Dual-Target Anticancer Drug Candidates: Rational Design and Simulation Studies

F. Keshavarz[a] and D. Mohammad-Aghaie[b],*

[a]*Department of Chemistry, College of Sciences, Shiraz University, Shiraz, 71454, Iran*
[b]*Department of Chemistry, Shiraz University of Technology, Shiraz, 71555-313, Iran*

This study aims to design some dual-target anticancer candidates, capable to act as an alkylating agent as well as a thymidylate synthase (TS) inhibitor. The designed scaffold is a combination of nucleobase, amino acid and aziridine structures. The candidates are docked into TS and three DNA double strand structures and evaluated based on their binding interaction energies and ligand efficiencies, compared to several reference drugs. The ADME properties of the alkylating agents are also predicted. The designed ligands exhibit improved interaction energies and lower ligand efficiencies with respect to the reference drugs. Among the ligands, **L4** is the best DNA binding agent and **L2**, **L5** and **L6** are the best TS inhibitors. In addition, the thioTEPA based Ser and Met analogues are the strongest and poorest alkylating agents, respectively. Further, molecular dynamics simulations on the best ligand-target complexes, *i.e.* **L4**-4AWL and **L6**-TS systems, provide evidences for the potential **L6** and **L4** anti-proliferation activities.

INTRODUCTION

Cancer cells are associated with uncontrolled proliferation, invasiveness and metastases development. Concerning these characteristics, different types of anticancer drugs have emerged which are classified as: alkylating agents, cytotoxic antibiotics, antimetabolites, microtubule inhibitors, monoclonal antibodies, and steroid hormones and their antagonists [1]. Among the mentioned categories, alkylating agents and antimetabolites are the most commonly utilized drugs in cancer chemotherapies.

Alkylation of DNA is one of the simplest methods to accomplish interfering with replication that is the major mechanism by which most of the clinically relevant anti-cancer agents kill tumor cells [2]. They mainly complete their task by forming cytotoxic lesions through DNA interstrand crosslinks [3]. In fact, such agents impose their cytotoxic and antitumor effects through their ability in binding covalently to DNA [4]. DNA alkylation takes place

in a two step process; the first showing modest regioselectivity with almost no sequence selectivity while the second has modest sequence selectivity for mustards [3] plus very high regio- and sequence specificity for larger chiral DNA alkylating agents [5].

N,N',N''-triethylene thiophosphoramide, thioTEPA, and its major metabolite, *i.e.* TEPA (Fig. 1), are trifunctional alkylating agents which were developed in the 1950s. They contain a four-coordinated phosphorus atom and three active aziridine (ethylene imine) moieties [6]. ThioTEPA with a wide spectrum of antitumor activity was soon
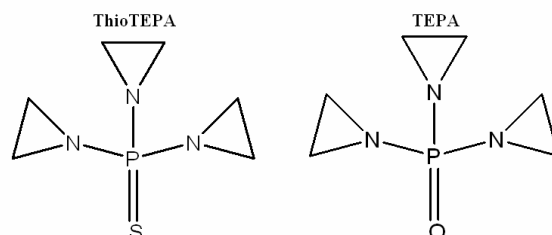

Fig. 1. Chemical structures of thioTEPA and TEPA.

*Corresponding author. E-mail: d_aghaie@sutech.ac.ir

recognized as an effective therapy for breast and ovarian cancers, and also solid tumors [7-10]. In conjunction with other alkylating agents, thioTEPA has experienced renewed interest due to being among the most effective anticancer drugs in high dose regimens [10,11].

Based on the viewpoint of the quantum theory of atoms in molecules (QTAIM), thioTEPA interacts mainly with DNA guanine base through non-covalent interactions that are responsible for the complex stability with hydrogen bonds of $NH\cdots N$ or $CH\cdots N$ types [12]. From the biological pathway perspective, the interaction of thioTEPA with DNA can follow two different pathways [10]: 1) thioTEPA can induce cell death by forming cross-links within DNA [13]. These cross-links cause the direct nucleophilic ring opening of two aziridinyl groups [14]; and 2) hydrolysis of thioTEPA to aziridine can happen due to a nucleophilic attack of water at the phosphorus atom and the N-P bond cleavage. In both of the pathways, the ring opening reaction of aziridine is initiated by its protonation, which makes aziridine the main target of nucleophilic attack [15]. Musser et al. have shown that thioTEPA leads to depurination and formation of aminoethyl adducts of guanine and adenine [16], so the nucleophilic attack should be advanced by guanine and adenine bases.

Antimetabolites are the other commonly utilized drug category in cancer chemotherapies. They exert their cytotoxic effects by blocking or subverting the pathways of DNA synthesis. Folate antagonists, often called antifolates, along with the purine and pyrimidine analogues are the main two classes of antimetabolites that respectively interfere with nucleotide synthesis and incorporate into DNA to influence the cell cycle [1].

Thymidylate synthase (TS) employs the co-factor of 5,10-methylene-5,6,7,8-tetrahydrofolate (mTHF) as both a methylene donor and a reductant to catalyze the methylation of deoxyuridine-5′-monophosphate (dUMP) to thymidine-5′-monophosphate (thymidine; one of the four nucleotides required for DNA synthesis) [17,18]. Hence, it has been considered as an important target for the development of antimetabolites [19,20]. TS related agents can toxify rapidly dividing cells [19] through inhibition or forming covalent complex with TS, and incorporation into DNA and/or RNA [21].

TS related antimetabolites impose different properties concerning folate transport, polyglutamylation and polyglutamylate accumulation and resistance patterns [20]. For example, the newest TS inhibitors are peptides like LSCQLYQR and its analogues that specifically target the monomer-monomer interface of TS to stabilize its inactive form [22]. However, the common feature in TS inhibition is leading directly to depletion of dTMP, thymidine monophosphate, and subsequently of dTTP, deoxythymidine triphosphate, and indirectly to an accumulation of dUMP, uracil monophosphate, that will end to dUTP, deoxythymidine triphosphate, incorporation into DNA due to lack of the natural substrate *i.e.* dTTP [23].

This study aims to design novel anticancer drug candidates able to act as an alkylating agent as well as a TS inhibitor. In this respect, some amino acid based entities were designed initially by inspiring from the structure and mechanism of thioTEPA, and their ADME properties were analyzed. The best designed alkylating agents along with an insight into the TS active site were later considered to propose chemically modified peptides with dual-targeting activities. This strategy has been shown to be successful in anti-proliferation activity against MCF-7 cell lines by amino acid based Schiff bases [24] and has the following advantages: minimizing drug resistance by combining drugs [1]; minimizing drug-drug interactions, offering lower drug accumulation in tissues, lower toxicity and biological diversity as a peptide drug [25], reducing risk of immunogenic effects and enhancing peptide stability as a chemically modified peptide [26]. The next step involved evaluation of the proposed structures by docking study of their binding to TS and three different DNA sequences equilibrated through molecular dynamics (MD) simulation under physiological conditions. Finally, ligand-target MD simulations were performed on the best TS inhibitor and DNA binding ligand to analyze their functions in a dynamic cell-like system.

METHODS

Preparation of DNA Structures

Three DNA structures taken from protein data bank (PDB) [27] were chosen as the target molecules to investigate binding of the ligands. It was tried to choose DNA structures that regulate different roles in gene transcription.

Accordingly, 4HC9 pdb entry code refering to DNA double strand structure bound to a GATA transcription factor [28], 4AWL [29] corresponding to a complex of DNA with NF-Y transcription factor and 3US0 [30] being the structure of p63 DNA binding domain in complex with an A/T rich DNA were selected. Table 1 demonstrates the list of mentioned structures with their chain lengths and sequences.

The selected DNA structures were submitted to GROMACS 4.5.0 simulation package [31] to obtain the equilibrated structures under conditions similar to that of the biological environment. As the first step, the required topology and interaction parameters were created by employing the AmberGS force field [32]. The initial simulation structures were generated by immersing DNA in a cubic box containing extended simple point charge (ESPC) water molecules [33]. Due to the presence of phosphorylated backbones, DNA structures carry negative charges. Considering the following characteristics, zinc ions were used to neutralize the systems. Zinc is the most abundant trace intracellular element which plays important roles in genetic stability and function [34]. It is considerably present in the cell nucleus [35], and has a significant impact on DNA replication, transcription and repair [36,37]. In addition, many resolved crystallographic DNA structures include zinc cofactors. The negative net charges of DNA structures in 4HC9, 4AWL and 3US0 systems were neutralized by addition of 19, 24 and 21 Zn^{2+} counter ions, respectively. The resulting systems were composed of a

DNA molecule, zinc ions and 41074, 49319 and 40261 explicit water molecules, respectively for 4HC9, 4AWL and 3US0. In the next step, the systems were energy minimized with 2000 steepest descent steps, while the cut-off values for van der Waals and short-range Coulomb forces were set to 14 and 9 Å, respectively. Thereafter, the systems experienced 40 ps of position restraint simulation to allow gradual relaxation of the initially used harmonic restraints on the macromolecular atoms while the solvent molecules were free to move. Finally, the molecular dynamics simulations were performed on the systems for 10 ns in the isobaric-isothermal (conserved NPT) ensemble [38]. Temperature and pressure were kept constant at 310 K and 1.0 bar by coupling the systems to the external baths with Berendsen thermostat and barostat, respectively [39]. The initial atomic velocities were taken from Maxwellian distribution [40,41] and where the required numerical integrations were done by the velocity Verlet algorithm [42]. During the course of simulations, the intermolecular (non-bonded) potentials were represented as sum of Lennard-Jones (LJ) forces and pairwise Coulomb interactions. The long-range electrostatic forces were computed using the particle mesh Ewald (PME) summation method [43,44]. Full periodic boundary conditions were used and the equations of motion were integrated by the means of leap-frog algorithm [45]. To further analyze the stability and equilibrium state of systems, evolution of the root mean square deviation (RMSD) of each DNA was plotted based on the trajectories produced during MD

Table 1. DNA pdb Entry Codes, Sequences and Lengths as Reported in Protein Data Bank

PDB code	Sequence	Length
4HC9	TTCCTAAATCAGAGATAACC	20
	AAGGTTATCTCTGATTTAGG	
4AWL	TTCTGAGCCAATCACCGAGCTCGAT	25
	ATCGAGCTCGGTGATTGGCTCAGAA	
3US0	AAACATGTTTATAAACATGTTT	22
	AAACATGTTTATAAACATGTTT	

simulations.

