Reports of NAS RA
ՀՀ ԳԱԱ Զեկույցներ
NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF ARMENIA
ՀԱՅԱՍՏԱՆԻ ՀԱՆՐԱՊԵՏՈՒԹՅԱՆ ԳԻՏՈՒԹՅՈՒՆՆԵՐԻ ԱԶԳԱՅԻՆ ԱԿԱԴԵՄԻԱ

2026 **Volume 126** **Issue 2** **Article No. 2**

Category: Biology
Type of Paper: Original Research Article
Received: April 8, 2026, **Revised:** May 8, 2026, **Accepted:** May 14, 2026
Published: June 16, 2026

DOI: 10.54503/0321-1339-2026.126.2-2

Systematic Analysis of the HIV-1 Protease Active-Site Conformational Space Across 690 Crystal Structures

Hamlet Khachatryan[1,*]

[1]Institute of Molecular Biology, Yerevan, 0014, Armenia

[*]Correspondence: hamletkhachatryann@gmail.com

Abstract

Human Immunodeficiency Virus-1 protease (HIV-1 PR) is among the most extensively studied drug targets in the Protein Data Bank (PDB), with more than 600 structural models predominantly derived by X-ray crystallography. This study presents a comprehensive analysis of the binding-site conformational space across the available structural record: 690 crystal structures deposited in the PDB with ≥90% sequence identity and resolution better than or equal to 2.50 Å, of which 684 were successfully featurized by pipeline. The structural dataset covers wild-type enzyme, crystallographic stabilization mutants, drug-resistant variants, and 452 distinct inhibitor binders. Each binding site was featurized as a volume-filling point cloud with six descriptors (electrostatic potential, lipophilicity, and pharmacophoric features) and represented as a geodesic distance matrix with further embedding in spectral distance space. Affinity propagation clustered all pockets into 16 discrete conformational states, with four dominant states accounting for 85% of all structures. The chemical space of the 452 extracted inhibitors is largely continuous, and its partitioning does not correspond to the receptor's conformational landscape; yet the molecular weight of binders correlates with cavity volume: heavier inhibitors disproportionately populate the largest closed-flap state, while lighter molecules favour a distinct, tighter binding envelope. This analysis provides a structural basis and insights for ensemble-based virtual screening, *de novo* generation of inhibitors, and the systematic interpretation of evidenced structure–ligand data across the HIV-1 PR inhibitor landscape.

Keywords: HIV-1 protease, conformational space, binding-site analysis, virtual screening, ensemble docking

1. Introduction

HIV-1 protease (HIV-1 PR) is an essential enzyme in the viral life cycle. It cleaves the Gag and Gag-Pol polyproteins at nine distinct sites, converting immature virions into infectious particles [1, 2]. PR is one of the earliest and heavily validated antiretroviral targets [3] as in the case of loss of functional activity, progeny virions remain non-infectious. Since the approval of saquinavir in 1995, eight protease inhibitors (PIs) have entered clinical use, reducing AIDS-related mortality dramatically [4]. Darunavir, the most recent and broadly active PI, achieves broad resistance resilience through a hydrogen-bond network with backbone carbonyl and amide atoms of Asp29 and Asp30 on both chains [5, 6].

Despite this therapeutic success, resistance remains a persistent clinical challenge. More than 50 mutations distributed across the protease sequence have been attributed to PI-resistant types, ranging from active-site residues (D30N, V32I, I47V/A, G48V, I50V/L, V82A/F/T/S, I84V) that directly reshape the inhibitor-binding interface, to distal residues (L10I, M36I, L63P, A71V) that act through allosteric conformational perturbations on active-site geometry [7, 8]. The broad, evolutionarily emerged landscape of mutations causing drug resistance, together with modern early drug discovery pipelines that rely on structural information, motivates this systematic description of the physico-chemical parameter distribution and geometry of the binding site across clinically and structurally characterized genotypes.

The structure of the PR is a C2-symmetric homodimer of two identical 99-residue subunits. The substrate-binding cavity sits at the dimer interface, bounded above by two β-hairpin flap loops (residues 47–55 on each chain) and below by the catalytic dyad Asp25/Asp25', one residue from each subunit [9, 10]. NMR relaxation dispersion experiments characterize the flaps as intrinsically flexible regions that, in the apo form of the enzyme, predominantly occupy a semi-open state, with a substrate-accessible transiently open conformation observed on the microsecond-to-millisecond timescale [11]. Molecular dynamics simulations confirm spontaneous flap opening on the nanosecond timescale [12], and crystallographically evidenced flap geometries range from fully closed inhibitor-bound structures to wide-open states.

The PDB now contains more than 600 HIV-1 PR crystal structures, making this enzyme one of the most densely sampled conformational ensembles in structural biology. Previous computational studies have examined this dataset in several ways. RMSD-based clustering of backbone coordinates has established the taxonomy of closed, semi-open, and open flap states [12]. Normal-mode analysis and elastic-network models have mapped the low-frequency motions linking flap opening to active-site geometry changes [8]. Pharmacophore models derived from single representative structures have been used for virtual screening campaigns, but the sensitivity of docking accuracy to receptor conformation [13, 14] motivates the use of conformational ensembles rather than single structures. Ensemble docking has been shown to increase the effectiveness of virtual screening and improve pose plausibility for HIV-1 PR [13, 14], but it requires a principled way to select a minimal set of receptor conformational states that adequately covers the relevant landscape.

