Article

Computational molecular docking analysis of Doxifluridine and its metabolites to identify potential hits for PDHK1

Neeta Azad,[1,2] Rita Kakkar[1*]

[1]*Computational Chemistry Group, Department of Chemistry, University of Delhi, Delhi-110007, India.*
[2]*Department of Chemistry, Atma Ram Sanatan Dharam College, University of Delhi, Dhaula Kuan, New Delhi-110021, India*

Supporting Information

rkakkar@chemistry.du.ac.in

Computational Details

Protein Preparation

First, the protein structure was refined using the protein preparation workflow in the Schrödinger software. The bond order of the residues was adjusted. The missing loops of the protein were filled using the "Fill loops" option of the "Protein Preparation wizard". Pre-co-crystallized water molecules beyond 5 Å of the active site were removed, and hydrogen atoms were added to the structures. Protein structures were first optimized using molecular mechanics calculations with the OPLS-2005 force field[1]. Minimization was performed until the average root mean square deviation (RMSD) of the nonhydrogen atoms reached 0.3 Å. Additional states for the inhibitors within the proteins were generated (using "Generate Het States"), and the most appropriate states for all ligands were selected. Refinement of the side chains was performed using the "Prime side chain refinement" option.

QM/MM Calculations

Mixed QM/MM calculations provide better results than pure MM calculations[2–6]. However, treating the entire protein using quantum mechanics is computationally expensive and time-consuming. Hence, we applied QM to the active site and MM to the remainder of the protein. QSite was used for QM/MM calculations. The ligand and active site residues were treated using density functional theory (DFT), and the rest of the protein was treated using the OPLS-2005 force field. The B3LYP/LACVP* basis set was used to treat the active site residues. LACVP* employs the 6-31G* basis set for nontransition elements in the active site. For QM treatment, the ligand AZD7545, 12 water molecules trapped inside the protein cavity, and residues Leu194, Gln197, Hie198, Leu201, Phe202, Leu57, Gln61, Phe62, Asp64, Phe65, Thr74, Ser75, Phe78, and Leu79 were selected by the side chain selection method. Non-bonded cut-offs were employed, and continuum solvation with the Poisson Boltzmann Solver model was used, along with a constant dielectric electrostatic treatment with a dielectric constant of 1.0. Force-field checks were skipped, and without applying any MM or QM constraints, a single-point calculation was performed on the protein.

Glide Docking

Glide calculations were performed on the QM/MM-treated protein to estimate the ligand-binding energies[7–9]. To soften the potential of the nonpolar parts of the receptor, the van der Waals radii of the receptor atoms were scaled by 1.00 with a partial atomic charge of 0.25. A grid box of size $56 \times 56 \times 56$ Å^3 with coordinates X = -5.52, Y = 41.09, and Z = 8.53 Å was generated at the centroid of the ligand. Ligands were docked into the active site using Glide-XP (Glide "extra precision"). Glide generates internal conformations and passes them through a series of filters. The first filter places the ligand center at various grid positions of 1 Å and rotates it around three Euler angles. At this stage, crude score values and geometrical filters weeded out unlikely binding modes. The next filter stage involves a grid-based force

field evaluation and refinement of docking solutions, including torsional and rigid body movements of the ligands. The OPLS-2005 force field was used for this purpose. A small number of surviving docking solutions can then be subjected to the Monte Carlo procedure to minimize the energy score. The final energy evaluation was performed using a Glide Score (GScore), and a single best pose was generated as the output for a specific ligand.

$$Gscore = a \times vdW + b \times Coul + Lipo + Hbond + Metal + BuryP + RotB + Site$$,

where vdW = van der Waals energy, $Coul$ = Coulomb energy, $Lipo$ = lipophilic contact term, $Hbond$ = hydrogen-bonding term, $Metal$ = metal-binding term, $BuryP$ = penalty for buried polar groups, $RotB$ = penalty for freezing rotatable bonds, and $Site$ = polar interaction at the active site. The coefficients of vdW and $Coul$ were $a = 0.065$ and $b = 0.130$, respectively.

The best docked structure for each ligand was chosen using the model energy score (E_{model}), which combines the Glide score, non-bonded interaction energy, and excess internal energy of the generated ligand conformation.

The Gibbs energy of binding between the receptor and ligand was predicted using the Prime MM/GBSA application in Schrödinger. MM/GBSA combines the OPLS molecular mechanics energy (E_{MM}), the surface-generalized Born solvation model for polar solvation (G_{SB}), and a nonpolar solvation term (G_{NP}). The G_{NP} term comprises the nonpolar solvent-accessible surface area and van der Waals interactions. The total Gibbs energy of binding was calculated as follows:

$\Delta G_{binding} = G_{complex} - (G_{protein} + G_{ligand})$ (1)
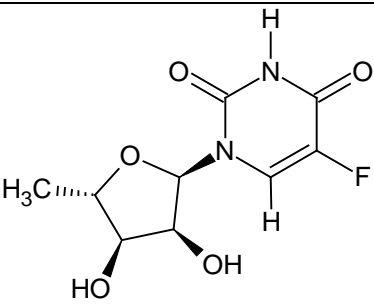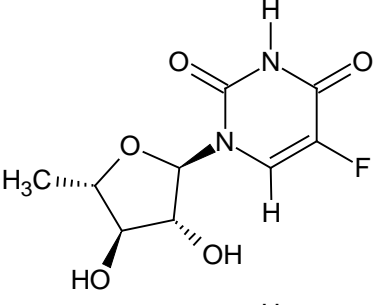
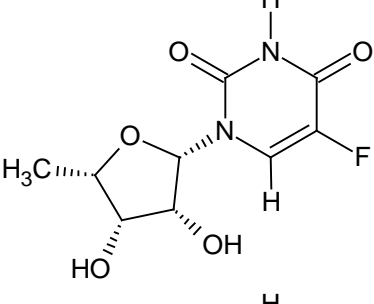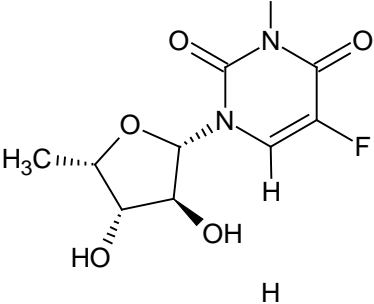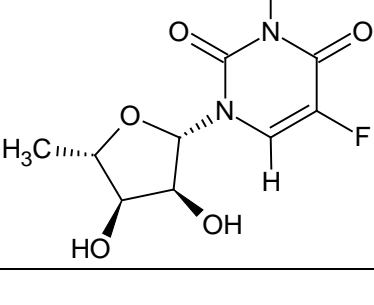$G = E_{MM} + G_{SB} + G_{NP}$