Preparation of the Thymidylate Synthase Structure

The initial structure of TS was retrieved from 1HVY [46] pdb entry code, chain A. This structure is in fact the structure of human thymidylate synthase, complexed with dUMP and Raltitrexed, an antifolate drug, in the closed conformation. In order to obtain the equilibrated TS structure in a cell-like environment, a MD simulation procedure, similar to that explained in the previous section, was employed. The GROMOS96 43a1 force field [47], as the most commonly used force field for proteins, was applied to the system. The TS structure, immersed in a box of 24751 ESPC water molecules, was neutralized by adding two Na^+ counter ions, since sodium cations are the most commonly used ions in protein simulations. Again, the RMSD values were plotted for further equilibrium analysis. It is noteworthy that both GROMOS96 and AMBER force fields have the same form for valence angles meanwhile they follow, respectively, $kw^2/2$ and $k(1-cos2w)$ forms for improper dihedral angles, in which w is the dihedral angle. In both applied force fields, non-bonded interactions are represented by LJ and Columbic potentials while nucleobase related parameters are not available in GROMOS force fields.

Reference Drugs

ThioTEPA and TEPA were used as the reference ligands for evaluating the activity of the designed amino acid based alkylating agents. Also, the binding results of CB3717, ZD9331, Raltitrexed, AG331, Pemetrexed, AG337, GS7904L and Capecittabine [21] were compared to the final dual-target drugs in analyzing the TS inhibitory efficiency. Please, refer to the outlined reference for their chemical structures.

Ligand Preparation

Structures of the designed and reference agents were generated by HyperChem Professional 7.0 [48] and were subjected to energy minimization by DFT calculations. The optimization was operated at the B3LYP/6-311G [49] level using Gaussian 03 [50]. The B3LYP method uses Becke's three-parameter hybrid exchange functional [51] and the correlation functional of Lee, Yang, and Parr and contains

both local and nonlocal terms [52].

Molecular Docking Simulation

Molecular docking studies were carried out by Molegro Virtual Docker (MVD) [53]. Having the three-dimensional structures of the receptor and the ligands, MVD determines the most likely conformation of the ligand-macromolecule complex. It performs flexible ligand and target side chain docking to specify the optimal geometry of each ligand during the docking simulation where the Moldock along with the Rerank scores [53] are generated by MVD for each different docking pose to give the best pose with the highest scores as the final docked configuration. The equilibrated DNA and TS structures at 310 K and the optimized ligands were submitted to MVD, allowing the program to assign bonds, bond orders, explicit hydrogen atoms, charges and flexible torsions if they were missing. The potent binding sites, also referred to as cavities, were distinguished using an algorithm implanted in the MVD software. Then, spherical grid of 0.3 Å resolution was centered at the major and minor grooves of the DNA molecule or the main active site of TS predetermined from 1HVY crystal structure. The Moldock SE search algorithm was used while setting the number of searching runs to 10. The other adjusted docking parameters included: 2000 max iterations, population size of 50, and an energy threshold of 100. Consequently, multiple poses, resembling different potential binding modes of each ligand were obtained, while at each step, MVD algorithm selected the conformation producing the lowest binding energy. After docking simulation, resulted poses were sorted based on their evaluated interaction energies and ligand efficiencies (LE1 and LE3). In addition, it was a priority to choose poses showing molecular orientations and binding sites, similar to 5,10-methylenetetrahydrofolate (D16) and UMP for TS inhibitors. In the case of the DNA structure, poses with aziridine rings oriented towards Adenine or Guanine bases were preferred. In MVD, LE1 and LE3 correspond to the Moldock and Rerank scores divided by heavy atoms count, respectively, where Moldock and Rerank scores are two-score functions of MVD that are computed during the docking iterations, by the program.

Ligand-Target Molecular Dynamics Simulation

L4-DNA MD simulation. L4 binding to 4AWL was analyzed as the most efficient ligand-DNA complex system.

The B3LYP optimized **L4** structure was assigned with Tripos charges by HyperChem Professional and imported to AmberTools 13 [54] to generate its General Amber Force Field [55] parameters and topologies. The missing parameters were assigned, manually. Then, the AnteChamber PYthon Parser interfacE (ACPYPE) code [56] was used to provide GROMACS compatible topology and coordinate files. The coordinates and AMBER topology parameters were incorporated in the GROMACS coordinate and topology files referring to the DNA double strand structure as directly taken from the 4AWL PDB structure. The simulation procedure of section "Preparation of DNA structures" was repeated for the DNA-**L4** system except that the final MD step was performed for 20 ns.

L6-TS MD simulation. L6 was recognized as the best TS inhibitor during docking simulations. Topology parameters of **L6** for GROMOS force field and its GROMACS coordinates were generated by the Dundee PRODRG2.5 server [57] according to the **L6** energy minimized structure at the B3LYP/6-311G level. The obtained data were embedded to the topology parameters and coordinates of TS, retrieved directly from 1HVY PDB structure. All the simulation processes explained in "Preparation of the thymidylate synthase structure" section, was repeated for the **L6**-TS system. However, the MD step were prolonged for 20 ns.

ADME Prediction

Absorption, distribution, metabolism, and excretion (ADME) properties were estimated for the designed alkylating agents, thioTEPA and TEPA, using the acd/i-Lab software (version 8.14 for Solaris; ACD Labs, Toronto, Canada). Since the major constituents of the designed dual-target drugs are amino acids and nucleobases, ADME predictions were focused on the alkylating amino acid analogues. The calculated properties consisted of Lipinski type, solubility, bioavailability, blood-brain distribution and plasma binding parameters. The properties were evaluated using the optimized ligand structures.

RESULTS AND DISCUSSION

Structural Stability of the Receptors

Conformations of the DNA strands and TS, in a water medium, were obtained by performing 10 ns of MD simulations. The resulting structures provided DNA and TS conformations in a cell-like environment. To investigate the stability of the final simulation structures, RMSD values were examined as a function of time and plotted with respect to the initial structures. Analysis of the RMSD values, Fig. 2, shows that the DNAs and TS reach a quasi-equilibrium state after about 6 ns. However, the
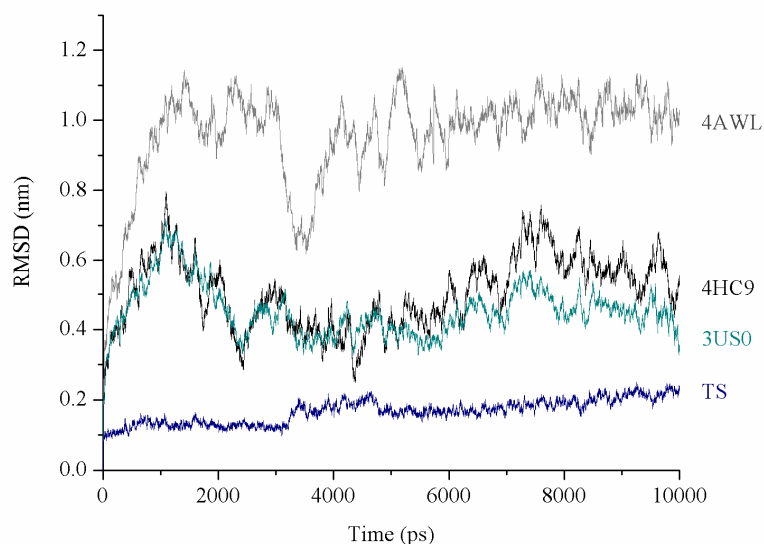

Fig. 2. Evolution of the backbone RMSD of the structures with time.

considerable magnitude of fluctuations around the corresponding average RMSD values for the DNA molecules indicates that these systems do not show high stabilities. The higher RMSD values for 4AWL correspond to the highest deviation from its starting structure.

The RMSD values for TS do not deviate significantly from those of its initial state and continue to oscillate around a low value of 0.166 nm. This means that the TS structure resides in a high stability mode and its conformational changes remain almost constant with respect to the initial structure taken from protein data bank. Overall analysis confirms equilibration status for the final DNA and TS conformations which makes them suitable for the use in molecular docking simulations.

