. Adams, J. S. Fraser, EMRinger: Side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods* **12**, 943–946 (2015). [doi:10.1038/nmeth.3541](#) [Medline](#)

44. Y. Koh, K. A. Matreyek, A. Engelman, Differential sensitivities of retroviruses to integrase strand transfer inhibitors. *J. Virol.* **85**, 3677–3682 (2011). [doi:10.1128/JVI.02541-10](#) [Medline](#)

45. M. C. Shun, J. E. Daigle, N. Vandegraaff, A. Engelman, Wild-type levels of human immunodeficiency virus type 1 infectivity in the absence of cellular emerin protein. *J. Virol.* **81**, 166–172 (2007). [doi:10.1128/JVI.01953-06](#) [Medline](#)

46. W. Li, M. H. Lee, L. Henderson, R. Tyagi, M. Bachani, J. Steiner, E. Campanac, D. A. Hoffman, G. von Geldern, K. Johnson, D. Maric, H. D. Morris, M. Lentz, K. Pak, A. Mammen, L. Ostrow, J. Rothstein, A. Nath, Human endogenous retrovirus-K contributes to motor neuron disease. *Sci. Transl. Med.* **7**, 307ra153 (2015). [doi:10.1126/scitranslmed.aac8201](#) [Medline](#)

47. D. R. Langley, H. K. Samanta, Z. Lin, M. A. Walker, M. R. Krystal, I. B. Dicker, The terminal (catalytic) adenosine of the HIV LTR controls the kinetics of binding and dissociation of HIV integrase strand transfer inhibitors. *Biochemistry* **47**, 13481–13488 (2008). [doi:10.1021/bi801372d](#) [Medline](#)

48. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004). [doi:10.1002/jcc.20035](#) [Medline](#)

49. C. M. Breneman, K. B. Wiberg, Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **11**, 361–373 (1990). [doi:10.1002/jcc.540110311](#)

50. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennnuci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda *et al*., Gaussian 09, Revision E.01.

51. J. D. Chai, M. Head-Gordon, Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **10**, 6615–6620 (2008). doi:10.1039/b810189b Medline

52. F. Weigend, R. Ahlrichs, Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297–3305 (2005). doi:10.1039/b508541a Medline

53. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005). doi:10.1002/jcc.20289 Medline

54. R. B. Best, X. Zhu, J. Shim, P. E. Lopes, J. Mittal, M. Feig, A. D. Mackerell Jr., Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ(1) and χ(2) dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012). doi:10.1021/ct300400x Medline

55. A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993). doi:10.1063/1.464913

56. R. Ditchfield, W. J. Hehre, J. A. Pople, P. J.A., Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **54**, 724–728 (1971). doi:10.1063/1.1674902

57. W. J. Hehre, R. Ditchfield, J. A. Pople, Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **56**, 2257–2261 (1972). doi:10.1063/1.1677527

58. M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, J. A. Pople, Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **77**, 3654–3665 (1982). doi:10.1063/1.444267

59. B. R. Brooks, C. L. Brooks 3rd, A. D. Mackerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, M. Karplus, CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614 (2009). doi:10.1002/jcc.21287 Medline

60. Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio Jr., H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A.

W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. W. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, M. Head-Gordon, Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **113**, 184–215 (2015). doi:10.1080/00268976.2014.952696

61. C. M. Smith, G. G. Hall, The approximation of electron densities. *Theor. Chim. Acta* **69**, 63–69 (1986). doi:10.1007/BF00526293

62. A. E. Reed, R. B. Weinstock, F. Weinhold, Natural population analysis. *J. Chem. Phys.* **83**, 735–746 (1985). doi:10.1063/1.449486

63. A. E. Reed, L. A. Curtiss, F. Weinhold, Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* **88**, 899–926 (1988). doi:10.1021/cr00088a005

64. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014). doi:10.1093/nar/gku316 Medline