ADME Properties

The QikProp program[10] was used to determine the absorption, distribution, metabolism, and excretion (ADME) properties of the analogs. It predicts both physically significant descriptors and pharmaceutically relevant properties of drugs.
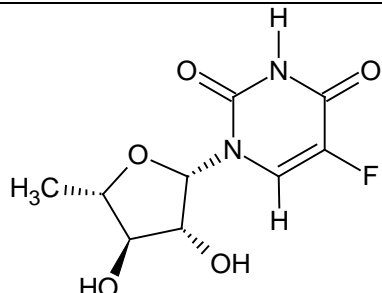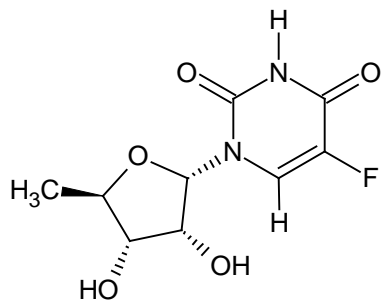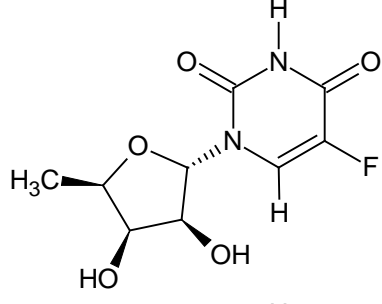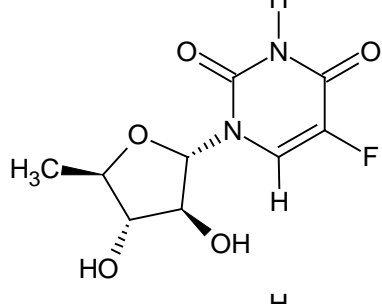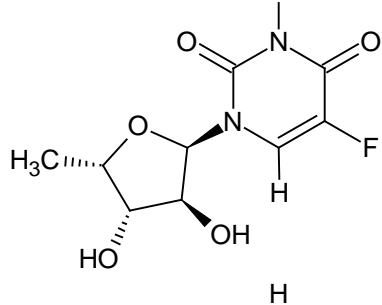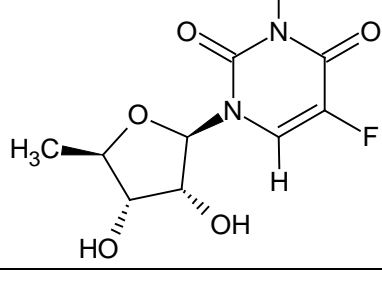
The program was processed in normal mode, and forty-four properties were predicted for different states of doxifluridine and its metabolites. These properties consisted of principal descriptors and physicochemical properties with a detailed analysis of the predicted octanol/water partition coefficient (logPo/w), octanol/gas partition coefficient (logPoct), water/gas partition coefficient (logPw), polarizability in cubic Ångstrom (Polrz), % human absorption in the intestines (QP%), brain/blood partition coefficient (logBB), IC_{50} value for the blockage of HERG K^+ channels (logHERG), skin permeability (logK$_P$), prediction of binding to human serum albumin (logKhsa), apparent Caco-2 cell permeability in mm/s (Caco), and apparent MDCK cell permeability in mm/s (MDCK). Caco-2 cells are a model of the gut-blood barrier. MDCK cells are considered good mimics of the blood-brain barrier. It also evaluates the acceptability of the analogs based on Lipinski's rule of five[11,12], which is essential for rational drug design. Poor absorption or permeation is more likely when a ligand violates Lipinski's rule of five, that is, it has more than five hydrogen donors, the molecular weight is over 500, the log$P$ is over 5, and the sum of N and O is over 10.

