Indian Journal of Chemistry
Vol. 63, September 2024, pp. 882-893
DOI: 10.56042/ijc.v63i9.11122

Modelling DNA-ligand interactions through variable force field based MD simulations

Anwesh Pandey*[a] & Anupriya Adhikari[b]

[a] Department of Physics, Babasaheb Bhimrao Ambedkar University, Lucknow, Uttar Pradesh 226 025, India

[b] Department of Chemistry, Graphic Era Hill University, Dehradun, Uttarakhand 248 002, India

E-mail: apdapbbau@gmail.com

Received 13 May 2024; accepted (revised) 30 August 2024

In the field of MD simulations, force field refers to as the combination of a set of mathematical formulae and incorporated parameters that are used to evaluate the protein energy as a function of its atomic coordinates over the passage of time. And therefore, the choice of suitable force field to carry out MD simulation is of utmost importance. Apart from literature survey and previously reported results, there is no direct evidence regarding the choice of appropriate force field for a particular system. This entirely depends upon the type of system that is to be simulated and the parameters to be analysed. In the current research article, analysis of molecular dynamics trajectories for ligand-DNA interactions over two such force fields, AMBER03 and CHARMM27 have been done. The functional forms including the parameterization protocols of both the force fields are also studied and well discussed in this research article. It was found that energetics related to CHARMM27 force field are in better agreement with the RMSD of the complex than that of the AMBER03 force field and hence CHARMM27 force field can be a suitable choice to model DNA-ligand complexes for shorter simulations as the convergence criterions are met earlier with minimum or no deviations in the RMS plots.

Keywords: AMBER03, CHARMM27, Force field, Molecular dynamics

The double stranded Deoxyribonucleic acid (DNA) is a hereditary material that has long been one of the exciting molecular targets for the therapeutic treatment of a variety of diseases on account of type of cells undergoing proliferation[1]. Drugs that target DNA typically work by intercalating between base pairs, alkylating to form covalent bonds, cleaving the DNA strand, or attaching to minor grooves[2]. The DNA minor groove binders (MGBs) drugs have become an interesting area for the development of new drugs possessing various potential as it binds either reversibly or irreversibly *via* non-covalent or covalent complex formation, respectively[3]. During reversible binding the drug interacts with the DNA backbone through electrostatic interactions, H bonding (HBs), or van der Waals bonds. DNA Minor Groove Binders (–MGBs) share various characteristics in common with one another, including being cationic, having multiple aromatic rings, and having the capacity to make hydrogen bonds. The hydrophobic aromatic rings present in the structure displace the spine of hydration along the minor groove and thus disturbs the protein-DNA interactions and therefore found multiple application in the therapy of different diseases. These small molecules mostly bind to adenine-thymine (AT) rich sequences of the DNA minor groove[4]. Distamycin and pentamidine are two molecules that have made significant strides in the creation of DNA minor groove binders with therapeutic action. They have also laid the foundation for development of more potent treatments. Many developments have been made, and the DNA minor groove continues to be a very relevant target for therapeutic purposes despite issues like selectivity still existing.

The well-established molecular dynamics has long been employed for describing the dynamic properties of macromolecules such as protein and DNA. This computer simulation can provide information on dynamics of DNA at atomistic level where movement of separate atoms and molecules can be followed in definite time frame which was not possible with experimental procedures such as X ray diffraction (XRD) and Nuclear magnetic resonance (NMR) techniques. Primarily the information provided from molecular simulation hinge on the kind of force field used to define the interactions. The quality of information obtained from molecular simulations

depends critically on the force field used to describe interactions in the simulated system. Therefore, it is imperative to identify the best force field used for DNA sequences for coarse-grained models to assess the validation and comparison of different DNA properties. Earlier researches have revealed 2,5-diarylfuran derivatives to be better minor groove binders and serve as potent antimicrobial agents[5-8,10]. The current work is in the same line for the development of a computational protocol for the enhancement of existing ligands and towards the search of novel antimicrobial agents[11-14]. Here, in particular 2,5-diarylfuran derivatives labelled as Mol-1, Mol-2 and Mol-3 selected from the literature[6] were geometrically optimized to attain a minimum potential with electrostatically optimized chemical structures. Afterwards all the ligands are subjected to molecular docking for analysing the interactions between DNA (PDB Id: 2MNE) base pairs and ligand and for the identification of possible binding site on the basis of previously reported research. It is already known that minimum the site binding energy, better will be the docked pose. Consequently, the docked pose having least binding site energy means having maximum binding affinity will be put to molecular dynamics trajectories over two force fields, AMBER03 & CHARMM27 for the 20 ns time duration. Later, through obtained result the conformational stability of ligand-DNA complex and time dependency of interaction and stability of complex is determined using both the force fields[7]. Such study is intended to recognize the best force field for studying the DNA-ligand interactions that will eventually aid-in drug design and also in studying drug metabolism through enzymes[8].

Methodology

System Selection and Preparation

Previous research work has established that furan-based derivatives possess the antimicrobial potential through binding with DNA[63,64]. Also, the dynamic stability of diarylfurans with DNA has been well-known[12]. Therefore, keeping in mind these facts, lead ligands (Mol-1, Mol-2 and Mol-3) were selected from literature having chemical structure shown in Fig. 1[6]. The DNA sequence having PDB Id 2MNE obtained from Protein Data Bank was used for the investigation[15,16]. The complete details about the structural data and base pair sequence attained from discovery studio software are stated in Table 1[17]. Before initiating the docking calculations, water molecules were removed from the DNA sequence using UCSF Chimera[18].

Molecular Geometry Optimization

Prior to docking and other computational calculations, geometry optimization for the ligands is carried out[9]. In the present work, geometry optimization of DNA groove binding ligands, *viz*., Mol-1, Mol-2 and Mol-3 has been carried out using Gaussian 09 software[18] using B3LYP hybrid functional at 6-31G** level of theory. Later these optimized ligands were subjected to molecular docking to selected DNA sequence, 2MNE. The main idea behind the descent of density functional theory is the exact determination of ground state energy and

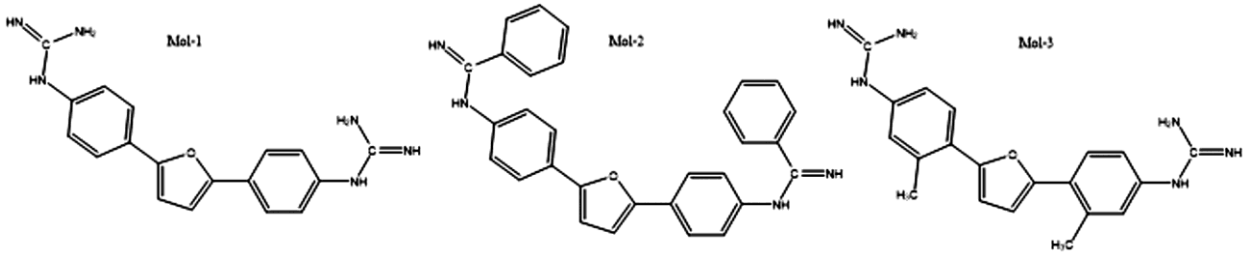

Fig. 1 — The chemical structures of selected ligands (Mol-1, Mol-2 and Mol-3).

