Tetraasteranes as Homologues of Cubanes: Effective Scaffolds for Drug Discovery

*Xiaokun Zhang [a], Chaochun Wei [a], Keli Zong [a], Qidi Zhong [b] and Hong Yan[a,]**

[a]*College of Chemistry and Life Science, Beijing University of Technology, Beijing, P. R. China*

[b]*School of Pharmacy, North China University of Science and Technology, Tangshan, P. R. China*

**Corresponding Author: hongyan@bjut.edu.cn*

Content

Methods

1 DFT calculations

Molecular geometry optimizations and vibrational analyses of the compounds in a vacuum were carried out using the M06-2X function[1] associated with the def2-TZVP[2] basis set. Confirming the absence of an imaginary frequency validated the achievement of local geometric minima.

The Multiwfn 3.8 program[3] facilitated all wavefunction analyses, and Visual Molecular Dynamics (VMD) software[4] generated isosurface maps for various orbitals and real space functions, utilizing files exported from Multiwfn. Gaussian 09 packages[5] were employed to complete all calculations.

2 *In silico* studies

2.1 Molecular docking

All computational modeling utilized the Schrodinger Suite, Version 12.4, Release 2021–2 (Schrodinger, LLC, New York, NY). Crystal Structures of the NMDA receptor (PDB ID: 5WEJ), histone deacetylase (HDAC, PDB ID: 4QA0), and Human σ1 receptor (PDB ID: 5HK1) were obtained from the Protein Data Bank. All proteins were prepared using the default settings in the Protein Preparation Wizard. Ionization states for side chain heteroatoms were applied using Epik at pH 7.4. Ligands were built in Maestro and prepared for docking using LigPrep with protonation states assigned using Epik. Conformational search and minimization utilized Macromodel the OPLS4 force field, and molecular docking was performed with the Glide program with standard precision.

2.2 Molecular dynamics simulations

Molecular dynamics simulations of 200 ns were conducted on the selected complexes to evaluate the stability of the complexes. GROMACS 2022.5 software[6] with the Amber99sb-ILDN force field[7] was selected for performing simulation studies. The Sobtop tool[8] was used to generate ligands topology files under the Gaff force field and the RESP2 charges were calculated by the Multiwfn[3] and Gaussian 09 software[5]. The docked complexes were solvated in a cube box using a transferable intermolecular potential 3P (TIP3P) water model[9]. The gap between the protein chain and the edges of the box was fixed to be 1 Å. The appropriate number of ions was added to neutralize the respective systems. The steepest descent algorithm performed energy minimization to remove the undesirable steric clashes. Long-range electrostatic interactions within the cut-off of 1.2 nm were calculated using Particle Mesh Ewald (PME) approximation[10]. Short-range columbic and Van der Waals interactions with the cut-offs of 1.2 nm were selected. The Linear Constraint Solver (LINCS)[11] technique was used to compute the covalent bond constraints. NVT and NPT ensembles performed the position restraint equilibration for 1 ns each. During simulation, the constant temperature (310K) was maintained using the Berendsen algorithm[12], and the constant pressure (1 bar) was maintained by employing Parrinello-Rahman Barostat[13]. Subsequently, the production run of 200 ns was conducted on the docked complexes at the NPT ensemble using the same parameters mentioned above. For the analysis, the coordinates of the simulated complexes were captured every 2 fs time-frame. The converged complexes were subjected to free energy calculations using the MM-PBSA method. The hydrogen bond frequency was calculated by MDAnalysis.[14]

2.3 Binding free energy calculations

The molecular mechanics-based Poisson-Boltzmann surface area (MM-PBSA) continuum solvation model was used to calculate the binding free energy for the simulated protein-ligand complex. The gmx_MMPBSA package[15] was used to calculate the various components of binding free energy, like electrostatic energy, van der Waals interaction energy, and solvation-free energy (polar + non-polar). Based on the RMSD plots, the last 20000 frames were selected to cover the most stabilized conformational states generated during the simulation run of 200 ns.

Table S1. Structure parameters of compound **1** calculated at M06-2X/def2-TZVP level.

Atoms		Distance (Å)	Atoms			Angle (°)
1	2	1.560	1	2	3	89.992
1	4	1.560	1	2	6	90.018
1	5	1.560	1	4	3	90.001
2	3	1.560	1	4	8	90.016
2	6	1.560	1	5	6	90.021
3	4	1.560	1	5	8	89.999
3	7	1.560	2	1	4	90.002
4	8	1.560	2	1	5	89.976
5	6	1.560	2	3	4	90.006
5	8	1.561	2	3	7	89.990
6	7	1.560	2	6	5	89.985
7	8	1.560	2	6	7	89.995
			3	2	6	90.007
			3	4	8	89.994
			3	7	6	90.008
			3	7	8	89.985
			4	1	5	89.998
			4	3	7	90.013
			4	8	5	89.987
			4	8	7	90.008
			5	6	7	90.032
			6	5	8	89.974
			6	7	8	89.982
			5	8	7	90.012

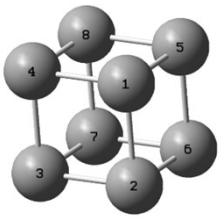


Table S2. Structure parameters of compound **2** calculated at M06-2X/def2-TZVP level.