Molecular Dynamics Calculations

The OPLS3 force field[13] was used for the simulation. The NPT ensemble was used at a temperature of 300 K and a pressure of 1 bar. Long-range electrostatic interactions were computed using the particle mesh Ewald approach[14], the long-range electrostatic interactions were computed. 9.0 Å was set as the cutoff radius for the Coulomb interactions. The Martyna–Tuckerman–Klein chain coupling technique[15] and a relaxation duration of 2.0 ps were combined with the barostat approach for pressure management. The Nosé–Hoover chain coupling technique[16] thermostat with a relaxation duration of 1.0 ps, was used to control the temperature. Using an r-RESPA integrator, which updated the long-range forces every three steps and the short-range forces at every step, the nonbonded forces were computed. A 2 fs time step was employed. For an additional investigation, the trajectories were stored at intervals of 2 fs. The simulation interaction diagram tool of the Desmond MD package was used to investigate the interactions between ligands and proteins. The stability of the MD simulations was tracked by analyzing the Root Mean Square Deviation (RMSD) of the ligand and protein atom locations over time. The systems were relaxed in accordance with Desmond's default methodology before MD computations. Prime MM-GBSA computations were used to determine the binding energies of the ligand-receptor complexes in all 1000 structures sampled during the simulation.
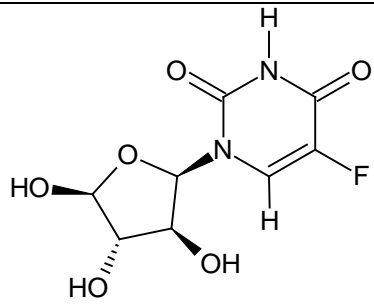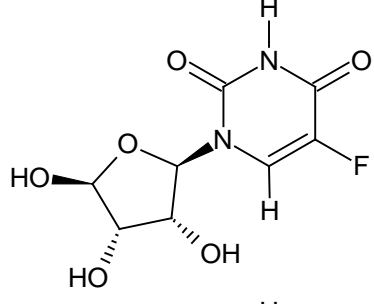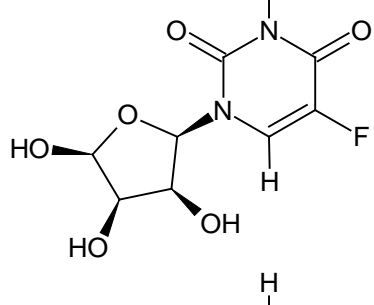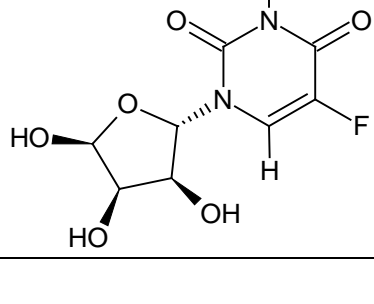
Table S1: Stereoisomers of doxifluridine, 5-fluorouridine and 5-fluorouracil
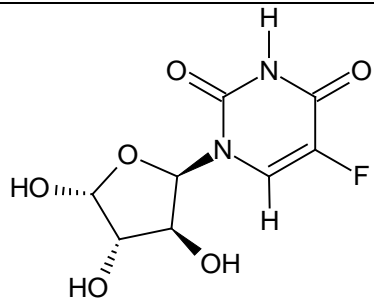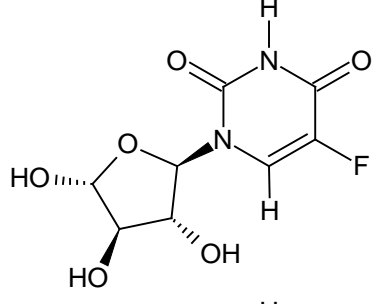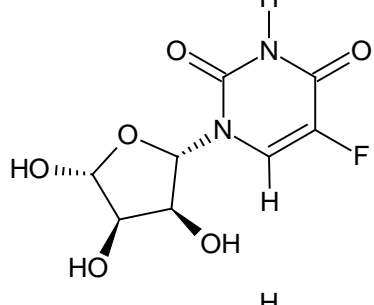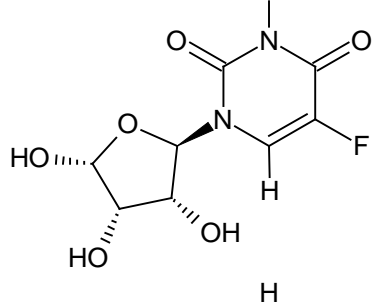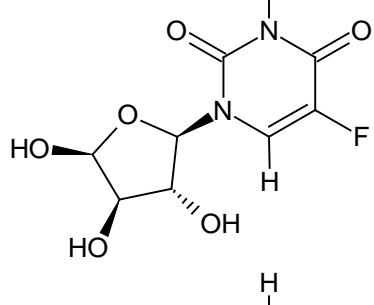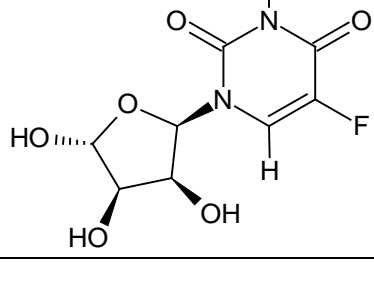
S. No.	Pose name	Stricture	Chirality			
1	D1		11S	12S	13S	14R
2	D2		11S	12R	13S	14R
3	D3		11R	12R	13R	14R
4	D4		11R	12S	13R	14R
5	D5		11R	12S	13S	14R
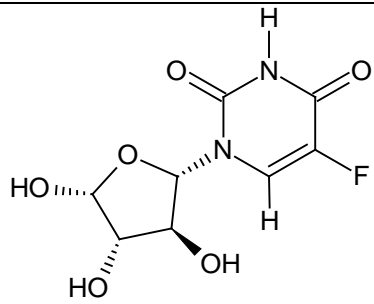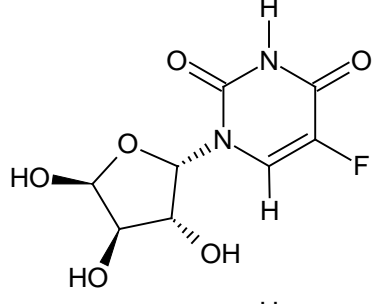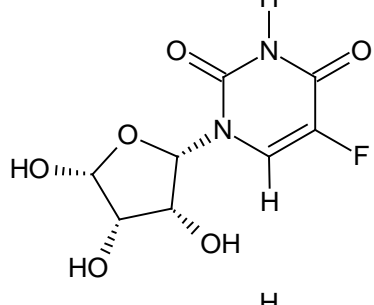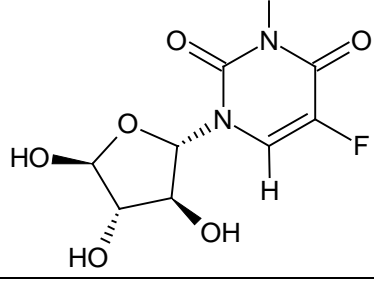
6	D6		11R	12R	13S	14R

7	D7		11R	12R	13R	14S

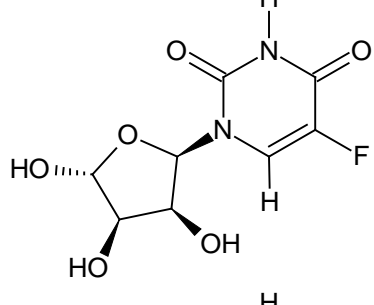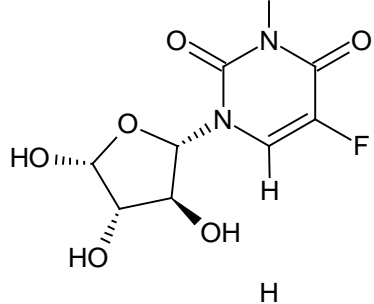| 8	D8		11R	12S	13S	14S