Table 1 — Nucleic Acid Report for 2MNE retrieved from Discovery Studio Software

S. No.	PDB Id.	Nucleic Acid Sequence from Existing Structure	Fundamental Molecule Information
1.	2MNE	B-DNA Dodecamer (5'-D(CGCGTTAACGCG)-3') Chain A: CGCGTTAACG CG Chain B: CGCGTTAACG CG	• Cell Space Group: 19 (P212121 Origin-1 Choice: 1) • Parameters: 25.49 40.87 67.02 90.0 90.0 90.0 • Crystallographic Resolution: 2.30 Å • Molecular Weight of Nucleic Acid Chains: 7270.82 • Number of Nucleic Acids: 24 • Experimental pH: 7.50

other molecular properties with the help of the density of electrons[19]. The ground state energy of the system is expressed as the sum of three energies which are the kinetic energy of the system, the potential energy component, and the third one is the exchange-correlation energy[20] and all these three components can be expressed as a function of electronic density; so in short, the ground state energy is a function of the density of electrons[21] which is, in turn, the function of electronic coordinates and so named density functional theory. The lack of exact functional for exchange correlational energy paved the path for different approximations like local density approximation (LDA), generalized gradient approximation (GGA), and some hybrid functional[21]. B3LYP is a kind of hybrid functional of Hartree-Fock exchange functional with gradient and local corrections for exchange-correlation terms[22]. In B3LYP, the exchange-correlation functional is of the form of:

$$E_{XC}^{B3LYP} = (1 - a_0)E_X^{LSDA} + a_0 E_X^{HF} + a_X \Delta E_X^{B88} + a_C E_C^{LYP} + (1 - a_C)E_C^{VWN} \quad \ldots (1)$$

On the right-hand side of equation (1), the first term $(1 - a_0)E_X^{LSDA}$ is the exchange term with Local Spin Density Approximation, the second term $a_0 E_X^{HF}$ is the Hartree-Fock exchange functional, third term $a_X \Delta E_X^{B88}$ is Becke's exchange energy; the fourth term $a_C E_C^{LYP}$ is the gradient correction term proposed by Lee Yang Parr and the last term $(1 - a_C)E_C^{VWN}$ corresponds to the local correction term of Vosco Wilk and Nussair[23].

Molecular Docking Calculations

The extensively used computational technique of molecular docking primarily proposes the binding of a potent inhibitor with target biomolecule such as DNA with a main aim to assess the binding potential of ligand with the target biomolecule[10]. Molecular docking is a method to predict the orientation and position of the ligand interacting with the macromolecule (Protein, DNA, *etc.*)[24]. This is done by predicting the binding affinity (or strength of connection) of different ligand orientations at different positions of target DNA or protein with the aid of scoring function. The ligand outside the target explores different orientations, conformations, and positions in order to find the active site of the target[25]. Molecular docking aims to find the optimized conformation of the ligand-macromolecule complex so that the free energy of the complex system is minimized[26]. Interactions forces between the ligand and the target are generally divided into four categories that are Electrostatic forces, Electrodynamics forces (Van der Walls interaction), Steric forces, and Solvent related forces (hydrophobic interactions and hydrogen bonds)[27-28]. Molecular docking has two major sections: first is search algorithm, and the second is scoring function. Search algorithms are employed to achieve optimal conformations of ligand[29]. There are several search algorithms like Monte Carlo[29], Fragment-based, Systematic searches Genetic Algorithms and many more which are available in different docking softwares[30-31]. LGA[25] is a kind of genetic algorithm. The scoring function is considered to be a tool to rank different binding modes or different ligands. In principle, the score must correspond to binding affinity. Scoring functions can be knowledge-based, empirical, or molecular mechanics based. Accordingly, the binding mode, or the ligand with the best score, is considered to be the best binder[26]. The force field incorporated in the AutoDock4 implement the following function to evaluate the free energy,

$$\Delta G = W_{vdw} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) + W_{hbond} \sum_{i,j} E(t)\left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right) + W_{elec} \sum_{i,j} \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} + W_{sol} \sum_{i,j} \left(S_i V_j + S_j V_i\right) e^{\left(\frac{-r_{ij}^2}{2\sigma^2}\right)} + W_{conf} N_{tors}$$

In the above expression of free energy, the weighting constants W are the optimized parameters to reproduce empirical free energies of experimentally characterized complexes. The first term on the right-hand side of the above equation is 6/12 LJ potential which models dispersion and repulsion interaction where parameters were taken from AMBER[32]. The second term corresponds to the hydrogen bond term, E(t) is a directional parameter that depends on the angle[33]. The third term is the screened Coulomb potential similar to that used in AutoDock3[34]. The fourth term is the desolvation term that depends upon volume (V), solvation parameter (S), and exponentially depends upon the distance[35]. The last term predicts the loss of torsional entropy upon binding, where N_{tors} is the number of rotatable bonds.

In the current work, the molecular docking was performed through Autodock4 software[36]. The Lamarckian Genetic Algorithm (LGA) was employed for carrying the docking calculations. In short, initially the Gasteiger charges were added to the drug-

DNA complex through Autodock Tools (ADT). This was followed by a grid box preparation along the three coordinate axes to enclose drug and DNA. For each drug-DNA complex, a 20 LGA run with a maximum cycle of 2500000 energy evaluations was executed. In last, the least binding energy docked pose was mined and aligned with the DNA for further investigation.

Molecular Dynamics Simulation (MDS)

MDS is a widespread computational method for examining protein folding, unwinding, and other structural changes, such as complex stability, over a definite time period. It has gained much popularity due to a paucity of resources for experiments resources[37].

The main feature of Molecular Dynamics Simulation is to integrate Newton's equations of motion

$$m_i \frac{d^2 r_i}{dt^2} = -\nabla_i \left[U\left(r_1, r_{2,} r_{3,\ldots} r_N\right)\right], i = 1,2,3, \ldots N$$

in order to predict the dynamical properties and time evolution of the system of interest. To obtain the solution of the equation mentioned above, the most basic approach is to use the Verlet algorithm[38]. There are several versions of Verlet algorithm, like velocity Verlet algorithm[39] and Leapfrog form[40]. According to the original Verlet algorithm, the position at time t+Δt can be approximated to

$$r_i(t + \Delta t) = 2r_i(t) - r_i(t - \Delta t) - \frac{\nabla U}{m_i}\Delta t^2$$

U is the interaction energy that is governed by the force field.

