doi.org/10.36721/PJPS.2024.37.5.REG.949-959.1

New dynamic scoring method for deep evaluation of naloxegol as β-tubulin binding inhibitor

Hamdullah Khadim Sheikh[1,2]*, Jose M Padron[2], Tanzila Arshad[1], Uzma Habib[3], Shahnila Jamil[1], Haroon Khan[4] and Khurshid Ayub[5]

[1]Department of Pharmacognosy, Faculty of Pharmacy and Pharmaceutical Sciences, University of Karachi, Karachi, Pakistan
[2]BioLab, Instituto Universitario de Bio-Orgánica Antonio González (IUBO-AG), Universidad de La Laguna, San Cristóbal de La Laguna, Spain
[3]School of Interdisciplinary Engineering and Sciences (SINES), National University of Sciences and Technology (NUST), Islamabad, Pakistan
[4]Department of Pharmacy, Abdul Wali Khan University, Mardan, Pakistan
[5]Department of Chemistry, COMSATS University, Islamabad, Pakistan

Abstract: We report a new scoring method for rating the performance of ligands on same protein, using their extensive dynamic flexibility properties, binding with protein, and impact on receptor protein. Based on molecular dynamics (MD), this method is more accurate than single-point energy calculations. This method identified an ideal FDA-approved drug as β-tubulin microtubule inhibitor with improved attributes compared to commercial microtubule disassembly inhibitor, Paclitaxel (PTX). We started with virtual screening (VS) of FDA-approved drugs inside PTX's binding pocket (A) of human β-tubulin protein. Screened ligands (>80% score) were evaluated for non-permeation through blood-brain barrier (BBB) as targets were body cancers, gastrointestinal absorption, Lipinski, non-efflux from central nervous system (CNS) by p-glycoprotein (Pgp) and ADMET analysis. This identified FDA-approved Naloxegol drug with superior attributes compared to PTX. Pocket (A) specific docking of chain length variable derivatives of Naloxegol gave docked poses that underwent MD run to give a range of properties and their descriptors (RMSD, RMSF, RoG, H-bonds, hydrophobic interaction, and SASA). QSPR validated that MD properties dependent upon $[-CH_2-CH_2-O-]_{n=0-7}$ chain length of Naloxegol. MD data underwent normalization, PCA analysis, and scoring against PTX. One Naloxegol derivative scored higher than PTX as a potential microtubule disassembly inhibitor.

Keywords: Computer-aided drug design, β-tubulin, ADMET, molecular dynamics simulations, paclitaxel, QSPR modeling.

INTRODUCTION

Microtubules (MTs) are assembled of dimers of α and β-tubulin. MTs constitute the third principal component of the cytoskeleton and are rigid hollow rods that undergo continuous assembly and disassembly within the cell. MTs determine shape and function of cell movements, the intracellular transport of organelles and separation of chromosomes during mitosis through depolymerization. The depolymerization starts with the hydrolysis of guanosine triphosphate (GTP), which is bound to β-tubulin. GTP gets hydrolyzed to guanosine diphosphate (GDP), which weakens the binding affinity of tubulin with adjacent molecules, thereby favoring depolymerization (Gudimchuk *et al*., 2021). During mitosis, these free tubulins reassemble themselves to form mitotic spindles that are further responsible for the chromosome separation of the daughter cells during mitosis. Thus, MTs disassembly plays an important part during cell division. Because of this, several chemotherapeutic drugs target different binding sites of human β-tubulin such as PTX, colchicine and the vinca alkaloids, etc. (Cermak *et al*., 2020). PTX belongs to the category of cytoskeletal drugs that stabilize the microtubule polymer and prevent the disassembly process. As a result, chromosomes are unable to achieve a metaphase spindle configuration, which blocks the progression of mitosis of the cell cycle for cell division (Cheng *et al*., 2020). This mechanism enables PTX to hinder the multiplication of cancer cells of many types such as pancreatic, ovarian cancer, breast, lung and cervical cancer. However, being a naturally obtained molecule with a large and non-rigid structure, PTX has many side effects due to its nonspecificity and multiple absorption, distribution, metabolism, excretion and Toxicity (ADMET) violations. Some other side effects are also due to the use of excipients such as Cremophor EL, polyoxyethylated castor oil, or co-administered drugs like cyclosporine and teniposide, but most of the side effects are due to non-specificity of the structure of PTX itself. Moreover, PTX is not ingestible so cannot be taken orally, and is effluxed by Pgp. PTX also has multiple ADMET violations. Thus, there remains the need for new *structure-based* designed ligands for inhibition of β-tubulin particularly focusing on PTX's specific binding pocket (A) without all these shortcomings. The structure-based drug design approach has three major categories. The first one is VS for the identification of ligands by

*Corresponding author: e-mail: hamdullah.khadim.sheikh@gmail.com

searching through massive databases of known drugs to find those that fit within the binding pocket of the receptor using quick approximation docking. The second category belongs to the de novo design of new ligand molecules that are built inside the determined binding pocket by sequentially assembling fragments (Batool *et al*., 2019). The third method is the structural optimization of already known ligands. In this work, we used VS and lead optimization approaches on the tubulin protein containing PTX ligand. Candidate drugs were screened to fit best inside the PTX pocket using the VS approach with a high affinity. To reduce the side effects of the ligands and keep them from crossing the blood-brain barrier (BBB) while being ingestible, we performed the structural optimization of the VS resultant ligand. We intended to find candidates with a high affinity for PTX's pocket (A) while not crossing the BBB, having less ADMET violations, capable of being ingestible and not effluxed by Pgp. The better binding may lead to higher drug performance and better ADMET attributes can result in reduced side effects. There exist many examples of structure-based drug discovery approach being used in medicinal chemistry for development of newer drug candidates or the optimization of structures (Sabe *et al*., 2021). After deep screening, we used ligand scoring method based on MD resultant ligand's dynamic properties and their binding interactions. The properties included root mean square deviation (RMSD) of ligand and protein, root mean square fluctuation (RMSF) of the ligand, radius of gyration (RoG), solvent accessible surface area (SASA), Lennard Jones (E_{LJ}), Coulombic energy (E_C), H-bonds and hydrophobic interactions. The corresponding descriptors for these properties, such as average values and standard deviations are normalized to ensure a consistent ranking of ligands. Subsequently, Principal Component Analysis (PCA) is employed to assign weights or loadings to these descriptors, ranking them into a comprehensive scoring system. This scoring system effectively rates the ligands against PTX, offering a dynamic assessment of their potential as microtubule disassembly inhibitors. No such study on this specific site, pocket (A) has been performed before in similar depth. Through this extensive screening we found the Naloxegol drug as lead and its mildly modified derivatives with better ADMET character and good docking score. Naloxegol is not proposed as a potential chemotherapeutic candidate before this study either.