9	D9		11R	12S	13R	14S

10	D10		11S	12S	13R	14R

11	D11		11S	12R	13R	14S

12	D12		11S	12R	13S	14S
13	D13		11S	12S	13S	14S
14	D14		11S	12S	13R	14S
15	D15		11S	12R	13R	14R
16	D16		11R	12R	13S	14S

5-Fluorouridine stereoisomers

17	U1		11S	12S	13R	14S
18	U2		11S	12R	13R	14S
19	U3		11S	12S	13S	14R
20	U4		11R	12S	13R	14R
21	U5		11S	12S	13S	14S
22	U6		11R	12S	13S	14S

23	U7		11S	12S	13R	14R
24	U8		11S	12R	13S	14R
25	U9		11R	12S	13S	14R
26	U10		11S	12R	13R	14R
27	U11		11S	12R	13S	14S
28	U12		11R	12R	13R	14S

29	U13		11R	12R	13S	14R
30	U14		11R	12R	13S	14S
31	U15		11R	12R	13R	14R
32	U16		11R	12S	13R	14S

Table S2: ADME properties of doxifluridine states

Title	QPpolrz	QPlogPC16	QPlogPoct	QPlogPw	QPlogPo/w	QPlogS
D1	20.03	6.82	16.17	14.67	-0.84	-1.70
D2	20.67	6.94	16.77	14.72	-0.82	-1.81
D3	21.08	6.97	17.00	14.84	-0.84	-1.93
D4	21.37	6.97	16.70	14.82	-0.78	-1.98
D5	20.82	6.87	16.62	14.69	-0.73	-1.84
D6	20.99	7.04	17.28	14.65	-0.64	-1.78
D7	19.92	6.72	16.03	14.67	-0.87	-1.73
D8	21.13	6.88	16.70	14.75	-0.77	-1.95
D9	20.35	6.82	16.02	14.70	-0.85	-1.79
D10	21.17	7.01	17.11	14.86	-0.88	-1.95
D11	19.88	6.72	16.75	14.59	-0.94	-1.72
D12	20.46	6.82	17.72	14.66	-0.87	-1.81
D13	20.58	6.98	16.45	14.72	-0.83	-1.77
D14	20.55	6.96	16.35	14.82	-1.04	-1.85
D15	21.11	6.89	17.68	14.78	-0.89	-1.98
D16	21.31	7.05	16.64	14.83	-0.86	-1.95

Table S3: Descriptors for doxifluridine states

Title	HERG	Caco	logBB	MDCK	logKp	logKhsa	QP%	PSA	glob
D1	-2.86	154.87	-0.92	115.02	-4.71	-0.75	61.25	112.34	0.93
D2	-2.99	129.65	-1.01	95.94	-4.87	-0.74	59.93	117.90	0.92
D3	-3.33	116.18	-1.12	84.27	-4.94	-0.76	58.97	121.44	0.90
D4	-3.39	128.61	-1.10	94.98	-4.85	-0.75	60.14	119.45	0.90
D5	-3.12	171.99	-0.93	129.88	-4.61	-0.75	62.70	115.13	0.91
D6	-2.85	190.46	-0.84	145.16	-4.52	-0.71	63.98	121.46	0.93
D7	-3.01	154.24	-0.95	114.51	-4.71	-0.77	61.05	111.87	0.92
D8	-3.34	143.72	-1.05	107.06	-4.77	-0.76	61.06	117.83	0.90
D9	-3.03	137.86	-1.00	102.82	-4.82	-0.76	60.27	117.04	0.92
D10	-3.33	98.29	-1.19	71.03	-5.10	-0.76	57.43	123.80	0.90
D11	-2.82	121.51	-1.02	88.21	-5.00	-0.75	58.77	116.74	0.92
D12	-3.02	125.43	-1.04	91.49	-4.94	-0.75	59.40	116.48	0.91
D13	-2.90	127.32	-1.00	93.87	-4.89	-0.73	59.75	122.58	0.93
D14	-3.07	69.83	-1.28	48.61	-5.45	-0.75	53.84	127.26	0.91
D15	-3.35	99.76	-1.20	72.32	-5.13	-0.76	57.49	120.94	0.89
D16	-3.25	97.59	-1.18	70.60	-5.12	-0.74	57.51	124.30	0.90

Table S4: ADME properties of 5-fluorouracil

Polrz	logPC16	logPoct	logPw	logPo/w	LogS	Rule of five	dipole	CNS
10.65	**3.93**	9.06	8.39	-0.89	-1.05	0	5.02	-1.00

Table S5: Important Descriptors for 5-fluorouracil

HERG	Caco	logBB	MDCK	logKp	logKhsa	QP%	PSA	glob
-2.66	190.00	-0.67	145.41	-4.67	-0.74	62.52	87.35	0.95

Table S6: ADME properties of 5-fluorouridine states

Title	polrz	logPC16	logPoct	logPw	logPo/w	logS
U1	19.07	7.28	18.83	18.01	-1.84	-1.68
U2	18.80	7.23	18.92	17.95	-1.83	-1.65
U3	18.80	7.29	18.84	17.95	-1.82	-1.63
U4	18.75	7.17	18.85	17.97	-1.85	-1.66
U5	19.27	7.37	19.60	18.02	-1.84	-1.69
U6	18.84	7.22	18.65	17.99	-1.85	-1.66
U7	18.76	7.24	18.04	17.90	-1.79	-1.63
U8	18.17	6.94	19.45	17.75	-1.71	-1.59

|						
U9	18.10	7.07	17.80	17.89	-1.86	-1.59
U10	19.45	7.33	19.62	18.03	-1.89	-1.74
U11	19.07	7.21	18.72	18.01	-1.90	-1.71
U12	18.86	7.17	19.46	17.94	-1.85	-1.67
U13	18.16	7.08	18.01	17.92	-1.85	-1.60
U14	18.11	7.06	17.93	17.91	-1.85	-1.60
U15	19.25	7.30	18.06	18.08	-1.88	-1.72
U16	18.67	7.08	19.39	17.85	-1.75	-1.64

