

Hierarchical computational prioritization of *Uvaria chamae* phytochemicals targeting Ebola virus glycoprotein and SARS-CoV-2 PLpro

Table S1: *Uvaria chamae* P. Beauv compounds selected for *in silico* study [4–7,13–24]

N°	Name	PubChem/Coconut ID	Molecular formula	Part of the plant	References
Acetogenins					
1	Annotemoyin-1	73029717	C 35 H 64 O 5	Root	[13]
2	Bullatencin	44577080	C 37 H 66 O 5	Root	[13,14]
3	Chamuvarinin	11342455	C 37 H 64 O 6	Root, seeds	[6–16]
4	Cis-bullatencin	CNP0239276.2	C 37 H 66 O 5	Root	[13]
5	Cis-uvariamicin-I	14759336	C 37 H 68 W 5	Root	[13]
6	Desacetyluvaricin	127149	C 37 H 66 W 6	Root, seeds	[15,16]
7	Dieporeticanin 1	102064358	C 37 H 66 W 4	Seeds	[15]
8	Dieporeticanin 2	102064359	C 37 H 66 W 4	Seeds	[15]
9	Dieporeticenin	101420981	C 37 H 64 W 4	Seeds	[15]
10	Joolanin	-	C 37 H 64 O 7	Seeds	[15]
11	Neoannonin	76315048	C 35 H 62 O 6	Root	[16]
12	Reticulatacin	10438442	C 37 H 68 O 5	Root	[13]
13	Solamin	11376469	C 35 H 64 O 5	Root	[13]
14	Squamocin	441612	C 37 H 66 O 7	Root, seeds	[15,16]
15	Tripoxyrollin	131753021	C 37 H 64 O 5	Seeds	[15]
16	Uvariamicin-I	44577078	C 37 H 68 O 5	Root	[13]

17	Uvariamicin-II	101392149	C 38 H 70 O 5	Root	[13]
18	Uvariamicin-III	44577079	C 37 H 68 O 5	Root	[13]
Alkaloids					
19	(+)-Armepavine	680292	C19H23NO3	Leaves	[17]
20	Corydine	CNP0157012.2	C 20 H 23 NO 4	Leaves	[17]
21	Nantenine	197001/CNP0165224.1	C 20 H 21 NO 4	Leaves	[17]
22	Nornantenine	3084228	C 19 H 19 NO 4	Leaves	[17]
23	O,O-Dimethylcoclaurine	10829011	C19H23NO3	Leaves	[17]
Flavonoids					
24	Chametín/Chamanetin	21721821	C 22 H 18 O 5	Stem, Root	[14,18,19]
25	Chamuvaritin	100418	C 29 H 24 O 5	Root	[20,21]
26	Dichamanetin	181193	C 29 H 24 O 6	Stem, Root	[14,18,19]
27	Diuvaretin	3085222	C30H28O6	Stem, Root	[14,19,12]
28	Isochamanetin	5318528	C 22 H 18 O 5	Stem, Root	[14,19]
29	Isouvaretin	151670	C 23 H 22 O 5	Stem, Root	[14,18,19]
30	Pinocembrin	68071	C15H12O4	Stem	[19]
31	Pinostrobin	73201	C16H14O4	Root	[19]
32	Uvangoletin	6483649	C16H16O4	Root	[7]
33	Uvaretin	73447	C 23 H 22 O 5	Stem, Root	[18,21]
34	Uvarinol	21721823	C 36 H 30 O 7	Stem, root	[14,22]
Phenols					
35	Caffeic acid	689043	C9H8O4	Root	[23]
36	Ellagic acid	5281855	C14H6O8	Root	[23]
37	Proanthocyanidin	108065	C 31 H 28 W 12	Root	[23]
38	Resorcinol	5054	C6H6O2	Root	[23]
Essential oils, sterols and terpenes					
39	Beta-sitosterol	222284	C 29 H 50 W	Stem	[24]
40	Stigmasterol	5280794	C 29 H 48 W	Stem	[24]
41	1-epi-cubebol	91753170	C15H26O	Leaves, Root	[6,4]
42	1-nitro-2-phénylethane	80208	C8H9NO2	Leaves	[4]
43	3-Carene	26049	C 10 H 16	Leaves	[4]
44	3Z-Hexenylbenzoate	101687121	C13H16O2	Leaves	[4]
45	4-nitrophenyl laurate	74778	C 18 H 27 NW 4	Stem	[24]
46	6,9-Guaiadiene	527113	C 15 H 24	Leaves	[4]
47	Alloaromadendrene	10899740	C 15 H 24	Leaves	[4]
48	Alpha-cadinol	10398656	C15H26O	Leaves	[4]
49	Alpha-copaene	19725	C 15 H 24	Leaves	[6,4]