Different force fields incorporate different potential functions. But most of the biomolecular force fields include class 1 type of potential formulation in original or slightly form[41].

$$U = \sum_{bonds} K_b\left(b - b_{eq}\right)^2 + \sum_{angles} K_\theta\left(\theta - \theta_{eq}\right)^2 + \sum_{dihedrals} K_\varphi[1 + cos\ cos(n\varphi - \sigma)] + \sum_{nonbondedpairs} \left(\varepsilon_{ij}\left[\left(\frac{R_{min,ij}}{r_{ij}}\right)^{12} - 2 * \left(\frac{R_{min,ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{r_{ij}}\right)$$

AMBER and CHARMM also implement Class 1 potential function in slightly different forms. The first three terms of the above equation are the interactions between the bonded atoms, and the last term models the interaction between the nonbonded pairs of atoms. First term on right-hand side models stretching of bonds about equilibrium bond length, second term models bending of bond angles about equilibrium bond angle, third term models rotation of dihedrals, and the fourth term models dispersion, repulsion, and electrostatic interaction.

In the above expression K_b and K_θ are the parameters that represent stiffness towards stretching of bonds and bending of bond angles. b_{eq} and θ_{eq} are the equilibrium parameters. K_φ is the barrier height corresponding to the rotation of dihedrals, φ is the dihedral angle, n and σ are periodicity and phase, respectively. There are two parts in the interaction energy of nonbonded interaction, first models LJ interaction, and the second models electrostatic interaction. ε is the well depth, R_{min} is the separation between nonbonded pairs of atoms for the minimum energy. Instead of R_{min}, some force fields use the separation between atoms at zero energy level, *i.e.*, σ[41]. AMBER and CHARMM both implement the R_{min} form of the LJ potential. The atoms separated by more than three bonds, referred to as 1-4 interaction, are also incorporated in nonbonded interaction but divided by some scaling factor. The value of this scaling factor for AMBER03 is 2.00, corresponding to LJ interaction, and the same is 1.20 corresponding to the LJ interaction[42]. q is the partial charge of individual atoms. 1-4 interactions are not scaled for the case of CHARMM force fields, *i.e.*, the scaling factor can be considered to be 1[41]. Partial charges differ from actual charges in such a way that actual charges can have only integer values, but the partial charges can have non-integer values in such a way that the net molecular charge should be the actual charge of the molecule. Above mentioned parameters are optimized to reproduce the set of experimental and computed quantum mechanical data sets. These datasets differ for different force fields. Although the target data to be reproduced to optimize the stretching parameters are the infrared spectra, this approach is more or less similar for all force fields[43]. AMBER and CHARMM both incorporates the Lorentz-Bertholet rule to compute LJ parameters, according to which the ε_{ij} is the geometric mean of the well depths of individual atoms ε_i&ε_j and $R_{min,ij}$ is the arithmetic mean of the separation at minimum energies of individual atoms $R_{min,i}$&$R_{min,j}$. In the AMBER03 force field, the electrostatic potential was computed in condensed phase using B3LYP/cc-pVTZ//HF/631-G**, and the partial charges were calculated by fitting

to Restrained Electrostatic Potential (RESP)[42,44]. The approach to compute charge parameters in the CHARMM force field was to obtain charge model from fits to dimer energetics of solute water and optimize it to reproduce slightly greater dipole moment[45]. Unlike AMBER, CHARMM has an additional term, Urey-Bradley angle term in the energy representation having the quadratic dependence upon the distance between the terminal atoms of 1-3 interactions. In AMBER, the improper dihedral rotation is modelled by the dihedral term of class 1 potential, but in CHARMM, the same is modelled *via* an additional quadratic term with a quadratic dependence on the improper dihedral angle[41]. CHARMM27 force field was developed by readjusting the backbone torsion potentials for nucleic acids[46]. GROMACS 5.1.1 (Groningen Machine for Chemical Simulations) software package[47] was used to carry out the molecular dynamics simulations in the current research work. A total of two drug-DNA complexes were created, after docking simulations for molecular dynamic simulations *viz*., (2MNE: mol-1, mol-2 & mol-3). These drug-DNA complexes were put to 5000ps (5ns) time scale simulation. There are many studies for the comparison of force fields for the nucleic acids but AMBER force fields[48] seems to be good for nucleic acid simulation due to the presence of specific topologies for the terminal nucleotides. AMBER03 force field already embedded in GROMACS software suite was used to generate the topology for the selected DNA sequence and antechamber module of AMBER program was used to generate topology of selected ligands through a python script: 'acpype.py'[49]. The Ligand-DNA complex was solvated in a box of varying dimensions using TIP3P water model at 298K[50]. Sodium ions were then added to the solvated box containing the DNA-ligand complex by randomly replacing the water molecules in order to neutralize the system. Particle Mesh Ewald (PME) was used to handle long-range electrostatic interactions in periodic boundary conditions[51]. Energy minimization of the whole system was carried out in 25000 steps using Steepest Descent leap- Frog Integration Method followed by NVT ensemble equilibration at a constant temperature of 300K for 50s using Berendsen thermostat[52]. The system was then equilibrated with NPT ensemble at a constant pressure of 1atm in 25000 steps using steepest descent leap-frog integrator[51]. Particle Mesh Ewald (PME) was used to handle long-range electrostatic interactions in periodic boundary conditions and all the bonds involving hydrogen atoms were constrained using the LINCS algorithm[53]. Graphs were plotted using XMgrace software[54].

Free Energy Calculations

To predict the most accurate and detailed description of the dynamics of a molecule, high-level quantum mechanical methods are considered the best method. But such an approach becomes quite complex and time-consuming for macromolecular systems such as biomolecules. For macromolecules, classical approaches are generally employed due to their reasonable accuracy. One such widely used classical simulation approach is Molecular Mechanics Poisson-Boltzmann Surface Area (MMPBSA) approach[55].

MMPBSA method is generally applied to calculate the binding free energy (ΔG_{bind}) of the ligand-biomolecule complexes. PBSA model is based upon approximating the solvation contribution to free energy using the continuum solvent model. In an aqueous solvent, the approximate binding free energy ($\Delta G_{bind,aq}$) can be written as

$$\Delta G_{bind\,,aq} = \Delta H - T\Delta S \approx \Delta E_{MM} + \Delta G_{bind\,,solv} - T\Delta S$$

ΔE_{MM} is the change in gas phase molecular mechanical energy, and is given by,

$$\Delta E_{MM} = \Delta E_{bond} + \Delta E_{angle} + \Delta E_{torsion} + \Delta E_{electrostatic} + \Delta E_{vdW}$$

$\Delta G_{bind\,,solv}$ is the change in solvation free energy and has two components,

$$\Delta G_{bind\,,solv} = \Delta G_{polar} + \Delta G_{non-polar}$$

$-T\Delta S$ is the measure of change in the conformational entropy upon binding.

Ensemble average over the set of sampled conformations is used to compute all these changes. In the MMPBSA approach, snapshots of the complex at various times during MDS production are taken and used to calculate average values of various quantities.