**Table S7:** Descriptors for 5-fluorouridine states

Title	HERG	Caco	logBB	MDCK	logKp	logKhsa	QP%	PSA	glob
U1	-3.08	40.95	-1.51	27.56	-5.74	-0.85	45.03	142.24	0.92
U2	-2.96	44.55	-1.45	30.22	-5.67	-0.84	45.76	139.97	0.92
U3	-2.87	43.56	-1.43	29.50	-5.69	-0.82	45.61	146.67	0.93
U4	-3.09	44.97	-1.48	30.40	-5.66	-0.86	45.73	138.09	0.91
U5	-3.02	38.05	-1.52	25.47	-5.80	-0.84	44.49	145.84	0.92
U6	-3.07	42.66	-1.49	28.72	-5.70	-0.85	45.30	140.03	0.92
U7	-2.84	50.50	-1.38	34.49	-5.57	-0.82	46.97	141.02	0.93
U8	-2.88	85.99	-1.19	61.23	-5.10	-0.83	51.57	132.93	0.93
U9	-2.86	48.04	-1.40	32.62	-5.62	-0.84	46.16	138.24	0.93
U10	-3.18	32.70	-1.63	21.60	-5.96	-0.86	43.02	142.81	0.91
U11	-3.19	34.95	-1.61	23.09	-5.90	-0.87	43.43	143.28	0.91
U12	-3.05	43.00	-1.49	29.03	-5.72	-0.86	45.36	141.16	0.91
U13	-2.93	50.24	-1.39	34.22	-5.56	-0.84	46.58	136.28	0.92
U14	-2.91	50.81	-1.39	34.64	-5.55	-0.84	46.67	140.84	0.93
U15	-3.24	35.54	-1.60	23.54	-5.85	-0.87	43.69	141.60	0.91
U16	-2.99	65.26	-1.32	45.68	-5.34	-0.84	49.20	139.57	0.92

Molecular Dynamics

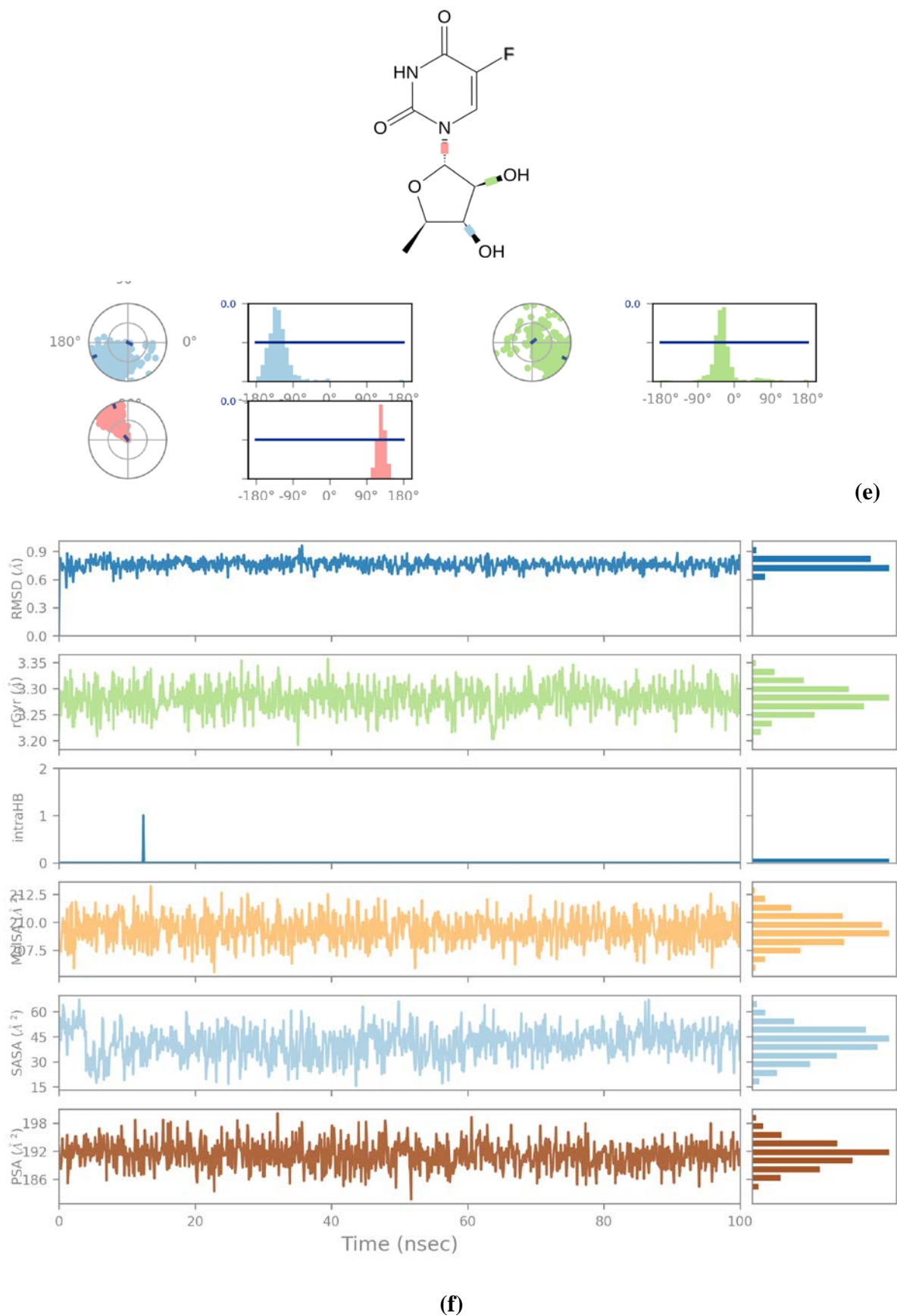(a)