50	Alpha-cubebene	442359	C 15 H 24	Leaves	[4]
51	Alpha-farnesene	5281516	C 15 H 24	Leaves	[4]
52	Alpha-Muurolene	12306047	C 15 H 24	Leaves	[4]
53	Alpha-phellandrene	7460	C 10 H 16	Leaves	[4]
54	Alpha-santalene	94 164		Stem	[24]
55	Alpha-Santalone	162952798	C15H22O	Leaves	[4]
56	Alpha-terpinene	7462	C 10 H 16	Leaves	[4]
57	Alpha-terpineol	17100	C10H18O	Leaves	[4]
58	Baldrinal	159846	C12H10O4	Stem	[24]
59	benzaldehyde	240	C7H6O	Leaves	[4]
60	Benzeneacetonitrile	8794	C 8 H 7 N	Leaves	[4]
61	Benzyle benzoate	2345	C14H12O2	Leaves	[4]
62	beta-bourbonene	62566	C 15 H 24	Leaves	[4]
63	beta-copaene	57339298	C 15 H 24	Leaves	[4]
54	beta-cubebene	93081	C 15 H 24	Leaves	[4]
65	Beta-elemene	6918391	C 15 H 24	Leaves	[4]
66	beta-Maaliene	101596917	C 15 H 24	Stem	[24]
67	beta-ocimene	18756	C 10 H 16	Leaves	[4]
68	beta-selinene	442393	C 15 H 24	Leaves	[4]
69	Bicyclogermacrene	13894537	C 15 H 24	Leaves	[4]
70	Borneol	64685	C10H18O	Leaves	[4]
71	Camphene	6616	C 10 H 16	Leaves	[4]
72	Caryophyllene	5281515	C 15 H 24	Leaves	[4]
73	Caryophyllene oxyde	1742210	C15H24O	Leaves	[4]
74	Copaene	12303902	C 15 H 24	Leaves	[4]
75	Cubebol	11276107	C15H26O	Leaves	[4]
76	delta-cadinene	441005	C 15 H 24	Leaves	[4]
77	Elemol	92138	C15H26O	Leaves	[4]
78	E-phytol	5280435	C20H40O	Leaves	[4]
79	Epi-cubebol	91 753 433	C15H24O	Leaves	[4]
80	Epicubenol	12,046,149	C15H26O	Leaves	[4]
81	Eucalyptol	2758	C10H18O	Leaves, Root	[6,4]
82	gamma-Muurolene	12313020	C 15 H 24	Leaves	[4]
83	gamma-terpinene	7461	C 10 H 16	Leaves	[4]
84	Germacrene A	9548705	C 15 H 24	Leaves	[4]
85	Germacrene B	5281519	C 15 H 24	Leaves	[4]
86	Germacrene D	5317570	C 15 H 24	Leaves, Root	[4]

87	Gleenol	6429080	C15H26O	Leaves	[4]
88	Humulene	5281520	C 15 H 24	Leaves, Root	[4]
89	Humulene epoxide II	10704181	C15H24O	Leaves	[4]
90	Isopropyle linoleate	5352860	C 21 H 38 O 2	Stem	[24]
91	Limonene	22311	C 10 H 16	Leaves	[4]
92	Linalol	6549	C10H18O	Leaves, Root	[6,4]
93	Lupeol	259846	C30H50O	Stem	[24]
94	Myrcene	31253	C 10 H 16	Leaves	[4]
95	Nerolidol	5284507	C15H26O	Leaves	[4]
96	P-cymene	7463	C10H14	Leaves	[4]
97	Pentadecanal	17697	C15H3O	Leaves	[4]
98	Spathulenol	92231	C15H24O	Leaves	[4]
99	Squalene	638072	C30H50	Stem	[24]
100	tau-Cadinol	160799	C15H26O	Leaves	[4]
101	tau-Muurolol	6432221	C15H26O	Leaves	[4]
102	Terpinen-4-ol	11230	C10H18O	Leaves	[4]
103	trans-Cadina-1(6),4-diene	10798255	C15H24	Leaves	[4]
104	trans-Cadina-1,4-diene	6430869	C15H24	Leaves	[4]
105	trans-Muurola-4(14),5-diene	91747125	C15H24	Leaves	[4]
Fatty acids					
106	Palmitic acid	985	C16H32O2	Leaves	[4]

Molecular dynamics simulations and quantum-chemical characterization

Five solvated systems were simulated for 100 ns each. For both viral targets the prioritized phytochemical was simulated together with a reference inhibitor bound to the identical receptor, so that every quantitative comparison is internally controlled. Because MM-GBSA binding free energies are not transferable from one receptor to another, no comparison is drawn between the two targets anywhere in this material.

Table S2: Systems simulated by molecular dynamics.

N°	Ligand	Role	Receptor	Residues
1	Proanthocyanidin	candidate	EBOV GP1–GP2 (5F1B, chains A+B)	248
2	Toremifene	reference inhibitor	EBOV GP1–GP2 (5F1B, chains A+B)	248
3	Proanthocyanidin	receptor-construct control	EBOV GP1 only (5F1B, chain A)	158
4	Uvangoletin	candidate	SARS-CoV-2 PLpro (7CMD, chain A)	316
5	GRL0617	reference inhibitor	SARS-CoV-2 PLpro (7CMD, chain A)	316

