CHAPTER 7

Computational insights on the mechanistic role of functionalized calixarene and fullerene

Article in press: *Journal of Molecular Modeling*.

DOI: 10.1007/s00894-013-1820-1

**Manishkumar B. Patel, Sivakumar Prasanth Kumar, Nikunj N. Valand, Yogesh T. Jasrai, Shobhana K. Menon**

*Synthesis, Biological Evaluation of Cationic Fullerene Quinazolinone Conjugates and its Binding Mode with Modeled M. tuberculosis Hypoxanthine-guanine Phosphoribosyltransferase Enzyme.*
Abstract

Novel calixarene and fullerene derivatives linked heterocyclic scaffolds have been synthesized and thoroughly characterized as described in the previous chapters. These molecules have biological applications as antimicrobial agents, antitubercular agents, DNA cleaving agents, anti-HIV agents and antimalarial agents. The mechanistic details for the activities observed were investigated using various bioinformatic tools. The antimicrobial activity of 1,3,4 oxadiazole derivative of calixarene was studied using molecular dynamic simulations while the antitubercular activity of fullerene-quinazolinone was studied using homology modeling. The DNA cleaving ability of the cationic fullerene s-triazine derivatives was better understood using DNA-ligand docking analysis. QSAR models were developed for the fullerene-quinoline conjugates exhibiting anti-HIV and antimalarial activity. Molecular docking and MD simulation techniques were used to study the interactions between PSC[4]R and LMN in the inclusion complex formed.
7.1 Molecular Dynamic studies of calix[4]arne based 1,3,4-oxadiazole

The synthesis and antibacterial study of calix[4]arne based 1,3,4-oxadiazole was reported in Chapter 2. Molecular Dynamic studies is discussed in this section.

7.1.1 Materials and Methods

The calix[4]arene functionalized with 1,3,4-oxadiazole moiety at the lower rim was selected as the initial structure as shown in the Figure 1A. Avagadro 1.1.0 program [1] was used to optimize the structure geometrically and utilized as the starting conformation for molecular dynamic (MD) simulations using XenoView v3.4 package [2]. The atom type of the molecular system was defined by Amber03 force field [3] and the atomic charges were added. Energy minimization of the molecular system was initially executed using Amber03 force field with tolerance 1.00e-4 kcal/mol and 100 steps of iterations. The MD simulations was performed with constant volume and shape ensemble in which the center of mass was fixed with 50000 as number of steps and 7 as cut off radius. The simulations were monitored upto 25 picoseconds (ps) at 298 K room temperature wherein an equilibrated structure was recovered at different time intervals to better understand the stability of the calix[4]arene 1,3,4-oxadiazole conjugate.

7.1.2 Results and Discussion

Geometry optimization of calix[4]arene coupled with 1,3,4-oxadiazole moiety was performed from Avagadro 1.1.0 program using Amber force field and a lowest energy conformer was retrieved which was subsequently deployed as the starting conformation for MD simulations in XenoView v3.4 package. Prior to MD, the structure was subjected to energy minimization to remove atomistic false bonds and steric clashes (Figure 1B). An ensemble of conformations were inspected over the generated MD trajectory which indicates the chemical moieties whose degrees of freedom are higher were subjected to acquire various chemical space leading to a series of conformations over potential energy surface. The methyl group at the upper rim of the calix[4]arene and the azole group of 1,3,4-oxadiazole group was observed to acquire different conformations owing to torsional angles.
An equilibrated structure was retrieved at 10 ps from the MD trajectory which depicts that the 1,3,4 oxadiazole group adopts twisted conformation in clockwise (+90°) and anticlockwise (-90°) direction with no change over projected plane (Figure 2A). This twisted conformation is due to the flexibility of carbon chains connecting azole at one end and lower rim of the calix[4]arene at another end. It should be noted that the aromatic connectivity of the calix[4]arene was stable during the simulation while the methyl group at its upper rim was observed to translate in 3D space. The energy plots and the pressure scaling indicates the structural stability of aromatic cage of calix[4]arene with various conformations owing to flexibility of functionalized 1,3,4-oxadiazole group (Figure 2B & C). We found that the geometrically optimized conformation of calix[4]arene 1,3,4-oxadiazole conjugate was similar to that of equilibrated conformation recovered at 10 ps. The intrannular region of calix[4]arene was noticed to obtain various volumetric cavity while no change over at upper rim while lower rim was kindled with clockwise and anticlockwise torsional rotations. We anticipate that functionalization of 1,3,4 oxadiazole was prominent for biological activity while the hydrophobic nature of calix[4]arene as carriers makes the supramolecule superior towards enhancing interaction with molecular targets crossing membrane barriers.

Figure 1 (A). 2D sketch of geometry optimized conformation of calix[4]arene 1,3,4-oxadiazole conjugate. (B). Energy minimized conformation retrieved using energy minimization module of XenoView.
Figure 2 (A). The molecular system of equilibrated calix[4]arene 1,3,4-oxadiazole conjugate retrieved at 10 ps. The plots of energy terms (B) and the internal pressure (C) calibrated up to 25 ps from MD trajectory.
7.2 Homology modeling and Molecular docking studies of cationic fullerene quinazolinone conjugates

The synthesis and antibacterial study of Cationic fullerene quinazolinone conjugates were reported in Chapter 3. Homology modeling and Molecular docking studies is discussed in this section.

![Cationic Fullerene quinazolinone conjugates](image)

**7.2.1 Materials and Methods**

**7.2.1.1 Sequence alignment and secondary structure comparison**

Homologous proteins of *M. tuberculosis* HGPRT was retrieved using Blastp [4] search over RCSB PDB [5] with BLOSUM62 scoring matrix to identify templates for modeling. The best template was identified by its E value and its sequence range. The modeling was executed in ExPASy (Expert Protein Analysis System) Swiss Model program with the specification of chosen template.

Swiss Model [6] has four progressive stages to efficiently model protein structures. These include (i) template selection: selection of PDB coordinate structures files with quality indicators, (ii) target template alignment: alignments are optimized by initially superimposing using an iterative least square algorithm and then, a local pairwise alignment was performed and optimized with a heuristic approach, (iii) model building: high scoring sequence similarities regions are first modeled and loop assignments and modeling are performed using constraint space programming (CSP) and at a later stage, side chains are fixed using rotamer library and (iv) evaluation: the deviations in the generated model are settled by descent energy minimization using the GROMOS96 force field [7].
The disorderness prediction was carried out using Disopred program [8] hosted by UCL to distinguish the regions of disordered structural profile which in turn can be deployed in secondary structure alignment in order to remove spurious alignment between template and the sequence inputted.