Alkylating Activity

Design of the alkylating agents. As explained in the "Introduction" section, the reactivity of thioTEPA depends on cross link formation with DNA which takes place by the nucleophilic opening of aziridine groups [13,15] regardless of the differences between the two possible pathways. Since it has been verified that presence of two aziridinyl arms is necessary in formation of the required cross links [14], it was decided to insert two aziridine moieties in some template structures. The main concern was to choose a proper template able to accept loading of two aziridine groups, in addition to having intrinsic abilities for binding to DNA molecules. Luscomble *et al.* have studied the interactions between protein residues and DNA and have specified arginine to display an affinity for guanine, glutamine for adenine and thymine, threonine for thymine, and phenylalanine, histidine and proline for adenine. Also, they have characterized cysteine to have a high propensity to contact the DNA backbone [60]. These outlined conclusions and their additional findings encouraged us to design several alkylating agents by utilizing amino acids as the carrier template for aziridinyl. Following this purpose, our designed agents are displayed in Fig. 3.
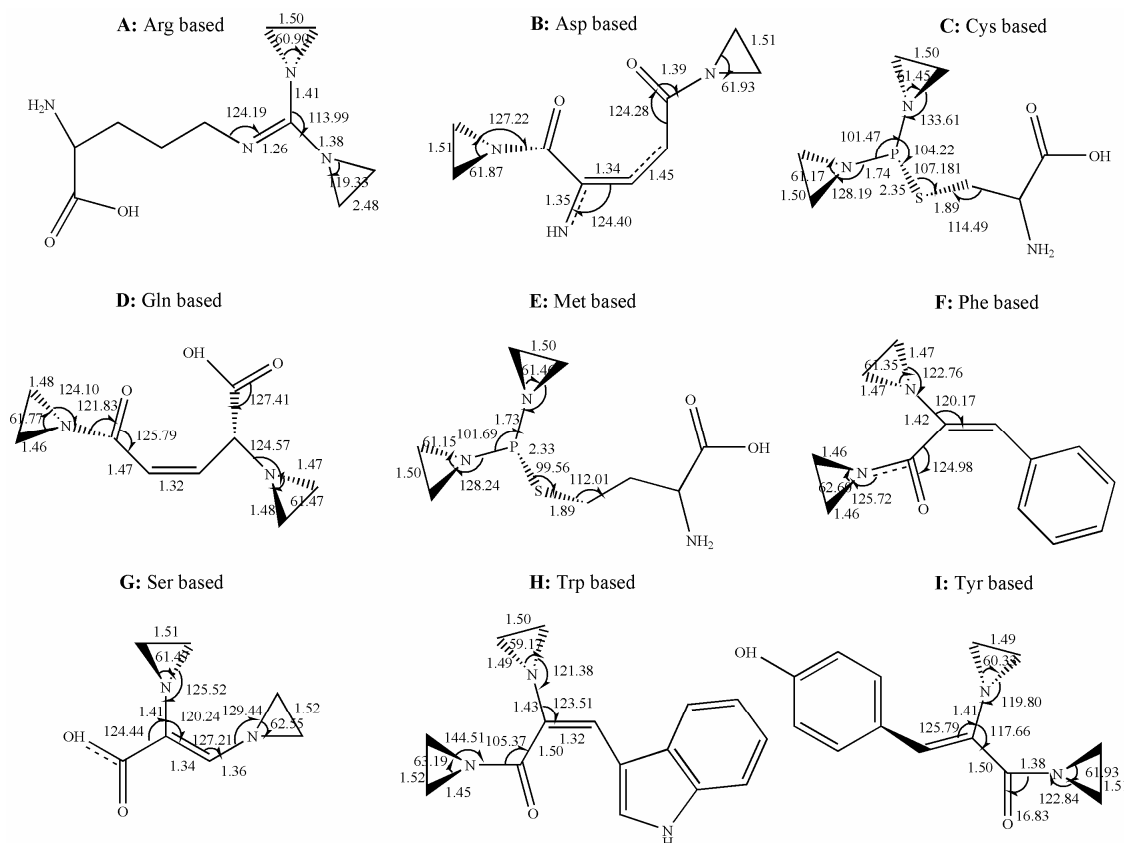

Fig. 3. Chemical structures of the designed alkylating agents.

Evaluation of the designed agents. Molecular docking simulations of the designed ligands were performed on the three equilibrated DNA structures. The simulations revealed no hydrogen bonding, short- or long-range electrostatic forces for any of the ligands and determined that the major contribution to their binding comes from steric interactions. Docking studies also helped to investigate the binding modes, interaction energies and ligand efficiencies of the designed and reference compounds where the results are summarized in Table 2.

Table 2. Docking Results for Binding of the Designed and Reference Alkylating Agents to the DNA Structures

	Ligand	MolDock score	Rerank score	Interaction energy	LE1	LE3
4HC9	Arg	-100.505	-57.2591	-107.022	-6.70036	-3.81728
	Asp	-97.1692	-59.993	-102.211	-7.47456	-4.61485
	Lys	-94.3932	-52.9405	-95.7655	-6.29288	-3.52937
	Gln	-100.227	-57.7524	-99.5136	-7.15907	-4.12517
	Met	-89.063	-47.4665	-101.065	-5.93753	-3.16443
	Phe	-99.0931	-57.7212	-110.06	-6.19332	-3.60757
	Ser	-83.498	-48.3013	-84.0001	-7.59073	-4.39102
	TEPA	-70.474	-37.4657	-71.7409	-6.40673	-3.40597
	ThioTEPA	-72.0762	-38.0982	-68.5931	-6.55238	-3.46347
	Trp	-117.059	-69.0898	-131.752	-6.16099	-3.6363
	Tyr	-98.5574	-51.5215	-124.781	-5.7975	-3.03068
4AWL	Arg	-112.744	-67.2603	-115.926	-7.51626	-4.48402
	Asp	-95.1014	-57.6495	-98.3416	-7.31549	-4.43458
	Lys	-112.15	-63.066	-117.667	-7.47664	-4.2044
	Gln	-100.636	-56.9941	-101.794	-7.18829	-4.07101
	Met	-104.212	-54.0853	-108.221	-6.94746	-3.60569
	Phe	-108.259	-60.8919	-110.447	-6.76617	-3.80574
	Ser	-92.6423	-53.5766	-93.766	-8.42203	-4.8706
	TEPA	-82.3632	-45.0488	-82.3246	-7.48756	-4.09535
	ThioTEPA	-77.6367	-40.2662	-76.9286	-7.05788	-3.66056
	Trp	-119.97	-69.4564	-130.628	-6.31423	-3.6556
	Tyr	-106.435	-61.22	-113.558	-6.26088	-3.60117
3US0	Arg	-125.373	-74.5271	-129.978	-8.35819	-4.96847
	Asp	-109.035	-66.3585	-112.514	-8.3873	-5.1045
	Lys	-115.776	-67.1525	-121.396	-7.7184	-4.47683
	Gln	-98.8953	-57.3883	-111.326	-7.06395	-4.09916
	Met	-108.647	-51.4404	-129.464	-7.24312	-3.42936
	Phe	-116.991	-69.5586	-131.842	-7.31197	-4.34741
	Ser	-94.1892	-56.1939	-98.7818	-8.56265	-5.10853
	TEPA	-72.6872	-38.7986	-74.0373	-6.60793	-3.52714
	ThioTEPA	-77.6919	-41.4916	-74.2922	-7.0629	-3.77196
	Trp	-136.877	-81.1394	-151.771	-7.20407	-4.27049
	Tyr	-118.926	-70.0011	-131.084	-6.99563	-4.11771

Since the predicted interaction energies are proportional to the number of ligand atoms, the best ligand should not be only inspected by the interaction energies. Ligand efficiency (LE) indices have been suggested as auxiliary criteria to assess quality of hits. These indices have distinct formulations whereas they are basically quantified in the form of binding affinity measures over physical molecular properties, *e.g.* number of the non-hydrogen atoms and total polar surface area [59]. In this respect the LE1 and LE3 values, in addition to the interaction energies, were plotted for the ligands, (see Fig. 4). It should be mentioned that lower LE value refers to higher ligand efficiency.

As it could be expected, the interaction energy of each ligand differs from one DNA to another. Even the order of interaction energies is not similar for the three targets. Among the ligands, the Trp based agent shows the highest affinity for all the sequences, while thioTEPA and TEPA show the lowest affinity scores. The molecular weight of Trp based agent is high and its structure is almost rigidified. Therefore, observation of lower interaction energy is quite logical. On the other hand, binding of both TEPA and thioTEPA structures is restricted by steric hindrance originating from the umbrella like geometry of the aziridine arms. Consequently, utilization of just two aziridine groups has enhanced the binding ability of the agents through reducing steric hindrance.

On the basis of LE1 and LE3 ligand efficiencies, the Ser based agent can bind to the three DNA double strands more efficiently though its interaction energies are not as noticeable as the Trp based ligand. Meantime, LE1 and LE3 ligand efficiencies for Tyr and Met based entities are the poorest, respectively.

Binding modes of the agents were also analyzed. Figure 5 displays superposition of the best conformations, *i.e.* producing the lowest interaction energies for the ligands docked to the DNA molecules. All the ligands have been bound to the intrastrand bases by the formation of cross links. It is exactly what an alkylating agent requires for exhibiting a successful alkylating activity. Ligands do not bind to the same bases at all DNA molecules. As an example, the binding modes of thioTEPA, the Ser and Met based ligands were analyzed. In the case of 3US0, thioTEPA binds to TA bases of one strand where its other aziridine arm interacts with the confronting pentose. The
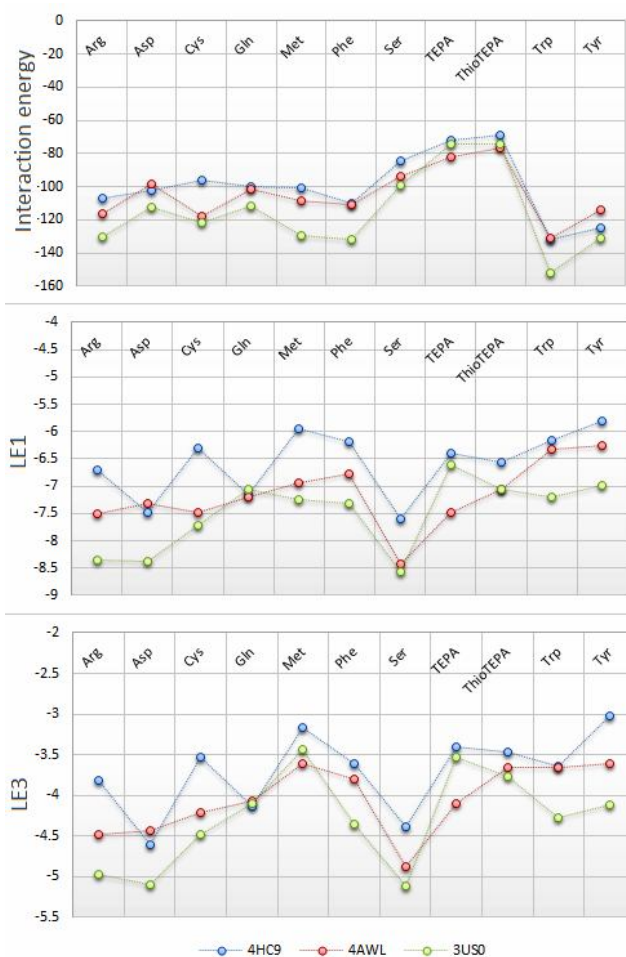

Fig. 4. Interaction energies, LE1 and LE3 ligand efficiencies of the alkylating agents.