This systematic characterization of the entire evidenced structural space describes not only backbone geometry but also the chemical character of the binding site across all structures. Using binding-site volumetric representation and spectral distance analysis as an alignment-agnostic comparison framework, 690 HIV-1 PR crystal structures were characterised, including holo states with 452 distinct inhibitors, 38 drug-resistant variants, 76 crystallographically stabilised forms, and 58 apo structures.

2. Materials and Methods

2.1. Dataset assembly

To retrieve the studied dataset from PDB, the QPocket tool was used with the following parameters:

- Reference structure: 1HSG (HIV-1 PR bound to indinavir, 1.9 Å [15])
- Lowest resolution threshold: ≤2.50 Å
- Sequence identity: ≥90%

Overall 740 structures were retrieved, 50 of which were discarded because they did not contain both homodimer chains required for pocket formation, and the remaining 690 were used for further analysis. Each subunit (99 residues, UniProt P04585) was aligned using local alignment and renumbered according to the UniProt sequence: chain A residues 1–99, chain B residues 501–599.

Structures were classified as Bound State (co-crystallized with small-molecule inhibitor, peptide-substrate, or peptide-analogue) or Unbound State. Residues corresponding to water molecules, crystallization buffer components (acetate, sulfate, glycerol, etc.), and metal ions were treated as biologically irrelevant and were not considered as binders. Overall 58 proteins were in Unbound state. Within bound states, 4 structures contained an ABA (2-aminobutyric acid)-based C2-symmetric peptidomimetic analogue (7HVP, 8HVP, 1B6J, 1CPI), and 8 structures contained indinavir (1HSG, 1C6Y, 1SDT, 1SDU, 1SDV, 2AVO, 2AVS, 2AVV). The pocket featurisation and clustering were carried out using the original QPocket utilities.

2.2. Structural alignment

Structures were superposed onto 1HSG by sequence-guided Cα superposition using UniProt-derived residue mapping. The mean Cα RMSD across all 690 structures was 0.41 ± 0.09 Å, with inhibitor-bound and apo/substrate structures showing indistinguishable RMSD distributions, confirming that the global β-fold is conserved across all structures (Figure 1A).

2.3. Pocket identification and featurization

The binding pockets were defined as a point cloud of cropped grid points (0.70 Å spacing) embedded within a solvent-accessible area and were calculated using CavitOmiX (v. 1.0, 2022, Innophore GmbH). Six of the 690 aligned structures were removed from the downstream analysis because of the failure of CavitOmiX to detect pockets, resulting in 684 structures in the final dataset. Each grid point was assigned six binding-site descriptors: electrostatic potential (APBS [16, 17]), lipophilic potential, lone-pair potential, H-bond donor probability, H-bond acceptor probability, and aromatic character. This per-point descriptor captures the chemical environment of the site independently of backbone RMSD, so structures with identical global fold but with binding cavities with physicochemically distinct properties arising from resistance mutations, alternative conformers, or different crystal contacts, are assigned distinct pocket descriptors.

2.4. Volume estimation

Binding-site volume (Å³) was estimated from the featurized point cloud using a sphere-union volatilization: each grid point was padded with a sphere of radius 1.40 Å on a grid with spacing 0.23 Å, and occupied voxels were summed. Formally, the number of voxels Nv falling within the union of spheres is counted, and volume = Nv × d³ where d = 0.23 Å. This produces an absolute cavity volume in Å³ and is independent of the 0.70 Å point-cloud spacing.

2.5. Spectral distance matrix and clustering

QPocket constructs a geometric affinity graph for each pocket by connecting adjacent grid points within 2.0 Å. The normalised graph Laplacian

$$L = I - D^{-1/2}\_deg \cdot A \cdot D^{-1/2}\_deg$$

is computed for each pocket, and its first k = 100 non-trivial eigenvectors are extracted. Per-structure spectral descriptors integrate pocket geometry and chemical features, pairwise Euclidean distances between these descriptors yield the 684 × 684 spectral distance matrix used for all downstream analyses. Affinity propagation [18] (damping 0.9, 300 iterations) was applied, resulting in 16 conformational clusters.

2.6. Dimensionality reduction for visualization

Each structure was represented by its L2-normalised row in the spectral distance matrix (684-dimensional vector encoding distances to all other featurized structures). The matrix was standardised column-wise and projected to 2D using UMAP (n_neighbors = 10, min_dist = 0.02, Euclidean metric) [19] and t-SNE (perplexity = 40, 1000 iterations, Euclidean metric).

2.7. Interaction and residue analysis

Contact frequencies within all evidenced small-molecule binder–protein pairs were analyzed with PLIP [20] and recorded by interaction type (H-bond donor/acceptor, hydrophobic, salt bridge, halogen bond). Residues of the same position from both homodimer chains were merged. QPocket does not sort structures by genotype, so frequencies at resistance-mutation positions mix wild-type and mutant contributions.

