**Supporting information for**

# Exploring the binding pathway of novel non-peptidomimetic plasmepsin V inhibitors

Raitis Bobrovs[1]*, Laura Drunka[1], Iveta Kanepe[1], Aigars Jirgensons[1], Amedeo Caflisch[2], Matteo Salvalaglio[3], Kristaps Jaudzems[1]

*[1] Latvian Institute of Organic Synthesis, Aizkraukles 21, Riga, LV1006, Latvia*
*[2] Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057, Zurich, Switzerland*
*[3] Thomas Young Centre and Department of Chemical Engineering, University College London, London WC1E 7JE, United Kingdom*

E-mail: raitis.bobrovs@osi.lv



## 1 Methods

### *1.1 Open flap plm V structure preparation*

Plm V crystal structure 4ZL4[1] was prepared using Maestro Protein Preparation Wizard[2] by adding missing side chains using Prime[3], adjusting side chain protonation states at pH 7.0, and minimising heavy atoms with convergence up to 0.30 Å. Plm V crystal structure with inhibitor and water molecules removed was aligned with plm II crystal structure 4Z22[4] in Schrodinger Maestro software[5]. The coordinates of the plm II non-peptidomimetic amino quinazolinone inhibitor were copied to plm V crystal structure, and prepared complex was minimised using Schrodinger Prime[3]. Further, the amino quinazolinone DR720 that was experimentally verified to inhibit plm V was aligned with the inhibitor of the prepared complex, and complex that contains plm V and DR720 coordinates was obtained. This complex was also minimised using Schrodinger Prime.

Further, the prepared plm V-DR720 complex was subjected to 100 ns molecular dynamics (MD) simulation to optimise binding site residue and inhibitor position. Restraints were applied to ligand core and aspartic dyad intermolecular distances to ensure that the ligand remains bound in the binding site throughout the simulation. Harmonic restraints were applied to $5B*B:_Amino group nitrogen, $5VW:_Amino group nitrogen, and $5VW:_Pyrimidine nitrogen interatomic distances at 2 Å with a force constant of 25 kcal/mol. MD simulation system setup and parameters used are described in section 1.3.

### *1.2 High-Throughput Virtual Screening (HTVS)*

MolPort in-stock screening compound library of more than 6 million compounds (2020) was prepared using LigPrep[6] by desalting the molecules, generating possible tautomers and ionisation states at pH 7.0 ± 2.0. The stereochemistry of the compounds was retained as specified in the library. The prepared library was docked in the open flap plm V structure generated.

Molecular docking was performed using Glide[7], with scaling of the van der Waals radii set to 0.9 for protein and ligand heavy atoms, and docking compounds flexibly. The top-scoring 3000 compounds were clustered to 300 representative compounds by calculating the Linear Fingerprints from Daylight invariant atom types and evaluating compound similarity using Tanimoto similarity metrics. The top-scoring compound was retained for each cluster. The top-ranked 300 representative compounds were visually inspected for their ability to form interaction similar to DR720, with molecules showing internal strains or unsatisfied hydrogen bond donors being deprioritised. A total of 28 potential plm V binders were selected for purchase. Docked poses were visualised using PyMOL[8].

### *1.3 Molecular dynamics (MD) simulations*

The MD simulation systems were prepared by placing the complex in dodecahedral boxe with at least 1.5 nm distance to the box walls. The TIP3P water model was used to solvate the complexes. Sodium and chloride ions were added to neutralise the system and reach 150 mM salt concentration. Forcefields for the inhibitors were based on the general AMBER force field (GAFF) and were generated using Ambertools[9]. Amber03 forcefield parameters were used for protein[10,11]. The prepared systems were relaxed through an energy minimisation, which was performed using the steepest descent algorithm with a tolerance of 100 kJ/mol·nm. After minimisation, systems were equilibrated in the NVT and then NPT ensembles for 5 ns. The MD (leapfrog) integration scheme with an integration time step of 2 fs was employed for equilibration and production runs. The particle mesh Ewald (PME) approach was used to calculate long-range electrostatic interactions with a cut-off of 0.8 nm. Both Lennard–Jones and Coulomb interactions were explicitly calculated up to 0.8 nm. The LINCS algorithm[12] was applied at each step to preserve the hydrogen bond lengths. NPT equilibration was performed employing a Berendsen barostat[13] with a coupling constant of 2 ps and reference pressure 1.0 bar. Velocity-rescale thermostat[14] with a coupling constant of 2 ps and reference temperature 298.0 K was used for equilibration and production simulations. The production run was performed in the NPT ensemble. The potential energy minimization and MD simulations were carried out with the software package Gromacs 2021[15,16] patched with Plumed 2.7[17].

### *1.4 Collective variables (CVs)*

#### *Path CV*

Here inhibitor binding/unbinding process was simulated using Path metadynamics (Path MetaD) approach, where simulation is biased along a predefined path. The reference path, consisting of 15 equally spaced frames, was prepared from a preliminary ligand unbinding simulation using Plumed pathtools, and RMSD-based PathCV was used to describe the position of a point in configurational space relative to the reference path. Reference path was created by selecting only coordinates of : gatoms of hsheets within 15 Å from Ser87 and two ligand core atoms (hydrogen bond donor and acceptor interacting with catalytic dyad). This allowed to reduce the computational cost associated with RMSD calculation and provided an opportunity to use the same path for all ligands.

The two CVs used in Path MetaD are: *s* – the progress along the predefined reference path; and *z* – the distance orthogonal to the reference path. Introduction of the CV z allows to explore configurations that differ from the reference path, thus, if reference path provided is not completely accurate, system is able to deviate from the predefined path and explore states that are more favourable than the defined ones.

***Funnel CV***

The preliminary ligand unbinding trajectory was generated using Funnel metadynamics (FunMetaD) method, where space explored by the ligand is limited using funnel shaped restraint potential[18]. More in-depth description of FunMetaD application to study inhibitor binding to aspartic proteases is described elsewhere[19]. In short, the parameters defining the funnel were as follows: the funnel anchor point was set to $ 5 V Wdwrp, the funnel axis *z*, was defined in the space by two points, one corresponding to the centre of mass of $ 5 V Wdqg $ 5 B * Bdqg the other close to position of b < . * P the distance where the restraint potential switches from a cone shape into a cylinder ($z_{cc}$) was set to 2.0 nm; the angle M gHdt defines the amplitude of the cone was set to 0.6 rad; the radius of the cylindrical section of the funnel ($R_{cyl}$) was set to 0.1 nm. Restraining walls were applied at both ends of the funnel axis at 4.5 nm and j W GnQ. Input files for the Funnel MetaD were partly prepared using FMAP tool[20].