Ser based agent crosses between one G on the first strand and one G of the other, while its carboxylic terminal orients against the DNA backbone. For the Met based agent, the two aziridine arms lay on one strand, but its terminal interacts with one A from the first strand and a G base from the other. For 4AWL, the conditions change and three ligands overlay each other between G of one strand and AC bases of the other. The next DNA, 4HC9, also involves with different binding sites; thioTEPA interacts with a G base *via* one of its aziridine moieties while its other parts are sandwiched between the DNA backbone of the two strands; carboxylic terminal of the Ser based agent interacts with T and the DNA backbones; and one aziridine of the Met based structure interacts with A where its other moieties almost
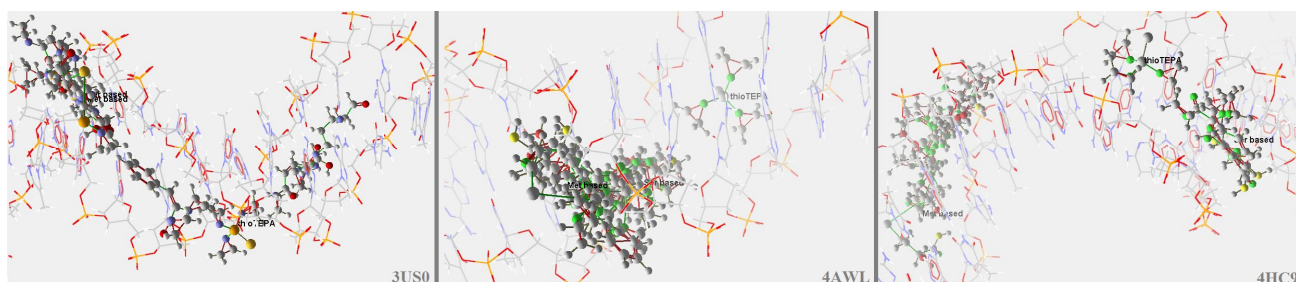
Fig. 5. Superimposition of the best docked conformation of each ligand on the DNA structures. The thioTEPA, Ser and Met based agents are labeled as the reference, best, sample and poor ligands, respectively.

Table 3. The Predicted ADME Properties of the Amino Acid Based Alkylating Agents

Ligand	MW	HBD	HBA	TPSA	RB	logSw	MPA	FUP	LogBB	logP
Arg	210.3	2	5	67.95	7	1.14	37%	0.97	-0.04	0.71
Asp	185.2	4	5	128.1	2	1.86	45%	0.99	-0.05	-2.65
Lys	255.3	6	5	77.41	7	1.13	26%	1	-2	-1.95
Gln	204.3	5	5	57.29	4	1.7	18%	0.99	-0.06	-0.64
Met	257.3	6	5	63.82	7	1.13	25%	1	-2	-1.9
Phe	234.4	1	3	57.29	3	-1.4	100%	0.26	0.13	-0.16
Ser	162.2	5	4	40.22	2	1.28	51%	0.98	-0.07	-1.44
TEPA	175.2	2	4	32.34	3	1.31	70%	0.99	-0.06	-1.19
ThioTEPA	190.2	0	3	57.55	3	1.26	34%	1	-0.04	-0.54
Trp	265.4	2	4	55.59	3	-1.92	78%	0.69	-0.59	-2
Tyr	250.4	2	4	77.52	3	-0.86	98%	0.46	-0.06	-1.19

collapse against the front strand.

ADME properties of the alkylating agents. The physically descriptors and pharmaceutically relevant properties of the designed agents and the reference alkylating compounds were analyzed to estimate drug-like properties of the molecules. The reported ADME predictions (Table 3) are comprised of:

1. Lipinski type descriptors: molecular weight (MW), number of hydrogen bond donors (HBD) and acceptors (HBA), and number of rotatable bonds (RB);
2. Total polar surface area (TPSA);
3. Solubility (logSw/mM);
4. Maximum passive absorption (MPA);
5. Fraction unbound in plasma (FUP);
6. Blood-brain distribution (logBB); and
7. Octanol/water partition coefficient (logP).

The molecules were also analyzed for chemical stability, passive absorption across intestinal barrier and active transport. All the structures, except the Met based molecule, possess chemical stability even at pH < 2. Passive absorption is poor (<30%) for Gln based structure but the intestinal absorption is good (>70%) for Phe and Tyr based entities and moderate for the rest of the structures. Also, no active transport is predicted for any of the ligands.

According to Table 3, all the structures follow Lipinski's rule (MW < 650, -2 < logP < 6, and logSw > -7) [60] by having a molecular weight of 160-260 and -2 < logP < 1 values. The exception is the Asp based structure

(logP = -2.65). The logSw values are above -7 for all of the ligands, too. LogBB values should be in the range of -3 - 1.2 to impose the ability to pass through the brain-blood barrier.

Therefore, all the structures can target brain tumors. The FUP values indicate that none of the structures can be transported by plasma proteins, sufficiently. Overall, the Ser and Arg based agents, used in the designed dual-target drugs in Table 4, satisfy every considered drug like properties while the poorest alkylating agent in docking studies (Met based) exhibits chemical instability.

Antimetabolite Function and Dual-Targeting

Design of the TS inhibitors. The 1HVY pdb structure was analyzed to gain enough insight into the TS active site and binding modes of its ligands. In this respect, PDBsum [61,62] was used and the data regarding TS enzymatic activity and LigPlot [63] of the 5,10-methylenetetrahydrofolate (D16) and UMP in their binding sites were obtained, see Fig. 6. Also, the binding modes were analyzed in detail by MVD to investigate all hydrophobic, electrostatic and hydrogen bonding features, corresponding to D16 and UMP that are represented in Fig. 7.

Amino acids, mostly Asp, or amino acid analogues exist as a structural part of the known TS inhibitors, *e.g.* Raltitrexed and ZD9331. Consequently, amino acids were considered to design the new ligands. Analyzing specific configurations of the ligands in TS binding site and the physiochemical properties of the involved residues, according to Figs. 6 and 7, resulted in the design of dual-target compounds. The resulting structures consist of a nucleobase (A or C; utilization of A and G bases had to be refused since alkylating agents target these bases), a spacer amino acid (aa), *i.e.* Gly, a hydrophobic amino acid (Trp which has some structural similarity with D16), a small flexible amino acid (Ala or Thr), a sulfur containing amino acid (Cys or Met, sharing similarity with thiophene ring of D16) and finally an alkylating amino acid analogue with aziridinyl rings. Based on the results reported in Table 2 and Fig. 4, the Ser based alkylating agent shows the best ligand efficiencies. Arg based ligand also endows considerable ligand efficiencies and multiple nitrogen atoms which is a common structural characteristic in all TS inhibitors. Therefore, Ser and Arg based alkylating agents were

Table 4. The Designed Dual-Target Ligands.

Ligand	Structure
L1	U-Gly-Trp-Ala-Met-Arg$_{aziridine}$
L2	U-Gly-Trp-Ala-Met-Ser$_{aziridine}$
L3	U-Gly-Trp-Ala-Cys-Arg$_{aziridine}$
L4	U-Gly-Trp-Ala-Cys-Ser$_{aziridine}$
L5	U-Gly-Trp-Thr-Met-Arg$_{aziridine}$
L6	U-Gly-Trp- Thr -Met-Ser$_{aziridine}$
L7	U-Gly-Trp- Thr -Cys-Arg$_{aziridine}$
L8	U-Gly-Trp- Thr -Cys-Ser$_{aziridine}$
L9	C -Gly-Trp-Ala-Met-Arg$_{aziridine}$
L10	C -Gly-Trp-Ala-Met-Ser$_{aziridine}$
L11	C -Gly-Trp-Ala-Cys-Arg$_{aziridine}$
L12	C -Gly-Trp-Ala-Cys-Ser$_{aziridine}$
L13	C -Gly-Trp-Thr-Met-Arg$_{aziridine}$
L14	C -Gly-Trp- Thr -Met-Ser$_{aziridine}$
L15	C -Gly-Trp- Thr -Cys-Arg$_{aziridine}$
L16	C-Gly-Trp- Thr -Cys-Ser$_{aziridine}$

selected in designing novel TS inhibitors. In this way, the designed ligands (Table 4) were as Base- aa$_{spacer}$- aa$_{hydrophobic}$- flexible aa$_{methyl}$- aa$_S$- aa$_{aziridine}$. The spacer amino acid is substituted at C_5 of the base.

Docking studies of the designed ligands as TS inhibitors. The designed dual-target ligands, reference TS inhibitors, thioTEPA, TEPA, Ser and Arg based agents were docked into the TS at physiological condition. The results regarding the best pose of each ligand are summarized in Table 5. No short- or long-range electrostatic interactions were predicted for any of the ligands. Meanwhile, noticeable hydrogen bonding interactions were observed for all of them and their binding interactions were found to be dominated by steric interactions. Therefore, the important role of hydrogen bonding and the dominancy of steric interactions of TS inhibitors were imitated by the designed ligands. All the proposed structures bind to the TS active site, which increases the possibility of antimetabolite activity.