MATERIALS AND METHODS

Detailed and stepwise methodology is given in part A of fig. 1. The datasets of calculations performed in this study and the supplementary information are available at https://dx.doi.org/10.17632/56n4b67mbx.1

Reception site preparation & VS

The PDB (.pdb) file 6I2I contains the structure of real human HeLa cancer microtubule protein, downloaded from Protein Data Bank (Liu *et al*., 2021). This structure was generated by electron microscopy (EM) method with a resolution of 3.60Å. The quality of the 6I2I PDB receptor file was verified by Ramachandran plot (An *et al*., 2023). 82.2% of residues were in most favored regions, 17.8% were allowed, while no residue was present in disallowed regions making 6I2I PDB file useable. The attached ligands G2P (phosphomethylphosphonic acid guanylate ester), GTP (guanosine-5'-triphosphate), Mg^{+2} atoms, and TA1 (PTX) were removed along with H_2O molecules, and polar H atoms were added to the structure using Discovery Studio (Version 2021) (Wang *et al*., 2015). The α-tubulin part was also removed and the coordinates of the receptor pocket (A) where ligand TAI was bound were noted as x=38.67, y=58.61 and z=91.64 with radius of 9.81. This was designated as pocket "A". The PDB of 6I2I receptor before and after optimization is available in the dataset and in fig. 1B, while its Ramachandran plot is given in fig. S1 in supplementary information. The VS on the modified 6I2I.pdb was performed on e-LEA3D utility (Singh *et al*, 2021) with x=38.67, y=58.61 and z=91.64 coordinates from the reception site preparation step. We screened e-LEA3D's built-in library of FDA approved. An additional total 6894 drug ligands of which 1379 were FDA approved drugs, 2068 were drug molecules that were approved by non-FDA bodies and 3447 were approved drugs in major jurisdictions (approved by FDA & DrugBank) were screened as well. Libraries of drugs were downloaded from the ZINC database (Potlitz *et al*., 2023).

BBB, Gastrointestinal absorption, PGP, ADMET screening & Lead optimization

Among the entire screened ligands, only the built-in e-LEA3D library gave ligands with non-permeation through BBB, non-efflux by Pgp and gastrointestinal absorption. The screened ligands were further filtered based on the criteria of having more than 80% score. This gave ligands (1-30), available in dataset. Structures (1-30) from VS were first screened for Lipinski's rule of five violations using Swiss ADMET tools and non-permeation through the BBB, since our targets are cancer cells outside the brain. In BBB sphere presented in part C of fig. 1, the points located within the yellow spherical part represent the ligands that may permeate through the BBB. The ones in the white part are supposed to be absorbed by the gastrointestinal tract, suggesting oral availability. Further required attribute was non-efflux from the CNS by Pgp which ensured that ligand persists within the cancer cell and is not pumped out (Waghray and Zhang, 2018). The blue dots indicate that the ligand can be effluxed from CNS by the Pgp, whereas the red ones are predicted to be the opposite. Results of BBB permeation, gastrointestinal absorption and Pgp efflux are given in the BOILED-Egg plot shown in part C of fig. 1 (and fig. S1 in supplementary information).

Upon ADMET screening performed on ligands (1-30) through Swiss ADEMT utility (Sardar, 2023), only Naloxegol (18) fitted the criteria of low violations in ADMET, non-permeation through BBB, non-efflux by Pgp while being gastrointestinal absorbable, hence was chosen for further analysis. It was found that number of ADMET violations was the function of length of $[-CH_2-CH_2-O-]_{n=0-7}$ chain length on structure of Naloxegol (18). Hence to find the best chain length derivative of Naloxegol with lowest ADMET violations, variable polyethylene glycol (PEG) chain length $[-CH_2-CH_2-O-]_{n=0-7}$ were designed and designated as (18a-h) in which Naloxegol itself was designated (18h) with n=7 (fig. 1D). Ligands (18a-h) and PTX were again analyzed for ADMET, BBB and efflux by Pgp (fig. 1C). The data on ADMET of ligands (18a-h) and PTX are available in the table S1 of supplementary information..

Docking

Swiss docking facility (Grosdidier *et al*., 2011) was used to perform accurate mode docking of Naloxegol derivatives (18a-h) and PTX on the 6I2I reception site. This was done to create docked poses to be used in MD. Same PTX specific binding pocket (A) was described in terms of coordinates (as x=38.67, y=58.61 and z=91.64) to restrict ligand interaction within the desired region. After the docking assay, the predicted binding modes were visualized on UCSF Chimera (Jonathan *et al*., 2015) and were downloaded (ZIP file) containing PDB files of docked adducts selective to pocket (A) where TAI (PTX) drug molecule was located initially. Binding mode with highest binding energy (-Kcal/mol) was selected for further analysis through MD. The entire data of docked PDBs and binding modes is available in the dataset. The data on binding energies and docking-related modes of interaction of ligands (18a-h) and PTX are available in tables S2-S3 of supplementary information..

MD run & Calculation of descriptors

MD simulations were carried out using GROMACS software (Version 5.1) (Lemkul, 2019) to evaluate the stability of docking poses of ligands (18a-h) and PTX within the pocket (A) of 6I2I receptor obtained from docking. The topologies and coordinate files of reception site and ligands from the docking study were separately created and parameterized with the CHARM2021 force field. The receptor and ligand adduct was joined again into an adduct, contained within a cubical box in which the adduct was at least 1 nm from the edges of the box to maintain at least 2 nm distance. The adduct was solvated with TIP3P H_2O molecules and Na^+ counter ions to neutralize the solvated system. This was followed by energy minimization with the steepest descent minimization algorithm and conjugate gradient protocol until the maximum force becomes less than 10 kJ mol^{-1} nm^{-1}. Energy minimization was followed by restrained constant number of particles, volume and temperature isochoric-isothermal (NVT) equilibration at 300K and a constant number of particles, pressure (500 psi) and temperature (NPT) ensemble equilibration. This equilibrated ensemble from the resultant trajectory file was finally subjected to MD simulation for 10 ns, with electrostatic and van der Waals cut-off of 1.2 nm. The resultant trajectory files (.xtc) were used for calculation of properties; RMSD (fig. 2), RMSF, RoG, SASA, H-bonds, hydrophobic interactions and interaction energies (fig. S3-S7 in supplementary information). Further descriptors such as average and standard deviation were then calculated. MD-related data of calculated properties, and descriptors are available in the dataset.

QSPR validation of MD descriptors & Scoring functions

Quantitative Structure-Property Relationship (QSPR) was established to validate correlation between molecular MD descriptors and the chain lengths $[-CH_2-CH_2-O-]_{n=0-7}$ of Naloxegol. Only descriptors with strong coefficient of determination (R^2>1) values were chosen for use in scoring functions. The RMSD related QSPR plots are presented in fig. 3 while the rest of the QSPR plots related to RMSF, RoG, H-bonds, hydrophobic interaction and SASA descriptors are given in S7-S9 in supplementary information. Descriptors derived from MD simulations underwent normalization using two normalization methods. For descriptors where a higher value indicates better performance, equation 1 was used, meanwhile, where a lower value is better equation 2 was used.

Normalized Value = (Descriptor Value - Min Value)/(Max Value - Min Value)…. Eq 1.

