**Computational analysis of Amsacrine resistance in human Topoisomerase II alpha mutants
(R487K and E571K) using homology modeling, docking and all-atom molecular dynamics
simulation in explicit solvent**

Safaa Sader and Chun Wu*

College of Science and Mathematics, Rowan University, Glassboro, NJ, 08028 USA

Abstract

Amsacrine is an effective Topoisomerase II enzyme inhibitor in acute lymphatic leukemia. Previous experimental studies have successfully identified two important mutations (R487K and E571K) conferring 100 and 25 fold resistance to Amsacrine respectively. Although the reduction of the cleavage ligand-DNA-protein ternary complex has been well thought as the major cause of drug resistance, the detailed energetic, structural and dynamic mechanisms remain to be elusive. In this study, we constructed human topoisomerase II alpha (hTop2 α) homology model docked with Amsacrine based on crystal structure of human Top2 β in complex with etoposide. This wild type complex was used to build the ternary complex with R487K and E571K mutants. Three 500ns molecular dynamics simulations were performed on complex systems of wild type and two mutants. The detailed energetic, structural and dynamic analysis were performed on the simulation data. Our binding data indicated a significant impairment of Amsacrine binding energy in the two mutants compared with the wild type. The order of weakening (R487K > E571K) was in agreement with the order of experimental drug resistance fold (R489K > E571K). Our binding energy decomposition further indicated that weakening of the ligand-protein interaction rather than the ligand-DNA interaction was the major contributor of the binding energy difference between R487K and E571K. In addition, key residues contributing to the binding energy (ΔG) or the decrease of the binding energy ($\Delta\Delta G$) were identified through the energy decomposition analysis. The change in ligand binding pose, dynamics of protein, DNA and ligand upon the mutations were thoroughly analyzed and discussed. Deciphering the molecular basis of drug resistance is crucial to overcome drug resistance using rational drug design.

Introduction

Topoisomerase II (Top2) is an essential enzyme that manages DNA supercoils and other topological problems by cleaving two strands of DNA duplex, passing another intact DNA helix through a transitory double-stranded break, and subsequently re-ligating the broken DNA helix. Top2 has two subtypes: alpha (Top2 α) and beta (Top2 β). It forms homo-dimer under physiological conditions (Figure 1). Because hTop2 β is required for normal physiological function such as transcription and its inhibition result in adverse effects and even secondary cancers[1], hTop2 α rather than hTop2 β is a preferred drug target[2]. More importantly, hTop2 α is highly abundant in rapidly growing cancer cells which prioritizes its role as anticancer drug target [3]. Due to its central role as DNA topology manager, hTop2 enzyme has been a target to many clinically important inhibitors which can be classified into enzyme catalytic inhibitors and enzyme poisons. Furthermore, enzyme poisons are subclassified into intercalators such as Amsacrine and non intercalators.

Amsacrine is an acridine derivative and considered as the first Top2-target anticancer drug effective in curing acute lymphatic leukemia and myeloid leukemia [4-7]. It acts as a Top2 intercalator poison [8, 9], killing tumorous cells by decreasing rates of DNA relegation and thus increases levels of Top2-DNA covalent complexes. It has relatively weak DNA binding affinity of 10^4 M^{-1} compared to other intercalators such as Adriamycin, Daunorubicin and Actinomycin-D which possess affinity range between 10^5 - 10^6 M^{-1} [10]. Unlike etoposide which inhibit selectively beta subtype, Amsacrine is equally effective in both alpha and beta subtypes of Top 2[8]. Amsacrine molecule consists of two moieties (Figure 1): the intercalative acridine moiety attached to 4'-amino-methanesulfon-m-anisidide head group [11]. The inhibitory effect of Amsacrine relies on the coordinated interplay of those two moieties where acridine moiety functions as DNA anchoring agent, and the head group contributes to the poisoning effect via establishing ligand-protein interactions. Intriguingly, head group has shown be effective in stimulating DNA session and formation of Top2-DNA covalent complex as standalone molecule without the help of intercalating body[11]. Although this was happening at higher

group and Top2 protein is critically important in its mechanism as Top2 poison [11]. Collectively, Amsacrine stabilizes cleavage complex with DNA via both its head group-top2 interaction [12] and DNA intercalation [13].

structure is not shown. Two point mutations are located in TOPRIM region represented in red. hTop2 α complex model was based on the crystal structure of Top2 β (PDB id: 3QX3). (C) Amsacrine-DNA complex showing intercalation mode in WT enzyme.

However, cancer cells can circumvent the inhibitory mechanisms via mutations which ultimately confer drug resistance [14, 15]. Fisher and coworkers have successfully identified two important point