Atoms		Distance (Å)	Atoms			Angle (°)
1	2	1.534	1	2	3	109.339
1	6	1.533	1	2	7	109.355
2	3	1.534	1	6	5	109.363
2	7	1.533	1	6	10	109.334
3	4	1.534	2	1	6	109.702
4	5	1.533	2	3	4	109.717
4	9	1.534	2	7	8	109.703
5	6	1.534	3	2	7	109.385
6	10	1.534	3	4	5	109.351
7	8	1.534	3	4	9	109.330
8	9	1.534	4	5	6	109.724
8	10	1.533	4	9	8	109.756
			5	4	9	109.315
			5	6	10	109.369
			7	8	9	109.324
			7	8	10	109.358
			9	8	10	109.353
			6	10	8	109.700

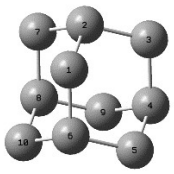


Table S3. Structure parameters of compound **3** calculated at M06-2X/def2-TZVP level.

Atoms		Distance (Å)	Atoms			Angle (°)
1	2	1.548	1	2	3	103.224
1	10	1.545	1	2	6	102.621
1	11	1.528	1	10	4	107.377
2	3	1.531	1	10	9	102.982
2	6	1.576	2	1	10	100.592
3	4	1.525	2	1	11	104.407
4	8	1.563	2	3	4	100.745
4	10	1.556	2	6	5	102.626
5	6	1.548	2	6	7	111.759
5	9	1.545	3	2	6	111.758
5	11	1.527	3	4	8	112.095
6	7	1.530	3	4	10	104.572
7	8	1.525	4	8	7	112.093
8	9	1.556	4	8	9	89.905
9	10	1.558	4	10	9	90.096
			5	6	7	103.217
			5	9	8	107.372
			5	9	10	102.986
			6	5	9	100.596
			6	5	11	104.401
			6	7	8	100.746
			7	8	9	104.572
			8	4	10	89.905
			8	9	10	90.094
			9	5	11	103.541
			10	1	11	103.541
			1	11	5	94.984

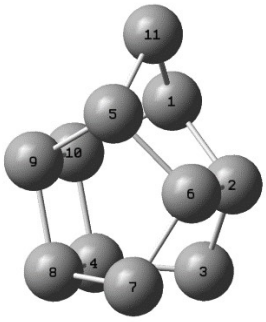


Table S4. Structure parameters of compound **4** calculated at M06-2X/def2-TZVP level.

Atoms		Distance (Å)	Atoms			Angle (°)
1	2	1.558	1	2	3	113.940
1	6	1.524	1	2	8	90.007
1	7	1.557	1	6	5	109.957
2	3	1.524	1	7	8	89.992
2	8	1.557	1	7	12	114.022
3	4	1.523	2	1	6	113.939
4	5	1.558	2	1	7	90.008
4	10	1.557	2	3	4	109.956
5	6	1.523	2	8	7	89.992
5	11	1.557	2	8	9	114.021
7	8	1.558	3	2	8	113.865
7	12	1.523	3	4	5	113.939
8	9	1.523	3	4	10	114.021
9	10	1.524	4	5	6	113.939
10	11	1.558	4	5	11	89.992
11	12	1.524	4	10	9	113.865
			4	10	11	90.008
			5	4	10	89.992
			5	11	10	90.008
			5	11	12	113.864
			6	1	7	113.864
			6	5	11	114.022
			7	8	9	113.939
			8	7	12	113.940
			8	9	10	109.956
			9	10	11	113.940
			10	11	12	113.939
			7	12	11	109.957

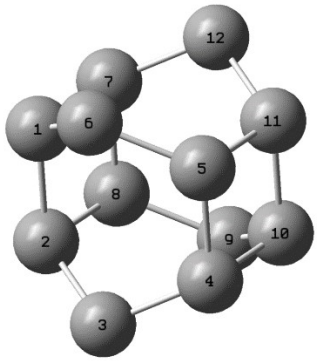


Table S5. Structure parameters of compound **5** calculated at M06-2X/def2-TZVP level.

Atoms		Distance (Å)	Atoms			Angle (°)
1	2	1.553	1	2	3	113.359
1	6	1.562	1	2	7	89.881
1	11	1.456	1	6	7	89.879
2	3	1.522	1	6	10	113.261
2	7	1.562	1	11	5	111.902
3	4	1.522	2	1	6	90.074
4	5	1.553	2	1	11	112.068
4	8	1.562	2	3	4	108.778
5	9	1.562	2	7	6	90.073
5	11	1.456	2	7	12	116.450
6	7	1.553	3	2	7	113.262
6	10	1.522	3	4	5	113.359
7	12	1.456	3	4	8	113.262
8	9	1.553	4	5	9	90.074
8	12	1.456	4	5	11	112.068
9	10	1.522	4	8	9	90.073
			4	8	12	116.450
			5	4	8	89.881
			5	9	8	89.879
			5	9	10	113.261
			6	1	11	116.448
			6	7	12	112.068
			6	10	9	108.782
			7	6	10	113.357
			8	9	10	113.357
			9	5	11	116.448
			9	8	12	112.068
			7	12	8	111.903

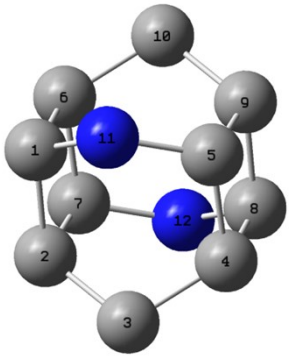


Table S6. Structure parameters of compound **6** calculated at M06-2X/def2-TZVP level.