2.8. Ligand chemical space analysis

Morgan circular fingerprints (radius 2, 2048 bits) were generated for all 452 unique inhibitor SMILES strings, and pairwise Tanimoto distances (1 − similarity) were computed with RDKit [21], yielding a 452 × 452 chemical distance matrix. Chemical space was visualised by t-SNE of this precomputed matrix. Molecular weight (MW), logP, H-bond donor/acceptor counts, topological polar surface area, and rotatable bond count were computed for all 452 ligands using RDKit.

3. Results

3.1. Dataset composition

Within the studied dataset 628 (91.0%) are co-crystallised with a small-molecule inhibitor, 58 (8.4%) are apo forms and 4 (0.6%) contain an ABA-based C2-symmetric peptidomimetic analogue (7HVP, 8HVP, 1B6J, 1CPI). These three subsets are used throughout the analysis. Of the 690 aligned structures, 684 pockets were successfully identified by CavitOmiX and featurized. The six unsuccessfully preprocessed structures were excluded from all downstream analyses, as described in Section 2.3.

Based on sequence annotation, 571 structures (82.8%) were classified as wild-type, 76 (11.0%) containing substitutions for crystallographic stabilisation (predominantly Q7K and/or L33I), 38 (5.5%) as drug-resistant mutants (at least one PI resistance mutation from the Stanford HIV Drug Resistance Database [22]), and 5 (0.7%) as site-directed engineered mutants. The observed resistance mutations were V82T (n = 11), L63P (n = 9), I50V (n = 8), V32I (n = 4), I54M (n = 2), V32I+I47V (n = 2), and single instances of L10I, I47V, I50L, and L90M.

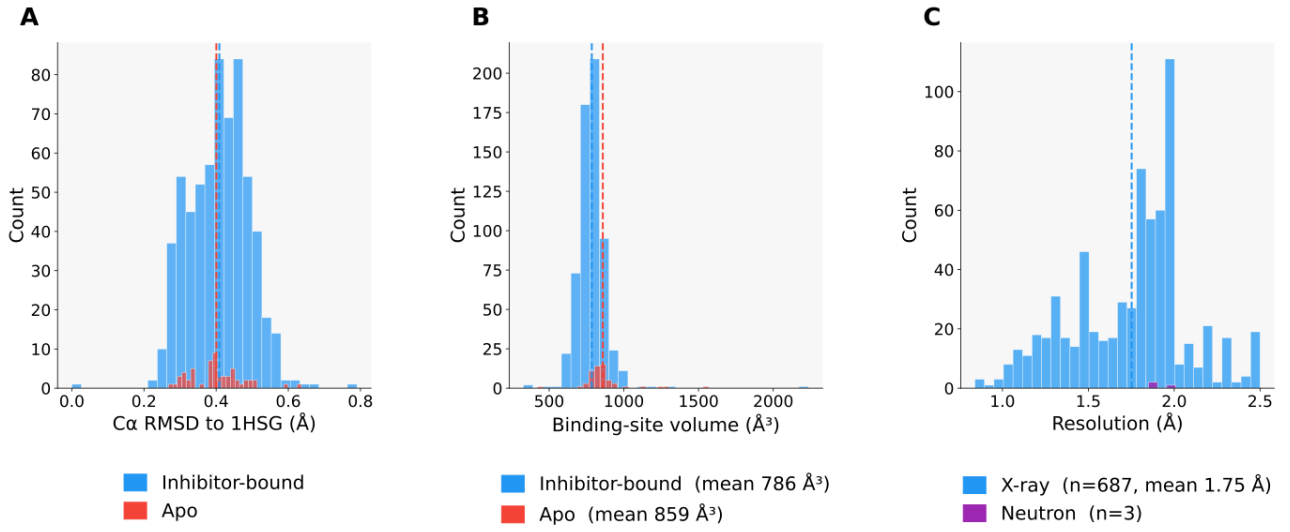

Figure 1. Dataset overview. **(A)** Distribution of Cα RMSD values to 1HSG for the three binding-state categories: holo (blue, n = 628), apo (red, n = 58), and holo with peptide-bound (green, n = 4). Dashed lines mark group means. **(B)** Binding-site volume distributions for the same categories: apo structures have a higher mean volume (859 Å³) than inhibitor-bound structures (786 Å³). **(C)** Crystallographic resolution of all 690 structures: 687 X-ray (blue) and 3 neutron diffraction (purple). Dashed line: mean X-ray resolution 1.75 Å.

Distribution of Cα RMSD after backbone-based alignment to 1HSG is uniform across all three categories, with maximum RMSD of 0.79 Å (mean 0.41 Å; Figure 1A), confirming global fold conservation within the evidenced conformational landscape. Binding-site volumes measured based on point clouds range from 328 to 2236 Å³ (mean 792 Å³; Figure 1B). Apo structures have a higher mean volume (859 Å³) than holo structures (786 Å³). All 690 structures were determined by X-ray crystallography (687) or neutron diffraction (3; PDB entries 2ZYE, 4JEC, 5T8H), with X-ray resolutions varying between 0.84 and 2.50 Å (mean 1.75 Å; Figure 1C).

3.2. Evidenced conformational landscape

After applying dimensionality reduction via UMAP (Figure 2A) and t-SNE (Figure 2B), the HIV-1 PR active-site conformational landscape shows that the four dominant states (C08, C09, C11, C12) form a compact central mass, with minor and exceptional states arranged at the periphery. Apo structures scatter across the interior and margins of the central mass indicating conformational selection rather than induced fit as the dominant binding mechanism [12].