mutations (R487K and E571K) in hTop2 α that confer >100-fold and >25-fold resistance to Amsacrine, respectively [16]. Their experimental data had confirmed the reduced cleavable complex formation as resistance mechanism for these two mutants [16]. Although a high resolution structure of Top2 β in complex with etoposide (PDB id: 3Qx3)[17] and an apo form of hTop2 α (PDB id: 4FM9 [18]) have been recently solved, the complex structures of those hTop2 α mutants in complex with Amsacrine are not available to answer how these mutations reduce the stability of the cleavable ligand-DNA-protein complex at molecular level in similar or different ways. Molecular docking, modeling and molecular dynamics (MD) simulations have been previously utilized to study drug resistance mechanism on both anti-infective agents quinolones [19] and Top1 inhibitor topotecan [20]. Molecular Mechanism-General Born Surface Area (MMGBSA) binding energies has been utilized to detect the mutational impact on ligand-protein [20] and protein-protein interactions [21] with statistically significant correlation between the predicted and experimental binding energies. However, DNA energetics and structural dynamics as a main contributor to drug resistance were not highlighted. Additionally, the concept of ligand-DNA (LD) and ligand-protein (LP) individual contribution to binding energy had not been explored. In this study, we performed a 500ns MD simulation to study these mutation effects on the structure and stability of the cleavable complex of hTop2 α . A comprehensive analysis of the resistance on ligand, DNA and protein was done based on MD simulation trajectories of WT and two mutants systems. Unlike hTop2 β which has been resolved experimentally in complex with an inhibitor (PDB id:3QX3), hTop2 α crystal structure (PDB id: 4FM9) only exist as an apo form bound to DNA (lacking the inhibitor) rendering it unusable in current study. Therefore, we used homology modeling to construct Amsacrine-DNA-hTop2a complex based on the complex structure of hTop2 β with etoposide. The wild type complex structure was used to introduce the two mutants R487K and E571K (Figure 1). Each of the three complexes was subjected to 500 ns molecular dynamics simulation. Our MMGBSA binding energy calculations indicated a significant weakening of Amsacrine binding affinity toward the two mutants when compared to its wild type. Binding energy decomposition showed that while LD interaction was reduced similarly for

both mutants, the reduction of LP interaction in R487K was much larger than that in E571K. Our clustering analysis indicates the changes of the binding poses upon the two point mutations, leading to different ligand-protein and ligand-DNA interactions. In particular, DNA intercalation of ligands was completely and partially lost in R487K and E571K complex structure respectively and a ligand was diffused away from its binding site in R487K. Additionally, the Root Mean Square Fluctuation (RMSF) results for DNA and protein receptors were referring to higher structural and conformational changes in the mutants complex structures compared with wild type.

Computational materials and methods

hTop2 α homology Modelling and docking. Crystal structure of hTop2 β dimer in complex with two etoposides (PDB id: 3QX3) was used as a template to build hTop2 α homology model. First, Prime module implemented in Maestro [22, 23] was used to build hTOP2 α homology model using hTop2 α FASTA sequence (P11388) from Uniprot [24]. Second, the complex structure was obtained by docking two prepared Amsacrine molecules into hTop2 α homolog using Glide XP [25, 26]. To check the accuracy of the docking protocol, we docked back the two crystal ligands of etoposides into hTop2 β and compared the docked poses with the crystal poses (Figure S1). The small Root Mean Square Deviation /RMSD (0.27/0.92Å) indicated the accuracy of our docking protocol. Third, the complex was prepared, optimized and minimized using maestro protein preparation wizard with default parameters. Fourth, this wild type complex structure was used to introduce two mutants (R487K and E571K) in both A and B chains of hTop2 α dimer.

MD Simulation System Setup. All three molecular dynamic simulation systems were built using the prepared and refined receptor-ligand complexes from the Glide XP docking as input files. In each complex, System was built using SPC as water solvent model [27] using orthorhombic solvent box with 6Å water buffer. System was neutralized using Na⁺ ions, and was added with a salt concentration

of 0.15 M NaCl. After the system was successfully solvated, OPLS2005 force field [28] was used to represent the receptor-ligand. The total number of atoms are 127308, 127060 and 127308 for the wide type, R487K and E571K, respectively.

Relaxation and Simulation Protocols. Desmond simulation package [28] was used to run all simulations, systems were first relaxed using the default relaxation protocol. This relaxation protocol consists of multiple stages: 1). Brownian dynamics simulation under the NVT ensemble with temperature 10 K for 100 ps. 2). Simulation under the NVT ensemble with temperature 10 K, small time steps and restraints on solute heavy atoms for 12 ps. 3). Simulation under the NVT ensemble with temperature 10 K and restraints on solute heavy atoms for 12 ps. 4). Simulation under the NPT ensemble with temperature 10 K, pressure 1 bar and restraints on solute heavy atoms for 12 ps. 5). Simulation under NPT with temperature 300 K, pressure 1 bar and no restraints for 1.5 ns. After the relaxation, a 500 ns production run was conducted for each of the three systems under the NPT ensemble (300 K and 1 bar) using the default protocol. In details, temperature was controlled using the Nosé-Hoover chain coupling scheme [29] with a coupling constant of 1.0 ps. Pressure was controlled using the Martyna-Tuckerman-Klein chain coupling scheme [29] with a coupling constant of 2.0 ps. M-SHAKE [30] was applied to constrain all bonds connecting hydrogen atoms, enabling a 2.0 fs time step in the simulations. The k-space Gaussian split Ewald method [31] was used to treat long-range electrostatic interactions under periodic boundary conditions (charge grid spacing of ~ 1.0 Å, and direct sum tolerance of 10^{-9}). The cutoff distance for short-range non-bonded interactions was 9 Å, with the long-range van der Waals interactions based on a uniform density approximation. To reduce the computation, non-bonded forces were calculated using an r-RESPA integrator [32] where the short range forces were updated every step and the long range forces were updated every three steps. The trajectories were saved at 40.0 ps intervals for analysis. It took over 2 months to complete the simulations using 3 GPUs (NVIDIA Tesla M2050).

Simulation interaction diagram (SID) analysis. Simulation Interaction Diagram tool implemented in Desmond was used to analyze the behavior and interaction of proteins and ligands during the course of a simulation. Analysis included Root Mean Square Deviation (Figure S2); protein-ligand contacts (Figure S4-S10) including H-bonds, hydrophobic, ionic and water-bridge contacts; Root Mean Square Fluctuation (Figure S11) measures; secondary structure changes (Figure S12) and ligand torsional profile (Figure S13-S14).