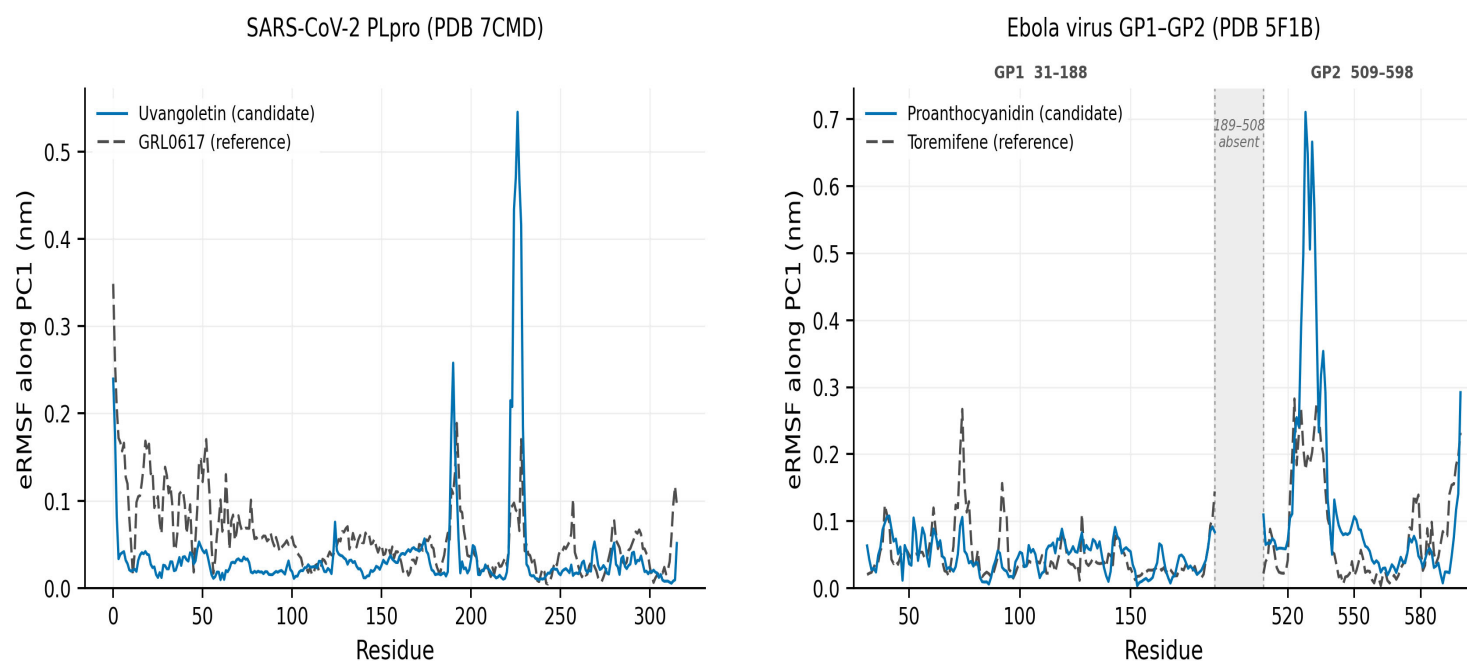


Figure S1: Essential-dynamics fluctuation (eRMSF) projected on the first principal component, for each viral target. Candidates are shown as solid lines and reference inhibitors as dashed lines. Unlike the conventional RMSF, this quantity isolates the amplitude of the slowest collective motion of the protein instead of mixing it with fast local vibrations [41].

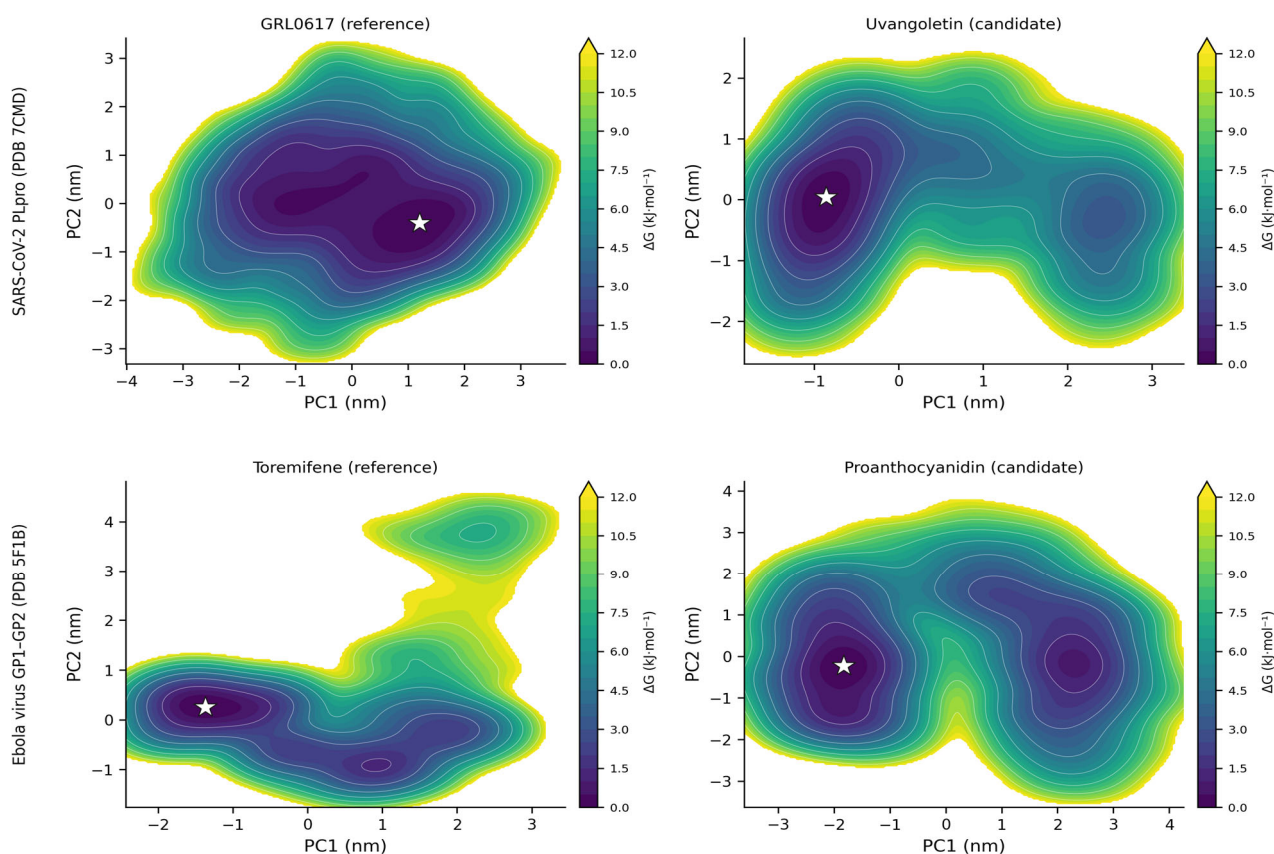


Figure S2: Gibbs free-energy landscapes of the four complexes over the first two principal components. Surfaces were rebuilt from the 2001 per-frame projections with a Gaussian kernel density estimate, $\Delta G = -k_B T \ln(P/P_{\max})$ at 298.15 K, and capped at $12 \text{ kJ}\cdot\text{mol}^{-1}$; the star marks the global minimum. The kernel estimate is used because a 2001-frame trajectory does not support a fine histogram grid. Each candidate samples two distinct basins separated by a low barrier, whereas the reference inhibitor of the PLpro occupies a single broad basin.