Change in polar solvation free energy ΔG_{polar} is computed using Generalised Born (GB) pairwise approximation or finite difference solution of Poisson-Boltzmann equation (PBE)[56,57]. And the change in non-polar solvation free energy $\Delta G_{non-polar}$ is calculated using the relation,

$$\Delta G_{non-polar} = \Delta G_{dispersion} + \gamma * SAV + b$$

The correction terms b and γ are set to be constant, and for the AMBER package, the constant values are –0.569 kcal/mol and 0.0378 kcal/mol-Å^2, respectively[58]. SAV is the Solvent Accessible Volume enclosed by the Solvent Accessible Surface Area

(SASA). $\Delta G_{dispersion}$ can be computed by solvent-accessible volume or surface area integration.

The quasi-harmonic analysis or normal mode analysis is used to approximate the configurational entropy (S). A different approach to S was proposed by Duan *et al.*[59,60] which was based upon interaction entropy (IE) and was defined as,

$$-T\Delta S = k_B T \; ln \; ln \; \langle e^{\beta \Delta E_{pl}^{int}} \rangle$$

ΔE_{pl}^{int} is the fluctuation of protein-ligand interaction energy, and the ensemble average $\langle e^{\beta \Delta E_{pl}^{int}} \rangle$ can be extracted from MD simulation[59,60]. In the current research work, G-MMPBSA method was applied to calculate the binding free energies of all receptor ligand complexes.

Results and Discussion

Geometry Optimization

Geometry optimization is a procedure to restrain the molecule in its lowest possible energy state. So, it is essential to start with the geometrically optimized structure for docking process, so as to have the QM charges[61]. Fig. 2 represents the optimized geometries of the selected ligands.

DFT Studies of the selected ligands

By means of B3LYP/6-31G(d,p) level in DFT calculations, geometry optimization and the molecular orbital energy values of all four ligands was studied and their theoretical values were compared with each other. The different geometrical parameters; such as Energy Gap (Eg), Electron Affinity (EA), Ionization Potential (IP), Electronegativity (μ), Chemical Potential (χ), Chemical Hardness (η), Chemical Softness (σ) and Nucleophilicity Index (ω), are arranged in Table 2. The negative values of HOMO and LUMO energy values for all the three ligands clearly suggest their structural stability. However, the lesser energy band gap between HOMO and LUMO of Mol-3 specifies higher feasibility of molecular charge transfer interaction compared to mol-1 and mol-2, as shown in Fig. 3. Henceforth mol-3 might demonstrate better DNA binding essential for virtuous biological activity.

Molecular docking

The computational estimation of structures of ligand-DNA complexes within a targeted binding site of DNA through molecular docking is of utmost importance. In the present study, the optimized structures of the certain ligands (mol 1, mol 2 and mol 3) were docked to 2MNE to obtain the best docked posed complex of ligand and 2MNE. The results of molecular docking for 2MNE DNA sequences are summarised in Table 3.The best docked posed structures of the selected DNA and optimized ligand structures are shown in Fig. 4.

It is clearly distinct from the structures of all the complexes that all ligands are minor groove binders for the selected DNA sequences. The most significant finding from the docking analysis is that by introduction of bulkier groups to the base structure of furan leads to lesser number of hydrogen bonding interactions due to increase in stearic hindrance. Our observation is also strengthened by the fact that add-on of bulky groups decreases the number of hydrogen bonding interactions obtained by Discovery Studio software[62]. The various interactions such as H-bonds, π-π (hydrophobic), and non-polar π -cation (noncovalent) interactions (Fig. 5) were used to explain the dockingresults and was found similar to previously reported minor groove binders.

Hydrogen Bonding Analysis

In the Fig. 6, respective binding site and equivalent Hydrogen donor/acceptor clouded regions have been

Table 2 — Table representing the DFT descriptors of the selected molecules

S. No.	Electronic Parameters (eV)	Selected Ligands		
		Mol-1	Mol-2	Mol-3
1	E_{HOMO}	–7.17	–7.70	–7.74
2	E_{LUMO}	–7.03	–5.25	–4.80
3	Band Gap (ΔE)	–0.14	–2.45	–2.94
4	Electron Affinity (EA)	7.03	5.25	4.80
5	Ionization Potential (IP)	7.17	7.70	7.74
6	Electronegativity (μ)	7.10	6.481	6.27
7	Chemical Potential (χ)	–7.10	–6.48	–6.27
8	Chemical Hardness (η)	0.07	1.23	1.47
9	Chemical Softness (σ)	14.28	0.81	0.68
10	Nucleophilicity Index (ω)	360.07	17.06	13.37

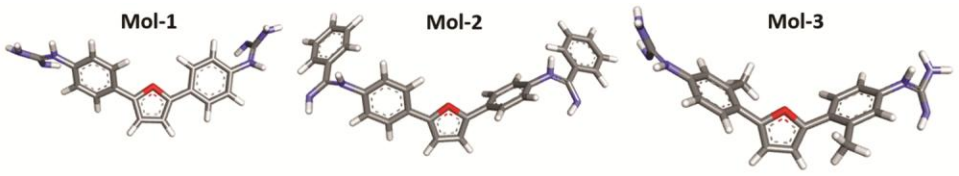

Fig. 2 — Optimized geometry of the selected ligands at B3LYP/6-31G(d,p) level

Table 3 — The hydrogen donor and the acceptor base pair residues involved in the formation of hydrogen bond between the DNA and ligand atoms

S. No.	DNA Sequence-Drug	No of H-bonds formed	Interacting Species	H-Bond Length (Å)	Docking Score (kcal/mol)
1	Complex-1	5	LIG0:H - A:DG7:O4	2.725	–8.35
			LIG0:H - A:DA6:N3	1.991	
			A:DG7:H22 - LIG0:O	2.170	
			LIG0:H - B:DC14:OP1	2.073	
			LIG0:H - B:DT15:OP1	2.149	
2	Complex-2	3	B:DG17:H22 - LIG0:O	2.098	–9.61
			LIG0:H - B:DT18:O2	2.904	
			LIG0:H - B:DC19:O4'	2.199	
3	Complex-3	3	LIG0:H - A:DA6:O3'	2.032	–8.77
			LIG0:H - A:DC4:O4'	2.189	
			LIG0:H - A:DC4:O5'	2.021	