The most isolated point in projections is the metallacarborane complex 1ZTZ (C07), whose cavity (2236 Å³) is the largest in the dataset. It shows the experimentally evidenced extreme of flap opening. The semi-open apo cluster C02 (mean volume 1001 Å³) and the dominant inhibitor-bound states (mean volume 777–804 Å³) occupy clearly separated volume ranges (Figure 2C). Within the four dominant states the volume distributions largely overlap (740–820 Å³), indicating that their differences in pocket character result from shape and physicochemical parameters rather than total cavity size.

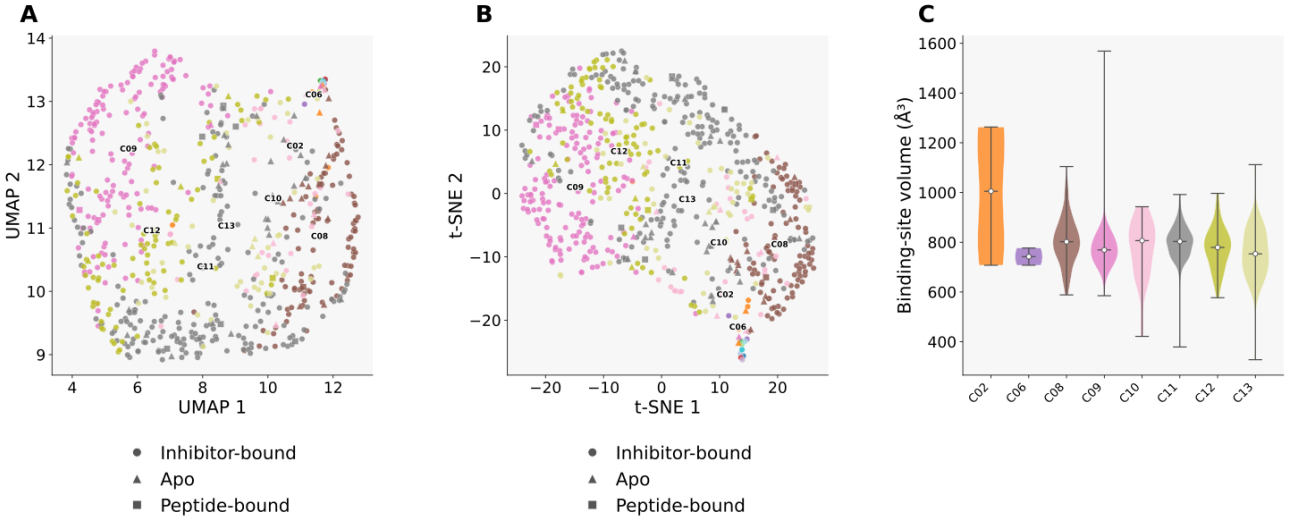

Figure 2. Conformational landscape of the HIV-1 PR active site. **(A)** UMAP and **(B)** t-SNE projections of per-structure spectral distance fingerprints, coloured by conformational state. Circles: inhibitor-bound, triangles: apo/peptide-bound. Stars mark state exemplars selected as representatives. Cluster labels shown for states with ≥8 members. **(C)** Distributions of binding-site volumes (Å³) per cluster.

3.3. The four dominant conformational states

Based on clustering results, four major clusters (conformational states) account for 583 of 684 featurized structures (85%) and represent the range of the conformational states resulting after binding of the vast majority of clinical and investigational inhibitors (Table 1).

C11 (exemplar 3VFB, n = 233, 34%) is the canonical darunavir-class geometry. In this conformational cluster, both flap loops adopt a closed configuration, the S1/S1′ hydrophobic subsite is compressed, and the catalytic dyad assumes a bifurcated hydrogen-bonding arrangement. **C09** (exemplar 2QI0, n = 155) corresponds to a closed-flap conformation characterized by relatively low

conformational variability within the cluster. **C08** (exemplar 2IDW, n = 103) likewise adopts a closed-flap architecture but displays the highest variance within the cluster, implying that it accommodates conformations that differ in the structural and physicochemical properties of the flap-tip sub-pockets as well as within the S3/S3′ subsite region. **C12** (exemplar 4Q1Y, n = 92) is the most internally homogeneous of the four clusters. Among the twelve minor states, C13 (n = 52) contains most of the drug-resistant variants (17.3% versus 5.5% overall). C10 (n = 33) occupies a position between the dominant mass and the wide-open outlier and includes intermediate flap geometries. C02 (n = 6) is the cluster where the majority of the structures are in the apo state and corresponds to the semi-open conformation used as a starting point in MD simulations [12].

Table 1. Cluster statistics.