Convergence of molecular dynamic simulations. To check the convergence of MD simulations, we investigated the protein C α and ligand RMSD plots for each trajectory (Figure S2). Clearly, the relatively flat plots within last 200 ns indicates that the complex systems have reached a steady state.

Trajectory Clustering Analysis. Desmond trajectory clustering tool [33] was used to group complex structures from the simulations. Backbone RMSD matrix was used as structural similarity metric, the hierarchical clustering with average linkage [33] was selected as the clustering method. The merging distance cutoff was set to be 2Å. The centroid structure (i.e. the structure having the largest number of neighbors in the structural family) was used to represent the structural family. The centroid structures of populated structural families (>2% of total structure population) are shown in Figure S3 of the supporting material.

MMGBSA Binding energy Calculations. MMGBSA binding energies were calculated using frames from the last 100ns of each system trajectory. OPLS2005 force field, VSGB 2.0 solvation model [34] and the default Prime procedure was used for the MMGBSA calculation. The performance of this method had been evaluated by previous studies done on diverse protein-ligand complexes (855 complexes). A significant correlation coefficient of 0.63 had been calculated between the experimental and predicated binding affinity [35]. Moreover, compared to Molecular Mechanics/Poisson–Boltzmann Surface Area (MMPBSA), MMGBSA displayed better correlation factor and more flexibility in terms of applications to multi-target comparisons [36]. The default procedure consists of three steps: Receptor alone (minimization), Ligand alone (minimization), Receptor-ligand complex

(minimization). The total binding free energy equation is: $\Delta G_{(bind)} = [E_{complex(minimized)}] - [E_{ligand(minimized)} + E_{receptor(minimized)}]$. Binding energy calculations were performed on three complexes: Ligand-Protein-DNA (LPD), Ligand-Protein and Ligand-DNA. To gain a more detailed understanding of binding nature, the individual interaction terms (Coulombic + H-bond + GB solvation+ van der Waals + pi-pi packing + self-contact + lipophilic) were grouped into three categories: (1) $E_{electrostatics}$, (2) E_{vdW} , and (3) $E_{lipophilic}$, where $E_{electrostatics} = (H_{bond} + E_{coulomb} + E_{GB_solvation})$, $E_{vdW} = (E_{vdW} + E_{pi-pi} + E_{self-contact})$ and $E_{lipophilic}$. Note that considering that the solute conformational entropy is not included in MMGBSA analysis, the binding energies by MMGBSA may over-estimate the true binding free energy (i.e. the binding affinity). But whenever the entropic terms are comparable, the MMGBSA binding energies can be used to estimate the relative binding free energy and thus rank the binding drugs.[37] Extensive benchmarking studies up to 1864 crystal complexes have shown that MMGBSA is a powerful tool in ranking ligands.[38-42]

Results

Table 1 Average MMGBSA binding energies of Amsacrine with TOP2 α WT and two Mutants.

ΔG Term	WT	R487K	E571K
-----------------	----	-------	-------

(kcal/mol)	LDP	LD	LP	LDP	LD	LP	LDP	LD	LP
¹ ΔG	-66.4±4.1	-31.9±3.3	-32.7±3.3	-37.0±6.5	-21.5±6.6	-14.9±8.8	-48.2±6.5	-20.9±6.5	-27.1±5.8
² ΔΔG	0	0	0	29.4±6.5	10.4±6.6	17.7±8.8	18.2±6.5	11.1±6.5	5.6±5.8
³ ΔVDW	-56.3±2.5	-33.8±2.2	-22.1±1.4	-42.5±4.2	-27.5±6.2	-15.0±5.5	-46.4±3.2	-27.5±3.4	-18.9±2.7
ΔΔVDW	0	0	0	13.8±4.2	6.3±6.2	7.1±5.5	10±3.2	6.3±3.4	3.2±2.7
⁴ ΔLIPO	-25.9±2.6	-7.9±1.4	-17.6±1.6	-8.6±4.1	-2.2±2.5	-6.4±4.8	-19.1±3.5	-2.7±3.0	-16.3±3.2
ΔΔLIPO	0	0	0	17.3±4.1	5.8±2.5	11.2±4.8	6.8±3.5	5.2±3.0	1.3±3.2
⁵ ΔGBELE	15.8±2.9	9.8±1.5	7.0±2.2	14.1±6.8	8.2±2.9	6.5±2.9	17.3±3.8	9.4±4.3	8.2±1.9
ΔΔGBELE	0	0	0	-1.7±6.8	-1.6±2.9	-0.6±2.9	1.4±3.8	-0.4±4.3	1.1±1.9

¹ ΔG: MMGBSA binding energy (Complex – Receptor – Ligand).
² ΔΔG: Relative binding energy with reference to the WT complex.
³ ΔVDW: Change of van der Waals energy (VDW + Pi-Pi stacking +Self-contact correction) in gas phase upon complex formation
⁴ ΔLIPO: Change of lipophilic term (Lipophilic energy) upon complex formation.
⁵ΔGBELE: Change of electrostatic interactions (GB/Generalized Born electrostatic solvation energy+ ELE/Coulomb energy +Hydrogen-bonding) upon complex formation.
LD: Ligand-DNA Interaction, LP: Ligand-Protein Interaction, LPD: Ligand-Protein-DNA interaction
Note: Binding energy calculated based on average values from two subunits A and B (See Table S1).