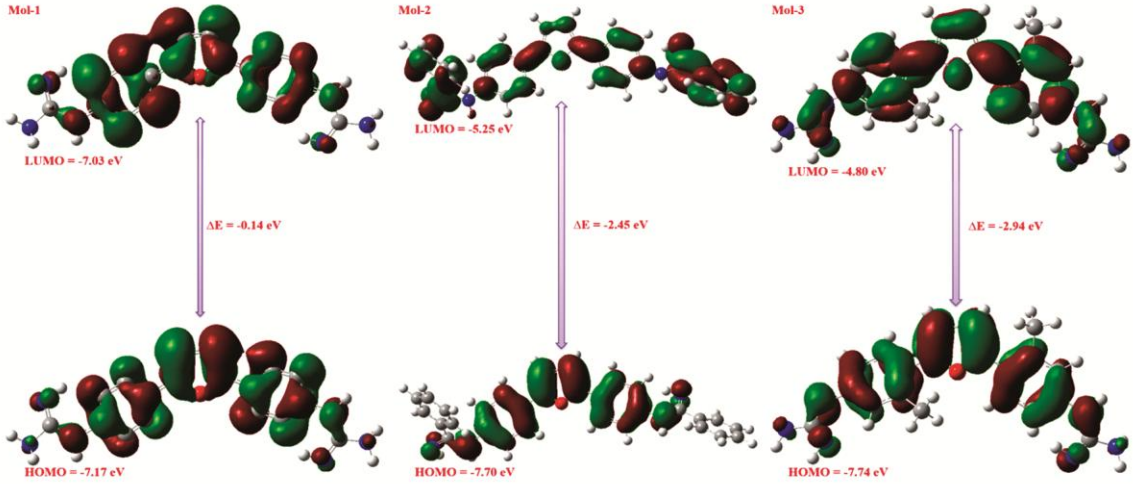

Fig. 3 — HOMO-LUMO orbitals of the selected ligands

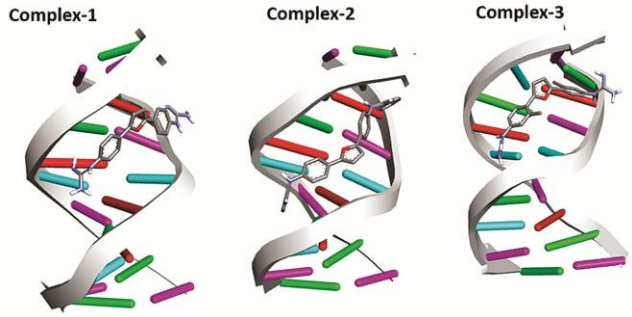

Fig. 4 — Docked poses of the selected ligands with 2MNE

shown clearly. Table 3 lists the interacting base pair species, the number of hydrogen bonds formed, and the length of the hydrogen bonds. The complex formed by three ligands namely mol-1, mol-2 and mol-3 with the 2MNE shows 5, 3 and 3 hydrogen bonds with base pairs respectively. Mol-1 forms strong hydrogen bond of bond length 1.99Å with DA6 and the same base pair also forms hydrogen bond in case of mol-3. Therefore, as per the maximum number of hydrogen bonds formed, mol-1 complex with DNA sequence 2MNE is of maximum stability.

Molecular Dynamics Simulation

Structural stability with time evolution can be studied using molecular dynamics simulation studies.

Using such methods computational studies for the structure and dynamics can be performed by mimicking the real time environments for a definite time duration. Results of molecular dynamics simulations are obtained in terms of various parameters. These parameters were studied and discussed for both the force fields involved, as follows.

Energy Variations

We performed a 5ns MD using TIP3P water model with AMBER03 and CHARMM27 force fields. The energies obtained are represented in Fig. 7. Clearly

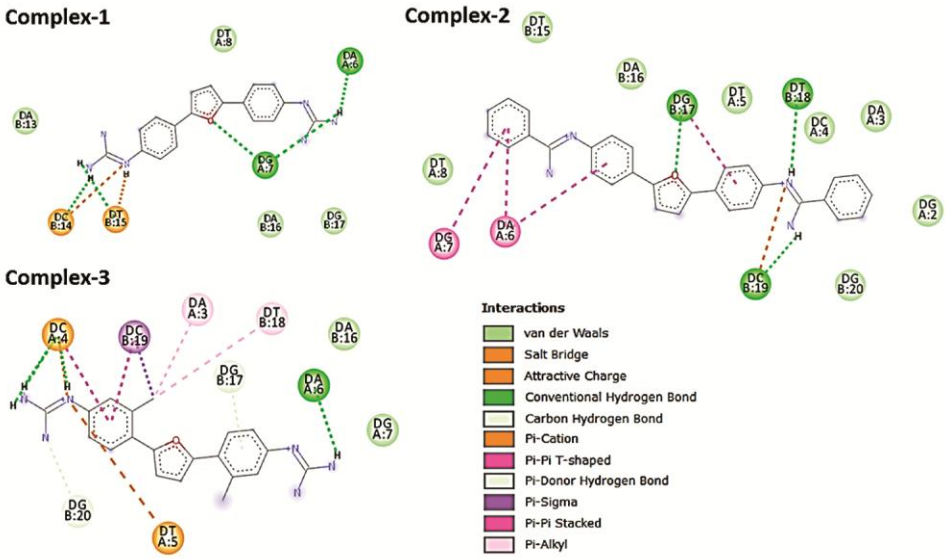

Fig. 5 — Different types of interactions between selected ligands with 2MNE

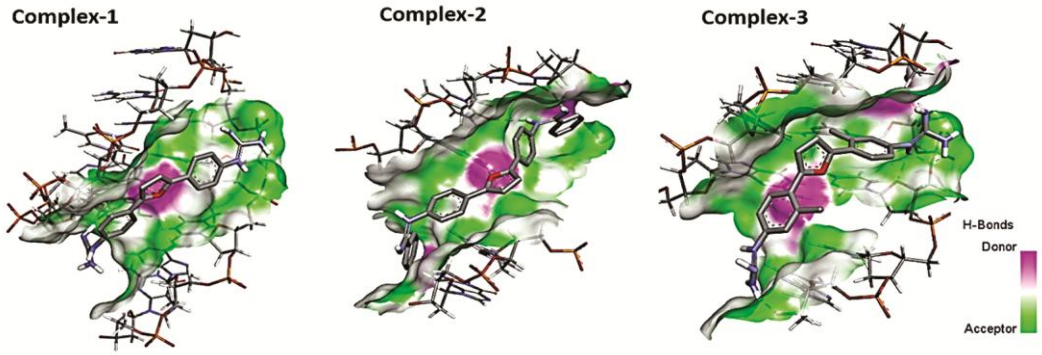

Fig. 6 — Different sites of H-bonding between selected ligands and 2MNE

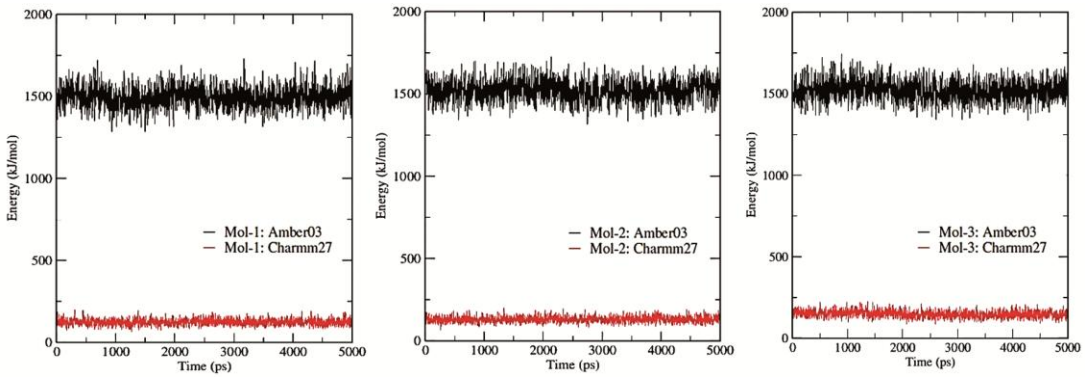

Fig. 7 — Energy variation upon MD simulation

both the force fields tend to converge the energetics of the complex; however, the CHARMM27 energies are lower than that of the AMBER03 energies. Along with that the range of variations in energies for AMBER03 is higher than that of CHARMM27. So, we may say that CHARMM27 holds an edge over AMBER03 with TIP3P water model for the simulation of DNA-ligand systems.