(b)

(c)

(d)

(e)

(f)
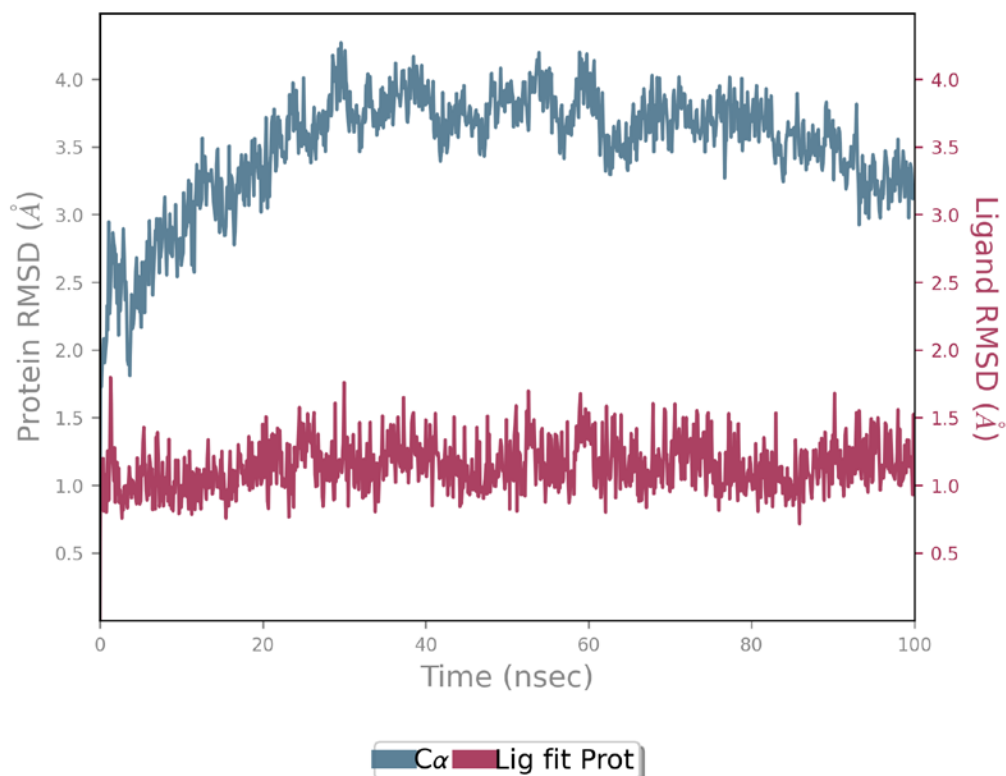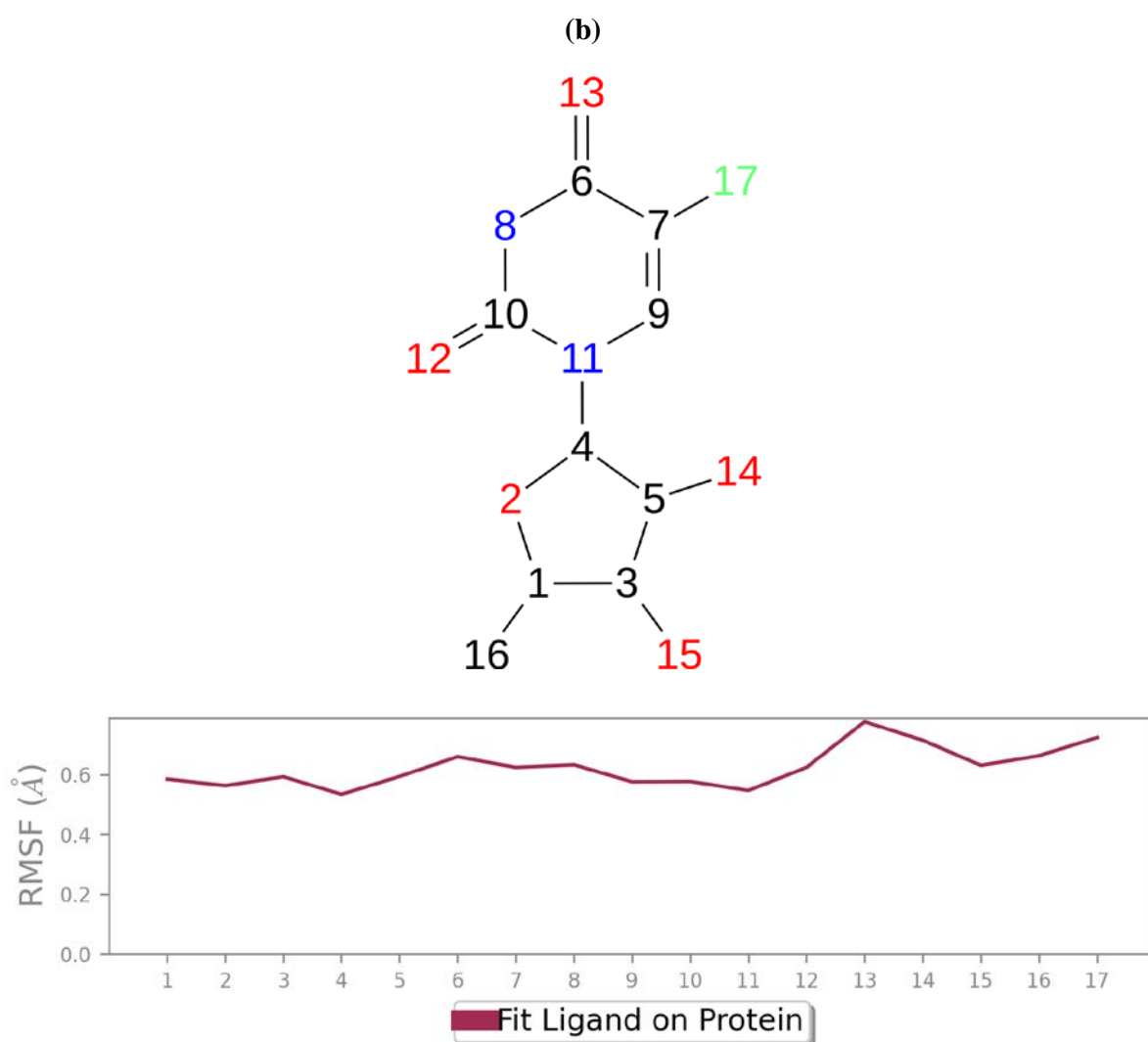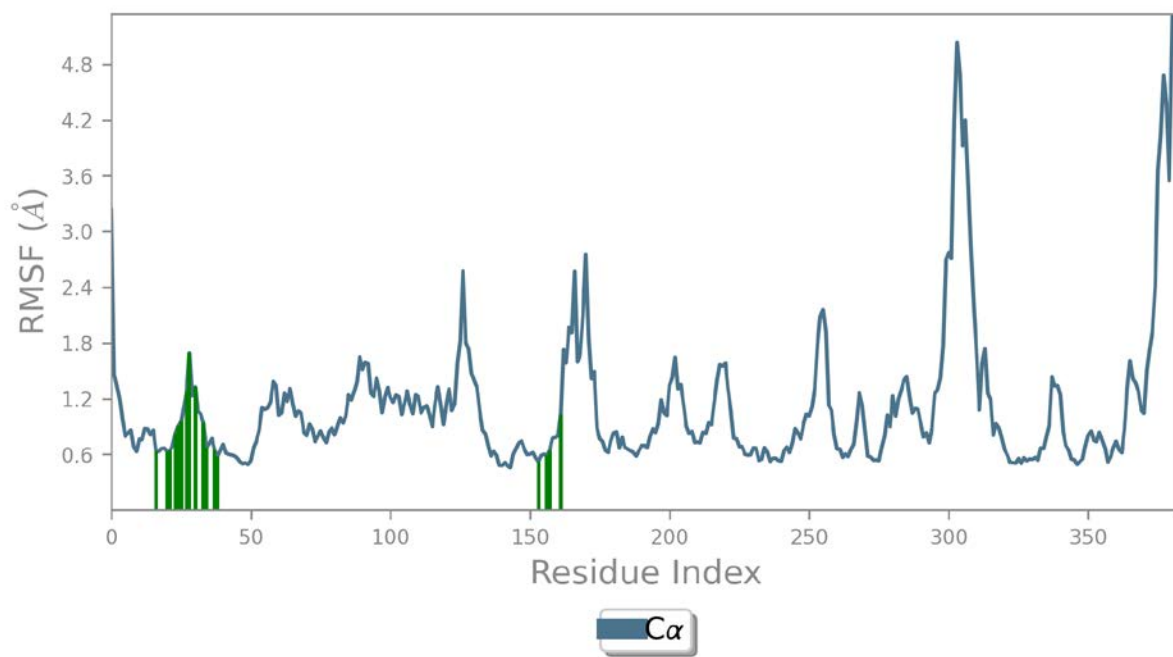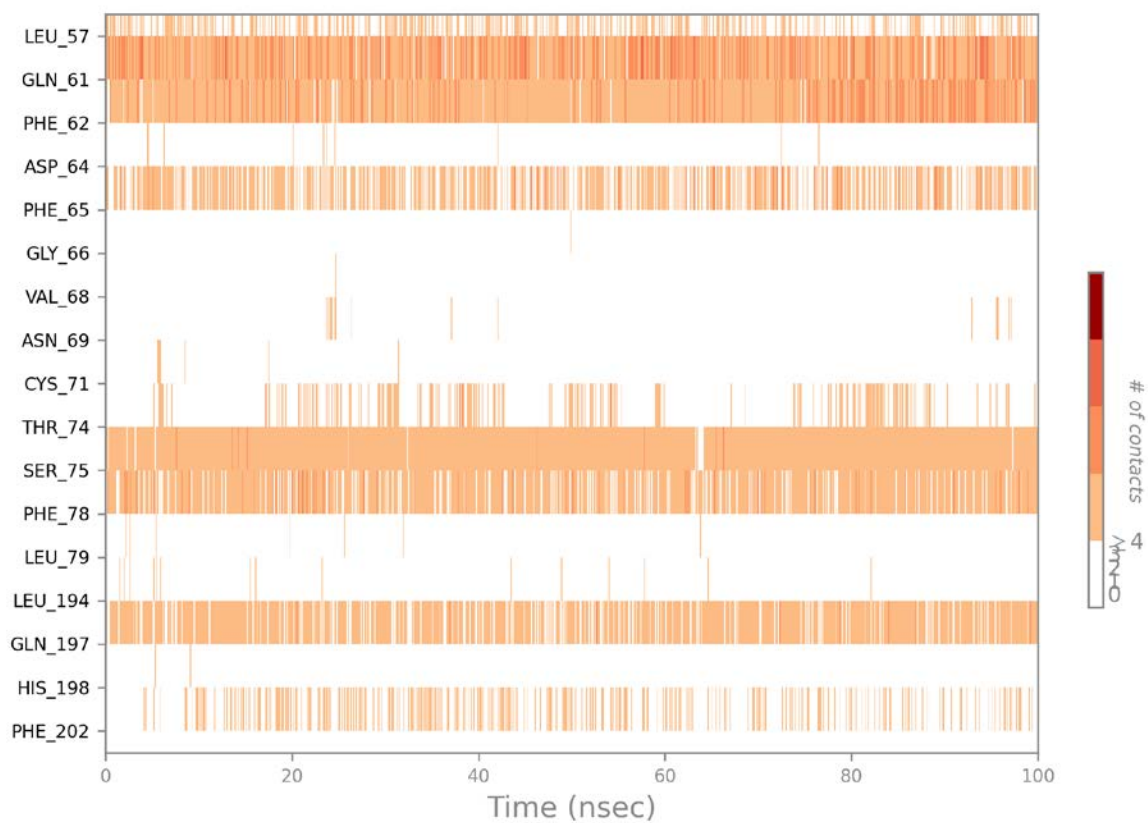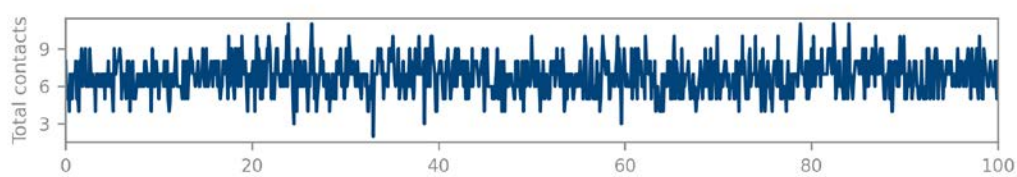