Cluster	Exemplar	Cluster Size	Holo Form	Apo Form	Pocket Volume (Å^3)	RMSD (Å)
C00	1A9M	1	1	0	774	0.44
C01	1BV7	1	1	0	777	0.48
C02	1F7A	6	4	2	1237	0.39
C03	1K6P	1	1	0	1284	0.34
C04	1MET	1	1	0	804	0.46
C05	1MRX	1	1	0	744	0.36
C06	1MSN	2	2	0	742	0.43
C07	1ZTZ	1	1	0	2236	0.79
C08	2IDW	103	90	13	804	0.40
C09	2QI0	155	151	4	769	0.43
C10	3NDW	33	27	6	807	0.37
C11	3VFB	233	198	35	803	0.40
C12	4Q1Y	92	86	6	779	0.42
C13	6IXD	52	48	4	753	0.39
C14	6OPT	1	1	0	765	0.46
C15	6OPU	1	1	0	794	0.40

3.4. The binding-site core

Across all 684 featurized structures, the following 12 residues are exposed within the volume of the binding pockets in ≥80% of cases (Figure 3B): Leu23, Ala28, Asp29, Asp30, Val32, Ile47, Gly48, Ile50, Pro81, Val82, and Ile84 (both chains), together with Gly27. These residues form all established sub-pockets: the catalytic cleft (Asp29/30), S1/S1′ hydrophobic pocket (Ala28, Leu23, Val32), flap tips (Ile50, Gly48), and S2/S2′ wall (Pro81, Val82, Ile84). Those amino acids form regions where the majority of the established direct-contact resistance-mutation sites are [6].

Observed interaction frequencies in the small molecular inhibitor-bound subset (Figure 3A) show that Asp29A/B carries the highest hydrogen-bond frequency (58.6% and 51.1% for chains A and B). Hydrophobic contacts concentrate at Ala28A/B and Ile50A/B. The chain-A/B asymmetry arises from the mix of symmetric and asymmetric inhibitor scaffolds in the dataset. Residues located in peripheral positions (Thr80, Phe53, Lys45) engage fewer than 30% of structures and they form outer boundary reached by inhibitors within evidenced space.

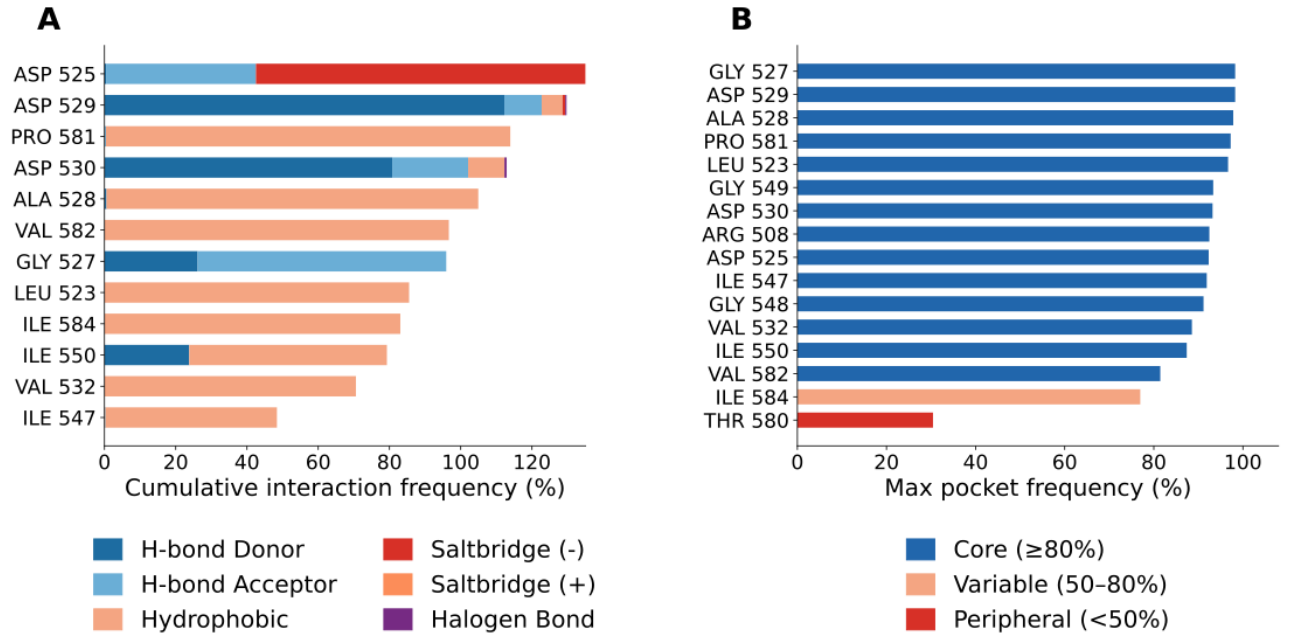

Figure 3. Frequency of the interacting residues in binding site. **(A)** Observed interaction frequencies for the top-12 residues, stacked by type (merged for both chains). **(B)** Potential interaction frequency: fraction of all 684 structures in which each residue falls within the pocket boundary. Blue: core (≥80%); gold: variable (50–80%); red: peripheral (<50%). All direct-contact resistance-mutation sites lie within the core.

3.5. Inhibitor chemical space

The studied 452 chemically unique inhibitors resolved by structural methods are the result of more than three decades of HIV-1 PR medicinal chemistry.