Atoms		Distance (Å)	Atoms			Angle (°)
1	2	1.554	1	2	3	113.071
1	6	1.554	1	2	7	89.524
1	11	1.414	1	6	7	89.522
2	3	1.520	1	6	10	113.077
2	7	1.554	1	11	5	113.101
3	4	1.520	2	1	6	90.361
4	5	1.554	2	1	11	114.489
4	8	1.554	2	3	4	107.884
5	9	1.554	2	7	6	90.363
5	11	1.413	2	7	12	114.495
6	7	1.554	3	2	7	113.079
6	10	1.520	3	4	5	113.067
7	12	1.413	3	4	8	113.080
8	9	1.554	4	5	9	90.360
8	12	1.413	4	5	11	114.492
9	10	1.520	4	8	9	90.364
			4	8	12	114.494
			5	4	8	89.524
			5	9	8	89.523
			5	9	10	113.075
			6	1	11	114.491
			6	7	12	114.490
			6	10	9	107.886
			7	6	10	113.071
			8	9	10	113.071
			9	5	11	114.493
			9	8	12	114.490
			7	12	8	113.098

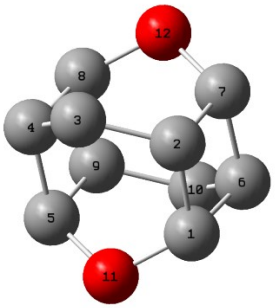


Table S7. Solvation free energy (kcal/mol) of compounds **1**-**6** calculated at M06-2X/def2-TZVP level.

Compounds	Water	Ethyl acetate	Cyclohexane
1	-0.73	-6.07	-6.57
2	1.32	-4.86	-5.04
3	0.50	-5.82	-6.04
4	0.83	-5.91	-6.16
5	-8.61	-8.43	-6.36
6	-6.14	-7.74	-6.31

Table S8. Dipole and quadrupole of compounds **1**-**6** calculated at M06-2X/def2-TZVP level.

Compounds	Dipole (log D)	Quadrupole (a.u.)
1	-3.70	1.91
2	-3.43	2.04
3	-1.57	2.06
4	-5.11	2.10
5	-3.36	2.09
6	-4.06	2.09