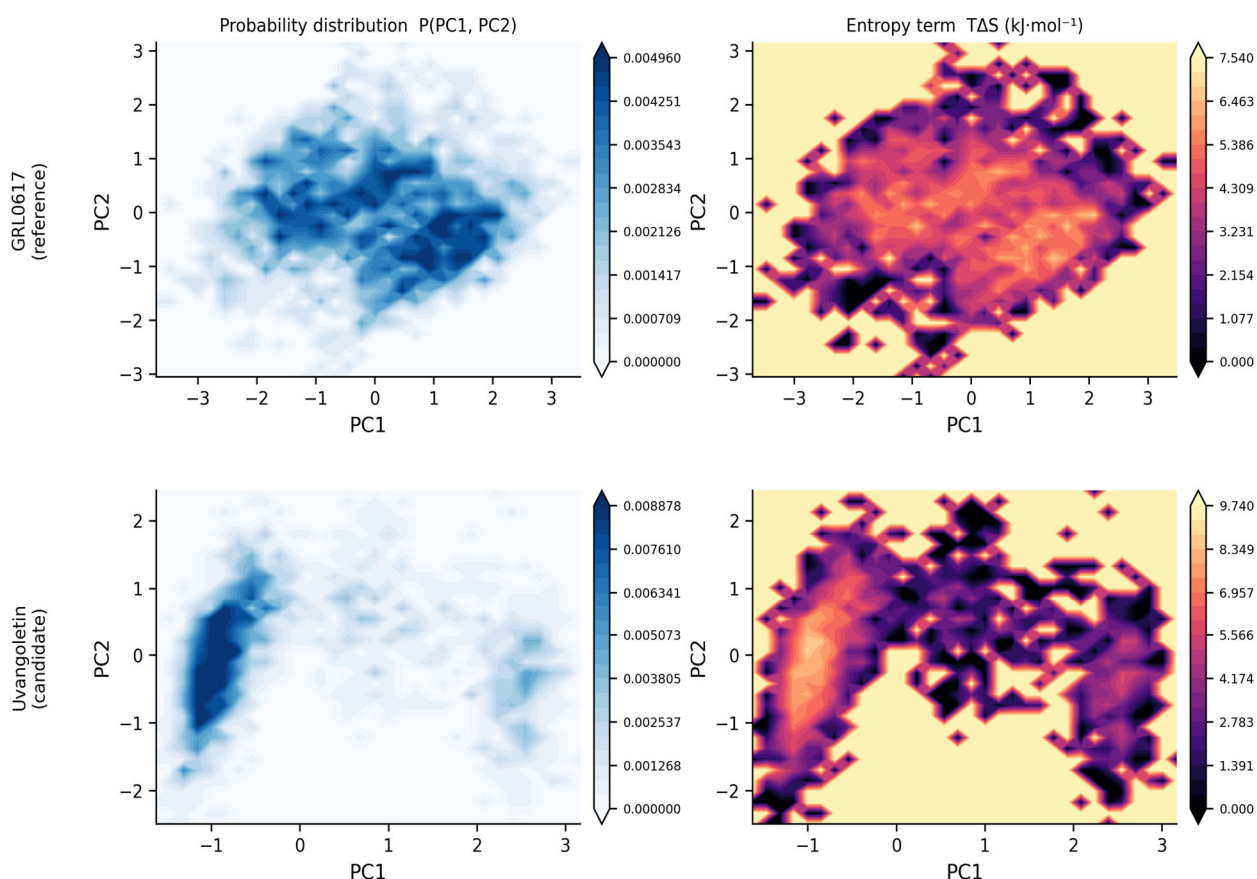


Figure S3: Conformational sampling of the SARS-CoV-2 PLpro complexes over the first two principal components: probability density and entropic term, for the reference inhibitor GRL0617 (top) and the candidate uvangoletin (bottom). No enthalpy map is reported. The original gmx sham run was performed without an energy file, so the map it wrote contains a binary occupancy mask rather than an enthalpy; and when the calculation is repeated using the per-frame potential energies recovered from the GROMACS log, the potential energy of the solvated box is found to be uncorrelated with the two collective coordinates ($R^2 = 0.0004$; the solvent fluctuation, $\sigma = 1293$ kJ·mol⁻¹, is fifty times larger than the part explained by PC1 and PC2, $\sigma = 26$ kJ·mol⁻¹). An enthalpy–entropy decomposition of this landscape would therefore report solvent noise, and only the free energy is reported.