### *1.5 Metadynamics (metaD)*

The Plumed plugin[17] was used to carry out metadynamics calculations. The bias was added to PathCV components *s* and *z*, and the respective Gaussian widths were set to 0.1 and 0.001 Å. Gaussians were deposited every 1 ps in the well-tempered scheme[21] with a bias factor of 10 and initial Gaussian height set at 3 kJ/mol. A soft harmonic restraining bias was applied on the *z* variable at 0.1 nm to prevent ligand deviation too far away from the reference path, while enabling the possibility for the system to explore conformational space different from the original path. The RMSD between the reference path provided and system studied was calculated after coordinate superimposition every 1 ps.

Multiple ligand binding/unbinding events were observed within each simulation, and converged FES was typically obtained after 1000 ns long simulation. Simulation on 1GPU and 6 CPU cores did run at a speed of ~135 ns/day, thus converged simulation could be obtained after ~1 week of calculations. The simulations were reweighed[22] as a function of selected variables using Plumed driver tool. Trajectories were analysed using VMD software[23] and figures were prepared using Pymol[8] and Matplotlib[24] software.

For the FunMetaD simulation, the bias was added on a distance CV *d1* with a Gaussian width set to 0.03 nm. Gaussians were deposited every 1 ps in the well-tempered scheme[27] with a bias factor of 15 and initial Gaussian height set at 5 kJ/mol. A soft harmonic restraining bias was applied on the *d1* variable at 4.0 nm to prevent the ligand escape from the funnel.

### *1.6 Sketch-map*

2D projections showing ligand binding modes and their connectivity were generated using nonlinear dimensionality reduction algorithm sketch-map[25,26]. 2D projections were generated from interatomic distances between ligand transition state mimetic group centre of mass and each protein binding site atom that was used in PathMetaD reference path (51 distances). Only frames where ligand-catalytic site distance was less than 3.0 nm were used in sketch-map generation (a set of 13069 frames extracted from 7 simulations of verified plm V inhibitors; extracted using Plumed). The workflow of sketch-map generation followed protocol described elsewhere[25,26]. In brief, the dissimilarities between the frames were computed, and farthest point sampling was used to select 50 landmark points. Then, the

nonlinear sketch-map optimisations were performed iteratively on landmark points until convergence of the low-dimensional projections. At the end, the remaining frames of the trajectory were projected on the optimised sketch-map using out of sample embedding. Mapping of verified inhibitor trajectories on top of reference map were performed in a similar manner, with an exception that landmark points and weights of reference map were used. The final 2D configuration space was coloured by the ligand-catalytic dyad interatomic distance. Sketch-map calculations were performed using Plumed development version (03.2023). Plumed input files are available via PLUMED-NEST[27] (https://www.plumed-nest.org), the public repository for the Plumed consortium, using the project plumID: 23.019.

### *1.7 Protein expression and purification*

Plasmepsin (plm) II and IV was expressed and purified as described by Beyer *et al.*[28]. Briefly, pET3a plasmid containing plm gene was transformed into BL21(DE3) *E. coli* strain and cultured on agar plate containing 100 m 6 ` a2m-picillin overnight at 37 °C. One colony was inoculated in 25 mL of LB medium containing 100 m 6 ` a2m-picillin and grown overnight at 30 °C, 200 rpm. 10 mL of overnight culture were transferred to 1 L LB medium containing 100 m 6 ` a2m-picillin and incubated at 37 °C, 200 rpm until $OD_{600}$ is around 0.6. Protein expression was induced with 0.3 mM IPTG and incubated for additional 3 h. Bacteria were harvested by centrifugation (15 min at 6000 g, 4 °C), resuspended in lysis buffer A1 (10 mM Tris-HCl, pH 8.0, 20 mM $MgCl_2$, 5 mM $CaCl_2$) and lysed by ultrasonication in ice cold bath. Inclusion bodies were harvested by centrifugation over 27% sucrose cushion (30 min at 12000 g, 4 °C), washed subsequently with resuspension buffer B1 (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1 mM DTT, 100 mM NaCl) and resuspension buffer C1 (50 mM Tris-HCl, pH 8.0, 5 mM EDTA, 2.5 mM DTT, 0.5% Triton X-100). Inclusion bodies were solubilized in buffer D (50 mM CAPS, pH 10.5, 8 M urea, 100 mM DTT, 5 mM EDTA), refolded by dialysis against 5 L refolding buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 2.5 mM DTT) and applied to HiTrap Q HP anion exchange column equilibrated with chromatography buffer A (20 mM Tris-HCl, pH 8.0, 2.5 mM DTT). Fractions containing the protein were concentrated and added to activation buffer (100 mM sodium acetate, pH 5.0). Mixture was incubated at room temperature with agitation for 1 h, centrifuged to remove precipitant (5 min at 10000 g, 25 °C) and supernatant was applied to HiLoad 16/600 Superdex 200 pg column equilibrated in buffer C (20 mM Tris-Cl, pH 8.0, 150 mM NaCl, 4 mM DTT). Collected fractions were analysed with SDS-PAGE. Protein containing fractions were concentrated and used for further experiments.