Variation in Radius of Gyration

R_g variations help us to understand the compactness of DNA helices. From Fig. 8we can see that the R_g

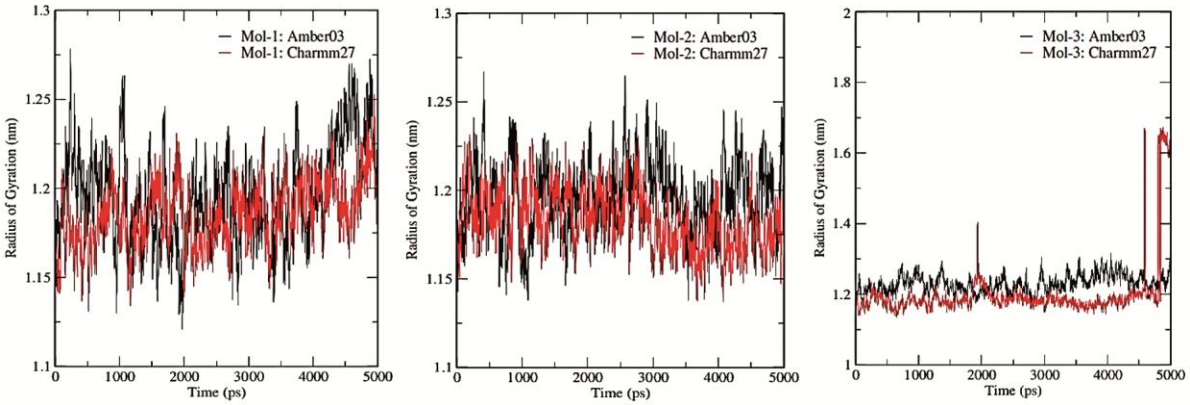

Fig. 8 — Radius of gyration variation upon MD simulation

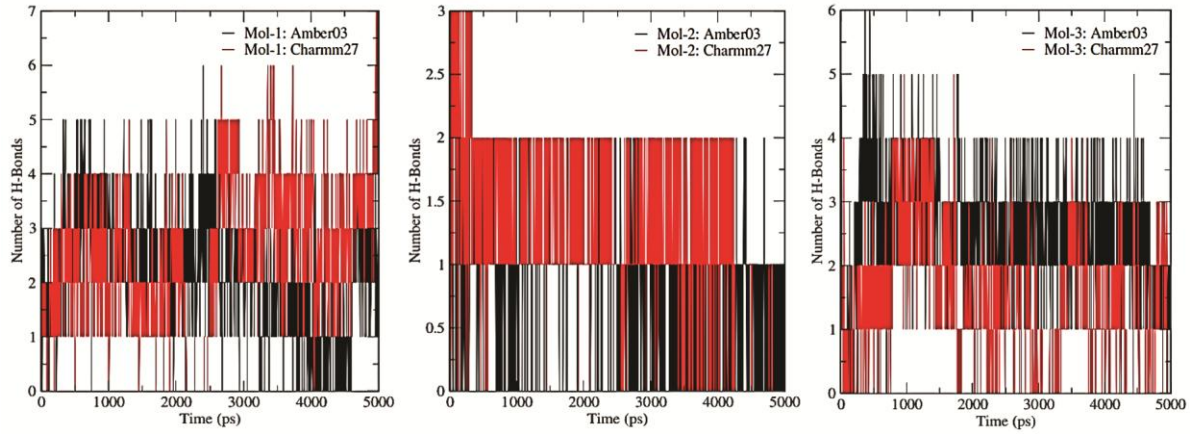

Fig. 9 — H-bond variation upon MD simulation

variations are comparable for all the three DNA-ligand systems for the 5ns MD using tip3p water model with AMBER03 and CHARMM27 force fields.

Variations in Number of Hydrogen Bonds

The non-covalent interaction, hydrogen bond formation between DNA (molecular target) and ligand is one of the main factors for the stability of complex. The stability and strength of the complex is directly proportional to the number of hydrogen bonds formed between DNA and the ligand. Fig. 9 demonstrates the number of hydrogen bonds formed and broken between ligand and DNA *vs*. simulation time for the two force fields; CHARMM27 and AMBER03 using TIP3P water model.

Table 4 consolidates the numbers of H-bonds formed during the MD simulation. Through the data obtained for number of H bond in MD simulation, it can be depicted that for two out of three systems CHARMM27 force field reports maximum number of H-bonds than AMBER03. Henceforth it may be deduced that CHARMM27 force field with TIP3P water model gives good results for H-bonding than AMBER03 force field in short DNA sequence.

Table 4 — Number of H-bonds formed during the MD simulation

System/FF	Complex-1	Complex-2	Complex-3
AMBER03	6	2	6
CHARMM27	7	3	5

Root Mean Square Deviation (RMSD)

RMSD gives us a brief idea about the conformational stability of the system. The RMSD plots of all the drug - DNA complexes simulated using tip3p water model at AMBER03 and CHARMM27 force fields are represented in Fig. 10. From the variations shown below, clearly both the force fields tend to converge the RMSD of the complexes, however, the CHARMM27 plots are more convergent than that of the AMBER03. Along with that the range of variations in energies for AMBER03 is higher than that of CHARMM27. So, we may say that CHARMM27 converge the RMSD of the DNA-

ligand systems even for a shorter simulation of 5ns than AMBER03.

Root Mean Square Fluctuation

The root mean square fluctuation (RMSF) in molecular simulation analysis determines the variation of each subunit of molecular target such as DNA. In the present work, we performed RMSF for the atoms involved in the DNA-ligand complexes. Fig. 11 signifies RMSF graph plotted among the RMSF value and the atom number. In the graph, the peaks represent that the fluctuations of set of atoms is higher than the variations of other atoms present in the molecule. It is clearly observed from the graph that in case of CHARMM27 force field with TIP3P water model, the fluctuations in the atoms remain minimal than that of AMBER03 force field with tip3p water model. Therefore, CHARMM27 should be the preferred force field for the simulations of such systems.

Free energy calculations

As CHARMM27 force field gave the more promising results for DNA-drug complex in terms of the energetics and the RMS plots, the component-wise free energy contribution of each drug-DNA complexes obtained from CHARMM27 trajectories was studied and shown in Table 5. The residue wise total energy as calculated by residue decomposition analysis of G-MMPBSA module is represented in Fig. 12.