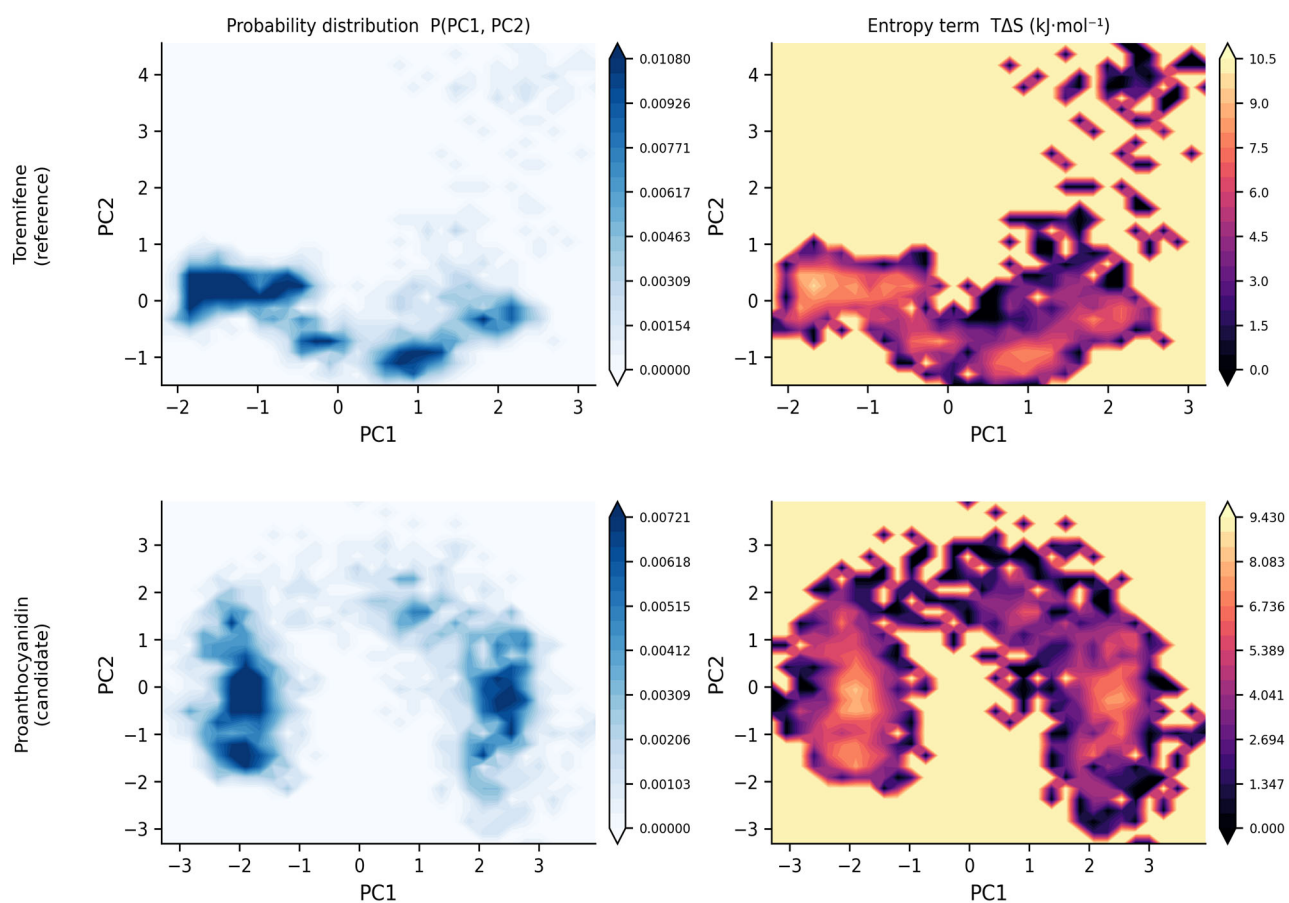


Figure S4: Conformational sampling of the Ebola virus GP1-GP2 complexes, for the reference inhibitor toremifene (top) and the candidate proanthocyanidin (bottom); same convention as Figure S3.