The interaction energies of the designed dual-target entities, reference TS inhibitors, Ser and Arg based and reference alkylating agents are about -220 to -170, -170 to
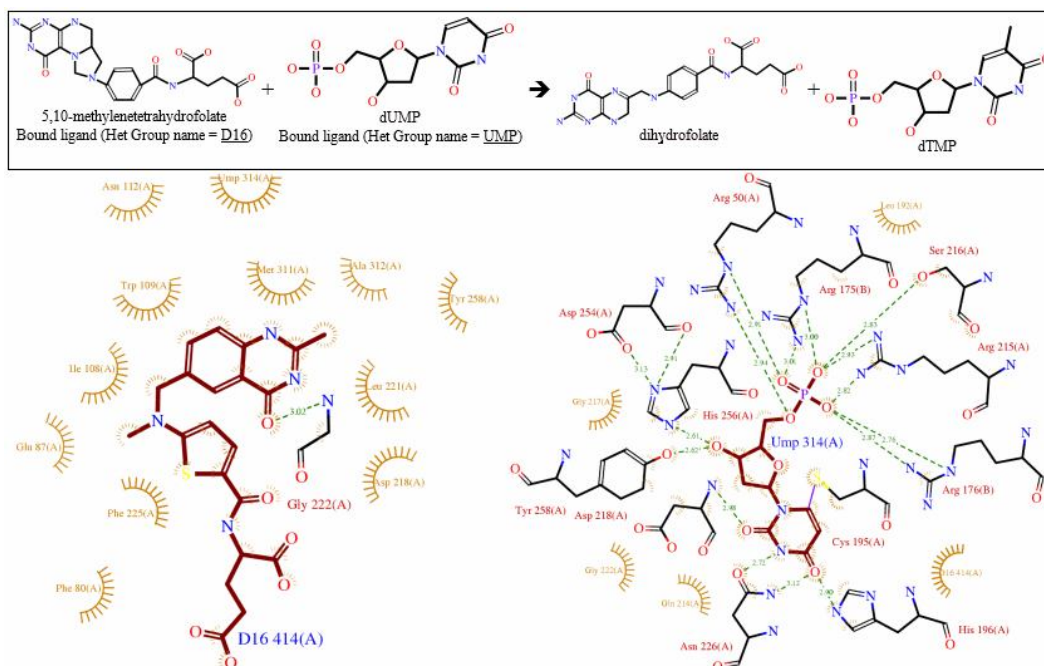
Fig. 6. Enzymatic pathway of thymidylate synthase and the binding sites for 5,10-methylenetetrahydrofolate (D16) and UMP as reported by PDBsum.
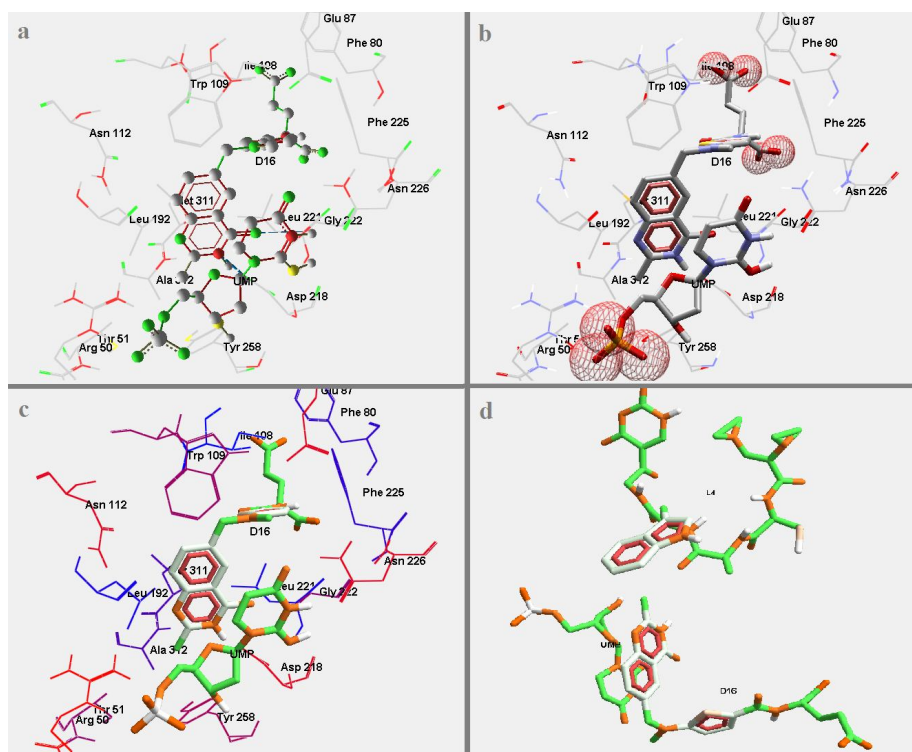

Fig. 7. Different views of D16 and UMP in TS active site; a) hydrogen bonding view with H-bonds shown by dashed blue lines, b) Electrostatic view, c) hydrophobic view as green and orange colors corresponding to hydrophilic and hydrophobic nature, respectively, and d) Hydrophobic view of D16 and UMP backbones as compared to **L4**, after docking to TS, as a sample designed compound.

Table 5. Docking Results for Binding of the Designed Dual-Target Ligands, Ser and Arg
Based Alkylating Agents, Reference Alkylating Agents and the TS Inhibitors to TS

Ligand	MolDock sScore	Rerank score	Interaction energy	HBond	LE1	LE3
L1	-162.193	-112.54	-197.936	-12.011	-3.00357	-2.08408
L2	-180.932	-140.986	-211.042	-8.60846	-3.69249	-2.87726
L3	-174.359	-128.764	-188.44	-5.98043	-3.35305	-2.47623
L4	-170.433	-23.643	-176.914	-12.7621	-3.62623	-0.50304
L5	-196.097	-158.131	-219.936	-17.2833	-3.50172	-2.82376
L6	-207.354	-140.343	-217.073	-8.35694	-4.06577	-2.75182
L7	-161.818	-70.4826	-189.061	-7.98869	-2.99662	-1.30523
L8	-167.923	-57.2951	-190.998	-14.5165	-3.427	-1.16929
L9	-171.087	-111.822	-183.431	-11.1577	-3.16827	-2.07077
L10	-181.857	-127.091	-178.071	-11.7321	-3.71137	-2.59369
L11	-154.874	-124.273	-173.734	-4.17506	-2.97835	-2.38987
L12	-167.083	-104.982	-163.144	-12.6151	-3.55497	-2.23365
L13	-158.719	-53.8146	-206.931	-19.9259	-2.83427	-0.96098
L14	-182.801	-84.0378	-184.581	-6.46084	-3.26431	-1.50067
L15	-160.381	-101.354	-181.996	-16.0103	-2.97002	-1.87693
L16	-171.582	-124.871	-193.474	-12.6699	-3.50167	-2.54839
AG337	-101.495	-72.9434	-117.224	0	-5.07473	-3.64717
CB3717	-136.486	-98.0372	-159.105	-2.5	-3.89959	-2.80106
Capecitabine	-107.305	-57.1794	-129.742	-3.4666	-4.2922	-2.28718
GS7904L	-147.595	-113.92	-165.071	-15.0813	-3.98907	-3.07891
Pemetrexed	-141.394	-61.539	-148.676	-14.5714	-4.56111	-1.98513
Raltitrexed	-154.83	-43.9814	-167.349	-16.0438	-4.83843	-1.37442
ZD9331	-140.172	-111.919	-166.347	-10.8324	-3.59417	-2.86973
Arg based	-85.1432	-65.4167	-94.1202	-3.34535	-5.67621	-4.36112
Ser based	-64.6512	-53.3396	-71.7824	-8.89791	-5.87738	-4.84906
thioTEPA	-70.4598	-52.4698	-69.4013	-4.99397	-6.40543	-4.76999
TEPA	-66.9629	-49.3453	-66.851	-3.8069	-6.08753	-4.48594

-120, -71, and -94 and -69, respectively. These values indicate the lowest interaction energies in favor of the designed structures. However, ligand efficiencies are decreased by introduction of the novel ligands so that LE1 scores reach to -4 to -2. The values of -5 to -3, -5.9 and -5.7, -6 and -6.4 are respectively resulted from binding of the reference TS inhibitors, Ser and Arg based and reference alkylating agents. The main reason, according to the docking results, might be due to the failure of the designed compounds to stretch inside the binding site, compared to UMP and D16. Please, refer to Fig. 7 for a graphical view.

Docking studies of the designed ligands as alkylating agents. Binding of the designed dual-target structures and reference alkylating agents to the three DNA double strands

were evaluated by docking simulations. The interaction between all the ligands and DNA structures are mediated by steric interaction, and neither electrostatic nor hydrogen bonding contributions. The results of docking scores, binding energies and ligand efficiencies are reported in Table 6. The interaction energies of the designed dual-target ligands are considerably greater than those in reference alkylating agents or those in amino acid based agents, which

Table 6. Docking Results for Binding of the Designed Dual-Target Ligands and the Reference Alkylating Agents to the DNA Structures