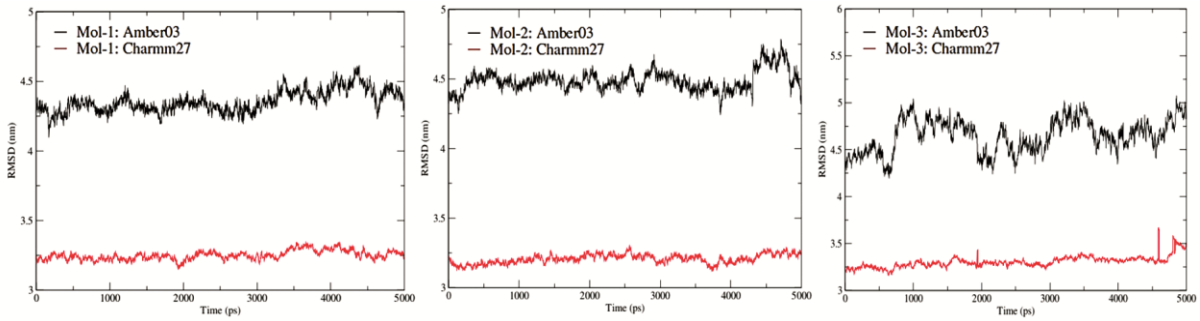

Fig. 10 — RMSD variation upon MD simulation

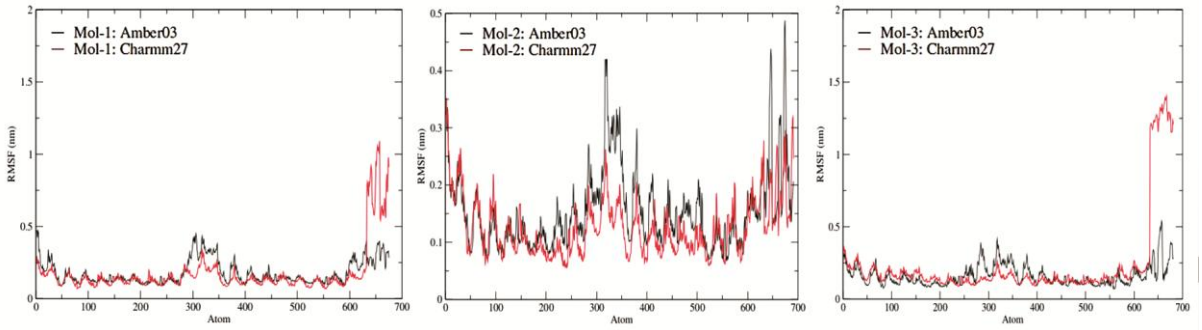

Fig. 11 — RMSF variation upon MD simulation

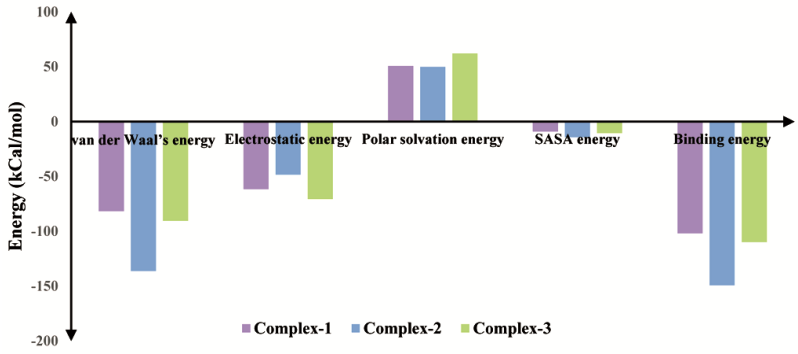

Fig. 12 — Variation in various free energy components

Table 5 — Various free energy components

S. No.	Free Energy Terms	Complex-1(kJ/mol)	Complex-2(kJ/mol)	Complex-3(kJ/mol)
1	van der Waal's energy	–81.800	–136.351	–90.639
2	Electrostatic energy	–61.758	–48.491	–70.851
3	Polar solvation energy	50.795	49.883	62.129
4	SASA energy	–9.265	–14.388	–10.643
5	Binding energy	–102.029	–149.346	–110.004

Conclusions

Development of DNA minor groove binders for treatment of various diseases has long been the keen area of interest. And for the same assistance through computational techniques have been very beneficial that aid-in studying the binding site, interactions involved, complex stability and dynamics of molecular target where motion of each atom can be tracked in particular time duration. As it is well-known that for the molecular simulation analysis, selection of force field is very crucial, specific and not just random. We tried to evaluate two force fields to model DNA-ligand interactions for shorter lengths of simulations. Diverse parameters are studied for observing the compactness of DNA, complex stability such as radius of gyration, RMSD, and RMSF. It was found that although both the force fields had their own pros and cons, CHARMM27 was found to have met the desired convergence criterions in terms of the energetics and the RMS plots obtained through it had minimum or no deviations at all than that of AMBER03. We further evaluated the free energies of the CHARMM27 trajectories and found the binding free energy to be maximum, followed by van der Waal's and electrostatic components. We may therefore predict that CHARMM27 is a suitable force field to model DNA-ligand interactions for shorter lengths of simulations.

Acknowledgement

AP would like to acknowledge University Grants Commission, New Delhi, India for the fellowship during the course of this research work. We would also like to thank Dr. Devesh Kumar, Professor, Dept. of Physics, BBAU, Lucknow for computational help and scientific discussions. AP would like to dedicate this article to the memory of Dr. Hemant Kumar Srivastava, Asso. Prof., Dept. of Medicinal Chemistry, NIPER, Guwahati, India.