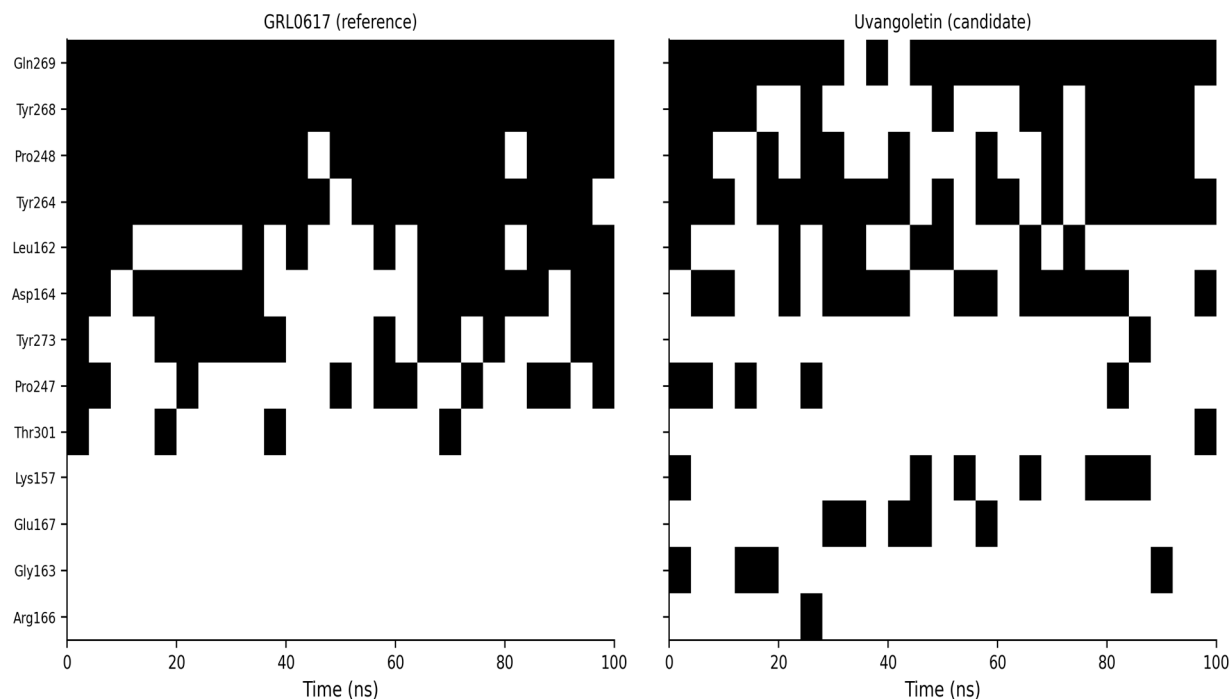


Figure S5: Persistence of the binding-site contacts of the SARS-CoV-2 PLpro complexes along the trajectory. A black cell indicates that at least one non-covalent interaction was detected by PLIP [18] in that frame (25 frames, one every 4 ns). Both panels share the same residue ordering, so the candidate and the reference binding modes can be read against each other row by row.

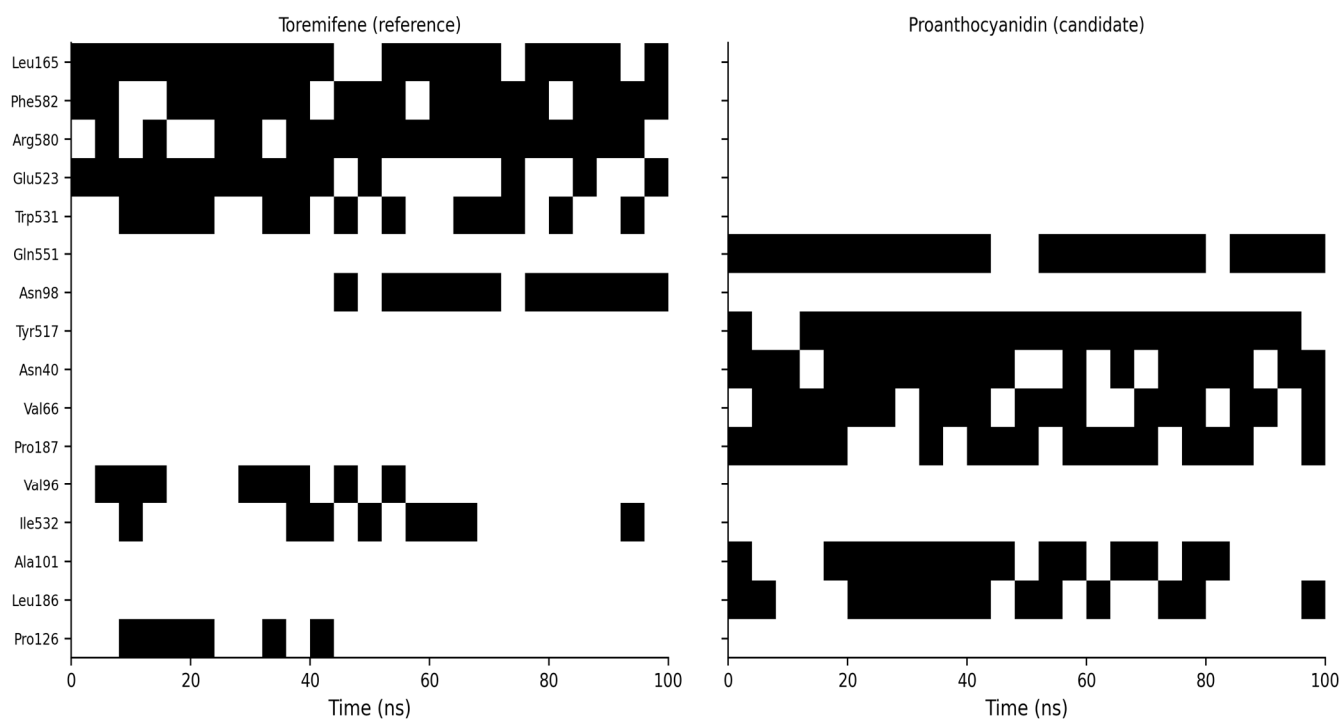


Figure S6: Persistence of the binding-site contacts of the Ebola virus GP1-GP2 complexes; same convention as Figure S5.

Table S3: Summary of the molecular dynamics analyses. Means over the 100 ns trajectories ($\pm$ standard deviation); MM-GBSA binding free energy over the last 20 ns ($\pm$ standard error of the mean) [37]. Configurational entropy was not computed, so the

reported ΔG values are effective, enthalpy-dominated binding energies, valid for ranking complexes that share an identical receptor and not as absolute affinities.

System	Role	Backbone RMSD (nm)	Rg (nm)	SASA (nm ²)	ΔG_{bind} (kcal·mol ⁻¹)
Proanthocyanidin – EBOV GP1–GP2	candidate	0.30 ± 0.04	1.90 ± 0.02	133.26 ± 2.99	-37.16 ± 0.27
Toremifene – EBOV GP1–GP2 (reference)	control	0.34 ± 0.04	1.87 ± 0.02	132.29 ± 3.66	-49.71 ± 0.27
Proanthocyanidin – EBOV GP1 (chain A)	candidate	0.37 ± 0.04	1.70 ± 0.02	96.55 ± 2.26	-29.94 ± 0.39
Uvangoletin – SARS-CoV-2 PLpro	candidate	0.25 ± 0.04	2.36 ± 0.02	163.17 ± 2.65	-26.78 ± 0.18
GRL0617 – SARS-CoV-2 PLpro (reference)	control	0.19 ± 0.03	2.39 ± 0.02	168.72 ± 2.56	-33.72 ± 0.18

• Ligand RMSD of the uvangoletin–PLpro system

The ligand RMSD written by the automated analysis compares each frame with a reference structure without re-imaging the ligand. When the ligand sits in a neighbouring periodic image of the simulation box, this returns a large and essentially constant value even though the ligand never leaves the pocket, which is what occurred for the uvangoletin–PLpro trajectory. The quantity was therefore recomputed for all four complexes directly from the 101 trajectory snapshots of each: the protein chain and the ligand are first made whole again across the periodic boundaries, each frame is then superimposed on the first one by a Kabsch fit on the backbone N, CA, C and O atoms, and the heavy-atom RMSD of the ligand is measured against the reference pose.