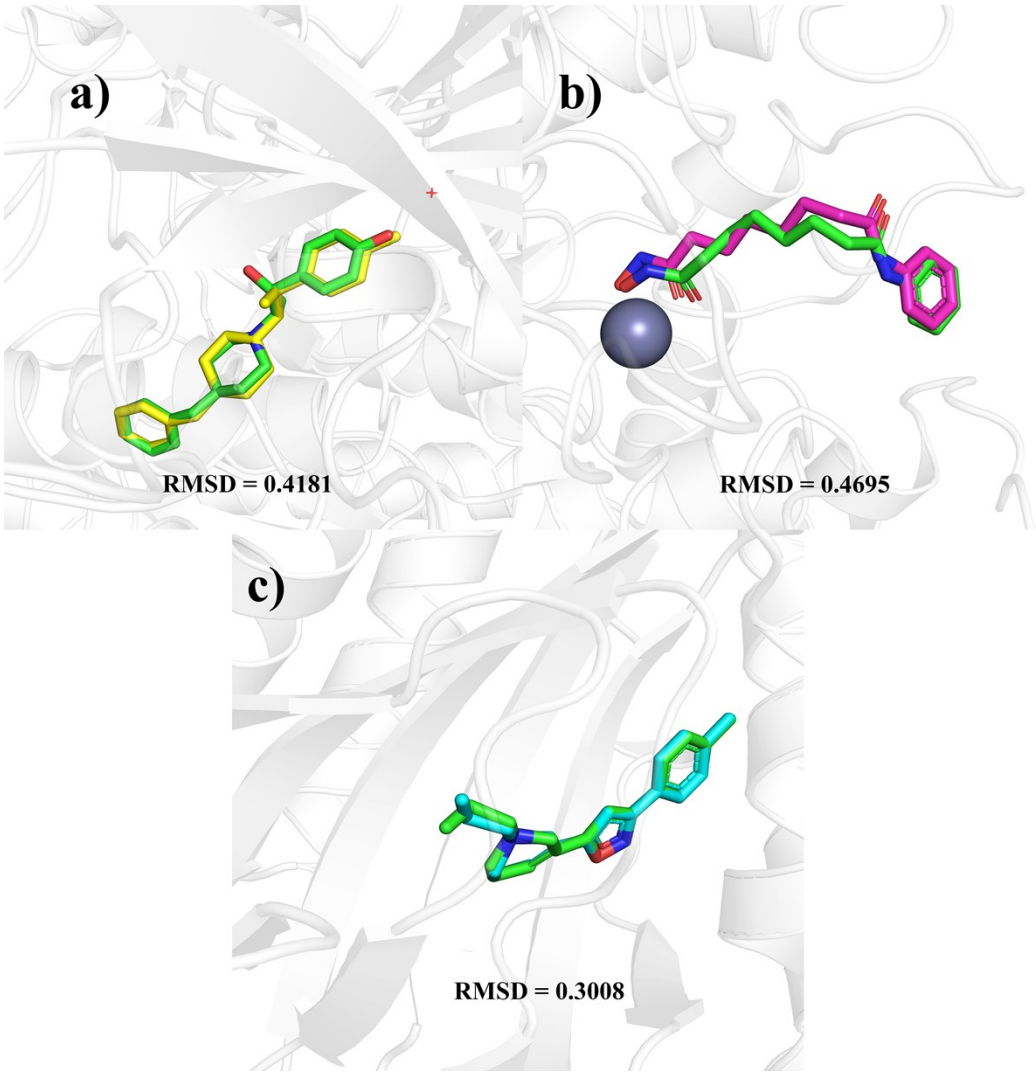
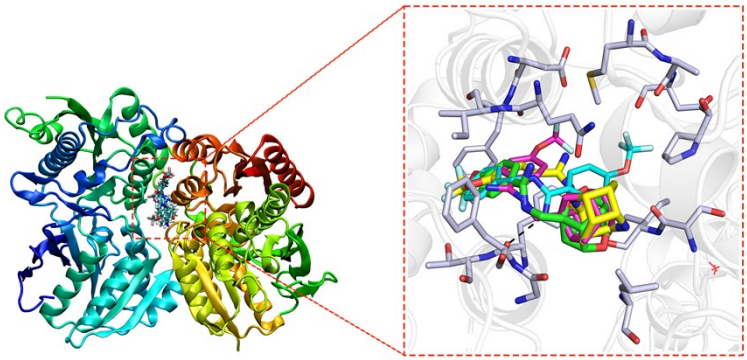
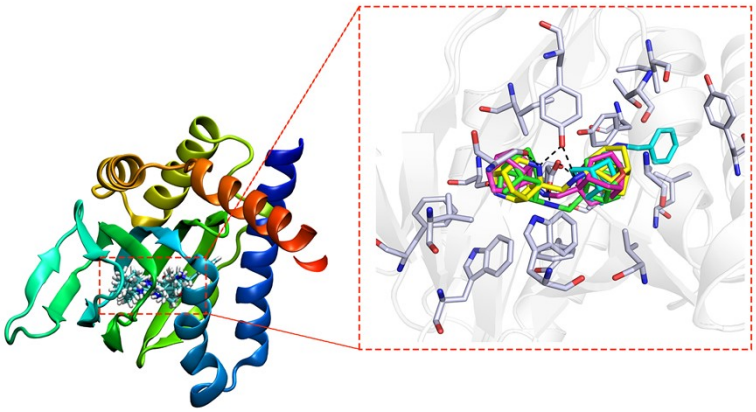


Figure S1. Pictorial representation of the redocked pose, crystallised pose and RMSD values of a) NMDA, b) HDLP and c) σ1 receptor.



Series 1



Series 2

Figure S2. Predicted binding mode of **Series 1** and **2**.