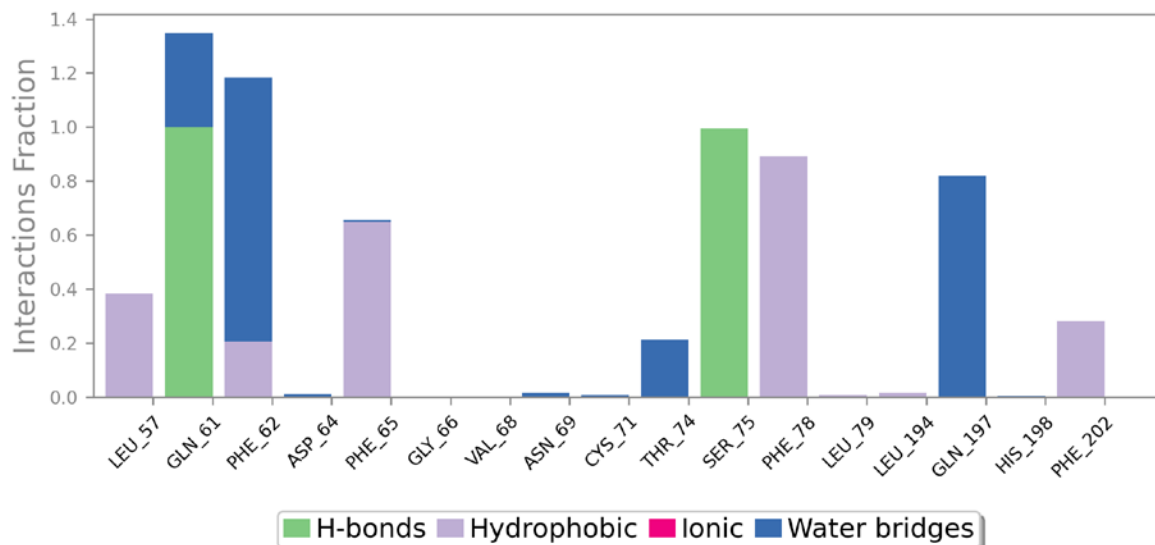
Figure S1. Results of MD simulations. (a) Protein-ligand RMSD (b) Protein RMSF (c) Ligand RMSF (d) Protein-ligand contacts (e) Ligand torsion profile (f) Ligand properties