MMGBSA binding energy of Amsacrine indicated significant binding affinity weakening in two mutants compared to wild type Top2 α . hTop2 α homodimer contain two ligand binding sites; A and B which required two separate binding energy calculations (Table S1). The binding energy (ΔG) of WT Amsacrine-DNA-Top2 α complex (LDP) were -71.8 kcal/mol and -60.9 kcal/mol in binding site (A) and (B) respectively. R487K mutation attenuated the binding energy to be -39.1 kcal/mol and -34.8 kcal/mol in sites A and B respectively. Similarly, E571K mutation impaired LDP binding affinity to be -42.9 kcal/mol and -53.5 kcal/mol in site A and B respectively. Taking the average over two sites (Table 1), the LDP binding energies of WT, R487K and E571K were -66.4 kcal/mol, -37.0 kcal/mol and -48.2 kcal/mol, respectively. Therefore, R487K and E571K mutation reduced the LDP binding energy by 29.4 kcal/mol and 18.2 kcal/mol, respectively. To gain insight on the relative contribution of ligand-DNA (LD) and ligand-protein (LP) interactions, additional MMGBSA calculations were performed (Table S1). Taking the average over two sites, LD and LP of WT were -31.9 kcal/mol and -32.7 kcal/mol, respectively. R487K reduced LD and LP to -21.5 kcal/mol and -14.9 kcal/mol, respectively. E571K reduced LD and LP to -20.9 kcal/mol and -27.1 kcal/mol. On one hand, R487K and E571K had very similar weakening effects on LD interaction by 10.4 kcal/mol and 11.1 kcal/mol. On the other hand, R487K caused much bigger reduction on LP interaction than E571K (17.7 kcal/mol VS 5.6 kcal/mol). We also decomposed the binding energy into Van der Waals (VDW) and lipophilic (LIPO) and electrostatic interactions (GBELE) in Table 1. Clearly, VDW and lipophilic terms favored the binding for all complexes, whereas GBELE was agonist the binding.

Table 2 Top residues of WT and two mutants contributing to the binding energy and the binding energy change.

WT	$\Delta G(\text{kcal/mol})$	R487K	$\Delta G(\text{kcal/mol})$	R487K	$\Delta\Delta G(\text{kcal/mol})$	E571K	$\Delta G(\text{kcal/mol})$	E571k	$\Delta\Delta G(\text{kcal/mol})$
DT9	-9.3	DA12	-4.5	DT9	9.3	M762	-6.42	DT9	5.57
M762	-6.61	G488	-3.64	M762	6.61	DG13	-4.72	M766	5.41
M766	-5.48	DG13	-2.89	M766	5.49	DT9	-3.73	S763	3.35
S763	-4.06	DC8	-2.73	S763	4.07	G462	-2.05	M765	1.63
M765	-1.74	K487	-1.74	M765	1.74	R804	-1.94	D463	1.17
DA12	-1.69	M762	-1.5	DA12	1.73	S464	-1.37	DC11	1.14
R487	-1.43	DG7	-0.57	K487	1.43	DA12	-1.01	DA12	0.68
DC11	-1.27	k436	-0.55	DC11	1.39	R487	-0.99	DG10	0.65
R804	-1.19			R804	1.2	S763	-0.71		
DG13	-0.84			DG13	0.86	Y805	-0.53		
				K489	0.66	DC8	-0.53		

$\Delta\Delta G$: Relative binding energy with reference to the WT complex.

Figure 2 Graphical representations of Amsacrine-hTop2 α complex showing DNA and protein interacting regions with the ligand. Key residues contributing to binding energy (Δ G) and relative binding energy ($\Delta\Delta$ G) are shown for each binding site. Three complexes were aligned and superimposed using Maestro software.

Binding energy per residue contribution showed essential residues participating in Amsacrine

binding in WT and two mutants. To understand the energetic basis of resistance at residue level, we retrieved the raw decomposition data of MMGBSA and identify the essential residues taking part in ligand binding based on cutoff values (-0.5 kcal/mol for ΔG , 0.5 kcal/mol for $\Delta\Delta G$) as can be seen in Table 2 and Figure 2. In WT, top ten residues contributed -33.61 kcal/mol of total binding energy, including 4 DNA residues and 6 protein residues. Whereas most of these residues are in direct contact with the ligand, positively charged R804 and DT9 are slightly away from the ligand. In R487K mutant, top 8 residues were contributing to -18.12 kcal/mol of total binding energy, including 4 DNA residues and 4 protein residues. Among these residues, only positively charged K436 is not in direct contact with the ligand. In E571K mutant, top 11 residues contributed to -23.5 kcal/mol of total binding energy including 4 DNA residues and 7 protein residues. Among these residues, Y805, R804 and DT9 are slightly away from the ligand. Moreover, we also identified the key residues that weaken the binding energy relative to WT and thus contribute to the drug resistance. Most of these residues are close to the ligand.

Figure 3 (A-C) Representative structure of the most abundant structure families of the three complexes. **(D)** Superimposition of the three complex structures. **(E)** Superposition of ligand-DNA complexes (WT, E571K and R487K are in green, blue and yellow).