Plm V expression and purification protocol was based on Loymunkong *et al.*[29] with several modifications. BL21(DE3) *E. coli* cells were transformed with the vector containing plm V and ampicillin resistance genes and then grown overnight on agar plates containing 100 µg/mL ampicillin. One fresh colony was inoculated in 25 mL LB medium containing 100 µg/mL ampicillin and incubated overnight at 37 °C, 200 rpm. Further, 10 mL of overnight culture were transferred to 3 L LB medium containing 100 µg/mL ampicillin and grown at 37 °C, 200 rpm until $OD_{600}$ was 0.4-0.5, and then temperature was reduced to 16 °C. Cells were induced with 0.2 mM IPTG when $OD_{600}$ was 0.7-0.8, and grown for 20 h. Cells were harvested by centrifugation (15 min at 7000 g, 4 °C). Pellets were resuspended in buffer A1 (50 mM Tris-Cl, pH 8.5, 0.1% Triton (w/v), 500 mM NaCl, 4 mM DTT) in a ratio 1 g of cells per 10 mL buffer, and lysed by ultrasonication in ice cold bath. Lysis solution was centrifuged for 40 min at 30000 g, 4 °C. Supernatant was collected and purified using nickel affinity HisTrap HP column, with buffer A (50 mM Tris-Cl, pH 8.5, 500 mM NaCl, 4 mM DTT) as equilibration buffer. Protein was eluted with linear gradient (0% to 100% for 40 min) against buffer B (50 mM Tris-Cl, pH 8.5, 500 mM NaCl, 500 mM Imidazole, 4 mM DTT). Fractions were collected and analysed by SDS-PAGE gel. Fractions containing plm V were concentrated to 2 mL or less at 4 °C, and applied to size exclusion chromatography using

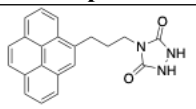
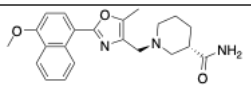
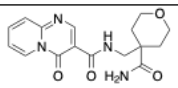
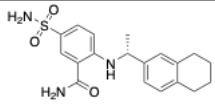
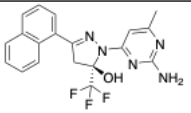
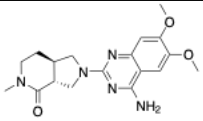
HiLoad 16/600 Superdex 200 pg column column equilibrated in Buffer C (50 mM Tris-Cl, pH 7.5, 300 mM NaCl, 10 mM 2-beta-mercaptoethanol). Flow speed was 0.6 mL/min. Collected fractions were analysed by SDS-PAGE gel, and fractions containing plm V monomer were concentrated to around 0.5-1 mg/mL and used for further experiments.

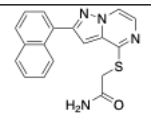
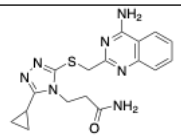
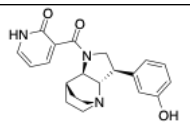
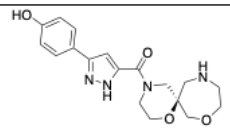
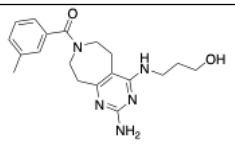
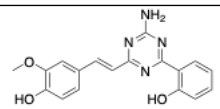
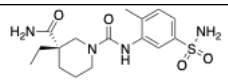
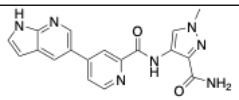
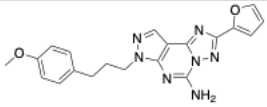
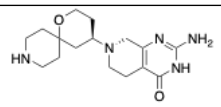
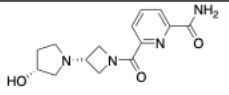
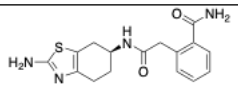
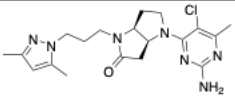
### *1.8 Enzymatic assay*

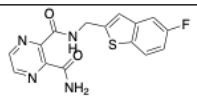
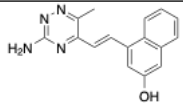
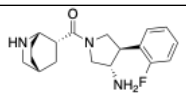
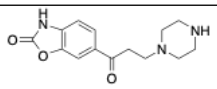
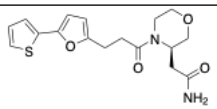
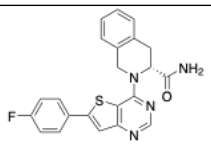
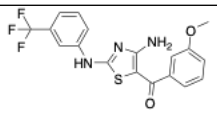
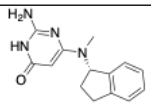
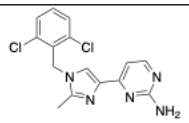
A fluorescence resonance energy transfer (FRET) assay was performed to evaluate ability of compounds to inhibit plm II, IV, V and catD. A solution of compounds for testing on white 96 well plate was added to the purified recombinant enzyme plm II, IV or catD (Sigma, cat.nr. C8696) in 0.1M NaOAc buffer, pH 4.5, 10% glycerol. Recombinant plm V reaction buffer was 25 mM Tris, 25 mM MES, pH 6.4, 5mM DTT, 0.005% Tween20. The mixture was incubated for 30 min at 37 °C. Substrate DABCYL-Glu-Arg-Nle-Phe-Leu-Ser-Phe-Pro-EDANS for plm II, IV, catD, and DABCYL-Leu-Asn-Lys-Arg-Leu-Leu-His-Glu-Thr-Gln-EDANS for plm V (AnaSpec Inc) was then added to reach a final concentration of 5 µM for plm II, IV and catD, and 10 µM for plm V. Hydrolysis of the substrate was detected as an increase in fluorescence (Em 490 nm, Ex 336 nm) at 37 °C. The data points were collected every 1 min over the period of 15 min (60 min for plm V). For the rate calculation, only linear interval was used, which was slightly different for each enzyme. Inhibitors, dissolved in DMSO, were added to reaction to reach 100 µM concentration (4% DMSO in the final solution), and were tested in duplicate. $IC_{50}$ values were determined for compounds with a higher than 50% inhibitory effect. Compounds were tested in three repeated triplicate experiments. $IC_{50}$ values were calculated using software Graph Pad Prism 5.0.

## 2 Additional data

*Table SI1. The inhibition potency of commercially obtained HTVS hits against plm II, IV V and human cathepsin D.*

| | | **$IC_{50}$, % &** | | | |
|---|---|---|---|---|---|
| **No** | **Compound** | **plm V** | **plm II** | **plm IV** | ***h*catD** |
| 1 |  MolPort-002-904-606 | 4.4±0.7 | 41±6 | 87±11 | 40±8 |
| 2 |  MolPort-019-900-307 | 6.7±0.5 | 17±4 | 26±5 | 8±2 |
| 3 |  MolPort-020-062-340 | 14.7±1.0 | >100 | 71±2 | 21±3 |
| 4 |  MolPort-023-187-757 | 14.8±0.7 | 78±5 | 35±4 | 24±4 |
| 5 |  MolPort-000-124-439 | 16.1±1.4 | >100 | >100 | >100 |
| 6 |  | 23.7±2 | >100 | >100 | >100 |