7.2.1.2 Model validation and stereochemistry corrections

The modeled \textit{M. tuberculosis} HGPRT was subjected to standard energy minimization protocol using AMBER92 ff99SB force field with 1,000 steps of steepest gradient optimization procedures in Chimera software [academic version]. AMBER92 ff99SB is a refined force field with main emphasis on improvisation of amino acids side-chain torsions ($\chi_1$) [9]. The modeled protein was validated by Ramachandran plot [10] (stereochemistry check) and QMEAN [11] (quality measure). The model structure file was specified in PROCHECK program available over SAVES (Structural Analysis and Verification Server) server [12]. QMEAN6 Z-score was studied using QMEAN server for model quality estimation.

7.2.1.3 $pK_a$ predictions

As the modeled \textit{M. tuberculosis} HGPRT was known to constitute charged residues, we calculated the $pK_a$ values to understand the local environment of the residues at physiological pH and its H bonding potential. H++ [13] works on standard generalized Born (GB) or the Poisson-Boltzmann (PB; considered in this study) methodology to compute $pK_a$ values of ionizable residues. The analysis was carried out in an explicit solvent with endurance in dielectric ionic strength (internal: 6, external: 80, pH: 6.5, salinity: 0.15) with one Mg$^{2+}$ counter ion. Structure file was prepared with AMBER ff10 [14] topology descriptions and assignment of missing hydrogens.

7.2.1.4 Binding site prediction

It is well known that HGPRTs possesses catalytic site for its enzymatic activity [15]. Q-SiteFinder was employed to identify ligand binding sites which utilize an interaction energy scheme to locate energetically favorable binding sites between the protein and a simple van der Waals probe [16]. Spatial proximal scheme was used to
Cluster energetically favorable sites and then ranked by accounting its estimated interaction energies. The top scoring predicted cavities were deployed in active site directed docking studies.

7.2.1.5 Molecular docking of fullerene derivatives and its energetic profile

Prior to docking, the ligand dataset was energy minimized using Chem3D all-atom force field [17] engineered in PyRx program [18] for Scripps Research Institute. Docking simulations were performed on a single machine equipped with Intel Dual-Core processor with 4 GB RAM and 150 GB hard disk operating on a Windows 7 operating system using iGEMDOCK version 2.1 program [19]. iGEMDOCK computes a ligand conformation and relative orientations owing to search space relative to the protein target binding site by applying genetic evolutionary methods (GA). Screening was carried out with the following specifications: population size: 500, number of generations: 100 and number of solutions: 10. Followed by generation of interaction profiles, the pharmacological interactions were mined and assigned a Z score. We sorted the interaction profiles in the following order: binding energy, H bonding term and van der Waals term with electrostatic term remained undisturbed. The ligand-receptor interaction profile was further investigated using PEARLS [20]. web program which extensively makes use of AMBER empirical force field [21] to estimate various energy descriptors

7.2.2 Results and discussion

7.2.2.1 Sequence alignment and secondary structure comparison

*M. tuberculosis* HGPRT is composed of 202 amino acids and possesses phosphoribosyl transferase (PRT type I) domain [22]. Blastp [4] search over RCSB Protein Data Bank (PDB) [5] enabled us to identify the homologous proteins in various organisms. These include HGPRTs from *Thermoanaerobacter tengcongensis* (identities: 50.5%; similarities: 71.3%), *Bacillus anthracis* (46%; 61.1%), *Escherichia coli* (42.3%; 60.1%), *Salmonella typhimurium* (41.8%; 59.6%), *Vibrio cholerae* (41.5%; 60.5%), etc. Hence, we considered *T. tengcongensis* (PDB entry: 2geb, Expect (E) value for alignment: le-62) as template for modeling the *M. tuberculosis* HGPRT using Swiss
Chapter 7

Model [6]. The structural features of the present model were discussed based on the template crystallographic and biological data.

Upon careful inspection of pair wise secondary structure alignment generated from template and \textit{M. tuberculosis} HGPRT, all the structure assignments were proven to be consistent with the template and resembled the realistic protein structure (Figure 3). The N-terminal 13 amino acids were not modeled due to the unavailability of alignment corresponding to the region and merely represent a disordered structure profile. This disorderiness was manually validated using UCL (University College London) Disopred program [8] and found to be correlated with the template in a confidence window of 8 and 9.

![Secondary structure alignment of T. tengcongensis (template) and M. tuberculosis (model) HGPRTs.](image)

**Figure 3.** Secondary structure alignment of \textit{T. tengcongensis} (template) and \textit{M. tuberculosis} (model) HGPRTs. Active site residues and double serine repeat are highlighted in light blue and green colors. The secondary structure elements such as \(\alpha\)-helices and \(\beta\)-sheets are represented in cyan and purple colors.
7.2.2.2 Model validation and stereochemistry corrections

The modeled *M. tuberculosis* HGPRT was energy minimized using AMBER92 ff99SB [9] force field with 1,000 steps of steepest gradient optimization protocol to return a plausible model. Ramachandran plot [10] exhibited 91% amino acids in core areas of ϕ-ψ distributions with 9% dispensed in additional allowed regions and no outliers (Figure 4a). This stereochemistry quality measure can be comparable to the used template. *T. tengcongensis* HGPRT reported 97.3% in favorable regions, 2.7% in additional allowed regions and absence of any outlier. Besides, QMEAN6 Z-score, a composite scoring function consisting of the following pseudo energy terms: $C_\beta$ interaction, all-atom pairwise, solvation, torsional angle and agreement measures: secondary structure and solvent accessibility [11], for modeled HGPRT protein was observed to be 0.709 that is quite close to the experimental X-ray structures (Figure 4b). Thus, the model was proved to be structurally realistic and implicated in further analysis.

![Figure 4 (a). Ramachandran plot showed that amino acid distributions reflects a reliable structure. (b). The quality of model (a black dot in the Z-score plot) resembles the experimental structure statistics.](image-url)
7.2.2.3 Structural features of modeled *M. tuberculosis* HGPRT

The active form of *T. tengcongensis* HGPRT exists in dimer/tetramer form in solution [23]. The monomeric information of template was utilized to identify the structurally conserved regions (SCRs) and active sites in the present model. *M. tuberculosis* HGPRT constitutes 3 α-helices and 9 β-sheets (Figure 5). The hood domain consists of 3 β-sheets organized antiparallelly whereas the core domain enclosed by a network of 6 β strands. Structural comparisons with other members revealed that 4 loops designated by the Roman letters I, II, III and IV are actively participated in structural integrity of the enzyme. Loop I is known for its dimer interactions and exert intense influence on enzymatic efficiencies in an indirect fashion. It is otherwise known as PP$_i$ (inorganic pyrophosphate) loop for its binding affinity towards it. Loop II is structurally annotated as flexible loop for its elasticity in bringing the active site residues and oligomerization contacts, a prerequisite for stable conformation and enzymatic activity together with loop I. Loops III and IV offers its amino acids to engage in substrate-binding and catalysis. Another structural feature in the modeled protein is its double serine repeat with the sequence pattern, SSXXSSXXSS in the position span of 91-100 and forms an element of loop II. It is recognized for its thermostability activity. Chen and coworkers proposed that this loop will undergo large conformation upon binding to its substrate, α-D-phosphoribosyl-1-pyrophosphate (PRPP) and thereby sealing off the active site cleft with the intention of shielding reactive transition species from bulk solvent [15]. It is also suggested that this loop will espouse a 3-10 helix conformation, a more stable structural form in contrast to loop at a half closed position.