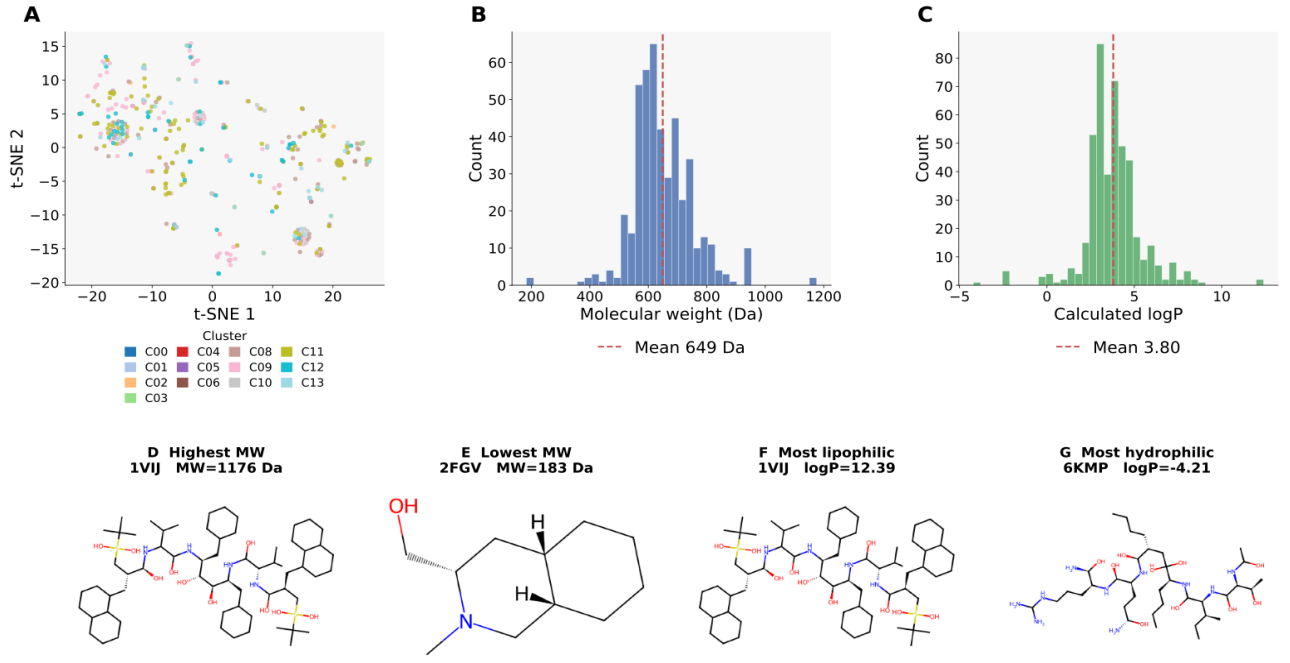

Figure 4. Chemical space of the structurally resolved inhibitors. **(A)** t-SNE of the 452 × 452 Tanimoto distance matrix, coloured by receptor cluster. Distributions of **(B)** molecular weight and **(C)** logP across all 452 inhibitors. **(D)** Lowest-MW inhibitor (2FGV, 183 Da). **(E)** Highest-MW inhibitor (1VIJ, 1176 Da). **(F)** Most lipophilic inhibitor (1VIJ, logP = 12.4). **(G)** Most hydrophilic inhibitor (6KMP, logP = −4.2).

Pairwise Tanimoto distances within the studied structures have a mean of 0.74 ± 0.13, with more than 80% of ligand pairs below a similarity of 0.30, showcasing broad scaffold diversity. Molecular weights range from 183 Da (2FGV, Figure 4D) to 1176 Da (1VIJ, Figure 4E), with a mean of 436 Da (Figure 4B). Calculated logP ranges from −4.2 (6KMP, Figure 4G) to 12.4 (1VIJ, Figure 4F), with a mean of 2.9 (Figure 4C).

Dimensionality reduction of the Tanimoto distance matrix with t-SNE (Figure 4A) shows continuous and undifferentiated chemical space in the context of binding pocket clusters. Cluster labels are mixed throughout the projection and no conformational state occupies a distinct region. However, states differ in the size of inhibitor they accommodate. C11 predominantly contains the heaviest inhibitors (highest median MW). C09 accommodates a lower and narrower MW range. Lipophilicity overlaps broadly across all four dominant states, so logP is not a determinant of conformational state resulting after the binding event.

3.6. Representative pocket conformations

Three structures showcasing the full conformational range sampled by the dataset are shown in Figure 5.

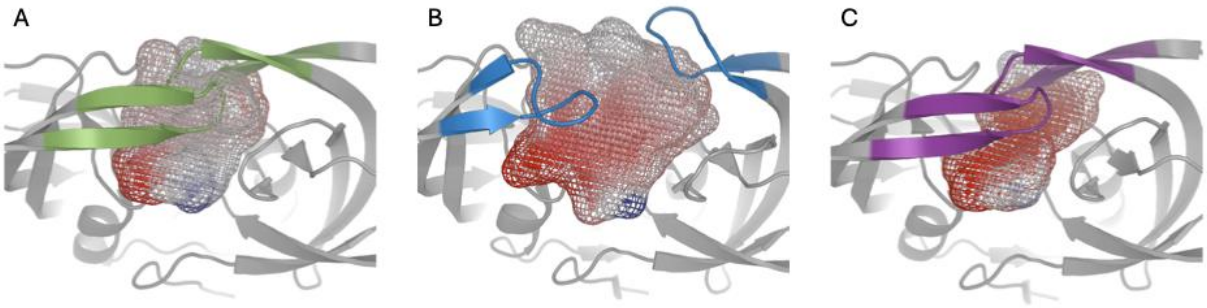

Figure 5. Representative binding-site conformations. Flap loops are coloured per panel; the point cloud representing binding site volume is coloured by electrostatic potential (red: negative, blue: positive, white: neutral). **(A)** 1F7A (C02) — apo semi-open state, flaps (green) displaced outward, cavity volume 1237 Å³ (n_cluster = 6). **(B)** 1ZTZ (C07) — metallacarborane-bound, maximally open flaps (blue), cavity volume 2236 Å³ (n_cluster = 1). **(C)** 3VFB (C11) — dominant darunavir-class closed-flap geometry (purple), cavity volume 728 Å³ (n_cluster = 233).