Normalized Value = 1 - (Descriptor Value - Min Value) / (Max Value - Min Value) …. Eq 2.

Principal Component Analysis (PCA) transformed the normalized descriptors into a set of linearly uncorrelated variables called principal components. Weights were assigned to each descriptor and scoring functions were created by equation 3.

Ligand Score = Σ(Normalized Descriptor × Weight) …. Eq 3.

Ligands were ranked based on sum of their scores. A score higher than PTX indicates a better performance, while a lower score suggests a less favorable performance than PTX. The entire dataset of scoring functions (.xls) is available in supplementary information available in dataset.

RESULTS

The data on descriptors derived from MD properties of ligands (18a-h) and PTX, such as RMSD, RMSF, RoG, SASA, E_{Total}, E_{LJ}, E_C, H-bonds (within 0.35 nm) and hydrophobic interaction, are given below in tables 1-7. While, trends in QSPR of MD properties related descriptors against variable chain length $[-CH_2-CH_2-O-]_{n=0-7}$ and final scores of ligands (18a-h) and PTX are provided in tables 8-9.

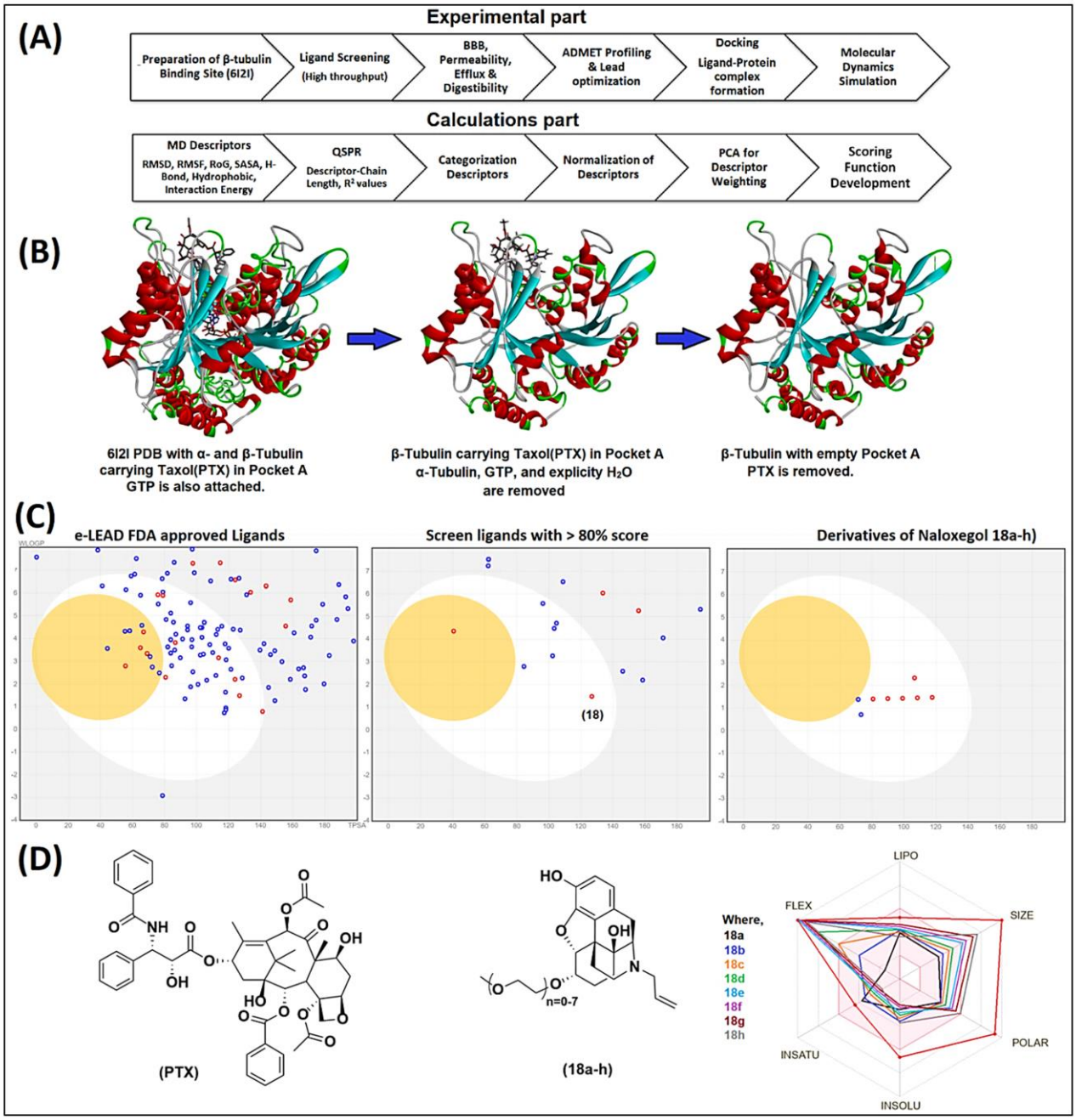

Fig. 1: (A) Entire methodology (B) Original 6I2I protein PDB, 6I2I with only b-tubulin and PTX and 6I2I with no ligand attached, for Virtual screening. (C) ADMET and BBB for Naloxegol and derived ligands (18a-h). (D) Molecular structures of ligands (18a-h) and PTX.

DISCUSSION

Target site preparation, VS & Lead optimization

For structure-based drug design, we used VS and lead optimization approach for pocket (A). Only the Neloxegol (18h) was found to have non-permeation through the BBB, absorption by the gastrointestinal tract, non-efflux from the CNS by Pgp and Lipinski's rule compliance as shown in fig. 1C. Naloxegol is a μ-opioid receptor antagonist, used for opioid-induced constipation (OIC). It is a pegylated derivative of the opioid antagonist naloxone, designed to reduce the gastrointestinal side effects of opioids. The PEG chain increases the molecular size and polarity of the molecule, thus reducing its ability to permeate the BBB thus restricting it to peripheral tissues. The $[-CH_2-CH_2-O-]_n$ chain length was varied to n=0-7 (18a-h), to make the molecule reduced in size, less flexible, thus violations in ADMET in the non-optimized Naloxegol (18h) structure were reduced. The final ADMET results can be seen in fig. 1C.