Two mutants adopted different binding poses compared with wild type. To explore mutational impact on Amsacrine binding pose and ligand-receptor interaction, we have clustered the trajectory as described in the method section (Figure S2), extracted the most abundant structure of each complex, aligned them using Prime implemented in Maestro (Figure 3) and generated 2D ligand-receptor interaction diagram (Figure S3) and 3D ligand-receptor interaction diagram (Figure 4). The three

ligands in the superimposed complexes was shown in Figure 3E and Figure 4 to visualize the differences between the wild type and mutants structures. Two Mutants exhibited significant DNA intercalation changes compared with wild DNA. In particular, the loss of intercalation state of Amsacrine partially (in case of E571K) and completely (in case of R487K) as seen in Figure 4. The 2D ligand-receptor interaction diagrams (Figure S3) have also shown the significant change in binding interactions between Amsacrine and DNA-protein in two mutants compared with wild type.

Figure 4 Graphical representations of most abundant structures of Amsacrine-DNA interactions. **A:** WT. **B1&B2:** Front and rear views in R487K. **C:** E571K. Residues within 3Å cutoff distance from ligand (blue) are shown in cyan.

To quantify the differences, the ligand-protein contact profile was calculated from the trajectory using Desmond (SID) tool and was shown in Figure S4-S9. This profile contains key protein residues that are in contact with ligand and their timelines. To facilitate the comparison, we tabulated the key residues contacting Amsacrine in the three complexes (Table 3). Obviously, Amsacrine lost some of the interacting residues in case of two mutants as seen in table and sought to bind to some residues significantly different from the wild type. This was clearly noticeable in R487K mutation which exhibited the highest LP binding energy weakening. We noticed that Amsacrine in chain B in this mutation started forming new interactions in the 400 residue range of the protein which was absent in case of the wild type. Consistently, the key residues from our binding energy decomposition analysis (Table 2) are among the residues from this ligand-protein contact analysis (Table 3).

Table 3 Interacting residues with Amsacrine from the SID analysis (see Figure S5-S10).

WT-A	WT-B	R487K-A	R487K-B	E571K-A	E571K-B
			ARG 434		
			ILE 435		
LYS 436		LYS 436	LYS 436		
			GLY 437		
			ILE 438		
			PRO 439		
			LYS 440		LYS 440
			LEU 441		
GLH 461	GLH 461	GLH 461	GLH 461		GLH 461
			GLY 462		
ASP 463	ASP 463	ASP 463	ASP 463		ASP 463
	SER 464	SER 464	SER 464		SER 464
			PRO 485		
LEU 486	LEU 486	LEU 486	LEU 486		LEU 486
ARG 487	ARG 487			ARG 487	ARG 487
		LYS 487	LYS 487		
		GLY 488	GLY 488		
		LYS 489			
		ASN 504			
			ALA 505		
			GLH 506		
		GLU 506			
		LYS 743			
		TYR 757			
		HIS 759			
GLY 760		GLY 760			
GLU 761		GLU 761		GLU 761	
MET 762	MET 762	MET 762	MET 762	MET 762	MET 762
SER 763	SER 763	SER 763			SER 763
LEU 764					
MET 765	MET 765			MET 765	
ILE 766					
	MET 766	MET 766		MET 766	MET 766
	THR 767				
MET 769	ILE 769				
		LYS 798	LYS 798		
SER 800	SER 800			SER 800	
ALA 801	ALA 801			ALA 801	
SER 802	SER 802	SER 802	SER 802	SER 802	
PRO 803	PRO 803			PRO 803	
ARG 804	ARG 804	ARG 804		ARG 804	ARG 804
TYR 805		TYR 805	TYR 805	TYR 805	TYR 805

A and B: Chain A and B.

Figure 5 DNA structure analysis. (A) 2D diagram showing 4 DNA chains simulated in this study (CDEF). (B) DNA 3D representation showing cleavage points and nucleotides sequence in each strand. (C) DNA-RMSF diagrams of WT and two mutants. Three complexes were aligned and superimposed using Maestro software. RMSFs calculations were based on phosphate atoms initial position as reference.

DNA backbone RMSF analysis supported higher structure fluctuation in both mutants

compared to wild type DNA. DNA phosphate RMSF for the three systems was calculated and shown in Figure 5. Four DNA strand ends show larger fluctuation than the middle residues. A small peak was observed at the lesioned position between chain C and D. Over the full residue range, the magnitude of fluctuation was in the following order R487K>E571K>WT. The reduction of ligand-DNA interactions might contribute to the larger DNA structural fluctuation.

Figure 6 Protein C α RMSF diagrams of the hTop2 α WT and two mutants complexes. Residues from 1 to 645 represent the chain A of the enzyme, while residues from 646 to 1330 belong to chain B.

Protein C α RMSF analysis indicated larger conformational fluctuations in both mutants

compared to wild type enzyme. Protein C α RMSF was calculated using Desmond (SID) tool to detect structural fluctuations along the protein chain. Diagrams of the three complexes were combined in Figure 6 for comparison purpose. Results clearly shows that for the most part, R487K mutation reflected the highest fluctuation pattern, followed by E571K mutation which in turn was higher than WT level. The increase of structural fluctuation might be contributed by the decrease of the ligand-protein binding energy. Furthermore, the dominant fluctuation pattern in both binding sites was

consist with the secondary structure distribution of the protein structure where loops and turns are exhibiting the highest dynamic effects compared to helix and β -strand.

Figure 7 Amsacrine (L-RMSF) in WT and two mutants. L-RMSF values represent the average of two binding sites (A and B).

Ligand RMSF indicated larger structural fluctuations in two mutants compared to wild type.