	Ligand	MolDock score	Rerank score	Interaction energy	LE1	LE3
4HC9	L1	-129.344	-60.0832	-154.391	-2.39526	-1.11265
	L2	-151.993	-76.0972	-161.243	-3.1019	-1.553
	L3	-129.327	-63.0568	-163.745	-2.48706	-1.21263
	L4	-178.29	-98.4474	-196.057	-3.7934	-2.09462
	L5	-169.379	-85.4142	-218.54	-3.02462	-1.52525
	L6	-186.407	-90.3739	-181.831	-3.65504	-1.77204
	L7	-156.615	-72.417	-169.01	-2.90028	-1.34106
	L8	-146.287	-70.0864	-152.143	-2.98545	-1.43033
	L9	-163.403	-88.4183	-199.451	-3.02599	-1.63738
	L10	-163.646	-80.7616	-181.734	-3.33971	-1.6482
	L11	-165.265	-87.5671	-202.124	-3.17817	-1.68398
	L12	-151.317	-76.2673	-167.652	-3.21952	-1.62271
	L13	-167.836	-69.9303	-157.157	-2.99706	-1.24875
	L14	-154.677	-71.2207	-162.439	-2.7621	-1.2718
	L15	-193.626	-93.9755	-200.793	-3.58568	-1.74029
	L16	-164.734	-79.5866	-162.852	-3.36193	-1.62422
	TEPA	-69.7591	-37.2267	-67.2187	-6.34174	-3.38425
	ThioTEPA	-73.263	-37.6099	-72.0919	-6.66027	-3.41908
4AWL	L1	-201.95	-102.829	-219.204	-3.73981	-1.90424
	L2	-166.137	-73.2311	-145.527	-3.39055	-1.49451
	L3	-187.266	-100.302	-225.317	-3.60127	-1.92889
	L4	-210.137	-117.205	-228.63	-4.471	-2.49372
	L5	-134.67	-58.0397	-130.01	-2.40482	-1.03642
	L6	-184.175	-90.461	-175.936	-3.61128	-1.77375
	L7	-178.247	-87.0248	-197.26	-3.30088	-1.61157
	L8	-176.298	-90.6598	-192.64	-3.59791	-1.8502
	L9	-165.156	-83.2802	-207.345	-3.05845	-1.54223
	L10	-119.929	-64.8362	-151.69	-2.44752	-1.32319
	L11	-184.576	-96.3873	-219.234	-3.54953	-1.8536
	L12	-167.494	-84.9703	-202.184	-3.56369	-1.80788
	L13	-123.776	-36.3746	-193.669	-2.21029	-0.64955
	L14	-169.003	-84.0992	-195.828	-3.01791	-1.50177
	L15	-146.636	-65.0993	-156.467	-2.71549	-1.20554
	L16	-198.136	-103.49	-222.349	-4.04359	-2.11203
	TEPA	-82.3556	-45.0193	-82.2297	-7.48687	-4.09266
	ThioTEPA	-76.5901	-39.9124	-75.3196	-6.96274	-3.6284

Table 6. Continued

	L1	-157.07	-80.5702	-170.802	-2.90871	-1.49204
	L2	-155.626	-76.3973	-163.288	-3.17604	-1.55913
	L3	-200.712	-101.437	-206.22	-3.85984	-1.9507
	L4	-187.138	-97.2997	-194.386	-3.98166	-2.07021
	L5	-135.856	-72.2628	-160.513	-2.426	-1.29041
	L6	-198.684	-100.299	-202.805	-3.89577	-1.96665
	L7	-172.714	-85.8711	-203.321	-3.19841	-1.59021
	L8	-152.839	-81.4396	-176.821	-3.11916	-1.66203
3US0	L9	-167.555	-90.0479	-181.508	-3.10287	-1.66755
	L10	-206.123	-103.529	-208.444	-4.20658	-2.11284
	L11	-182.227	-81.5881	-189.843	-3.50437	-1.569
	L12	-198.705	-105.889	-206.861	-4.22776	-2.25296
	L13	-163.954	-79.0635	-200.428	-2.92775	-1.41185
	L14	-164.497	-72.3136	-178.239	-2.93745	-1.29131
	L15	-192.693	-98.3004	-219.509	-3.56839	-1.82038
	L16	-126.931	-60.7344	-134.076	-2.59042	-1.23948
	TEPA	-76.1285	-40.905	-75.8129	-6.92078	-3.71864
	ThioTEPA	-74.0666	-39.3105	-70.9018	-6.73333	-3.57368
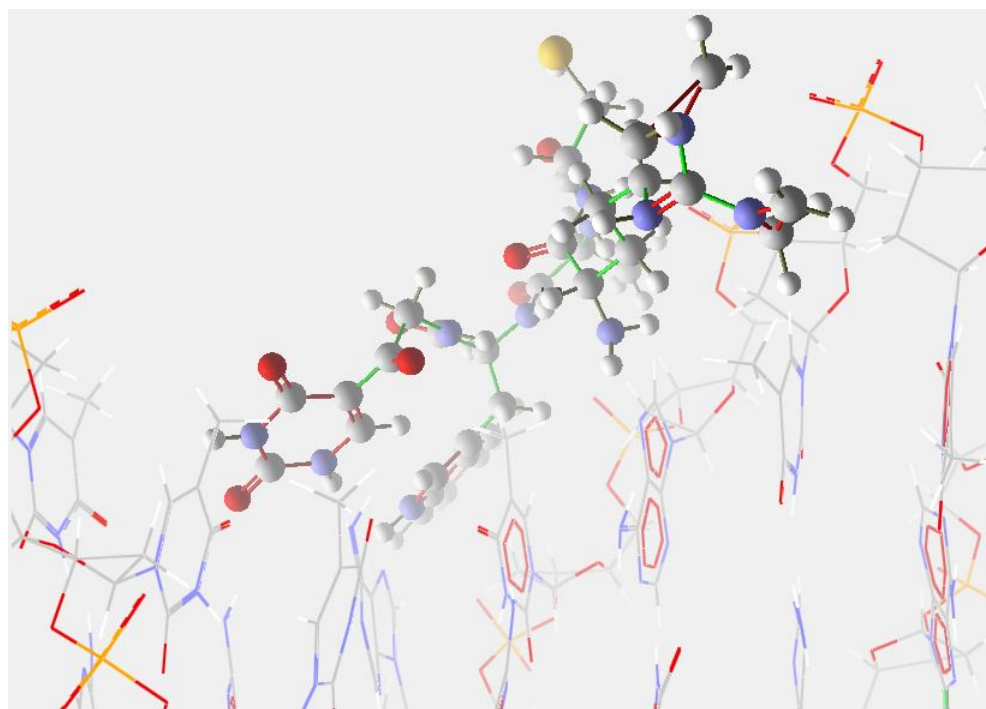
Fig. 8. Binding mode of **L3**-3US0 conformation; a) top and b) close view.

can be attributed to their higher molecular weight. However, the ligand efficiencies have been increased; from about -7 for the reference agents to about a minimum of -4.5 for the drug candidates.

The binding mode of the best entity **L3**-3US0 conformation is presented in Fig. 8 as a randomly selected example. Analysis of the binding modes implies that the designed ligands overlay on the phosphate backbone towards the nucleotide or on the intrastrand bases. The flexibility provided by the introduced ligands facilitates their maximum stretching on the DNA helical structure and positioning them in the major grooves of the DNA molecules. This condition enables the ligand to act as an alkylating agent if the aziridine moieties encounter A or G nucleobases or block the DNA grooves from being complexed with transcription factors or other fatal proteins in cell replication.

Finally, based on the reported results in Fig. 9, on the criteria of interaction energies and also LE1 and LE3 ligand efficiencies, **L6**, **L5** and **L2** were found to be the best TS inhibitors. On the other hand, according to Fig. 10, the interaction energies were approximately the same for all the ligands while **L4** is the best in DNA binding with regard to LE values (see Fig. 11 for the chemical structures of **L2**, **L4**, **L5** and **L6** as the best ligands).

Dual-Targeting Profile Verification

TS inhibition. L6 was recognized as the best TS binding ligand. MD simulation of **L6**-TS system, starting from the configuration shown in Fig. 12 revealed that **L6** rapidly binds to D16-UMP binding region after 0.5 ns. **L6** maintains stability in TS active site, during the 20 ns of the MD simulation. The final averaged configuration of **L6**-TS by GROMACS and the 1HVY structure were analyzed with MVD to determine the contributing forces in **L6** binding compared with D16 and UMP.

As illustrated in Figs. 6 and 7, D16 and UMP are involved in steric (hydrophobic and van der Waals), electrostatic and hydrogen bonding interactions towards Ala 312, Arg *50*, 78, 126, *215* and 271, Asn 112 and 226, Asp *48*, *49*, 110 and 218, Cys *195*, Gln 214, Glu 30, 86, 87 and 310, Gly 222, His 196, 256 and 261, Ile 108 and 307, Leu 192 and 221, Lys *47*, 77, 82, 93, 99, 104, 107, 266 and 308, Met 311, Phe 80 and 225, Pro *193*, Ser *216*, Thr 51, Trp 109
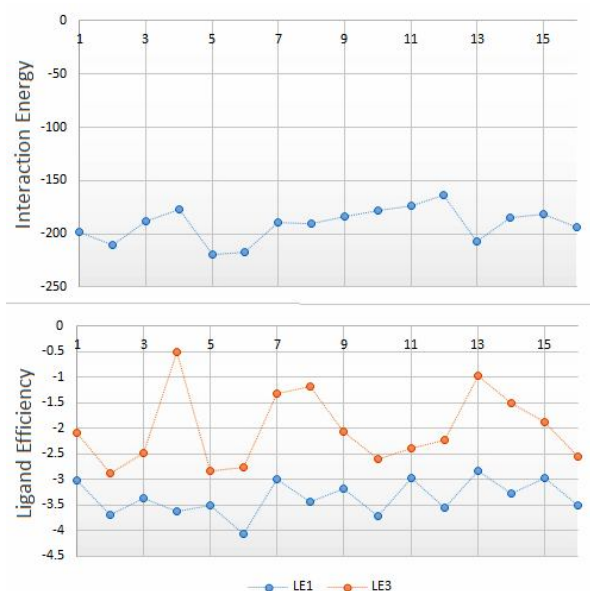