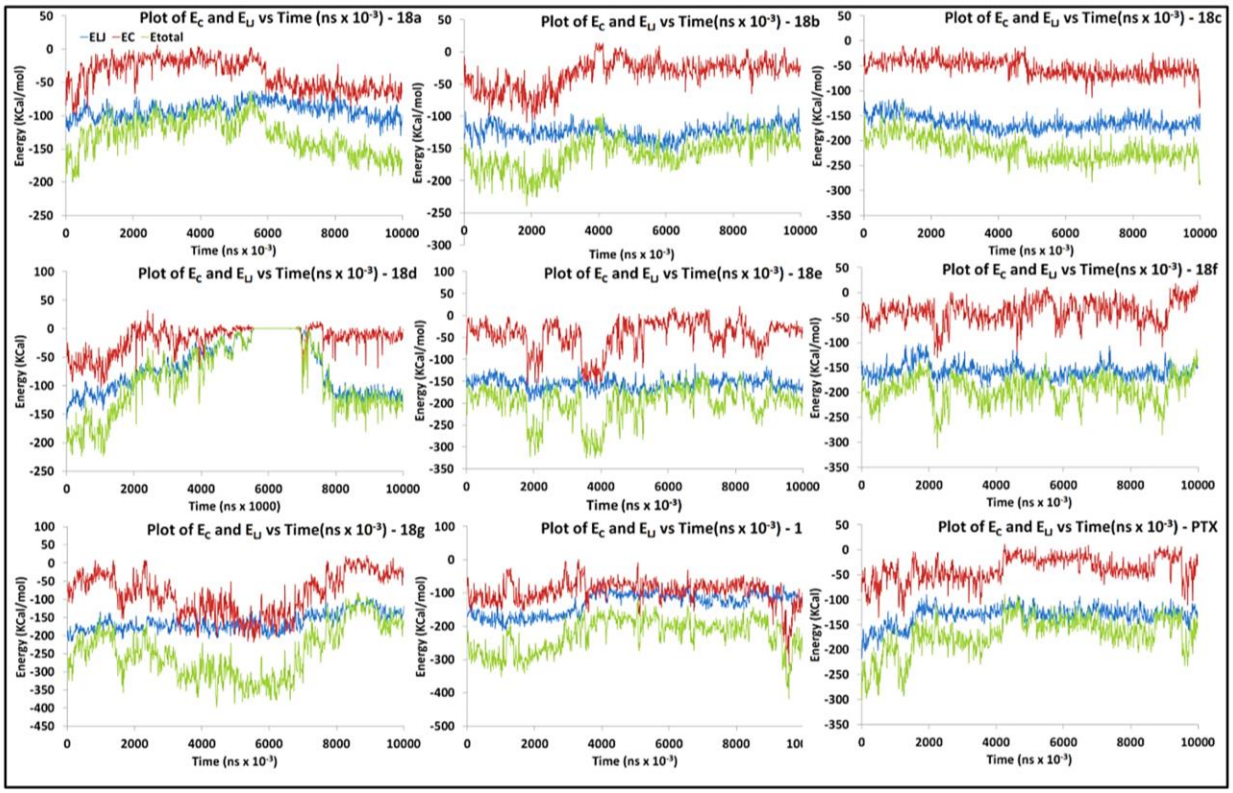

Fig. 2: RMSD plots of ligands (18a-h) and PTX on time scale of 10 ns.

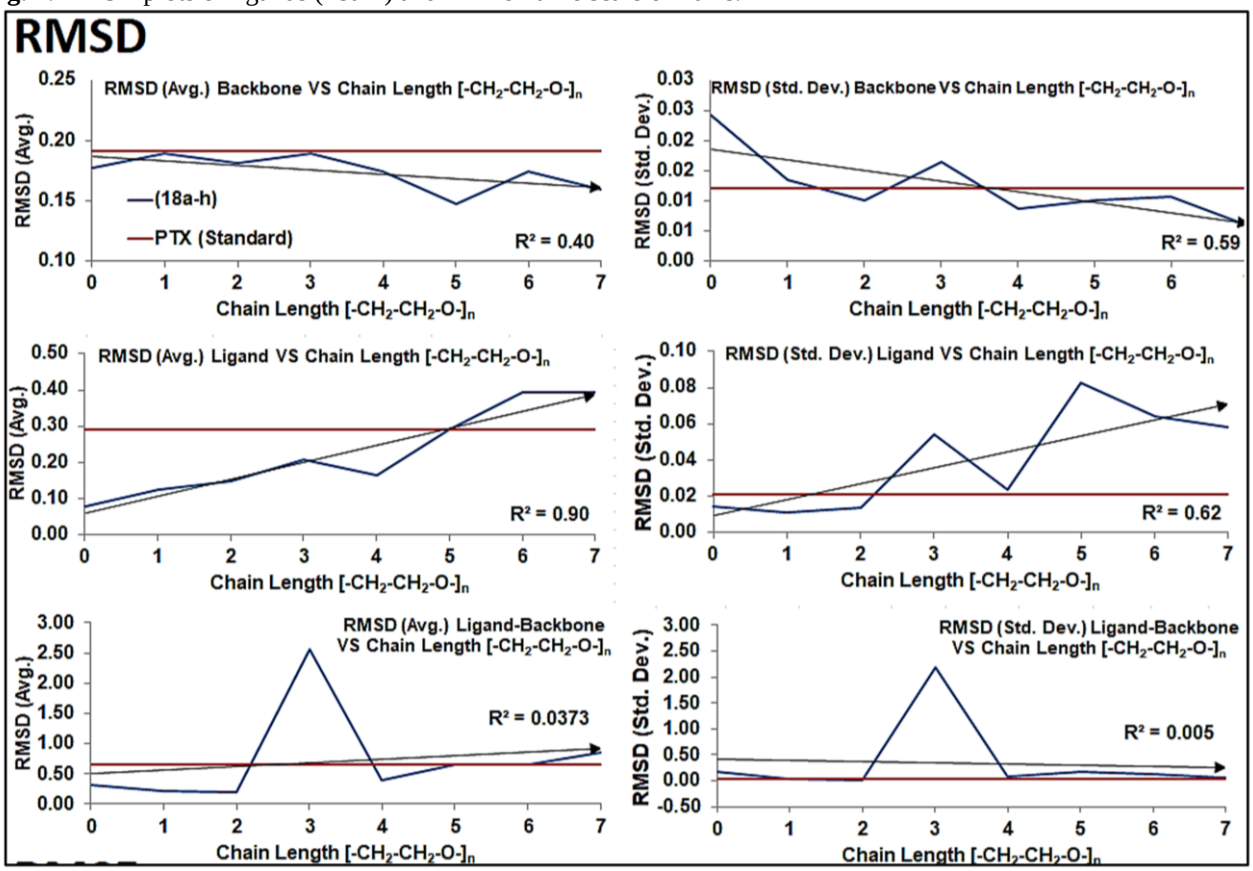

Fig. 3: QSPR plots of RMSD and RMSF related descriptors of ligands (18a-h) and PTX against variable chain length [-CH_2-CH_2-O-]$_n$.

Table 1: Root mean square deviation (RMSD) related descriptors.

Ligand	$[-CH_2-CH_2-O-]_n$	Avg. RMSD	Std. Dev.	Equilibrium Period	Avg. RMSD	Std. Dev.	Equilibrium Period	Avg. RMSD	Std. Dev.	Equilibrium Period
		Backbone			Ligand			Ligand-Backbone		
PTX		0.191	0.012		0.289	0.021		0.644	0.038	
18a	0	0.178	0.024		0.078	0.015		0.314	0.165	
18b	1	0.189	0.014		0.123	0.011		0.216	0.030	
18c	2	0.181	0.014		0.149	0.014		0.195	0.016	
18d	3	0.189	0.016		0.208	0.054		2.551	2.183	
18e	4	0.175	0.009		0.165	0.023		0.397	0.076	
18f	5	0.147	0.010		0.290	0.083		0.649	0.177	
18g	6	0.175	0.011		0.393	0.064		0.641	0.128	
18h	7	0.160	0.006		0.395	0.058		0.855	0.070	

Table 2: Root mean square fluctuation (RMSF) related descriptors.