Structural comparison with template yielded significant conservativeness at the active site enriched with positive and negatively charged residues. These include Leu65, Lys66, Glu122, Asp123, Asp126, Lys154, Phe175, Asp182 and Arg188. We believed that the occurrence of charged residues essentially plays a vital role in catalysis at physiological pH. Theoretical p$K_a$ values were calculated using H++ program [13] by specifying the input of modeled protein. The predictions are as follows: Lys66 = 11.6, Glu122 = 6.8, Asp123 = 4.3, Asp126 = 1.2, Lys154 = 12.0, Asp182 = 1.3 and Arg188 =
12.0 making up a protein’s total charge of -5 at pH 6.5. Of note, the active site residues are largely distributed by loops of close proximity to core domain with extensions from hood domain. Three member aromatic rings composed of Phe175, Tyr183 and Tyr 201 were observed to be oriented towards each other and seems to stabilize Phe175 for purine binding and catalysis in the present model, which is in good correspondence with the experimental structure of *T. tengcongensis* HGPRT proposed by Chen *et al* [15].

![Image of protein structure](image)

**Figure 5.** The *M. tuberculosis* HGPRT model

### 7.2.2.4 Interaction of fullerene and its derivatives in *M. tuberculosis* HGPRT

A cavity with active site residues along with adjacent neighbors within the proximity of 6 Å was constructed to implement as docking grid. A grid space with a site volume of 197 Å³ predicted by Q-SiteFinder [16] was utilized in the docking simulations. The ligand dataset consisted of 6 fullerene derivatives together with the parent
compound, fullerene was energy minimized using Ghemical all-atom force field [17] and optimized by 1,000 steps of conjugate gradients protocol with the help of PyRx program [18].

Active site directed docking studies was performed using iGEMDOCK version 2.1[19]. We tried to investigate the specificity of fullerene moiety preliminarily to distinguish the floor to place the carbon cage in the vicinity of the active site and to allow quinazolinone side arm to interact with active site residues. Fullerene cage prefers to interact with two floors encompassed by the following architecture: β1-β8-β9 and β5-β6-β7 (Figure 6a). Upon examining the scoring function, fullerene moiety appears to interact with non-covalent forces in the β1-β8-β9 structural arrangement with van der Waals term of -48.8825 kJ/mol overall contributing $E_{\text{total}}$ of -49.9251 kJ/mol. Hence, it is evident that the major pharmacological activity of fullerene is its lipophilicity as reflected by the major contribution of van der Waals term (97.91 %) in the total binding energy function.

The occupancy of the fullerene moiety on the structural floor adjacent to the active site was a crucial finding to make the quinazolinone moiety interact with the active site residues (Figure 6b). As the quinazolinone moiety is enriched with H bond acceptors, H bond donors and hydrophobic pharmacophore groups and the active site possess chargeable amino acids, it is essential to understand the molecular interaction. Moreover, the preference of quinazolinone moiety of fullerene binding to more uncharged amino acids was observed which can be distinguished into two main aspects. (i). We strongly believe that the loops I and II are the chief protein folding unit to bring the active site residues very closer to each other in comparison to crystallographic template data. Hence, we foresee that the H bond acceptors and donors of quinazolinone may participate in H bonding with active site amino acids along with its neighbors. (ii). The aromatic pharmacophore of quinazolinone may aid in interaction largely via $\pi-\pi$ interactions by aromatic amino acids to flexibly place the aromatic ligand groups initially and then will optimize its H bonding patterns. As stated by first aspect, we found that aliphatic amino acids are better interacted with H bond donors and acceptor groups of quinazolinone side...
arm. Then again, the aromatic moiety of side arm prefers to interact with aromatic hydrophobic side chains of cavity residues. The second prospect deals about the \( \pi-\pi \) interactions, wherein the active site of *M. tuberculosis* HGPRT constitutes only one aromatic amino acid (Phe175) and intriguingly, it is populated by charged residues viz. Lys66, Glu122, Asp123, Asp126, Lys154, Asp182 and Arg188.

![Figure 6 (a). The docked poses of compounds 5a-f in the *M. tuberculosis* HGPRT enzyme. (b). A closer view of the molecules docked poses at the active site.](image)

The compounds in the dataset second binding energy in the range of -115.2901 to -134.6001 kJ/mol which shows that the molecules interact in a similar fashion. Fullerene moiety of compounds 5a, 5c and 5d was found to bind at the first floor whereas compounds 5b, 5e and 5f was observed in the second floor with all the quinazolinone side arm of the compounds set interacted in the common active site loop elements Table 1. Compound 5c had secured a total energy of -128.9799 kJ/mol with a summed function of H bonding term and van der Waals term of -126.6426 kJ/mol. Compound 5f was found to attribute an energetic representation having a total energy of -126.6058 kJ/mol and summed function of -123.2172 kJ/mol. The contribution of summed descriptor by
compounds, 5c and 5f was found to be 98.18 % and 98.88 % which is in good concordance with the fullerene moiety’s van der Waals term (97.91 %). In addition, the experimental inhibitory measure, MIC of these compounds (12.5 µg/mL and 3.125 µg/mL) was in good agreement as compared to its computational binding energy predictions.

The following are the amino acids of *M. tuberculosis* HGPRT associated with compound 5f H bonding: Ala38 (no. of H bonds: 1), Gly41 (1), Ile62 (1), Leu70 (1), Thr73 (2), Asp74 (3), Leu75 (1), Ala76 (1), Arg77 (1) and Tyr179 (1). Hence, it is evident that most of the aliphatic amino acids were preferred by fullerene side arm to establish H bonding with few charged residues (Asp74, Arg77) (Figure 7). Owing to the conformational entropy exerted by the better activity molecules, compounds 5c and 5f was estimated to be 8.96 and 9.33 kcal/mol using PEARLS (Program for Energetic Analysis of Ligand Receptor Systems) program [20], respectively. Thus, the binding of fullerene compounds can be summarized briefly that the conformation entropy required by fullerene moiety over the floor adjacent to active site and accelerates the proper placement of aromatic side arm in the cavity preferably with aromatic amino acids (for instance, Tyr179) and finally constructs a network of H bonds with aliphatic and charged amino acids located proximally.
Figure 7. 2D representation of compound 5f interaction with *M. tuberculosis* HGPR.T key residues. The H bonding and charged residues are designated in green and blue colored circles. The receptor surface area exhibiting π contacts are shown in pink arc.

Table 1. Energetic contributions of fullerene and its derivatives (5a-f).