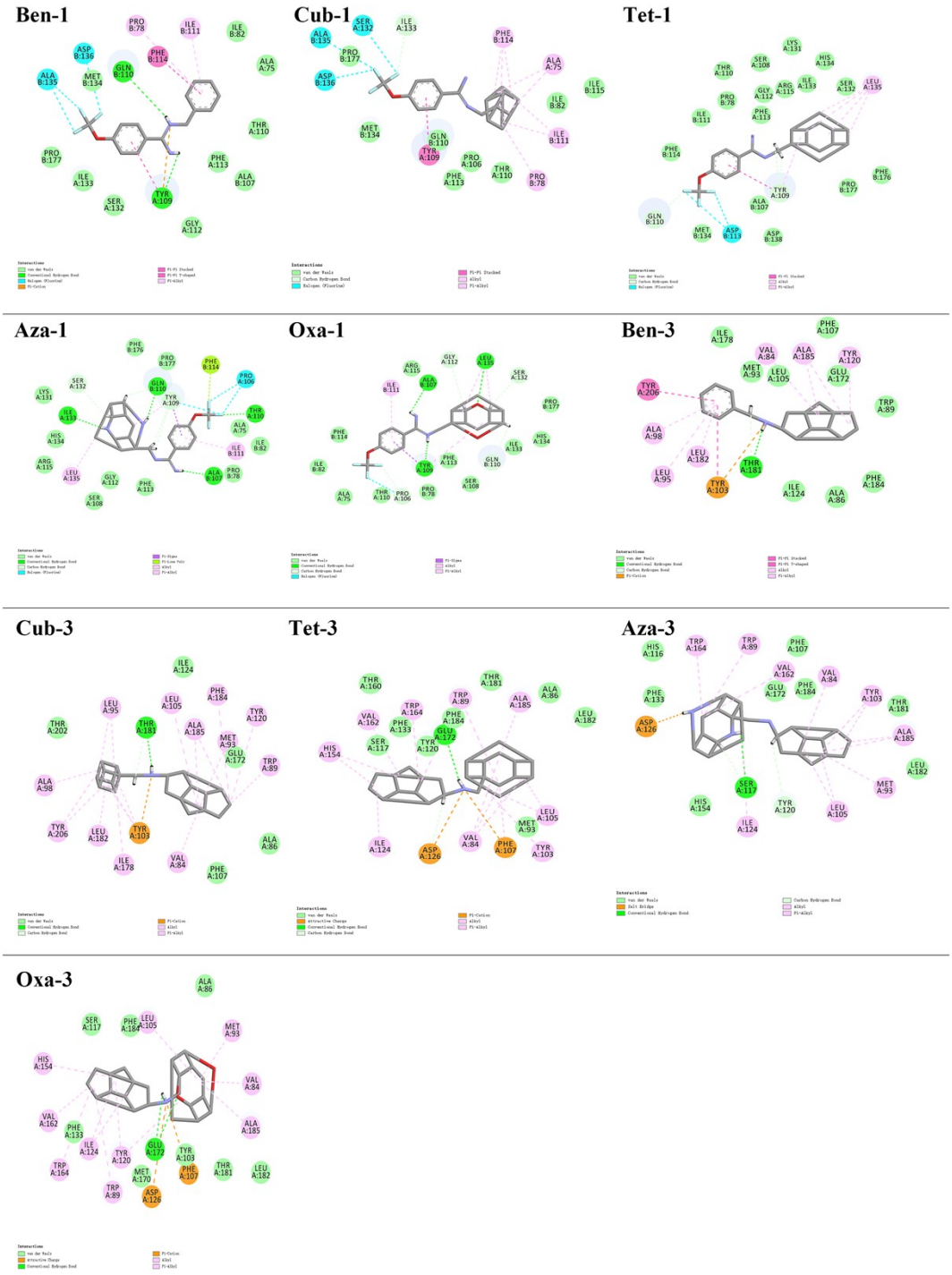


Figure S3. 2D interaction diagrams of **Series 1** and **3**.

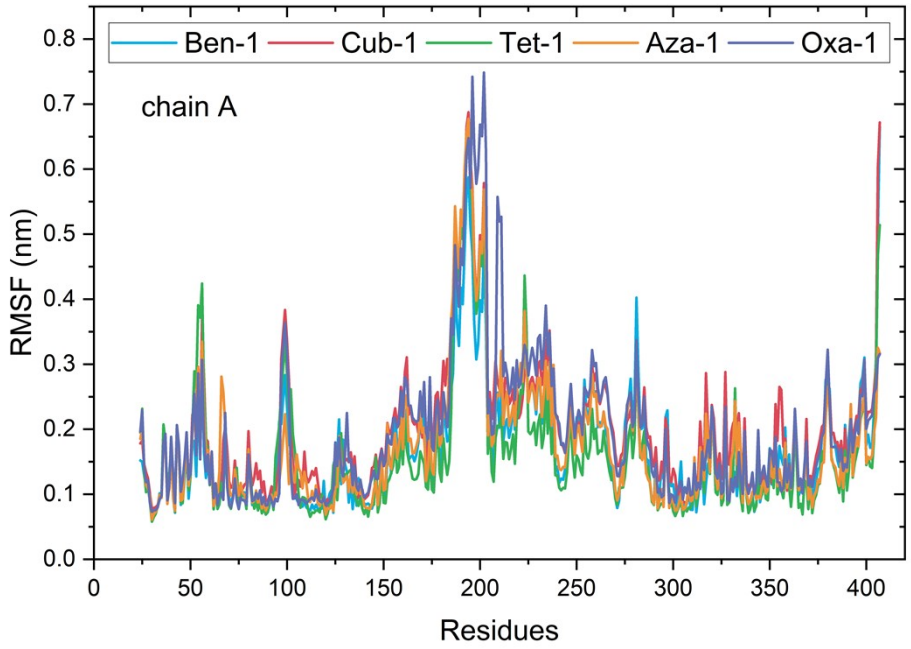


Figure S4. Graphical representation of the protein-RMSF plot of chain A for protein of **Series 1**.

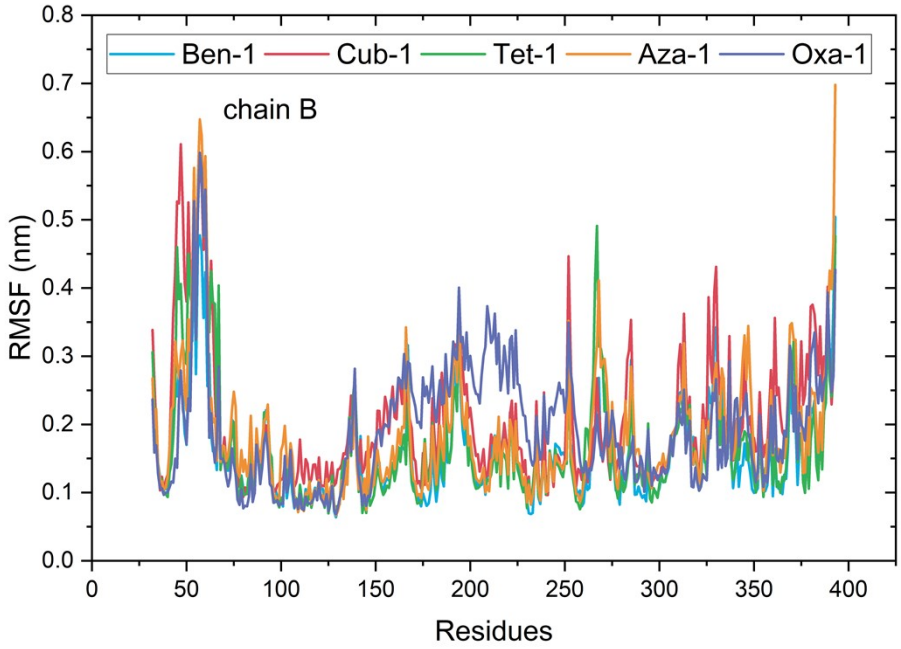


Figure S5. Graphical representation of the protein-RMSF plot of chain B for protein of **Series 1**.