Ligand	$[-CH_2-CH_2-O-]_n$	Mean RMSF	Max. RMSF	Number of Peaks > Avg. RMSF	Std. Dev. of RMSF
PTX		0.135	0.525	5	0.105
18a	0	0.046	0.193	5	0.044
18b	1	0.066	0.188	6	0.050
18c	2	0.079	0.240	5	0.060
18d	3	0.157	0.349	8	0.066
18e	4	0.096	0.278	9	0.050
18f	5	0.181	0.489	10	0.097
18g	6	0.270	0.598	13	0.113
18h	7	0.270	0.522	13	0.980

Table 3: Radius of gyration (RoG) related descriptors.

Ligand	$[-CH_2-CH_2-O-]_n$	Max. RoG	Min. RoG	Avg. RoG	Std. Dev.	Max. RoG	Min. RoG	Avg. RoG	Std. Dev.	Max. RoG	Min. RoG	Avg. RoG	Std. Dev.
		Ligand-Protein				Ligand				Protein			
PTX		2.18	2.14	2.16	0.0070	0.47	0.39	0.440	0.0100	2.19	2.14	2.16	0.0079
18a	0	2.16	2.12	2.14	0.0054	0.33	0.31	0.322	0.0039	2.17	2.13	2.15	0.0056
18b	1	2.18	2.14	2.15	0.0064	0.40	0.35	0.386	0.0071	2.18	2.14	2.15	0.0063
18c	2	2.18	2.14	2.16	0.0079	0.39	0.47	0.440	0.0106	2.19	2.14	2.16	0.0079
18d	3	2.34	2.13	2.16	0.0290	0.58	0.39	0.500	0.0320	2.18	2.12	2.15	0.0100
18e	4	2.18	2.14	2.16	0.0063	0.44	0.39	0.440	0.0200	2.18	2.14	2.16	0.0063
18f	5	2.17	2.13	2.15	0.0066	0.70	0.42	0.570	0.0620	2.17	2.13	2.15	0.0070
18g	6	2.17	2.13	2.15	0.0083	0.78	0.43	0.550	0.0840	2.17	2.13	2.15	0.0065
18h	7	2.18	2.13	2.15	0.0068	0.79	0.43	0.540	0.0600	2.17	2.13	2.15	0.0070

Table 4: Solvent accessible surface area (SASA) related descriptors.

Ligand	Avg. Ligand SASA	Avg. Polar Ligand SASA	Avg. nonpolar Ligand SASA	Avg Protein SASA	Avg SASA of Protein-Ligand
PTX	10.1	4.35	4.35	191.84	192.64
18a	5.24	2.63	5.79	189.28	187.98
18b	6.36	2.8	6.59	190.7891	188.48
18c	7.15	3.12	7.28	191.33	188.56
18d	7.79	3.25	7.81	186.92	189.64
18e	7.58	3.51	8.33	191.59	189.02
18f	9.39	3.8	9.24	188.61	189.22
18g	9.65	3.98	9.48	189.55	188.45
18h	9.63	4.57	10.17	190.94	191.313

Table 5: Interaction energy (E_I) along with Lennard Jones (E_{LJ}) and Coulombic energy (E_C) related descriptors.

Ligand	Avg. E_J	Std. Dev.	Avg. E_C	Std. Dev.	$E_I = E_{LJ} + E_C$
PTX	-131.64	19.00	-38.40	24.84	-170.04
18a	-91.65	12.27	-38.01	22.94	-129.66
18b	-123.71	11.69	-33.92	22.90	-157.63
18c	-163.94	14.39	-53.08	17.06	-217.02
18d	-69.09	45.18	-17.62	23.55	-86.72
18e	-155.98	13.73	-44.14	35.56	-200.12
18f	-157.41	15.07	-37.47	24.00	-194.88
18g	-84.00	55.84	-162.4	25.17	-246.43
18h	-131.18	33.43	-95.79	36.35	-226.97

Table 6: H-bonds (within 0.35 nm) between ligand-protein related descriptors.

Ligand	Total H-bonds	Avg. H-bonds	Std. Dev. (Stability of H-Bonds)
PTX	1516	1.51	1.36
18a	1472	1.47	1.75
18b	593	0.59	0.93
18c	1019	1.01	1.1
18d	957	0.95	2.28
18e	1520	1.51	1.72
18f	2736	2.73	1.65
18g	2047	2.07	1.64
18h	458	0.45	0.83

Table 7: Hydrophobic interaction between ligand-protein related descriptors.

Ligand	Minimum Distance (Avg.)	Std. Dev.
PTX	0.524	0.099
18a	0.785	0.207
18b	0.493	0.089
18c	0.432	0.063
18d	0.778	0.318
18e	0.479	0.069
18f	0.486	0.078
18g	0.454	0.082
18h	0.513	0.051

Table 8: Trends in QSPR of MD properties related descriptors against variable chain length [-CH_2-CH_2-O-]$_{n=0-7}$, explained as: Descriptor increases/decreases/fluctuates/stable with 'n'; R^2 gives strength of correlation of QSPR; specific derivatives with chain length (n) being higher or lower than the PTX standard line.