| | | | | | |
|---|---|---|---|---|---|
| | MolPort-046-754-050 | | | | |
| 7 |  MolPort-010-720-952 | 70.0±3 | 65±4 | >100 | 54±4 |
| 8 |  MolPort-007-247-852 | >100 | | | |
| 9 |  MolPort-021-769-369 | >100 | | | |
| 10 |  MolPort-020-225-228 | >100 | | | |
| 11 |  MolPort-021-747-521 | >100 | | | |
| 12 |  MolPort-004-973-679 | >100 | | | |
| 13 |  MolPort-029-897-916 | >100 | | | |
| 14 |  MolPort-046-536-743 | >100 | | | |
| 15 |  MolPort-023-276-442 | >100 | | | |
| 16 |  MolPort-047-485-514 | >100 | | | |
| 17 |  MolPort-046-900-588 | >100 | | | |
| 18 |  MolPort-046-467-283 | >100 | | | |
| 19 |  | >100 | | | |

| | | |
|---|---|---|
| | MolPort-047-388-850 | |
| 20 |  MolPort-027-901-656 | >100 |
| 21 |  MolPort-003-873-222 | >100 |
| 22 |  MolPort-040-820-532 | >100 |
| 23 |  MolPort-008-721-604 | >100 |
| 24 |  MolPort-021-769-041 | >100 |
| 25 |  MolPort-005-607-857 | >100 |
| 26 |  MolPort-002-576-690 | >100 |
| 27 |  MolPort-046-913-712 | >100 |
| 28 |  MolPort-002-853-158 | >100 |

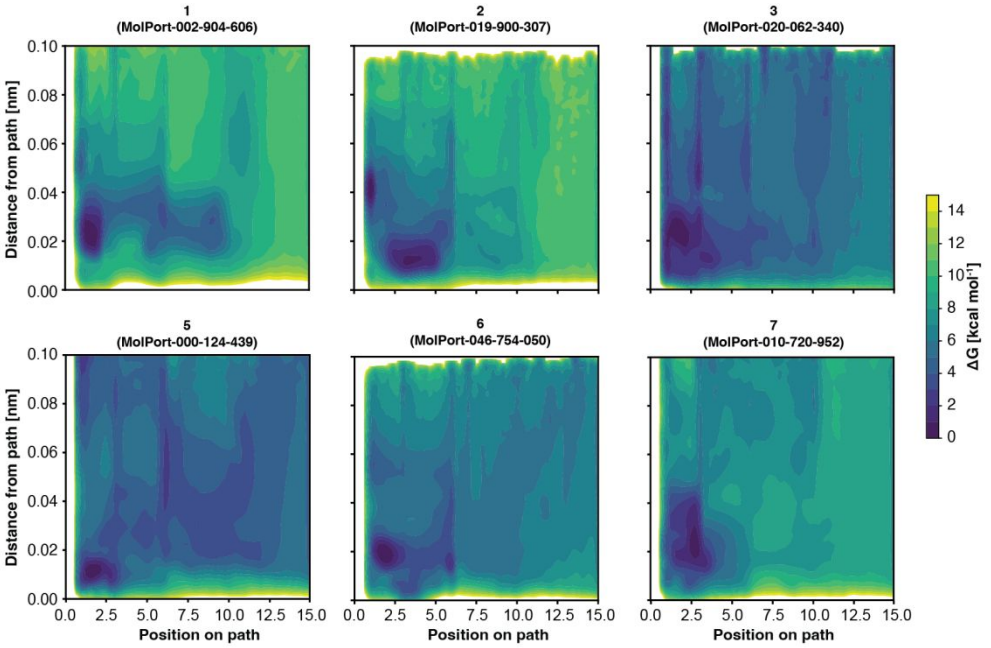



*Figure SI1. The FES of verified plm V inhibitors obtained after reweighing the path metaD simulations. Isosurfaces are shown every 1 kcal/mol.*

*Table SI2. The binding free energy X Y Z of the HTVS hits, and their estimated binding energy (calculated from $IC_{50}$ values as Y Z \ G R T ln($IC_{50}$)).*

| | Comp. ID | A B calculated | A B from $IC_{50}$ values |
|---|---|---|---|
| active comp. | MolPort-002-904-606 | 9.0±0.3 | 7.3 |
| | MolPort-019-900-307 | 8.0±0.3 | 7.1 |
| | MolPort-020-062-340 | 5.5±0.5 | 6.6 |
| | MolPort-023-187-757 | 7.9±0.5 | 6.6 |
| | MolPort-000-124-439 | 3.7±0.2 | 6.5 |
| | MolPort-046-754-050 | 5.6±0.2 | 6.3 |
| | MolPort-010-720-952 | 6.6±0.3 | 5.7 |
| inactive comp. | MolPort-002-576-690 | 4.7±0.4 | <5.5 |
| | MolPort-002-853-158 | 5.5±0.3 | <5.5 |
| | MolPort-003-873-222 | 4.6±0.2 | <5.5 |
| | MolPort-004-973-679 | 3.0±0.2 | <5.5 |
| | MolPort-005-607-857 | 5.7±0.3 | <5.5 |
| | MolPort-007-247-852 | 2.4±0.2 | <5.5 |
| | MolPort-008-721-604 | 4.6±0.4 | <5.5 |
| | MolPort-020-225-228 | 3.9±0.3 | <5.5 |
| | MolPort-021-747-521 | 6.3±0.3 | <5.5 |
| | MolPort-021-769-041 | 1.1±0.2 | <5.5 |
| | MolPort-021-769-369 | 6.1±0.3 | <5.5 |
| | MolPort-023-276-442 | 4.2±0.3 | <5.5 |

| | | |
|---|---|---|
| MolPort-027-901-656 | 5.4±0.3 | <5.5 |
| MolPort-029-897-916 | 4.0±0.2 | <5.5 |
| MolPort-040-820-532 | 3.7±0.3 | <5.5 |
| MolPort-046-467-283 | 0.76±0.12 | <5.5 |
| MolPort-046-536-743 | 4.2±0.3 | <5.5 |
| MolPort-046-900-588 | 2.3±0.2 | <5.5 |
| MolPort-046-913-712 | 2.1±0.3 | <5.5 |
| MolPort-047-388-850 | 5.1±0.6 | <5.5 |
| MolPort-047-485-514 | 6.7±0.4 | <5.5 |