<table>
<thead>
<tr>
<th>Compounds</th>
<th>$\Delta E_{\text{total}}^{a}$ (kJ/mol)</th>
<th>H bond term and van der Waals term$^{a}$ (kJ/mol)</th>
<th>Electrostatic term$^{a}$ (kJ/mol)</th>
<th>Ligand internal energy$^{a}$ (kJ/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fullerene</td>
<td>$-49.9251$</td>
<td>$-48.8825$</td>
<td>$0.00$</td>
<td>$-1.0426$</td>
</tr>
<tr>
<td>5a</td>
<td>$-125.8729$</td>
<td>$-123.5979$</td>
<td>$-0.6224$</td>
<td>$-1.6526$</td>
</tr>
<tr>
<td>5b</td>
<td>$-115.2901$</td>
<td>$-113.4021$</td>
<td>$-0.1264$</td>
<td>$-1.7616$</td>
</tr>
<tr>
<td>5c</td>
<td>$-128.9799$</td>
<td>$-126.6426$</td>
<td>$-0.6000$</td>
<td>$-1.7372$</td>
</tr>
<tr>
<td>5d</td>
<td>$-134.6001$</td>
<td>$-130.0406$</td>
<td>$-3.1334$</td>
<td>$-1.4261$</td>
</tr>
<tr>
<td>5e</td>
<td>$-122.9621$</td>
<td>$-121.2018$</td>
<td>$-0.2967$</td>
<td>$-1.4035$</td>
</tr>
<tr>
<td>5f</td>
<td>$-124.6058$</td>
<td>$-128.2172$</td>
<td>$-0.1523$</td>
<td>$-1.2363$</td>
</tr>
</tbody>
</table>

Conformational entropy (kcal/mol)$^{b}$

<table>
<thead>
<tr>
<th></th>
<th>Conformational entropy (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5e</td>
<td>8.96</td>
</tr>
<tr>
<td>5f</td>
<td>9.33</td>
</tr>
</tbody>
</table>

$^{a}$ Calculated using iGEMDOCK version 2.1.

$^{b}$ Estimated using PEARLS program.
7.4 DNA-ligand docking analysis using cationic fullerene s-triazine derivatives with developed pBR322 DNA structure model.

The synthesis and DNA cleavage study of cationic fullerene s-triazine conjugates were reported in Chapter 4. DNA-ligand docking analysis is discussed in this section.

7.4.1 Materials and method
7.4.1.1 DNA structure modeling and docking

Canonical B-DNA structure was constructed using 3D-DART (3DNA-Driven DNA Analysis and Rebuilding Tool) web server [24]. A randomly selected region from pBR322 nucleotide sequence [25] was given as input. 3D-DART uses 3DNA software suite [26], consisting of various modules to execute DNA modeling. The ‘fiber’ module developed a canonical B-DNA structure file whereas a corresponding base pair (step) parameter file was created by ‘find_pair’ and ‘analyze’ programs. Subsequently, ‘local’ and ‘global’ bends were introduced in the DNA structure file with default settings of roll, twist and tilt. The DNA structure was reconstructed using ‘rebuild’ module and the structure file in PDB format was recovered.

The ligand dataset (C₆₀, 5a-f and 7a-f) was drawn and optimized in Marvin Sketch 5.11.4. The DNA structure file (treated as receptor) and ligand set was specified as inputs in Hex 6.3. Hex docking algorithm relies on structure complementarity matching generated from polar spherical coordinates. The molecular surface was developed by utilizing several radial expansions of spherical harmonic basis functions [27]. The following parameter settings were provided to run docking simulations: (i) correlation type = shape and electrostatics, (ii) Fast Fourier Transformation (FFT) = 5D, (iii) grid dimension = 0.6, (iv) receptor and ligand range = 180, (v) twist range = 360 and (vi).
distance range = 40. The docked conformers were then energy minimized in post-processing step to generate orientations by docking energy. The best docked conformers were retrieved by clustering the docking energy.

7.4.2 Results and Discussion
7.4.2.1 Interaction of C₆₀
Several studies have presented that carbon cage of C₆₀ specifically interact with guanosines (G; nucleotide of DNA) of minor groove without any significant DNA sequence selectivity [28]. It is also known that minor groove forms the basis for such structural integration. Further, the G sites in minor groove are oxidized by \(^1\text{O}_2\) produced during fullerene photoirradiation [29]. Hence, it is clear that fullerene binding to minor groove may facilitate intercalation by oxidizing G sites of minor groove via photoexcitation and specifically interact with that sites owing to its structural complementarity regions [29]. Beside, atomistic simulations with DNA revealed that fullerene does not prefer to interact with major groove [30]. In addition, the association of C₆₀-DNA is predominantly driven by the hydrophobic interaction of carbon cage with that of DNA and such C₆₀-DNA stable hybrids are not associated by electrostatic interaction. We report here the binding mode of fullerene (devoid of side arm) with developed pBR322 ds (double standed) DNA. Molecular docking simulations yielded a precise hydrophobic interaction with minor groove dominated by Gs (forward strand: G₈₁, G₈₃ and G₈₅; reverse strand: G₃₃) with an interaction energy of -189.94 Kcal/mol (Figure 8). Hence, it is demonstrated that fullerene act as minor groove binder.

Figure 8. Docked conformation of fullerene in DNA minor groove.
7.4.2.2 Interaction of substituted triazine derivatives (intermediates).

We examined the association of substituted triazine derivatives (side arm) with ds DNA model to understand the binding of side arm and its contribution towards DNA intercalation. Consistent with recent study, fulleropyrrolidine structure completely fits on the DNA minor groove along with counterpart interaction (Figure 9). In compounds 5d-f, the morpholine attached to triazine moiety was found to stack inside the minor groove and establish π-interaction with nitrogenous bases of DNA. Schiff base which forms the linker between fulleropyrrolidine ring and substituted triazine moiety was appropriately fitted inside the minor groove. Substituted aromatic amines of s-triazine derivatives which contributes electron-donating group was found to interact electrostatically with phosphate group of DNA. However, compounds 5c and 5f bearing electron-withdrawing group revealed no interaction with DNA phosphate groups and directed outside the cavity. It should be noted that all the intermediates (5a-f) interact preferentially with DNA minor groove with an interaction energy in the range of -189.94 to -285.98 Kcal/mol and anticipated that it will act as minor groove binders.