Main Property	MD related Property Descriptors vs Chain Length [-CH_2-CH_2-O-]$_n$					
RMSD	Avg. RMSD of Ligand	Std. Dev. Avg. RMSD of Ligand	Avg. RMSD of Backbone	Std. Dev. of RMSD of Backbone	Avg. RMSD of Ligand-Backbone	Std. Dev. of RMSD Ligand-Backbone
QSPR Trend	Increases, R^2=0.90; n=0-5 < PTX, n=6-7 > PTX	Increases, R^2=0.62; n=2, 6 > PTX, n=0-5, 7 < PTX	Stable, R^2=0.40; n=0-7 < PTX	Decreases, R^2=0.60; n=0-1, 3 > PTX n=2, 4-7 < PTX	Stable, R^2=0.03; n=0-7 < PTX Spike at n=3	Stable, R^2=0.005; n=0-7 < PTX Spike at n=3
RMSF	Avg. RMSF of Ligand	Max. RMSF of Ligand	Number of Peaks > Avg. RMSF of Ligand	Std. Dev. of RMSF of Ligand		
QSPR Trend	Increases, R^2=0.86; n=1-4 < PTX, n=6-7 > PTX	Increases, R^2=0.84; n=0-5 < PTX n=6-7 > PTX	Increases, R^2=0.92; n=0-5 < PTX n=6-7 > PTX	Increases, R^2=0.40; n=0-6 < PTX n=7 > PTX		
RoG	Avg. RoG of Ligand	Std. Dev. Avg. RoG Of Ligand	Avg. RoG of Protein	Std. Dev. of Avg. RoG of Protein	Avg. RoG of Ligand-Backbone	Std. Dev. of Avg. RoG of Ligand-Backbone
QSPR Trend	Increases, R^2=0.80; n=0-2 < PTX n=3-7 > PTX	Increases, R^2=0.80; n=0-2 < PTX n=3-7 > PTX	Fluctuates, R^2=0.01; ; n=0-7 < PTX	Stable, R^2=0.01; n=0-7 < PTX Spike at n=3	Stable, R^2=0.02; N=0-7 < PTX	Stable, R^2=0.01; n=0-7 < PTX Spike at n=3
SASA	Avg. Ligand SASA	Avg. Protein SASA	Avg. Ligand-Protein SASA			
QSPR Trend	Increases, R^2=0.92; n=0-7 < PTX	Fluctuates, R^2=0.0012; n=0-7 < PTX	Fluctuates, R^2=0.47; n=0-7 < PTX			
Interaction Energy	Avg. E_{LJ}	Avg. E_C	$E_{Total} = E_{LJ} + E_C$	Std. Dev. of E_{Total}		
QSPR Trend	Decreases, R^2=0.29; n=2-7 < PTX n=0-1 > PTX	Decreases, R^2=0.40; n=2, 4-7 < PTX n=0-1, 3 > PTX	Decreases, R^2=0.40; n=2,4-7 < PTX n=0-1,3 > PTX	Increases, R^2=0.39; n=0-2,4-5 < PTX n=3,6-7 > PTX		
H-Bond	Total H-Bond	Avg. H-Bond				
QSPR Trend	Increases, R^2=0.31; n=0-4,7 < PTX n=5-6 > PTX	Increases, R^2=0.13; n=1-2,7 < PTX n=0, 3-7 > PTX				
Hydrophobic Interaction	Avg. Min. Distance	Std. Dev. of Avg. of Min. Distance				
QSPR Trend	Decreases, R^2=0.21; n=1-2,4-7< PTX n=0, 3 > PTX	Decreases, R^2=0.17; n=1-2,4-7 < PTX n=0, 3 > PTX				

Table 9: Final scores of ligands (18a-h) and PTX.

Ligand Name	[-CH_2-CH_2-O-]$_n$	Score 1: Ligand Dynamic Properties	Ligand Name	[-CH_2-CH_2-O-]$_n$	Score 2: Interaction Energies	Ligand Name	[-CH_2-CH_2-O-]$_n$	Score 3: Impact on Protein
18b	1	-2.492	18c	2	-2.156	18a	0	0.214
18c	2	-2.304	18h	7	-2.030	18d	3	0.346
18e	4	-2.166	18f	5	-1.822	18b	1	0.444
PTX	-	-1.892	18e	4	-1.801	18c	2	0.466
18a	0	-1.650	18g	6	-1.789	PTX		0.472
18d	3	-1.513	18b	1	-1.708	18g	6	0.782
18f	5	-0.962	PTX		-1.598	18e	4	0.861
18g	6	-0.595	18a	0	-0.708	18h	7	1.216
18h	7	-0.510	18d	3	-0.043	18f	5	1.254

For comparison purposes, PTX molecule does not fit within the ADMET radar and has multiple violations against the ADMET attributes. PTX also does not show gastrointestinal absorption and thus cannot be taken orally while the designed ligands show gastrointestinal absorption. One of the added advantages of Naloxegol derivatives (18a-h) is planar structures with lesser rotatable bonds compared to PTX, which reduces its conformational flexibility and enhances stability inside the receptor pocket (A), which was further verified through MD. Compared to this, flexible structure can change its conformations which will result in unstable interactions with the receptor site within the pocket. Ligands (18a-h) also contain less heavy atoms. Although PTX being larger has higher number of H-bond acceptor/donor sites which enhances H_2O solubility. Lipophilicity of ligands (18a-h) with PTX is similar. Ligands (18a-h) have higher gastrointestinal absorption than PTX as can be seen in the fig. 1C. This implies that these ligands can be taken as candidates for oral ingestion. Because of extensive screening, the ligands (18a-h) have no Lipinski violation while PTX structure has two. In the data of ADMET given in the table S1 in supplementary information, lipophilicity (log P values, iLOGP, XLOGP3 and WLOGP) determines cell permeation with higher values implying better membrane penetration. Molecular size, defined by molecular weight and heavy atom count, controls BBB permeation and metabolic stability. Polarity (topological polar surface area, TPSA) evaluates solubility in H_2O. High polarity resists passage through nonpolar bio-membranes. Solubility (ESOL Log S and Silicos-IT LogSw) ensures adequate bioavailability, as compounds need to be soluble enough to reach therapeutic levels yet not so soluble that they are rapidly excreted. Flexibility (rotatable bonds) influences binding conformations, excessive flexibility might reduce binding stability. Lastly, unsaturation in molecule contributes to structural rigidity. Higher rigidity enhances binding stability but may hinder fitting into the binding site.

Docking

Docking of ligands (18a-h) and PTX in pocket (A) gave docked adducts for the MD run. The Swiss Dock facility (http://www.swissdock.ch) automatically sets up the receptor protein and ligand. Swiss Dock allows *local docking* inside a pre-determined binding pocket. In contrast, *blind docking* is performed on all of the available sites on the entire protein. Swiss Dock tool functions on EADock DSS docking software. The calculation algorithm creates favorable 5000-15000 binding modes within 2A^o in pocket, ranked in descending order. The solvent effects are based on the FACTS implicit solvation model. Docking analysis is carried out in the CHARMM22/27 all-H force field (FF). During calculation, PDB structure of the protein is divided into protein and non-protein parts. The protein part is further decomposed into CHARMM segments. SwissDock itself derives the CHARMM topology, parameters and coordinates, from the Mol2 file of the ligand using the Merck molecular force field (MMFF). The dihedral angles and bond lengths are retained while the charges are taken from MMFF. Van der Waals parameters are picked from the closest types of atom in CHARMM22 (Croitoru, 2021). This is followed by user defining three docking preset modes i.e. very fast, fast and accurate. The binding energies of Naloxegol, and derived ligands (18a-h), and PTX in the pocket (A) from docking are given in table S2 in supplementary information, along with clusters (pocket) and elements (binding mode). Ligands (18a-h), gave good level of docking resultant binding energies in -7.5 to -9.5 Kcal/mol range. Lower binding energy suggests a stronger ligand-protein interaction, increased inhibitory activity and stability of the ligand-protein adduct. PTX showed binding energy of -9.25 Kcal/mol. All the ligands (18a-h) and PTX showed no ionic interaction. The details of 2D interactions are given in table S3 in supplementary information. The trend indicates increase in binding energy as the chain length increases from n=0-7, interaction between the drug and the beta-tubulin becomes stronger.