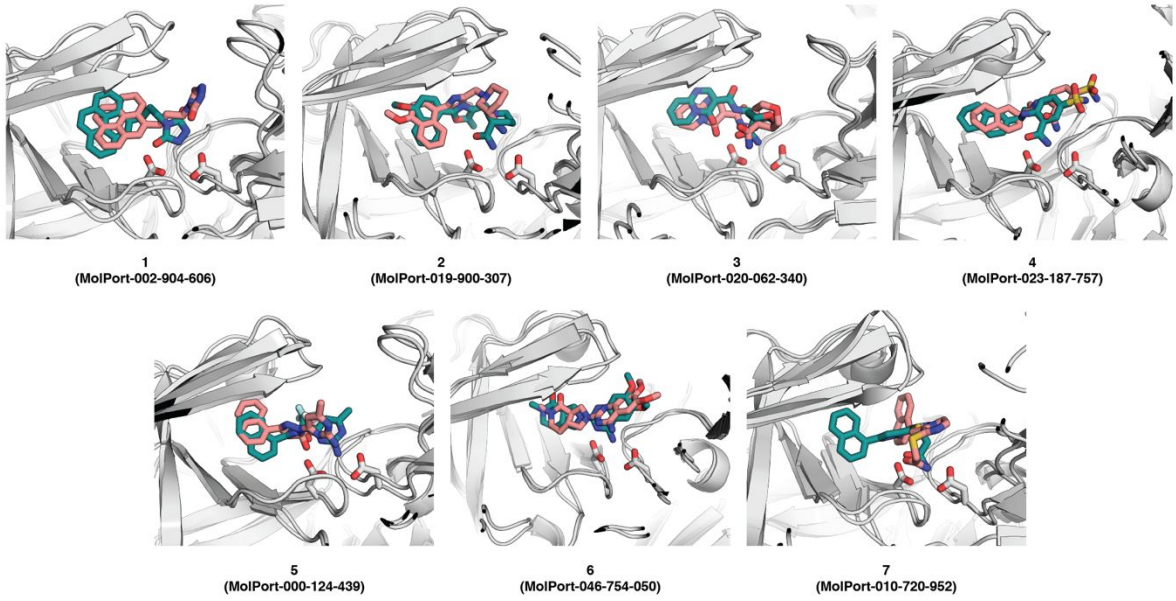



*Fig SI2. The estimated binding modes of the verified plm V inhibitors identified using docking (teal) and metadynamics (salmon) studies. The ligand and catalytic dyad residue side chains are shown as sticks. Hydrogens are omitted for clarity.*

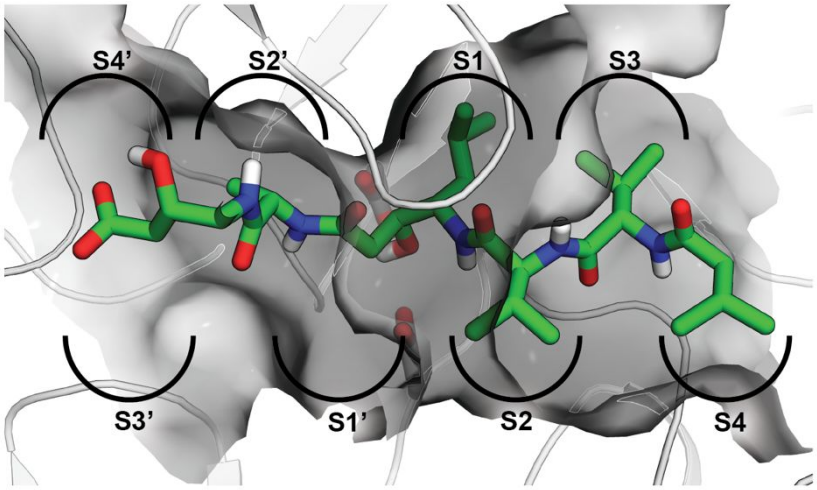


*Figure SI3. Aspartic protease binding sub-site naming scheme. Substrate mimetic pepstatin A is shown as green sticks; aspartic protease (plasmepsin II; PDB ID: 1SME) as cartoon; catalytic dyad as white sticks. Hydrogens are omitted for clarity.*

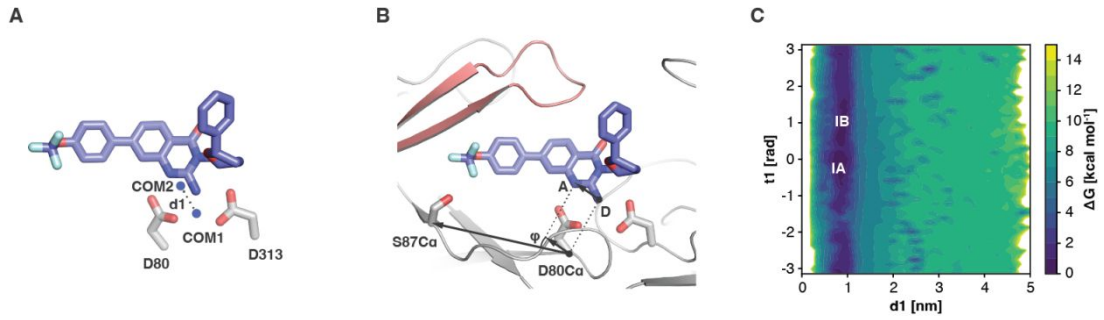
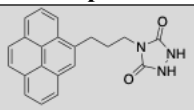
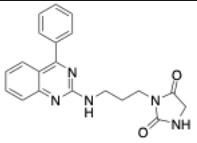



*Figure SI4. **A** Distance (d1) between the centre of mass of the E atoms of catalytic dyad residues (Asp80 and Asp313, COM1) and centre of mass of ligand transition state mimetic group (COM1). **B** Torsion b between the ligand transition mimetic group and flap pocket axis defined as a vector between Asp80 and Ser87 E catoms. **C** FES of compound MolPort-023-187-757 binding to plm V reweighed as a function of the ligand core-catalytic dyad distance d1 and torsion b (ligand alignment with respect to the flap pocket axis). Isosurfaces are shown for every 1 kcal/mol. The deepest FES basins are indicated as IA and IB, and corresponding binding modes are shown in main text Fig. 6.*

### *2.1 Compound 1, 5 and 6 structure-activity relationship (SAR) studies*

Of the seven hits identified, commercially available analogues were purchased for compounds **1**, **5** and **6** (12, 24 and 15 analogues, respectively) due to their selectivity over human cathepsin D and higher medicinal chemistry development potential. The analogues selected had modifications in the expected flap pocket substituent position, whereas the transition state mimetic part of the molecule was not changed. The majority of the compound **1** and **5** analogues showed lower plm V inhibitory potency than the original hits (see Tables SI2, SI3), and only MolPort-019-894-150 was slightly more potent ($IC_{50}$ (plm V) = 13.0±0.8 m@tNan the parent compound. However, for compound **6**, nearly all tested analogues showed higher plm V inhibition potency than the parent compound. The most active compound was compound MolPort-035-715-983 with an $IC_{50}$ of 5.0±0.3 m@(see Table SI4).