1F7A (C02, panel A) is the exemplar apo semi-open state: both flap β-hairpins are displaced outward, leaving the catalytic cleft fully accessible and the measured pocket volume at 1237 Å³. **1ZTZ** (C07, panel B) binds a metallacarborane cluster that forces maximal flap opening [23] (cavity volume of 2236 Å³) and represents the upper bound of accessible volume in the evidenced space. **3VFB** (C11, panel C) is the dominant darunavir-class geometry (34% of all structures): the flap loops compress against the inhibitor, the S1/S1′ hydrophobic subsite is fully enclosed, and the Asp29/30 backbone forms the hydrogen-bond contacts exploited for resistance resilience [5]. Its pocket volume (728 Å³) reflects the closed conformation. Together these three structures bound the conformational landscape and illustrate the range of cavity geometries that any ensemble-based drug design campaign should consider.

4. Discussion

In this study, the identified 12-residue binding-site core, covering the catalytic cleft, the S1/S1′ hydrophobic subsite, the flap tips, and the S2/S2′ wall, is a contact network where every node is both

an inhibitor interaction region and a structurally or catalytically essential residue. The core largely overlaps with all established direct-contact resistance-mutation sites and explains why the HIV-1 PR active site resists complete evolutionary escape from inhibition: any single mutation weakens one contact while the inhibitor holds at the remaining 11 nodes. Darunavir anchors to backbone atoms of Asp29/30 rather than side-chain groups, exploiting contacts that survive any point mutation at those positions [5, 6].

The four dominant states differ in the size of inhibitor they accommodate but not in scaffold topology. Within any cluster, the mean pairwise Tanimoto distance (0.73–0.76) matches the overall dataset mean (0.74), so no state filters for a particular chemical framework. States C09 and C12 are geometrically compact and favour lower-MW inhibitors, C08 encompasses the broadest internal variation and probably covers multiple sub-states that each deserve a representative in ensemble docking campaigns, C11 is the largest state and accommodates the small molecules with the widest MW range. Ligands above roughly 600 Da are disproportionately found in C11, while fragments below 400 Da preferentially populate C09 and C12. This is a practical argument for receptor selection in ensemble virtual screening. The active site is therefore amenable to *de novo* design without requiring a new receptor conformation to emerge, since diverse scaffolds access the same four dominant states.

5. Conclusions

In this study HIV-1 PR active-site conformational space has been analyzed across 690 crystal structures and includes the full diversity of clinical inhibitor scaffolds and resistance genotypes. Affinity propagation of a spectral distance metric clustered 684 featurized pockets into 16 states, four of which account for 85% of all structures. These four states share a 12-residue binding-site core that overlaps with every direct-contact resistance-mutation site. Conformational states differ in the cavity volumes they present (328–2236 Å^3 across the dataset) and in the MW range of inhibitors they accommodate, but chemical space across 452 inhibitors is continuous and does not segregate by receptor state. The dataset and distance metrics reported here provide a ready-made receptor ensemble for virtual screening and a structural reference for resistance-aware inhibitor design, highlighting the importance of the initial structure selection for the numerical methods behind in silico screening campaigns.

References

[1] N. E. Kohl, E. A. Emini, W. A. Schleif, L. J. Davis, J. C. Heimbach, R. A. F. Dixon, E. M. Scolnick, and I. S. Sigal. Active human immunodeficiency virus protease is required for viral infectivity. Proc. Natl. Acad. Sci. USA, 85(13):4686–4690, 1988. https://doi.org/10.1073/pnas.85.13.4686.

[2] C. Peng, B. K. Ho, T. W. Chang, and N. T. Chang. Role of human immunodeficiency virus type 1-specific protease in core protein maturation and viral infectivity. J. Virol., 63(6):2550–2556, 1989. https://doi.org/10.1128/jvi.63.6.2550-2556.1989.

[3] A. Wlodawer and J. W. Erickson. Structure-based inhibitors of HIV-1 protease. Annu. Rev. Biochem., 62:543–585, 1993. https://doi.org/10.1146/annurev.bi.62.070193.002551.

[4] E. J. Arts and D. J. Hazuda. HIV-1 antiretroviral drug therapy. Cold Spring Harb. Perspect. Med., 2(4):a007161, 2012. doi: https://doi.org/10.1101/cshperspect.a007161.

[5] N. M. King, M. Prabu-Jeyabalan, E. A. Nalivaika, and C. A. Schiffer. Combating susceptibility to drug resistance: lessons from HIV-1 protease. Chem. Biol., 11(10):1333–1338, 2004. https://doi.org/10.1016/j.chembiol.2004.08.010.