Ligand RMSF was also calculated to characterize the ligand position fluctuations within the complex (Figure S11). Figure 7 shows the average ligand RMSF over the two binding sites for the three complexes. Clearly, the general pattern of fluctuation is significantly higher for the mutants compared to the wild type. More importantly, mutant R487K showed higher ligand fluctuation than mutant E571K. Another worth mentioning is the region of the Amsacrine molecule exhibiting higher

fluctuation. Obviously, the highest fluctuations are seen in the first and the last part of L-RMSF plot which corresponds to the 3-methoxy group of m-AMSA and the 4'-amino-methanesulfon-m-anisidide head group which are flexible groups. However, unexpectedly, another high fluctuation was observed at the middle of the plot which corresponds to the third ring of acridine moiety in Amsacrine responsible for active intercalation inside DNA grove in WT and E571K systems.

Discussions

Previous *in vivo* studies [16, 43] identified two point mutations (R487K and E571K) of hTop2 α that caused more than 100-fold and 25-fold resistance to Amsacrine, respectively. Reduced cleavable complex formation was identified to be the resistance mechanism. However, detailed molecular and structural mechanisms underlying this resistance remains to be elusive. We have applied homology modeling and MD simulations to mechanically understand mutational impact on the complex structure and stability.

Our MMGBSA binding data indicated the significant decrease of the binding energy in the two mutant complex compared to wild type (Table 1). The binding energy decrease caused by R487K mutation (29.4 kcal/mol) is larger than that by E571K (18.2 kcal/mol), indicating that R487K can cause higher fold of drug resistance than E571K by a bigger stability reduction of the ternary complex. This fold change order is qualitatively consistent with the experimental result which indicated 100 and 25 fold drug resistance for R487K and E571K respectively. When decomposing the overall ligand binding into ligand-DNA and ligand-protein interactions, we observed different contribution pattern for these two mutants. In R487K, a $\Delta\Delta G$ value of 17.7 kcal/mol in LP suggested significant impairment of ligand-protein interaction. This is likely the direct impact of the mutation on the protein's ability to form intermolecular interaction with Amsacrine, because R487 is one of the major contributors to Amsacrine protein interaction in the wild type[12]. This also emphasizes the role of Amsacrine

molecule head group in forming the ternary complex [11]. On other hand, $\Delta\Delta G$ of LD in R487K (i.e. 10.4 kcal/mol) is lower than change in LP (i.e. 17.7 kcal/mol), indicating that ligand-DNA interaction contributes less than ligand-protein interaction to overall ligand binding energy decrease. In E571K, decrease of LP (5.6 kcal/mol) is much smaller than decrease of LD (11.0 kcal/mol), indicating that change of ligand-DNA interaction plays bigger role in the overall decrease of the ligand binding energy. This might be explained by the fact that E571K is located far away from the ligand-protein binding site and thus cause smaller perturbation on ligand-protein interaction. When comparing R487 with E571, the larger LP weakening rather than LD in R487 explains its relatively larger fold resistance than E571K. This result is consistent with the role of Amsacrine as top2 α poison [11], LP interaction play more important role than ligand-DNA in the drug action. Furthermore, to understand the molecular nature of resistance at residue level, we investigated the residual breakdown of binding energy of three systems. The enzyme WT and two mutants displayed different set of residues contributing majorly in energy of binding as indicated by table 2. Data from ligand 2D interaction histogram suggests that in WT, DA12 and DC11 were making Pi-Pi and H-bonds with Amsacrine respectively. M762 and M766 contributed by H-bond, hydrophobic and water bridges, M763 contributed by water bridges, M765 contributed by hydrophobic and water bridges and R487 contributed by H-bonds and water bridges. In R487K mutant, acridine moiety in Amsacrine is making Pi-Pi stacking and H-bonding with DG13. And other ligand-protein interactions were observed such as K487 which contributed by H-bond, hydrophobic, water bridges and ionic forces. Moreover, in E571K mutant, R804 formed Pi-Pi stacking and DG13 was making H-bonding with Amsacrine. And S464 was making H-bond and water bridges and M762 was making hydrophobic and water bridges. We further identified the essential residues in R487K and E571K mutants contributed less binding energy of than those of WT to bind to Amsacrine. This lower contribution of binding energy can be seen in terms of $\Delta\Delta G$ from two mutants as tabulated in table 2 which illustrates the highly contributing residues in reducing binding energy.

Consistent with the binding energy reduction caused by the two mutations, our structural data have shown that although the secondary structure (Figure S12) and the tertiary scaffold (Figure 2D) were maintained, the Amsacrine-DNA intercalation changed significantly (Figure 2E and Figure 3). For example, the Amsacrine Molecule moiety was diffused away and relocated to be outside the DNA binding groove as seen in E571K (Figure 4 C) . As such, both LD and LP contacts changed significantly, leading to unstable ligand-DNA-protein complex and higher structural fluctuation (Figure 5-7). This causes Amsacrine lose its inhibitory effect, because hTop2 α could be re-ligated and thus recovered. Our WT structural data agrees the previously published experimental studies that established the role of acridine moiety in Amsacrine to anchor it into the DNA groove [10, 11, 13].

Conclusion

MD simulation enabled us to probe how mutations lead to drug resistance through energetic, structural and dynamic analysis. Our energetic data have shown that the two mutants (R487K and E571K) significantly decreased the binding energy of Amsacrine to hTop2 α -DNA covalent complex, which was consist with the experimental observation of reduction of the ternary complex formation due to the two mutations as the drug resistance mechanism. More importantly, larger drug resistance of R487K than E571K was caused by larger weakening of LP interaction rather than LD interaction. This agrees with the established Amsacrine inhibitory mechanism which emphasize the crucial role of protein in maintaining ternary complex with ligand. In both mutants, weakening of LD interaction also contributed to destabilization of the ternary complex, and partial and complete loss of DNA intercalation was observed in R487K and E571K. Our DNA, protein and ligand RMSF data suggested that the decrease of the binding energy of Amsacrine in WT, E571K and R487K was correlated to the increase of structural fluctuation of protein (C α RMSF), ligand (L-RMSF) and DNA. The successful analysis here support the idea that MD simulation could prioritizes and characterizes the most

significant mutation quantitatively and qualitatively. That is crucial for future drug design of effective ligands to counteract these mutations.