*Table SI3. The inhibition potency of commercially obtained MolPort-002-904-606 analogues against plm II, IV V and human cathepsin D.*

| No | Compound | $IC_{50}$, % & | | | |
|---|---|---|---|---|---|
| | | **plm V** | **plm II** | **plm IV** | ***h*catD** |
| Hit 1 |  MolPort-002-904-606 | 4.4±0.7 | 41±6 | 87±11 | 40±8 |
| 1 |  MolPort-035-742-529 | 17±0.8 | >100 | >100 | 116±4 |

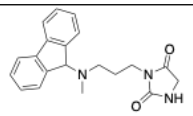
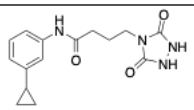
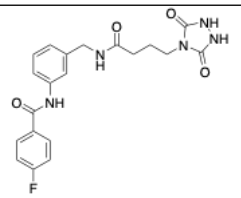
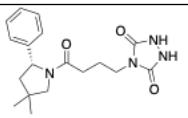
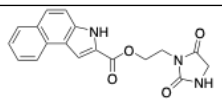
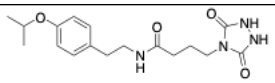
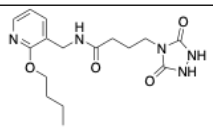
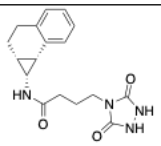
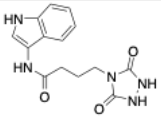
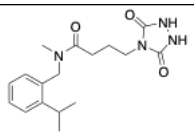
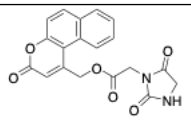
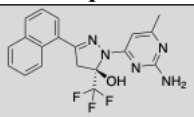
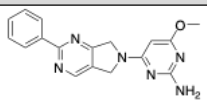
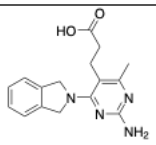
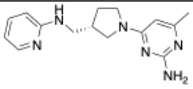
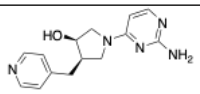
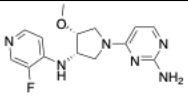
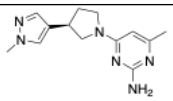
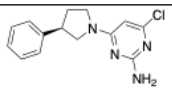
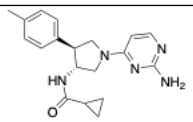
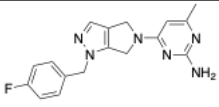
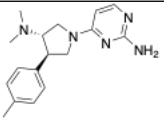
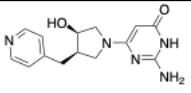
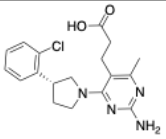
| | | |
|---|---|---|
| 2 |  MolPort-028-305-898 | >100 |
| 3 |  MolPort-046-827-503 | >100 |
| 4 |  MolPort-046-907-266 | >100 |
| 5 |  MolPort-047-472-186 | >100 |
| 6 |  MolPort-035-834-467 | >100 |
| 7 |  MolPort-044-534-139 | >100 |
| 8 |  MolPort-044-534-140 | >100 |
| 9 |  MolPort-046-495-512 | >100 |
| 10 |  MolPort-047-570-307 | >100 |
| 11 |  MolPort-047-758-654 | >100 |
| 12 |  MolPort-005-323-470 | >100 |

*Table SI4. The inhibition potency of commercially obtained MolPort-000-124-439 analogues against plm II, IV V and human cathepsin D.*

**$IC_{50}$,** % &

| No | Compound | plm V | plm II | plm IV | *h*catD |
|---|---|---|---|---|---|
| Hit 5 |  MolPort-000-124-439 | 16.1±1.4 | >100 | >100 | >100 |
| 1 |  MolPort-019-894-150 | 13.0±0.8 | >100 | >100 | >100 |
| 2 |  MolPort-028-598-146 | 49±2 | >100 | >100 | >100 |
| 3 |  MolPort-030-003-057 | 100 | >100 | >100 | >100 |
| 4 |  MolPort-046-557-088 | >100 | | | |
| 5 |  MolPort-046-922-478 | >100 | | | |
| 6 |  MolPort-046-075-586 | >100 | | | |
| 7 |  MolPort-027-585-343 | >100 | | | |
| 8 |  MolPort-023-254-453 | >100 | | | |
| 9 |  MolPort-019-799-967 | >100 | | | |
| 10 |  MolPort-023-253-194 | >100 | | | |
| 11 |  MolPort-046-575-849 | >100 | | | |
| 12 |  MolPort-028-580-527 | >100 | | | |