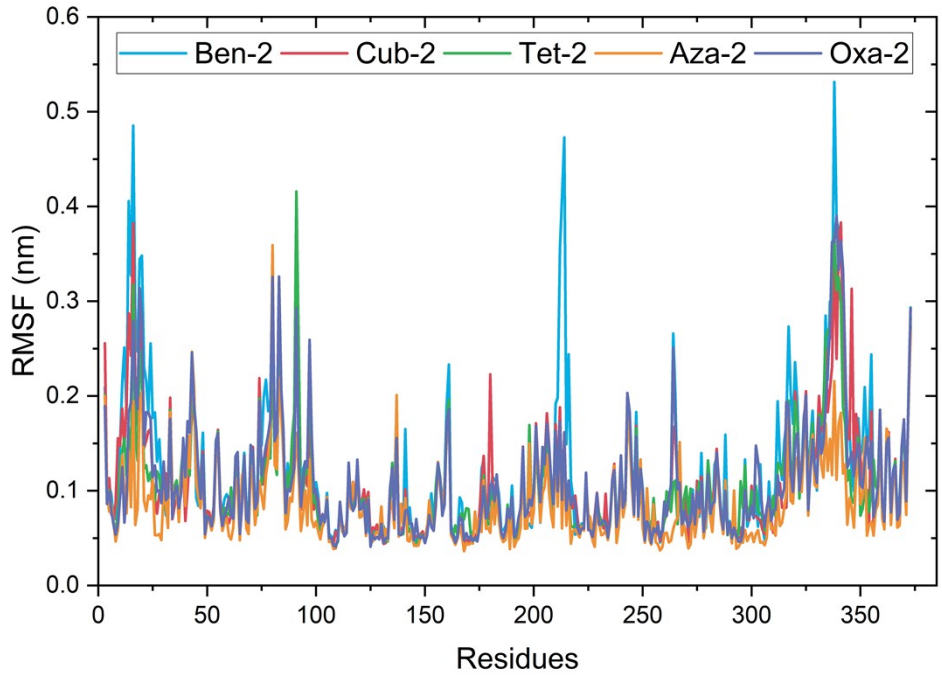


Figure S6. Graphical representation of the protein-RMSF plot of **Series 2**.

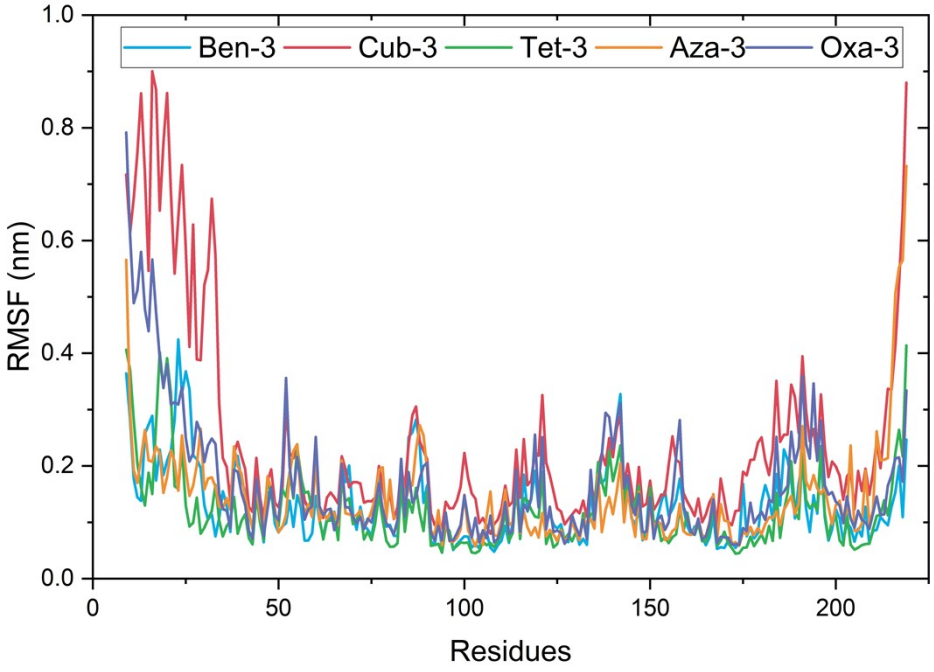


Figure S7. Graphical representation of the protein-RMSF plot of **Series 3**.