Acknowledgements

This work was supported by Rowan Startup and SEED grant and the National Science Foundation under Grant NSF ACI-1429467 and XSEDE MCB160004/160164.

Reference

- [1] Haffner, M.C., Aryee, M.J., Toubaji, A., Esopi, D.M., Albadine, R., Gurel, B., et al. Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements. *Nature Genetics*. 2010, 42, 668-U45.
- [2] Thakur, D.S. Topoisomerase II inhibitors in cancer treatment. *international journal of pharmaceutical sciences and nanotechnology*. 2011, 3, 1173.
- [3] Dingemans, A.M.C., Pinedo, H.M., Giaccone, G. Clinical resistance to topoisomerase-targeted drugs. *Biochimica Et Biophysica Acta-Genes, Structure and Expression*. 1998, 1400, 275-88.
- [4] Nitiss, J.L. Targeting DNA topoisomerase II in cancer chemotherapy. *Nature reviews. Cancer*. 2009, 9, 338-50.
- [5] Tewey, K.M., Chen, G.L., Nelson, E.M., Liu, L.F. Intercalative Antitumor Drugs Interfere With The Breakage-Reunion Reaction Of Mammalian DNA Topoisomerase-II. *Journal of Biological Chemistry*. 1984, 259, 9182-7.
- [6] Pommier, Y., Minford, J.K., Schwartz, R.E., Zwelling, L.A., Kohn, K.W. Effects Of The DNA Intercalators 4'-(9-Acridinylamino)Methanesulfon-Meta-Aniside And 2-Methyl-9-Hydroxyellipticin On Topoisomerase-II Mediated DNA Strand Cleavage And Strand Passage. *Biochemistry*. 1985, 24, 6410-6.
- [7] Minford, J., Pommier, Y., Filipinski, J., Kohn, K.W., Kerrigan, D., Mattern, M., et al. Isolation Of Intercalator-Dependent Protein-Linked DNA Strand Cleavage Activity From Cell-Nuclei And Identification As Topoisomerase-II. *Biochemistry*. 1986, 25, 9-16.
- [8] Marsh, K.L., Willmore, E., Tinelli, S., Cornarotti, M., Meczes, E.L., Capranico, G., et al. Amsacrine-promoted DNA cleavage site determinants for the two human DNA topoisomerase II isoforms alpha and beta. *Biochemical Pharmacology*. 1996, 52, 1675-85.
- [9] Huff, A.C., Ward, R.E., Kreuzer, K.N. Mutational Alteration Of The Breakage Rejoining Subunit Of Bacteriophage-T4 DNA Topoisomerase Confers Resistance To Antitumor Agent m-AMSA. *Molecular & General Genetics*. 1990, 221, 27-32.
- [10] E.Graves*, R.M.W.a.D. Thermodynamics of the interactions of m-AMSA and o-AMSA with nucleic acids: influence of ionic strength and DNA base composition. *Nucleic acids research*. 1989, 17, 9933.
- [11] Ketron, A.C., Denny, W.A., Graves, D.E., Osheroff, N. Amsacrine as a Topoisomerase II Poison: Importance of Drug-DNA Interactions. *Biochemistry*. 2012, 51, 1730-9.
- [12] Farsani, F.M., Ganjalikhany, M.R., Dehbashi, M., Naeini, M.M., Vallian, S. Structural basis of DNA topoisomerase II-alpha (Top2-alpha) inhibition: a computational analysis of interactions between Top2-alpha and its inhibitors. *Medicinal Chemistry Research*. 2016, 25, 1250-9.
- [13] Jangir, D.K., Dey, S.K., Kundu, S., Mehrotra, R. Assessment of amsacrine binding with DNA using UV-visible, circular dichroism and Raman spectroscopic techniques. *Journal of Photochemistry and Photobiology B-Biology*. 2012, 114, 38-43.
- [14] Dong, K.C., Berger, J.M. Structural basis for gate-DNA recognition and bending by type IIA topoisomerases. *Nature*. 2007, 450, 1201-U4.