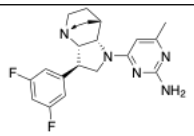
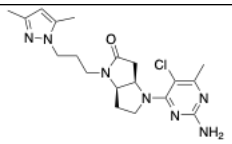
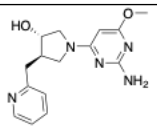
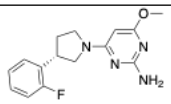
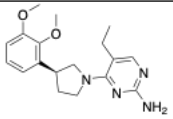
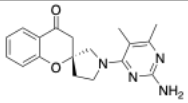
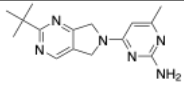
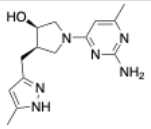
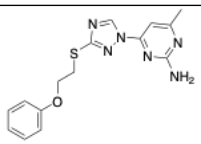
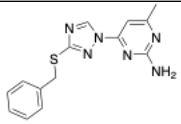
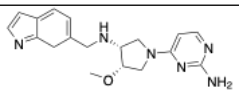
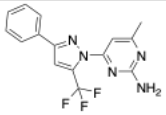
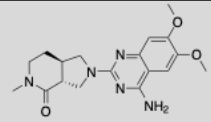
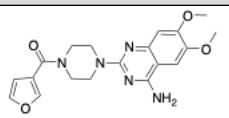
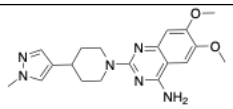
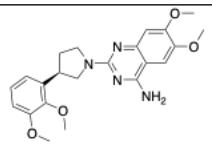
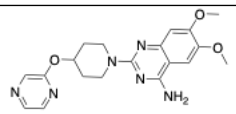
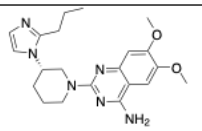
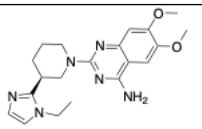
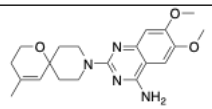
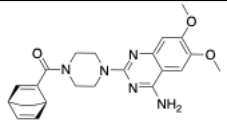
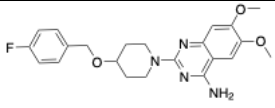
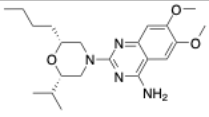
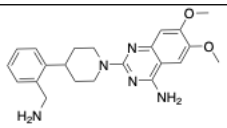
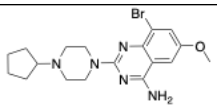
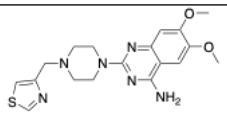
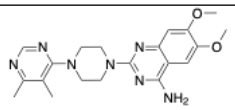
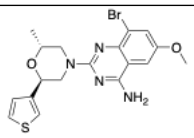
| | | |
|---|---|---|
| 13 |  MolPort-020-186-233 | >100 |
| 14 |  MolPort-047-388-850 | >100 |
| 15 |  MolPort-046-571-298 | >100 |
| 16 |  MolPort-028-739-203 | >100 |
| 17 |  MolPort-027-861-426 | >100 |
| 18 |  MolPort-027-863-316 | >100 |
| 19 |  MolPort-019-806-219 | >100 |
| 20 |  MolPort-046-555-894 | >100 |
| 21 |  MolPort-007-703-633 | >100 |
| 22 |  MolPort-007-703-630 | >100 |
| 23 |  MolPort-046-921-939 | >100 |
| 24 |  MolPort-000-124-433 | >100 |

*Table SI5. The inhibition potency of commercially obtained MolPort-046-754-050 analogues against plm II, IV V and human cathepsin D.*

| No | Compound | $IC_{50}$, % & | | | |
|---|---|---|---|---|---|
| | | **plm V** | **plm II** | **plm IV** | ***h*catD** |
| Hit 6 |  MolPort-046-754-050 | 23.7±2 | >100 | >100 | >100 |
| 1 |  MolPort-035-715-983 | 5.0±0.3 | >100 | 115±5 | 48±3 |
| 2 |  MolPort-046-074-271 | 5.6±0.4 | 9.5±1.6 | 11.1±1.6 | 11.7±1.3 |
| 3 |  MolPort-047-516-522 | 6.4±0.4 | 54±3 | 48±3 | 42±4 |
| 4 |  MolPort-046-812-445 | 7.0±0.4 | 96±5 | 33±4 | 31±4 |
| 5 |  MolPort-047-626-663 | 7.0±0.5 | >100 | >100 | >100 |
| 6 |  MolPort-028-733-980 | 7.2±0.4 | 61±4 | 97±5 | >100 |
| 7 |  MolPort-046-739-746 | 8.1±0.5 | 61±4 | 72±5 | >100 |
| 8 |  MolPort-003-849-966 | 8.5±0.5 | 118±4 | 66±3 | >100 |
| 9 |  MolPort-046-805-638 | 8.7±0.5 | 41±2 | 31±2 | >100 |
| 10 |  MolPort-046-520-834 | 9.3±0.6 | 23±3 | 25±3 | 13.4±1.0 |

| | | | | | |
|---|---|---|---|---|---|
| 11 |  MolPort-047-626-662 | 11.5±0.6 | 23±3 | 20.2±1.4 | 16±2 |
| 12 |  MolPort-030-037-821 | 12.7±0.6 | 62±3 | 50±3 | 28±3 |
| 13 |  MolPort-046-440-107 | 17.0±0.7 | 101±4 | 86±4 | 31±5 |
| 14 |  MolPort-046-529-437 | 29.0±1.2 | >100 | >100 | 120±5 |
| 15 |  MolPort-047-554-734 | >100 | >100 | 142±5 | 97±4 |

## 3 References


(1) Hodder, A. N.; Sleebs, B. E.; Czabotar, P. E.; Gazdik, M.; Xu, Y.; O'Neill, M. T.; Lopaticki, S.; Nebl, T.; Triglia, T.; Smith, B. J.; Lowes, K.; Boddey, J. A.; Cowman, A. F. Structural Basis for Plasmepsin v Inhibition That Blocks Export of Malaria Proteins to Human Erythrocytes. *Nat Struct Mol Biol* **2015**, *22* (8), 590–596. https://doi.org/10.1038/nsmb.3061.

(2) Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J Comput Aided Mol Des* **2013**, *27* (3), 221–234. https://doi.org/10.1007/s10822-013-9644-8.

(3) Schrödinger LLC. Prime, Schrödinger Release 2020-4. New York, NY 2020.

(4) Rasina, D.; Otikovs, M.; Leitans, J.; Recacha, R.; Borysov, O. V.; Kanepe-Lapsa, I.; Domraceva, I.; Pantelejevs, T.; Tars, K.; Blackman, M. J.; Jaudzems, K.; Jirgensons, A. Fragment-Based Discovery of 2-Aminoquinazolin-4(3H)-Ones As Novel Class Nonpeptidomimetic Inhibitors of the Plasmepsins I, II, and IV. *J Med Chem* **2016**, *59* (1), 374–387. https://doi.org/10.1021/acs.jmedchem.5b01558.

(5) Schrödinger LLC. Maestro, Schrödinger Release 2020-4. New York, NY 2020.