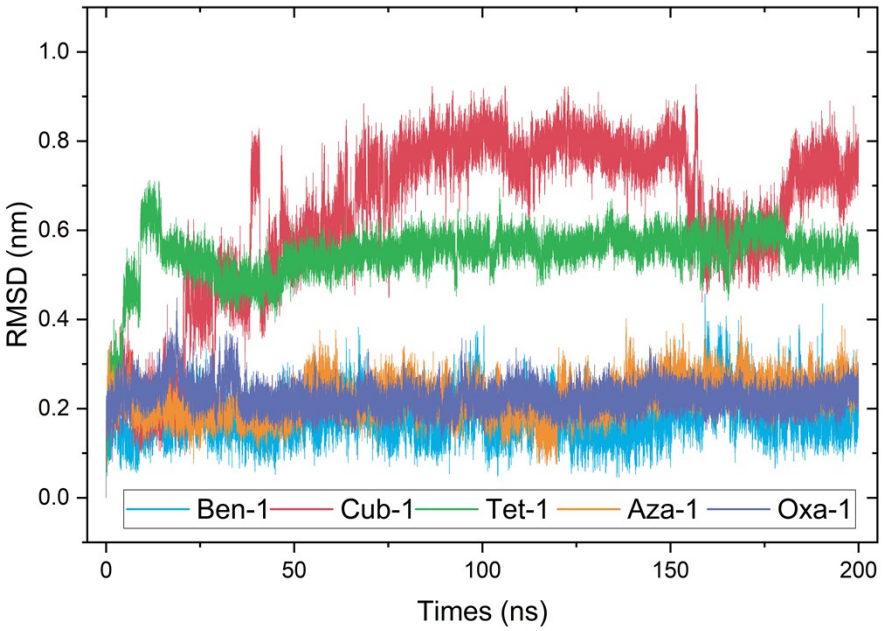


Figure S8. Graphical representation of the ligand-RMSD plot of **Series 1**.

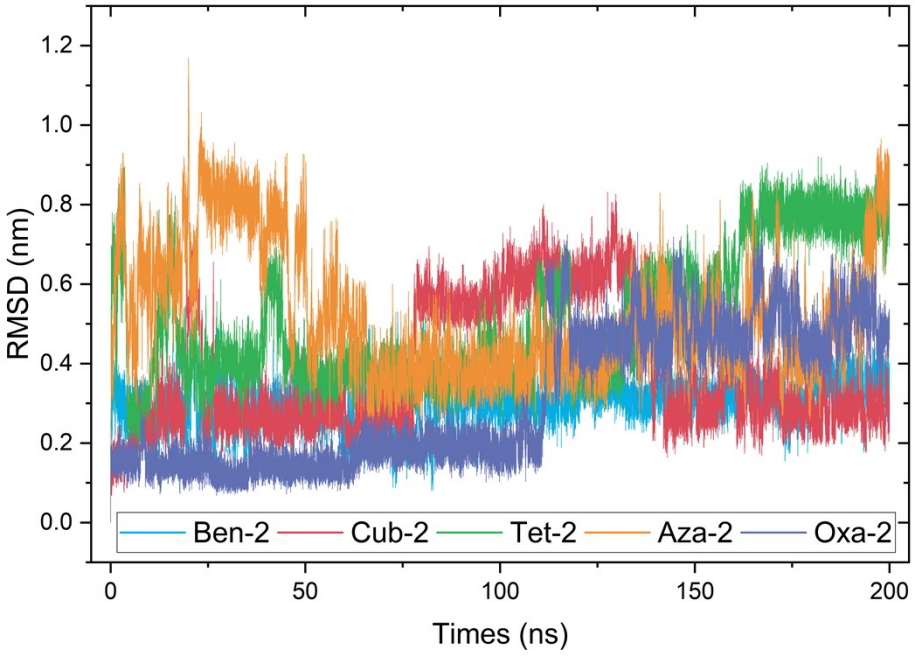


Figure S9. Graphical representation of the ligand-RMSD plot of **Series 2**.

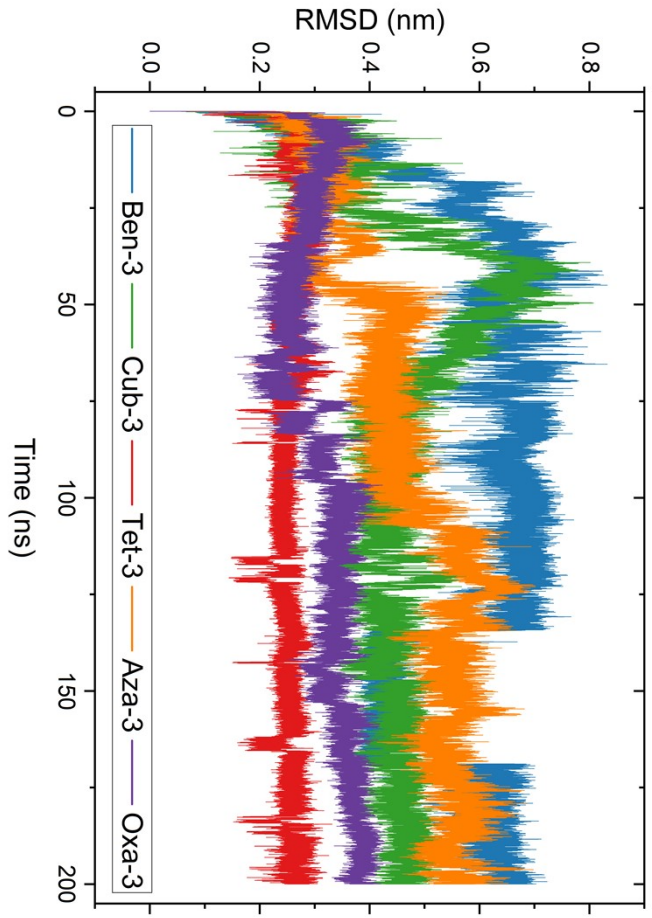


Figure S10. Graphical representation of the ligand-RMSD plot of **Series 3**.