Figure 9. Docked conformation of 7c and 7f in DNA minor groove.
7.4.2.3 Interaction of cationic fullerene substituted s-triazines (7a-f):
The interaction of cationic fullerene substituted s-triazine derivatives can be explained by its counterpart interaction discussed above. As stated earlier, the fullerene of fulleropyrrolidine ring occupies the minor groove cavity hydrophobically (forward-strand: G19 and G23; reverse-strand: G93, G94 and G95) whereas the attached quartarnary salt interacts with phosphates of nearby nucleotides (C22). Schiff base acting as linker was found to be buried inside the minor groove facilitating appropriate fitting of pocket and substituted s-triazine. In addition, the triazine moiety prefers to establish H-bonding (A24, G25, and T26) with phosphate backbone using its nitrogen and oxygen atoms (Figure 10). It is obvious that the contribution of counterpart interaction enhanced the interaction energy of fullerene conjugates. Molecules containing electron donating groups (7d-f) scored better interaction energy (7d: -404 Kcal/mol, 7e: -400 Kcal/mol, 7f: -398 Kcal/mol) than the electron-withdrawing molecules (7b:-359 Kcal/mol, 7c: -395 Kcal/mol). The ranking of interaction energy follows the cleaving ability of 7a-f.

Figure 10. Atomic details on the interaction of 7c and 7f in DNA minor groove.
7.5 Molecular docking assisted SVM-based QSAR model to understand the inhibitory mechanism of cationic fullerene-quinoline conjugate towards HIV-PR enzyme.

The synthesis and anti-HIV activity of cationic fullerene quinaoline conjugates were reported in Chapter 5. Molecular docking assisted SVM-based QSAR model is discussed in this section.

7.5.1 Materials and Method

The crystal structure of HIV-PR co-crystallized with 4-(4-chloro-phenyl)-1-{3-[2-[4-fluoro-phenyl]-[1,3] dithiolan-2-yl-propyl]-piperidin-4-ol (UCSF8) [PDB entry: 1 aid] selected as the target protein [31] and fullerene-based HIV-PR inhibitors which were collected from published data [32] were considered as ligand dataset. The target protein was prepared by removing crystallographic waters and added missing heavy and H atoms using PYMOL software [33] The protein was then energy minimization using Amber force field (1000 iteration steps) [34] in Tripos Benchware program [35]. The ligand dataset was drawn in Marvin Sketch 5.11.4 and saved in structure data (SD) format.

The lowest energy conformer of each ligand was retrieved using an energy minimization approach using Ghemical all-atom force field [17] (1000 steepest gradient steps) engineered in PyRx package [18]. Multiple reiterated docking studies were carried out in iGemDock v 2.1 [19] with HIV-PR structure (1 aid) as receptor molecule while inhibition dataset (52 molecules) as ligand, respectively. A support vector machine (SVM) model
was developed using SVM-light software version 6.01 [36], with residue-specific pharmacological interaction as descriptors (independent variables) while experimentally known binding affinities were converted into logarithmic units which were employed as dependent variable inconsistent pharmacological interaction (variables having 5% standard deviation, variables having more zero values) were removed.

7.5.2 Results and Discussion

We considered the crystal structure of HIV-PR in complex with 4-(4-chloro-phenyl)-1-{3-[2-[4-fluoro-phenyl]-1,3]dithiolan-2-yl-propyl]-piperidin-4-ol (PDB entry = 1 aid) due to the unavailability of HIV-PR complex with fullerene or its derivatives in the PDB.[5] The crystal structure of HIV-PR consisted of two identical chains (A and B) with UCSF8 inhibitor bound at active site. The water molecules and the inhibitors were removed using PYMOL software [33]. To enable docking of molecules, the experimentally characterized cavity was enumerated as the binding pocket where the residue Asp25 (chain A) and Asp25' (chain B) lies on the vertex of the pocket. We developed an inhibitor data set of 52 molecules by combining experimentally known HIV-PR fullerene derivatives [32] along with the synthesized molecule discussed in the present study. Primarily, the automated docking procedure was validated by self-docking approach wherein the co-crystallized ligand, UCSF8 was redocked at the defined binding pocket and measured root mean squared deviation (RMSD) value with the specification of co-crystallized as reference structure or template and the docked pose of UCSF8 as the superimposing molecule. The RMSD value of 2.01 Å indicated that the docking pocket adopted in the present study was able to reproduce the experimental observation.

The Inhibitor dataset was drawn in Marvin Sketch 5.11.4 and exported in SDF. The lowest energy conformer of each molecule in the dataset was received using an energy minimization approach where in Ghemical all-atom force field[17] was applied and endured 1000 iterative steps of steepest gradient optimization scheme implemented in PyRx package.[18]
The energy minimized dataset was specified as ligand molecule in iGemDock v 2.1,[19] a molecular docking software. HIV-PR binding pocket was developed using the co-crystallized inhibitor, UCSF8 in the protein structure (1aid) which acted as centroid to sort out the interacting amino acid along with its neighbour residue within the distance of 15 Å to evade blind docking. It is very important to mention that iGemDock after docking produces amino acid specific pharmacological interaction with its interaction energies. The pharmacological interaction represent conserved interacting residue which often form binding pocket with specific physico-chemical properties, to play the essential function of a target protein [37]. The pharmacological interactions were distinguished by the type of contact with residues in question and categorized as H-bonding, van der Waals and electrostatic interaction, respectively. The binding energy of each ligand was computed by summing up all its residue-specific energy terms. Better interacting ligand were recognized by lowest binding energy of the docked pose among the 10 generated poses. This selection criterion helped us to retrieve good docked pose per each ligand in the dataset for which the pharmacological interaction profile was predicted.

ILE-47-B, V-S-ILE-47-B, V-M-GLY-48-B, V-M-GLY-49-B, V-M-ILE-50-B, V-S-ILE-50-B, V-S-PRO-81-B, V-M-VAL-82-B, V-S-VAL-82-B and V-M-ILE-84-B. These sites are expressed in a simple notion for its readability in which the first letter indicates ‘H’ for H bonding; V for van der Waals contacts, the second letter indicates ‘M’ for main chain; ‘S’ for side chain, the third abbreviation stands for standardized 3 letter amino acid codes, the number at the forth position indicates the amino acid position in the HIV protein structure and the fifth character denote the type of chain viz. A and B.

We considered a linear kernel for SVM development using SVM-light software version 6.01 distributed by Thorsten Joachims from Cornell University[36]. Initially, the ligand dataset (52 molecules) were divided into training (40 molecules), validation (5 molecules) and test sets (7 molecules). It was ensured that data partition must accompany diverse activity range and chemical space. We made an extensive analysis on the reélection of ‘C’ and ‘ε’ parameters which affects the SVM model complexity and employed 10,000 interactions with default settings of (minimum = 0; maximum = 1,000,000) and ε (minimum: 0; maximum =1), respectively. The optimized models were then sorted sequentially and selected a better SVM model possessing significant statistical values : MSE = 1.522866, C = 0.359996 and ε = 0.002228. MSE >> C >> ε.