(6) Schrödinger LLC. LigPrep, Schrödinger Release 2020-4. Schrödinger, LLC: New York 2020.

(7) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J Med Chem* **2004**, *47* (7), 1750–1759. https://doi.org/10.1021/jm030644s.

(8) Schrödinger LLC. *The PyMOL Molecular Graphics System, Version 2.0*; 2020.

(9) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J Comput Chem* **2005**, *26* (16), 1668–1688. https://doi.org/10.1002/jcc.20290.

(10) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J Comput Chem* **2004**, *25* (9), 1157–1174. https://doi.org/10.1002/jcc.20035.

(11) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber Ff99SB Protein Force Field. *Proteins: Structure, Function, and Bioinformatics* **2010**, *78* (8), 1950–1958. https://doi.org/10.1002/prot.22711.

(12) Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J Chem Theory Comput* **2008**, *4* (1), 116–122. https://doi.org/10.1021/ct700200b.

(13) Berendsen, H. J. C. C.; Postma, J. P. M. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J Chem Phys* **1984**, *81* (8), 3684–3690. https://doi.org/10.1063/1.448118.

(14) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J Chem Phys* **2007**, *126* (1), 014101. https://doi.org/10.1063/1.2408420.

(15) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J Comput Chem* **2005**, *26* (16), 1701–1718. https://doi.org/10.1002/jcc.20291.

(16) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindah, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

(17) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers for an Old Bird. *Comput Phys Commun* **2014**, *185* (2), 604–613. https://doi.org/10.1016/j.cpc.2013.09.018.

(18) Limongelli, V.; Bonomi, M.; Parrinello, M. Funnel Metadynamics as Accurate Binding Free-Energy Method. *Proc Natl Acad Sci U S A* **2013**, *110* (16), 6358–6363. https://doi.org/10.1073/pnas.1303186110.

(19) Bobrovs, R.; Basens, E. E.; Drunka, L.; Kanepe, I.; Matisone, S.; Velins, K. K.; Andrianov, V.; Leitis, G.; Zelencova-Gopejenko, D.; Rasina, D.; Jirgensons, A.; Jaudzems, K. Exploring Aspartic Protease Inhibitor Binding to Design Selective Antimalarials. *J Chem Inf Model* **2022**, *62* (13), 3263–3273. https://doi.org/10.1021/acs.jcim.2c00422.

(20) Raniolo, S.; Limongelli, V. Ligand Binding Free-Energy Calculations with Funnel Metadynamics. *Nat Protoc* **2020**, *15* (9), 2837–2866. https://doi.org/10.1038/s41596-020-0342-4.

(21) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys Rev Lett* **2008**, *100* (2), 1–4. https://doi.org/10.1103/PhysRevLett.100.020603.

(22) Bonomi, M.; Barducci, A.; Parrinello, M. Reconstructing the Equilibrium Boltzmann Distribution from Well-Tempered Metadynamics. *J Comput Chem* **2009**, *30* (11), 1615–1621. https://doi.org/10.1002/jcc.21305.

(23) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J Mol Graph* **1996**, *14* (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

(24) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* **2007**, *9* (3), 90–95. https://doi.org/10.1109/MCSE.2007.55.

(25) Tribello, G. A.; Ceriotti, M.; Parrinello, M. Using Sketch-Map Coordinates to Analyze and Bias Molecular Dynamics Simulations. *Proc Natl Acad Sci U S A* **2012**, *109* (14), 5196–5201. https://doi.org/10.1073/pnas.1201152109.

(26) Ceriotti, M.; Tribello, G. A.; Parrinello, M. Simplifying the Representation of Complex Free-Energy Landscapes Using Sketch-Map. *Proc Natl Acad Sci U S A* **2011**, *108* (32), 13023–13028. https://doi.org/10.1073/pnas.1108486108.

(27) Bonomi, M.; Bussi, G.; Camilloni, C.; Tribello, G. A.; Banáš, P.; Barducci, A.; Bernetti, M.; Bolhuis, P. G.; Bottaro, S.; Branduardi, D.; Capelli, R.; Carloni, P.; Ceriotti, M.; Cesari, A.; Chen, H.; Chen, W.; Colizzi, F.; De, S.; De La Pierre, M.; Donadio, D.; Drobot, V.; Ensing, B.; Ferguson, A. L.; Filizola, M.; Fraser, J. S.; Fu, H.; Gasparotto, P.; Gervasio, F. L.; Giberti, F.; Gil-Ley, A.; Giorgino, T.; Heller, G. T.; Hocky, G. M.; Iannuzzi, M.; Invernizzi, M.; Jelfs, K. E.; Jussupow, A.; Kirilin, E.; Laio, A.; Limongelli, V.; Lindorff-Larsen, K.; Löhr, T.; Marinelli, F.; Martin-Samos, L.; Masetti, M.; Meyer,

R.; Michaelides, A.; Molteni, C.; Morishita, T.; Nava, M.; Paissoni, C.; Papaleo, E.; Parrinello, M.; Pfaendtner, J.; Piaggi, P.; Piccini, G.; Pietropaolo, A.; Pietrucci, F.; Pipolo, S.; Provasi, D.; Quigley, D.; Raiteri, P.; Raniolo, S.; Rydzewski, J.; Salvalaglio, M.; Sosso, G. C.; Spiwok, V.; Šponer, J.; Swenson, D. W. H.; Tiwary, P.; Valsson, O.; Vendruscolo, M.; Voth, G. A.; White, A. Promoting Transparency and Reproducibility in Enhanced Molecular Simulations. *Nat Methods* **2019**, *16* (8), 670–673. https://doi.org/10.1038/s41592-019-0506-8.

(28) Beyer, B. B.; Goldfarb, N. E.; Dunn, B. M. Expression, Purification, and Characterization of Aspartic Endopeptidases: Plasmodium Plasmepsins and “Short” Recombinant Human Pseudocathepsin. *Curr Protoc Protein Sci* **2003**, *32* (1). https://doi.org/10.1002/0471140864.ps2114s32.

(29) Loymunkong, C.; Sittikul, P.; Songtawee, N.; Wongpanya, R.; Boonyalai, N. Yield Improvement and Enzymatic Dissection of Plasmodium Falciparum Plasmepsin V. *Mol Biochem Parasitol* **2019**, *231* (May), 111188. https://doi.org/10.1016/j.molbiopara.2019.111188.