The procedure was validated by applying it to all four complexes. In the three where the automated value is sound it reproduces it: proanthocyanidin 0.175 ± 0.042 against 0.18 ± 0.04 nm, toremifene 0.467 ± 0.156 against 0.47 ± 0.15 nm, GRL0617 0.176 ± 0.066 against 0.18 ± 0.06 nm. The backbone RMSD, which the periodic-image problem does not affect, is reproduced in all four systems to within 0.01 nm. For the uvangoletin complex the recomputation returns 0.278 ± 0.111 nm. The candidate therefore stays bound throughout the trajectory, somewhat more mobile than the reference inhibitor, with a group of excursions between 45 and 65 ns over which the mean rises to 0.40 nm and reaches 0.58 nm at 64 ns, against 0.25 nm over the rest of the trajectory. That excursion is not resolved by the backbone principal components — its correlation with PC1 and PC2 is only -0.31 and $+0.36$ — so it is a local rearrangement of the ligand inside the pocket, distinct from the secondary basin of Figure S2, which contains only frames of the first 23 ns. Two

independent observations agree: the hydrogen bond with Gln269 is present in 92% of the analysed frames, and the MM-GBSA binding free energy is stable over the whole sampling window.

Table S4: Ligand RMSD of the four complexes: value written by the automated analysis and value recomputed from the trajectory snapshots with the periodic images corrected. The backbone RMSD, which the periodic-image problem does not affect, is given as a control of the recomputation.

Complex	Ligand RMSD, automated (nm)	Ligand RMSD, recomputed (nm)	Backbone RMSD, automated (nm)	Backbone RMSD, recomputed (nm)
Proanthocyanidin – EBOV GP1–GP2 (candidate)	0.18 ± 0.04	0.175 ± 0.042	0.30	0.303
Toremifene – EBOV GP1–GP2 (reference)	0.47 ± 0.15	0.467 ± 0.156	0.34	0.336
Uvangoletin – PLpro (candi- date)	11.45 ± 0.09	0.278 ± 0.111	0.25	0.245
GRL0617 – PLpro (reference)	0.18 ± 0.06	0.176 ± 0.066	0.19	0.195

Quantum-chemical characterization of the ligands

Ligand geometries were taken from the last frame of each trajectory and optimized in implicit aqueous phase at the B3LYP-D3BJ/6-31G(d,p) level with the CPCM solvation model, using ORCA 6.1.1 [44]. An analytical harmonic frequency calculation was run on every optimized geometry; none of the four structures shows an imaginary frequency, so each is a true minimum on the potential energy surface and the thermal corrections at 298.15 K are valid. Because the 6-31G(d,p) basis set has no diffuse functions and therefore describes the lowest unoccupied orbital poorly, a single-point calculation was then run on every optimized geometry at the B3LYP-D3BJ/6-311+G(d,p) level with the same solvation model, and the frontier orbital energies were taken from it. Reactivity descriptors were derived from those energies through Koopmans’ theorem. Toremifene was modelled in its amine-protonated form (net charge +1), consistent with the molecular dynamics setup in which the simulation box was neutralized with chloride ions only; its dipole moment is not reported, because for a species of non-zero net charge the dipole depends on the choice of coordinate origin and is not a well-defined molecular property. Comparisons within the Ebola virus pair are accordingly restricted to the HOMO–LUMO gap and the chemical hardness, both of which are differences of orbital energies.