References

1. W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc* **1996**, 118 (45), 11225–11236.

2. A.E. Cho, D. Rinaldo. Extension of QM/MM docking and its applications to metalloproteins. *J Comput Chem* **2009**, 30 (16), 2609–2616.

3. S.-H. Cho, J.-Y. Kim, J.-H. Chun, J.-D. Kim. Ultrasonic formation of nanobubbles and their zeta-potentials in aqueous electrolyte and surfactant solutions. *Colloids Surf A Physicochem Eng Asp* **2005**, 269 (1–3), 28–34.

4. Y.E. Chung, M.-J. Kim, Y.N. Park, et al. Varying Appearances of Cholangiocarcinoma: Radiologic-Pathologic Correlation. *RadioGraphics* **2009**, 29 (3), 683–700.

5. A. Khandelwal, V. Lukacova, D. Comez, et al. A Combination of Docking, QM/MM Methods, and MD Simulation for Binding Affinity Estimation of Metalloprotein Ligands. *J Med Chem* **2005**, 48 (17), 5437–5447.

6. L. Tian, R.A. Friesner. QM/MM Simulation on P450 BM3 Enzyme Catalysis Mechanism. *J Chem Theory Comput* **2009**, 5 (5), 1421–1431.

7. R.A. Friesner, J.L. Banks, R.B. Murphy, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J Med Chem* **2004**, 47 (7), 1739–1749.

8. T.A. Halgren, R.B. Murphy, R.A. Friesner, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J Med Chem* **2004**, 47 (7), 1750–1759.

9. E. Krovat, T. Steindl, T. Langer. Recent Advances in Docking and Scoring. *Current Computer Aided-Drug Design* **2005**, 1 (1), 93–102.

10. E.M. Duffy, W.L. Jorgensen. Prediction of Properties from Simulations: Free Energies of Solvation in Hexadecane, Octanol, and Water. *J Am Chem Soc* **2000**, 122 (12), 2878–2888.

11. C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* **1997**, 23 (1–3), 3–25.

12. C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings 1PII of original article: S0169-409X(96)00423-1. The article was originally published in Advanced Drug Delivery Reviews 23 (1997) 3–25. 1. *Adv Drug Deliv Rev* **2001**, 46 (1–3), 3–26.

13. E. Harder, W. Damm, J. Maple, et al. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J Chem Theory Comput* **2016**, 12 (1), 281–296.

14. A.Y. Toukmaji, J.A. Board. Ewald summation techniques in perspective: a survey. *Comput Phys Commun* **1996**, 95 (2–3), 73–92.

15. G.J. Martyna, D.J. Tobias, M.L. Klein. Constant pressure molecular dynamics algorithms. *J Chem Phys* **1994**, 101 (5), 4177–4189.

16. G.J. Martyna, M.L. Klein, M. Tuckerman. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *J Chem Phys* **1992**, 97 (4), 2635–2643.