MD

Docked conformations of ligands (18a-h) and PTX was run on Gromacs package (Dolatkhah *et al*., 2017) for a period of 10 ns. MD run on pre-docked conformation transformed static model to a dynamic one over time as the ligand is positioned into energetically favorable configuration within the binding site. From MD run we extracted the MD related properties descriptors (fig. 2 and fig. S3-S6 in supplementary information). From the MD

related properties, further descriptors were calculated to give one numerical value to a dynamic range of data (fig. 3 and fig. S7-9 in supplementary file). RMSD measures the changing distance between the atoms of the ligand and binding site during simulation. A lower average value or standard deviation of RMSD indicates a stable ligand and ligand-backbone structure during MD run, leading to higher efficacy of drug. Average or standard deviation of RMSD of backbone shows changes in protein structure upon ligand binding. Higher variability might have a connection with multiple drug resistance (MDR). Average and standard deviation of RMSD of ligand-backbone evaluates the stability of the ligand-protein complex relative to their original position. Consistent position is desirable for sustained binding. Similarly, RMSF assesses the flexibility of the ligand atoms. Lower average and standard deviation values of RMSF indicate less fluctuation leading to more consistent interactions with the target protein. Standard deviation of RoG of ligand evaluates the change in size of the ligand over 10 ns. Stable size of ligand implies stable interaction with target protein. Standard deviation of RoG of protein and ligand-protein indicates changes in size of protein and ligand-protein complex. Less change in size indicates stable ligand-protein interaction. Standard deviation of SASA of ligand measures the change in surface area of ligand accessible to solvent present in surroundings. Stable values mean ligand is not changing its solvent interactive profile. Standard deviation of SASA of protein measures the change surface area of the protein available to interact with solvent molecules or the ligand. To understand the nature and strength of the dynamic ligand-protein interactive properties E_{Total} was calculated as sum of E_J (Van Der Waal strength) and E_C (electrostatic interactions). We also measured the number and strength of H-bonds between ligands and protein. Hydrophobic interaction of ligand with the hydrophobic sites of the protein was also measured.

Validation of descriptors through R^2 in QSPR

QSPR plots of descriptors of MD properties are given in fig. 3 and fig. S7-S9 in supplementary information. The R^2>1 values in QSPR plots validated the MD properties related descriptors to be dependent upon chain length [-CH_2-CH_2-O-]$_n$ (table 1) and thus to be used in calculation of scoring functions that can rate ligands (18a-h) against PTX. In QSPR plots of average and standard deviation of RMSD (fig. 3) of ligand and backbone, increasing [-CH_2-CH_2-O-] units exhibited higher flexibility leading to unstable binding interactions. High R^2 (0.90, 0.60) values for the QSPR of average RMSD of the ligand and backbone (0.40, 0.60) show strong correlation between chain length and ligand and backbone flexibility, but not the ligand-backbone indicated by the lower R^2 values (0.003, 0.005) showing that the complex remains stable and not capable of contributing to scoring calculations. Ligands (18a-h) displayed more stable interaction than PTX. The RMSD descriptors of ligand contributed to *scoring function 1* related to ligand dynamic flexibility, while the backbone RMSD contributed to *scoring function 3* related to protein stability. Derivative (18c) had better stability compared to PTX in all ligand, backbone, and ligand-backbone RMSD. The *scoring function 1* further included RMSF related descriptors of ligand such as average, standard deviation, number of peaks above average, and maximum RMSF (fig. S7 in supplementary information). The correlation between these descriptors and the chain length (n) was validated by higher R^2 values (0.86, 0.84, 0.92 and 0.40 respectively) and an increase in chain length increased the RMSF, decreasing the stability of ligand-protein interactions. Ligand (18c) was again more favorable, compared to PTX. The RoG (fig. S8 in supplementary information) of ligand also contributed to *scoring function 1*. The average and standard deviation of RoG of ligand exhibited a strong correlation with the chain length with R^2 (0.80, 0.80). Smaller chain lengths showed less changes in RoG compared to PTX, suggesting a stable ligand-protein interactions, compared to higher chain lengths. Lower R^2 values for average and standard deviation of RoG of protein, ligand-protein (0.01, 0.01, 0.02 and 0.02) imply weaker relationship between chain length and protein or ligand-protein volume, hence were not considered in scoring functions. High R^2 (0.93) for the QSPR of average SASA (fig. S8 in supplementary) of the ligand showed its dependence upon chain length, it contributed to *scoring function 1*. QSPR of Average SASA of protein against chain length (n) showed low R^2 (0.0012) values.

The R^2 of QSPR of average Lennard-Jones (E_{LJ}) (0.29), Coulombic (E_C) (0.40), E_{Total} (0.40) and standard deviation of E_{total} (0.39) demonstrated correlation with chain length (n) (fig. S9 in supplementary information). More negative values corresponded to more favorable interactions. Shorter chain lengths exhibited higher interaction energies compared to PTX, suggesting stronger binding affinities compared to longer chains indicating favorable conformational entropy and effective intermolecular binding forces within the pocket. Ligand (18c) showed the most favorable interaction energies across E_{LJ}, E_C and E_{Total} compared to PTX. Binding energy profile, H-bonding and hydrophobic interactions contributed to interaction *scoring function 2*. R^2 values for total and average H-bonds exhibited relationship with chain length (0.31, 0.13) (fig. 9). Longer chains showed increased number of H-bonds because of higher number of H-bonds accepting O atoms. Shorter chains such as ligand (18c) because of having fewer O atoms wielded fewer H-bonds compared to PTX. This could result in more selective or nonspecific interaction profile, minimizing off-target effects. R^2 of average minimum distance and standard deviation (0.21, 0.17) of hydrophobic interaction showed correlation with chain length as well (fig. S8 in supplementary information). Shorter chain lengths such as in ligand (18c) resulted in stronger hydrophobic interactions.

Ligand scoring functions

Normalization scaled the descriptors of variable magnitudes to one range for comparison across ligands (18a-h) and PTX. This was required because MD calculated properties and their further descriptors had different scales and units. Principal Component Analysis (PCA) and weight/loading transformed normalized values into a set of linearly uncorrelated variables known as principal components (Jolliffe and Cadima, 2016). The first components found major variance in the dataset, providing representation of ligand-protein interactions. Using the weights from PCA analysis, *scoring of ligands* provided a quantitative measure for potential effectiveness of ligands as a β-tubulin inhibitor in terms of following three scoring functions (table 2).

Scoring function 1 (ligand dynamic properties)

This scoring function included descriptors of RMSD, RMSF, RoG and SASA of ligands. Ligand (18a-c) with n=0-2 showed stable interaction with the β-tubulin compared to PTX. Longer chain ligands (18d-h) with chain length n=3-7 have lower scores, implying decrease in binding stability. In comparison, PTX is in the middle of the list. Increasing the chain length of the ligands initially leads to better scores upto n=2 after which the performance begins to decline.

Scoring function 2 (Interaction energy)

This scoring function is composed of E_{Total}, E_J, E_C, H-bonds and hydrophobic interaction between ligands and protein. Ligands (18a-c) with n=0-2 showed better electrostatic, Van der Waals, H-bonding and hydrophobic interactions than PTX. As the chain length increases in ligands (18d-h), H-bonds and hydrophobic interactions start to weaken, suggesting that longer chains might hinder binding due to increased flexibility.