References

1 Watson J D & Crick F H C, *Nature*, 171 (1953) 737.
2 Pindur U, M Jansen &Lemster T, *Curr Med Chem*, 12 (2005) 2805.
3 Yadav U, Yadav S K & Yadav R K, *J Mol Liq*, 280 (2019) 135.
4 Wilson W D, Tanious F A, Mathis A, Tevis D, Hall J E & Boykin D W, *Biochimie*, 90 (2008) 999.
5 Yadava U, Yadav S K & Yadav R K, *DNA Repair (Amst)*, 60 (2017) 9.
6 Tidwell R R& Boykin D W, "Dicationic DNA minor groove binders as antimicrobial agents," *Small Mol DNA RNA Bind Syn nuc acid com*,(2002) 414–460,
7 Pandey A, Misra M & Yadav A K, *Eur J Biol Res*, 11 (2021) 14.
8 Yadav R, Pandey A, Awasthi N & Shukla A, *Adv Sci Eng Med*, 12 (2019) 83.
9 Pandey A, Upadhyaya A, Kumar S & Yadav A K, *Drug Des*, 10 (2020) 7.
10 Pandey A, Yadav R, Shukla A & Yadav A K, *Adv Sci Eng Med*, 12 (2020) 40.
11 Tiwari P, Chapagain P P, Seddek A, Annamalai T, Üren A & Tse-Dinh Y-C, *Chem Med Chem*, 2020.
12 Shukla A K, Shrivas M K, Pandey A & Pandey J, *Bioorg Chem*, 109 (2021) 104687.
13 Adhikari A, Pandey A, Kumar D & Tiwari A K, *Lett Drug Design Disc*, 19 (2022) 549.
14 Shaik S, Kumar D, Visser S P D, Altun A & Thiel W, *Chem Rev*, 105 (2005) 2279.
15 Alniss H Y, Marie-Virginie S, Sadikov M, Golovchenko I, Anthony N G, Khalaf A I, MacKay S P, Suckling C J & Parkinson J A, *Chem Bio Chem*, 15 (2014) 1978.
16 Berman H M, Westbrook J, Feng Z, Gilliland G, Bhat T N, Weissig H, Shindyalov I N & Bourne P E, *Nucleic Acids Res*, 28 (2000) 235.
17 D S Biovia, *Discovery studio modeling environment*, Release, 2017.
18 Pettersen E F, Goddard T D, Huang C C, Couch G S, Greenblatt D M, Meng E C, Ferrin T E, *J Comp Chem*, 25 (2004) 1605.
19 Hohenberg P & Kohn W, *Phys Rev*, 136 (1964) B864.
20 W Kohn & Sham L J, *Phy Rev*, 140 (1965) A1133.
21 Tandon H, Chakraborty T & Suhag V, *Studies*, 39 (2019) 46.
22 Stephens P J, Devlin F J, Chabalowski C F & Frisch M J, *J Phys Chem*, 98 (1994) 11623.
23 Vosko S H, Wilk L & Nusair M, *Can J Phys*, 58 (1980) 1200.
24 Lengauer T & Rarey M, *Curr Opin Struc Biol*, 6 (1996) 402.
25 Morris G M, Goodsell D S, Halliday R S, Huey R, Hart W E, BELEW R K, OLSON A J, *J Comp Chem*, 19 (1998) 1639.
26 Chaudhary K K & Mishra N, *Databases*, 3 (2016) 4.
27 Goodsell D S & Olson A J, *Pro Struc Func Bioinform*, 8 (1990) 195.
28 Kuntz I D, Blaney J M, Oatley S J, Langridge R & Ferrin T E, *J Mol Biol*, 161 (1982) 269.
29 Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H & Teller E, *J Chem Phys*, 21 (1953) 1087.

30 Rarey M, Kramer B & Lengauer T, *J Comp Aided Mol Des*, 11 (1997) 369.
31 Schulz-Gasch T & Stahl M, *J Mol Model*, 9 (2003) 47.
32 Weiner S J, *J Am Chem Soc*, 106 (1984) 765.
33 Goodford P J, *J Med Chem*, 28 (1985) 849.
34 Mehler E L & Solmajer T, *Protein Eng Des Sel*, 4 (1991) 903.
35 Stouten P F W, Frömmel C, Nakamura H & Sander C, *Mol Simul*, 10 (1993) 97.
36 Morris G M, Huey R, Lindstrom W, Sanner M F, Belew R K, Goodsell D S & Olson A J *J Comp Chem*, 30 (2009) 2785.
37 Zhao H &Caflisch A, *Eur J Med Chem*, 91 (2015) 4.
38 Verlet L, *Phys Rev*, 159 (1967) 98.
39 Swope W C, Andersen H C, Berens P H & Wilson K R, *J Chem Phys*, 76 (1982) 637.
40 Allen M P, *Comp Soft Matt Synth Poly Pro*, 23 (2004) 1.
41 Guvench O & MacKerell A D, *Molecular Modeling of Proteins*, (Springer) 2008, p. 63.
42 Duan Y, Wu C, Chowdhury S, Lee M C, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J & Kollman P, *J Comp Chem*, 24 (2003) 1999.
43 MacKerell A D, Bashford D, Bellott M, Dunbrack R L, Evanseck J D, Field M J, Fischer S, Gao J, Guo H, Joseph-McCarthy S D, Kuchnir L, Kuczera K, Lau F T K, Mattos C, Michnick S, Ngo T, Nguyen D T, Prodhom B, Reiher W E, Roux B, Schlenkrich M, Smith J C, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D & Karplus M, *J Phys Chem B*, 102 (1998) 3586.
44 Bayly C I, Cieplak P, Cornell W & Kollman P A, *J Phys Chem*, 97 (1993) 10269.
45 Neria E, Fischer S & Karplus M, *J Chem Phys*, 105 (1996) 1902.
46 Foloppe N, MacKerell J & Alexander D, *J Comp Chem*, 21 (2000) 86.
47 Abraham M J, Murtola T, Schulz R, Páll S, Smith J C, Hess B & Lindahl E, *Software X*, 1 (2015) 19.
48 Ponder J W & Case D A, *Adv Pro Chem*, 66 (2003) 27.
49 Silva A W S D & Vranken W F, *BMC Res Notes*, 5 (2012) 367.
50 Mark P & Nilsson L, *J Phys Chem A*, 105 (2001) 9954.
51 Darden T, York D & Pedersen L, *J Chem Phys*, 98 (1993) 10089.
52 Ryckaert J-P, Ciccotti G & Berendsen H J C, *J Comp Phys*, 23 (1977) 327.
53 Hess B, Bekker H, Berendsen H J C &Fraaije J G E M, *J Comp Chem*, 18 (1997) 1463.
54 Turner P J, *XMGRACE, Version 5.1. 19, Cent Coast Land-Margin Res Oregon Grad Inst Sci Tech Beaverton, OR*, 2005.
55 Srinivasan J, Cheatham T E, Cieplak P, P A Kollman & D A Case, *J Am Chem Soc*, 120 (1998) 9401.
56 A Nicholls *et al.*, "Predicting small-molecule solvation free energies: an informal blind test for computational chemistry," *J Med Chem*, 51 (2008) 769.
57 Shivakumar D, Deng Y & Roux B, *J Chem Theory Comp*, 5 (2009) 919.
58 Wang C, Greene D, Xiao L, Qi R & Luo R, *Front Mol Biosci*, 4 (2018) 87.
59 Duan L, Liu X & Zhang J Z H, *J Am Chem Soc*, 138 (2016) 5722.
60 Duan L, Feng G, Wang X, Wang L & Zhang Q, *Phys Chem Chem Phys*, 19 (2017) 10140.
61 Pandey A, Mishra R & Yadav A, *Int J Anal Exp Modal Anal*, 12 (2020) 1300.
62 Wallace A C, Laskowski R A & Thornton J M, *Protein Eng Des Sel*, 8 (1995) 127.
63 Pandey A, Mishra R, Shukla A, Yadav A K & Kumar D, In-silico docking studies of 2,5-bis(4-amidinophenyl) furan and its derivatives, ISAFBM Conference Proceedings (2019) ISBN: 978-93-5351-824-0.
64 Shukla A, Mishra R, Pandey A, Dwivedi A K & Kumar D, Interaction of Flavonols with DNA: Molecular Docking Studies, ISAFBM Conference Proceedings (2019) ISBN: 978-93-5351-824-0.