Fig. 9. Interaction energy and ligand efficiency plots for the designed ligands when docked into TS active site. The numbers on the vertical axis correspond to the entity number.
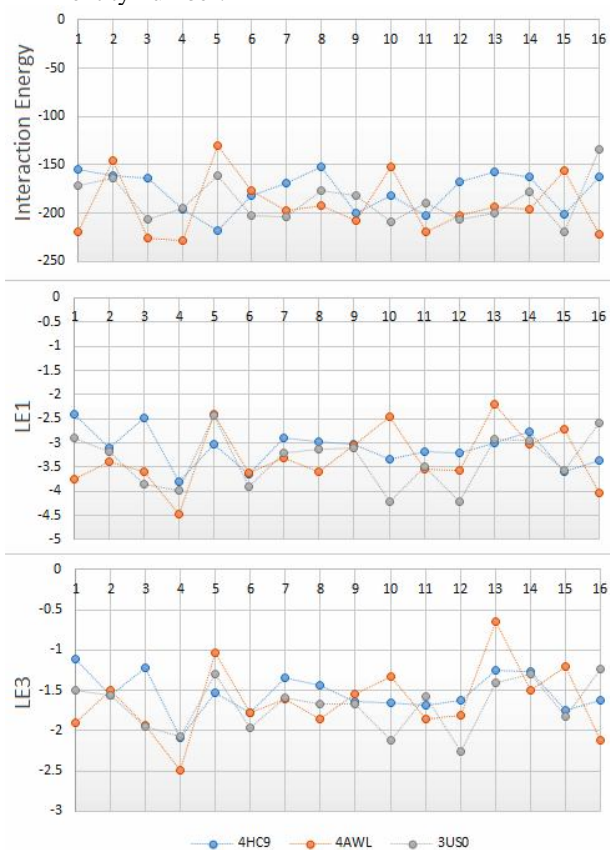

Fig. 10. Interaction energy and ligand efficiency plots of the ligands after docking to the DNA structures.
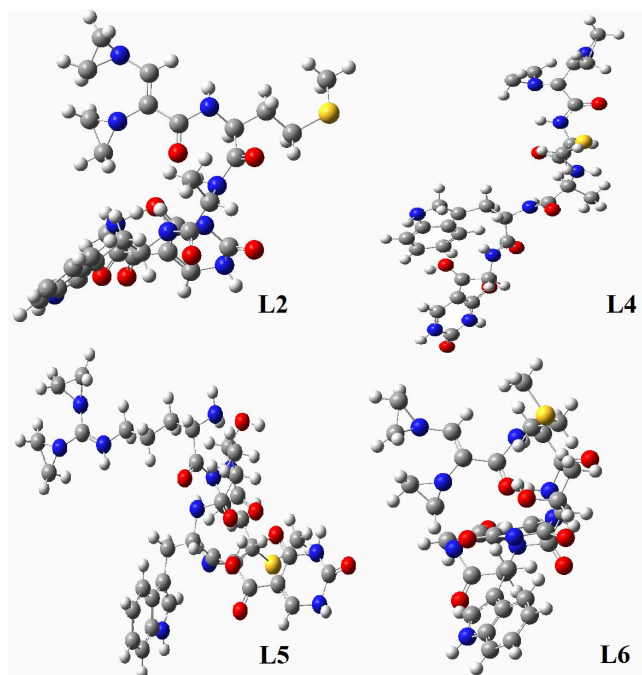
Fig. 11. Optimized structures of **L2**, **L4**, **L5** and **L6** at the B3LYP/6-311G level of theory.

and Tyr 135 and 258, and Val 223. On the other hand, **L6** is bound to Arg 50 and 215, Asp 48, 49, 119 and 254, Cys 195, Gly 52 and 122, Leu 121 and 187, Lys 47, Pro 193 and 194, Ser 120 and 216, Thr 51 and Trp 182. Binding site residues of **L6** and UMP are mainly common but **L6** also exposes force fields to D16 binding region. The ID of the

common residues is indicated in italic format.

Total energy, steric, hydrogen bonding, short- and long-range electrostatic interaction *unitless* values are respectively -115.22, -120.48, -3.51, -2.18 and -4.78 for D16, -102.64, -68.82, -27.09, -14.29 and -3.54 for UMP, and -75.21, -101.89, -5.40, 0.00 and 0.00 for **L6**. Therefore, the interaction energies of **L6** are comparable to D16 and UMP. Though **L6** does not conduct any electrostatic interaction, its hydrogen bonding potential is stronger than that of UMP. A significant feature about **L6** is that it provides noticeable steric hindrance that enables withdrawing UMP and D16 to interrupt the TS metabolic activities. Furthermore, since **L6** has proved to be resistant in its binding site for 20 ns, it can guarantee its anti-metabolic role in cancer cells by residing in the common D16 and UMP binding domain.

DNA alkylation. **L4**-4AWL system was simulated under cell-like condition, as the best ligand-DNA complex; refer to Fig. 13 for substantial MD simulation snapshots. Initially, **L4** and 4AWL are at large separation. Then, **L4** starts moving towards and backwards 4AWL, gradually. At 12.5 ns, it reaches to one terminal of 4AWL. It starts binding to 4AWL poorly at 14 ns by approaching the phosphate backbone *via* its aziridine arms while the U-Gly-Trp part of **L4** (UGT) almost suspends in water molecules. The aziridine moieties sustain their positions for about 2 ns. Meantime, UGT advances to form an intra-strand bridge at 16.5 ns (Depicted in Fig. 13). In this step, Trp and the aziridine arms are pointed at phosphate groups of opposite
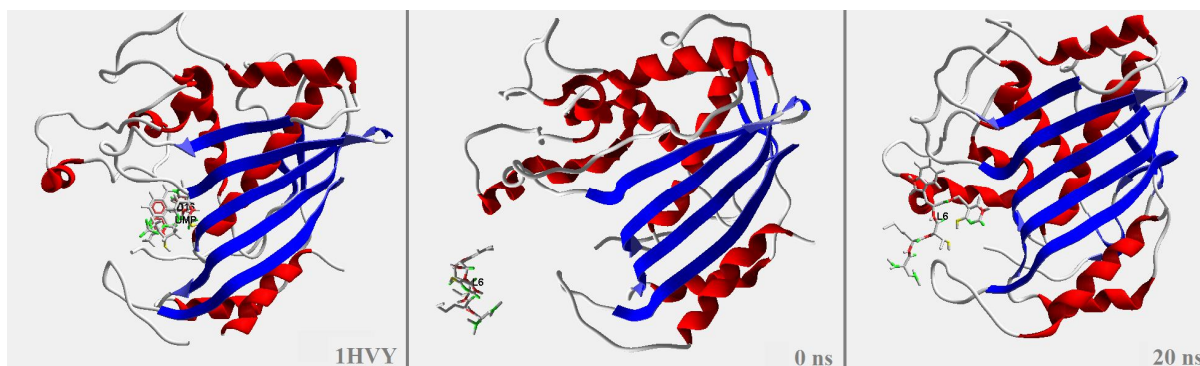

Fig. 12. D16 and UMP ligands in 1HVY PDB structure, the simulated 1HVY-**L6** system at 0 ns of the molecular dynamics simulation and the average structural configuration of 1HVY- **L6** after 20 ns.
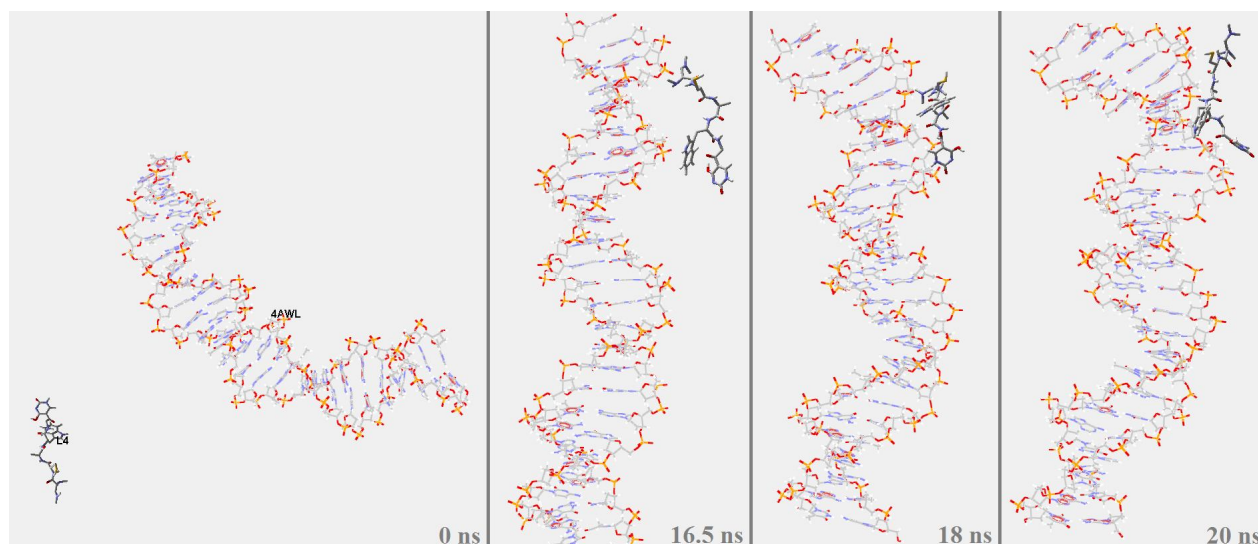
Fig. 13. Substantial snapshots of the 4AWL-**L4** system MD simulation from 0 to 20 ns. The ligand and 4AWL are colored based on their elements.

strands. At 18 ns, the aziridine functionalities seem to be oriented to the adenine and guanine rich region of 4AWL. However, π-π interaction, arisen from U and Trp interaction with the nucleobases stacks UGT on 4AWL and weakened aziridine-DNA interaction suspends the aziridine head among the solvent molecules. At 20 ns, **L4** completely overlays in the 4AWL minor groove, but the aziridine is not headed toward A or G nucleobases. Consequently, **L4** is able to bind to DNA and slide on the strands, to allocate itself at the best position. Meanwhile, the dominant van der Waals interactions do not permit the aziridine arms to orient towards the A/G nucleotides, properly. Since DNA-protein interactions are mainly driven by van der Waals potentials [56], the presence of **L4** can weaken the interactions driven by transcription proteins and the DNA binding enzymes.