[6] A. M. J. Wensing, N. M. van Maarseveen, and M. Nijhuis. Fifteen years of HIV protease inhibitors: raising the barrier to resistance. Antiviral Res., 85(1):59–74, 2010. https://doi.org/10.1016/j.antiviral.2009.10.003.

[7] M. W. Tang and R. W. Shafer. HIV-1 antiretroviral resistance: scientific principles and clinical applications. Drugs, 72(9):e1–e25, 2012. https://doi.org/10.2165/11633630-000000000-00000.

[8] D. A. Ragland, E. A. Nalivaika, M. N. L. Nalam, K. L. Prachanronarong, H. Cao, R. M. Bandaranayake, Y. Cai, N. Kurt-Yilmaz, and C. A. Schiffer. Drug resistance conferred by mutations outside the active site through alterations in the dynamic and structural ensemble of HIV-1 protease. J. Am. Chem. Soc., 136(34):11956–11963, 2014. https://doi.org/10.1021/ja504096m.

[9] M. A. Navia, P. M. D. Fitzgerald, B. M. McKeever, C.-T. Leu, J. C. Heimbach, W. K. Herber, I. S. Sigal, P. L. Darke, and J. P. Springer. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. Nature, 337:615–620, 1989. https://doi.org/10.1038/337615a0.

[10] A. Wlodawer, M. Miller, M. Jáskolski, B. K. Sathyanarayana, E. Baldwin, I. T. Weber, L. M. Selk, L. Clawson, J. Schneider, and S. B. H. Kent. Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. Science, 245(4918):616–621, 1989. https://doi.org/10.1126/science.2548279.

[11] R. Ishima, D. I. Freedberg, Y. X. Wang, J. M. Louis, and D. A. Torchia. Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function. Structure, 7(9):1047–1055, 1999. https://doi.org/10.1016/S0969-2126(99)80172-5.

[12] V. Hornak, A. Okur, R. C. Rizzo, and C. Simmerling. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. Proc. Natl. Acad. Sci. USA, 103(4):915–920, 2006. https://doi.org/10.1073/pnas.0508452103.

[13] J.-H. Lin, A. L. Perryman, J. R. Schames, and J. A. McCammon. Computational drug design accommodating receptor flexibility: the relaxed complex scheme. J. Am. Chem. Soc., 124(20):5632–5633, 2002. https://doi.org/10.1021/ja0260162.

[14] R. E. Amaro, J. Baudry, J. Chodera, O. Demir, J. A. McCammon, Y. Miao, and J. C. Smith. Ensemble docking in drug discovery. Biophys. J., 114(10):2271–2278, 2018. https://doi.org/10.1016/j.bpj.2018.02.038.

[15] A. Krohn, S. Redshaw, J. C. Ritchie, B. J. Graves, and M. H. Hatada. Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (R)-hydroxyethylamine isostere. J. Med. Chem., 34(11):3340–3342, 1991. https://doi.org/10.1021/jm00115a028. 1HSG: HIV-1 PR/indinavir complex, PDB 1994.

[16] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA, 98(18):10037–10041, 2001. https://doi.org/10.1073/pnas.181342398.

[17] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, D. W. Gohara, T. Dolinsky, R. Konecny, D. R. Koes, J. Erik Nielsen, T. Head-Gordon, W. Geng, R. Krasny, G.-W. Wei, M. J. Holst, J. A. McCammon, and N. A. Baker. Improvements to the APBS biomolecular solvation software suite. Protein Sci., 27(1):112–128, 2018. https://doi.org/10.1002/pro.3280.

[18] B. J. Frey and D. Dueck. Clustering by passing messages between data points. Science, 315(5814):972–976, 2007. https://doi.org/10.1126/science.1136800.

[19] L. McInnes, J. Healy, N. Saul, and L. Großberger. UMAP: Uniform manifold approximation and projection for dimension reduction. J. Open Source Softw., 3(29):861, 2018. https://doi.org/10.21105/joss.00861.

[20] S. Salentin, S. Schreiber, V. J. Haupt, M. F. Adasme, and M. Schroeder. PLIP: fully automated protein–ligand interaction profiler. Nucleic Acids Res., 43(W1):W443–W447, 2015. https://doi.org/10.1093/nar/gkv315.

[21] G. Landrum. RDKit: Open-source cheminformatics software, 2006. URL https://www.rdkit.org.

[22] R. W. Shafer and J. M. Schapiro. HIV-1 drug resistance mutations: an updated framework for the second decade of HAART. AIDS Rev., 10(2):67–84, 2008.

[23] P. Řezáčová, J. Pokorná, J. Brynda, M. Kožíšek, P. Cígler, M. Lepšík, J. Fanfrlík, J. Řezáč, K. Grantz Šašková, I. Sieglová, J. Plešek, V. Šícha, B. Grüner, H. Oberwinkler, J. Sedláček, H.-G. Kräusslich, P. Hobza, V. Král, and J Konvalinka. Design of HIV protease inhibitors based on inorganic polyhedral metallacarboranes. J. Med. Chem., 52(22):7132–7141, 2009. https://doi.org/10.1021/jm9011388.

Data Availability Statement: All data are available on request from the author.
Conflicts of Interest: The author declares no conflict of interest.