Table S9. Physicochemical and ADMET of Series 1-3.

Investigated compounds												
Tet-1	Aza-1	Oxa-1	Ben-2	Cub-2	Tet-2	Aza-2	Oxa-2	Ben-3	Cub-3	Tet-3	Aza-3	Oxa-3
96.14	96.14	89.00	73.33	75.23	94.45	98.27	87.01	76.52	80.91	98.88	103.00	91.74
45.11	45.11	63.57	78.43	78.43	78.43	102.50	96.89	12.03	16.61	12.03	36.09	30.49
5.69	5.69	3.95	2.28	1.07	2.63	-0.52	1.10	2.77	1.06	3.65	0.30	1.91
4.64	4.64	3.08	1.92	1.37	2.59	0.67	1.18	3.27	2.16	4.28	2.07	2.65
-4.33	-3.13	-2.80	-2.70	-2.57	-3.57	-2.85	-2.37	-3.65	-3.13	-3.09	-3.03	-3.40
High	High	High	High	High	High	High	High	High	High	High	High	High
Yes	Yes	Yes	No	No	Yes	No	No	Yes	Yes	Yes	No	Yes
No	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No	No	No	No	No	No	No	No	No	No	No	No	No
Yes	Yes	No	No	No	No	No	No	No	No	Yes	No	No
Yes	Yes	No	No	No	No	No	No	No	No	No	No	No
No	No	Yes	No	No	No	No	No	Yes	No	No	No	No
No	No	No	No	Yes	Yes	No	Yes	No	No	No	No	No
Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes
No	No	No	No	No	No	No	Yes	No	Yes	No	No	No
-0.130	-0.050	-0.450	0.001	-1.120	-1.290	-0.314	-0.550	-0.590	0.590	0.420	0.775	0.635
No	No	No	No	No	No	No	No	No	No	No	No	No
Yes	Yes	No	No	No	No	Yes	No	No	Yes	Yes	No	No
3.119	2.827	3.107	1.720	1.906	2.294	2.161	2.085	2.509	2.541	2.921	2.552	2.598
0.997	1.770	0.731	1.981	1.842	1.555	1.808	0.993	2.122	0.949	2.344	1.944	1.873
Yes	No	No	No	Yes	Yes	No	No	Yes	No	Yes	Yes	No
-1.650	1.257	1.107	1.016	2.210	1.041	2.393	2.156	0.468	0.721	-0.660	0.919	0.572

		Ben-1	Cub-1
Molecular properties	Molar Refractivity	73.78	76.92
	TPSA (Az)	45.11	45.11
	Log P o/w (WLOGP)	4.81	4.13
	Consensus Log P o/w	3.70	3.47
	Water solubility (log mol/L)	-4.33	-4.23
Pharmacokinetics parameters	GI absorption	High	High
	BBB permeant	Yes	Yes
	P-gp substrate	No	Yes
	CYP1A2 inhibitor	Yes	No
	CYP2C19 inhibitor	Yes	Yes
	CYP2C9 inhibitor	Yes	No
	CYP2D6 inhibitor	Yes	Yes
	CYP3A4 inhibitor	No	No
Drug/lead likeness	Drug likeness (Lipiniski)	Yes	Yes
	Lead likeness	Yes	No
	Ames toxicity	Yes	No
Toxicity parameters	Max. tolerated dose (log mg/kg/day)	0.545	-1.010
	hERG I inhibitor	No	No
	hERG II inhibitor	Yes	No
	Oral rat acute toxicity (LD50) (mol/kg)	2.614	3.081
	Oral rat chronic toxicity (LOAEL)(log mg/kg_bw/day)	1.326	0.992
	Hepatotoxicity	No	No
	Minnow toxicity (log mM)	-0.050	1.220

References

(1) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-class Functionals and 12 other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215-241.

(2) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297-3305.

(3) Lu, T.; Chen, F. Multiwfn: a Multifunctional Wavefunction Analyzer. *J. Comput. Chem.* **2012**, *33*, 580-592.

(4) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33-38.

(5) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 09*; 2010.

(6) 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*, 1701-1718.

(7) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins Struct. Funct. Bioinf.* **2006**, *65*, 712-725.

(8) *Sobtop*; http://sobereva.com/soft/Sobtop, 2023. http://sobereva.com/soft/Sobtop (accessed 2023-09-09).

(9) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926-935.

(10) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089-10092.

(11) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463-1472.

(12) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684-3690.

(13) Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Phys. Rev. Lett.* **1980**, *45*, 1196-1199.

(14) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* **2011**, *32*, 2319-2327.

(15) Valdes-Tresanco, M. S.; Valdes-Tresanco, M. E.; Valiente, P. A.; Moreno, E. gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. *J. Chem. Theory Comput.* **2021**, *17*, 6281-6291.