To study its applicability, we predicted the binding affinities of the test set and found that the residual (difference between the measured and the predicted binding affinities) are distributed in single logarithmic unit. The cross-correlation coefficient (r²) recorded a significant value 0.966 which additionally indicates the predictive ability of the developed SVM model. (Figure 11)The developed SVM model mainly focuses on the interactions of fullerene side arm with HIV-PR active site lining residues. The pharmacological profiles yielded H-bonding ability of Asp25, Asp29 and Asp30 while other key residues viz. Thr26 and Gly27 made no significant contribution in interaction profile. Hence, negatively charged residues, especially Asp forms an important interaction element in H bond profiles. On the other hand, van der Waals contacts were observed with amino acids viz. Gly27, Ala28, Val30, Ile47, Gly48, Gly49, Ile 50, Val82, Ile84, and Pro 81. Surprisingly, Asp25' made van der Waals interaction with docked
poses while other aspartic acid including Asp25, Asp29 and Asp30 develops this contact using its side chain. We anticipate the strength of residue-specific energetics in interaction with fullerene side arm forms the basement for predicting HIV-PR binding affinities. The developed SVM model is in good agreement with enriched van der Waals term since the binding of HIV-PR fullerene based inhibitors is due to the hydrophobicity and steric compatibility of fullerenes [38].

The interaction energy of the synthesized compounds (5a, b, e-i) follows the trend of binding affinity. Compound 5i which have the maximum aliphatic chain length scored interaction energy -163.8 kcal/mol while compounds learning less aliphatic chain have larger energy values (5f: -160.9, 5g: -154.9 and 5h: -161.1 kcal/mol) indicating the length of the aliphatic chain increases the binding affinity (5g: >0.92 μg/ml, 5h: >0.78 μg/ml) 5i: >0.75-74.4 μg/ml). To understand the type of interaction made by aliphatic chain containing fullerene derivatives, we studied the docked pose of compound 5i which showed that compounds 5i and 6 [32] exhibit similar van der Waals interaction with Asp25 and Asp25’ residues, respectively. We anticipate that the enhancement of aliphatic chain length improved binding affinities by extensive van der Waals contacts with amino acids lining the HIV-PR binding pocket. (Figure 12)

![Figure 11. Scatter plot of observed vs. predicted pIC50 values.](image)
Figure 12. Docked pose of compounds 5i & 6 inside HIV-PR active site
7.6 3D-QSAR model of fullerene quinoline derivatives was developed using kNN-MFA method.

The synthesis and anti-malarial activity of fullerene quinaoline conjugates were reported in Chapter 5. 3D-QSAR model of fullerene quinoline derivatives is discussed in this section.

7.6.1 Materials and methods

All molecular modeling studies were performed using the VLife Molecular Design Suite (MDS suite 4.0, version 3.5 from VLife Sciences, Pune, India) implemented on Intel Workstation with a Xeon processor and Windows 7 Ultimate operating system. The structures of all compounds were drawn in Marvin Sketch 5.11.4. All structures were cleaned and 3D optimized.

The inhibitory concentration data (IC$_{50}$) of chloroquinoline-fullerene conjugates (Table 2) was considered as biology activity. The in vitro biological activities of the ligand dataset (IC$_{50}$) were transformed into its logarithmic scale, -log pIC$_{50}$ and defined as dependent variable in generating 3D QSAR. 3D-QSAR model was developed using ‘k Nearest Neighbor-Molecular Field Analysis (kNN-MFA)’ methodology wherein the training and the test set was selected in the ratio of 4:1. It was ensured that compounds selection should balance the activity span between the two sets and exhibit structural diversity. The synthesized compounds were divided into a training set (6 compounds) and a test set (3 compounds). Uni-Column statistics was applied to ensure better partition of training and test set. The ligand dataset under study was subjected to molecular alignments using ‘template-based alignment’ method with actives 4i chosen as template. Tripos force field [35] with the specification of Gasteiger and Marsili atom charge type
was applied to generate electrostatic, steric and hydrophobic field descriptors using a carbon probe (charge = +1, dielectric constant = 1.0) over calculated field grid points. This procedure yielded 5,040 descriptors among which statistically significant descriptors were identified using 'Remove invariant columns' option and parsed as QSAR variables to develop 3D-QSAR. The 3D-QSAR studies were performed by kNN method using Forward-Backward Stepwise Variable Selection as variable selection method. The kNN methodology is based on a simple distance learning approach whereby an unknown member is classified according to the majority of its kNNs in the training set. The nearness is estimated by an appropriate distance metrics (e.g., a molecular similarity measure calculated using field interactions of molecular structures). The standard kNN MFA method is comprised of the following steps: (i). The distances between an unknown object \((u)\) and all the objects in the training set was calculated. (ii). According to calculated distances, select \(k\) objects from the training set which appears most similar to object \(u\). (iii). The object \(u\) was classified to a group to which the majority of the \(k\) objects belong. (iv). Finally, an optimal \(k\) value was selected by optimizing the test set classification (as carried out in steps i , ii and iii) or by applying leave-one out cross (LOO) validation technique. Forward Stepwise Variable Selection method was utilized to select the QSAR field variables and optimal \(k\) values. Subsequently, the steric, electrostatic, and hydrophobic energies were computed at the lattice points of the grid using carbon methyl probe and transformed into interaction energy values which are encoded as independent QSAR variables for model development. The model was generated according to the following conditions: cross-correlation limit = 1.0, model selection criterion = \(q^2\) and \(F\)-test ‘in’ = 4.0. As some additional parameters, variance cutoff was set at 0.000 kcal/molÅ and the the distance-based weighted average was selected as prediction method.

7.6.2 Results and discussion

3D-QSAR model was developed using kNN-MFA method with hybrid chloroquinoline-fullerene derivatives in Vlife Molecular Design suite. The electrostatic, steric and hydrophobic field descriptors were calculated using Tripos force field [35] and
Gasteiger-Marsili [39] atomic charge. Various descriptors were calculated using a carbon atom with charge +1.0 in distance dependent dielectric environment.

‘Forward-Backward Stepwise variable selection’ method was applied to sort out letter descriptors which contributes to the biological activity. Ligand dataset was partitioned into training (6 molecules: 4b, 4c, 4e, 4f, 4h, 4i) and test set (3 molecules: 4a, 4d, 4g). It was ensured that the selection of test set should span the activity range and chemical diversity of the dataset. ‘Uni-Column’ statistics of training and test set reflected the correct selection of these sets since this parameter check correctness of selection criteria. The maximum and minimum activities of the training and test set were compared as follow.

i. The maximum value of (Predicted Inhibitory concentration) pIC50 (6.2680) of test set should be less than or equal to maximum value of pIC50 of training set (6.6020).

ii. The minimum value of pIC50 (0.4651) of test set should be higher than or equal to minimum value of pIC50 of training set (0.2941).