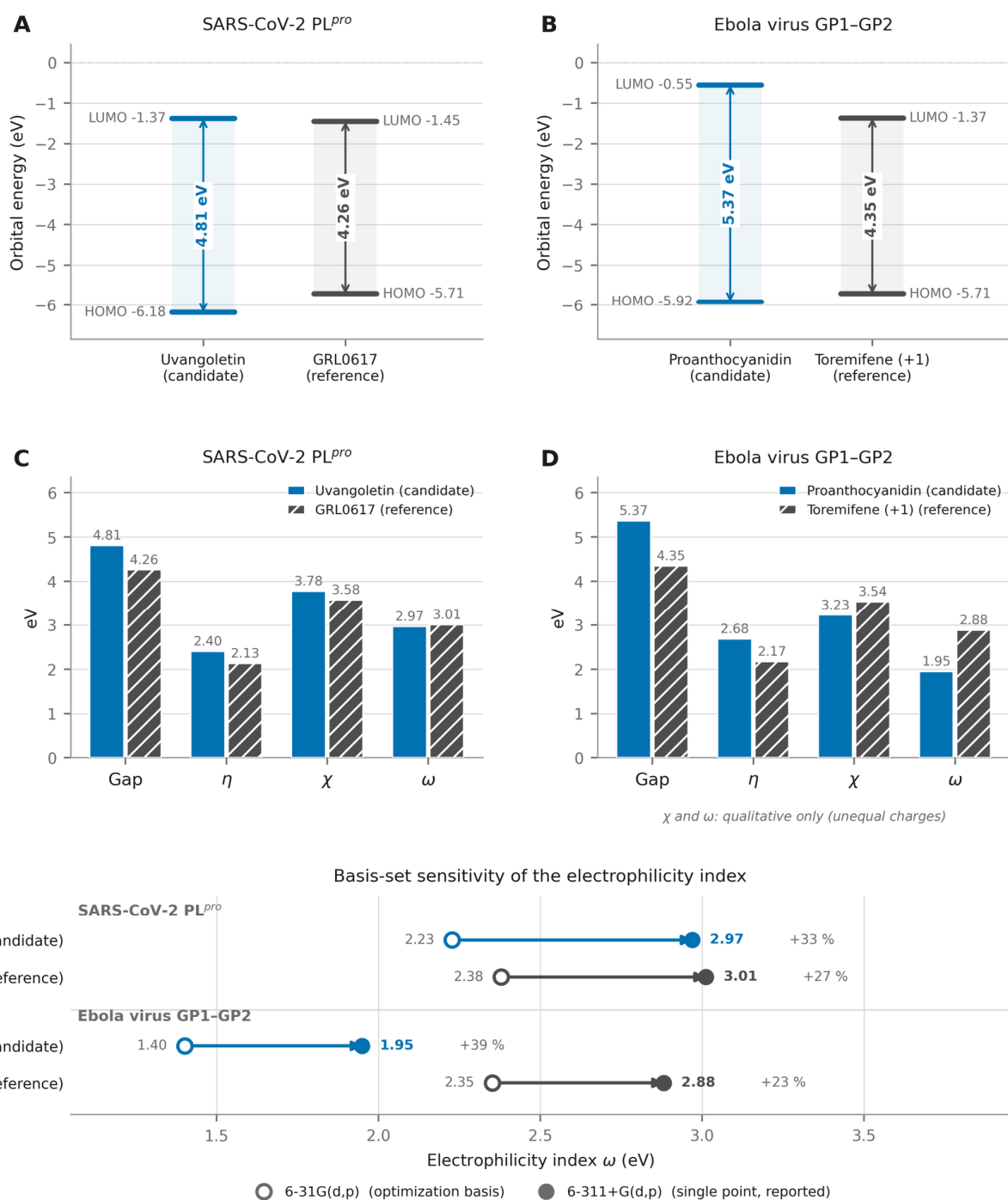


Figure S7: Conceptual-DFT reactivity profile of the four ligands. Each candidate is shown against the reference inhibitor of its own target; no comparison is drawn between targets. (A, B) Frontier orbital energies and HOMO–LUMO gap at the B3LYP-D3BJ/6-311+G(d,p) level in implicit water. (C, D) Reactivity descriptors derived from those energies through Koopmans’ theorem: HOMO–LUMO gap, chemical hardness η , electronegativity χ and electrophilicity index ω . For the Ebola virus pair the reference inhibitor is a cation, so χ and ω are read qualitatively and the dipole moment is not reported. (E) Methodological panel: effect of adding diffuse functions to the basis set on the electrophilicity index, shown for each ligand separately; open circles are the 6-31G(d,p) optimization basis, filled circles the 6-311+G(d,p) single point from which the reported values are taken. No ligand-to-ligand comparison is intended in that panel. Candidates are drawn in blue and reference inhibitors in dark grey with a hatched fill, Numerical values are given in Table S5.

Table S5: Conceptual-DFT reactivity descriptors of the four ligands, at both basis-set levels. Frontier orbital energies and the descriptors derived from them are reported from the 6-311+G(d,p) single point; the thermochemistry comes from the 6-31G(d,p) optimization and frequency calculation, which is the level at which the geometries and the Hessian were obtained. $IP = -E(\text{HOMO})$, $EA = -E(\text{LUMO})$, $\chi = (IP + EA)/2$, $\eta = (IP - EA)/2$, $S = 1/(2\eta)$, $\omega = \chi^2/(2\eta)$.

Descriptor	Uvangoletin (candidate)	GRL0617 (reference)	Proanthocyanidin (candidate)	Toremifene (reference)
Target	SARS-CoV-2 PLpro	SARS-CoV-2 PLpro	Ebola virus GP1-GP2	Ebola virus GP1-GP2
Net charge	0	0	0	+1
B3LYP-D3BJ/6-311+G(d,p) single point — reported values				
E(HOMO) (eV)	-6.1842	-5.7130	-5.9175	-5.7121
E(LUMO) (eV)	-1.3743	-1.4515	-0.5513	-1.3654
HOMO–LUMO gap (eV)	4.810	4.262	5.366	4.347
IP (eV)	6.184	5.713	5.918	5.712
EA (eV)	1.374	1.452	0.551	1.365
χ (eV)	3.779	3.582	3.234	3.539
η (eV)	2.405	2.131	2.683	2.173
S (eV ⁻¹)	0.2079	0.2347	0.1863	0.2301
ω (eV)	2.969	3.011	1.950	2.881
Dipole moment (D)	7.64	5.65	4.34	not reported (ion)
E (Eh)	-920.047942	-958.719744	-2100.491389	-1597.915735
B3LYP-D3BJ/6-31G(d,p) optimization and frequencies				
E(HOMO) (eV)	-5.7858	-5.3654	-5.5314	-5.4364
E(LUMO) (eV)	-0.8476	-1.0462	-0.0248	-1.0136
HOMO–LUMO gap (eV)	4.938	4.319	5.507	4.423
ω (eV)	2.228	2.379	1.402	2.352
E (Eh)	-919.792530	-958.476160	-2099.890040	-1597.612955
ZPE (kcal·mol ⁻¹)	181.77	224.21	346.88	310.53
H (Eh)	-919.483807	-958.097636	-2099.298287	-1597.089733

G (Eh)	-919.546334	-958.165181	-2099.399782	-1597.172369
Imaginary frequencies	0	0	0	0