CONCLUSIONS

In this study, a set of entities were introduced as dual-target anticancer drug candidates, capable to act as an alkylating agent as well as a TS inhibitor. These ligands are in fact chemically modified structures that are combinations of nucleobases, amino acids and aziridine rings, designed based on alkylating mechanism of thioTEPA and detailed

analysis of TS active site. The designed drugs were evaluated through docking into the equilibrated structures of TS and three DNA structures, in a cell like medium. Also, ADME properties of the alkylating agents were predicted. **L4** is introduced as the best DNA binding agent, in addition to **L2**, **L5** and **L6** as the best TS inhibitors. MD simulation on **L6**-TS system verified that **L6** binds to D16-UMP binding site with similar interaction energies and prevents D16 and UMP binding by eminent steric hindrance. Also, MD simulation on **L4**-4AWL system revealed that **L4** demands for a lag time to slide on the DNA double strand to find the most proper binding site and block the DNA minor groove. **L4** did not depict any alkylating activities in the meantime. It is also noteworthy that the larger size of the designed ligands has two main advantages: a) Even though ligand efficiencies are decreased, the blocking potency with an acceptable range of interaction energies would be a bonus in blocking TS active site and DNA grooves which ends up to anti-proliferation activity; and b) The vasculature in a growing tumor is leaky. So, as observed for drug delivery systems, the larger size of the drugs will make them tend to get trapped in the tumor and taken out of circulation, *i.e.* passive targeting that can enhance anti-tumor drug efficacies [64].

ACKNOWLEDGEMENTS

REFERENCES

[1]   I. Mullaney, An Introduction to Anticancer Drugs, in: Doggrell, Sheila (Ed.), Pharmacology in One Semester, 2012.

[2]   R.H. Hargreaves, J.A. Hartley, J. Butler, Front. in Biosci. 5 (2000) E172.

[3]   J.A. Hartley, C.C. Ohare, J. Baumgart, Brit. J. Cancer 79 (1999) 261.

[4]   K. Sanjiv, T.L. Su, S. Suman, R. Kakadiya, T.C. Lai, H.-Y. Wang, M. Hsiao, T.C. Lee, Int. J. Cancer 130 (2011) 1440.

[5]   W.A. Denny, Curr. Med. Chem. 8 (2001) 533.

[6]   S. Farber, R. Appleton, V. Downing, F. Heald, J. King, R. Toch, Cancer (Phila.) 6 (1953) 135.

[7]   E. van der Wall, J.H. Beijnen, S. Rodenhuis, Cancer Treat. Rev. 21 (1995) 105.

[8]   M.J. van Maanen, I.M. Tijhof, J.M.A. Damen, C. Versluis, J.J.K. van den Bosch, *et al.*, Cancer Res. 59 (1999) 4720.

[9]   M.J. van Maanen, K. Doesburg Smits, J.M.A. Damen, A.J.R. Heck, J.H. Beijnen, Int. J. Pharma. 200 (2000) 187.

[10]   M.V. Maanen, C.J. Smeets, J.H. Beijnen, Cancer Treat. Rev. 26 (2000) 257.

[11]   S.N. Wolff, R.H. Herzig, J.W. Fay, C.F. LeMaistre, R.A. Brown, D. Frei-Lahr, *et al.*, Sem. in Onco. 17 (1990) 2.

[12]   Z. Karagoz Genc, S. Selcuk, S. Sandal, N. Colak, S. Keser, M. Sekerci, M. Karatepe, Med. Chem. Res. 23 (2014) 2476.

[13]   N.A. Cohen, M.J. Egorin, S.W. Snyder, B. Ashar, B.E. Wietharn, *et al.*, Cancer Res. 5 (1991) 4360.

[14]   R.D. Gill, C. Cussac, R.L. Souhami, F. Laval, Cancer Res. 56 (1996) 3721.

[15]   J.A. Dunn, T.J. Bardos, Biochem. Pharma. 41 (1991) 885.

[16]   S.M. Musser, S. Pan, M.J. Egorin, D.J. Kyle, P.S. Callery, Chem. Res. Tox. 5 (1992) 95.

[17]   C.W. Carreras, D.V. Santi, Ann. Rev. in Biochem. 64 (1995) 689.

[18]   M. Friedkin, D. Roberts, J. Biol. Chem. 220 (1956) 653.

[19]   D.W. Begley, S. Zheng, G. Varani, Chem. Biol. Drug Des. 76 (2010) 218.

[20]   A.L. Jackman, A.H. Calvert, Ann. Onco. 6 (1995) 871.

[21]   E. Chu, M.A. Callender, M.P. Farrell, J.C. Schmitz, Cancer Chemo. Pharma. 52 (2003) 80.

[22]   M. Pelà, P. Saxena, R. Luciani, M. Santucci, S. Ferrari, G. Marverti, C. Marraccini, A. Martello, S. Pirondi, F. Genovese, S. Salvadori, D. D'Arca, G. Ponterini, M.P. Costi, R. Guerrini, J. Med. Chem. 57 (2014) 1355.

[23]   B.H.M. van Triest, P.G. Giaccone, G.J. Peters, Ann. Onco. 11 (2000) 385.

[24]   T.Y. Nikolaienko, L.A. Bulavin, L.F. Sukhodub, Mol. Inf. 33 (2014) 104.

[25]   Y. Yagi, K. Terada, T. Noma, K. Ikebukuro, K. Sode, BMC Bioinformatics 8 (2007) 11.

[26]   M. Kondoh, H. Uchida, T. Hanada, M. Hoshino, U.S. Patent Application 12 (2008) 576.

[27]   H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, Nucl. Acids Res. 28 (2000) 235. www.rcsb.org.

[28]   Y. Chen, D.L. Bates, R. Dey, P.-H. Chen, A.C. Dantas Machado, I.A. Laird-Offringa, R. Rohs, L. Chen, Cell Reps. 2 (2012) 1197.

[29]   M. Nardini, N. Gnesutta, G. Donati, R. Gatta, C. Forni, A. Fossati, at al., Cell 152 (2013) 132.

[30]   C. Chen, N. Gorlatova, O. Herzberg, J. Biol. Chem. 287 (2012) 7477.

[31]   S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, *et al*., Bioinformatics (2013). doi:10.1093/bioinformatics/btt055.

[32]   A.E. García, K.Y. Sanbonmatsu, Procc. Nat. Acad. Sci. 99 (2002) 2782.

[33]   H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunstetren, J. Hermans, Interaction models for water in relation to protein hydration, in: B. Pullman (Ed.), Intermolecular forces, Reidel: The Netherlands, 1981.

[34] I.E. Dreosti, Mut. Res. 475 (2001) 161.

[35] R.J. Cousins, Procc. Nut. Soc. 57 (1998) 307.

[36] E. Ho, J. Nut. Biochem. 15 (2004) 572.

[37] K.H. Falchuk, Mol. Cell Biochem. 188 (1998) 41.

[38] H.C. Andersen, J. Chem. Phys. 72 (1980) 2384.

[39] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, J. Chem. Phys. 81 (1984) 3684.

[40] K. Huang, Statistical mechanics, Wiley, New York, 1963.

[41] E.H. Kennard, Kinetic Theory of Gases, McGraw-Hill, New York, 1963.

[42] W.C. Swope, H.C. Andersen, P.H. Berens, K.R. Wilson, J. Chem. Phys. 76 (1982) 637.

[43] T. Darden, D. York, L. Pedersen, J. Chem. Phys. 98 (1993) 10089.

[44] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, J. Chem. Phys. 103 (1995) 8577.

[45] W.F. van Gunsteren, H.J.C. Berendsen, Mol. Sim. 1 (1988) 173.

[46] J. Phan, S. Koli, W. Minor, R.B. Dunlap, S.H. Berger, L. Lebioda, Biochemistry 40 (2001) 1897.

[47] W.F. van Gunsteren, S.R. Billeter, A.A. Eising, P.H. Hüenberger, P. Krüger, A.E. Mark, W.R.P. Scott, I.G. Tironi, Biomolecular Simulation: the GROMOS 96 Manual and User Guide, Switzerland, 1996.

[48] Hypercube Inc, HyperChem, Release 7.0 for windows, 2002.

[49] W.J. Hehre, R. Ditchfield, J.A. Pople, J. Chem. Phys. 56 (1972) 2257.

[50] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, *et al.*, Gaussian 03, revision C.02, Gaussian, Inc, Wallingford, CT, 2010.

[51] A.D. Becke, J. Chem. Phys. 98 (1993) 5648.

[52] C.T. Lee, W.T. Yang, R.G. Parr, Phys. Rev. B 37 (1988) 785.

[53] R. Thomsen, M.H. Christensen, J. Med. Chem. 4 (2006) 3315.

[54] J. Wang, W. Wang, P.A. Kollman, D.A. Case, J. Mol. Graph. & Mod. 25 (2006) 247.

[55] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, J. Comput. Chem. 25 (2004) 1157.

[56] A.W. Sousa da Silva, W.F. Vranken, E.D. Laue, BMC Res. Notes. 5 (2012) 367. http://acpype.googlecode.com, accessed June 2014.

[57] A.W. Schuttelkopf, D.M.F. Van Aalten, Acta Crysta. D Biol. Crysta. 60 (2004) 1355.

[58] N.M. Luscombe, R.A. Laskowski, J.M. Thornton, Nuc. Acids Res. 29 (2001) 2860.

[59] C. Abad-Zapatero, Exp. Opin. Drug Disco. 1 (2007) 469.

[60] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Adv. Drug Del. Rev. 64 (2012) 4.

[61] R.A. Laskowski, V.V. Chistyakov, J.M. Thornton, Nuc. Acids Res. 33 (2005) D266.

[62] M. Kondoh, H. Uchida, T. Hanada, M. Hoshino, U.S. Patent Application 12 (2008) 576.

[63] A.C. Wallace, R.A. Laskowski, J.M. Thornton, Prot. Eng. 8 (1996) 127.

[64] D. Peer, J.M. Karp, S. Hong, O.C. Farokhzad, R. Margalit, R. Langer, Nature Nanotech. 2 (2007) 751.