This statistics showed the selection of test set was interpolative and derived within the minimum-maximum range of training set. The cross correlation coefficient (q² = 0.9238) was considered as the primary model selection criterion to recognize statistically significant 3D-QSAR model. Hence, the developed 3D-QSAR model displays good predictivity in regular cross validation. (Figure 13)

In 3D-QSAR kNN-MFA model, 3D data points generated over aligned chloroquinoline- fullerene conjugates were used to optimize the electrostatic, steric and hydrophobic requirements of the molecules spread over grid points which strongly favour the 3D pharmacophore features responsible for biological activity. Molecular alignment generated 18 hydrophobic field points spread over the molecular surface. Since the cross-correlation of field points possessing >0.5 were forbidden, fullerene which forms the identical chemical moiety in the entire dataset was not accounted for field points displaying diversity across molecules and may function as important component in enhancing binding affinities. In the generated molecular alignment, all the molecules was found to be aligned over fulleropyrrolidine scaffold while diversity of 4i possessing long
chain of Schiff base makes it to occupy other grid points. The best field point generated in ‘Forward-Backward stepwise variable selection’ kNN-MFA model was H-121(0.4000, 0.416) at lattice point 121. (Figure 14) This field point suggested the distribution of hydrophobic surface over aligned Schiff base linker at a lattice point 121 common to all of the molecules under grid space. The hydrophobic field suggested the preference of lipophilic groups near the R group of the dataset. This hydrophobic environment indicates the role of long chain lipophilic group attached to fulleropyrrolidine ring. The selected kNN-MFA model exhibited significant cross-correlation ($q^2 = 0.9238$) with a low standard error (0.1284) indicating the reliability and predictive accuracy of the model.

We anticipate that the long Schiff base chain along with the hydrophobic fullerene ball may enhance its lipophilicity and thereby increase the chance of transporting across cell membranes which is in good agreement with the antimalarial activity observed.

Table 2. Statistical parameters deployed in kNN-MFA 3D-QSAR model.

<table>
<thead>
<tr>
<th>Training Set</th>
<th>Test Set</th>
<th>Significant Descriptor</th>
<th>$q^2$ (cross-correlation)</th>
<th>$q^2$ (cross-correlation standard error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>3</td>
<td>H-121</td>
<td>0.9238</td>
<td>0.1284</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.4000, 0.416) at lattice point 121</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 13. Scatter plot of observed vs. predicted $pIC_{50}$ values.
Table 3. Comparison of observed pIC$_{50}$ with predicted pIC$_{50}$ values generated using 3D QSAR model.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Observed pIC$_{50}$</th>
<th>Predicted pIC$_{50}$</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>4a</td>
<td>6.602</td>
<td>6.835</td>
<td>-0.233</td>
</tr>
<tr>
<td>4b</td>
<td>6.812</td>
<td>6.728</td>
<td>0.084</td>
</tr>
<tr>
<td>4c</td>
<td>6.854</td>
<td>6.707</td>
<td>0.147</td>
</tr>
<tr>
<td>4d</td>
<td>6.268</td>
<td>6.701</td>
<td>-0.433</td>
</tr>
<tr>
<td>4e</td>
<td>6.602</td>
<td>6.833</td>
<td>-0.231</td>
</tr>
<tr>
<td>4f</td>
<td>7.585</td>
<td>7.594</td>
<td>-0.009</td>
</tr>
<tr>
<td>4g</td>
<td>6.854</td>
<td>7.593</td>
<td>-0.739</td>
</tr>
<tr>
<td>4h</td>
<td>7.585</td>
<td>7.594</td>
<td>-0.009</td>
</tr>
<tr>
<td>4i</td>
<td>7.602</td>
<td>7.585</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Figure 14. Molecular alignment of 4a-i used to generate kNN-MFA 3D-QSAR model. The significant hydrophobic field point (H _ 121 (0.400, 0.416) is shown in yellow colour.
7.7 Molecular docking and MD simulation techniques of PSC[4]R-LMN inclusion complex

We prepared inclusion complex of PSC[4]R with poorly water soluble drug LMN, synthesis physicochemical property, spectral analysis and oral acute toxicity were reported in Chapter 6. Molecular docking and MD simulation is discussed in this section.

![PSC[4]R:LMN Inclusion complex](image)

7.7.1 Materials and Methods

A host-guest complex of PSC[4]R and LMN was developed using molecular docking approach. Geometry optimization of the saddle shaped PSC[4]R and LMN was carried out using all-atom Ghemical force field [17] engineered in Avagadro 1.1.0 program [1] with steepest gradient approach (number of steps = 500, convergence = 10e-7 kcal/mol) and deployed as starting receptor and ligand conformations for docking in Hex 6.3 software. Hex utilizes spherical polar Fourier (SPF) basis functions circumventing grid sampling to derive generating functions (GFs) for 1D, 3D and 5D orientations and employs fast Fourier transformation (FFT) rotational correlations over angular terms constrained search space to sort best pairwise interactions between two molecules [40].

The docked poses in the best cluster were retrieved and examined manually using Jmol structure visualizer [41]. The pose having lowest $\Delta E_{\text{total}}$ (interaction energy term of Hex software) was provided as the starting conformation for molecular dynamic (MD) simulations using XenoView v3.4 package [42]. The atom type of the molecular system...
was defined by Amber03 force field [14] and the atomic charges were added. Energy minimization of the molecular system was initially executed using Amber03 force field with tolerance $1.00 \times 10^{-4}$ kcal/mol and 100 steps of iterations. The MD simulations was performed with constant volume and shape ensemble in which the center of mass was fixed with 50000 as number of steps and 7 as cut off radius. The simulations were monitored upto 50 picoseconds (ps) at 298 K room temperature wherein an equilibrated structure was recovered at different time intervals to better understand the molecular interaction in the PSC[4]R-LMN inclusion complex.

7.7.2 Results and Discussion

The inclusion complex of PSC[4]R and LMN representing a host-guest complex (1:1 stoichiometry) was developed using molecular docking approach. A saddle shaped PSC[4]R was chosen as the starting molecule and geometrically optimized using Avagadro 1.1.0 program (Figure 15A). The structurally optimized PSC[4]R was specified as receptor while LMN was defined as ligand. A 3D flexible translation of LMN over the PSC[4]R surface was enabled to capture the best interaction delineated by Hex 6.3 energy scoring scheme.
Table 4. Energetic details of the computational PSC[4]R-LMN inclusion complex models.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Energy (kJ/mol)</th>
<th>Force field</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMN</td>
<td>$\Delta E_{\text{LMN}} = 42.3805$</td>
<td>Ghemical</td>
<td>Geometry optimized using Avagadro 1.1.0 program</td>
</tr>
<tr>
<td>PSC[4]R-LMN inclusion complex</td>
<td>$\Delta E_{\text{PSC[4]R-LMN}} = 11151.3$</td>
<td>Ghemical</td>
<td>Best host-guest complex ($\Delta E_{\text{total}} = -702.17$ kJ/mol) retrieved from molecular docking</td>
</tr>
</tbody>
</table>