- [15] Shahid, M., Sobia, E., Singh, A., Malik, A., Khan, H.M., Jonas, D., et al. Beta-lactams and Beta-lactamase-inhibitors in current- or potential-clinical practice: A comprehensive update. *Critical Reviews in Microbiology*. 2009, 35, 81-108.
- [16] Patel, S., Keller, B.A., Fisher, L.M. Mutations at arg486 and glu571 in human topoisomerase II alpha confer resistance to amsacrine: Relevance for antitumor drug resistance in human cells. *Molecular Pharmacology*. 2000, 57, 784-91.
- [17] Wu, C.C., Li, T.K., Farh, L., Lin, L.Y., Lin, T.S., Yu, Y.J., et al. Structural Basis of Type II Topoisomerase Inhibition by the Anticancer Drug Etoposide. *Science*. 2011, 333, 459-62.
- [18] Wendorff, T.J., Schmidt, B.H., Heslop, P., Austin, C.A., Berger, J.M. The Structure of DNA-Bound Human Topoisomerase II Alpha: Conformational Mechanisms for Coordinating Inter-Subunit Interactions with DNA Cleavage. *Journal of Molecular Biology*. 2012, 424, 109-24.
- [19] Cecylia S. Lupala, P.G.-G.a.J.J.P. Molecular Determinants of the Bacterial Resistance to Fluoroquinolones:A Computational Study. *Current Computer-Aided Drug Design*. 2013, 9, 281-8.
- [20] Pan, P., Li, Y., Yu, H., Sun, H., Hou, T. Molecular principle of topotecan resistance by topoisomerase I mutations through molecular modeling approaches. *Journal of chemical information and modeling*. 2013, 53, 997-1006.
- [21] Beard, H., Cholleti, A., Pearlman, D., Sherman, W., Loving, K.A. Applying physics-based scoring to calculate free energies of binding for single amino acid mutations in protein-protein complexes. *PloS one*. 2013, 8, e82849.
- [22] Jacobson, M.P., Pincus, D.L., Rapp, C.S., Day, T.J.F., Honig, B., Shaw, D.E., et al. A hierarchical approach to all-atom protein loop prediction. *Proteins-Structure Function and Bioinformatics*. 2004, 55, 351-67.
- [23] Jacobson, M.P., Friesner, R.A., Xiang, Z., Honig, B. On the Role of the Crystal Environment in Determining Protein Side-chain Conformations. *Journal of Molecular Biology*. 2002, 320, 597-608.
- [24] Magrane, M., Consortium, U. UniProt Knowledgebase: a hub of integrated protein data. *Database : the journal of biological databases and curation*. 2011, 2011, bar009.
- [25] Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., et al. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry*. 2006, 49, 6177-96.
- [26] Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., et al. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry*. 2004, 47, 1739-49.
- [27] Mark, P., Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *Journal of Physical Chemistry A*. 2001, 105, 9954-60.
- [28] Jorgensen, W.L., Maxwell, D.S., TiradoRives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*. 1996, 118, 11225-36.
- [29] Ikeguchi, M. Partial rigid-body dynamics in NPT, NPAT and NP gamma T ensembles for proteins and membranes. *Journal of Computational Chemistry*. 2004, 25, 529-41.
- [30] Bailey, A.G., Lowe, C.P. MILCH SHAKE: An Efficient Method for Constraint Dynamics Applied to Alkanes. *Journal of Computational Chemistry*. 2009, 30, 2485-93.
- [31] Shan, Y.B., Klepeis, J.L., Eastwood, M.P., Dror, R.O., Shaw, D.E. Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. *Journal of Chemical Physics*. 2005, 122.
- [32] Stuart, S.J., Zhou, R.H., Berne, B.J. Molecular dynamics with multiple time scales: The selection of efficient reference system propagators. *Journal of Chemical Physics*. 1996, 105, 1426-36.
- [33] Kevin J. Bowers, E.C., Huafeng Xu, Ron O. Dror, Michael P. Eastwood,, Brent A. Gregersen, J.L.K., Istvan Kolossvary, Mark A. Moraes, Federico D. Sacerdoti,, John K. Salmon, Y.S., David E. Shaw. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. 2006.
- [34] Li, J., Abel, R., Zhu, K., Cao, Y., Zhao, S., Friesner, R.A. The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling. *Proteins-Structure Function and Bioinformatics*. 2011, 79, 2794-812.

- [35] Greenidge, P.A., Kramer, C., Mozziconacci, J.C., Wolf, R.M. MM/GBSA Binding Energy Prediction on the PDBbind Data Set: Successes, Failures, and Directions for Further Improvement. *Journal of chemical information and modeling*. 2013, 53, 201-9.
- [36] Sun, H.Y., Li, Y.Y., Tian, S., Xu, L., Hou, T.J. Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. *Physical Chemistry Chemical Physics*. 2014, 16, 16719-29.
- [37] Kongsted, J., Soderhjelm, P., Ryde, U. How accurate are continuum solvation models for drug-like molecules? *Journal of Computer-Aided Molecular Design*. 2009, 23, 395-409.
- [38] Kollman, P., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., et al. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* 2000, 33, 889-97.
- [39] Hou, T., Wang, J., Li, Y., Wang, W. Assessing the Performance of the Molecular Mechanics/Poisson Boltzmann Surface Area and Molecular Mechanics/Generalized Born Surface Area Methods. II. The Accuracy of Ranking Poses Generated from Docking. *J. Comput. Chem.* 2010, 32, 866-77.
- [40] Hou, T., Wang, J., Li, Y., Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 2011, 51, 69-82.
- [41] Xu, L., Sun, H., Li, Y., Wang, J., Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 3. The Impact of Force Fields and Ligand Charge Models. *J. Phys. Chem. B*. 2013, 117, 8408-21.
- [42] Sun, H.Y., Li, Y.Y., Shen, M.Y., Tian, S., Xu, L., Pan, P.C., et al. Assessing the performance of MM/PBSA and MM/GBSA methods. 5. Improved docking performance using high solute dielectric constant MM/GBSA and MM/PBSA rescoring. *Physical Chemistry Chemical Physics*. 2014, 16, 22035-45.
- [43] Michael Hinds, K.D., Janice Mayes, Elizabeth Altschuler, Ruud Jansen, Fred D. Ledley, and Leonard A. Zwelling. Identification of a Point Mutation in the Topoisomerase II Gene from a Human Leukemia Cell Line Containing an Amsacrine-resistant Form of Topoisomerase II 1. *Cancer Research*. 1991, 51, 4729-31.