Scoring function 3 (Ligand-protein)

Ligands that significantly alter the dynamics and conformation of tubulin could promote MDR, known as MDR inducers, such as PTX. *Scoring function 3* consisted of Backbone and Ligand-Backbone RMSD. Ligands (18a-c) indicated fewer changes in structure of β-tubulin. As chain length increased, there was change in protein or Ligand-Protein structure.

Scoring functions 1, 2 and 3 revealed that ligand (18c) with n=2, demonstrated superior attributes compared to PTX in terms of structural stability, binding affinity with β-tubulin, and its impact on stability of the protein.

This scoring method, because of being dependent upon MD, is more accurate than existing static or single-point energy calculations methods such as molecular (MM) and quantum mechanics (QM) (Sheikh *et al*., 2020; Sheikh *et al*., 2022) that simulate the system at a single point in time, while realistic biological interactions are dynamic in nature involving changes in molecules over time. As a result, static methods are less accurate in evaluating ligand-protein interactions. In future, this dynamic scoring method can be used to compare any new ligands with existing commercial ones in terms of enhanced performance. This scoring method is highly flexible as by incorporating additional MD related properties and their descriptors, method can be fine-tuned further to evaluate any kind of performance desired from ligands such as superior binding attributes, stable interaction with protein, and impact of ligand on the protein etc.

CONCLUSION

Our primary goal was to identify an ideal β-tubulin interactive microtubule disassembly inhibitor for PTX's drug pocket (A) using extensive screening and scoring system. The VS gave a list of probable molecules which were further filtered for less ADMET and Lipinski violations, BBB non-permeability and non-efflux by Pgp analyses. Among the found ligands, Naloxegol derivatives fitted in the pocket (A) better than PTX itself while fulfilling the desired drug attributes. Ligand (18c) with $[-CH_2-CH_2-O-]_{n=2}$ chain, gave the best scores over PTX, in all three functions, indicating less flexibility and interaction energies with lesser impact on the protein structure (reduced MDR). Its oral bioavailability, no ADMET violations, impermeability to the BBB and resistance to Pgp efflux make it a potential candidate. Easy synthesis of ligand (18c) shows its viability for development as a commercial drug. We used a new scoring method based on MD simulation to rank optimized ligand (18c) higher than both PTX and Naloxegol. QSPR (R^2) between MD descriptors and $[-CH_2-CH_2-O-]_n$ chain length validated that the descriptors used in scoring functions. Thus through this new dynamic scoring method, ligand (18c) turns out to be worth pursuing for retrosynthesis, characterization, *in vitro* and *in vivo* analysis.

ACKNOWLEDGMENTS

School of Interdisciplinary Engineering and Sciences (SINES), National University of Sciences and Technology (NUST), Pakistan.

REFERENCES

An X, Zhang W, Rong C and Liu S (2023). Understanding Ramachandran plot for dipeptide: A density functional theory and information theoretic approach study. *Chin. J. Chem.*, **70**(3): 243-252.

Batool M, Ahmad B and Choi S (2019). A structure-based drug discovery paradigm. *Int. J. Mol. Sci.*, **20**(11): 2783.

Cermak V, Dostal V, Jelínek M, Libusova L, Kovar J, Rosel D and Brabek J (2020). Microtubule-targeting

agents and their impact on cancer treatment. *Eur. J. Cell Biol.,* **99**(4): 151075.

Cheng Z, Lu X and Feng B (2020). A review of research progress of antitumor drugs based on tubulin targets. *Transl. Cancer Res.,* **9**(6): 4020.

Croitoru A (2021). Additive CHARMM36 force field for nonstandard amino acids. *J. Chem. Theory Comput,* **17**(6): 3554-3570.

Dolatkhah Z, Javanshir S, Sadr AS, Hosseini J and Sardari S (2017). Synthesis, molecular docking, molecular dynamics studies, and biological evaluation of 4H-chromone-1,2,3,4-tetrahydropyrimidine-5-carboxylate derivatives as potential antileukemic agents. *J. Chem. Inf. Model.,* **57**(6): 1246-1257.

Grosdidier A, Zoete V and Michielin O (2011). Swiss dock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.,* **39**(2): W270-W277.

Gudimchuk NB and McIntosh JR (2021). Regulation of microtubule dynamics, mechanics and function through the growing tip. *Nat. Rev. Mol. Cell Biol.,* **22**(12): 777-795.

Jolliffe IT and Cadima J (2016). Principal component analysis: A review and recent developments. *Philos. Trans. Royal Soc. A.,* **374**(2065):

Jonathan EC, Conrad CH and Thomas EF (2015). RRDistMaps: A UCSF Chimera tool for viewing and comparing protein distance maps, Bioinformatics **31**(9): 1484-1486.

Lemkul JA (2019). From Proteins to Perturbed Hamiltonians: A suite of tutorials for the GROMACS-2018 molecular simulation package [Article v1.0]. *Living J Comp Mol Sci.,* **1**(1): 5068-5068.

Liu H, Gruber CW, Alewood PF, Moller A and Muttenthaler M (2020). The oxytocin receptor signaling system and breast cancer: A critical review. *Oncogene,* **39**(37): 5917-5932.

Potlitz F, Link A and Schulig L (2023). Advances in the discovery of new chemotypes through ultra-large library docking. *Expert Opin. Drug Discov.,* **18**(3): 303-313.

Sabe VT, Ntombela T, Jhamba LA, Maguire GE, Govender T, Naicker T and Kruger HG (2021). Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: A review. *Eur. J. Med. Chem.,* **224**: 113705.

Sardar H (2023). Drug like potential of Daidzein using Swiss ADME prediction: *In silico* approaches. *Phytonutrients,* **2**(1): 2-8.

Sheikh HK, Arshad T, Habib U, Merajoddin M, Mohammad ZS, Usman R and Hasan M (2020). Docking analysis of aryl derivatives of diepoxide alkylating agents. *Pak. J. Pharm. Sci.,* **33**(5): 2017-2021.

Sheikh HK, Arshad T, Habib U, Mohammad ZS, Ahmed M and Hasan M (2022). Colorimetric chromophoric rapid detection of SARS-Cov-2 through breath analysis. *Pak. J. Pharm Sci.,* **35**(1): 157-160.

Singh N, Chaput L and Villoutreix BO (2021). Virtual screening web servers: Designing chemical probes and drug candidates in the cyberspace. *Brief Bioinform.,* **22**(2): 1790-1818.

Waghray D and Zhang Q (2018). Inhibit or evade multidrug Resistance P-Glycoprotein in cancer treatment. *J Med Chem.,* **61**(12): 5108-5121.

Wang Q, He J, Wu D, Wang J, Yan J and Li H (2015). Interaction of α-cyperone with human serum albumin: Determination of the binding site by using discovery studio and via spectroscopic methods. *J. Lumin.,* **164**: 81-85.