An ensemble of best docked conformations were retrieved and analyzed the best cluster containing lowest energy docked poses representing relative 3D orientations using Jmol structure visualizer. The docked pose of PSC[4]R-LMN complex securing a total energy ($\Delta E_{\text{total}}$) of -702.17 kJ/mol was considered as the best host-guest complex and employed in subsequent studies. The intrannular region of PSC[4]R forms a larger volumetric cavity to accommodate the aromatic group of LMN. It is also noticed that these 3D organization of PSC[4]R-LMN complex was sorted in the best cluster (20
poses) ranked by Hex scoring system indicating the reliability of the docked poses. The amine group of LMN establishes two hydrogen bonds with one of the \( p \)-sulfonato group localized at the PSC[4]R upper rim. The triazine and phenyl moieties of LMN also form \( \pi-\pi \) interaction with aromatic moieties of PSC[4]R upper rim. It was also found that the dichlorophenyl group in LMN forbids its interaction with PSC[4]R \( p \)-sulfonato group owing to electronegative clouds (Figure 15B). The shape complementarity plays an essential role to fit the LMN molecule onto PSC[4]R annulus and PSC[4]R \( p \)-sulfonato group act as ‘supporter appendages’ preventing LMN dissociation from the inclusion complex. It may be assumed from the docked pose that PSC[4]R resembles the shape of a scissor holding LMN in its cutting edge whereas the lower rim \( p \)-sulfonato group functions as handle. Hence, the introduction of \( p \)-sulfonato group in PSC[4]R acts as one among the important molecular units facilitating inclusion complex formation.

To validate the retrieval of best conformations from docking procedure, we relied on the variation in molecular total energy and considered an independent molecular mechanics program for its calculation. Avagadro 1.1.0 program was used to measure the total energies in which the molecules were atom typed using Ghemical force field (Table 4). It is evident from the comparison of energy terms that the PSC[4]R-LMN inclusion complex scored \( (\Delta E_{PSC[4]R-LMN}) \) 11151.3 kJ/mol where its composite partners, PSC[4]R and LMN possessed intermolecular energy of 12622.9 kJ/mol \( (\Delta E_{PSC[4]R}) \) and 42.3805 kJ/mol \( (\Delta E_{LMN}) \), respectively.

To further assess the role of \( p \)-sulfonato group on PSC[4]R and its structural integrity with LMN, we performed small MD simulations (50 ps) of the docked PSC[4]R-LMN inclusion complex in vacuo. We focused on the internal pressure or vibrations inside PSC[4]R which in turn affects the orientation of \( p \)-sulfonato groups on the lower and upper rims. Hence, a constant shape and volume ensemble of inclusion complex was specified as one of the constraint with variable internal pressure at room temperature and parsed as parameters of the MD simulations. The docked PSC[4]R-LMN inclusion complex was initially energy minimized using Amber03 force field engineered in XenoView v3.4 MD package and employed as starting conformation (Figure 16).
An initial structure at 10 ps was recovered from MD trajectory to better understand the structural organization of PSC[4]R-LMN inclusion complex (Figure 17A). A very short 3D motion was captured on the PSC[4]R annular region in which LMN’s triazine and phenyl moieties reorient (87.9°; measured in accordance with the docked pose) itself while the closeness of LMN’s diamino group with p- sulfonato group of PSC[4]R was noticed. In addition, the electronegative clouds of the LMN dichlorophenyl moiety and PSC[4]R p- sulfonato group occludes the interactions between them as similarly ascertained in the docked pose. This steric clash enables the p- sulfonato group of PSC[4]R upper rim to move away from each other (11.13 Å) thereby causes an increase in internal pressure inside PSC[4]R molecule. This 3D motion also drives the p- sulfonato group on the lower rim to stretch (11.34 Å). Thus, the PSC[4]R p- sulfonato group essentially plays a key role in the structural integrity of PSC[4]R-LMN inclusion complex.

However, on increasing the internal pressure (15 ps) on PSC[4]R molecule, the p- sulfonato group at the upper rim gets contracted (10.17 Å) and decreased the volume of the annular cavity making the LMN to dissociate from the inclusion complex (Figure 17B). This contraction was also noticed in the lower rim where p- sulfonato group approach each other (10.05 Å). After 15 ps, the inclusion complex is unstable at longer times possibly due to the change in internal pressure and fragments into small carbon molecules. Hence, it is evident that the 3D arrangement of p- sulfonato group in PSC[4]R enhances the molecular interaction and structural integrity of PSC[4]R-LMN inclusion complex.
**Figure 15 (A).** Geometrically optimized PSC[4]R-LMN inclusion complex and (B) Docked pose of LMN on PSC[4]R showing 2 H bonds and 2 π-π contacts.

**Figure 16 (A).** The molecular system of PSC[4]R-LMN inclusion complex used in MD simulations and the plot of potential energy over 100 iterations (B), the composite energies (C) and the internal pressure (D) calibrated upto 25 ps from MD trajectory.
7.7.3 Conclusion

The mechanistic studies of the designed supramolecules were investigated using leading bioinformatics approaches. A molecular dynamic (MD) simulation of calix[4]arene 1,3,4 oxadiazole conjugate revealed that the functionalized group 1,3,4-oxadiazole was observed to acquire various conformations owing to its flexible carbon chain connectivity the lower rim of the carbon cage of calix[4]arene. It indicates that the stability of the functionalized calix[4]arene lies in the aromatic connectivity. The active role of cationic fullerene-quinazolinone conjugates towards antitubercular activity was studied using molecular docking. Using *M.tuberculosis* HGPRT homology model as template, the binding posses of fullerene-quinazolinone conjugated were enumerated which revealed that the conformational entropy exerted by fullerene over HGPRTs active site accelerated the development of H bond network.

DNA-ligand docking analysis was performed using cationic fullerene s-triazine derivatives with developed pBR322 DNA structure model. We found that s-triazine owing to its electronic properties may act as minor groove binders where as fullerene dictates A-T sequence specificity over minor groove. Fullerene upon functionalization was observed to enhance binding affinity in comparison to s-triazine and fullerene alone.
We developed a SVM-based molecular docking assisted QSAR model to understand the inhibitory mechanism of fullerene-quinoline conjugate towards HIV-PR enzyme. The interaction profile of the compound dataset predicted by docking studies was encoded as descriptors to generate SVM-based QSAR model. The developed QSAR model was validated using test set and found to be statistically meaningful ($q^2 = 0.966$) which indicates its predicted ability. The model can be easily extended to predict the biological activities of the newly synthesized fullerene-based HIV-PR inhibitors prior to synthesis and biological evaluation.

Chloroquinoline-fullerene conjugates were also subjected to 3D-QSAR modeling using kNN-MFA method. The developed QSAR model exhibited significant cross-correction ($q^2 = 0.977$) with a low standard error (0.1284) indicating the reliability and predictive accuracy of the model. We observed the long Schiff base chain along with the fullerene enhances the lipophilic nature while chloroquinoline may act as potent antimalarial chemical moiety in interacting with the molecular targets.

References


[33] The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.